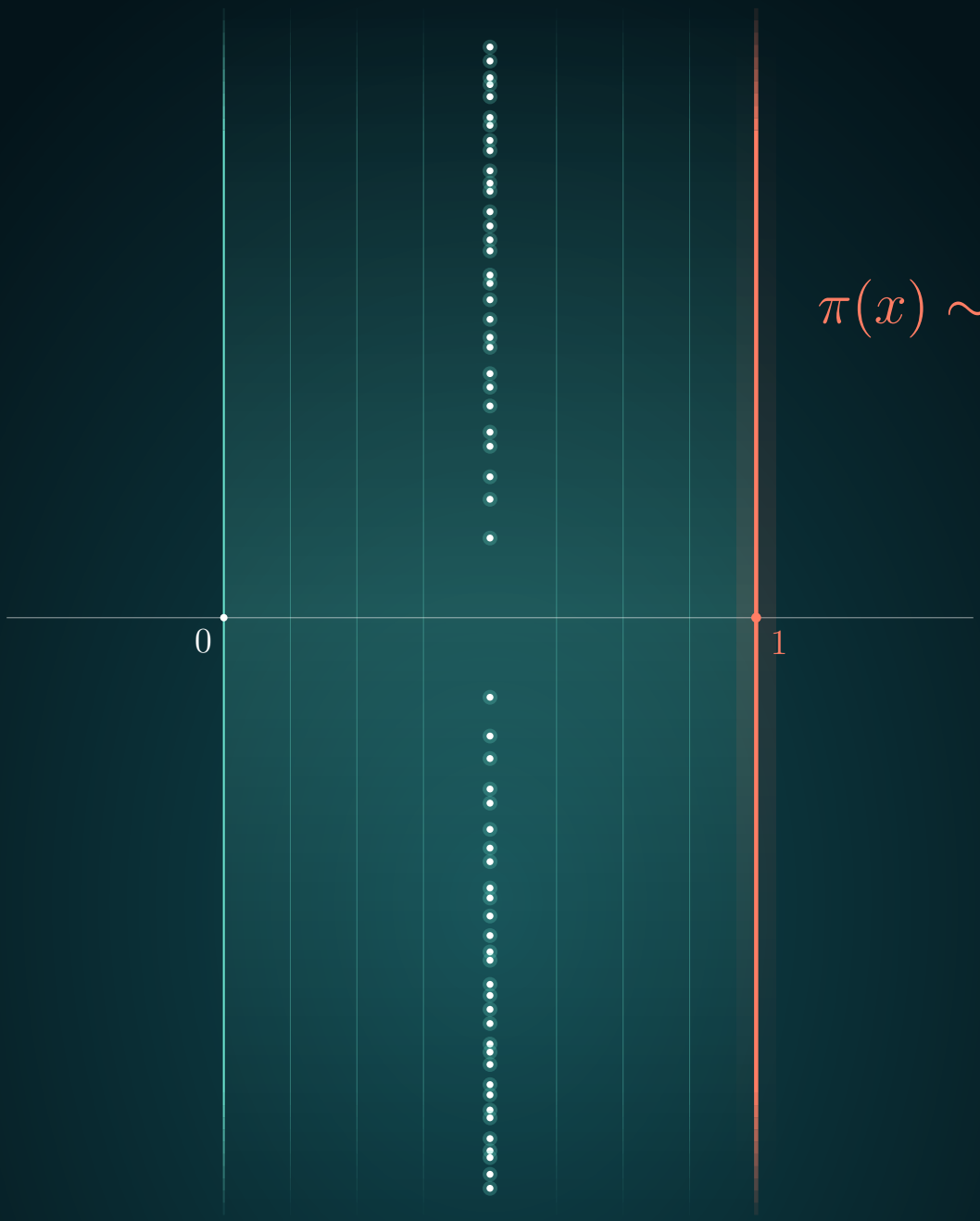


Teoria Analítica dos Números

Ramon M. Nunes

Versão de 16 de novembro de 2026

Página intencionalmente deixada em branco.

Teoria Analítica dos Números

Ramon M. Nunes

Versão de 16 de novembro de 2026

Página intencionalmente deixada em branco.

Sumário

Sobre este texto iii

I Métodos elementares 1

1 Funções aritméticas 2
1.1 Linguagem assintótica . . . 2
1.2 Exemplos de funções aritméticas . . . 3
1.3 A função de Möbius e a fórmula de inversão . . . 4
1.4 Funções multiplicativas . . . 5
1.5 Convolução de Dirichlet . . . 7
1.5.1 Inteiros relativamente primos com k . . . 9
1.6 Séries de Dirichlet formais . . . 11

2 Estimativas de somas aritméticas 13
2.1 Soma de Euler . . . 13
2.1.1 Primeiros exemplos . . . 13
2.1.2 Comparação com a integral . . . 15
2.1.3 Fórmula da soma de Euler . . . 15
2.1.4 A soma harmônica e a constante de Euler . . . 16
2.2 A soma dos logaritmos e a fórmula de Stirling . . . 16
2.3 Soma por partes . . . 18
2.4 O método da convolução . . . 18
2.5 O método da hipérbole . . . 21
2.5.1 A soma da função número de divisores . . . 21
2.5.2 O caso geral . . . 22
2.5.3 Pontos de coordenadas inteiras em um círculo . . . 23
2.6 Valores médios . . . 23
2.6.1 Média e média logarítmica . . . 23
2.6.2 A média implica a média logarítmica . . . 24

3 Distribuição dos primos: resultados elementares 27
3.1 As funções de Chebyshev e as estimativas de Chebyshev . . . 27
3.1.1 Demonstração para ψ . . . 27
3.1.2 Demonstração para θ e π; outras constantes . . . 29
3.2 Relações entre π, ψ e θ . . . 29
3.3 Os teoremas de Mertens . . . 30
3.4 Consequências do Teorema dos Números Primos . . . 32
3.5 O Teorema dos Números Primos e a função de Möbius . . . 33
3.5.1 Resultados auxiliares sobre μ . . . 33

3.5.2 O TNP implica $H_\mu = 0$ . . . 34
3.5.3 $\sum_{n\leq x} \mu(n) = o(x)$ implica o TNP . . . 35

II Métodos analíticos **38**

4 Séries de Dirichlet **39**
4.1 Convergência e analiticidade . . . 39
4.2 Representação integral . . . 39
4.3 Extensão analítica e a hipótese de Riemann . . . 40
4.3.1 A função ζ perto de $s = 1$ . . . 40
4.3.2 Extensão a partir de cotas de S_f . . . 40
4.3.3 A função de Möbius e a hipótese de Riemann . . . 42
4.3.4 Contagem de zeros . . . 42
4.4 A constante de Mertens . . . 43
4.4.1 Preparação . . . 43
4.4.2 O argumento de Hildebrand . . . 44
4.4.3 O argumento de Tenenbaum . . . 46
4.5 Inversão de Mellin . . . 48
4.6 A fórmula de Perron . . . 49
4.6.1 A fórmula de Perron para $S_f^1(x)$ . . . 49
4.6.2 A fórmula de Perron para $S_f(x)$ . . . 53
4.6.3 O caso em que x é inteiro . . . 56

5 O Teorema dos Números Primos **57**
5.1 A função zeta e o TNP: heurística . . . 57
5.2 Não-anulação de ζ na reta $\mathrm{Re}(s) = 1$ . . . 59
5.3 Cotas superiores para ζ e ζ' . . . 60
5.4 Cotas inferiores e região livre de zeros . . . 64
5.5 Prova do Teorema dos Números Primos . . . 67
5.5.1 A soma suavizada . . . 68
5.5.2 De Ψ para ψ . . . 71
5.6 Formas alternativas e o TNP para a função de Möbius . . . 72
5.7 Fórmulas explícitas e os zeros de ζ . . . 73

6 Primos em progressões aritméticas: o Teorema de Dirichlet **75**
6.1 Motivação . . . 75
6.1.1 Adaptações do argumento de Euclides . . . 75
6.1.2 A ideia de Dirichlet . . . 76
6.1.3 Analogia com séries de Fourier . . . 77
6.2 Caracteres de Dirichlet . . . 78
6.3 O grupo dos caracteres . . . 80
6.4 Relações de ortogonalidade . . . 84
6.4.1 Decomposição de funções periódicas . . . 85
6.5 Funções L e a prova do Teorema de Dirichlet . . . 86
6.5.1 Extensão à região $\mathrm{Re}(s) > 0$ . . . 86
6.5.2 Produto de Euler . . . 88
6.5.3 Redução à não-anulação e prova do Teorema de Dirichlet . . . 89
6.6 A não-anulação de $L(1, \chi)$ . . . 90
6.6.1 Caso χ não-real . . . 90
6.6.2 Caso χ real . . . 91

Referências Bibliográficas **94**

Sobre este texto

Este texto foi organizado por assunto a partir de notas de um curso de Teoria Analítica dos Números ministrado na Universidade Federal do Ceará (UFC). O texto foi feito a partir de fotos, que foram transformadas em LaTeX com a ajuda do Gemini e organizadas com a ajuda do Claude. Os trechos que resumem as notas de A. J. Hildebrand [Hil05] estão indicados ao longo do texto.

A fórmula explícita, a contagem de zeros e o termo de erro de de la Vallée Poussin são citados sem demonstração na Seção 5.7. Alguns resultados ficam como exercício.

Parte I

Métodos elementares

Capítulo 1

Funções aritméticas

1.1 Linguagem assintótica

Definição 1.1 ("O grande")**.** Dadas $f, g\colon \mathbb{R}_{>0} \to \mathbb{C}$ tais que $g(x) \geq 0$ para todo x, dizemos que $f(x) = O(g(x))$ se existem $C > 0$ e $x_0 \in \mathbb{R}_{>0}$ tais que, se $x > x_0$, então $|f(x)| \leq C\, g(x)$.

Exemplo 1.2.

- $x = O(x)$, com $C = 1$.
- $200x = O(x)$, com $C = 200$.
- $17x^2 + 1 + x = O(x^2)$, com $C = 19$. De fato, para $x > 1$ temos $|17x^2| \leq 17x^2$, $|x| \leq x^2$ e $|1| \leq x^2$; logo, pela desigualdade triangular,

$$|17x^2 + x + 1| \leq |17x^2| + |x| + |1| \leq 17x^2 + x^2 + x^2 = 19x^2.$$

- Para todo $m \in \mathbb{Z}_{>0}$, $x^m = O(e^x)$, pois $\lim\limits_{x\to+\infty} \dfrac{x^m}{e^x} = 0$.

Lema 1.3.

a) Se $f_1(x) = O(g(x))$ e $f_2(x) = O(g(x))$, então $f_1(x) + f_2(x) = O(g(x))$.

b) Se $f(x) = O(g(x))$ e $g(x) = O(h(x))$, então $f(x) = O(h(x))$.

Definição 1.4 ("o pequeno")**.** Dadas $f, g\colon \mathbb{R}_{>0} \to \mathbb{C}$, dizemos que $f(x) = o(g(x))$ se existe $x_0 > 0$ tal que, se $x > x_0$, então $g(x) \neq 0$, e além disso $\lim\limits_{x\to+\infty} \dfrac{f(x)}{g(x)} = 0$.

Exemplo 1.5.

- $x = o(x^2)$; mais geralmente, $x^m = o(x^n)$ se $m < n$.
- Para todo $m > 0$, $x^m = o(e^x)$.
- Para todos $m, n > 0$, $(\log x)^m = o(x^n)$, isto é, $\lim\limits_{x\to+\infty} \dfrac{(\log x)^m}{x^n} = 0$.

Definição 1.6 (Equivalência assintótica)**.** Dadas $f, g\colon \mathbb{R}_{>0} \to \mathbb{C}$, dizemos que $f(x) \sim g(x)$ se existe $x_0 > 0$ tal que, se $x > x_0$, então $g(x) \neq 0$, e além disso $\lim\limits_{x\to+\infty} \dfrac{f(x)}{g(x)} = 1$.

Exemplo 1.7.

- $x \sim x$.
- $x^2 + x + 8 \sim x^2$.
- $\sqrt{x+2} \sim \sqrt{x}$.

Definição 1.8. Dizemos que $f(x) \ll g(x)$ se $f(x) = O(g(x))$. Escrevemos também $f(x) \gg g(x)$ se $g(x) \ll f(x)$.

Por exemplo, $x + 1 \ll x$. A notação O também é usada dentro de fórmulas, como em $x^2 + x + 1 = x^2 + O(x)$.

Observação 1.9 (Estimativas válidas para todo $x \geq 1$). A notação O diz respeito apenas a x suficientemente grande. Em muitos casos, porém, uma estimativa $f(x) = O(g(x))$ pode ser estendida a todo $x \geq 1$, aumentando a constante. Basta que, no intervalo $[1, x_0]$, a função f seja limitada e g seja maior ou igual a uma constante positiva.

Isso acontece, por exemplo, se f é *contínua por partes* em $[1, x_0]$, isto é, contínua exceto em um número finito de pontos, nos quais tem limites laterais finitos (como as somas $\sum_{n \leq x} a_n$, que só saltam nos inteiros), e g é contínua e positiva em $[1, x_0]$. De fato, pelo Teorema de Weierstrass, uma função contínua em um intervalo fechado e limitado é limitada e atinge o seu mínimo. Aplicado a cada trecho entre dois pontos de descontinuidade, ele mostra que f é limitada em $[1, x_0]$; aplicado a g, que g tem mínimo positivo ali. Então $|f(x)| \leq C' g(x)$ para todo $x \geq 1$, com

$$C' = \max\Big(C,\ \frac{\sup_{[1,x_0]} |f|}{\min_{[1,x_0]} g}\Big).$$

A hipótese sobre g é necessária: com $g(x) = \log x$, que se anula em $x = 1$, isso não funciona.

Esta observação é usada quando uma estimativa é aplicada a uma variável que não é necessariamente grande, como $y = x/d$ com d próximo de x, e os termos O são depois somados: nesse caso precisamos de uma única constante válida para todo $y \geq 1$. Por exemplo, $\sum_{m \leq y} m = \frac{y^2}{2} + O(y)$ vale para todo $y \geq 1$, com constante 1.

Definição 1.10 (Mesma ordem de grandeza)**.** Dadas $f, g \colon \mathbb{R}_{>0} \to \mathbb{C}$, dizemos que $f(x) \asymp g(x)$ se $f(x) = O(g(x))$ e $g(x) = O(f(x))$.

1.2 Exemplos de funções aritméticas

Definição 1.11. Uma *função aritmética* é uma função $f \colon \mathbb{Z}_{>0} \to \mathbb{C}$ (ou seja, essencialmente uma sequência).

Exemplo 1.12.

1) *Função constante*: dado $c \in \mathbb{C}$, $f(n) = c$ para todo n. Em particular, quando $c = 1$, escrevemos por abuso de linguagem 1 para essa função: $1(n) = 1$.

2) *Função "unidade"*: $e(n) = \begin{cases} 1, & n = 1 \\ 0, & n \neq 1 \end{cases}$ ou seja, $e(n) = \delta_{n=1}$.

3) *Função identidade*: $\mathrm{id}(n) = n$.

4) *Função* log: $\log(n) = \log n$, na base natural.

5) *Função número de divisores*: $d(n) = \#\{d \in \mathbb{Z}_{>0};\ d \mid n\}$. Temos $d(1) = 1$, $d(2) = 2$, $d(p) = 2$ se p é primo, e $d(30) = 8$.

   Na Seção 2.5 veremos que $\dfrac{1}{x} \displaystyle\sum_{n \leq x} d(n) \sim \log x$.

6) *Função soma dos divisores*: $\sigma(n) = \sum_{d|n} d$. Aqui usamos a notação usual para somatórios:

   - $\sum_{i=a}^{b} f(i) = f(a) + f(a+1) + \cdots + f(b)$;
   - $\sum_{n \text{ par}} f(n) = f(2) + f(4) + f(6) + \ldots$;
   - $\sum_{d|n}$ indica a soma sobre todos os divisores positivos d de n.

   Temos $\sigma(1) = 1$, $\sigma(2) = 1 + 2 = 3$, $\sigma(3) = 1 + 3 = 4$, $\sigma(4) = 1 + 2 + 4 = 7$, $\sigma(6) = 1 + 2 + 3 + 6 = 12$ e $\sigma(10) = 1 + 2 + 5 + 10 = 18$.

7) *Função ϕ*: $\phi(n) = \#\{1 \leq k \leq n;\ (k, n) = 1\}$, onde (k, n) é o máximo divisor comum de k e n. Temos $\phi(3) = 2$, $\phi(6) = 2$ e $\phi(15) = 8$.

8) *Função μ (de Möbius)*:

$$\mu(n) = \begin{cases} 0 & \text{se } n \text{ não é livre de quadrados,} \\ (-1)^k & \text{se } n = p_1 p_2 \ldots p_k \text{ (primos distintos).} \end{cases}$$

   Temos $\mu(1) = 1$, $\mu(2) = -1$, $\mu(3) = -1$, $\mu(4) = 0$, $\mu(5) = -1$, $\mu(6) = 1$, $\ldots$, $\mu(30) = -1$.

9) *Função ω*: $\omega(n)$ é o número de divisores primos de n. Temos $\omega(1) = 0$, $\omega(p) = 1$, $\omega(p^2) = 1$, $\ldots$, $\omega(p_1^{\alpha_1} \ldots p_k^{\alpha_k}) = k$.

10) *Função Ω*: $\Omega(n)$ é o número de divisores primos de n contados com multiplicidade. Temos $\Omega(1) = 0$, $\Omega(p) = 1$, $\Omega(p^2) = 2$, $\Omega(p_1^{\alpha_1} \ldots p_k^{\alpha_k}) = \alpha_1 + \alpha_2 + \cdots + \alpha_k$.

11) *Função de Liouville*: $\lambda(n) = (-1)^{\Omega(n)}$.

12) *Função de von Mangoldt*:

$$\Lambda(n) = \begin{cases} \log p, & n = p^\alpha,\ \alpha \geq 1, \\ 0, & \text{caso contrário.} \end{cases}$$

   Temos $\Lambda(1) = 0$, $\Lambda(2) = \log 2$, $\Lambda(3) = \log 3$, $\Lambda(4) = \log 2$, $\ldots$, $\Lambda(6) = 0$, $\ldots$

O motivo de se estudar $\Lambda(n)$ é que ela aparece naturalmente ao se tentar estudar os números primos usando a função ζ de Riemann, $\zeta(s) = \sum_{n=1}^{+\infty} \frac{1}{n^s}$.

1.3 A função de Möbius e a fórmula de inversão

Seja μ a função de Möbius definida no Exemplo 1.12, e seja e a função unidade.

Proposição 1.13 (Fórmula da inversão de Möbius)**.**

$$\sum_{d|n} \mu(d) = \begin{cases} 1 & \textit{se } n = 1, \\ 0 & \textit{caso contrário.} \end{cases}$$

Ou seja, $\sum_{d|n} \mu(d) = e(n)$.

Demonstração (inclusão-exclusão ou binômio de Newton). Tratamos $n = 1$ e os $n > 1$ livres de quadrados; o caso geral reduz-se a estes pelo Exercício 1.14.

Se $n = 1$: $\sum_{d|1} \mu(d) = \mu(1) = 1$.

Se $n = p$: $\sum_{d|p} \mu(d) = \mu(1) + \mu(p) = 1 - 1 = 0$.

Se $n = p_1 p_2$, com p_1, p_2 distintos: $\sum_{d|p_1p_2} \mu(d) = \mu(1)+\mu(p_1)+\mu(p_2)+\mu(p_1p_2) = 1-1-1+1 = 0$.

Se $n = p_1 p_2 \dots p_k$, com $k \geq 1$ e os p_j distintos:

$$\sum_{d|p_1\dots p_k} \mu(d) = \sum_{i=0}^{k} \sum_{\substack{d|p_1p_2\dots p_k \\ d \text{ igual ao produto de } i \\ \text{primos distintos}}} \mu(d) = \sum_{i=0}^{k} \sum_{\substack{d|p_1\dots p_k \\ d=\text{produto de } i \\ \text{primos distintos}}} (-1)^i$$

$$= \sum_{i=0}^{k} (-1)^i \sum_{\substack{d|p_1\dots p_k \\ d=\text{prod. } i \text{ primos} \\ \text{distintos}}} 1.$$

A quantidade de divisores de $p_1 \dots p_k$ iguais ao produto de i primos distintos é $\binom{k}{i}$. Portanto, pelo binômio de Newton,

$$\sum_{d|p_1\dots p_k} \mu(d) = \sum_{i=0}^{k} (-1)^i \binom{k}{i} = (1-1)^k = 0. \qquad \square$$

Exercício 1.14. Mostrar que, se $\alpha_1, \dots, \alpha_k \geq 1$,

$$\sum_{d|p_1^{\alpha_1}\dots p_k^{\alpha_k}} \mu(d) = \sum_{d|p_1\dots p_k} \mu(d).$$

Com a linguagem da convolução de Dirichlet, a Proposição 1.13 se escreve $e = \mu * 1$, e daí sai a forma usual da inversão: $f * 1 = g$ implica $g * \mu = f$. Esse enunciado é o Corolário 1.27 da Seção 1.5, onde a convolução é definida.

1.4 Funções multiplicativas

Definição 1.15. Seja f uma função aritmética. Dizemos que f é *multiplicativa* se

$$f(mn) = f(m)f(n) \quad \text{sempre que } (m,n) = 1. \tag{\#}$$

Definição 1.16. Uma função é dita *completamente multiplicativa* se (#) é válida para todos $m, n \in \mathbb{Z}_{>0}$.

Observação 1.17. Se $(m,n) = 1$, a aplicação $(d,b) \mapsto db$ é uma bijeção entre os pares ordenados (d,b) com $d \mid m$ e $b \mid n$ e os divisores de mn. (Que db divide mn vale sempre; a bijeção depende de $(m,n) = 1$.)

Exemplo 1.18. Dentre as funções da Seção 1.2:

- $e(n)$, $1(n)$ e $\mathrm{id}(n)$ são completamente multiplicativas.
- $\mu(n)$ é multiplicativa (exercício), mas não completamente multiplicativa: $\mu(2) = -1$ e $\mu(4) = 0$, logo $\mu(4) \neq \mu(2)^2$.

- $\phi(n)$ é multiplicativa (fato conhecido da teoria elementar dos números), mas não completamente: $\phi(2) = 1$ e $\phi(4) = 2$, logo $\phi(4) \neq \phi(2)^2$.
- $d(n)$ e $\sigma(n)$ são multiplicativas (a ideia é a bijeção da Observação 1.17), mas não completamente multiplicativas:

$$d(2) = 2,\ d(4) = 3 \ \Rightarrow\ d(4) \neq d(2)^2,$$
$$\sigma(2) = 3,\ \sigma(4) = 7 \ \Rightarrow\ \sigma(4) \neq \sigma(2)^2.$$

- $\omega(n)$ não é multiplicativa: $1 \cdot 1 = \omega(2)\omega(3) \neq \omega(6) = 2$.
- $\Omega(n)$ não é multiplicativa: $1 \cdot 1 = \Omega(2)\Omega(3) \neq \Omega(6) = 2$.
- $\lambda(n)$ é completamente multiplicativa, como veremos na Observação 1.21.
- $\Lambda(n)$ não é multiplicativa: basta tomar $m = 2$, $n = 3$.

Definição 1.19. Seja f uma função aritmética. Dizemos que f é *aditiva* se

$$f(mn) = f(m) + f(n) \quad \text{sempre que } (m, n) = 1.$$

Definição 1.20. Uma função aritmética f é *completamente aditiva* se $f(mn) = f(m) + f(n)$ para todos $m, n \in \mathbb{Z}_{>0}$.

Observação 1.21.

- ω é aditiva, mas não é completamente aditiva.
- Ω é completamente aditiva.
- Se f é uma função completamente aditiva e $g(n) = a^{f(n)}$, com $a \in \mathbb{R}_{>0}$, então g é completamente multiplicativa: $g(mn) = a^{f(m)+f(n)} = g(m)g(n)$.
- Se $f(n) \in \mathbb{Z}_{\geq 0}$ para todo n, a mesma conclusão vale com $a \in \mathbb{C}$. Em particular, $\lambda(n) = (-1)^{\Omega(n)}$ é completamente multiplicativa.

Proposição 1.22.

a) Se f, g são multiplicativas, então $f \cdot g$ é multiplicativa.

b) Se f, g são multiplicativas e $g(n) \neq 0$ para todo n, então $\dfrac{f}{g}$ é multiplicativa.

c) Os itens anteriores continuam válidos trocando "multiplicativa" por "completamente multiplicativa".

Demonstração. Exercício. □

Uma estratégia para provar identidades entre funções aritméticas é mostrar que os dois lados da identidade definem funções multiplicativas e, em seguida, que a identidade vale para $n = p^k$, $k \geq 1$. Ela se justifica pelo lema a seguir; exemplos de uso estão na Seção 1.5.

Lema 1.23. *Sejam f e g funções multiplicativas com $f(1) = g(1) = 1$ tais que $f(p^k) = g(p^k)$ para todo p primo e todo $k \geq 1$. Então $f = g$.*

Demonstração. Para todo $n = p_1^{\alpha_1} \cdots p_r^{\alpha_r} > 1$,

$$\begin{aligned} f(n) &= f(p_1^{\alpha_1} \cdots p_r^{\alpha_r}) = f(p_1^{\alpha_1}) \cdots f(p_r^{\alpha_r}) = g(p_1^{\alpha_1}) \cdots g(p_r^{\alpha_r}) \\ &= g(p_1^{\alpha_1} \cdots p_r^{\alpha_r}) = g(n). \end{aligned}$$

□

1.5 Convolução de Dirichlet

Definição 1.24. Dadas duas funções aritméticas f e g, definimos sua *convolução* (de Dirichlet) $f * g$ por

$$(f * g)(n) = \sum_{d|n} f(d)\, g\Big(\frac{n}{d}\Big).$$

Podemos também escrever a convolução de f e g da seguinte maneira:

$$(f * g)(n) = \sum_{ab=n} f(a)\, g(b). \tag{1.1}$$

Em particular,

$$(f * g)(1) = f(1)g(1), \qquad (f * g)(p) = f(1)g(p) + f(p)g(1),$$
$$(f * g)(p^2) = f(1)g(p^2) + f(p)g(p) + f(p^2)g(1).$$

Exemplo 1.25.

1) $d = 1 * 1$.

2) $\sigma = \mathrm{id} * 1$.

3) $e = \mu * 1$ (Proposição 1.13).

4) $\phi = \mu * \mathrm{id}$.

5) $\log = \Lambda * 1$.

O item 5) fica como exercício.

Proposição 1.26 (Propriedades da convolução)**.**

i) $f * e = e * f = f$.

ii) $f * g = g * f$.

iii) $f * (g * h) = (f * g) * h$.

iv) Se $f(1) \neq 0$*, então existe uma única* g *tal que* $f * g = e$.

Demonstração. i) É imediato.
ii) Basta usar a expressão (1.1).
iii) Basta ver que os dois lados são iguais a $\sum_{abc=n} f(a)g(b)h(c)$.
iv) Por indução forte. Dada f, queremos definir $g(1), g(2), \ldots$ de modo que

$$f(1)g(1) = 1, \qquad f(1)g(2) + f(2)g(1) = 0,$$
$$f(1)g(3) + f(3)g(1) = 0, \ \ldots, \ f(1)g(n) + \cdots + f(n)g(1) = 0.$$

Mostra-se por indução que existe uma única escolha que torna as n primeiras igualdades verdadeiras. O passo indutivo consiste em escrever, para $n > 1$,

$$g(n) = -\frac{1}{f(1)} \sum_{\substack{d>1 \\ d|n}} f(d)\, g\Big(\frac{n}{d}\Big). \qquad \square$$

A função g do item iv) é a *inversa de f com respeito à convolução de Dirichlet* e será denotada por f^{-1}.

Corolário 1.27. *Se $f * 1 = g$, então $g * \mu = f$.*

Demonstração. De fato, $f * 1 = g \;\Rightarrow\; f * 1 * \mu = g * \mu \;\Rightarrow\; f * e = g * \mu \;\Rightarrow\; f = g * \mu$. □

Proposição 1.28 (Convolução de multiplicativas)**.**

1) Se f, g são multiplicativas, então $f * g$ é multiplicativa.

2) Se f é multiplicativa e $f(1) \neq 0$, então a inversa de f com respeito à convolução de Dirichlet é multiplicativa.

3) Se $f * g = h$, com f e h multiplicativas e $f(1) \neq 0$, então g é multiplicativa.

4) Se h é completamente multiplicativa, então $h(f * g) = (hf) * (hg)$.

Demonstração. 1) Sejam $m, n \in \mathbb{Z}_{>0}$ com $(m, n) = 1$. Queremos mostrar que $(f * g)(mn) = (f * g)(m)(f * g)(n)$. Temos

$$(f * g)(m) = \sum_{d|m} f(d)\, g\Big(\frac{m}{d}\Big), \qquad (f * g)(n) = \sum_{b|n} f(b)\, g\Big(\frac{n}{b}\Big),$$

donde

$$(f * g)(m)(f * g)(n) = \sum_{d|m} \sum_{b|n} f(d) f(b)\, g\Big(\frac{m}{d}\Big) g\Big(\frac{n}{b}\Big).$$

Agora, $(d, b) = 1$ e $(\frac{m}{d}, \frac{n}{b}) = 1$. Como f, g são multiplicativas, a bijeção $(d, b) \leftrightarrow db$ da Observação 1.17 dá

$$(f * g)(m)(f * g)(n) = \sum_{d|m} \sum_{b|n} f(db)\, g\Big(\frac{mn}{db}\Big) = \sum_{c|mn} f(c)\, g\Big(\frac{mn}{c}\Big) = (f * g)(mn).$$

2) Vamos provar, por indução no valor de $K = mn$, que se $(m, n) = 1$ então $f^{-1}(mn) = f^{-1}(m) f^{-1}(n)$.

Caso inicial: $K = 1$. Sabemos que $f(1) f^{-1}(1) = 1$. Como f é multiplicativa, $f(1) = f(1)^2$, e como $f(1) \neq 0$, temos $f(1) = 1$; logo $f^{-1}(1) = 1$ e

$$f^{-1}(1) = 1 = f^{-1}(1)\, f^{-1}(1).$$

Hipótese de indução: suponha que, se $mn \leq K$ com $(m, n) = 1$, então $f^{-1}(mn) = f^{-1}(m) f^{-1}(n)$.

Passo indutivo: suponha $(m, n) = 1$ e $mn = K + 1$. Como $K \geq 1$, temos $e(K + 1) = 0$. Por definição de f^{-1},

$$\begin{aligned} 0 &= \sum_{c|mn} f(c)\, f^{-1}\Big(\frac{mn}{c}\Big) = \sum_{d|m} \sum_{b|n} f(db)\, f^{-1}\Big(\frac{m}{d} \cdot \frac{n}{b}\Big) \\ &= f^{-1}(mn) + \sum_{\substack{d|m,\ b|n \\ (d,b)\neq(1,1)}} f(db)\, f^{-1}\Big(\frac{m}{d} \cdot \frac{n}{b}\Big). \end{aligned}$$

Se $(d,b) \neq (1,1)$, então $db > 1$ e, portanto, $\frac{mn}{db} < K+1$, isto é, $\frac{mn}{db} \leq K$. Logo, pela hipótese de indução,

$$\begin{aligned}
&\sum_{\substack{d|m,\ b|n \\ (d,b)\neq(1,1)}} f(db)\, f^{-1}\Big(\frac{m}{d}\Big) f^{-1}\Big(\frac{n}{b}\Big) \\
&\qquad = \sum_{d|m} \sum_{b|n} f(d)f(b)\, f^{-1}\Big(\frac{m}{d}\Big) f^{-1}\Big(\frac{n}{b}\Big) - f^{-1}(m)f^{-1}(n) \\
&\qquad = \Big(\sum_{d|m} f(d) f^{-1}\Big(\frac{m}{d}\Big)\Big)\Big(\sum_{b|n} f(b) f^{-1}\Big(\frac{n}{b}\Big)\Big) - f^{-1}(m)f^{-1}(n) \\
&\qquad = e(m)\, e(n) - f^{-1}(m)f^{-1}(n).
\end{aligned}$$

Como $mn = K+1 > 1$, temos $m > 1$ ou $n > 1$; logo $e(m) = 0$ ou $e(n) = 0$, e assim $e(m)e(n) = 0$. Conclusão: $0 = f^{-1}(mn) - f^{-1}(m)f^{-1}(n)$.

3) Segue dos itens 1) e 2), pois $g = f^{-1} * h$.

4) É imediato. □

Observação 1.29. Se f, g são multiplicativas, $f * g$ é multiplicativa. *Mas* se f, g são completamente multiplicativas, $f * g$ não precisa ser completamente multiplicativa: por exemplo, $1 * 1 = d$, e d não é completamente multiplicativa.

Vejamos a estratégia do Lema 1.23 em dois exemplos.

Exemplo 1.30. Vamos mostrar que $\mathrm{id} = \phi * 1$. Como os dois lados são funções multiplicativas, basta mostrar que $p^k = (\phi * 1)(p^k)$ para todo p primo e todo $k \geq 1$. Mas, por soma telescópica,

$$(\phi * 1)(p^k) = \sum_{d|p^k} \phi(d) = \sum_{i=0}^{k} \phi(p^i) = 1 + (p-1) + (p^2 - p) + \cdots + (p^k - p^{k-1}) = p^k.$$

Pelo Corolário 1.27, isso dá $\phi = \mu * \mathrm{id}$, o item 4) do Exemplo 1.25.

Exemplo 1.31. Considere a identidade $\displaystyle\sum_{d|n} \frac{\mu^2(d)}{\phi(d)} = \frac{n}{\phi(n)}$. O lado direito é uma função multiplicativa (quociente de duas funções multiplicativas). O lado esquerdo também: μ é multiplicativa, logo μ^2 é multiplicativa; como μ^2 e ϕ são multiplicativas, $\frac{\mu^2}{\phi}$ é multiplicativa, e portanto $\frac{\mu^2}{\phi} * 1$ é multiplicativa. Assim, pelo Lema 1.23, basta verificar a identidade para $n = p^k$.

Exercício 1.32. Encontrar uma fórmula "legal" para a inversa de Dirichlet de $\lambda = (-1)^{\Omega}$ e para a de $\lambda * 1$.

1.5.1 Inteiros relativamente primos com k

Sabemos que $\phi(k)$ é igual ao número de inteiros em $\{1, 2, \ldots, k\}$ que são relativamente primos com k. Dado x um parâmetro suficientemente grande, quantos números n relativamente primos com k existem tais que $1 \leq n \leq x$? Ou seja, queremos estudar

$$A_k(x) = \sum_{\substack{n \leq x \\ (n,k)=1}} 1.$$

Denotamos por $\lfloor y \rfloor$ a parte inteira de y e por $\{y\} = y - \lfloor y \rfloor \in [0,1)$ a sua parte fracionária; em particular, $\lfloor y \rfloor = y + O(1)$.

Tome $k = 60$. Se $x = 60m$, em cada intervalo $\{60j + 1, \ldots, 60(j + 1)\}$ temos exatamente $\phi(60) = 16$ números inteiros relativamente primos com 60. Logo, neste caso,

$$\sum_{\substack{n\leq x\\(n,60)=1}} 1 = 16 \cdot m = 16\,\frac{x}{60}.$$

Mais geralmente, para k qualquer, se $x = km$ com $m \in \mathbb{Z}_{>0}$,

$$\sum_{\substack{n\leq x\\(n,k)=1}} 1 = \frac{\phi(k)}{k}\,x.$$

Em geral, podemos limitar x entre dois múltiplos de k:

$$km \leq x < k(m + 1), \qquad m = \Big\lfloor \frac{x}{k} \Big\rfloor.$$

Daí

$$\sum_{\substack{n\leq km\\(n,k)=1}} 1 \;\leq\; A_k(x) \;\leq \sum_{\substack{n\leq k(m+1)\\(n,k)=1}} 1,$$

isto é, $\phi(k)\,m \leq A_k(x) \leq \phi(k)\,(m + 1)$, e portanto

$$A_k(x) = \phi(k)\,m + E_1(x), \quad \text{onde } |E_1(x)| \leq \phi(k).$$

Agora, $m = \lfloor \frac{x}{k} \rfloor = \frac{x}{k} - \{\frac{x}{k}\}$, de modo que

$$A_k(x) = \frac{\phi(k)}{k}\,x + E_2(x) + E_1(x), \quad \text{onde } E_2(x) = -\phi(k)\Big\{\frac{x}{k}\Big\}.$$

Como $|E_2(x)| \leq \phi(k)$, temos $|E_1(x) + E_2(x)| \leq 2\phi(k)$. Ou seja, $E(x) := E_1(x) + E_2(x) = O(\phi(k))$.

Se k é uma constante, esse termo de erro é $O(1)$. Porém, se em um problema específico quisermos saber o que acontece quando $x, k \to +\infty$, podemos fazer melhor. Por exemplo, tome $x = p^3 + 1$ e $k = p^2$. Neste caso, mostramos que

$$A_{p^2}(p^3 + 1) = \frac{p - 1}{p}(p^3 + 1) + O(p^2).$$

Mas podemos fazer melhor neste caso, usando que $p \mid n$ se, e só se, $n = pm$:

$$\begin{aligned}\sum_{\substack{n\leq p^3+1\\(n,p^2)=1}} 1 &= \sum_{\substack{n\leq p^3+1\\(n,p)=1}} 1 = \sum_{n\leq p^3+1} 1 - \sum_{\substack{n\leq p^3+1\\p\mid n}} 1 = p^3 + 1 - \sum_{m\leq \frac{p^3+1}{p}} 1\\ &= p^3 + 1 - \Big\lfloor \frac{p^3 + 1}{p} \Big\rfloor = p^3 + 1 - \frac{p^3 + 1}{p} + O(1) = \frac{p - 1}{p}(p^3 + 1) + O(1).\end{aligned}$$

No caso geral, a ideia para obter uma fórmula melhor é usar a Proposição 1.13. Temos

$$A_k(x) = \sum_{\substack{n\leq x\\(n,k)=1}} 1 = \sum_{n\leq x} e((n,k)) = \sum_{n\leq x}\sum_{d\mid(n,k)} \mu(d) = \sum_{d\mid k}\mu(d)\sum_{\substack{n\leq x\\d\mid n}} 1.$$

Mudando de variável na segunda soma,

$$\sum_{\substack{n\leq x\\d\mid n}} 1 = \sum_{m\leq \frac{x}{d}} 1 = \Big\lfloor \frac{x}{d} \Big\rfloor = \frac{x}{d} + O(1).$$

Substituindo,

$$A_k(x) = \sum_{d|k} \mu(d)\Big(\frac{x}{d} + O(1)\Big) = \Big(\sum_{d|k} \frac{\mu(d)}{d}\Big)x + O\Big(\sum_{d|k} |\mu(d)|\Big).$$

Pelo item 4) do Exemplo 1.25, o coeficiente de x é $\phi(k)/k$, como antes. Agora,

$$\sum_{d|k} |\mu(d)| = \#\{d \mid k;\ d \text{ é livre de quadrados}\} = 2^{\omega(k)} \leq d(k).$$

Assim, o erro é $O(d(k))$, ou melhor ainda, $O(2^{\omega(k)})$.

Isso é melhor do que $O(\phi(k))$ porque $d(k) \ll k^\varepsilon$ para todo $\varepsilon > 0$, ao passo que $\phi(k) \gg k^{1-\delta}$ para todo $\delta > 0$ (exercícios a seguir).

Exercício 1.33. Mostre que, para todo $\delta > 0$, existe $C_\delta > 0$ tal que $\phi(k) \geq C_\delta\, k^{1-\delta}$ para todo $k \geq 1$, e conclua que $d(k) = o(\phi(k))$ quando $k \to +\infty$. (Sugestão: compare $\phi(p^a) = p^a(1 - \frac{1}{p})$ com $p^{a(1-\delta)}$, separando os primos $p \geq 2^{1/\delta}$ dos demais.)

Exercício 1.34. Mostre que, para todo $\varepsilon > 0$, $d(k) = O_\varepsilon(k^\varepsilon)$. (Sugestão: escreva $\frac{d(k)}{k^\varepsilon} = \prod_{p^a \| k} \frac{a+1}{p^{a\varepsilon}}$ e observe que os fatores com $p \geq 2^{1/\varepsilon}$ são no máximo 1.)

1.6 Séries de Dirichlet formais

Como exemplo de manipulação de séries vindo do Cálculo: para $|x| < 1$ temos $1+x+x^2+\cdots = \dfrac{1}{1-x}$, e daí

$$1 + 2x + 3x^2 + \cdots = \frac{d}{dx}(1 + x + x^2 + \dots) = \frac{d}{dx}\Big(\frac{1}{1-x}\Big) = \frac{1}{(1-x)^2}.$$

Passamos a uma manipulação análoga, por enquanto sem nos preocuparmos com a convergência.

Definição 1.35. Dada uma função aritmética f, definimos sua *série de Dirichlet formal* por

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

Aqui s não tem significado: é um símbolo, sujeito à regra $m^s n^s = (mn)^s$.

O ponto crucial é a seguinte fórmula:

$$\Big(\sum_{n\geq 1} \frac{f(n)}{n^s}\Big)\Big(\sum_{n\geq 1} \frac{g(n)}{n^s}\Big) = \sum_{n\geq 1} \frac{1}{n^s} \underbrace{\sum_{n_1 n_2 = n} f(n_1)g(n_2)}_{(f*g)(n)} = D_{f*g}(s).$$

Outro ponto crucial é o seguinte. Se f é multiplicativa e não identicamente nula, e se escrevemos cada n como $n = p_1^{\alpha_1} \cdots p_k^{\alpha_k}$, então

$$\begin{aligned} D_f(s) &= \sum_{n\geq 1} \frac{f(n)}{n^s} = \sum_{n\geq 1} \frac{f(p_1^{\alpha_1} \cdots p_k^{\alpha_k})}{(p_1^{\alpha_1} \cdots p_k^{\alpha_k})^s} \\ &= \sum_{n\geq 1} \frac{f(p_1^{\alpha_1})}{p_1^{\alpha_1 s}} \frac{f(p_2^{\alpha_2})}{p_2^{\alpha_2 s}} \cdots = \prod_p \Big(1 + \frac{f(p)}{p^s} + \frac{f(p^2)}{p^{2s}} + \dots\Big). \end{aligned}$$

Subentende-se que, ao calcular o produto, o número 1 deve ser escolhido com finitas exceções. Por fim, se f é completamente multiplicativa (e não identicamente nula),

$$\begin{aligned} 1 + \frac{f(p)}{p^s} + \frac{f(p^2)}{p^{2s}} + \ldots &= 1 + \frac{f(p)}{p^s} + \Big(\frac{f(p)}{p^s}\Big)^2 + \Big(\frac{f(p)}{p^s}\Big)^3 + \ldots \\ &= \frac{1}{1 - \frac{f(p)}{p^s}}. \end{aligned}$$

Em particular, para $f = 1$ obtemos $D_1(s) = \zeta(s)$ e, pelo Teorema Fundamental da Aritmética e pela soma da progressão geométrica de termo inicial 1 e razão $\frac{1}{p^s}$,

$$\zeta(s) = \sum_{n\geq 1} \frac{1}{n^s} = \prod_p \Big(1 + \frac{1}{p^s} + \frac{1}{p^{2s}} + \ldots\Big) = \prod_p \frac{1}{1 - \frac{1}{p^s}} = \prod_p \Big(1 - \frac{1}{p^s}\Big)^{-1}.$$

De maneira não rigorosa, é assim que surge a função μ: invertendo,

$$\Big(\sum_{n\geq 1} \frac{1}{n^s}\Big)^{-1} = \prod_p \Big(1 - \frac{1}{p^s}\Big) = \sum_{n\geq 1} \frac{\mu(n)}{n^s},$$

onde a última igualdade é obtida por uma conta muito parecida com a da primeira igualdade acima. As questões de convergência são tratadas no Capítulo 4.

Capítulo 2

Estimativas de somas aritméticas

Dada uma função aritmética f e $x \in \mathbb{R}_{>0}$, queremos entender a soma

$$S_f(x) = \sum_{n \le x} f(n).$$

Neste capítulo reunimos as técnicas básicas para isso: a comparação com a integral e a fórmula de soma de Euler, a soma por partes, o método da convolução e o método da hipérbole. Ao final, usamos essas estimativas para discutir valores médios de funções aritméticas.

2.1 Soma de Euler

2.1.1 Primeiros exemplos

Lembre que, para x real, $\lfloor x \rfloor$ denota a parte inteira de x e $\{x\}$ a sua parte fracionária, de modo que

$$\lfloor x \rfloor = x - \{x\}, \qquad 0 \le \{x\} < 1.$$

Vamos começar pelos casos mais simples.

Exemplo 2.1 ($f = 1$)**.** Temos

$$S_f(x) = \sum_{n \le x} 1 = \#\{n \in \mathbb{Z}_{>0};\ n \le x\} = \#\{n \in \mathbb{Z};\ 0 < n \le x\} = \lfloor x \rfloor.$$

O inteiro $\lfloor x \rfloor$ é o único que satisfaz $\lfloor x \rfloor \le x < \lfloor x \rfloor + 1$. Em particular, $\lfloor x \rfloor \le x$ implica $\lfloor x \rfloor = O(x)$. Também $0 \le x - \lfloor x \rfloor < 1$, ou seja,

$$x - \lfloor x \rfloor = O(1) \ \Leftrightarrow\ \lfloor x \rfloor = x + O(1).$$

Concluímos que $\sum_{n \le x} 1 = O(x)$ e $\sum_{n \le x} 1 = x + O(1)$.

Exemplo 2.2 ($f(n) = n$)**.** Temos

$$\begin{aligned} S_f(x) = \sum_{n \le x} n &= 1 + 2 + \cdots + \lfloor x \rfloor = \frac{\lfloor x \rfloor (1 + \lfloor x \rfloor)}{2} = \frac{\lfloor x \rfloor^2}{2} + \frac{\lfloor x \rfloor}{2} \\ &= \frac{\lfloor x \rfloor^2}{2} + O(x) = \frac{(x - \{x\})^2}{2} + O(x) = \frac{x^2 - 2x\{x\} + \{x\}^2}{2} + O(x) \\ &= \frac{x^2}{2} - \underbrace{x\{x\}}_{O(x)} + \underbrace{\frac{\{x\}^2}{2}}_{O(1)} + O(x) = \frac{x^2}{2} + O(x). \end{aligned}$$

A mesma conta pode ser feita assim:

$$\frac{\lfloor x\rfloor^2}{2}+O(x)=\frac{(x+O(1))^2}{2}+O(x)=\frac{x^2}{2}+O(x).$$

Usamos que, se $f_1(x)=O(g_1(x))$ e $f_2(x)=O(g_2(x))$, então $f_1(x)f_2(x)=O(g_1(x)g_2(x))$; ou que, se $h(x)\geq 0$ e $f(x)=O(g(x))$, então $f(x)h(x)=O(g(x)h(x))$.

Exercício 2.3. Tentar encontrar fórmulas assintóticas para $\sum_{n\leq x} n^2$, $\sum_{n\leq x} n^3$, $\sum_{n\leq x} n^4$, ...

O que é uma fórmula assintótica? Exemplos são $\sum_{n\leq x} 1 = x+O(1)$ e $\sum_{n\leq x} n = \frac{x^2}{2}+O(x)$, quando $x\to+\infty$. Como resposta provisória, uma fórmula assintótica para $S_f(x)$ é algo como

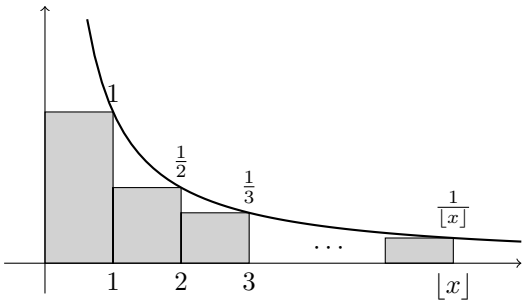
$$S_f(x)=A(x)+O(B(x)),$$

onde A é relativamente simples e $B(x)=o(A(x))$. Adiantando a resposta do exercício acima,

$$\sum_{n\leq x} n^k = \frac{x^{k+1}}{k+1}+O(x^k).$$

Observação 2.4. A constante implícita em O pode depender de k. Às vezes usaremos a notação $O_k(x^k)$.

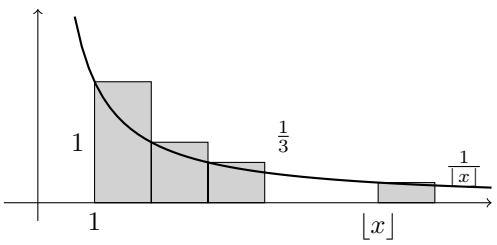
Exemplo 2.5 ($f(n)=\frac{1}{n}$)**.** Seja agora $f(n)=\dfrac{1}{n}$ para todo $n\in\mathbb{Z}_{>0}$, de modo que $S_f(x)=\sum_{n\leq x}\dfrac{1}{n}$. Esta soma é a soma das áreas dos retângulos hachurados no gráfico de $f(t)=\dfrac{1}{t}$, $t\in\mathbb{R}_{>0}$:



Excetuando-se o primeiro, todos os outros retângulos cabem, sem sobreposição, na região abaixo do gráfico entre 1 e $\lfloor x\rfloor$. Portanto a soma dessas áreas é no máximo $\int_1^{\lfloor x\rfloor}\dfrac{dt}{t}$, ou seja,

$$\frac{1}{2}+\frac{1}{3}+\cdots+\frac{1}{\lfloor x\rfloor}\leq\int_1^{\lfloor x\rfloor}\frac{dt}{t}=\log\lfloor x\rfloor\leq\log x\ \Rightarrow\ S_f(x)-1\leq\log x\ \Leftrightarrow\ S_f(x)\leq\log x+1.$$

Considere agora os seguintes retângulos:



Agora os retângulos cobrem toda a área sob o gráfico entre 1 e $\lfloor x \rfloor + 1$. Portanto

$$S_f(x) \geq \int_1^{\lfloor x \rfloor + 1} \frac{dt}{t} = \log(\lfloor x \rfloor + 1) > \log x.$$

Juntando as duas estimativas, para $x > 1$ obtemos

$$\log x < S_f(x) \leq \log x + 1 \;\Rightarrow\; S_f(x) = \log x + O(1).$$

2.1.2 Comparação com a integral

O argumento do último exemplo sugere o resultado geral a seguir.

Teorema 2.6. *Seja $f\colon [1, +\infty) \to \mathbb{R}$ monótona. Então*

$$S_f(x) = \int_1^x f(t)\,dt + O\big(\max(|f(1)|, |f(x)|)\big).$$

Demonstração. Exercício. □

Exercício 2.7. Mostrar que

- $\max(|f(x)|, |g(x)|) = O(|f(x)| + |g(x)|)$;
- $|f(x) + g(x)| = O(\max(|f(x)|, |g(x)|))$.

Em particular, o termo de erro do Teorema 2.6 também pode ser escrito como $O(|f(1)| + |f(x)|)$.

2.1.3 Fórmula da soma de Euler

Teorema 2.8 (Soma de Euler)**.** *Sejam $0 < a \leq b$ e $f\colon [a, b] \to \mathbb{R}$ uma função C^1. Então*

$$\sum_{a < n \leq b} f(n) = \int_a^b f(t)\,dt + \int_a^b \{t\} f'(t)\,dt + \{a\} f(a) - \{b\} f(b).$$

Demonstração para a e b inteiros. Suponhamos que a e b são inteiros. Nesse caso temos que provar que

$$\sum_{a < n \leq b} f(n) = \int_a^b f(t)\,dt + \int_a^b \{t\} f'(t)\,dt.$$

Considere o termo $\displaystyle\int_a^b \{t\} f'(t)\,dt$. Decompondo em pedaços e integrando por partes,

$$\begin{aligned}
\int_a^b \{t\} f'(t)\,dt &= \int_a^{a+1} \{t\} f'(t)\,dt + \cdots + \int_{b-1}^b \{t\} f'(t)\,dt = \sum_{k=a}^{b-1} \int_k^{k+1} (t-k) f'(t)\,dt \\
&= \sum_{k=a}^{b-1} \Big(\big((t-k) f(t)\big)\Big|_k^{k+1} - \int_k^{k+1} f(t)\,dt \Big) \\
&= \sum_{k=a}^{b-1} \Big(f(k+1) - \int_k^{k+1} f(t)\,dt \Big) = \sum_{a < n \leq b} f(n) - \int_a^b f(t)\,dt.
\end{aligned}$$

Isso conclui a demonstração para $a, b \in \mathbb{Z}$. □

Exercício 2.9. Completar a demonstração para o caso em que a e b não são inteiros.

2.1.4 A soma harmônica e a constante de Euler

Proposição 2.10. *Existe uma constante γ tal que*

$$\sum_{n\le x}\frac{1}{n} = \log x + \gamma + O\Big(\frac{1}{x}\Big).$$

Demonstração. Aplicamos o Teorema 2.8 com $f(t)=\frac{1}{t}$, $a=1$, $b=x$. Como $\{1\}=0$,

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

Somando 1 na igualdade,

$$\sum_{n\le x}\frac{1}{n} = \log x + 1 - \int_1^x \frac{\{t\}}{t^2}\,dt - \frac{\{x\}}{x} = \log x + \Big(1 - \int_1^x \frac{\{t\}}{t^2}\,dt\Big) + O\Big(\frac{1}{x}\Big).$$

Como $\Big|\frac{\{t\}}{t^2}\Big| \le \frac{1}{t^2}$ e $\displaystyle\int_1^{+\infty}\frac{dt}{t^2}$ converge, o mesmo vale para $\displaystyle\int_1^{+\infty}\frac{\{t\}}{t^2}\,dt$. Então podemos escrever

$$1 - \int_1^x \frac{\{t\}}{t^2}\,dt = \underbrace{1 - \int_1^{+\infty}\frac{\{t\}}{t^2}\,dt}_{=:\,\gamma} + \int_x^{+\infty}\frac{\{t\}}{t^2}\,dt.$$

Além disso,

$$\Big|\int_x^{+\infty}\frac{\{t\}}{t^2}\,dt\Big| \le \int_x^{+\infty}\frac{dt}{t^2} = \frac{1}{x},$$

o que conclui a demonstração. □

Observação 2.11. A constante γ é chamada de constante de Euler. Ela pode ser escrita como

$$\gamma = \lim_{x\to+\infty}\Big(\sum_{n\le x}\frac{1}{n} - \log x\Big).$$

2.2 A soma dos logaritmos e a fórmula de Stirling

Em geral, se f possui extensão integrável a $(1,x)$, podemos comparar $S_f(x)$ com $\displaystyle\int_1^x f(t)\,dt$; é o que faz o Teorema 2.6 quando f é monótona. Outro exemplo de função desse tipo, isto é, que possui extensão aos reais, é o logaritmo. A primeira coisa que pode ser feita é usar esse teorema e obter

$$\begin{aligned} S_{\log}(x) &= \int_1^x \log t\,dt + O(|\log x|) = [t\log t - t]_1^x + O(\log x)\\ &= x\log x - x + 1 + O(\log x) = x\log x - x + O(\log x) \qquad (x\ge 2). \end{aligned}$$

Tendo em vista aplicações futuras, vamos encontrar uma fórmula mais precisa, usando a fórmula de soma de Euler (Teorema 2.8).

Proposição 2.12. *Existe uma constante C' tal que, para $x\ge 1$,*

$$S_{\log}(x) = x\log x - x + \Big(\frac{1}{2} - \{x\}\Big)\log x + C' + O\Big(\frac{1}{x}\Big).$$

Em particular, se $x = N \in \mathbb{Z}_{>0}$,

$$S_{\log}(N) = N\log N - N + \frac{1}{2}\log N + C' + O\Big(\frac{1}{N}\Big).$$

Demonstração. Fazendo $f(t) = \log t$, $a = 1$ e $b = x$ no Teorema 2.8, e lembrando que $\log 1 = 0$ e $\{1\} = 0$,

$$S_{\log}(x) = \int_1^x \log t\, dt + \int_1^x \{t\}\frac{1}{t}\, dt - \{x\}\log x.$$

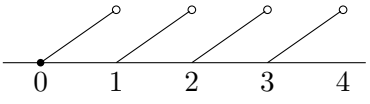
Vamos integrar por partes a segunda integral, com

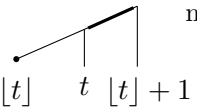
$$u'(t) = \{t\}, \quad v(t) = \frac{1}{t} \quad \Big(\Rightarrow u(t) = \int_1^t \{s\}\, ds\Big).$$

Assim,

$$\int_1^x \frac{\{t\}}{t}\, dt = \Big[\frac{u(t)}{t}\Big]_1^x + \int_1^x \frac{u(t)}{t^2}\, dt.$$

Antes de continuar, vamos analisar o gráfico de $\{t\}$:





Pela figura, cada intervalo de comprimento 1 contribui com $\frac{1}{2}$, e é fácil perceber que

$$u(t) = \int_1^{\lfloor t\rfloor} \{s\}\, ds + \int_{\lfloor t\rfloor}^t \{s\}\, ds = \frac{\lfloor t\rfloor - 1}{2} + O(1) = \frac{t}{2} + O(1). \tag{2.1}$$

Como $u(1) = 0$, segue que

$$\Big[\frac{u(t)}{t}\Big]_1^x = \frac{u(x)}{x} - \frac{u(1)}{1} = \frac{\frac{x}{2} + O(1)}{x} = \frac{1}{2} + O\Big(\frac{1}{x}\Big).$$

Graças a (2.1), temos que

$$\int_1^x \frac{u(t)}{t^2}\, dt = \frac{1}{2}\int_1^x \frac{dt}{t} + \int_1^x \frac{E_1(t)}{t^2}\, dt = \frac{\log x}{2} + \int_1^x \frac{E_1(t)}{t^2}\, dt, \quad \text{onde } E_1(t) = O(1).$$

Como $E_1(t) = O(1)$, temos $\frac{E_1(t)}{t^2} = O(\frac{1}{t^2})$, donde $\int_1^{+\infty} \Big|\frac{E_1(t)}{t^2}\Big|\, dt < +\infty$. Como convergência absoluta implica convergência simples, $C := \int_1^{+\infty} \frac{E_1(t)}{t^2}\, dt$ existe. Além disso,

$$\int_1^x \frac{E_1(t)}{t^2}\, dt = C - \int_x^{+\infty} \frac{E_1(t)}{t^2}\, dt \quad \text{e} \quad \int_x^{+\infty} \frac{E_1(t)}{t^2}\, dt \ll \int_x^{+\infty} \frac{dt}{t^2} = \frac{1}{x}.$$

Juntando todas as peças,

$$S_{\log}(x) = x\log x - x + 1 + \frac{1}{2} + \frac{\log x}{2} + C - \{x\}\log x + O\Big(\frac{1}{x}\Big),$$

que é a fórmula do enunciado, com $C' = \frac{3}{2} + C$. Se $x = N$ é inteiro, $\{x\} = 0$. □

Observação 2.13. $S_{\log}(N)$ é a mesma coisa que $\log(N!)$. Daí o seguinte corolário.

Corolário 2.14 (Fórmula de Stirling)**.** *Existe uma constante $c > 0$ tal que*

$$n! = c\sqrt{n}\,\Big(\frac{n}{e}\Big)^n\Big(1 + O\Big(\frac{1}{n}\Big)\Big).$$

Demonstração. Exponenciando a fórmula da Proposição 2.12,

$$N! = e^{N\log N}e^{-N}e^{\frac{1}{2}\log N}e^{C'}e^{O(\frac{1}{N})} = N^N e^{-N}\sqrt{N}\,c\,e^{O(\frac{1}{N})},$$

onde $c = e^{C'}$. Resta provar que $e^{O(\frac{1}{N})} = 1 + O(\frac{1}{N})$ (igualdade que deve ser lida da esquerda para a direita). Isso segue da série de Taylor da função e^x: se x é suficientemente pequeno, $e^x = 1+O(x)$. Mais precisamente, $e^x = 1+xe^y$ para algum y entre 0 e x; para x suficientemente pequeno, podemos supor $|x| < 1$. □

Observação 2.15. O valor da constante do Corolário 2.14 é $c = \sqrt{2\pi}$, mas isso não será demonstrado aqui.

2.3 Soma por partes

A soma por partes é quase a mesma coisa que a integração por partes, só que com soma, e é uma generalização da fórmula de soma de Euler (Teorema 2.8).

Teorema 2.16 (Soma por partes)**.** *Seja f uma função aritmética e g uma função C^1 em $[y, x]$, com $0 \le y < x$. Então*

$$\sum_{y<n\le x} f(n)g(n) = S_f(x)g(x) - S_f(y)g(y) - \int_y^x S_f(t)g'(t)\,dt.$$

Ideia da demonstração. Primeiro, considere o caso em que x e y são inteiros. Neste caso, escreva

$$g(n) - g(y) = \int_y^n g'(t)\,dt \iff g(n) = g(y) + \int_y^n g'(t)\,dt$$

e depois troque a ordem da soma com a integral. O caso geral fica como exercício. □

Quando f é uma função aritmética, $S_f(t) = 0$ para $t < 1$. Assim, tomando $0 < y < 1$ no Teorema 2.16, a soma passa a incluir $n = 1$ e a integral pode ser tomada a partir de 1. Vamos usar esta forma nos dois casos seguintes, que serão empregados na Seção 2.6.

Exemplo 2.17. Sejam f uma função aritmética, $s \in \mathbb{R}_{>0}$, $S(x) = \sum_{n\le x} f(n)$ e $T(x) = \sum_{n\le x} \frac{f(n)}{n^s}$. Aplicando o teorema à função $n \mapsto \frac{f(n)}{n^s}$ e a $g(t) = t^s$,

$$S(x) = \sum_{n\le x} \frac{f(n)}{n^s}\,n^s = T(x)x^s - \int_1^x T(t)st^{s-1}\,dt. \tag{2.2}$$

Aplicando-o à função f e a $g(t) = \frac{1}{t}$,

$$\sum_{n\le x} \frac{f(n)}{n} = \frac{S_f(x)}{x} + \int_1^x S_f(t)\frac{dt}{t^2}. \tag{2.3}$$

2.4 O método da convolução

Até agora só estudamos funções que possuem extensão aos reais. Isso deixa de fora as funções verdadeiramente aritméticas: $d(n)$, $\phi(n)$, $\Lambda(n)$, a função característica dos primos, $\mu(n)$, ...

Temos $\phi(n) = n\prod_{p|n}\big(1-\frac{1}{p}\big)$. Veremos que esta função não difere tanto da função $f(n) = n$, e isso nos permitirá estudar $S_\phi(x)$. Outro exemplo é a função característica dos números

livres de quadrados, $f(n) = |\mu(n)| = \mu(n)^2$: para calcular $S_f(x)$, vamos comparar f com a função 1.

Essas comparações vão se dar na forma de convoluções: vamos dizer que f e g são semelhantes ou "próximas" se $f = g * h$ com h "pequena". Como sabemos (Seções 1.3 e 1.5), pela inversão de Möbius,

$$\phi * 1 = \mathrm{id} \iff \phi = \mathrm{id} * \mu.$$

Já para a função μ^2, queremos uma função g tal que $\mu^2 = 1 * g$. Ela sempre existe, pois isso equivale a $g = \mu * \mu^2$. Como μ e μ^2 são funções multiplicativas, basta calcular $\mu * \mu^2(p^k)$ para p primo e $k \geq 1$:

$$\begin{aligned}
\mu * \mu^2(p) &= \mu(1)\mu^2(p) + \mu(p)\mu^2(1) = 0,\\
\mu * \mu^2(p^2) &= \mu(1)\mu^2(p^2) + \mu(p)\mu^2(p) + \mu(p^2)\mu^2(1) = -1,\\
\mu * \mu^2(p^3) &= \mu(1)\mu^2(p^3) + \mu(p)\mu^2(p^2) + \mu(p^2)\mu^2(p) + \mu(p^3)\mu^2(1) = 0.
\end{aligned}$$

Em geral, $\mu * \mu^2(p^k) = 0$ se $k \geq 3$, de modo que

$$g(p^k) = \begin{cases} 1 & \text{se } k = 0\\ 0 & \text{se } k = 1\\ -1 & \text{se } k = 2\\ 0 & \text{se } k \geq 3. \end{cases}$$

Então, se $n = p_1^{\alpha_1} \cdots p_k^{\alpha_k}$, com $\alpha_i \geq 1$,

$$g(n) = \begin{cases} (-1)^k & \text{se } \alpha_1 = \alpha_2 = \cdots = \alpha_k = 2\\ 0 & \text{caso contrário.} \end{cases}$$

Ou seja,

$$g(n) = \begin{cases} \mu(m) & \text{se } n = m^2\\ 0 & \text{caso contrário.} \end{cases}$$

Informalmente, podemos escrever $g(n) = \delta_{n=\square}\, \mu(\sqrt{n})$.

O mesmo se vê usando séries de Dirichlet (Seção 1.6): sejam $D_\mu(s) = \sum \dfrac{\mu(n)}{n^s}$ e $D_{\mu^2}(s) = \sum \dfrac{\mu^2(n)}{n^s}$. Como as funções μ e μ^2 são multiplicativas, temos uma fórmula como produtório:

$$D_\mu(s) = \prod_p \Big(1 + \frac{\mu(p)}{p^s} + \ldots\Big), \quad D_{\mu^2}(s) = \prod_p \Big(1 + \frac{\mu^2(p)}{p^s} + \ldots\Big).$$

Assim, $D_{\mu*\mu^2}(s) = \displaystyle\prod_p \Big(1 - \frac{1}{p^{2s}}\Big)$.

Proposição 2.18. *Para $x \geq 2$, $\displaystyle\sum_{n\leq x} \phi(n) = \frac{3}{\pi^2}x^2 + O(x \log x)$.*

Demonstração. Como $\phi = \mu * \mathrm{id}$, temos $\phi(n) = \displaystyle\sum_{d|n} \mu(d)\frac{n}{d}$, donde

$$\sum_{n\leq x} \phi(n) = \sum_{n\leq x} \sum_{d|n} \mu(d)\frac{n}{d} = \sum_{d\geq 1} \sum_{\substack{n\leq x\\ d|n}} \mu(d)\frac{n}{d}.$$

Como $d \mid n$ e $1 \le n \le x$ implicam $d \le x$, podemos escrever, usando o Exemplo 2.2 na penúltima igualdade, que vale para todo $y \ge 1$ com a mesma constante (Observação 1.9),

$$\sum_{n\le x}\phi(n) = \sum_{d\le x}\frac{\mu(d)}{d}\sum_{\substack{n\le x\\ d\mid n}} n = \sum_{d\le x}\frac{\mu(d)}{d}\sum_{m\le \frac{x}{d}} dm = \sum_{d\le x}\mu(d)\sum_{m\le\frac{x}{d}} m$$
$$= \sum_{d\le x}\mu(d)\Big(\frac{1}{2}\Big(\frac{x}{d}\Big)^2 + O\Big(\frac{x}{d}\Big)\Big) = \frac{x^2}{2}\sum_{d\le x}\frac{\mu(d)}{d^2} + O\Big(x\sum_{d\le x}\frac{|\mu(d)|}{d}\Big).$$

Como $\sum \frac{\mu(d)}{d^2}$ é convergente, podemos escrever

$$\sum_{d\le x}\frac{\mu(d)}{d^2} = S - \sum_{d>x}\frac{\mu(d)}{d^2}, \qquad S = \sum_{d=1}^{+\infty}\frac{\mu(d)}{d^2}.$$

Agora,

$$\Big|\sum_{d>x}\frac{\mu(d)}{d^2}\Big| \le \sum_{d>x}\frac{1}{d^2} = O\Big(\int_x^{+\infty}\frac{dt}{t^2}\Big) = O\Big(\frac{1}{x}\Big).$$

De maneira semelhante,

$$\Big|\sum_{d\le x}\frac{|\mu(d)|}{d}\Big| \le \Big|\sum_{d\le x}\frac{1}{d}\Big| = O(\log x) \qquad (x \ge 2).$$

Concluímos que

$$\sum_{n\le x}\phi(n) = \frac{S}{2}x^2 + O(x) + O(x\log x) = \frac{S}{2}x^2 + O(x\log x),$$

restando apenas provar que $\displaystyle\sum_{n=1}^{+\infty}\frac{\mu(n)}{n^2} = \frac{6}{\pi^2}$. Isso segue de $\sum \frac{1}{n^2} = \frac{\pi^2}{6}$ e da fórmula $(\sum \frac{\mu(n)}{n^2})(\sum \frac{1}{n^2}) = 1$. Essa fórmula foi obtida no Capítulo 1 apenas de maneira formal. Ela vale aqui porque as duas séries convergem absolutamente, o que permite multiplicá-las e reagrupar os termos: o produto é $\sum_n (\mu * 1)(n)/n^2 = 1$. A convergência das séries de Dirichlet é tratada na Seção 4.1; o fato geral de que o produto de duas séries de Dirichlet absolutamente convergentes corresponde à convolução está em [Hil05, Teorema 4.1]. □

Proposição 2.19. $\displaystyle\sum_{n\le x}\mu^2(n) = \frac{6}{\pi^2}x + O(\sqrt{x})$.

Demonstração. Como vimos, $\mu^2(n) = \displaystyle\sum_{d\mid n} g(d)$, onde $g(n) = \delta_{n=\square}\,\mu(\sqrt{n})$. Ou seja,

$$\mu^2(n) = \sum_{d^2\mid n}\mu(d) \;\Rightarrow\; S_{\mu^2}(x) = \sum_{n\le x}\sum_{d^2\mid n}\mu(d).$$

Trocando a ordem da soma, a segunda soma restringe a primeira, e com uma mudança de variável obtemos

$$S_{\mu^2}(x) = \sum_{d\ge 1}\mu(d)\sum_{\substack{n\le x\\ d^2\mid n}} 1 = \sum_{d\le\sqrt{x}}\mu(d)\sum_{\substack{n\le x\\ d^2\mid n}} 1 = \sum_{d\le\sqrt{x}}\mu(d)\sum_{m\le\frac{x}{d^2}} 1.$$

Portanto,

$$S_{\mu^2}(x) = \sum_{d\le\sqrt{x}} \mu(d)\Big(\frac{x}{d^2} + O(1)\Big) = x \sum_{d\le\sqrt{x}} \frac{\mu(d)}{d^2} + O\Big(\sum_{d\le\sqrt{x}} |\mu(d)|\Big).$$

Pelo que vimos na demonstração da Proposição 2.18, $\displaystyle\sum_{d\le\sqrt{x}} \frac{\mu(d)}{d^2} = S + O(\frac{1}{\sqrt{x}})$, e além disso $\displaystyle\sum_{d\le\sqrt{x}} |\mu(d)| = O(\sqrt{x})$. Logo $S_{\mu^2}(x) = Sx + O(\sqrt{x})$. Como $S = \frac{6}{\pi^2}$,

$$S_{\mu^2}(x) = \frac{6}{\pi^2}x + O(\sqrt{x}). \qquad \square$$

2.5 O método da hipérbole

2.5.1 A soma da função número de divisores

Temos $d = 1 * 1$. Um primeiro cálculo de $S_d(x)$ é

$$S_d(x) = \sum_{n\le x} d(n) = \sum_{n\le x}\Big(\sum_{d|n} 1\Big) = \sum_{d\le x}\sum_{\substack{d|n\\ n\le x}} 1 = \sum_{d\le x}\sum_{m\le \frac{x}{d}} 1 = \sum_{d\le x}\Big(\frac{x}{d} + O(1)\Big)$$

$$= x\sum_{d\le x}\frac{1}{d} + O(x) = x(\log x + O(1)) + O(x) = x\log x + O(x).$$

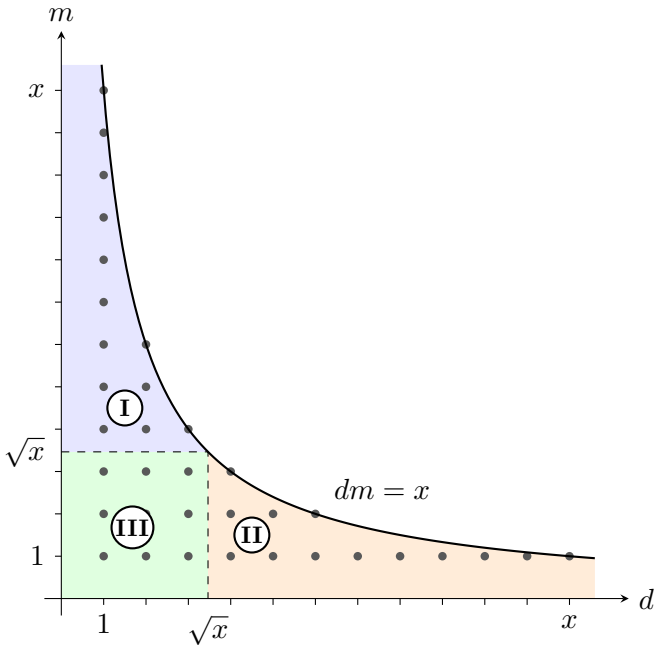
Dá para fazer melhor, com o método da hipérbole.

A ideia é a seguinte: na condição $dm \le x$, as duas variáveis não podem ser grandes ao mesmo tempo. Podemos dividir a região onde $dm \le x$ em três regiões (em todas elas, $dm \le x$):

I. região onde $d \le \sqrt{x}$ e $m > \sqrt{x}$ (d pequeno, m grande);

II. região onde $m \le \sqrt{x}$ e $d > \sqrt{x}$ (m pequeno, d grande);

III. região onde $d, m \le \sqrt{x}$ (d e m pequenos).



Exemplo com $x = 12$: os pontos são os pares (d, m) de inteiros positivos com $dm \le x$.

Note que não estamos calculando uma integral: estamos contando pontos de coordenadas inteiras.

Observe primeiro que $\sum_{d\le x}\sum_{m\le \frac{x}{d}} 1 = \sum_{dm\le x} 1$. Pela decomposição em três regiões acima,

$$\sum_{dm\le x} 1 = \sum_{d\le\sqrt{x}}\ \sum_{\sqrt{x}<m\le\frac{x}{d}} 1 + \sum_{m\le\sqrt{x}}\ \sum_{\sqrt{x}<d\le\frac{x}{m}} 1 + \sum_{\substack{d\le\sqrt{x}\\ m\le\sqrt{x}}} 1.$$

Veja que, na Região III, a condição $dm \le x$ é redundante, pois $d \le \sqrt{x}$ e $m \le \sqrt{x}$ já implicam $dm \le x$.

Para a primeira parcela do lado direito, veja que

$$\sum_{\sqrt{x}<m\le\frac{x}{d}} 1 = \left\lfloor \frac{x}{d} \right\rfloor - \lfloor\sqrt{x}\rfloor = \Big(\frac{x}{d} + O(1)\Big) - (\sqrt{x} + O(1)) = \frac{x}{d} - \sqrt{x} + O(1).$$

Com isso, usando a Proposição 2.10 (a fórmula para a soma harmônica com erro $O(\frac{1}{x})$),

$$\begin{aligned}\sum_{d\le\sqrt{x}}\ \sum_{\sqrt{x}<m\le\frac{x}{d}} 1 &= \sum_{d\le\sqrt{x}}\Big(\frac{x}{d} - \sqrt{x} + O(1)\Big) = x\sum_{d\le\sqrt{x}}\frac{1}{d} - \sqrt{x}\sum_{d\le\sqrt{x}} 1 + O(\sqrt{x})\\ &= x\Big(\log\sqrt{x} + \gamma + O\Big(\frac{1}{\sqrt{x}}\Big)\Big) - \sqrt{x}(\sqrt{x} + O(1)) + O(\sqrt{x})\\ &= \frac{x\log x}{2} + \gamma x - x + O(\sqrt{x}).\end{aligned}$$

A segunda parcela tem a mesma forma:

$$\sum_{m\le\sqrt{x}}\ \sum_{\sqrt{x}<d\le\frac{x}{m}} 1 = \frac{x\log x}{2} + \gamma x - x + O(\sqrt{x}).$$

Por fim,

$$\sum_{d\le\sqrt{x}}\ \sum_{m\le\sqrt{x}} 1 = (\sqrt{x} + O(1))(\sqrt{x} + O(1)) = x + O(\sqrt{x}).$$

Juntando tudo, temos o seguinte resultado.

Teorema 2.20. $S_d(x) = \sum_{dm\le x} 1 = x\log x + (2\gamma - 1)x + O(\sqrt{x})$.

A maneira como fizemos o método da hipérbole é pouco usual. A decomposição mais usual é

$$\sum_{dm\le x} 1 = \sum_{d\le\sqrt{x}}\ \sum_{m\le\frac{x}{d}} 1 + \sum_{m\le\sqrt{x}}\ \sum_{d\le\frac{x}{m}} 1 - \sum_{d\le\sqrt{x}}\ \sum_{m\le\sqrt{x}} 1.$$

Observação 2.21. Provamos que $S_d(x) = x\log x + (2\gamma - 1)x + \Delta(x)$, onde $\Delta(x) = O(\sqrt{x})$. Essa estimativa para Δ não é a melhor possível: ela pode ser melhorada.

2.5.2 O caso geral

O método da hipérbole pode ser generalizado para convoluções mais gerais que $d = 1 * 1$ e com noções diferentes de pequeno e grande. Mais precisamente, se $y \le x$ e $dm \le x$, então não podemos ter simultaneamente $d > y$ e $m > \frac{x}{y}$. Com isso, temos a seguinte fórmula:

$$\begin{aligned}S_{f*g}(x) = \sum_{n\le x}(f*g)(n) = &\sum_{d\le y} f(d)\sum_{m\le\frac{x}{d}} g(m) + \sum_{m\le\frac{x}{y}} g(m)\sum_{d\le\frac{x}{m}} f(d)\\ &- \sum_{d\le y} f(d)\sum_{m\le\frac{x}{y}} g(m).\end{aligned} \tag{2.4}$$

Exercício 2.22. Demonstrar a fórmula (2.4), seguindo os mesmos passos anteriores.

2.5.3 Pontos de coordenadas inteiras em um círculo

Para $x > 0$, seja

$$C(x) = \#\{(a,b) \in \mathbb{Z}^2;\ a^2 + b^2 \leq x\}$$

o número de pontos de coordenadas inteiras no disco fechado de centro 0 e raio $\sqrt{x}$.

Proposição 2.23. $C(x) = \pi x + O(\sqrt{x})$.

Demonstração. A cada um desses pontos associamos o quadrado de lado 1 centrado nele, e seja $\delta = \frac{\sqrt{2}}{2}$ a metade da diagonal desses quadrados. Os quadrados têm interiores disjuntos e área total $C(x)$. Todo ponto do plano está a distância no máximo δ do centro do quadrado que o contém; por isso a união dos quadrados contém o disco de raio $\sqrt{x}-\delta$ e está contida no disco de raio $\sqrt{x}+\delta$. Comparando as áreas, para $x \geq \delta^2$ obtemos $\pi(\sqrt{x}-\delta)^2 \leq C(x) \leq \pi(\sqrt{x}+\delta)^2$.

Expandindo os quadrados,

$$\begin{aligned} \pi(x - 2\sqrt{x}\,\delta + \delta^2) \leq C(x) \leq \pi(x + 2\sqrt{x}\,\delta + \delta^2), \\ \text{isto é,} \quad \pi x - 2\pi\sqrt{x}\,\delta + \pi\delta^2 \leq C(x) \leq \pi x + 2\pi\sqrt{x}\,\delta + \pi\delta^2, \\ \text{isto é,} \quad -2\pi\sqrt{x}\,\delta + \pi\delta^2 \leq C(x) - \pi x \leq 2\pi\sqrt{x}\,\delta + \pi\delta^2. \end{aligned}$$

Para x grande, o limitante inferior é negativo e o superior é positivo; logo $|C(x) - \pi x|$ é no máximo a diferença entre eles:

$$\begin{aligned} C(x) - \pi x &= O\big((2\pi\sqrt{x}\,\delta + \pi\delta^2) - (-2\pi\sqrt{x}\,\delta + \pi\delta^2)\big) \\ &= O(4\pi\sqrt{x}\,\delta) = O(\sqrt{x}). \end{aligned}$$

□

2.6 Valores médios

2.6.1 Média e média logarítmica

Definição 2.24. Seja f uma função aritmética. Dizemos que f possui *média* ou *valor médio* se o limite

$$M_f := \lim_{x\to+\infty} \frac{1}{x} \sum_{n\leq x} f(n)$$

existe. Neste caso, a média de f é o limite M_f.

Definição 2.25. Seja f uma função aritmética. Dizemos que f possui *média logarítmica* se o limite

$$L_f := \lim_{x\to+\infty} \frac{1}{\log x} \sum_{n\leq x} \frac{f(n)}{n}$$

existe. Neste caso, a média logarítmica de f é o limite L_f.

Exemplo 2.26.

- Se $f(n) = c$ para todo $n \in \mathbb{N}$, então $M_f = L_f = c$.
- $M_{\mu^2} = \frac{6}{\pi^2}$.
- d e ϕ não possuem média.

Lema 2.27 (Kronecker)**.** *Seja f aritmética e seja $s \in \mathbb{R}_{>0}$ tal que $\sum_{n=1}^{+\infty} \frac{f(n)}{n^s}$ converge. Então*

$$\lim_{x\to+\infty} \frac{1}{x^s} \sum_{n\leq x} f(n) = 0.$$

Demonstração do Lema 2.27. Escrevamos $S(x) = \sum_{n\le x} f(n)$ e $T(x) = \sum_{n\le x} \frac{f(n)}{n^s}$. Suponha que $T(x) \to T$ quando $x \to +\infty$. Dado $\varepsilon > 0$, existe $x_0 > 0$ tal que $x > x_0 \Rightarrow |T - T(x)| < \varepsilon$. Usando a soma por partes na forma (2.2),

$$\begin{aligned} S(x) &= T(x)x^s - \int_1^x T(t)st^{s-1}\,dt \\ &= T(x)\int_0^x st^{s-1}\,dt - \int_1^x T(t)st^{s-1}\,dt = \int_0^x T(x)st^{s-1}\,dt - \int_0^x T(t)st^{s-1}\,dt. \end{aligned}$$

Usamos aqui que $x^s = \int_0^x st^{s-1}\,dt$ e que $T(t) = 0$ para todo $t \in (0,1)$. De modo que, para $x > x_0$,

$$S(x) = \int_0^x (T(x) - T(t))st^{s-1}\,dt = \underbrace{\int_0^{x_0} (T(x) - T(t))st^{s-1}\,dt}_{A} + \underbrace{\int_{x_0}^x (T(x) - T(t))st^{s-1}\,dt}_{B}.$$

Em B temos $x_0 \le t \le x$. Portanto $|T(x) - T(t)| \le |T(x) - T| + |T - T(t)| < 2\varepsilon$. Por outro lado, em A,

$$\begin{aligned} |T(x) - T(t)| &\le |T(x) - T| + |T - T(t)| \le \varepsilon + |T| + |T(t)| \\ &\le \varepsilon + |T| + \sum_{n\le t}\left|\frac{f(n)}{n^s}\right| \le \varepsilon + |T| + \sum_{n\le x_0}\left|\frac{f(n)}{n^s}\right| =: M(x_0,\varepsilon). \end{aligned}$$

O principal ponto desta última estimativa é que $M = M(x_0,\varepsilon)$ não depende de x. Segue que:

- $|A| \le \int_0^{x_0} |T(x) - T(t)|st^{s-1}\,dt \le M \cdot x_0^s$;
- $|B| \le \int_{x_0}^x |T(x) - T(t)|st^{s-1}\,dt \le 2\varepsilon \int_{x_0}^x st^{s-1}\,dt = 2\varepsilon(x^s - x_0^s) \le 2\varepsilon x^s$.

Logo $\frac{|S(x)|}{x^s} \le \frac{Mx_0^s}{x^s} + 2\varepsilon$ para $x > x_0$ e, como M e x_0 não dependem de x, $\limsup_{x\to+\infty} \frac{|S(x)|}{x^s} \le 2\varepsilon$. Como $\varepsilon > 0$ é arbitrário, $\lim_{x\to+\infty} \frac{S(x)}{x^s} = 0$. □

Em particular, quando $s = 1$, temos $M_f = 0$.

Observação 2.28. Podemos tomar $s \in \mathbb{C}$ com $\operatorname{Re} s > 0$ (com a mesma demonstração, estimando $|st^{s-1}| = |s|\, t^{\operatorname{Re} s - 1}$).

2.6.2 A média implica a média logarítmica

Teorema 2.29. *Se M_f existe, então L_f também existe e $L_f = M_f$.*

Demonstração. Seja $A = M_f$ e seja $T_f(x) = \sum_{n\le x} \frac{f(n)}{n}$. Usando a soma por partes na forma (2.3),

$$T_f(x) = \frac{S_f(x)}{x} + \int_1^x S_f(t)\frac{dt}{t^2}.$$

Como $\frac{S_f(x)}{x} \to A$, segue que

$$\frac{S_f(x)}{x\log x} \to 0 \quad \text{quando } x \to +\infty. \tag{2.5}$$

Seja $\varepsilon > 0$. Por hipótese, existe $x_0 = x_0(\varepsilon) > 1$ tal que, se $t > x_0$, então $\left|\frac{S_f(t)}{t} - A\right| < \varepsilon$. Se $1 \le t \le x_0$,

$$\left|\frac{S_f(t)}{t} - A\right| \le \sum_{n\le x_0} |f(n)| + |A| =: K_0 \quad (K_0 \text{ só depende de } \varepsilon).$$

Com isso, para $x > x_0$,

$$\begin{aligned}\left|\int_1^x \frac{S_f(t)}{t^2}\,dt - A\log x\right| &= \left|\int_1^x \frac{S_f(t)}{t^2}\,dt - \int_1^x \frac{A}{t}\,dt\right| = \left|\int_1^x \Big(\frac{S_f(t)}{t} - A\Big)\frac{dt}{t}\right| \\ &\le \int_1^{x_0}\left|\frac{S_f(t)}{t} - A\right|\frac{dt}{t} + \int_{x_0}^x \left|\frac{S_f(t)}{t} - A\right|\frac{dt}{t} \\ &\le K_0\int_1^{x_0}\frac{dt}{t} + \varepsilon\int_{x_0}^x \frac{dt}{t} = K_0\log x_0 + \varepsilon(\log x - \log x_0) \\ &\le K_0\log x_0 + \varepsilon\log x,\end{aligned}$$

donde

$$\left|\frac{1}{\log x}\int_1^x \frac{S_f(t)}{t^2}\,dt - A\right| \le \frac{K_0\log x_0}{\log x} + \varepsilon.$$

Claramente, existe x_1 suficientemente grande tal que $x > x_1 \Rightarrow \frac{K_0\log x_0}{\log x} < \varepsilon$. Ou seja, se $x > \max\{x_0, x_1\}$, temos

$$\left|\frac{1}{\log x}\int_1^x \frac{S_f(t)}{t^2}\,dt - A\right| \le 2\varepsilon.$$

Como ε é arbitrário, $\lim\limits_{x\to+\infty}\dfrac{1}{\log x}\displaystyle\int_1^x \frac{S_f(t)}{t^2}\,dt = A$. Esse limite, junto com (2.5), nos diz que $\frac{T_f(x)}{\log x} \to A$ quando $x \to +\infty$. Ou seja, L_f existe e é igual a A. □

Exercício 2.30. Mostrar que

$$\sum_{p\le x} 1 \sim \frac{x}{\log x} \;\Rightarrow\; \frac{1}{\log x}\sum_{p\le x}\frac{1}{p} \to 0.$$

Exercício 2.31. Suponha que f é uma função aritmética tal que $S_f(x) \sim \frac{x}{\log x}$. Mostre que $\displaystyle\sum_{n\le x}\frac{f(n)}{n} \sim \log\log x$.

O exemplo a seguir mostra que uma função pode ter média logarítmica sem ter média.

Exemplo 2.32. Considere a função

$$f(n) = \begin{cases} n & \text{se } n = 2^k \\ 0 & \text{caso contrário.}\end{cases}$$

Seja $y = \left\lfloor \frac{\log x}{\log 2}\right\rfloor$. Então

$$S_f(x) = \sum_{n\le x} f(n) = \sum_{2^k\le x} 2^k = \sum_{k\le\frac{\log x}{\log 2}} 2^k = 1 + 2 + \cdots + 2^y = 2^{y+1} - 1.$$

Suponha $x = x_m = 2^m$; nesse caso $y = m$. Então $2^{y+1} - 1 = 2^{m+1} - 1 = 2x_m - 1$, donde

$$\frac{S_f(x_m)}{x_m} = 2 - \frac{1}{x_m} \xrightarrow[m\to+\infty]{} 2.$$

Seja agora $x'_m = 2^m - 1$. Nesse caso $y = \left\lfloor \frac{\log(2^m-1)}{\log 2}\right\rfloor = m - 1$, e

$$2^{y+1} - 1 = 2^m - 1 = x'_m \;\Rightarrow\; \frac{S_f(x'_m)}{x'_m} = 1 \to 1 \quad (m \to +\infty).$$

Ou seja, temos duas subsequências para as quais os limites são números diferentes, e portanto M_f não existe. Por outro lado,

$$\sum_{n\leq x}\frac{f(n)}{n}=\sum_{2^k\leq x}1=\sum_{k\leq\frac{\log x}{\log 2}}1=y+1.$$

Mas sabemos que $\lfloor z\rfloor=z+O(1)$. Portanto $\displaystyle\sum_{n\leq x}\frac{f(n)}{n}=\frac{\log x}{\log 2}+O(1)$, e

$$L_f=\lim_{x\to+\infty}\frac{1}{\log x}\sum_{n\leq x}\frac{f(n)}{n}=\frac{1}{\log 2}.$$

Capítulo 3

Distribuição dos primos: resultados elementares

3.1 As funções de Chebyshev e as estimativas de Chebyshev

Definição 3.1. Definimos, sendo p sempre primo e Λ a função de von Mangoldt:

- $\pi(x) := \sum_{p \le x} 1 = \#\{p \le x;\ p \text{ é primo}\}$;
- $\theta(x) := \sum_{p \le x} \log p$;
- $\psi(x) := \sum_{n \le x} \Lambda(n)$.

Vamos estudar essas três quantidades ao mesmo tempo. A função $\pi(x)$ é a mais natural, $\psi(x)$ aparecerá naturalmente no contexto de séries de Dirichlet, pois

$$\sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s} = -\frac{\zeta'(s)}{\zeta(s)} \quad (\Lambda = \log * \mu),$$

e $\theta(x)$ é um ponto intermediário.

Teorema 3.2 (Chebyshev)**.** *Temos:*

(i) $\pi(x) \asymp \frac{x}{\log x}$ *(isto é,* $\frac{x}{\log x} \ll \pi(x) \ll \frac{x}{\log x}$*);*

(ii) $\theta(x) \asymp x$ $(x \ll \theta(x) \ll x)$*;*

(iii) $\psi(x) \asymp x$ $(x \ll \psi(x) \ll x)$*.*

3.1.1 Demonstração para ψ

Começamos pelo item (iii); os itens (ii) e (i) serão deduzidos dele na Subseção 3.1.2.

Demonstração do item (iii). Suponha $x > 10$. Seja $S(x) = \sum_{n \le x} \log n$ e $D(x) = S(x) - 2S(\frac{x}{2})$.

Vimos, por comparação com a integral (Seção 2.2), que $S(x) = x \log x - x + O(\log x)$. Logo

$$D(x) = x \log x - x - x \log\Big(\frac{x}{2}\Big) + x + O(\log x) = x \log 2 + O(\log x)$$
$$\Rightarrow \lim_{x \to +\infty} \frac{D(x)}{x} = \log 2 \in \Big(\frac{1}{2}, 1\Big).$$

Portanto, existe $x_0 > 100$ tal que, se $x > x_0$,

$$\frac{x}{2} < D(x) < x. \tag{$\oplus$}$$

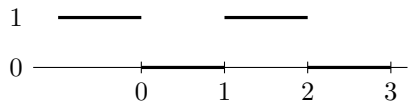
Por outro lado, $\log = \Lambda * 1$; portanto, para $x \geq 1$,

$$S(x) = \sum_{n\leq x} \log n = \sum_{n\leq x}\sum_{d|n} \Lambda(d) = \sum_{d\leq x} \Lambda(d) \sum_{\substack{n\leq x\\ d|n}} 1 = \sum_{d\leq x} \Lambda(d) \left\lfloor \frac{x}{d} \right\rfloor.$$

Assim, como $\lfloor \frac{x}{2d} \rfloor = 0$ se $d > \frac{x}{2}$,

$$\begin{aligned} D(x) = S(x) - 2S(x/2) &= \sum_{d\leq x} \Lambda(d) \left\lfloor \frac{x}{d} \right\rfloor - 2 \sum_{d\leq \frac{x}{2}} \Lambda(d) \left\lfloor \frac{x}{2d} \right\rfloor \\ &= \sum_{d\leq x} \Lambda(d) \left\lfloor \frac{x}{d} \right\rfloor - 2 \sum_{d\leq x} \Lambda(d) \left\lfloor \frac{x}{2d} \right\rfloor = \sum_{d\leq x} \Lambda(d)\, \Omega\Big(\frac{x}{d}\Big), \end{aligned}$$

onde $\Omega(y) = \lfloor y \rfloor - 2\lfloor \frac{y}{2} \rfloor$. Vejamos o gráfico de Ω:



Em particular,

$$\Omega(t) = \begin{cases} 1 & \text{se } t \in [1,2), \\ \in \{0,1\} & \text{em geral,} \end{cases}$$

de modo que

$$D(x) \leq \sum_{d\leq x} \Lambda(d) = \psi(x), \tag{I}$$

$$D(x) \geq \sum_{\frac{x}{2}<d\leq x} \Lambda(d) = \psi(x) - \psi(x/2). \tag{II}$$

Usando (I) e ($\oplus$), segue que $\psi(x) \geq D(x) > \frac{x}{2}$ para $x > x_0$. Ou seja, $x \ll \psi(x)$.

Agora, por (II), $\psi(x) \leq D(x) + \psi(\frac{x}{2})$. Do mesmo modo, $\psi(x/2) \leq D(x/2) + \psi(x/4)$, e assim por diante. Continuamos enquanto $\frac{x}{2^k} > x_0$. Assim, se k_0 é tal que $\frac{x}{2^{k_0+1}} \leq x_0 < \frac{x}{2^{k_0}}$, temos

$$\begin{aligned} \psi(x) &\leq D(x) + \psi(x/2), \\ \psi(x/2) &\leq D(x/2) + \psi(x/4), \\ &\vdots \\ \psi(x/2^{k_0}) &\leq D(x/2^{k_0}) + \psi(x/2^{k_0+1}). \end{aligned}$$

Somando essas desigualdades e usando $D(y) \leq y$ para $y > x_0$,

$$\begin{aligned} \psi(x) &\leq D(x) + D\Big(\frac{x}{2}\Big) + \cdots + D\Big(\frac{x}{2^{k_0}}\Big) + \psi\Big(\frac{x}{2^{k_0+1}}\Big) \\ &\leq x + \frac{x}{2} + \cdots + \frac{x}{2^{k_0}} + \underbrace{\psi(x_0)}_{\substack{\text{não depende}\\ \text{de } x}} \\ &\leq 2x + \psi(x_0) \ll x. \end{aligned}$$

□

Observação 3.3 (Sobre as constantes implícitas). Em $x \ll \psi(x)$, mostramos que $x \le 2\psi(x)$ para x suficientemente grande. Em $\psi(x) \ll x$, conseguimos $\psi(x) \le (2+\varepsilon)x$ para todo $\varepsilon > 0$ e x suficientemente grande. Para isto usamos $\frac{1}{2} < \log 2 < 1$. Trocando 1 por $\log 2 + \frac{\varepsilon}{2}$ e $\frac{1}{2}$ por $\log 2 - \varepsilon$, é possível mostrar, com o mesmo argumento, que para x suficientemente grande

$$\psi(x) \ge (\log 2 - \varepsilon)x \quad \Big(\text{isto é, } x \le \frac{1}{\log 2 - \varepsilon}\,\psi(x)\Big) \qquad \text{e} \qquad \psi(x) \le (2\log 2 + \varepsilon)x.$$

3.1.2 Demonstração para θ e π; outras constantes

Provado que $\psi(x) \asymp x$, seguem as outras duas estimativas.

Separando as potências de primos com expoente $m \ge 2$, que só ocorrem para $p \le \sqrt{x}$,

$$0 \le \psi(x) - \theta(x) = \sum_{p \le \sqrt{x}} \log p \sum_{2 \le m \le \frac{\log x}{\log p}} 1 \le \sqrt{x}\log x, \tag{3.1}$$

o que transfere as cotas de ψ para θ. Para π, a cota inferior vem de $\pi(x) \ge \theta(x)/\log x$. A superior vem de separar os primos em $p \le \sqrt{x}$ e $\sqrt{x} < p \le x$, o que dá $\pi(x) \le \sqrt{x} + \frac{2}{\log x}\theta(x)$.

Hildebrand cita um argumento alternativo pelo binomial central. Todo primo $n < p \le 2n$ divide $\binom{2n}{n} \le 2^{2n}$, logo $\theta(2n) - \theta(n) \le (2\log 2)n$, e iterando chega-se a $\theta(x) \le (4\log 2)x$.

O argumento de ψ vale com quaisquer $c_1 < \log 2$ e $c_2 > 2\log 2$ para x grande (Observação 3.3). Chebyshev obteve $c_1 = 0{,}92\ldots$ e $c_2 = 1{,}10\ldots$ com outra combinação linear de $S(x/k)$, onde $S(x) = \sum_{n \le x} \log n$.

Diamond e Erdős mostraram que dá para aproximar as constantes de 1 por esse caminho, mas a prova usa o próprio Teorema dos Números Primos (TNP), de modo que isso não dá uma prova nova do TNP.

3.2 Relações entre π, ψ e θ

As três funções podem ser comparadas de forma precisa.

Teorema 3.4 ([Hil05, Teorema 3.2]). *Para $x \ge 2$,*

$$\theta(x) = \psi(x) + O(\sqrt{x}), \qquad \pi(x) = \frac{\psi(x)}{\log x} + O\Big(\frac{x}{\log^2 x}\Big).$$

Ideia da prova. Na primeira, estima-se (3.1) por $\pi(\sqrt{x})\log x \ll \sqrt{x}$, agora usando o Teorema 3.2. Na segunda, a soma por partes dá

$$\pi(x) = \frac{\theta(x)}{\log x} + \int_2^x \frac{\theta(t)}{t\log^2 t}\,dt,$$

e, com $\theta(t) \ll t$ e separando a integral em $\sqrt{x}$, o último termo é $\ll x/\log^2 x$. Por fim, troca-se θ por ψ pela primeira parte.

Corolário 3.5 ([Hil05, Corolário 3.3]: formas equivalentes do TNP). *Quando $x \to \infty$, são equivalentes: $\pi(x) \sim \frac{x}{\log x}$, $\theta(x) \sim x$ e $\psi(x) \sim x$.*

As três funções $\psi(x)$, $\theta(x)$ e $\pi(x)\log x$ diferem por $O(x/\log x)$, menor que os termos principais.

3.3 Os teoremas de Mertens

Os teoremas de Mertens contam os primos com pesos logarítmicos.

Teorema 3.6 (Mertens)**.**

(1) $\displaystyle\sum_{n\le x}\frac{\Lambda(n)}{n}=\log x+O(1)$;

(2) $\displaystyle\sum_{p\le x}\frac{\log p}{p}=\log x+O(1)$;

(3) $\displaystyle\sum_{p\le x}\frac{1}{p}=\log\log x+A+O\Big(\frac{1}{\log x}\Big)$;

(4) $\displaystyle\prod_{p\le x}\Big(1-\frac{1}{p}\Big)=\frac{e^{-\gamma}}{\log x}\Big(1+O\Big(\frac{1}{\log x}\Big)\Big)$.

Observação 3.7. Nesta seção só provaremos que o lado esquerdo de (4) é da forma $\frac{C}{\log x}(1+O(\frac{1}{\log x}))$ para alguma constante C. A prova de que $C=e^{-\gamma}$ é feita na Seção 4.4.

Demonstração dos itens (1), (2) e (3). Item (1). Na demonstração do Teorema 3.2 mostramos que

$$x\log x-x+O(\log x)=\sum_{n\le x}\log n=\sum_{n\le x}\sum_{d\mid n}\Lambda(d)=\sum_{d\le x}\Lambda(d)\left\lfloor\frac{x}{d}\right\rfloor.$$

Usando que $\lfloor x\rfloor=x+O(1)$,

$$x\log x-x+O(\log x)=x\sum_{d\le x}\frac{\Lambda(d)}{d}+O\Big(\sum_{d\le x}\Lambda(d)\Big).$$

Pelo Teorema 3.2, $\sum_{d\le x}\Lambda(d)=O(x)$. Logo

$$x\sum_{d\le x}\frac{\Lambda(d)}{d}=x\log x+O(x)\Rightarrow\sum_{d\le x}\frac{\Lambda(d)}{d}=\log x+O(1),$$

o que prova (1).

Item (2). Assim como em (3.1), podemos usar a definição de Λ para escrever

$$\begin{aligned}\sum_{n\le x}\frac{\Lambda(n)}{n}&=\sum_{p\le x}\frac{\log p}{p}+\sum_{p^2\le x}\frac{\log p}{p^2}+\sum_{p^3\le x}\frac{\log p}{p^3}+\dots\\&=\sum_{p\le x}\frac{\log p}{p}+\sum_{k\ge2}\sum_{p^k\le x}\frac{\log p}{p^k}.\end{aligned}$$

Basta mostrar que $\sum_{k\ge2}\sum_{p^k\le x}\frac{\log p}{p^k}=O(1)$. Completando a soma, mudando a ordem das somas e somando a progressão geométrica,

$$\sum_{k\ge2}\sum_{p^k\le x}\frac{\log p}{p^k}\le\sum_p\log p\sum_{k\ge2}\frac{1}{p^k}=\sum_p\log p\cdot\frac{1/p^2}{1-\frac{1}{p}}=\sum_p\frac{\log p}{p^2-p}<+\infty,$$

ou seja, $O(1)$.

Item (3). Para estudar $\sum_{p\le x}\frac{1}{p}$, podemos usar a soma por partes:

$$\sum_{p\le x}\frac{1}{p}=\sum_{p\le x}\frac{\log p}{p}\cdot\frac{1}{\log p}=\frac{\sum_{p\le x}\frac{\log p}{p}}{\log x}+\int_2^x\Big(\sum_{p\le t}\frac{\log p}{p}\Big)\frac{dt}{t(\log t)^2}$$
$$=\frac{\log x+O(1)}{\log x}+\int_2^x\Big(\sum_{p\le t}\frac{\log p}{p}\Big)\frac{dt}{t(\log t)^2}.$$

Defina $E(t)=\sum_{p\le t}\frac{\log p}{p}-\log t$. Pelo item (2), $E(t)=O(1)$. Isso quer dizer que existem $x_0>0$ e $C>0$ tais que

$$t>x_0\Rightarrow|E(t)|\le C. \tag{α}$$

Por outro lado, se $2\le t\le x_0$,

$$\left|\sum_{p\le t}\frac{\log p}{p}-\log t\right|\le\left|\sum_{p\le t}\frac{\log p}{p}\right|+|\log t|\le\sum_{p\le x_0}\frac{\log p}{p}+\log x_0\le C', \tag{β}$$

onde C' pode depender de x_0. Juntando (α) e (β), existe $C''>0$ tal que $|E(t)|<C''$ para todo $t\in[2,+\infty)$. Com isso,

$$\int_2^x\Big(\sum_{p\le t}\frac{\log p}{p}\Big)\frac{dt}{t(\log t)^2}=\int_2^x\log t\cdot\frac{dt}{t(\log t)^2}+\int_2^x\frac{E(t)}{t(\log t)^2}\,dt.$$

O primeiro termo é

$$\int_2^x\frac{dt}{t\log t}=\log\log t\Big|_2^x=\log\log x-\log\log 2.$$

Em relação ao segundo termo, $\int_2^{+\infty}\frac{E(t)}{t(\log t)^2}\,dt$ é absolutamente convergente, por comparação com $\int_2^{+\infty}\frac{C''}{t(\log t)^2}\,dt<+\infty$. Portanto,

$$\int_2^x\frac{E(t)}{t(\log t)^2}\,dt=\underbrace{\int_2^{+\infty}\frac{E(t)}{t(\log t)^2}\,dt}_{=:A_0}-\int_x^{+\infty}\frac{E(t)}{t(\log t)^2}\,dt$$

e

$$\int_x^{+\infty}\frac{E(t)}{t(\log t)^2}\,dt\ll\int_x^{+\infty}\frac{dt}{t(\log t)^2}=\frac{1}{\log x}.$$

Conclusão:

$$\sum_{p\le x}\frac{1}{p}=\log\log x+\underbrace{(1-\log\log 2+A_0)}_{=:A}+O\Big(\frac{1}{\log x}\Big). \qquad \square$$

Para passar de (3) a (4) usaremos o seguinte fato.

Lema 3.8. *Se $y\in\left[-\frac{1}{2},\frac{1}{2}\right]$, então $e^y=1+O(y)$. Isto é, existe $c>0$ tal que $|e^y-1|\le c|y|$ para todo $y\in\left[-\frac{1}{2},\frac{1}{2}\right]$.*

Demonstração. Parte-se da série

$$e^y=1+y+\frac{y^2}{2}+\frac{y^3}{3!}+\dots$$

e compara-se com uma progressão geométrica. $\square$

Demonstração do item (4), a menos da constante. A fórmula (4) equivale a

$$\sum_{p\le x}\log\Big(1-\frac{1}{p}\Big)=-\gamma-\log\log x+\log\Big(1+O\Big(\frac{1}{\log x}\Big)\Big),$$

ou seja, a

$$\sum_{p\le x}\log\Big(1-\frac{1}{p}\Big)=-\gamma-\log\log x+O\Big(\frac{1}{\log x}\Big).$$

Vamos provar essa fórmula com uma constante c_0 no lugar de $-\gamma$. Como $\log(1-y)\approx -y$, escrevemos

$$\log\Big(1-\frac{1}{p}\Big)=-\frac{1}{p}+R(p),$$

onde

$$R(p)=\log\Big(1-\frac{1}{p}\Big)+\frac{1}{p}=-\frac{1}{p}-\frac{1}{2p^2}-\frac{1}{3p^3}-\cdots+\frac{1}{p}=-\frac{1}{2p^2}-\frac{1}{3p^3}-\dots$$

Daí

$$|R(p)|=\frac{1}{2p^2}+\frac{1}{3p^3}+\cdots<\frac{1}{2p^2}+\frac{1}{2p^3}+\cdots=\frac{1}{2p^2}\Big(\frac{1}{1-\frac{1}{p}}\Big)=\frac{1}{2p(p-1)}=O\Big(\frac{1}{p^2}\Big).$$

Logo,

$$\sum_{p\le x}\log\Big(1-\frac{1}{p}\Big)=\sum_{p\le x}\Big(-\frac{1}{p}+R(p)\Big)=-\sum_{p\le x}\frac{1}{p}+\sum_{p\le x}R(p).$$

Como $R(p)=O(\frac{1}{p^2})$, pelo critério de comparação a série $\sum_p R(p)$ converge, digamos para c, e $\sum_{p\le x}R(p)=c-\sum_{p>x}R(p)$. Daí, pelo item (3),

$$-\sum_{p\le x}\frac{1}{p}+\sum_{p\le x}R(p)=-\log\log x-A+O\Big(\frac{1}{\log x}\Big)+c-\sum_{p>x}R(p).$$

Como

$$\sum_{p>x}R(p)=O\Big(\int_x^{+\infty}\frac{1}{t^2}\,dt\Big)=O\Big(\frac{1}{x}\Big),$$

provamos, com $c_0=c-A$, que

$$\sum_{p\le x}\log\Big(1-\frac{1}{p}\Big)=c_0-\log\log x+O\Big(\frac{1}{\log x}\Big).$$

Exponenciando ambos os termos,

$$\prod_{p\le x}\Big(1-\frac{1}{p}\Big)=\frac{e^{c_0}}{\log x}\,e^{O(\frac{1}{\log x})},$$

e o Lema 3.8 dá a forma anunciada na Observação 3.7, com $C=e^{c_0}$. □

3.4 Consequências do Teorema dos Números Primos

Por enquanto, os teoremas de Mertens são o mais próximo do Teorema dos Números Primos a que se consegue chegar. A prova desse teorema fica para o Capítulo 5. Vamos analisar suas consequências.

Teorema 3.9 (Teorema dos Números Primos — TNP)**.** $\pi(x) \sim \dfrac{x}{\log x}$.

Pelo Corolário 3.5, isto é equivalente a $\psi(x) \sim x$ ou a $\theta(x) \sim x$.

Corolário 3.10 (Consequências do TNP)**.**

(1) Se numerarmos os primos em ordem crescente, $p_1 = 2$, $p_2 = 3$, $p_3 = 5$, $p_4 = 7, \ldots$, então $p_n \sim n \log n$.

(2) $\omega(n)$ tem ordem máxima $\dfrac{\log n}{\log\log n}$, isto é, $\limsup \dfrac{\omega(n)}{\log n/\log\log n} = 1$.

(3) Para todo $\varepsilon > 0$, existe $x_0 = x_0(\varepsilon)$ tal que, se $x > x_0$, existe p primo com $x < p \leq (1+\varepsilon)x$. (O Postulado de Bertrand é o caso $\varepsilon = 1$, $x_0 = 1$.)

(4) As frações do tipo $\dfrac{p}{q}$, onde p e q são primos, formam um conjunto denso em $\mathbb{R}_{>0}$.

(5) Sejam $a_1, a_2, \ldots, a_m \in \{0, 1, \ldots, 9\}$ com $a_1 \neq 0$. Então existe um número primo p cujos m primeiros dígitos na base 10 são, nessa ordem, $a_1, a_2, \ldots, a_m$ (isto é, $p = a_1 \ldots a_m\, b_{m+1} \ldots b_M$).

Exercício 3.11. Demonstre o Corolário 3.10.

3.5 O Teorema dos Números Primos e a função de Möbius

Teorema 3.12. *O TNP vale se, e somente se,* $\displaystyle\sum_{n\leq x} \mu(n) = o(x)$.

3.5.1 Resultados auxiliares sobre μ

Antes da demonstração, alguns resultados auxiliares sobre μ.

Teorema 3.13 (Mertens para μ)**.** *Para todo $x \geq 1$,*

$$\Big|\sum_{n\leq x} \frac{\mu(n)}{n}\Big| \leq 1. \quad \Big(\textit{Logo } \sum_{n\leq x} \frac{\mu(n)}{n} = O(1).\Big)$$

Demonstração. A ideia é a seguinte:

$$\begin{aligned} 1 = \sum_{n\leq x} e(n) &= \sum_{n\leq x}\sum_{d|n} \mu(d) = \sum_{d\leq x} \mu(d) \sum_{m\leq \frac{x}{d}} 1 \\ &= \sum_{d\leq x} \mu(d)\Big\lfloor \frac{x}{d}\Big\rfloor = x\sum_{d\leq x} \frac{\mu(d)}{d} - \sum_{d\leq x} \mu(d)\Big\{\frac{x}{d}\Big\}. \end{aligned}$$

Portanto

$$x\sum_{d\leq x} \frac{\mu(d)}{d} = 1 + \sum_{d\leq x} \mu(d)\Big\{\frac{x}{d}\Big\} \;\Rightarrow\; \Big|x\sum_{d\leq x} \frac{\mu(d)}{d}\Big| \leq 1 + x \;\Rightarrow\; \Big|\sum_{d\leq x} \frac{\mu(d)}{d}\Big| \leq 1 + \frac{1}{x}.$$

Para obter a cota 1, trocamos x por $\lfloor x \rfloor$ nas contas, de modo que

$$\lfloor x\rfloor \sum_{d\leq \lfloor x\rfloor} \frac{\mu(d)}{d} = 1 + \sum_{d\leq \lfloor x\rfloor} \mu(d)\Big\{\frac{\lfloor x\rfloor}{d}\Big\}.$$

Como o termo $d = 1$ da última soma se anula, pois $\{\lfloor x \rfloor\} = 0$,

$$\Big|\lfloor x\rfloor \sum_{d\le\lfloor x\rfloor} \frac{\mu(d)}{d}\Big| \le 1 + \lfloor x\rfloor - 1 \;\Rightarrow\; \Big|\underbrace{\sum_{d\le\lfloor x\rfloor} \frac{\mu(d)}{d}}_{=\sum_{d\le x}\frac{\mu(d)}{d}}\Big| \le 1. \qquad \square$$

Observação 3.14. Essa troca de x por $\lfloor x \rfloor$ raramente traz algum ganho.

Corolário 3.15. *$L_\mu = 0$ e, se M_μ existe, $M_\mu = 0$.*

Aqui M_f e L_f são os valores médios da Seção 2.6: $L_\mu = 0$ porque a soma do Teorema 3.13 é limitada, e $M_\mu = L_\mu$ quando M_μ existe.

Lema 3.16. *Seja f uma função aritmética e defina, se existir,*

$$H_f = \lim_{x\to+\infty} \frac{1}{x\log x} \sum_{n\le x} f(n)\log n.$$

Então H_f existe se, e somente se, M_f existe. Nesse caso, $H_f = M_f$.

Demonstração. Usa-se soma por partes. Com $S_f(t) = \sum_{n\le t} f(n)$ e $S_{f\cdot\log}(t) = \sum_{n\le t} f(n)\log n$,

- $\displaystyle\sum_{n\le x} f(n)\log n = S_f(x)\log x - \int_1^x \frac{S_f(t)}{t}\,dt$;
- $\displaystyle\sum_{n\le x} f(n) = f(1) + \frac{S_{f\cdot\log}(x)}{\log x} + \int_2^x S_{f\cdot\log}(t)\frac{dt}{t(\log t)^2}$.

Suponha que $M_f = A$ existe, isto é, $S_f(t) = At + o(t)$. Então $\int_1^x \frac{S_f(t)}{t}\,dt = O(x) = o(x\log x)$, e a primeira fórmula dá $S_{f\cdot\log}(x) = Ax\log x + o(x\log x)$, ou seja, $H_f = A$. Reciprocamente, suponha que $H_f = B$ existe, isto é, $S_{f\cdot\log}(t) = Bt\log t + o(t\log t)$. Então $\frac{S_{f\cdot\log}(x)}{\log x} = Bx + o(x)$ e

$$\int_2^x S_{f\cdot\log}(t)\frac{dt}{t(\log t)^2} = O\Big(\int_2^x \frac{dt}{\log t}\Big) = O\Big(\frac{x}{\log x}\Big) = o(x),$$

e a segunda fórmula dá $S_f(x) = Bx + o(x)$, ou seja, $M_f = B$. $\square$

Portanto, $\sum_{n\le x}\mu(n) = o(x) \Leftrightarrow M_\mu = 0 \Leftrightarrow H_\mu = 0$. Com isso, precisamos provar que o TNP equivale a $H_\mu = 0$.

3.5.2 O TNP implica $H_\mu = 0$

Demonstração de que o TNP implica $H_\mu = 0$. Suponha o TNP na forma $\psi(x) \sim x$. Vamos usar a fórmula $\Lambda = \mu * \log$. Com ela,

$$\begin{aligned}\psi(x) = \sum_{n\le x}\Lambda(n) &= \sum_{n\le x}\sum_{d\mid n}\mu(d)\log\Big(\frac{n}{d}\Big)\\ &= \sum_{n\le x}\log n\sum_{d\mid n}\mu(d) - \sum_{n\le x}\sum_{d\mid n}\mu(d)\log d = -\sum_{n\le x}\sum_{d\mid n}\mu(d)\log d.\end{aligned}$$

Assim, $\Lambda = f * 1$, onde $f(n) = -\mu(n)\log n$. Por inversão de Möbius,

$$-\mu(n)\log n = (\mu * \Lambda)(n). \qquad \text{(A)}$$

Seja $T(x) = \sum_{n\le x} \mu(n)\log n$. Usando a convolução (A) e a mudança de variável $n = d\cdot m$,

$$T(x) = -\sum_{d\le x}\mu(d)\sum_{\substack{n\le x\\ d\mid n}}\Lambda(n/d) = -\sum_{d\le x}\mu(d)\sum_{m\le \frac{x}{d}}\Lambda(m) = -\sum_{d\le x}\mu(d)\psi\Big(\frac{x}{d}\Big).$$

Queremos mostrar que $\frac{T(x)}{x\log x}\to 0$ quando $x\to+\infty$. Seja $S = \sum_{d\le x}\mu(d)\psi(\frac{x}{d})$, de modo que $|T(x)| = |S|$. Na soma S, mesmo que x seja um número grande, $\frac{x}{d}$ pode assumir valores pequenos (se d for grande). Como $\psi(x)\sim x$, para todo $\varepsilon>0$ existe $x_0 = x_0(\varepsilon)\ge 3$ tal que, se $t\ge x_0$, então $\big|\frac{\psi(t)}{t}-1\big|<\varepsilon$, isto é, $|\psi(t)-t|<\varepsilon t$. Escreva

$$S = \underbrace{\sum_{d\le \frac{x}{x_0}}\mu(d)\psi\Big(\frac{x}{d}\Big)}_{S_1} + \underbrace{\sum_{d>\frac{x}{x_0}}\mu(d)\psi(x/d)}_{S_2}.$$

Para S_1,

$$S_1 = \sum_{d\le \frac{x}{x_0}}\mu(d)\frac{x}{d} + \sum_{d\le\frac{x}{x_0}}\mu(d)\Big(\psi(x/d)-\frac{x}{d}\Big),$$

donde, pelo Teorema 3.13 e como $\sum_{d\le x/x_0}\frac{1}{d}\le 1+\log\frac{x}{x_0}\le\log x$,

$$|S_1|\le x\Big|\sum_{d\le\frac{x}{x_0}}\frac{\mu(d)}{d}\Big| + \varepsilon x\sum_{d\le\frac{x}{x_0}}\frac{|\mu(d)|}{d}\le x+\varepsilon x\log x.$$

Para $S_2 = \sum_{\frac{x}{x_0}<d\le x}\mu(d)\psi(\frac{x}{d})$, como ψ é crescente,

$$|S_2|\le \sum_{\frac{x}{x_0}<d\le x}|\mu(d)|\psi(x_0)\le\psi(x_0)\sum_{d\le x}|\mu(d)|\le\psi(x_0)x.$$

Logo $|T(x)|\le(1+\psi(x_0))x+\varepsilon x\log x$ e

$$\Big|\frac{T(x)}{x\log x}\Big|\le\frac{1+\psi(x_0)}{\log x}+\varepsilon\le 2\varepsilon$$

para x suficientemente grande. Como ε é arbitrário, concluímos. □

3.5.3 $\sum_{n\le x}\mu(n) = o(x)$ implica o TNP

Demonstração de que $\sum_{n\le x}\mu(n)=o(x)$ implica o TNP. Seguimos [Hil05, pp. 99–102]. A demonstração parte da identidade $\Lambda = \mu * \log$, que segue de $\log = \Lambda * 1$ por inversão de Möbius. A aplicação direta dessa identidade, porém, não funciona: trocando a ordem das somas,

$$\psi(x) = \sum_{n\le x}\Lambda(n) = \sum_{d\le x}\log d\sum_{n\le\frac{x}{d}}\mu(n),$$

e uma cota da forma $\varepsilon x/d$ para a soma interna daria apenas $\varepsilon\sum_{d\le x}(\log d)\frac{x}{d}$, que é da ordem de $\varepsilon x(\log x)^2$ e não fornece sequer uma cota do tipo de Chebyshev para $\psi(x)$. Mesmo cotas mais fortes para $\sum_{n\le x}\mu(n)$, da forma $O(x/(\log x)^A)$, dariam por esse caminho, no máximo, $\psi(x)\ll x$.

Para contornar essa dificuldade, aproximamos a função $\log n$, que é suave, por uma função aritmética que nada tem de suave, mas tem boas propriedades aritméticas: $d(n)-2\gamma$, onde

$d = 1 * 1$ é a função divisor e γ é a constante de Euler. O motivo da escolha é que as somas parciais das duas funções são muito próximas. De fato, pelo Teorema 2.20,

$$\sum_{n\le x}\big(d(n)-2\gamma\big) = x\log x + (2\gamma-1)x + O(\sqrt{x}) - 2\gamma x + O(1) = x(\log x - 1) + O(\sqrt{x}),$$

e, pela Proposição 2.12,

$$\sum_{n\le x}\log n = x(\log x - 1) + O(\log x) \qquad (x \ge 2).$$

Assim, definindo a função r por $\log = 1 * 1 - 2\gamma \cdot 1 + r$, temos, para $x \ge 1$,

$$R(x) := \sum_{n\le x} r(n) = O(\sqrt{x}). \tag{3.2}$$

Substituindo em $\Lambda = \mu * \log$ e usando $\mu * 1 = e$,

$$\Lambda = \mu * (1 * 1 - 2\gamma \cdot 1 + r) = (\mu * 1) * 1 - 2\gamma(\mu * 1) + \mu * r = 1 - 2\gamma e + \mu * r.$$

Logo

$$\psi(x) = \sum_{n\le x} 1 - 2\gamma + \sum_{n\le x}(\mu * r)(n) = x + O(1) + E(x), \qquad \text{onde} \quad E(x) = \sum_{n\le x}(\mu * r)(n),$$

e resta mostrar que $E(x) = o(x)$.

Vale a pena comparar $E(x)$ com a soma $\psi(x) = \sum_{n\le x}(\mu * \log)(n)$ de que partimos. As duas são somas de convoluções com μ; a diferença é que $\log n$ foi trocada por $r(n)$, que, por (3.2), é muito menor em média. Isso faz toda a diferença: escrevendo $E(x) = \sum_{d\le x}\mu(d)\sum_{n\le x/d} r(n)$ e usando (3.2) na soma interna,

$$E(x) \ll \sum_{d\le x}\sqrt{x/d} = \sqrt{x}\sum_{d\le x}\frac{1}{\sqrt{d}} \ll x,$$

o que é melhor, por um fator $(\log x)^2$, do que o mesmo argumento daria com $\log n$ no lugar de $r(n)$, e já basta para obter a cota de Chebyshev para $\psi(x)$.

Resta melhorar essa cota de $O(x)$ para $o(x)$, usando a hipótese, que ainda não foi usada. Seja $M(x) = \sum_{n\le x}\mu(n)$. Fixado y com $1 \le y \le x$, a fórmula (2.4), com $f = r$ e $g = \mu$, dá

$$E(x) = \Sigma_1 + \Sigma_2 - \Sigma_3,$$

onde

$$\Sigma_1 = \sum_{d\le y} r(d)\, M\Big(\frac{x}{d}\Big), \qquad \Sigma_2 = \sum_{m\le x/y}\mu(m)\, R\Big(\frac{x}{m}\Big), \qquad \Sigma_3 = R(y)\, M\Big(\frac{x}{y}\Big).$$

Seja $\varepsilon > 0$. Pela hipótese, existe $x_0 = x_0(\varepsilon)$ tal que $|M(t)| \le \varepsilon t$ para $t \ge x_0$; e, por (3.2), existe uma constante c tal que $|R(t)| \le c\sqrt{t}$ para todo $t \ge 1$. Então, para $x \ge x_0 y$,

$$|\Sigma_3| \le c\sqrt{y}\,\varepsilon\,\frac{x}{y} \le c\varepsilon x$$

e

$$|\Sigma_2| \le \sum_{m\le x/y} c\sqrt{\frac{x}{m}} = c\sqrt{x}\sum_{m\le x/y}\frac{1}{\sqrt{m}} \le \frac{2cx}{\sqrt{y}},$$

pois $\sum_{m\le t} m^{-1/2} \le 1 + \int_1^t s^{-1/2}\,ds \le 2\sqrt{t}$ para $t \ge 1$. Por fim,

$$|\Sigma_1| \le \sum_{d\le y} |r(d)|\,\varepsilon\,\frac{x}{d} = C(y)\,\varepsilon x, \qquad \text{onde} \quad C(y) = \sum_{d\le y} \frac{|r(d)|}{d}$$

depende apenas de y. Juntando as três cotas,

$$|E(x)| \le x\Big(c\varepsilon + C(y)\varepsilon + \frac{2c}{\sqrt{y}}\Big), \qquad \text{donde} \qquad \limsup_{x\to+\infty} \frac{|E(x)|}{x} \le \varepsilon\big(c + C(y)\big) + \frac{2c}{\sqrt{y}}$$

para quaisquer $\varepsilon > 0$ e $y \ge 1$ fixados. Como ε é arbitrário, o lim sup é no máximo $2c/\sqrt{y}$; como y pode ser tomado arbitrariamente grande, ele é 0. Logo $E(x) = o(x)$, isto é, $\psi(x) \sim x$. (Para o argumento funcionar, é essencial que a constante c não dependa de ε nem de y, e que $C(y)$ não dependa de ε.) □

Parte II

Métodos analíticos

Capítulo 4

Séries de Dirichlet

Na Seção 1.6, as séries de Dirichlet foram tratadas de modo formal. Neste capítulo, elas são usadas como funções de $s = \sigma + it$. Três fatos fazem essa passagem: a convergência e a analiticidade em semiplanos (Seção 4.1), a representação integral (Seção 4.2) e o comportamento da função ζ perto de $s = 1$ (Seção 4.3).

Notação:

$$D_f(s) = \sum_{n\geq 1} \frac{f(n)}{n^s}, \qquad S_f(x) = \sum_{n\leq x} f(n), \qquad T_f(x) = \sum_{n\leq x} \frac{f(n)}{n}.$$

A Seção 4.3 usa a representação integral para estender D_f a partir de cotas de S_f. A Seção 4.4 aplica a segunda fórmula do Teorema 4.3 à série de $\log\zeta$ e usa $\zeta(s) = \frac{1}{s-1} + O(1)$ para achar a constante de Mertens. As Seções 4.5 e 4.6 fazem o caminho inverso, de D_f para S_f, com a inversão de Mellin e a fórmula de Perron.

4.1 Convergência e analiticidade

Toda série de Dirichlet tem uma abscissa de convergência absoluta $\sigma_a = \sigma_a(f)$ e uma abscissa de convergência $\sigma_c = \sigma_c(f)$, com $\sigma_c \leq \sigma_a \leq \sigma_c + 1$. O ponto central é que convergir em s_0 implica convergir em todo s com $\sigma > \sigma_0$, uniformemente em compactos desse semiplano, o que sai de soma por partes e do critério de Cauchy.

Teorema 4.1 ([Hil05, Teorema 4.7])**.** *D_f é analítica em $\sigma > \sigma_c$, e pode ser derivada termo a termo: $D_f'(s) = -\sum_{n\geq 1} f(n)(\log n)n^{-s}$.*

Para funções definidas por integrais, como as da Seção 4.2, usaremos sem demonstração o critério a seguir.

Lema 4.2. *Sejam $U \subseteq \mathbb{C}$ aberto e $F\colon (0,+\infty)\times U \to \mathbb{C}$ contínua, tal que, para todo $x \in (0,+\infty)$, a função $s \mapsto F(x,s)$ é holomorfa em U. Se, para todo compacto $K \subset U$, a integral $\int_0^{+\infty} F(x,s)\,dx$ converge uniformemente em K, então ela define uma função holomorfa em U.*

4.2 Representação integral

A passagem de $S_f(x)$ para $D_f(s)$ é a transformada de Mellin.

Teorema 4.3 ([Hil05, Teorema 4.10]: transformada de Mellin)**.** *Se σ_c é finita, então*

$$D_f(s) = s\int_1^\infty S_f(x)\,x^{-s-1}\,dx \qquad (\sigma > \max(0,\sigma_c)),$$
$$D_f(s) = (s-1)\int_1^\infty T_f(x)\,x^{-s}\,dx \qquad (\sigma > \max(1,\sigma_c)).$$

Ideia da demonstração. A primeira fórmula é uma soma por partes. A segunda é a primeira aplicada a $f(n)/n$ no ponto $s-1$, já que $D_{f(n)/n}(s-1) = D_f(s)$.

Observação 4.4. Como S_f e T_f se anulam em $x < 1$, as integrais do Teorema 4.3 também podem ser escritas a partir de 0.

4.3 Extensão analítica e a hipótese de Riemann

4.3.1 A função ζ perto de $s = 1$

Com $f = 1$ e $S_f(x) = \lfloor x\rfloor = x - \{x\}$, a primeira fórmula do Teorema 4.3 e o fato de que $\sum_{n\le x} 1 = x + O(1)$ dão a extensão meromorfa de ζ à região $\{s \in \mathbb{C};\ \mathrm{Re}(s) > 0\}$.

Teorema 4.5 ([Hil05, Teorema 4.11 e Corolário 4.12])**.** *A função ζ se estende a $\sigma > 0$ por*

$$\zeta(s) = \frac{s}{s-1} - s\int_1^\infty \{x\}\,x^{-s-1}\,dx,$$

analítica exceto por um polo simples em $s = 1$, de resíduo 1. Para $|s-1| \le \frac{1}{2}$, $s \neq 1$,

$$\zeta(s) = \frac{1}{s-1} + \gamma + O(|s-1|).$$

A integral converge absoluta e uniformemente em $\sigma \ge \varepsilon$, pois é limitada por $1/\varepsilon$, e daí vem a analiticidade.

4.3.2 Extensão a partir de cotas de S_f

O mesmo argumento se aplica a uma função aritmética qualquer, desde que se tenha uma cota para S_f.

Teorema 4.6 ([Hil05, Teorema 4.13])**.** *Seja f uma função aritmética.*

(i) Se $S_f(x) = O(x^\theta)$ para algum $\theta \ge 0$, então D_f possui extensão holomorfa ao semiplano $\mathrm{Re}(s) > \theta$.

(ii) Se $S_f(x) = Ax^\alpha + O(x^\theta)$, com $A \in \mathbb{C}$ e $\alpha > \theta \ge 0$, então $D_f(s) - \frac{As}{s-\alpha}$ possui extensão holomorfa ao semiplano $\mathrm{Re}(s) > \theta$. Em particular, D_f possui extensão meromorfa a esse semiplano, com no máximo um polo, simples, em $s = \alpha$, de resíduo αA.

Demonstração do Teorema 4.6. Observação inicial. Como $S_f(x) = O(x^\theta)$, temos

$$f(n) = S_f(n) - S_f(n-1) = O(n^\theta + (n-1)^\theta) = O(n^\theta).$$

Isso garante que, se $\sigma > \theta + 1$, então

$$\frac{|f(n)|}{n^\sigma} = O\Big(\frac{1}{n^{\sigma-\theta}}\Big),$$

de modo que $\sum |f(n)|n^{-\sigma}$ converge por comparação. Portanto, $\sigma_c(f) \le \sigma_a(f) \le \theta + 1 < +\infty$.

Item (i). Pelo Teorema 4.3,

$$D_f(s) = s\int_0^{+\infty} S_f(x)x^{-s-1}\,dx \quad \text{para } \mathrm{Re}(s) \text{ suficientemente grande.}$$

Agora, basta mostrar que o lado direito define uma função holomorfa em $\{s;\ \mathrm{Re}(s) > \theta\}$.

Como $S_f(x) = 0$ para todo $x < 1$, podemos tomar a integral a partir de $x = 1$. E, como $S_f(x) = O(x^\theta)$,

$$S_f(x)x^{-s-1} = O(x^{\theta-\mathrm{Re}(s)-1}).$$

Logo, por comparação, $\int_1^{+\infty} S_f(x)x^{-s-1}\,dx$ converge em $\{s \in \mathbb{C};\ \mathrm{Re}(s) > \theta\}$.

A função $s \mapsto S_f(x)x^{-s-1}$ é claramente holomorfa. Para garantir que $\int_1^{+\infty} S_f(x)x^{-s-1}\,dx$ é holomorfa, basta mostrar, pelo Lema 4.2, que a convergência é uniforme em compactos. Para ver isto, note que todo compacto K de $\{s \in \mathbb{C};\ \mathrm{Re}(s) > \theta\}$ está contido em um conjunto da forma $\{s \in \mathbb{C};\ \mathrm{Re}(s) \geq \theta_0\}$ para algum $\theta_0 > \theta$. Assim, se $s \in K$,

$$\begin{aligned}\int_Y^{+\infty} |S_f(x)x^{-s-1}|\,dx &\ll \int_Y^{+\infty} x^{\theta-\mathrm{Re}(s)-1}\,dx \\ &\leq \int_Y^{+\infty} x^{\theta-\theta_0-1}\,dx,\end{aligned}$$

e esta última cota não depende de $s \in K$ e tende a 0 quando $Y \to +\infty$, pois $\theta - \theta_0 < 0$.

A função $F(x, s) = S_f(x)x^{-s-1}$ não é contínua em x, pois S_f é uma função escada, de modo que o Lema 4.2 não se aplica literalmente. O que se usa é que, em cada intervalo $[n, n+1]$, ela coincide, salvo no extremo direito, com a função contínua $S_f(n)x^{-s-1}$. Assim, cada $\int_n^{n+1} F(x, s)\,dx$ é holomorfa, e a integral é o limite, uniforme em compactos, das somas parciais dessas funções.

Item (ii). Como $S_f(x) = Ax^\alpha + O(x^\theta)$, em particular $S_f(x) = O(x^\alpha)$. Com isso, podemos proceder como acima para garantir que $\sigma_c(f) \leq \sigma_a(f) < +\infty$.

Novamente pelo Teorema 4.3,

$$D_f(s) = s\int_0^{+\infty} S_f(x)x^{-s-1}\,dx \quad \text{para } \mathrm{Re}(s) \text{ suficientemente grande}$$

(em particular, tome $\mathrm{Re}(s) > \alpha + 1$). Agora, basta mostrar que a função obtida ao subtrair $\frac{As}{s-\alpha}$ do lado direito é holomorfa em $\{s \in \mathbb{C};\ \mathrm{Re}(s) > \theta\}$.

Mais uma vez usando que $S_f(x) = 0$ se $x < 1$, o lado direito é igual a

$$\begin{aligned}&s\int_1^{+\infty} S_f(x)x^{-s-1}\,dx \\ &\quad= s\int_1^{+\infty} Ax^{\alpha-s-1}\,dx + s\int_1^{+\infty} (S_f(x) - Ax^\alpha)x^{-s-1}\,dx \\ &\quad= As\,\frac{x^{\alpha-s}}{\alpha-s}\Big|_{x=1}^{x=+\infty} + s\int_1^{+\infty} (S_f(x) - Ax^\alpha)x^{-s-1}\,dx \\ &\quad= \frac{As}{s-\alpha} + s\int_1^{+\infty} (S_f(x) - Ax^\alpha)x^{-s-1}\,dx.\end{aligned}$$

O fato de que o segundo termo acima define uma função holomorfa em $\{s \in \mathbb{C};\ \mathrm{Re}(s) > \theta\}$ segue como no item (i), pois $S_f(x) - Ax^\alpha = O(x^\theta)$. □

Exercício 4.7. Na demonstração acima supusemos, implicitamente, que $\sigma_c(f) > -\infty$. Mostre que o caso $\sigma_c(f) = -\infty$ (por exemplo, $f(n) = 1/n!$) também está coberto: o Teorema 4.3 continua valendo, e a demonstração não muda.

4.3.3 A função de Möbius e a hipótese de Riemann

Exemplo 4.8. Sejam $f = \mu$ e $D_\mu(s) = \dfrac{1}{\zeta(s)}$, com $\sigma_a(\mu) = 1$. Se soubéssemos que

$$S_\mu(x) = O\big(x^{\frac{1}{2}+\varepsilon}\big) \quad \forall \varepsilon > 0,$$

pelo item (i) do Teorema 4.6, isto significaria que $\frac{1}{\zeta(s)}$ possui extensão holomorfa à região $\{\mathrm{Re}(s) > \frac{1}{2}\}$. Ou seja, $\frac{1}{\zeta}$ não possui polos nesta região; mas

$$\mathrm{polo}\Big(\frac{1}{\zeta}\Big) = \mathrm{zero}(\zeta).$$

Isto significaria, portanto, que ζ não possui zero na região $\{\mathrm{Re}(s) > \frac{1}{2}\}$. Esta é a hipótese de Riemann.

Na verdade, não se sabe se existe algum $\theta < 1$ tal que

$$S_\mu(x) = O(x^\theta).$$

O melhor que se consegue é ganhar um pouco mais do que algumas potências de $\log x$, ou coisa que o valha:

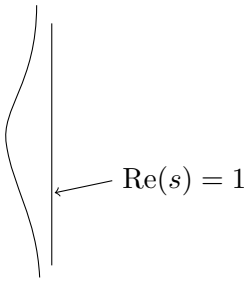
$$S_\mu(x) = O\big(xe^{-c\sqrt{\log x}}\big) \quad (\exists c) \qquad \text{(de la Vallée Poussin, 1899)},$$

$$S_\mu(x) = O\Big(xe^{-c\frac{(\log x)^{3/5}}{(\log\log x)^{1/5}}}\Big) \qquad \text{(Vinogradov–Korobov, 1958)},$$

Perceba que

$$(\log x)^A \ll e^{c\sqrt{\log x}} \ll x^\varepsilon \quad \forall \varepsilon \ \forall A.$$

Essas cotas vêm de regiões livres de zeros de ζ à esquerda da reta $\mathrm{Re}(s) = 1$, como no esboço abaixo.



4.3.4 Contagem de zeros

Para $T > 0$, seja $N(T)$ o número de zeros de ζ na faixa crítica com parte imaginária entre 0 e T, *contados com multiplicidade*:

$$N(T) = \sum_{\substack{\rho:\ \zeta(\rho)=0 \\ 0<\mathrm{Re}(\rho)<1,\ 0<\mathrm{Im}(\rho)\le T}} m(\rho),$$

onde $m(\rho)$ é a multiplicidade do zero ρ. Então sabe-se que

$$N(T) \sim cT\log T, \qquad c = \frac{1}{2\pi},$$

e que uma proporção positiva destes zeros possui parte real igual a $\frac{1}{2}$. A contagem com multiplicidade é a natural aqui, pois é ela que o princípio do argumento fornece. Não se sabe se todos os zeros são simples: se fossem, $N(T)$ seria simplesmente o número de zeros. Esta é a mesma contagem do Teorema 5.15.

4.4 A constante de Mertens

Na Seção 3.3, provamos que existe uma constante c tal que

$$\prod_{p\le x}\Big(1-\frac{1}{p}\Big)=\frac{e^{-c}}{\log x}\Big(1+O\Big(\frac{1}{\log x}\Big)\Big).$$

Vamos identificá-la.

Teorema 4.9. *A constante acima é $c=\gamma$.*

Daremos duas demonstrações, depois de uma preparação comum.

4.4.1 Preparação

A fórmula acima é equivalente a

$$-\sum_{p\le x}\log\Big(1-\frac{1}{p}\Big)=\log\log x+c+O\Big(\frac{1}{\log x}\Big). \tag{4.1}$$

Além disso,

$$\log\Big(1-\frac{1}{p}\Big)=-\sum_m\frac{1}{mp^m}\quad\Rightarrow\quad -\sum_{p\le x}\log\Big(1-\frac{1}{p}\Big)=\sum_{p\le x}\sum_m\frac{1}{mp^m}.$$

A contribuição dos termos em que $p^m>x$ é $O(\frac{1}{\log x})$. De fato, como $p\le x$, esses termos têm $m\ge 2$; para cada $p\le\sqrt{x}$ eles somam no máximo $\sum_{m:\,p^m>x}p^{-m}\le\frac{2}{x}$, e para cada $p>\sqrt{x}$ somam no máximo $\sum_{m\ge2}p^{-m}=\frac{1}{p(p-1)}$, de modo que o total é $\le\sqrt{x}\cdot\frac{2}{x}+\sum_{n>\sqrt{x}}\frac{1}{n(n-1)}\ll\frac{1}{\sqrt{x}}$.

Com isso,

$$\begin{aligned}-\sum_{p\le x}\log\Big(1-\frac{1}{p}\Big)&=\sum_{p^m\le x}\frac{1}{mp^m}+O\Big(\frac{1}{\log x}\Big)\\&=\sum_{n\le x}\frac{f(n)}{n}+O\Big(\frac{1}{\log x}\Big),\end{aligned}$$

onde

$$f(n)=\begin{cases}\dfrac{1}{m} & \text{se } n=p^m,\\ 0 & \text{caso contrário.}\end{cases}$$

Daí,

$$T_f(x)=\sum_{n\le x}\frac{f(n)}{n}=\log\log x+c+O\Big(\frac{1}{\log x}\Big). \tag{4.2}$$

Além disso, $D_f(s)=\sum_{n\ge1}\frac{f(n)}{n^s}$ converge para $\mathrm{Re}(s)>1$ e

$$\log\zeta(\sigma)=\sum_p\sum_m\frac{1}{mp^{m\sigma}}=\sum_n\frac{f(n)}{n^\sigma}=D_f(\sigma)\qquad(\sigma>1).$$

Nas duas demonstrações, será preciso comparar expressões para $\log\zeta(\sigma)$ quando $\sigma\to1^+$, e para isso precisamos de um termo de erro que tenda a zero. Saber apenas que $\lim_{\sigma\to1^+}\zeta(\sigma)(\sigma-1)$ existe daria somente

$$\log\zeta(\sigma)=\log\Big(\frac{1}{\sigma-1}\Big)+O(1),$$

o que não basta. Usamos então o Teorema 4.5, do qual vem $\lim_{\sigma\to 1^+}\zeta(\sigma)(\sigma-1)=1$ e, mais precisamente,

$$\zeta(\sigma)=\frac{1}{\sigma-1}+O(1).$$

De fato, com isso, temos

$$\begin{aligned}\log\zeta(\sigma)&=\log\Big(\frac{1}{\sigma-1}+O(1)\Big)\\&=\log\Big(\frac{1}{\sigma-1}\Big)+\log(1+O(\sigma-1))\\&=\log\Big(\frac{1}{\sigma-1}\Big)+O(\sigma-1).\end{aligned}\tag{4.3}$$

4.4.2 O argumento de Hildebrand

Primeira demonstração do Teorema 4.9. Pela segunda fórmula do Teorema 4.3,

$$D_f(\sigma)=(\sigma-1)\int_1^{+\infty}\Big(\sum_{n\le x}\frac{f(n)}{n}\Big)x^{-\sigma}\,dx.$$

Como $f(1)=0$, podemos escrever

$$D_f(\sigma)=(\sigma-1)\int_2^{+\infty}\Big(\sum_{n\le x}\frac{f(n)}{n}\Big)x^{-\sigma}\,dx.$$

Usando (4.2),

$$D_f(\sigma)=(\sigma-1)\int_2^{+\infty}\Big(\log\log x+c+O\Big(\frac{1}{\log x}\Big)\Big)x^{-\sigma}\,dx.$$

Com a mudança de variável $u=\log x$,

$$D_f(\sigma)=(\sigma-1)\int_{\log 2}^{+\infty}\Big(\log u+c+O\Big(\frac{1}{u}\Big)\Big)e^{(1-\sigma)u}\,du.$$

Suponha que σ esteja próximo de 1, digamos $1<\sigma<1{,}49$ (talvez não precise tanto).

Vamos estimar o erro:

$$(\sigma-1)\int_{\log 2}^{+\infty}O\Big(\frac{1}{u}\Big)e^{(1-\sigma)u}\,du.$$

Com a mudança de variável $v=(\sigma-1)u$, ele fica

$$(\sigma-1)\int_{(\sigma-1)\log 2}^{+\infty}O\Big(\frac{1}{v}\Big)e^{-v}\,dv.$$

Dividindo a integral em duas,

$$\begin{aligned}&\ll(\sigma-1)\int_{(\sigma-1)\log 2}^{1}\frac{1}{v}e^{-v}\,dv+(\sigma-1)\int_1^{+\infty}\frac{1}{v}e^{-v}\,dv\\&\ll(\sigma-1)\int_{(\sigma-1)\log 2}^{1}\frac{dv}{v}+(\sigma-1)\int_1^{+\infty}e^{-v}\,dv\\&\le(\sigma-1)\Big(\log\Big(\frac{1}{(\sigma-1)\log 2}\Big)+1\Big)\\&\ll(\sigma-1)\log\Big(\frac{1}{\sigma-1}\Big).\end{aligned}\tag{4.4}$$

Ou seja, provamos que

$$D_f(\sigma) = (\sigma-1)\int_{\log 2}^{+\infty} (\log u + c)e^{(1-\sigma)u}\,du + O\Big((\sigma-1)\log\Big(\frac{1}{\sigma-1}\Big)\Big).$$

Completando a integral,

$$\begin{aligned} D_f(\sigma) = (\sigma-1)\int_{0}^{+\infty} (\log u + c)e^{(1-\sigma)u}\,du \\ + O\Big((\sigma-1)\log\Big(\frac{1}{\sigma-1}\Big)\Big) \\ + O\Big((\sigma-1)\int_0^{\log 2} (|\log u| + |c|)e^{(1-\sigma)u}\,du\Big). \end{aligned}$$

O último erro, com a mudança $u = 1/v$, é

$$\begin{aligned} O\Big((\sigma-1)\int_0^{\log 2} (|\log u| + |c|)\,du\Big) &= O\Big((\sigma-1)\int_{1/\log 2}^{+\infty} (\log v + |c|)\frac{dv}{v^2}\Big) \\ &= O(\sigma-1). \end{aligned}$$

Finalmente, temos

$$D_f(\sigma) = (\sigma-1)\int_{0}^{+\infty} (\log u + c)e^{(1-\sigma)u}\,du + O\Big((\sigma-1)\log\Big(\frac{1}{\sigma-1}\Big)\Big).$$

Fazendo $v = (\sigma-1)u$ nessa última integral, e usando que $\int_0^{+\infty} e^{-v}\,dv = 1$,

$$\begin{aligned} D_f(\sigma) &= \int_0^{+\infty} \Big(\log\Big(\frac{1}{\sigma-1}\Big) + \log v + c\Big)e^{-v}\,dv + O\Big((\sigma-1)\log\Big(\frac{1}{\sigma-1}\Big)\Big) \\ &= \log\Big(\frac{1}{\sigma-1}\Big) + I + c + O\Big((\sigma-1)\log\Big(\frac{1}{\sigma-1}\Big)\Big), \end{aligned}$$

onde $I = \int_0^{+\infty} (\log v)e^{-v}\,dv$.

Por outro lado, por (4.3),

$$D_f(\sigma) = \log\zeta(\sigma) = \log\Big(\frac{1}{\sigma-1}\Big) + O(\sigma-1).$$

Comparando as duas expressões e fazendo $\sigma \to 1^+$, encontramos $I + c = 0$, isto é, $c = -I = \gamma$. □

Esta demonstração só estaria completa se provássemos também a igualdade

$$\gamma = -\int_0^{+\infty} \log v\, e^{-v}\,dv,$$

que equivale a $\Gamma'(1) = -\gamma$ e não será demonstrada aqui. A segunda demonstração, a seguir, não depende dela.

O argumento seguinte prova $c = \gamma$ diretamente com a definição de γ dada na Seção 2.1, isto é,

$$\gamma = \lim_{x\to+\infty}\Big(\sum_{n\le x}\frac{1}{n} - \log x\Big).$$

4.4.3 O argumento de Tenenbaum

O argumento de Tenenbaum [Ten15] é semelhante, mas um pouco mais complicado, e envolve um truque.

Segunda demonstração do Teorema 4.9. O truque é o seguinte. Para x próximo de 0,

$$e^{-x} = 1 - x + O(x^2) \quad \Leftrightarrow \quad \frac{1-e^{-x}}{x} = 1 + O(x)$$

$$\Rightarrow \quad \log\Big(\frac{1}{x}\Big) - \log\Big(\frac{1}{1-e^{-x}}\Big) = O(x).$$

Usando essa observação em (4.3), obtemos

$$\begin{aligned}\log\zeta(\sigma) &= \log\Big(\frac{1}{\sigma-1}\Big) + O(\sigma-1)\\ &= \log\Big(\frac{1}{1-e^{1-\sigma}}\Big) + O(\sigma-1)\\ &= \sum_{n\geq 1}\frac{1}{ne^{n(\sigma-1)}} + O(\sigma-1).\end{aligned}$$

Usando soma por partes, a série acima é o mesmo que

$$(\sigma-1)\int_0^{+\infty}\Big(\sum_{n\leq t}\frac{1}{n}\Big)e^{-(\sigma-1)t}\,dt.$$

Considere agora a série de Dirichlet $\sum_p \frac{1}{p^s}$, que converge quando $\mathrm{Re}(s) > 1$. Mais uma vez por soma por partes,

$$\sum_p \frac{1}{p^\sigma} = (\sigma-1)\int_1^{+\infty}\Big(\sum_{p\leq u}\frac{1}{p}\Big)u^{-\sigma}\,du.$$

A ideia é comparar as duas séries. Lembre que

$$\sum_{n\leq t}\frac{1}{n} \sim \log t \quad \text{e} \quad \sum_{p\leq u}\frac{1}{p} \sim \log\log u.$$

Isso, e o fato de que a primeira integral tem exponencial e a segunda não tem, sugerem a mudança de variável $t = \log u$, isto é, $u = e^t$.

Fazendo isto, temos

$$\sum_p \frac{1}{p^\sigma} = (\sigma-1)\int_0^{+\infty}\Big(\sum_{p\leq e^t}\frac{1}{p}\Big)e^{-(\sigma-1)t}\,dt.$$

Comparando,

$$\log\zeta(\sigma) - \sum_p \frac{1}{p^\sigma} = (\sigma-1)\int_0^{+\infty} F(t)e^{-(\sigma-1)t}\,dt + O(\sigma-1),$$

onde

$$F(t) = \sum_{n\leq t}\frac{1}{n} - \sum_{p\leq e^t}\frac{1}{p}.$$

Seja A a constante de $\sum_{p\le x} 1/p = \log\log x + A + O(1/\log x)$, dada pela Seção 3.3. Se $t \ge \log 2$, então

$$\begin{aligned} F(t) &= \log t + \gamma + O\Big(\frac{1}{t}\Big) - \Big(\log t + A + O\Big(\frac{1}{t}\Big)\Big) \\ &= \gamma - A + O\Big(\frac{1}{t}\Big). \end{aligned}$$

Por outro lado, se $t < \log 2$, então $F(t) = 0$. Donde,

$$\begin{aligned} &\log\zeta(\sigma) - \sum_p \frac{1}{p^\sigma} \\ &\quad = (\sigma-1)\int_{\log 2}^{+\infty} \Big(\gamma - A + O\Big(\frac{1}{t}\Big)\Big) e^{-(\sigma-1)t}\,dt + O(\sigma-1). \end{aligned}$$

O termo de erro com $O(1/t)$ é, por (4.4), $O\big((\sigma-1)\log(\frac{1}{\sigma-1})\big)$.

Por outro lado, o restante é igual a

$$(\sigma-1)\int_{\log 2}^{+\infty} (\gamma - A) e^{-(\sigma-1)t}\,dt.$$

Fazendo $v = (\sigma-1)t$, isto é

$$(\gamma - A)\int_{(\sigma-1)\log 2}^{+\infty} e^{-v}\,dv.$$

A integral vale

$$\int_0^{+\infty} e^{-v}\,dv + O\Big(\int_0^{(\sigma-1)\log 2} 1\,dv\Big) = 1 + O(\sigma-1).$$

Conclusão até aqui:

$$\log\zeta(\sigma) - \sum_p \frac{1}{p^\sigma} = (\gamma - A) + O\Big((\sigma-1)\log\Big(\frac{1}{\sigma-1}\Big)\Big).$$

Lembrando que

$$\log\zeta(\sigma) = \sum_p \log\Big(\Big(1 - \frac{1}{p^\sigma}\Big)^{-1}\Big),$$

segue que

$$\begin{aligned} \gamma - A &= \lim_{\sigma\to 1^+} \Big(\log\zeta(\sigma) - \sum_p \frac{1}{p^\sigma}\Big) \\ &= \lim_{\sigma\to 1^+} \sum_p \Big(\log\Big(\Big(1-\frac{1}{p^\sigma}\Big)^{-1}\Big) - \frac{1}{p^\sigma}\Big) \\ &= \sum_p \Big(\log\Big(\Big(1-\frac{1}{p}\Big)^{-1}\Big) - \frac{1}{p}\Big). \end{aligned}$$

A troca da soma e do limite se justifica pois

$$\log\Big(\frac{1}{1-\frac{1}{p^\sigma}}\Big) - \frac{1}{p^\sigma} = O\Big(\frac{1}{p^{2\sigma}}\Big) = O\Big(\frac{1}{p^2}\Big).$$

Agora temos tudo de que precisamos.

Por (4.1),

$$c = \lim_{x\to+\infty} \Big(\sum_{p\le x} \log\Big(\Big(1-\frac{1}{p}\Big)^{-1}\Big) - \log\log x\Big).$$

Escreva isso como $\ell_1 + \ell_2$, onde

$$\ell_1 = \lim_{x\to+\infty} \sum_{p\le x} \Big(\log\Big(\Big(1-\frac{1}{p}\Big)^{-1}\Big) - \frac{1}{p}\Big)$$

e

$$\ell_2 = \lim_{x\to+\infty} \Big(\sum_{p\le x} \frac{1}{p} - \log\log x\Big).$$

Por definição, $\ell_2 = A$ e, pelo que mostramos, $\ell_1 = \gamma - A$. Portanto, $c = \ell_1 + \ell_2 = \gamma$. □

4.5 Inversão de Mellin

Vamos agora usar $D_f(s)$ para obter informações sobre $S_f(x)$:

$S_f(x) \longrightarrow D_f(s)$	Transformada de Mellin,
$D_f(s) \longrightarrow S_f(x)$	Inversão de Mellin.

O modelo é a transformada de Fourier,

$$\hat{f}(y) := \int f(x) e^{-2\pi i x y}\, dx,$$

com sua fórmula de inversão,

$$f(x) = \int \hat{f}(y) e^{2\pi i x y}\, dy.$$

Seja $f\colon (0,+\infty) \to \mathbb{C}$ uma função suave, de suporte compacto em $(0,+\infty)$. Então a transformada de Mellin de f,

$$\tilde{f}(s) = \int_0^{+\infty} f(x) x^{-s-1}\, dx,$$

converge para todo $s \in \mathbb{C}$, e vale a fórmula de inversão: para todo $\sigma \in \mathbb{R}$,

$$f(x) = \frac{1}{2\pi i} \int_{(\sigma)} \tilde{f}(s)\, x^s\, ds,$$

onde (σ) é a reta $\mathrm{Re}(s) = \sigma$, orientada para cima.

De fato, escreva $s = \sigma + i\tau$. Com a substituição $x = e^u$,

$$\begin{aligned}\tilde{f}(s) &= \int_{\mathbb{R}} f(e^u) e^{-su}\, du = \int_{\mathbb{R}} f(e^u) e^{-\sigma u}\cdot e^{-i\tau u}\, du \\ &= \hat{g}\Big(\frac{\tau}{2\pi}\Big),\end{aligned}$$

onde $g(u) = f(e^u)e^{-\sigma u}$. Pela inversão de Fourier,

$$\begin{aligned}g(u) &= \int_{\mathbb{R}} \hat{g}(v)\, e^{2\pi i u v}\, dv \\ &= \int_{\mathbb{R}} \tilde{f}(\sigma + 2\pi i v)\, e^{2\pi i u v}\, dv,\end{aligned}$$

isto é,

$$f(e^u)e^{-\sigma u} = \int_{\mathbb{R}} \tilde{f}(\sigma + 2\pi i v)\, e^{2\pi i u v}\, dv.$$

Voltando a x, com $u = \log x$,

$$\frac{f(x)}{x^\sigma} = \int_{\mathbb{R}} \tilde{f}(\sigma + 2\pi i v)\, x^{2\pi i v}\, dv,$$

ou seja,

$$f(x) = \int_{\mathbb{R}} \tilde{f}(\sigma + 2\pi i v)\, x^{\sigma + 2\pi i v}\, dv.$$

A mudança de variável $s = \sigma + 2\pi i v$ dá a fórmula enunciada.

A função $S_f(x)$ não tem suporte compacto em $(0, +\infty)$, mas tem suporte "longe" de 0. Mais precisamente,

$$\operatorname{supp}(S_f) \subseteq [1, +\infty).$$

Por isso, valerá uma fórmula de inversão de Mellin para σ suficientemente grande (ou pequeno, se mudarmos a convenção).

4.6 A fórmula de Perron

Pelo Teorema 4.3,

$$D_f(s) = s \int_0^{+\infty} S_f(x)\, x^{-s-1}\, dx.$$

Vamos provar que existe uma inversão desta fórmula, a fórmula de Perron (Teorema 4.14). Antes dela, vamos provar um resultado similar para a soma, um pouco mais suave,

$$S_f^1(x) = \sum_{n \le x} f(n)(x - n).$$

4.6.1 A fórmula de Perron para $S_f^1(x)$

Teorema 4.10 (Fórmula de Perron para $S_f^1(x)$)**.** *Seja f uma função aritmética e suponha que $\sigma_a = \sigma_a(f)$ seja finita. Então, para todo $c > \max(0, \sigma_a)$ e todo $x \ge 1$,*

$$S_f^1(x) = \frac{1}{2\pi i} \int_{(c)} D_f(s)\, \frac{x^{s+1}}{s(s+1)}\, ds.$$

Divagação: *sharp cut-off*

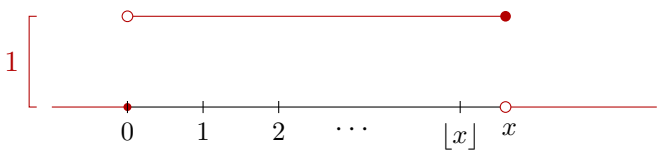
A soma $S_f(x) = \sum_{n \le x} f(n)$ pode ser escrita como

$$S_f(x) = \sum_{n \ge 1} f(n)\Omega(n),$$

onde $\Omega = \Omega_x$ é a função

$$\Omega(t) = \begin{cases} 1 & \text{se } 0 < t \le x, \\ 0 & \text{caso contrário,} \end{cases}$$

cujo gráfico é



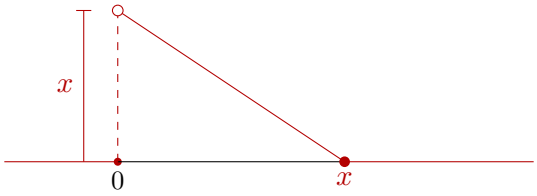
Se não quisermos que Ω dependa de x, basta observar que

$$\Omega_x(t) = \Omega_1\Big(\frac{t}{x}\Big).$$

Por outro lado,

$$\sum_{n\le x} f(n)(x-n) = \sum_{n\ge 1} f(n)\Omega(n),$$

onde agora $\Omega = \Omega_x$ é a função cujo gráfico é



Essa função já é um pouco melhor: não há descontinuidade em $t = x$. Mais uma vez, se não quisermos que Ω dependa de x, note que

$$\Omega_x(t) = x\,\Omega_1\Big(\frac{t}{x}\Big).$$

Divagação: *smooth cut-off*

Seja $\Omega\colon \mathbb{R}_{>0} \to \mathbb{C}$ uma função suave de suporte compacto em $(0, +\infty)$. Podemos considerar a soma

$$S(f,\Omega) = \sum_{n\ge 1} f(n)\Omega(n).$$

Se, em algum sentido, Ω aproximar a função característica de $[0, x]$, dizemos informalmente que $S(f,\Omega)$ é uma "versão suave" de $S_f(x)$.

Demonstração da fórmula para $S_f^1(x)$

A demonstração se apoia no lema seguinte.

Lema 4.11. *Para $c > 0$ e $y > 0$,*

$$\frac{1}{2\pi i}\int_{(c)} \frac{y^s}{s(s+1)}\,ds = \begin{cases} 1 - \dfrac{1}{y} & \text{se } y > 1, \\ 0 & \text{se } 0 < y \le 1. \end{cases} \qquad (\bigstar)$$

Demonstração do Teorema 4.10. Ignorando, por enquanto, problemas de convergência,

$$\begin{aligned} &\frac{1}{2\pi i}\int_{(c)} D_f(s)\,\frac{x^{s+1}}{s(s+1)}\,ds \\ &= \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \sum_{n\ge 1} \frac{f(n)}{n^s}\,\frac{x^{s+1}}{s(s+1)}\,ds \\ &= \frac{1}{2\pi i}\sum_{n\ge 1} f(n)\int_{(c)} \frac{(x/n)^s\,x}{s(s+1)}\,ds \\ &= \sum_{n\ge 1} f(n)\,x\Big(\frac{1}{2\pi i}\int_{(c)} \frac{(x/n)^s}{s(s+1)}\,ds\Big). \end{aligned}$$

Então, além de justificar a troca de ordem da série e da integral, basta aplicar o Lema 4.11.

Vamos justificar a troca de ordem da $\sum$ e da $\int$. Escrevendo $s = c + it$,

$$\frac{1}{2\pi i}\int_{(c)} \sum_{n\geq 1} \frac{f(n)}{n^s} \cdot \frac{x^{s+1}}{s(s+1)}\, ds$$
$$= \frac{1}{2\pi}\int_{\mathbb{R}} \sum_{n\geq 1} \frac{f(n)}{n^{c+it}} \cdot \frac{x^{c+it+1}}{(c+it)(c+it+1)}\, dt.$$

Para justificar a troca, basta mostrar que

$$\int_{\mathbb{R}} \sum_{n\geq 1} \Big| \frac{f(n)}{n^{c+it}} \cdot \frac{x^{c+it+1}}{(c+it)(c+it+1)} \Big|\, dt$$
$$= \int_{\mathbb{R}} \sum_{n\geq 1} \frac{|f(n)|\, x\,(x/n)^c}{|(c+it)(c+it+1)|}\, dt$$

converge. De fato, a combinação $\int \sum$ acima é limitada por

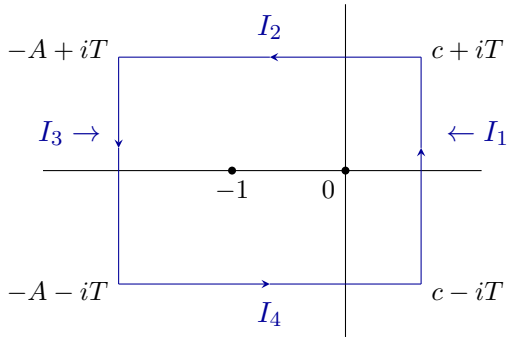
$$x^{c+1} \cdot \underbrace{\Big(\sum_{n\geq 1} \frac{|f(n)|}{n^c}\Big)}_{\substack{\text{converge pois} \\ c>\sigma_a(f)}} \cdot \underbrace{\int_{\mathbb{R}} \frac{dt}{|c+it|\cdot|c+it+1|}}_{\substack{\text{converge por} \\ \text{comparação com } \int_{-\infty}^{\infty} \frac{dt}{c^2+t^2}}} \cdot$$ □

Demonstração do Lema 4.11, caso $y > 1$. A identidade será provada usando o teorema dos resíduos para a função

$$s \longmapsto \frac{y^s}{s(s+1)}.$$

Esta função tem polos em $s = 0$ e $s = -1$, com resíduos iguais a 1 em $s = 0$ e a $-\dfrac{1}{y}$ em $s = -1$.

Seja R o contorno do retângulo (no sentido anti-horário) de vértices $c-iT$, $c+iT$, $-A+iT$ e $-A-iT$, onde A e T são constantes positivas com $A > 1$.



Pelo teorema dos resíduos,

$$\frac{1}{2\pi i}\int_R \frac{y^s}{s(s+1)}\, ds = 1 - \frac{1}{y}.$$

Escreva

$$\int_R \frac{y^s}{s(s+1)}\, ds = I_1 + I_2 + I_3 + I_4,$$

onde cada termo é a integral sobre o segmento indicado na figura.

Primeiro vamos estimar I_3:

$$I_3 = -i\int_{-T}^{T} \frac{y^{-A+it}}{(-A+it)(-A+1+it)}\, dt,$$

donde, se $A > 2$,

$$\begin{aligned}|I_3| &\le \int_{-T}^{T} \frac{y^{-A}}{|A-1|^2+t^2}\,dt \\ &\le y^{-A}\int_{-\infty}^{+\infty} \frac{dt}{|A-1|^2+t^2} \\ &\le y^{-A}\int_{-\infty}^{+\infty} \frac{dt}{1+t^2} \ll y^{-A}.\end{aligned}$$

Por outro lado,

$$I_2 = -\int_{-A}^{c} \frac{y^{x+iT}}{(x+iT)(x+1+iT)}\,dx,$$

donde

$$|I_2| \le \int_{-A}^{c} \frac{y^x}{T^2}\,dx \le \frac{1}{T^2}\int_{-\infty}^{c} y^x\,dx = \frac{y^c}{\log y\cdot T^2} \ll \frac{1}{T^2}.$$

A estimativa de I_4 é análoga: $|I_4| \ll \dfrac{1}{T^2}$. Assim, tomando A e T grandes,

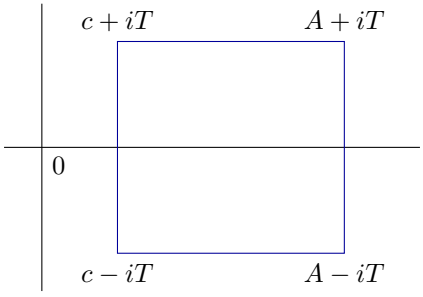
$$\begin{aligned}\frac{1}{2\pi i}\int_{c-iT}^{c+iT} \frac{y^s}{s(s+1)}\,ds &= 1 - \frac{1}{y} - \frac{1}{2\pi i}(I_2+I_3+I_4) \\ &= 1 - \frac{1}{y} + O\Big(y^{-A} + \frac{1}{T^2}\Big).\end{aligned}$$

Fazendo $A \to +\infty$,

$$\frac{1}{2\pi i}\int_{c-iT}^{c+iT} \frac{y^s}{s(s+1)}\,ds = 1 - \frac{1}{y} + O\Big(\frac{1}{T^2}\Big).$$

Fazendo $T \to +\infty$, temos o resultado para $y > 1$. □

Demonstração do Lema 4.11, caso $0 < y < 1$. Neste caso, tomamos um retângulo "para a direita", de vértices $c - iT$, $c + iT$, $A + iT$, $A - iT$ (no sentido horário).



Agora não há polos no interior de R. Portanto, o teorema dos resíduos diz que

$$\int_R \frac{y^s}{s(s+1)}\,ds = 0.$$

O resto segue de maneira análoga. □

Demonstração do Lema 4.11, caso $y = 1$. O argumento acima só funciona plenamente se $y > 1$ ou $y < 1$. No caso $y = 1$, fazemos a conta diretamente. Com $s = c + it$, $t \in [-T, T]$,

$$\begin{aligned}\int_{(c)} \frac{ds}{s(s+1)} &= \lim_{T\to+\infty}\int_{c-iT}^{c+iT}\Big(\frac{1}{s} - \frac{1}{s+1}\Big)ds \\ &= \lim_{T\to+\infty} i\int_{-T}^{T}\Big(\frac{1}{c+it} - \frac{1}{(c+1)+it}\Big)dt.\end{aligned}$$

Ora,

$$i\int_{-T}^{T}\Big(\frac{1}{c+it}-\frac{1}{(c+1)+it}\Big)dt = i\int_{-T}^{T}\frac{c-it}{c^2+t^2}\,dt - i\int_{-T}^{T}\frac{(c+1)-it}{(c+1)^2+t^2}\,dt,$$

e os termos com it no numerador são funções ímpares de t, logo suas integrais se anulam. Resta

$$ci\int_{-T}^{T}\frac{dt}{c^2+t^2} - (c+1)i\int_{-T}^{T}\frac{dt}{(c+1)^2+t^2}.$$

Com $t = c\cdot u$ na primeira integral e $t=(c+1)u$ na segunda, isto é

$$\begin{aligned} &ci\int_{-T/c}^{T/c}\frac{c\,du}{c^2+c^2u^2} - (c+1)i\int_{-T/(c+1)}^{T/(c+1)}\frac{(c+1)\,du}{(c+1)^2+(c+1)^2u^2}\\ &= i\int_{-T/c}^{T/c}\frac{du}{1+u^2} - i\int_{-T/(c+1)}^{T/(c+1)}\frac{du}{1+u^2}. \end{aligned}$$

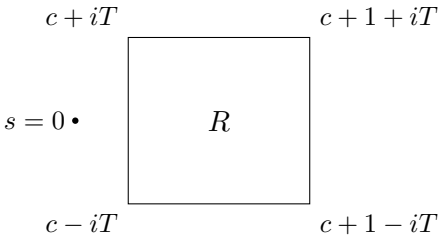
Agora, fazendo $T\to+\infty$,

$$\int_{(c)}\frac{ds}{s(s+1)} = i\int_{-\infty}^{+\infty}\frac{du}{1+u^2} - i\int_{-\infty}^{+\infty}\frac{du}{1+u^2} = 0. \qquad \square$$

Observação 4.12. Aqui não precisamos do valor, mas vale lembrar que

$$\int\frac{du}{1+u^2} = \operatorname{arctg}(u), \quad \text{logo} \quad \int_{-\infty}^{+\infty}\frac{du}{1+u^2} = \pi.$$

Observação 4.13 (Método alternativo para $y=1$). Sejam $c>0$, $g(s)=1/s$ e R o retângulo abaixo, que não contém o polo $s=0$ de g.



A ideia é aplicar o teorema dos resíduos a g em R, observando que

$$\int_{c-iT}^{c+iT}\frac{ds}{s+1} = \int_{c+1-iT}^{c+1+iT}\frac{ds}{s}.$$

4.6.2 A fórmula de Perron para $S_f(x)$

Teorema 4.14 (Fórmula de Perron para $S_f(x)$)**.** *Seja f uma função aritmética e suponha que a abscissa de convergência absoluta $\sigma_a = \sigma_a(f)$ seja finita. Então, para todo $c > \max(0,\sigma_a)$ e todo $x>1$, $x\notin\mathbb{Z}$,*

$$S_f(x) = \underbrace{\frac{1}{2\pi i}\int_{(c)} D_f(s)\,\frac{x^s}{s}\,ds}_{I(c)}.$$

Essa integral não converge absolutamente. Ela deve ser interpretada como

$$\lim_{T\to+\infty}\underbrace{\frac{1}{2\pi i}\int_{c-iT}^{c+iT} D_f(s)\,\frac{x^s}{s}\,ds}_{I(c,T)}.$$

Além disso, dado $T > 0$,

$$I(c) = I(c,T) + O\Big(\frac{x^c}{T}\sum_{n\geq 1}\frac{|f(n)|}{n^c|\log(x/n)|}\Big).$$

A demonstração requer um lema análogo ao Lema 4.11.

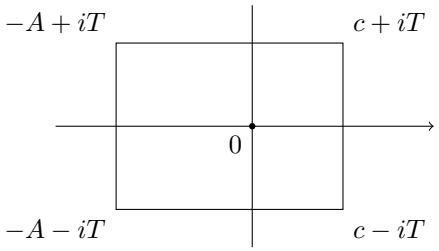
Lema 4.15. *Seja $c > 0$ e seja*

$$J(y,T) = \frac{1}{2\pi i}\int_{c-iT}^{c+iT}\frac{y^s}{s}\,ds.$$

Dados $T > 0$ e $y > 0$, $y \neq 1$, temos

$$J(y,T) = \begin{cases} 1 + O\Big(\dfrac{y^c}{T\log y}\Big) & \text{se } y > 1, \\ O\Big(\dfrac{y^c}{T|\log y|}\Big) & \text{se } 0 < y < 1. \end{cases}$$

Demonstração do Lema 4.15. Suponha primeiro $y > 1$. Considere o mesmo tipo de retângulo usado na demonstração do Lema 4.11:



A função $\dfrac{y^s}{s}$ só possui um polo, em $s = 0$, com resíduo igual a 1. Portanto, pelo teorema dos resíduos,

$$\frac{1}{2\pi i}\int_R \frac{y^s}{s}\,ds = 1.$$

Mais uma vez escrevendo

$$\int_R = \underbrace{\int_{c-iT}^{c+iT}}_{I_1} + \underbrace{\int_{c+iT}^{-A+iT}}_{I_2} + \underbrace{\int_{-A+iT}^{-A-iT}}_{I_3} + \underbrace{\int_{-A-iT}^{c-iT}}_{I_4},$$

queremos estimar I_2, I_3, I_4. Primeiramente,

$$I_2 = -\int_{-A+iT}^{c+iT}\frac{y^s}{s}\,ds = -\int_{-A}^{c}\frac{y^{x+iT}}{x+iT}\,dx,$$

donde

$$\begin{aligned} |I_2| &\leq \int_{-A}^{c}\frac{y^x}{T}\,dx \\ &= \frac{1}{T}\Big(\frac{y^x}{\log y}\Big)\Big|_{-A}^{c} \\ &\leq \frac{y^c}{T\log y}. \end{aligned}$$

Analogamente, $|I_4| \leq \dfrac{y^c}{T \log y}$. Por outro lado,

$$I_3 = -\int_{-A-iT}^{-A+iT} \frac{y^s}{s}\, ds = -i \int_{-T}^{T} \frac{y^{-A+it}}{-A+it}\, dt,$$

donde

$$|I_3| \leq \frac{2Ty^{-A}}{A}.$$

Juntando essas estimativas,

$$\frac{1}{2\pi i} \int_{c-iT}^{c+iT} \frac{y^s}{s}\, ds = 1 + O\Big(\frac{y^c}{T \log y} + \frac{2Ty^{-A}}{A}\Big).$$

Fazendo $A \to +\infty$,

$$\frac{1}{2\pi i} \int_{c-iT}^{c+iT} \frac{y^s}{s}\, ds = 1 + O\Big(\frac{y^c}{T \log y}\Big).$$

O caso $0 < y < 1$ é análogo: basta tomar o retângulo "para a direita", e não teremos polo dentro do retângulo. □

Demonstração do Teorema 4.14. Supondo que possamos trocar a $\sum$ e a $\int$,

$$\begin{aligned} I(c,T) &= \sum_{n \geq 1} f(n) \cdot \frac{1}{2\pi i} \int_{c-iT}^{c+iT} \frac{(x/n)^s}{s}\, ds \\ &= \sum_{n \geq 1} f(n) \cdot J(x/n, T). \end{aligned}$$

Pelo Lema 4.15, e lembrando que $x \notin \mathbb{Z}$ (o que significa que x/n nunca é igual a 1), temos

$$\begin{aligned} I(c,T) &= \sum_{n \leq x} f(n)\Big(1 + O\Big(\frac{(x/n)^c}{T|\log(x/n)|}\Big)\Big) \\ &\quad + \sum_{n > x} f(n) \cdot O\Big(\frac{(x/n)^c}{T|\log(x/n)|}\Big) \\ &= S_f(x) + O\Big(\frac{x^c}{T} \sum_{n \geq 1} \frac{|f(n)|}{n^c |\log(x/n)|}\Big). \end{aligned}$$

Para justificar a troca de $\sum$ e $\int$, basta mostrar que a integral

$$\int_{c-iT}^{c+iT} \sum \frac{f(n)}{n^s} \cdot \frac{x^s}{s}\, ds$$

converge absolutamente, ou seja, que é finito

$$x^c \int_{-T}^{T} \sum_{n \geq 1} \frac{|f(n)|}{n^c} \frac{1}{\sqrt{c^2+t^2}}\, dt = x^c \Big(\sum_{n \geq 1} \frac{|f(n)|}{n^c}\Big) \cdot \int_{-T}^{T} \frac{dt}{\sqrt{c^2+t^2}}.$$

Isso ocorre, pois a última integral é limitada por uma constante dependendo de T.

Isso prova a "versão truncada" da fórmula de Perron. Para obter a versão não truncada, basta fazer $T \to +\infty$. De fato, o termo de erro é $O(\frac{x^c}{T} \sum_{n \geq 1} \frac{|f(n)|}{n^c |\log(x/n)|})$, e a série converge. Para $n > 2x$ temos $|\log(x/n)| > \log 2$, e $\sum |f(n)| n^{-c} < \infty$ porque $c > \sigma_a$; os demais termos são em número finito e têm $\log(x/n) \neq 0$, porque $x \notin \mathbb{Z}$. Logo o termo de erro é $O(1/T)$, com constante dependendo de x, e tende a 0. Portanto,

$$S_f(x) = \frac{1}{2\pi i} \int_{(c)} D_f(s)\, \frac{x^s}{s}\, ds.$$ □

4.6.3 O caso em que x é inteiro

Quando $x \in \mathbb{Z}$, vale uma versão não truncada do Teorema 4.14, mas para isso é necessário fazer uma modificação. Defina

$$\sum_{n\leq x}{}' f(n) = \begin{cases} \displaystyle\sum_{n\leq x} f(n) & \text{se } x \notin \mathbb{Z}, \\ \displaystyle\sum_{n\leq x-1} f(n) + \frac{f(x)}{2} & \text{se } x \in \mathbb{Z}. \end{cases}$$

Com esta definição, vale

$$\sum_{n\leq x}{}' f(n) = \frac{1}{2\pi i}\int_{(c)} D_f(s)\,\frac{x^s}{s}\,ds.$$

Observação 4.16. Equivalentemente,

$$\sum_{n\leq x}{}' f(n) = \lim_{\varepsilon\to 0^+}\Big(\frac{S_f(x-\varepsilon)+S_f(x+\varepsilon)}{2}\Big).$$

Observação 4.17. Isto está relacionado ao fato de que

$$\frac{1}{2\pi i}\int_{(c)} \frac{y^s}{s}\,ds = \begin{cases} 1 & \text{se } y > 1, \\ \frac{1}{2} & \text{se } y = 1, \\ 0 & \text{se } 0 < y < 1. \end{cases}$$

Capítulo 5

O Teorema dos Números Primos

5.1 A função zeta e o TNP: heurística

As fórmulas de Mellin e de Perron (Seções 4.5 e 4.6) permitem relacionar

$$\sum_{n\leq x} f(n) \rightleftarrows D_f(s).$$

Assim, para estudar números primos, devemos estudar

$$P(s) = \sum_p \frac{1}{p^s}.$$

Porém esta função não possui continuação meromorfa em torno de $s = 1$.

Aqui entram as funções $p \mapsto \log p$ e $\Lambda(n)$, que nos levam a considerar

$$\sum_p \frac{\log p}{p^s} \quad \text{e} \quad \sum_{n\geq 1} \frac{\Lambda(n)}{n^s} = -\frac{\zeta'(s)}{\zeta(s)}$$

(lembre que $\Lambda = \log * \mu$). A segunda função possui extensão meromorfa, que segue da extensão de ζ (Seção 4.3). Ao contrário de $P(s)$, a primeira também se estende em torno de $s = 1$: a diferença

$$-\frac{\zeta'(s)}{\zeta(s)} - \sum_p \frac{\log p}{p^s} = \sum_p \sum_{m\geq 2} \frac{\log p}{p^{ms}}$$

é holomorfa em $\mathrm{Re}(s) > \frac{1}{2}$, de modo que $\sum_p \frac{\log p}{p^s}$ tem continuação meromorfa a essa região, com polo simples em $s = 1$.

Assim, o TNP em sua forma

$$\psi(x) = \sum_{n\leq x} \Lambda(n) \sim x$$

vai ser provado usando informações analíticas sobre $-\frac{\zeta'(s)}{\zeta(s)}$.

Em particular, pela fórmula de Perron,

$$\sum_{n\leq x} \Lambda(n) = \frac{1}{2\pi i} \int_{(c)} -\frac{\zeta'(s)}{\zeta(s)} \frac{x^s}{s}\, ds \quad (x \notin \mathbb{Z}).$$

Se, por acaso, soubéssemos que ζ não se anula na região $\mathrm{Re}(s) \geq 1 - \delta$, poderíamos mover o contorno de integração para a esquerda, de (c) para $(1 - \delta)$, passando por um único polo em $s = 1$.

Veja que

$$\zeta(s) = \frac{1}{s-1} + g(s) \quad \text{(onde } g(s) \text{ é holomorfa).}$$

Consequentemente,

$$\zeta'(s) = \frac{-1}{(s-1)^2} + g'(s).$$

Portanto, a derivada logarítmica é

$$-\frac{\zeta'(s)}{\zeta(s)} = -\frac{\frac{-1}{(s-1)^2} + g'(s)}{\frac{1}{s-1} + g(s)} = \frac{1}{s-1} + h(s) \quad \text{(onde } h(s) \text{ é holomorfa),}$$

e

$$\text{Res}_{s=1}\left(-\frac{\zeta'(s)}{\zeta(s)}\frac{x^s}{s}\right) = \lim_{s\to 1}(s-1)\left(\frac{1}{s-1} + h(s)\right)\frac{x^s}{s} = x.$$

Após mover o contorno, obtemos

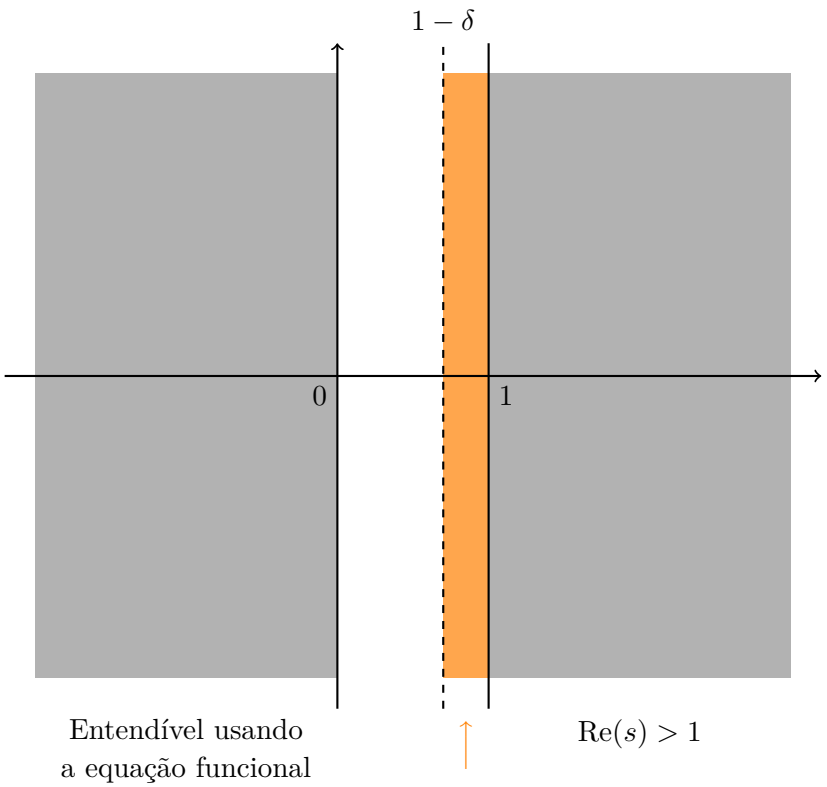
$$\sum_{n\leq x}\Lambda(n) = x - \frac{1}{2\pi i}\int_{(1-\delta)}\frac{\zeta'(s)}{\zeta(s)}\frac{x^s}{s}\,ds.$$

Supondo ainda que

$$\int_{\mathbb{R}}\frac{|\zeta'(1-\delta+it)|}{|\zeta(1-\delta+it)|}\,\frac{dt}{|1-\delta+it|} < +\infty,$$

teríamos

$$\psi(x) = x + O(x^{1-\delta}).$$

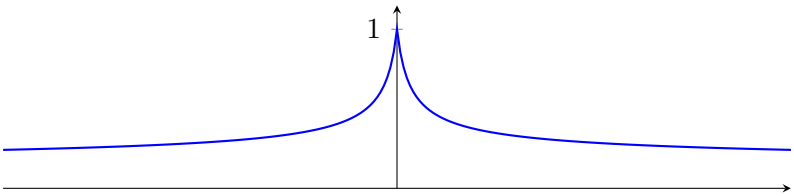


De la Vallée Poussin provou que existe $c_1 > 0$ tal que $\zeta(s) \neq 0$ para $s = \sigma + it$ na região

$$\sigma > 1 - \frac{c_1}{\log(|t| + e)},$$

cuja fronteira é controlada pela função $r(t) = \frac{1}{\log(|t|+e)}$:

Gráfico de $r(x) = \frac{1}{\log(|x|+e)}$



Isso leva a uma versão efetiva do TNP com um termo de erro:

$$\psi(x) = x + O\big(xe^{-c\sqrt{\log x}}\big).$$

Este resultado não será demonstrado aqui (ver a Seção 5.7). A comparação de ordens de magnitude é a seguinte: para todo $\varepsilon > 0$ e $A > 0$,

$$(\log x)^A \ll e^{c\sqrt{\log x}} \ll x^\varepsilon \quad (\text{para } x \to \infty).$$

Assim, para provar o TNP, basta mostrar que $\zeta(\sigma + it) \neq 0$ para valores de σ um pouco menores do que 1 e também para $\sigma = 1$. Vamos começar com o caso $\sigma = 1$.

5.2 Não-anulação de ζ na reta $\mathrm{Re}(s) = 1$

Teorema 5.1 (Hadamard, de la Vallée Poussin)**.**

$$\zeta(1 + it) \neq 0 \quad \textit{para todo} \quad t \neq 0.$$

A demonstração usa a seguinte desigualdade trigonométrica (3-4-1):

$$3 + 4\cos\theta + \cos 2\theta \geq 0.$$

Demonstração da desigualdade. Usando que $\cos 2\theta = 2\cos^2\theta - 1$,

$$\begin{aligned} 3 + 4\cos\theta + \cos 2\theta &= 2 + 4\cos\theta + 2\cos^2\theta \\ &= 2(1 + \cos\theta)^2 \geq 0. \end{aligned}$$

□

Motivação para o uso da desigualdade (Davenport). Relembre a fórmula da Seção 4.4, com $s = \sigma + it$, $\sigma > 1$:

$$\log\zeta(s) = \sum_p \sum_{m\geq 1} \frac{1}{m} p^{-ms} = \sum_p \sum_{m\geq 1} \frac{1}{m} p^{-m\sigma} e^{-imt\log p}. \tag{5.1}$$

Se $\zeta(1 + it) = 0$, fazendo $\sigma \to 1$, a parte real da série tenderia para $-\infty$. Isto significaria que para vários valores de p e m, $\cos(mt\log p)$ seria negativo. Se esse valor fosse suficientemente negativo, teríamos $\cos(2mt\log p)$ positivo, fazendo com que a parte real da série em $\sigma + 2it$ tendesse para $+\infty$. Mas isso indicaria um polo em $1 + 2it$, o que já sabemos que não ocorre.

Demonstração do Teorema 5.1. Considere a função

$$F(\sigma) = \zeta(\sigma)^3\zeta(\sigma+it)^4\zeta(\sigma+2it).$$

Tomando o logaritmo de $F(\sigma)$ e usando mais uma vez (5.1):

$$\begin{aligned}\log F(\sigma) &= 3\log\zeta(\sigma) + 4\log\zeta(\sigma+it) + \log\zeta(\sigma+2it)\\ &= \sum_p\sum_{m\geq 1}\frac{1}{m}p^{-m\sigma}\left(3+4e^{-imt\log p}+e^{-2imt\log p}\right).\end{aligned}$$

Agora, tomamos a parte real. Note que $\mathrm{Re}(e^{-ik\theta}) = \cos(k\theta)$. Logo

$$\mathrm{Re}(\log F(\sigma)) = \sum_p\sum_{m\geq 1}\frac{1}{m}p^{-m\sigma}\left(3+4\cos(mt\log p)+\cos(2mt\log p)\right).$$

Pela desigualdade 3-4-1 (com $\theta = mt\log p$), o termo entre parênteses é ≥ 0. Logo

$$\log|F(\sigma)| = \mathrm{Re}(\log F(\sigma)) \geq 0, \quad \text{isto é,} \quad |F(\sigma)| \geq 1 \quad (\sigma > 1,\ t\in\mathbb{R}). \tag{5.2}$$

Suponha agora que $\zeta(1+it) = 0$ para algum $t \neq 0$. Isto significa que para s próximo de $1+it$,

$$\zeta(s) = (s-1-it)g(s),$$

onde $g(s)$ é holomorfa. Assim, perto de $1+it$,

$$\zeta(s) = O(|s-1-it|).$$

Por outro lado, o polo simples de $\zeta(s)$ em $s=1$ implica que perto de $s=1$,

$$\zeta(s) = O\left(\frac{1}{|s-1|}\right).$$

Finalmente, como ζ é holomorfa em $s = 1+2it$, temos que perto de $1+2it$,

$$\zeta(s) = O(1).$$

Quando $|\sigma-1| < \varepsilon$, também $|(\sigma+it)-(1+it)|$, $|(\sigma+2it)-(1+2it)| < \varepsilon$.

Assim, para σ próximo de 1,

$$|F(\sigma)| = O\left(\frac{1}{|\sigma-1|^3}\right)\cdot O\left(|\sigma-1|^4\right)\cdot O(1) = O(|\sigma-1|).$$

Mas isto implicaria que

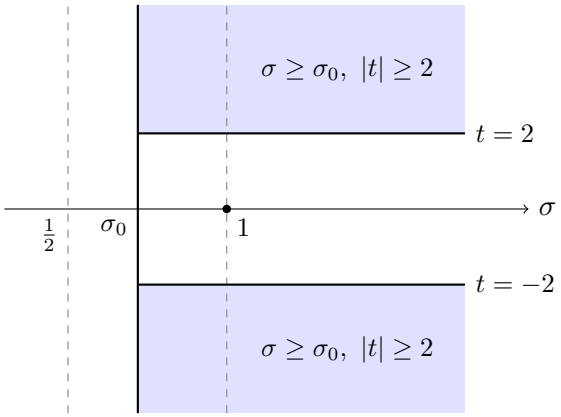
$$\lim_{\sigma\to 1^+}|F(\sigma)| = 0,$$

o que contradiz (5.2). Portanto, $\zeta(1+it)\neq 0$. □

5.3 Cotas superiores para ζ e ζ'

Voltando ao argumento heurístico da Seção 5.1, uma vez que tenhamos uma região sem zeros, também precisaremos estimar a integral de

$$-\frac{\zeta'(s)}{\zeta(s)}\frac{x^s}{s}.$$



A região sombreada é $\{\sigma \geq \sigma_0,\ |t| \geq 2\}$, onde valem as cotas do item (i) do teorema a seguir; o ponto marcado é o polo de ζ em $s = 1$.

Para esta parte do argumento, vamos usar o teorema a seguir, com a notação $s = \sigma + it$.

Teorema 5.2 (Cotas superiores para ζ e ζ')**.** *Temos as seguintes desigualdades:*

$$(i) \quad |\zeta(s)| \leq 4\frac{|t|^{1-\sigma_0}}{1-\sigma_0}, \qquad |t| \geq 2,\ \tfrac{1}{2} \leq \sigma_0 < 1,\ \sigma \geq \sigma_0,$$

$$(ii) \quad |\zeta(s)| \leq A_1 \log|t|, \qquad |t| \geq 2,\ \sigma \geq 1 - \frac{1}{4\log|t|},$$

$$(iii) \quad |\zeta'(s)| \leq A_2(\log|t|)^2, \qquad |t| \geq 2,\ \sigma \geq 1 - \frac{1}{12\log|t|},$$

onde A_1 e A_2 são constantes.

Demonstração. **Demonstração do item** (i)**.** Vamos usar a expressão

$$\zeta(s) = \frac{s}{s-1} - s\int_1^\infty \frac{\{u\}}{u^{s+1}}\,du \quad (\text{para } \sigma > 0).$$

Além disso, precisaremos da versão finita

$$\begin{aligned}\sum_{n\leq x}\frac{1}{n^s} &= \frac{\lfloor x\rfloor}{x^s} + s\int_1^x \frac{\lfloor u\rfloor}{u^{s+1}}\,du \\ &= \frac{\lfloor x\rfloor}{x^s} + s\int_1^x \frac{du}{u^s} - s\int_1^x \frac{\{u\}}{u^{s+1}}\,du \\ &= \frac{\lfloor x\rfloor}{x^s} + \frac{s}{1-s}\left(x^{1-s}-1\right) - s\int_1^x \frac{\{u\}}{u^{s+1}}\,du.\end{aligned}$$

Tomando $x = N \in \mathbb{Z}$ e subtraindo as expressões acima,

$$\zeta(s) = \sum_{n=1}^{N}\frac{1}{n^s} + \frac{N^{1-s}}{s-1} - s\int_N^\infty \frac{\{u\}}{u^{s+1}}\,du.$$

Pela desigualdade triangular, para $\sigma > 0$ e $t \neq 0$,

$$|\zeta(s)| \leq \sum_{n=1}^{N}\frac{1}{n^\sigma} + \frac{N^{1-\sigma}}{|t|} + |s|\int_N^\infty \frac{1}{u^{\sigma+1}}\,du.$$

(Aqui usamos que $|s-1| = |(\sigma-1)+it| \geq |t|$.)

Veja agora que

$$\int_N^\infty \frac{1}{u^{\sigma+1}}\,du = \left(\frac{-u^{-\sigma}}{\sigma}\right)\Bigg|_N^{+\infty} = \frac{N^{-\sigma}}{\sigma}.$$

Agora, se $\frac{1}{2} \le \sigma_0 < 1$ e $\sigma \ge \sigma_0$, como a soma é função decrescente de σ,

$$\sum_{n=1}^{N} \frac{1}{n^\sigma} \le \sum_{n=1}^{N} \frac{1}{n^{\sigma_0}} \le \int_0^N \frac{du}{u^{\sigma_0}} = \frac{N^{1-\sigma_0}}{1-\sigma_0}.$$

Logo, se $|t| \ge 2$, $\frac{1}{2} \le \sigma_0 < 1$ e $\sigma \ge \sigma_0$,

$$|\zeta(s)| \le \frac{N^{1-\sigma_0}}{1-\sigma_0} + \frac{N^{1-\sigma_0}}{|t|} + N^{-\sigma_0} + \frac{|t|}{\sigma_0}N^{-\sigma_0}.$$

Aqui usamos que $|s| \le |\sigma| + |t|$. Tomando $N = \lfloor |t| \rfloor \le |t|$,

$$\begin{aligned}|\zeta(s)| &\le \frac{|t|^{1-\sigma_0}}{1-\sigma_0}\left(1 + \frac{1-\sigma_0}{|t|} + \frac{1-\sigma_0}{\lfloor |t| \rfloor} + \frac{(1-\sigma_0)|t|}{\sigma_0 \lfloor |t| \rfloor}\right)\\ &\le \left(1 + \frac{1}{4} + \frac{1}{4} + \frac{3}{2}\right)\frac{|t|^{1-\sigma_0}}{1-\sigma_0} = 3\,\frac{|t|^{1-\sigma_0}}{1-\sigma_0} \le 4\,\frac{|t|^{1-\sigma_0}}{1-\sigma_0},\end{aligned}$$

pois $\frac{1-\sigma_0}{\sigma_0} \le 1$ e $\frac{|t|}{\lfloor |t| \rfloor} < \frac{3}{2}$ para $|t| \ge 2$. Isso conclui a demonstração do item (i).

Demonstração do item (ii)**.** Veja que

$$|t| \ge 2 \Rightarrow 4\log|t| \ge \log 16 > 2.$$

Isto quer dizer que

$$1 - \frac{1}{4\log|t|} \ge \frac{1}{2}.$$

Pelo item (i), definindo $\sigma_0 = 1 - \frac{1}{4\log|t|}$, temos

$$|\zeta(s)| \le \frac{4 \cdot |t|^{\frac{1}{4\log|t|}}}{\frac{1}{4\log|t|}} = 16e^{1/4}\log|t|.$$

Demonstração do item (iii)**.** Se $\sigma \ge 2$, tem-se

$$-\zeta'(s) = \sum_{n\ge 1} \frac{\log n}{n^s},$$

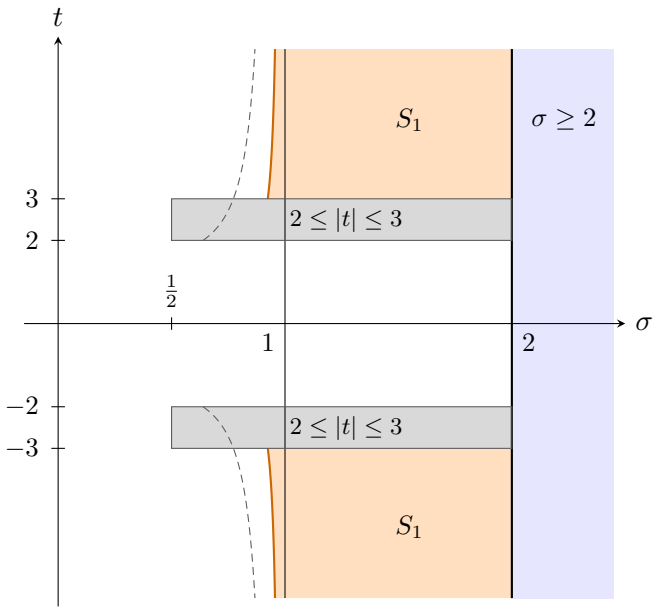
e disto deduzimos que

$$|\zeta'(s)| \le \sum_{n\ge 1} \frac{\log n}{n^2} = O(1).$$

Suponha agora que s esteja na região compacta

$$\{s \in \mathbb{C} :\ 2 \le |t| \le 3;\ \tfrac{1}{2} \le \sigma \le 2\}.$$

Como, nesta região, ζ' é contínua, ela é, portanto, limitada.



As três regiões da demonstração, com a fronteira de S_1, $\sigma = 1 - \frac{1}{12\log|t|}$, em laranja. A curva tracejada é a fronteira da região do item (ii), $\sigma = 1 - \frac{1}{4\log|t|}$.

Por fim, seja agora $s \in S_1 = \left\{s \in \mathbb{C} : |t| \geq 3,\ 1 - \frac{1}{12\log|t|} \leq \sigma \leq 2\right\}$. Considere o disco fechado de centro s e raio $\delta = \frac{1}{12\log|t|}$.

Para todo $s \in S_1$, este disco está contido na região do item (ii), ou seja, na região

$$\left\{s' \in \mathbb{C} : |\mathrm{Im}(s')| \geq 2,\ \mathrm{Re}(s') \geq 1 - \frac{1}{4\log|\mathrm{Im}(s')|}\right\}.$$

De fato, seja s' tal que $|s' - s| \leq \delta$. Como $\delta < 1$, temos $|\mathrm{Im}(s')| \leq |t| + 1 \leq |t|^{3/2}$, pois $|t| \geq 3$; logo $\log|\mathrm{Im}(s')| \leq \frac{3}{2}\log|t|$. Assim,

$$\begin{aligned}\mathrm{Re}(s') \geq \mathrm{Re}(s) - \delta &\geq \left(1 - \frac{1}{12\log|t|}\right) - \frac{1}{12\log|t|} \\ &= 1 - \frac{1}{6\log|t|} \geq 1 - \frac{1}{4\log|\mathrm{Im}(s')|}.\end{aligned}$$

Da mesma maneira, para a parte imaginária:

$$|\mathrm{Im}(s')| \geq |\mathrm{Im}(s)| - \delta \geq 3 - \frac{1}{12\log 3} \geq 2.$$

Finalmente, pela fórmula integral de Cauchy, para todo $s \in S_1$,

$$\zeta'(s) = \frac{1}{2\pi i}\int_{|z-s|=\delta} \frac{\zeta(z)}{(z-s)^2}\,dz.$$

Isto implica que

$$|\zeta'(s)| \leq \frac{1}{\delta} \cdot \sup_{|z-s|=\delta} |\zeta(z)|.$$

Pelo item (ii), sabemos que $|\zeta(z)| \leq A_1 \log|\mathrm{Im}(z)|$. Agora, para todo z no disco, $|\mathrm{Im}(z)| \leq |t| + \delta \leq |t| + 1$. Concluímos que

$$|\zeta'(s)| \leq 12\log|t| \cdot A_1 \log(|t|+1) \leq A_2(\log|t|)^2. \qquad \square$$

Observação 5.3. Para a k-ésima derivada, o mesmo argumento diz que existem $A_k, B_k > 0$ tais que

$$|\zeta^{(k)}(s)| \leq A_k(\log|t|)^{k+1}$$

na região

$$\left\{s \in \mathbb{C} : \ |t| \geq 2, \ \sigma \geq 1 - \frac{1}{B_k \log|t|}\right\}.$$

5.4 Cotas inferiores e região livre de zeros

Para o Teorema dos Números Primos, precisamos de estimativas para a derivada logarítmica $-\frac{\zeta'(s)}{\zeta(s)}$. Portanto, ainda precisamos de uma cota para $\frac{1}{\zeta(s)}$.

Teorema 5.4 (Cotas inferiores e região livre de zeros para ζ)**.** *As seguintes afirmações sobre a função ζ são válidas:*

(i) $\zeta(s) \neq 0$ *em* $\mathrm{Re}(s) \geq 1$.

(ii) Existem constantes $c, A > 0$ tais que $\zeta(s)$ não possui zeros em

$$\{\sigma > 1 - c, \ |t| \leq 2\}$$

e, nesta região, tem-se

$$\left|\frac{1}{\zeta(s)}\right| \leq A.$$

(iii) Existem constantes $c', A' > 0$ tais que $\zeta(s)$ não possui zeros em

$$\left\{\sigma > 1 - \frac{c'}{(\log|t|)^9}, \ |t| \geq 2\right\}$$

e, nesta região, tem-se

$$\left|\frac{1}{\zeta(s)}\right| \leq A'(\log|t|)^7.$$

Demonstração. **Demonstração do item** (i)**.** Quando $\mathrm{Re}(s) > 1$, temos o produto de Euler:

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

O produto converge absolutamente, logo $\zeta(s) \neq 0$. Quando $\mathrm{Re}(s) = 1$, $s \neq 1$, o Teorema 5.1 diz que $\zeta(s) \neq 0$.

Demonstração do item (ii)**.** Para $|t| \leq 2$, a ideia é usar a continuidade. O polo em $s = 1$ atrapalha um pouco. Para resolver o problema, considere $F(s) = (s-1)\zeta(s)$. Esta função possui extensão holomorfa a $\mathrm{Re}(s) > 0$ e não se anula em $\{\mathrm{Re}(s) = 1\}$ (veja que $F(1) = 1$).

Por continuidade e compacidade, existe uma região, como no enunciado, onde F não se anula. Consequentemente, ζ não se anula.

Dividindo a região $\{\sigma \geq 1 - c, \ |t| \leq 2\}$ em

$$\{1 - c \leq \sigma \leq 2, \ |t| \leq 2\} \cup \{\sigma \geq 2, \ |t| \leq 2\},$$

observamos que, na primeira região, por ser compacta, como $\frac{1}{\zeta}$ define uma função contínua em seu fecho, esta deve ser limitada aí. Por fim, na segunda região,

$$\frac{1}{\zeta(s)} = \sum_{n\geq 1} \frac{\mu(n)}{n^s} \Longrightarrow \left|\frac{1}{\zeta(s)}\right| \leq \sum_{n\geq 1} \frac{|\mu(n)|}{n^\sigma} \leq \sum_{n\geq 1} \frac{1}{n^2} = O(1).$$

Demonstração do item (iii)**.** Para $\sigma > 1$, temos

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

Assim, se $\sigma \geq 2$, então

$$\left|\frac{1}{\zeta(s)}\right| \leq \sum_{n\geq 1} \frac{1}{n^2} = O(1),$$

que é uma estimativa melhor que $O((\log|t|)^7)$.

Suponha agora que

$$1 - \frac{c'}{(\log|t|)^9} < \sigma < 2.$$

Para $1 < \sigma < 2$, usamos mais uma vez a desigualdade (5.2):

$$|\zeta^3(\sigma)\zeta^4(\sigma+it)\zeta(\sigma+2it)| \geq 1 \quad (\sigma > 1,\ t \in \mathbb{R}).$$

Obtemos assim a estimativa:

$$\left|\frac{1}{\zeta(\sigma+it)}\right| \leq |\zeta(\sigma)|^{\frac{3}{4}} |\zeta(\sigma+2it)|^{\frac{1}{4}}.$$

Usando que $\zeta(s)$ tem um polo simples em $s = 1$, vemos que

$$|\zeta(\sigma)| \leq C_0 \cdot \frac{1}{|\sigma - 1|} \quad (1 < \sigma < 2).$$

Pelo Teorema 5.2, temos

$$|\zeta(\sigma+2it)| \leq A_1 \log|2t| \leq 2A_1 \log|t|.$$

Juntando tudo, obtemos

$$\left|\frac{1}{\zeta(\sigma+it)}\right| \leq \left(\frac{C_0}{\sigma-1}\right)^{\frac{3}{4}} (2A_1 \log|t|)^{\frac{1}{4}} \quad (1 < \sigma < 2).$$

Suponha agora que s está na região

$$\left\{1 + \frac{c'}{(\log|t|)^9} \leq \sigma < 2,\ |t| \geq 2\right\}, \tag{5.3}$$

onde tomamos $c' < (\log 2)^9$, de modo que $1 + \frac{c'}{(\log|t|)^9} < 2$ para todo $|t| \geq 2$. Segue da desigualdade acima que

$$\left|\frac{1}{\zeta(s)}\right| \leq \left(\frac{C_0(\log|t|)^9}{c'}\right)^{\frac{3}{4}} (2A_1 \log|t|)^{\frac{1}{4}} = \frac{C_0^{3/4}(2A_1)^{1/4}}{(c')^{3/4}} (\log|t|)^7.$$

Ou seja,

$$|\zeta(s)| \geq \frac{(c')^{3/4}}{C_0^{3/4}(2A_1)^{1/4}} (\log|t|)^{-7}.$$

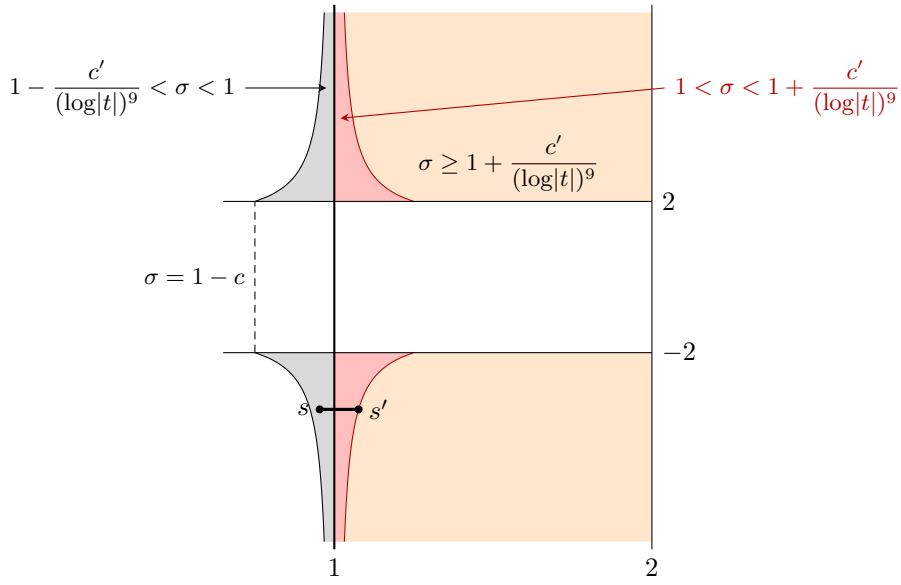
Considere agora s na região

$$\left\{1 - \frac{c'}{(\log|t|)^9} < \sigma < 1 + \frac{c'}{(\log|t|)^9},\ |t| \geq 2\right\}.$$

Considere o segmento horizontal ligando s ao ponto

$$s' = 1 + \frac{c'}{(\log|t|)^9} + it.$$

A figura a seguir mostra as regiões da demonstração, com o segmento ligando s a s'.



A largura das faixas está exagerada para facilitar a leitura.

Vamos mostrar uma cota inferior para $\zeta(s)$ utilizando uma cota inferior para $\zeta(s')$, uma cota superior para $\zeta'(z)$, $z \in [s, s']$, e a desigualdade do valor médio.

Primeiramente, como s' está na região (5.3), segue que

$$|\zeta(s')| \geq \frac{(c')^{3/4}}{C_0^{3/4}(2A_1)^{1/4}}(\log|t|)^{-7}.$$

Por outro lado,

$$\zeta(s') - \zeta(s) = \int_{\sigma}^{1+\frac{c'}{(\log|t|)^9}} \zeta'(x+it)\,dx.$$

Donde,

$$|\zeta(s') - \zeta(s)| \leq \frac{2c'}{(\log|t|)^9} \sup_{x \in [\sigma, 1+\frac{c'}{(\log|t|)^9}]} |\zeta'(x+it)|.$$

Pelo Teorema 5.2, para c' suficientemente pequeno,

$$|\zeta'(x+it)| \leq A_2(\log|t|)^2.$$

Com isso,

$$|\zeta(s') - \zeta(s)| \leq \frac{2c'A_2}{(\log|t|)^7}.$$

Portanto, pela desigualdade triangular,

$$|\zeta(s)| \geq |\zeta(s')| - |\zeta(s') - \zeta(s)| \geq \left(\frac{(c')^{3/4}}{C_0^{3/4}(2A_1)^{1/4}} - 2c'A_2\right)(\log|t|)^{-7}.$$

Como $c' \to 0$ mais rapidamente que $(c')^{3/4}$, tomando c' suficientemente pequeno, a constante fica positiva. Isso conclui a demonstração do item (iii). □

O teorema a seguir combina os Teoremas 5.2 e 5.4 em uma cota para $\frac{\zeta'(s)}{\zeta(s)}$.

Teorema 5.5 (Cotas para $-\frac{\zeta'(s)}{\zeta(s)}$)**.**

(i) Existem constantes $c, A > 0$ tais que $\zeta(s)$ não possui zeros em

$$\{\sigma > 1 - c,\ |t| \le 2\}$$

e, nesta região, tem-se

$$\left|-\frac{\zeta'(s)}{\zeta(s)}\right| \le A \max\left(1, \frac{1}{|\sigma - 1|}\right) \quad (\sigma \neq 1).$$

(ii) Existem constantes $c', A' > 0$ tais que $\zeta(s)$ não possui zeros em

$$\left\{\sigma > 1 - \frac{c'}{(\log|t|)^9},\ |t| \ge 2\right\}$$

e, nesta região, tem-se

$$\left|-\frac{\zeta'(s)}{\zeta(s)}\right| \le A'(\log|t|)^9.$$

Demonstração. O item (ii) segue diretamente dos Teoremas 5.2 e 5.4. Eles garantem que existe c' tal que, se $\sigma \ge 1 - \frac{c'}{(\log|t|)^9}$, $|t| \ge 2$, temos

$$\zeta'(s) \ll (\log|t|)^2, \qquad \frac{1}{\zeta(s)} \ll (\log|t|)^7,$$

e $2 + 7 = 9$.

Para o item (i), divide-se em 2 regiões: $\{1 - c \le \sigma \le 2\}$ e $\{\sigma > 2\}$. Observe que, por um lado, $-\frac{\zeta'(s)}{\zeta(s)}$ possui um polo de ordem 1 em $s = 1$ e, por outro lado, se $\mathrm{Re}(s) \ge 2$,

$$\left|-\frac{\zeta'(s)}{\zeta(s)}\right| \le \sum_{n\ge1} \frac{\Lambda(n)}{n^\sigma} \le \sum_{n\ge1} \frac{\Lambda(n)}{n^2} = O(1).$$

Resta a região compacta $\{1 - c \le \sigma \le 2,\ |t| \le 2\}$. Diminuindo c, se necessário, ela não contém zeros de ζ; logo $(s-1)\zeta'(s)/\zeta(s)$ é holomorfa numa vizinhança dela e, portanto, limitada nela. Isso dá $|\zeta'(s)/\zeta(s)| \ll \frac{1}{|s-1|} \le \frac{1}{|\sigma-1|}$ nessa região. □

5.5 Prova do Teorema dos Números Primos

Nesta seção e na seguinte, $R(x)$ denota uma função da forma

$$R(x) = \exp(-c(\log x)^{1/10}),$$

onde a constante $c > 0$ pode mudar de linha para linha. Nosso objetivo é demonstrar o seguinte teorema, que contém a fórmula assintótica $\psi(x) \sim x$.

Teorema 5.6 (Teorema dos Números Primos)**.** *Existe $c > 0$ tal que*

$$\psi(x) = x + O\big(x \exp(-c(\log x)^{1/10})\big).$$

5.5.1 A soma suavizada

Como a fórmula de Perron é melhor de ser aplicada quando temos uma soma mais suave (Seção 4.6), considere

$$\Psi(x) = \sum_{n \le x} \Lambda(n)(x-n).$$

Proposição 5.7. *Existe $c > 0$ tal que*

$$\Psi(x) = \frac{x^2}{2} + O\big(x^2 \exp(-c(\log x)^{1/10})\big). \tag{5.4}$$

Demonstração. Pela fórmula de Perron, temos

$$\Psi(x) = \frac{1}{2\pi i} \int_{(a)} -\frac{\zeta'(s)}{\zeta(s)} \frac{x^{s+1}}{s(s+1)}\, ds \quad (a > 1).$$

Seja $x \ge 3$ e tome T um parâmetro a ser escolhido mais tarde tal que $3 \le T \le x$. A escolha será $T = \exp\left((\log x)^{1/10}\right)$.

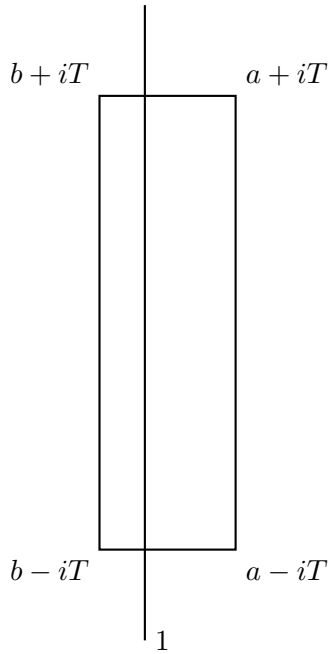
Tome $a = 1 + \frac{1}{\log x}$. Escreva

$$\int_{(a)} (\dots) = \int_{a-i\infty}^{a-iT} (\dots) + \int_{a-iT}^{a+iT} (\dots) + \int_{a+iT}^{a+i\infty} (\dots)$$

e chame as integrais do lado direito de I_-, I_0 e I_+, respectivamente. Para calcular I_0, vamos usar o teorema dos resíduos no retângulo de vértices $a - iT$, $a + iT$, $b + iT$, $b - iT$, onde

$$b = 1 - \frac{c_0}{(\log T)^9}$$

para uma constante $c_0 > 0$ suficientemente pequena em relação às constantes do Teorema 5.5, de modo que o retângulo fique na região livre de zeros.



Pelo teorema dos resíduos, temos a soma das integrais no contorno:

$$\int_{a-iT}^{a+iT}(\dots)+\int_{a+iT}^{b+iT}(\dots)+\int_{b+iT}^{b-iT}(\dots)+\int_{b-iT}^{a-iT}(\dots)$$
$$=2\pi i\cdot\operatorname{Res}_{s=1}\left(\left(-\frac{\zeta'(s)}{\zeta(s)}\right)\frac{x^{s+1}}{s(s+1)}\right)=2\pi i\cdot\frac{x^2}{2}.$$

A primeira integral do lado esquerdo é I_0. Chamando as outras de I_1, I_2 e I_3, respectivamente, obtemos a seguinte fórmula para $\Psi(x)$:

$$\Psi(x)=\frac{x^2}{2}+\frac{1}{2\pi i}\left(I_+ + I_- - I_1 - I_2 - I_3\right).$$

Vamos estimar as integrais I_+ e I_-. Para $\sigma>1$,

$$-\frac{\zeta'(s)}{\zeta(s)}=\sum_{n\geq 1}\frac{\Lambda(n)}{n^s}\quad\Rightarrow\quad\left|-\frac{\zeta'(s)}{\zeta(s)}\right|\leq\sum_{n\geq 1}\frac{\Lambda(n)}{n^\sigma}=-\frac{\zeta'(\sigma)}{\zeta(\sigma)}.$$

Como $-\frac{\zeta'}{\zeta}$ possui um polo de ordem 1 em $\sigma=1$, temos que

$$\left|-\frac{\zeta'(\sigma)}{\zeta(\sigma)}\right|\ll\frac{1}{\sigma-1}.$$

Logo, para $\operatorname{Re}(s)=a$, vale a estimativa

$$\left|-\frac{\zeta'(s)}{\zeta(s)}\right|\ll\frac{1}{a-1}=\log x.$$

Aplicando esta estimativa para $s=a+it$, $t>T$, obtemos:

$$I_+\ll\log x\int_T^\infty\frac{x^{a+1}}{|a+it||a+1+it|}\,dt$$
$$\ll x^2\log x\int_T^\infty\frac{dt}{t^2}=\frac{x^2\log x}{T}.$$

De maneira análoga, $I_-\ll\frac{x^2\log x}{T}$.

Para as integrais horizontais (por exemplo I_1), temos:

$$I_1=\int_b^a\frac{\zeta'(y+iT)}{\zeta(y+iT)}\frac{x^{y+1+iT}}{(y+iT)(y+1+iT)}\,dy.$$

O que leva à estimativa:

$$I_1\ll\frac{x^2}{T^2}(a-b)\sup_{y\in[b,a]}\left|-\frac{\zeta'(y+iT)}{\zeta(y+iT)}\right|.$$

Por um lado, pelo Teorema 5.5, para todo y no domínio de integração,

$$\left|-\frac{\zeta'(y+iT)}{\zeta(y+iT)}\right|\ll(\log T)^9.$$

Por outro lado,

$$a-b=\frac{1}{\log x}+\frac{c_0}{(\log T)^9}.$$

Portanto,

$$\begin{aligned} I_1 &\ll \frac{x^2}{T^2}(\log T)^9\left(\frac{1}{\log x}+\frac{1}{(\log T)^9}\right) \\ &= \frac{x^2}{T^2}\frac{(\log T)^9}{\log x}+\frac{x^2}{T^2}. \end{aligned}$$

Analogamente, provamos a mesma estimativa para I_3.

Finalmente, para a integral I_2, sabemos pelo Teorema 5.5 que para $s=b+it$ vale:

$$-\frac{\zeta'(s)}{\zeta(s)} \ll \begin{cases} 1+\frac{1}{|b-1|}, & |t|\le 2, \\ (\log|t|)^9, & |t|\ge 2. \end{cases}$$

Como $\frac{1}{1-b} \asymp (\log T)^9$, obtém-se na região considerada a cota uniforme

$$-\frac{\zeta'(s)}{\zeta(s)} \ll (\log T)^9.$$

Assim,

$$I_2 \ll (\log T)^9 x^{b+1}\int_{-T}^{T}\frac{dt}{|b+it||b+1+it|}.$$

Como $|b+it||b+1+it| \ge |b+it|^2$ e $b>1/2$ tomando c_0 suficientemente pequeno, a integral acima é

$$\ll \int_{-T}^{T}\frac{dt}{1+t^2} = O(1).$$

De modo que

$$I_2 \ll \frac{x^2(\log T)^9}{\exp\left(c_0\frac{\log x}{(\log T)^9}\right)}.$$

Juntando todas as estimativas acima, temos que

$$\Psi(x) = \frac{x^2}{2}+O\left(x^2\left(\frac{\log x}{T}+\frac{(\log T)^9}{T^2\log x}+\frac{1}{T^2}+\frac{(\log T)^9}{\exp\left(c_0\frac{\log x}{(\log T)^9}\right)}\right)\right). \tag{5.5}$$

Escolhendo $T=\exp\left((\log x)^{1/10}\right)$, que satisfaz $3\le T\le x$ para x grande, concluímos (5.4). □

Observação 5.8. Para mostrar que o termo de erro acima é admissível, precisamos utilizar as estimativas

$$(\log x)^A \ll \exp(c(\log x)^{1/10}) \ll x^{\varepsilon} \quad (A,c,\varepsilon>0).$$

Por exemplo, para $T=\exp(c(\log x)^{1/10})$,

$$\begin{aligned} \frac{\log x}{T} &= \log x\cdot\exp(-c(\log x)^{1/10}) \\ &\ll \exp(\tfrac{c}{2}(\log x)^{1/10}-c(\log x)^{1/10}) \\ &= \exp(-\tfrac{c}{2}(\log x)^{1/10}), \end{aligned}$$

$$\begin{aligned} \frac{(\log T)^9}{T^2\log x} &= c^9(\log x)^{-1/10}\cdot\exp(-2c(\log x)^{1/10}) \\ &\ll \exp(c(\log x)^{1/10}-2c(\log x)^{1/10}). \end{aligned}$$

Os outros casos são semelhantes.

Observação 5.9. O motivo para a escolha de T é algo não tão trivial e requer um pouco de prática. Primeiramente, por conta da observação acima, podemos perceber que as potências de $\log x$ e $\log T$ são pouco relevantes para a ordem de grandeza total. Sendo assim, o termo de erro é moralmente da ordem de

$$x^2\left(\frac{1}{T}+\frac{1}{T^2}+\frac{1}{T^2}+\frac{1}{\exp\left(c_0\frac{\log x}{(\log T)^9}\right)}\right)\ll x^2\left(\frac{1}{T}+\frac{1}{\exp\left(c_0\frac{\log x}{(\log T)^9}\right)}\right).$$

Neste ponto, observa-se que o primeiro termo é uma função decrescente de T e o segundo, uma função crescente. Segue imediatamente que pelo menos um dos termos é maior que o valor obtido quando as funções se igualam (por quê?). Forçando os dois termos a serem iguais, obtemos a cota acima, que por isso é ótima para esta expressão. Um mero cálculo mostra que

$$T=\exp\left(c_0\frac{\log x}{(\log T)^9}\right)\iff T=\exp\left((c_0\log x)^{1/10}\right).$$

Como o resultado final não diz nada de concreto sobre a constante aparecendo dentro da exponencial, a escolha $T=\exp\left((\log x)^{1/10}\right)$ é suficiente.

Observação 5.10. Se no lugar da estimativa $-\frac{\zeta'}{\zeta}\ll(\log T)^9$ e da região do Teorema 5.5, utilizarmos a estimativa $-\frac{\zeta'}{\zeta}\ll\log T$ e a região livre de zeros $\sigma>1-c/\log|t|$, devidas a de la Vallée Poussin, o que permite tomar $b=1-c_0/\log T$, o mesmo argumento acima nos daria a estimativa mais forte $O(x^2\exp(-c(\log x)^{1/2}))$.

5.5.2 De Ψ para ψ

Demonstração do Teorema 5.6. Pela Proposição 5.7, vale $\Psi(x)=\frac{x^2}{2}+O(x^2R(x))$. Observe que

$$\Psi(x)=\sum_{n\le x}\Lambda(n)(x-n)=\sum_{n\le x}\Lambda(n)\int_n^x dt.$$

Ou seja, $\Psi(x)=\int_2^x\psi(t)\,dt$. Donde,

$$\Psi(x)-\Psi((1-\delta)x)=\int_{(1-\delta)x}^{x}\psi(t)\,dt.$$

Como a função ψ é crescente, segue que

$$\delta x\,\psi(x)\ge\int_{(1-\delta)x}^{x}\psi(t)\,dt=\Psi(x)-\Psi((1-\delta)x).$$

Utilizando duas vezes (5.4), vemos que

$$\delta x\,\psi(x)\ge\frac{x^2}{2}-\frac{(1-\delta)^2x^2}{2}+O(x^2R(x))+O(x^2R((1-\delta)x)).$$

Supondo que $\delta\le\frac{1}{2}$,

$$\begin{aligned}R((1-\delta)x)&\le R\left(\frac{x}{2}\right)\\&=\exp\left(-c\left(\log\frac{x}{2}\right)^{1/10}\right)\\&\le\exp\left(-c'(\log x)^{1/10}\right),\end{aligned}$$

onde c' é uma constante independente de δ. Ou seja, usando nossa convenção (ou abusando da notação, como preferir),

$$R((1-\delta)x) = R(x).$$

Com isso, obtemos:

$$\delta x\psi(x) \geq (2\delta - \delta^2)\frac{x^2}{2} + O(x^2 R(x)).$$

Dividindo por δx:

$$\psi(x) \geq \left(1 - \frac{\delta}{2}\right) x + O\left(\frac{1}{\delta} x R(x)\right).$$

Mas isto implica que

$$\psi(x) \geq x + O\left(\delta x + \frac{xR(x)}{\delta}\right).$$

Tomando $\delta = \sqrt{R(x)}$, que satisfaz $\delta \leq \frac{1}{2}$ para x grande,

$$\psi(x) \geq x + O(x\sqrt{R(x)}), \quad \text{isto é,} \quad \psi(x) \geq x + O(xR(x)),$$

pois, mais uma vez por conta de nossa convenção,

$$\sqrt{R(x)} = R(x).$$

Analogamente, comparando $\delta x\, \psi(x)$ e $\int_x^{(1+\delta)x} \psi(t)\, dt = \Psi((1+\delta)x) - \Psi(x)$, segue que

$$\psi(x) \leq x + O(xR(x)). \qquad \square$$

5.6 Formas alternativas e o TNP para a função de Möbius

Corolário 5.11 (Formas alternativas do TNP)**.** *As seguintes fórmulas assintóticas sobre os números primos são verdadeiras:*

i) $\theta(x) = x + O(xR(x))$,

ii) $\pi(x) = \mathrm{Li}(x) + O(xR(x))$,

iii) $\sum_{p\leq x} \frac{\log p}{p} = \log x + c_1 + O(R(x))$,

iv) $\sum_{p\leq x} \frac{1}{p} = \log\log x + A + O(R(x))$,

v) $\prod_{p\leq x} \left(1 - \frac{1}{p}\right) = \frac{e^{-\gamma}}{\log x}(1 + O(R(x)))$,

onde $\mathrm{Li}(x) = \int_2^x \frac{dt}{\log t}$, c_1 *é uma constante e* A *é a constante que aparece na fórmula para* $\sum_{p\leq x} \frac{1}{p}$ *no Teorema de Mertens (Seção 3.3).*

Demonstração. Segue do Teorema 5.6 por soma por partes (Seção 2.3); no item (v), usa-se também o Teorema 4.9, que identifica a constante como γ. $\square$

Observação 5.12. A segunda fórmula nos diz em particular que é falsa a fórmula

$$\pi(x) = \frac{x}{\log x} + O(xR(x)).$$

Na verdade, tem-se a seguinte expansão assintótica para $\pi(x)$: para todo $k \in \mathbb{Z}_{\geq 1}$,

$$\pi(x) = \sum_{j=1}^{k} \frac{(j-1)!\, x}{(\log x)^j} + O\left(\frac{x}{(\log x)^{k+1}}\right).$$

Também temos um TNP para a função de Möbius:

Teorema 5.13.

$$\sum_{n\le x}\mu(n) = O(xR(x)).$$

Ideia da demonstração. A ideia é fazer o mesmo que fizemos antes, trocando Λ (von Mangoldt) por μ e $-\frac{\zeta'}{\zeta}$ por $\frac{1}{\zeta}$.

- Não-anulação de ζ e estimativas para $\frac{1}{\zeta(s)}$: já as temos (Teorema 5.4).
- Localização dos polos de $\frac{1}{\zeta(s)}$: não há polo em $s=1$ nem na região sem zeros. Por isso não há termo principal na fórmula.

A única parte do argumento que não funciona aqui é a parte final, em que usamos que a função $\psi(x)$ é crescente (a soma da função de Möbius oscila).

Essa parte final pode ser substituída pelo seguinte argumento. Sejam $M(x)=\sum_{n\le x}\mu(n)$ e $M_1(x)=\sum_{n\le x}\mu(n)(x-n)=\int_1^x M(t)\,dt$. O argumento acima dá $M_1(x)=O(x^2R(x))$. Para $0<\delta\le 1$ e $x\le t\le(1+\delta)x$, temos $|M(t)-M(x)|\le t-x+1\le \delta x+1$, pois $|\mu(n)|\le 1$. Logo

$$\Big|\delta x\,M(x)-\int_x^{(1+\delta)x} M(t)\,dt\Big|\le \delta x(\delta x+1),$$

e, como $\int_x^{(1+\delta)x} M(t)\,dt = M_1((1+\delta)x)-M_1(x)=O(x^2R(x))$,

$$|M(x)|\le \delta x+1+O\Big(\frac{xR(x)}{\delta}\Big).$$

Tomando $\delta=\sqrt{R(x)}$, obtemos $M(x)=O(x\sqrt{R(x)})$, e $\sqrt{R(x)}$ é novamente uma função da forma $R(x)$, com $c/2$ no lugar de c.

Outras maneiras possíveis de fazê-lo:

1. Usar a fórmula de Perron direta para $\sum_{n\le x}\mu(n)$ e não para $\sum_{n\le x}\mu(n)(x-n)$.
2. Tentar deduzir $\sum_{n\le x}\mu(n)=O(xR(x))$ partindo de $\sum_{n\le x}\Lambda(n)=x+O(xR(x))$, como fizemos na Seção 3.5 para mostrar que o TNP (sem termo de erro específico) implica

$$\sum_{n\le x}\mu(n)=o(x).$$

5.7 Fórmulas explícitas e os zeros de ζ

O TNP é o primeiro passo conectando as informações sobre $\zeta(s)$ na região $\{0<\mathrm{Re}(s)<1\}$ e $\pi(x)$ (ou $\psi(x)$). Quanto mais informações soubermos sobre ζ, mais informações teremos sobre os primos.

Os resultados desta seção são enunciados sem demonstração, assim como o termo de erro $O(x\exp(-c\sqrt{\log x}))$ da Seção 5.1. Lembre que $\Psi(x)$ é a função de (5.4).

Teorema 5.14 (Fórmula explícita para $\Psi(x)$)**.**

$$\Psi(x)=\frac{x^2}{2}-\sum_\rho\frac{x^{\rho+1}}{\rho(\rho+1)}-\frac{\zeta'(0)}{\zeta(0)}x+\frac{\zeta'(-1)}{\zeta(-1)}-\sum_{n\ge1}\frac{x^{1-2n}}{2n(2n-1)},$$

onde a soma sobre ρ percorre todos os zeros de ζ na região $\{0<\mathrm{Re}(s)<1\}$, cada um contado com a sua multiplicidade (o resíduo de $-\zeta'/\zeta$ em ρ é a multiplicidade de ρ). Os pontos $s=-2n$, que dão a última soma, são os zeros triviais de ζ.

Para tirar proveito da fórmula explícita, precisamos de algumas informações sobre o conjunto $\mathcal{Z} = \{\rho \in \mathbb{C};\ 0 < \mathrm{Re}(\rho) < 1 \text{ e } \zeta(\rho) = 0\}$ (ver também a Seção 4.3). Por exemplo:

Teorema 5.15. *Seja $T \geq 10$ e seja $N(T)$ o número de zeros $\rho \in \mathcal{Z}$ com $0 < \mathrm{Im}(\rho) \leq T$,* contados com multiplicidade*:*

$$N(T) = \sum_{\substack{\rho \in \mathcal{Z} \\ 0 < \mathrm{Im}(\rho) \leq T}} m(\rho),$$

onde $m(\rho)$ é a multiplicidade do zero ρ. Então vale a fórmula:

$$N(T) = \frac{T}{2\pi} \log\left(\frac{T}{2\pi}\right) - \frac{T}{2\pi} + O(\log T).$$

Como os zeros são simétricos em relação ao eixo real, a condição $|\mathrm{Im}(\rho)| \leq T$ dá o dobro dessa quantidade.

Como na Seção 4.3, a contagem com multiplicidade é a natural, pois é ela que o princípio do argumento fornece; não se sabe se todos os zeros de ζ são simples. Daqui em diante, somas sobre ρ contam cada zero com a sua multiplicidade.

Em particular, temos as seguintes estimativas:

$$\sum_{\rho} \frac{1}{|\rho|^2} < +\infty, \tag{A}$$

$$\sum_{|\mathrm{Im}(\rho)| \leq T} \frac{1}{|\rho|} \asymp (\log T)^2 \quad \text{para } T \text{ suficientemente grande.} \tag{B}$$

Temos também uma fórmula explícita para $\psi(x)$ (não suave):

$$\sum_{n \leq x}{}' \Lambda(n) = x - \sum_{\rho} \frac{x^\rho}{\rho} - \log(2\pi) - \frac{1}{2}\log(1 - x^{-2}).$$

Observação 5.16. Aqui $\sum'$ significa que, se x é inteiro, o termo $\Lambda(x)$ é contado com peso $1/2$. A fórmula vale para $x > 1$, e a soma sobre ρ é entendida como $\lim_{T \to +\infty} \sum_{|\mathrm{Im}(\rho)| \leq T} \frac{x^\rho}{\rho}$.

Estimando o rabo da série, podemos mostrar que, para $2 \leq T \leq x$,

$$\psi(x) = x - \sum_{|\mathrm{Im}(\rho)| \leq T} \frac{x^\rho}{\rho} + O\left(\frac{x(\log x)^2}{T}\right)$$

(ver [Dav00, Cap. 17]). Combinando esta fórmula, com $T = x$, com a estimativa (B) acima, segue que, se existe $\theta < 1$ tal que $\mathrm{Re}(\rho) \leq \theta$ para todo $\rho \in \mathcal{Z}$, então teríamos

$$\psi(x) = x + O\left(x^\theta (\log x)^2\right).$$

Em particular, a Hipótese de Riemann implicaria que

$$\psi(x) = x + O\left(x^{1/2} (\log x)^2\right).$$

Capítulo 6

Primos em progressões aritméticas: o Teorema de Dirichlet

6.1 Motivação

Consideremos os números primos até 100:

$$2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47,$$
$$53, 59, 61, 67, 71, 73, 79, 83, 89, 97.$$

Pergunta: Será que para todo $a \in \{1, 3, 7, 9\}$ existem infinitos primos cujo último dígito da expansão na base 10 é a?

Sabemos, com certeza, que os dígitos 0, 2, 4, 5, 6, 8 não podem aparecer infinitas vezes como dígito final de um número primo. De fato, se n termina em um dígito par, n deve ser par, e o único número primo par é 2. Se n termina em 5, n é múltiplo de 5, e o único primo múltiplo de 5 é o próprio 5.

O número 10 não tem nada de especial neste contexto. Poderíamos considerar a mesma pergunta para a representação em uma base q qualquer ou, o que é equivalente, olhar para classes de congruência módulo q.

Pergunta mais geral: Dados a, q, será que existem infinitos números primos na classe de congruência $\{n \equiv a \pmod{q}\}$?

Assim como no exemplo anterior, se houver $d > 1$ tal que $d \mid a$ e $d \mid q$, então d divide todos os elementos desta classe de congruência. Logo, esta classe não admitirá infinitos primos, pois um primo nela teria que ser múltiplo de d, isto é, igual a d.

Para todas as outras classes, temos:

Teorema 6.1 (Dirichlet)**.** *Sejam $a, q \in \mathbb{Z}$, com $q > 0$ e $(a, q) = 1$. Então existem infinitos primos $p \equiv a \pmod{q}$.*

6.1.1 Adaptações do argumento de Euclides

Como provar? Uma primeira ideia natural é adaptar a prova de Euclides de que existem infinitos primos, que, a partir de $p_1, \ldots, p_n$, considera $N = p_1 \cdots p_n + 1$. Isto significaria supor, por absurdo, que existam finitos primos de uma certa forma e, a partir deles, construir um novo número que não seja divisível por nenhum deles e mostrar que, por outro lado, este número *tem* que possuir um divisor primo desta forma.

Dados a, q como no enunciado, seja

$$\mathcal{P}(q, a) = \{p \text{ primo}; \ p \equiv a \pmod{q}\}.$$

- $\mathcal{P}(3,2)$ é infinito: Suponha que $\mathcal{P}(3,2)$ é finito. Como $2 \in \mathcal{P}(3,2)$, podemos escrever

  $$\mathcal{P}(3,2) = \{2, p_2, \ldots, p_m\}$$

  e considerar $N = 3p_2 \cdots p_m + 2$. O primo 2 fica fora do produto para que N não seja par. Como $N \equiv 2 \pmod 3$, algum fator primo p de N satisfaz $p \equiv 2 \pmod 3$; mas $p \neq 2$, pois N é ímpar, e $p \neq p_i$, pois nenhum p_i divide N. Isso contradiz a escolha da lista.

- $\mathcal{P}(3,1)$ é infinito: A ideia é usar resíduos quadráticos: por reciprocidade quadrática, para $p \neq 2, 3$,

  $$\left(\frac{-3}{p}\right) = 1 \iff p \equiv 1 \pmod 3.$$

  Suponha agora que $\mathcal{P}(3,1)$ é finito:

  $$\mathcal{P}(3,1) = \{p_1, \ldots, p_m\}$$

  e considere $N = (p_1 \cdots p_m)^2 + 3$. Como os p_i são ímpares, $N \equiv 4 \pmod 8$, de modo que N é par e $N/4$ é ímpar; além disso, $3 \nmid N$. Assim, olhamos para os divisores primos de $N/4$:

  $$p \mid \frac{N}{4} \implies \exists\, x \text{ tal que } p \mid x^2 + 3 \iff -3 \text{ é resíduo quadrático módulo } p.$$

  Como $N/4 > 1$, existe um tal primo p; ele é ímpar e diferente de 3, logo $p \equiv 1 \pmod 3$. Mas p não é nenhum dos p_i, pois caso contrário p dividiria 3. Isso contradiz a escolha da lista.

Outros casos semelhantes:

- $\mathcal{P}(6,1) = \mathcal{P}(3,1)$: nada a fazer.
- $\mathcal{P}(6,5) = \mathcal{P}(3,2) \setminus \{2\}$: nada a fazer.
- $\mathcal{P}(4,3)$, com $p_1 = 3$: $N = 4p_2 \cdots p_m + 3$.
- $\mathcal{P}(4,1)$: $N = (2p_1 \cdots p_m)^2 + 1$.
- $\mathcal{P}(5,1) \cup \mathcal{P}(5,4)$: $N = (p_1 \cdots p_m)^2 - 5$; aqui N é par e o que se usa é $N/4$, que é ímpar.
- $\mathcal{P}(5,2) \cup \mathcal{P}(5,3)$, com $p_1 = 2$: $N = 5p_2 \cdots p_m + 2$.

6.1.2 A ideia de Dirichlet

A ideia de Dirichlet está relacionada com outra prova da infinitude dos primos. Ela consiste em mostrar que a série

$$\sum_{p \equiv a\,(q)} \frac{1}{p}$$

diverge (aqui e adiante, (q) no índice abrevia "módulo q") ou, o que é equivalente, mostrar que

$$\lim_{\sigma \to 1^+} \sum_{p \equiv a\,(q)} \frac{1}{p^\sigma} = +\infty.$$

Denote por $P(\sigma, q, a)$ a série acima, e por $P(\sigma) = \sum_p 1/p^\sigma$ a série sobre todos os primos. Perceba que

$$P(\sigma) = \sum_{\substack{a \bmod q \\ (a,q)=1}} P(\sigma, q, a) + O(1),$$

onde o $O(1)$ vem dos primos que dividem q, que são em número finito.

Sejamos ousados: vamos tentar mostrar que, na soma acima, os termos $P(\sigma, q, a)$ divergem "igualmente". Isto é, queremos mostrar que, quando $\sigma \to 1^+$,

$$\begin{aligned} P(\sigma, q, a) &= \frac{1}{\phi(q)} P(\sigma) + O(1) \\ &= \frac{1}{\phi(q)} \log\left(\frac{1}{\sigma - 1}\right) + O(1). \end{aligned}$$

6.1.3 Analogia com séries de Fourier

Seja f uma função "legal" de período 2π. Então:

$$f(x) = \underbrace{a_0}_{\text{parte constante}} + \underbrace{\sum_{n\geq 1} b_n \cos(nx) + \sum_{n\geq 1} c_n \operatorname{sen}(nx)}_{\text{parte oscilante}}.$$

Na integração, as somas de cossenos e senos somem. Assim,

$$\frac{1}{2\pi}\int_0^{2\pi} f(x)\,dx = \frac{1}{2\pi}\int_0^{2\pi} (a_0 + \cdots)\,dx = a_0.$$

A ideia é fazer algum tipo de teoria de Fourier para a função característica dos números na classe de congruência $n \equiv a \pmod q$. Usaremos a notação

$$\mathbf{1}_X(x) = \begin{cases} 1 & \text{se } x \in X, \\ 0 & \text{caso contrário.} \end{cases}$$

Queremos escrever

$$\mathbf{1}_{n \equiv a\,(q)} = \text{parte constante} + \text{parte oscilante}.$$

Considerando a média, devemos ter parte constante $= \frac{1}{q}$. Assim, escrevendo P.O. para a parte oscilante,

$$\begin{aligned} P(\sigma, q, a) &= \sum_p \frac{1}{p^\sigma} \mathbf{1}_{p \equiv a\,(q)} \\ &\approx \sum_p \frac{1}{p^\sigma} \left(\frac{1}{q} + \text{P.O.}\right) \\ &\approx \frac{1}{q} P(\sigma) + E \end{aligned}$$

e, idealmente, o termo E é pequeno (por conta das oscilações). A passagem da segunda para a terceira linha é um salto de fé.

Temos um problema: apareceu $\frac{1}{q}$ e não $\frac{1}{\phi(q)}$. Isso acontece porque, como vimos, os números primos naturalmente pertencem às classes $\{n \equiv a \pmod q\}$ onde $(a, q) = 1$; então sua distribuição em *todas as classes* não é "bem comportada".

Na verdade, a decomposição $\mathbf{1}_{n \equiv a\,(q)} = \frac{1}{q} + \text{P.O.}$ é bem simples. Basta usar a fórmula

$$\frac{1}{q} \sum_{h \bmod q} e^{2\pi i h \frac{n-a}{q}} = \begin{cases} 1 & \text{se } n \equiv a \pmod q, \\ 0 & \text{caso contrário.} \end{cases}$$

Observação 6.2. Compare com

$$\frac{1}{2\pi}\int_0^{2\pi} e^{in\theta}\,d\theta = \begin{cases} 1 & \text{se } n = 0, \\ 0 & \text{se } n \in \mathbb{Z},\ n \neq 0. \end{cases}$$

O que Dirichlet trouxe de novidade foi um jeito de decompor a função $\mathbf{1}_{n\equiv a\,(q)}$ olhando só para os números com $(n, q) = 1$. Ou seja, precisamos estudar funções que são não nulas apenas quando $(n, q) = 1$.

6.2 Caracteres de Dirichlet

Definição 6.3. Seja $q \in \mathbb{Z}_{>0}$. Um *caractere de Dirichlet módulo q* é uma função aritmética χ com as seguintes propriedades:

i) χ é periódica módulo q:
$$\chi(n+q) = \chi(n) \quad \forall n \in \mathbb{Z}_{>0}.$$

ii) χ é completamente multiplicativa:
$$\chi(mn) = \chi(m)\chi(n) \quad \forall m, n \in \mathbb{Z}_{>0}.$$

iii) $\chi(n) \neq 0 \iff (n, q) = 1$.

Observação 6.4. Podemos estender naturalmente os caracteres de Dirichlet para os inteiros, usando a periodicidade. Por exemplo, para $q > 3$, $\chi(-3) := \chi(q-3)$. Daqui em diante, consideramos χ definida em $\mathbb{Z}$, com as propriedades acima valendo em $\mathbb{Z}$.

Exemplo 6.5 ($q = 1$)**.** $\chi(n) = 1$ para todo n é o único caractere módulo 1.

Exemplo 6.6 (q qualquer)**.** A função

$$\chi(n) = \begin{cases} 1 & \text{se } (n, q) = 1, \\ 0 & \text{caso contrário,} \end{cases}$$

é um caractere, chamado de *caractere principal* módulo q. Notação: $\chi_{0,q}$ (ou χ_0, se q estiver claro pelo contexto).

Exemplo 6.7 ($q = 2$)**.** $\chi(1) = 1$ pois χ é multiplicativa e $\chi(2) = 0$ pois $(2, 2) = 2$. Portanto, $\chi = \chi_0$ é o único caractere módulo 2.

Exemplo 6.8 ($q = 3$)**.** Como χ é periódico, os valores de $\chi(1), \chi(2)$ e $\chi(3)$ determinam χ. Temos

$$\chi(3) = 0 \text{ pois } (3, 3) = 3, \qquad \chi(1) = 1 \text{ pois } \chi \text{ é multiplicativo,}$$
$$\chi(2)\chi(2) = \chi(4) = \chi(1) = 1 \quad (\chi \text{ é completamente multiplicativo e periódico}).$$

Daí segue que

$$[\chi(2)]^2 = 1 \implies \chi(2) = \pm 1.$$

- Se $\chi(2) = 1$, então $\chi = \chi_0$.
- Se $\chi(2) = -1$, então $\chi = (\frac{\cdot}{3})$.

Exemplo 6.9 ($q = p$ primo ímpar)**.** $\chi(n) = \left(\frac{n}{p}\right)$ (símbolo de Legendre) é um caractere de Dirichlet.

Exemplo 6.10 (q ímpar)**.** $\chi(n) = \left[\frac{n}{q}\right]$ (símbolo de Jacobi) também é um exemplo.

Exemplo 6.11 ($q = 4$)**.** Mais uma vez, por periodicidade, os valores de $\chi(1), \chi(2), \chi(3)$ e $\chi(4)$ determinam χ. Além disso, é fácil ver que $\chi(1) = 1$, $\chi(2) = \chi(4) = 0$ e

$$\chi(3)^2 = \chi(9) = \chi(1) = 1 \implies \chi(3) = \pm 1.$$

Se $\chi(3) = 1$, então $\chi = \chi_0$. Se $\chi(3) = -1$, então $\chi(n) = 0$ para n par e, para n ímpar, escrevendo $n = p_1^{\alpha_1} \cdots p_k^{\alpha_k} q_1^{\beta_1} \cdots q_l^{\beta_l}$, com $p_i \equiv 1 \pmod 4$ e $q_j \equiv 3 \pmod 4$, temos

$$\begin{aligned}\chi(n) &= 1^{\alpha_1} \cdots 1^{\alpha_k} (-1)^{\beta_1} \cdots (-1)^{\beta_l} \\ &= \left(\frac{-1}{p_1}\right)^{\alpha_1} \cdots \left(\frac{-1}{p_k}\right)^{\alpha_k} \left(\frac{-1}{q_1}\right)^{\beta_1} \cdots \left(\frac{-1}{q_l}\right)^{\beta_l} \\ &= \left[\frac{-1}{n}\right],\end{aligned}$$

o símbolo de Jacobi "embaixo".

Exemplo 6.12 ($q = 5$)**.** Aqui $\chi(5) = 0$, e os valores de $\chi(1), \chi(2), \chi(3), \chi(4)$ determinam χ. Além disso, $\chi(1) = 1$ e $\chi(4)^2 = \chi(16) = \chi(1) = 1$, donde $\chi(4) = \pm 1$. Logo,

$$\chi(2)^2 = \chi(4) = \pm 1 \implies \chi(2)^4 = 1.$$

Temos também $\chi(2)^3 = \chi(8) = \chi(3)$. Resumindo, $\chi(2)^4 = 1$ e seu valor determina χ.

Exercício 6.13. Cada solução complexa de $\xi^4 = 1$ determina um caractere módulo 5 pondo $\chi(2) = \xi$.

Assim, módulo 5:

- Se $\chi(2) = 1$, então $\chi = \chi_0$.
- Se $\chi(2) = -1$, então $\chi = (\frac{\cdot}{5})$.
- Se $\chi(2) = \pm i$, obtemos 2 novos caracteres não-reais (complexos).

Para $\chi(2) = i$, por exemplo,

$$\chi(n) = \begin{cases} 0 & \text{se } n \equiv 0 \pmod 5, \\ 1 & \text{se } n \equiv 1 \pmod 5, \\ i & \text{se } n \equiv 2 \pmod 5, \\ -i & \text{se } n \equiv 3 \pmod 5, \\ -1 & \text{se } n \equiv 4 \pmod 5. \end{cases}$$

Este é o primeiro exemplo de um caractere cuja imagem não está contida em $\mathbb{R}$. Em geral, se $5 \nmid n$ e escrevemos $n = \prod p_{i1}^{\alpha_{i1}} \cdot \prod p_{i2}^{\alpha_{i2}} \cdot \prod p_{i3}^{\alpha_{i3}} \cdot \prod p_{i4}^{\alpha_{i4}}$, com $p_{ij} \equiv j \pmod 5$, então, para $\chi(2) = \pm i$,

$$\chi(n) = (\pm i)^{\sum \alpha_{i2}} (\mp i)^{\sum \alpha_{i3}} (-1)^{\sum \alpha_{i4}}.$$

Os valores de um caractere são sempre 0 ou raízes da unidade:

Proposição 6.14. *Dado $n \in \mathbb{Z}$, $\chi(n)$ é igual a 0 ou a uma raiz da unidade. Mais precisamente, se $\chi(n) \neq 0$, então $\chi(n) = e^{2\pi i \frac{k}{\phi(q)}}$ para algum $k \in \mathbb{Z}$ (que depende de n).*

Demonstração. Basta mostrar que se $\chi(n) \neq 0$, então $\chi(n)^{\phi(q)} = 1$. Ora, se $\chi(n) \neq 0$, pela propriedade iii) da Definição 6.3, $(n, q) = 1$. Então, pelo Teorema de Euler,

$$n^{\phi(q)} \equiv 1 \pmod{q} \iff n^{\phi(q)} = 1 + tq$$

para algum $t \in \mathbb{Z}$. Agora, pela propriedade i),

$$1 = \chi(1) = \chi(1 + tq) = \chi(n^{\phi(q)}).$$

Finalmente, pela propriedade ii),

$$\chi(n)^{\phi(q)} = \chi(n^{\phi(q)}) = 1. \qquad \square$$

6.3 O grupo dos caracteres

Proposição 6.15. *O conjunto $\mathfrak{X}_q := \{\chi \text{ caractere módulo } q\}$ forma um grupo abeliano com respeito à multiplicação ponto-a-ponto, e o inverso de um caractere é igual ao seu conjugado complexo $\overline{\chi}$, dado por $\overline{\chi}(n) = \overline{\chi(n)}$.*

Demonstração. A multiplicação é ponto-a-ponto:

$$(\chi_1 \cdot \chi_2)(n) := \chi_1(n) \cdot \chi_2(n).$$

- Associatividade ($(\chi_1\chi_2)\chi_3 = \chi_1(\chi_2\chi_3)$): segue da associatividade de $\mathbb{C}$.
- Elemento neutro ($\chi \cdot e = e \cdot \chi = \chi$): o elemento neutro é $e = \chi_0$. Temos que mostrar que, para todo χ,
$$\chi(n) \cdot \chi_0(n) = \chi(n), \quad \forall n \in \mathbb{Z}. \tag{6.1}$$
Se $(n, q) = 1$, então $\chi_0(n) = 1$. Logo,
$$\chi(n) \cdot \chi_0(n) = \chi(n) \cdot 1 = \chi(n).$$
Se $(n, q) > 1$, os dois lados de (6.1) são iguais a 0.
- Inverso ($\chi \cdot \chi^{-1} = e$): o elemento inverso é $\chi^{-1} = \overline{\chi}$. Temos que provar
$$\chi(n) \cdot \overline{\chi}(n) = \chi_0(n) \quad \forall n \in \mathbb{Z}. \tag{6.2}$$
Se $(n, q) = 1$, pela Proposição 6.14,
$$\chi(n) = e^{\frac{2\pi i k_n}{\phi(q)}}$$
para algum $k_n \in \mathbb{Z}$. Logo,
$$\overline{\chi}(n) = e^{-\frac{2\pi i k_n}{\phi(q)}},$$
de modo que $\chi(n) \cdot \overline{\chi}(n) = 1 = \chi_0(n)$. Se $(n, q) > 1$, os dois lados de (6.2) se anulam.
- Comutatividade ($\chi\psi = \psi\chi$): segue da comutatividade de $\mathbb{C}$. $\square$

Vamos usar o seguinte teorema sem demonstração, como uma *black box*:

Teorema 6.16 (Caracterização dos grupos abelianos finitos)**.** *Seja G um grupo abeliano finito. Então existem $m_1, m_2, \ldots, m_k$ tais que:*

$$(G, \cdot) \cong (\mathbb{Z}/m_1\mathbb{Z} \times \cdots \times \mathbb{Z}/m_k\mathbb{Z}, +).$$

Exemplo 6.17. Se p é um primo ímpar e $k \geq 1$, existe uma raiz primitiva ξ módulo p^k. Isto significa que existe um isomorfismo entre $(\mathbb{Z}/p^{k-1}(p-1)\mathbb{Z}, +)$ e $((\mathbb{Z}/p^k\mathbb{Z})^\times, \cdot)$:

$$\begin{aligned} \mathbb{Z}/p^{k-1}(p-1)\mathbb{Z} &\longrightarrow (\mathbb{Z}/p^k\mathbb{Z})^\times \\ j &\longmapsto \xi^j. \end{aligned}$$

Para evitar trocar de notação (multiplicativa para aditiva), denote por $(G_m, \cdot)$ o (único) grupo cíclico com m elementos. Ou seja,

$$G_m = \langle \xi \mid \xi^m = 1 \rangle = \{1, \xi, \ldots, \xi^{m-1}\}.$$

Com esta notação, podemos substituir $(\mathbb{Z}/m_1\mathbb{Z} \times \cdots \times \mathbb{Z}/m_k\mathbb{Z}, +)$ por $(G_{m_1} \times \cdots \times G_{m_k}, \cdot)$ no Teorema 6.16. A partir de agora todos os grupos serão multiplicativos, de modo que não vamos lembrar a operação.

Em particular, para todo q, existem $m_1, \ldots, m_k$ tais que:

$$(\mathbb{Z}/q\mathbb{Z})^\times \cong G_{m_1} \times \cdots \times G_{m_k}.$$

Exemplo 6.18. Temos $(\mathbb{Z}/8\mathbb{Z})^\times \cong G_2 \times G_2$ e, pelo Teorema Chinês dos Restos,

$$(\mathbb{Z}/15\mathbb{Z})^\times \cong (\mathbb{Z}/3\mathbb{Z})^\times \times (\mathbb{Z}/5\mathbb{Z})^\times \cong G_2 \times G_4.$$

Aplicando esta decomposição, chegamos ao principal resultado desta seção.

Teorema 6.19. *Seja $\mathfrak{X}_q$ o conjunto dos caracteres módulo q. Então $\#\mathfrak{X}_q = \phi(q)$. Além disso, se $(a, q) = 1$ e $a \not\equiv 1 \pmod q$, existe algum $\chi \in \mathfrak{X}_q$ tal que $\chi(a) \neq 1$.*

Demonstração de $\#\mathfrak{X}_q = \phi(q)$. Para cada $n \in \mathbb{Z}$, denote por $n \pmod q$ a sua classe de congruência módulo q. Observe que $n \pmod q \in (\mathbb{Z}/q\mathbb{Z})^\times$ se e somente se $(n, q) = 1$.

Sejam $m_1, \ldots, m_k$ tais que

$$(\mathbb{Z}/q\mathbb{Z})^\times \cong G_{m_1} \times \cdots \times G_{m_k}. \tag{6.3}$$

Sejam $\xi_1, \ldots, \xi_k$ geradores de $G_{m_1}, \ldots, G_{m_k}$, respectivamente. Seja

$$Z = \mu_{m_1} \times \mu_{m_2} \times \cdots \times \mu_{m_k},$$

onde

$$\begin{aligned} \mu_m &= \{\text{raízes } m\text{-ésimas da unidade}\} \\ &= \{\omega \in \mathbb{C};\ \omega^m = 1\} \\ &= \{e^{\frac{2\pi i j}{m}} \mid j = 0, 1, \ldots, m-1\}. \end{aligned}$$

Em particular, $\#\mu_m = m$ e, consequentemente,

$$\begin{aligned} \#Z &= m_1 m_2 \cdots m_k \\ &= \#\ (G_{m_1} \times \cdots \times G_{m_k}) \\ &= \#\ (\mathbb{Z}/q\mathbb{Z})^\times = \phi(q). \end{aligned}$$

Nosso objetivo é encontrar uma bijeção entre $\mathfrak{X}_q$ e Z.

Para cada $i = 1, \ldots, k$, seja $x_i \in (\mathbb{Z}/q\mathbb{Z})^\times$ correspondendo ao elemento

$$(1, 1, \ldots, 1, \underbrace{\xi_i}_{i\text{-ésima posição}}, 1, \ldots, 1).$$

Construa a seguinte função

$$\begin{aligned}\varphi : \mathfrak{X}_q &\longrightarrow Z \\ \chi &\longmapsto (\chi(n_1), \chi(n_2), \ldots, \chi(n_k)),\end{aligned}$$

onde $n_i \in \mathbb{Z}$ é tal que $n_i \pmod q = x_i$ ($\Longrightarrow (n_i, q) = 1$). Para verificar que φ está bem definida, é necessário mostrar que $\chi(n_i)^{m_i} = 1$. Como x_i corresponde a $(1, 1, \ldots, 1, \xi_i, 1, \ldots, 1)$ pelo homomorfismo (6.3) e $\xi_i^{m_i} = 1$, segue que $x_i^{m_i} = 1$ em $(\mathbb{Z}/q\mathbb{Z})^\times$. Ou seja,

$$n_i^{m_i} \equiv 1 \pmod q.$$

Em particular,

$$\chi(n_i)^{m_i} = \chi(n_i^{m_i}) = \chi(1) = 1,$$

o que mostra que φ está de fato bem definida.

Injetividade de φ. Para mostrar que φ é injetiva, vamos mostrar que $\varphi(\chi)$ determina todos os valores de χ. Em outras palavras, vamos mostrar como determinar o valor de $\chi(n)$ para qualquer $n \in \mathbb{Z}$ a partir dos valores de $\chi(n_1), \ldots, \chi(n_k)$.

Seja $\chi \in \mathfrak{X}_q$ e seja $(\omega_1, \ldots, \omega_k) = \varphi(\chi)$. Dado $n \in \mathbb{Z}$, se $(n, q) > 1$, então $\chi(n) = 0$. Se $(n, q) = 1$, seja

$$x = n \pmod q \in (\mathbb{Z}/q\mathbb{Z})^\times.$$

Pelo homomorfismo (6.3), x corresponde a um elemento de $G_{m_1} \times \cdots \times G_{m_k}$ da forma

$$\begin{aligned}&(\xi_1^{t_1}, \ldots, \xi_k^{t_k}) \\ &= \underbrace{(\xi_1, 1, \ldots, 1)}_{\text{corresponde a } x_1}{}^{t_1} \cdots \underbrace{(1, \ldots, 1, \xi_k)}_{\text{corresponde a } x_k}{}^{t_k}.\end{aligned}$$

Como (6.3) é homomorfismo, segue que $(\xi_1^{t_1}, \ldots, \xi_k^{t_k})$ corresponde a $x_1^{t_1} \cdots x_k^{t_k}$. Ou seja, $x = x_1^{t_1} \cdots x_k^{t_k}$. Como

$$x = n \pmod q, \quad x_i = n_i \pmod q,$$

segue que

$$n \equiv n_1^{t_1} \cdots n_k^{t_k} \pmod q.$$

Portanto,

$$\begin{aligned}\chi(n) &= \chi(n_1)^{t_1} \cdots \chi(n_k)^{t_k} \\ &= \omega_1^{t_1} \cdots \omega_k^{t_k},\end{aligned}$$

mostrando que $\chi(n)$ só depende de $\varphi(\chi) = (\omega_1, \ldots, \omega_k)$.

Sobrejetividade de φ. Para cada $(\omega_1, \ldots, \omega_k) \in Z$, precisamos construir $\chi \in \mathfrak{X}_q$ tal que $\chi(n_i) = \omega_i$, $i = 1, 2, \ldots, k$.

Defina χ da seguinte maneira:

- Se $(n, q) > 1$, $\chi(n) = 0$.
- Se $(n, q) = 1$, pelo argumento acima, existem $t_1, \ldots, t_k$ tais que

  $$n \equiv n_1^{t_1} \cdots n_k^{t_k} \pmod q.$$

  Então defina

  $$\chi(n) = \omega_1^{t_1} \cdots \omega_k^{t_k}.$$

Observação 6.20. Os números $t_1, \ldots, t_k$ não são únicos mas, como veremos, isto não será um problema.

Para ver que χ está bem definida, suponha que $t'_1, \ldots, t'_k$ também são tais que

$$n \equiv n_1^{t'_1} \cdots n_k^{t'_k} \pmod{q}.$$

Então temos

$$n_1^{t'_1} \cdots n_k^{t'_k} \equiv n_1^{t_1} \cdots n_k^{t_k} \pmod{q}.$$

Ou seja,

$$x_1^{t'_1} \cdots x_k^{t'_k} = x_1^{t_1} \cdots x_k^{t_k} \iff x_1^{t_1 - t'_1} \cdots x_k^{t_k - t'_k} = 1.$$

Pela correspondência (6.3), isto significa que

$$(\xi_1^{t_1 - t'_1}, \xi_2^{t_2 - t'_2}, \ldots, \xi_k^{t_k - t'_k}) = (1, \ldots, 1).$$

Isto quer dizer que

$$\xi_i^{t_i - t'_i} = 1, \quad i = 1, \ldots, k.$$

Como a ordem de ξ_i é m_i, segue que

$$t_i - t'_i = m_i \cdot s_i$$

para algum $s_i \in \mathbb{Z}$. Agora, como ω_i é uma raiz m_i-ésima da unidade,

$$\omega_i^{t_i - t'_i} = 1 \Longrightarrow \omega_i^{t_i} = \omega_i^{t'_i}.$$

De modo que

$$\omega_1^{t'_1} \cdots \omega_k^{t'_k} = \omega_1^{t_1} \cdots \omega_k^{t_k},$$

concluindo a prova de que χ está bem definida.

A construção acima só depende da classe de congruência de $n \pmod{q}$, e por construção $\chi(n) \neq 0 \iff (n, q) = 1$. De modo que as propriedades i) e iii) da Definição 6.3 são satisfeitas. Resta mostrar que a definição acima de χ fornece uma função completamente multiplicativa. Para ver isto, dividimos em dois casos:

- Se $(m, q) > 1$ ou $(n, q) > 1$, segue que $(mn, q) > 1$ e, portanto, $\chi(mn) = 0 = \chi(m)\chi(n)$.
- Se $(m, q) = (n, q) = 1$, basta notar que se

  $$\begin{aligned} m &\equiv n_1^{t_1} \cdots n_k^{t_k} \pmod{q} \\ n &\equiv n_1^{l_1} \cdots n_k^{l_k} \pmod{q} \end{aligned}$$

  então

  $$mn \equiv n_1^{t_1 + l_1} \cdots n_k^{t_k + l_k} \pmod{q},$$

  de modo que

  $$\begin{aligned} \omega_1^{t_1 + l_1} \cdots \omega_k^{t_k + l_k} &= (\omega_1^{t_1} \cdots \omega_k^{t_k})(\omega_1^{l_1} \cdots \omega_k^{l_k}) \\ &\iff \chi(mn) = \chi(m)\chi(n). \end{aligned}$$

Ou seja, χ é um caractere e, por construção, $\varphi(\chi) = (\omega_1, \ldots, \omega_k)$. Como as raízes da unidade ω_i podem ser escolhidas arbitrariamente, provamos que φ é sobrejetiva. □

Demonstração da parte final do Teorema 6.19. Seja a com $(a, q) = 1$ e $a \not\equiv 1 \pmod{q}$ e, com a notação da demonstração acima, escreva $a \equiv n_1^{t_1} \cdots n_k^{t_k} \pmod{q}$. Como a classe de a não é o elemento neutro de $(\mathbb{Z}/q\mathbb{Z})^\times$, existe j tal que $\xi_j^{t_j} \neq 1$, isto é, tal que $m_j \nmid t_j$. Como φ é sobrejetiva, existe $\chi \in \mathfrak{X}_q$ tal que $\varphi(\chi) = (1, \ldots, 1, e^{2\pi i/m_j}, 1, \ldots, 1)$, com $e^{2\pi i/m_j}$ na j-ésima posição. Então $\chi(a) = \chi(n_j)^{t_j} = e^{2\pi i t_j/m_j} \neq 1$. □

6.4 Relações de ortogonalidade

Vejamos como utilizar o Teorema 6.19 para provar o seguinte:

Teorema 6.21 (Ortogonalidade de caracteres)**.** *Seja $q \in \mathbb{Z}_{>0}$. Então valem:*

i) Para todo χ caractere módulo q,

$$\sum_{a=1}^{q} \chi(a) = \begin{cases} \phi(q) & \text{se } \chi = \chi_0, \\ 0 & \text{caso contrário.} \end{cases}$$

ii) Para todo $a \in \mathbb{Z}$,

$$\sum_{\chi \bmod q} \chi(a) = \begin{cases} \phi(q) & \text{se } a \equiv 1 \pmod q, \\ 0 & \text{caso contrário.} \end{cases}$$

A soma acima é tomada sobre todos os caracteres módulo q, isto é, sobre $\chi \in \mathfrak{X}_q$.

iii) Dados χ_1, χ_2 caracteres módulo q,

$$\sum_{a=1}^{q} \chi_1(a)\overline{\chi_2(a)} = \begin{cases} \phi(q) & \text{se } \chi_1 = \chi_2, \\ 0 & \text{caso contrário.} \end{cases}$$

iv) Dados $a_1, a_2 \in \mathbb{Z}$ com $(a_1, q) = (a_2, q) = 1$,

$$\sum_{\chi \bmod q} \chi(a_1)\overline{\chi(a_2)} = \begin{cases} \phi(q) & \text{se } a_1 \equiv a_2 \pmod q, \\ 0 & \text{caso contrário.} \end{cases}$$

Aqui, mais uma vez, a soma é tomada sobre todos os caracteres módulo q.

Observação 6.22. Nos itens i) e iii) podemos trocar a soma sobre $1 \leq a \leq q$ por $a \in S$, onde S é um Sistema Completo de Restos módulo q.

Demonstração. i) Se $\chi = \chi_0$ é o caractere principal,

$$\sum_{a=1}^{q} \chi_0(a) = \#\{1 \leq a \leq q : (a, q) = 1\} = \phi(q).$$

Seja agora $\chi \neq \chi_0$. Por definição, existe b tal que $(b, q) = 1$ e $\chi(b) \neq 1$. Segue de um fato básico de teoria dos números que o conjunto $\{b, 2b, \ldots, bq\}$ é um sistema completo de restos módulo q. Assim, seja

$$S = \sum_{a=1}^{q} \chi(a).$$

Então

$$\chi(b)S = \sum_{a=1}^{q} \chi(b)\chi(a) = \sum_{a=1}^{q} \chi(ab) = \sum_{a' \in \{b, \ldots, bq\}} \chi(a').$$

Pela Observação 6.22, temos que

$$\sum_{a' \in \{b, \ldots, bq\}} \chi(a') = S.$$

Ou seja,

$$\chi(b)S = S \implies (\chi(b) - 1)S = 0.$$

Como $\chi(b) \neq 1$, segue que $S = 0$.

ii) A prova é semelhante. Se $a \equiv 1 \pmod{q}$, então é claro que $\chi(a) = 1$ para todo caractere χ módulo q. Portanto, a soma é igual a

$$\#\{\text{caracteres módulo } q\} = \phi(q).$$

Se $(a, q) > 1$, todos os termos $\chi(a)$ são nulos. Seja agora $a \not\equiv 1 \pmod{q}$ com $(a, q) = 1$. Pelo Teorema 6.19, existe χ_1 tal que $\chi_1(a) \neq 1$. Seja

$$S = \sum_{\chi \bmod q} \chi(a).$$

Então

$$\chi_1(a)S = \sum_{\chi \bmod q} \chi(a)\chi_1(a) = \sum_{\chi \bmod q} (\chi\chi_1)(a).$$

Lembre que os caracteres de Dirichlet formam um grupo com respeito à multiplicação ponto-a-ponto (Proposição 6.15). Isso significa que o conjunto $\{\chi\chi_1 : \chi \bmod q\}$ é simplesmente o próprio conjunto dos caracteres módulo q. Assim,

$$\chi_1(a)S = \sum_{\chi \bmod q} (\chi\chi_1)(a) = \sum_{\chi \bmod q} \chi(a) = S.$$

Assim como no caso anterior, como $\chi_1(a) \neq 1$, segue que $S = 0$.

iii) Segue de i), utilizando mais uma vez que o conjunto de caracteres módulo q é grupo. Vale lembrar que o elemento neutro é χ_0 e $\chi_1\overline{\chi_2} = \chi_1\chi_2^{-1}$.

iv) Segue de ii) usando a estrutura de grupo de $(\mathbb{Z}/q\mathbb{Z})^\times$. □

Referência para caracteres: Serre, *A Course in Arithmetic*, Capítulo VI.

6.4.1 Decomposição de funções periódicas

As relações de ortogonalidade nos permitem construir uma "Teoria de Fourier" para funções $f : \mathbb{Z} \to \mathbb{C}$ com período q e tais que $f(n) = 0$ sempre que $(n, q) > 1$.

Exemplo 6.23. $f = \mathbf{1}_{(n,q)=1}$. Não há o que fazer: $f = \chi_0$.

Exemplo 6.24. Seja $q = p$ um primo ímpar e seja

$$f(n) = \begin{cases} 1 & \text{se } n \text{ é resíduo quadrático módulo } p, \\ 0 & \text{caso contrário.} \end{cases}$$

Por definição, 0 não é considerado resíduo quadrático.

Já sabemos que o símbolo de Legendre $\chi(n) = \left(\frac{n}{p}\right)$ é um caractere. Com essa escolha de χ,

$$\chi(n) = \begin{cases} 1 & \text{se } n \text{ é um resíduo quadrático,} \\ -1 & \text{se } n \text{ é um resíduo não-quadrático,} \\ 0 & \text{se } n \equiv 0 \pmod{p}. \end{cases}$$

Somando 1 e dividindo por dois, obtemos um resultado muito próximo da função f acima:

$$\frac{1 + \chi(n)}{2} = \begin{cases} f(n) & \text{se } n \not\equiv 0 \pmod{p}, \\ \frac{1}{2} & \text{se } n \equiv 0 \pmod{p}. \end{cases}$$

Na verdade, a função constante 1 não tem suporte nos números com $(n, q) = 1$. Seu parente mais próximo com essa propriedade é o caractere principal χ_0. E, de fato, podemos escrever

$$f(n) = \frac{1}{2}\chi_0(n) + \frac{1}{2}\chi(n).$$

Exemplo 6.25. Seja $(a,q)=1$ e considere

$$f(n)=\mathbf{1}_{n\equiv a\,(q)}.$$

Pelo item iv) do Teorema 6.21, se $(n,q)=1$,

$$\sum_{\chi \bmod q}\chi(n)\overline{\chi}(a)=\begin{cases}\phi(q) & \text{se } n\equiv a \pmod q,\\ 0 & \text{caso contrário.}\end{cases}$$

Se $(n,q)>1$, temos $\chi(n)=0$ para todo χ e $n\not\equiv a \pmod q$, de modo que a fórmula continua correta. Ou seja,

$$\mathbf{1}_{n\equiv a\,(q)}=\sum_{\chi \bmod q}\frac{\overline{\chi}(a)}{\phi(q)}\chi(n). \tag{6.4}$$

Na verdade, a partir deste exemplo podemos mostrar que toda função $f:\mathbb{Z}\to\mathbb{C}$ de período q com $f(n)=0$ para $(n,q)>1$ se escreve como soma de caracteres. De fato, para todo $n\in\mathbb{Z}$,

$$f(n)=\sum_{\substack{a=1\\(a,q)=1}}^{q} f(a)\cdot\mathbf{1}_{n\equiv a\,(q)}.$$

Usando a fórmula (6.4), podemos provar o seguinte resultado:

Teorema 6.26. *Seja $f:\mathbb{Z}\to\mathbb{C}$ de período q com $f(n)=0$ se $(n,q)>1$. Então existem números complexos c_χ, um para cada caractere de Dirichlet χ módulo q, tais que*

$$f(n)=\sum_{\chi \bmod q}c_\chi\chi(n).$$

Além disso, c_χ pode ser obtido da seguinte maneira:

$$c_\chi=\frac{1}{\phi(q)}\sum_{\substack{a=1\\(a,q)=1}}^{q} f(a)\overline{\chi}(a).$$

6.5 Funções L e a prova do Teorema de Dirichlet

A prova do Teorema de Dirichlet passa pelo estudo das funções L.

Definição 6.27 (Função L de Dirichlet)**.** A série de Dirichlet de um caractere χ, por motivos históricos, é chamada de *função L de Dirichlet* associada a χ e denotada por $L(s,\chi)$:

$$L(s,\chi)=\sum_{n=1}^{\infty}\frac{\chi(n)}{n^s},\quad \mathrm{Re}(s)>1.$$

Como $\chi(n)=0$ ou $\chi(n)$ é uma raiz da unidade, a série converge absolutamente para $\mathrm{Re}(s)>1$.

6.5.1 Extensão à região $\mathrm{Re}(s)>0$

Nosso objetivo é mostrar que estas funções possuem extensão meromorfa à região $\mathrm{Re}(s)>0$. Para isto, usamos o seguinte resultado:

Lema 6.28. *Seja $q\in\mathbb{Z}_{>0}$ e χ um caractere módulo q.*

i) Se $\chi \neq \chi_0$,

$$\Big|\sum_{n\leq x}\chi(n)\Big| \leq \phi(q).$$

ii) Se $\chi = \chi_0$,

$$\Big|\sum_{n\leq x}\chi_0(n) - \frac{\phi(q)}{q}x\Big| \leq 2\phi(q).$$

Demonstração. Pelas relações de ortogonalidade (Teorema 6.21), se $\chi \neq \chi_0$,

$$\sum_{n=1}^{q}\chi(n) = 0.$$

Pela periodicidade, temos, para todo $k \geq 1$,

$$\sum_{n=(k-1)q+1}^{kq}\chi(n) = 0.$$

Logo,

$$\begin{aligned}\sum_{n\leq x}\chi(n) &= \sum_{n=1}^{q}\chi(n) + \cdots + \sum_{n=(k-1)q+1}^{kq}\chi(n) + \sum_{kq+1\leq n\leq x}\chi(n)\\ &= \sum_{kq+1\leq n\leq x}\chi(n),\end{aligned}$$

onde $k = \left\lfloor \frac{x}{q}\right\rfloor$. Assim,

$$\begin{aligned}\Big|\sum_{n\leq x}\chi(n)\Big| &\leq \sum_{kq+1\leq n\leq x}|\chi(n)|\\ &= \#\{kq+1 \leq n \leq x;\ (n,q)=1\}.\end{aligned}$$

Como $x < (k+1)q$, esse conjunto está contido em $\{kq+1,\ldots,(k+1)q\}$, que é um Sistema Completo de Restos módulo q e, portanto, tem exatamente $\phi(q)$ elementos coprimos com q.

Para o caso $\chi = \chi_0$, a prova é semelhante, com a diferença que

$$\sum_{n=1}^{q}\chi_0(n) = \phi(q).$$

Repetindo o argumento acima, segue que, com o mesmo $k = \left\lfloor \frac{x}{q}\right\rfloor$,

$$\begin{aligned}\sum_{n\leq x}\chi_0(n) &= \left\lfloor\frac{x}{q}\right\rfloor\phi(q) + \sum_{kq+1\leq n\leq x}\chi_0(n)\\ &= \frac{\phi(q)}{q}x - \left\{\frac{x}{q}\right\}\phi(q) + \sum_{kq+1\leq n\leq x}\chi_0(n).\end{aligned}$$

É fácil ver que os dois últimos termos são limitados por $\phi(q)$: usa-se $|\{y\}| \leq 1$ para todo y e o argumento final do caso $\chi \neq \chi_0$, respectivamente. □

Para q fixo, o Lema 6.28 diz que

$$\sum_{n\le x}\chi(n)=\begin{cases}O(1) & \text{se } \chi\neq\chi_0,\\ \frac{\phi(q)}{q}x+O(1) & \text{se } \chi=\chi_0.\end{cases}$$

Pelo Teorema 4.6, $L(s,\chi)$ possui continuação meromorfa na região $\mathrm{Re}(s)>0$, com um único polo (simples) em $s=1$ se $\chi=\chi_0$ e, além disso,

$$\mathrm{Res}_{s=1}L(s,\chi_0)=\frac{\phi(q)}{q}.$$

Exercício 6.29. Sejam $\sigma_c(\chi)$ e $\sigma_a(\chi)$ as abscissas de convergência simples e absoluta de $L(s,\chi)$, respectivamente. Mostre que $\sigma_a(\chi)=1$ e

$$\sigma_c(\chi)=\begin{cases}0 & \text{se } \chi\neq\chi_0,\\ 1 & \text{se } \chi=\chi_0.\end{cases}$$

6.5.2 Produto de Euler

Como χ é completamente multiplicativa, vale o produto de Euler para $\mathrm{Re}(s)>1$:

$$L(s,\chi)=\prod_p\Big(1+\frac{\chi(p)}{p^s}+\frac{\chi(p^2)}{p^{2s}}+\cdots\Big)=\prod_p\Big(1-\frac{\chi(p)}{p^s}\Big)^{-1}.$$

Em particular, para $\chi=\chi_0$:

$$L(s,\chi_0)=\prod_p\Big(1-\frac{\chi_0(p)}{p^s}\Big)^{-1}=\prod_{p\nmid q}\Big(1-\frac{1}{p^s}\Big)^{-1}.$$

Obtemos

$$L(s,\chi_0)=\prod_p\Big(1-\frac{1}{p^s}\Big)^{-1}\prod_{p\mid q}\Big(1-\frac{1}{p^s}\Big)=\zeta(s)\prod_{p\mid q}\Big(1-\frac{1}{p^s}\Big).$$

Como $\prod_{p\mid q}(1-\frac{1}{p^s})$ é uma função inteira e ζ admite uma extensão meromorfa a $\mathrm{Re}(s)>0$ (Teorema 4.5), o mesmo vale para $L(s,\chi_0)$, utilizando a igualdade acima.

Observação 6.30. Na verdade, ζ admite extensão meromorfa a todo o plano complexo; esse fato clássico não é provado nem usado aqui.

Em particular,

$$\{\text{Zeros de } L(s,\chi_0)\}\subset\{\text{Zeros de } \zeta(s)\}\cup\Big\{\text{Zeros de }\Big(1-\frac{1}{p^s}\Big),\ p\mid q\Big\},$$

porém todos os zeros de $1-\frac{1}{p^s}$ estão em $\{\mathrm{Re}(s)=0\}$:

$$\begin{aligned}1-\frac{1}{p^s}=0 &\iff p^s=1\iff\\ &\iff e^{s\cdot\log p}=1\iff\\ &\iff s=\frac{2ik\pi}{\log p},\ k\in\mathbb{Z}.\end{aligned}$$

Portanto, perto de $s=1$, as funções $L(s,\chi_0)$ e $\zeta(s)$ são bem parecidas.

6.5.3 Redução à não-anulação e prova do Teorema de Dirichlet

Nosso objetivo é estudar a série

$$\sum_{p\equiv a\,(q)} \frac{1}{p^s}.$$

Pela fórmula (6.4), para $(a,q)=1$, sabe-se que

$$\sum_{p\equiv a\,(q)} \frac{1}{p^s} = \sum_{\chi \bmod q} \frac{\overline{\chi}(a)}{\phi(q)} \sum_p \frac{\chi(p)}{p^s}. \tag{6.5}$$

Isso nos conduz ao estudo de

$$\sum_p \frac{\chi(p)}{p^s}.$$

Lembre que, na Seção 4.4, estudamos a série $\sum_p \frac{1}{p^s}$ para s real. O mesmo cálculo vale para s complexo com $\mathrm{Re}(s)>1$ e mostra que

$$\sum_p \frac{1}{p^s} = \log\zeta(s) - \sum_p \sum_{m\geq 2} \frac{1}{mp^{ms}},$$

onde a segunda série converge absolutamente para $\mathrm{Re}(s)>1/2$. Em particular, para $\mathrm{Re}(s)>1$ (com $O(1)$ uniforme, pois o resto é limitado por $\sum_p \frac{1}{p(p-1)}$), temos

$$\sum_p \frac{1}{p^s} = \log\zeta(s) + O(1).$$

Como os caracteres χ são completamente multiplicativos, podemos fazer algo semelhante. Do produto de Euler,

$$\log L(s,\chi) = \sum_p \sum_{m\geq 1} \frac{\chi(p)^m}{mp^{ms}} \iff$$

$$\iff \sum_p \frac{\chi(p)}{p^s} = \log L(s,\chi) - \sum_p \sum_{m\geq 2} \frac{\chi(p)^m}{mp^{ms}}$$

$$\implies \sum_p \frac{\chi(p)}{p^s} = \log L(s,\chi) + O(1).$$

Isso nos leva ao estudo de $\log L(s,\chi)$.

Primeiramente, se $\chi=\chi_0$, da expressão de $L(s,\chi_0)$ em termos de $\zeta(s)$ segue que

$$\log L(s,\chi_0) = \log\zeta(s) + O(1).$$

Assim, separando a contribuição de $\chi=\chi_0$ na fórmula (6.5) e usando que $\overline{\chi_0}(a)=1$ pois $(a,q)=1$, temos

$$\sum_{p\equiv a\,(q)} \frac{1}{p^s} = \frac{1}{\phi(q)}\log\zeta(s) + \frac{1}{\phi(q)} \sum_{\substack{\chi \bmod q \\ \chi\neq\chi_0}} \overline{\chi}(a)\log L(s,\chi) + O(1).$$

Queremos tomar $s=\sigma>1$ e fazer $\sigma\to 1^+$. A ideia é que o primeiro termo à direita seja a "parte principal" e o segundo termo seja a parte oscilante (contribua para o termo de erro). Para isto, iremos mostrar que, para $\chi\neq\chi_0$, o limite $L_\chi = \lim_{s\to 1} L(s,\chi)$ existe e é diferente de 0. Isto garante que

$$\frac{1}{\phi(q)} \sum_{\chi\neq\chi_0} \overline{\chi}(a)\log L(s,\chi) = O(1).$$

De fato, aqui $\log L(s,\chi)$ é a série dupla acima. Para $s=\sigma>1$ real, ela é um logaritmo contínuo de $L(\sigma,\chi)$, pois, pelo produto de Euler, a sua exponencial é $L(\sigma,\chi)$. Pelo Teorema 6.31, $L(\cdot,\chi)$ é contínua em $[1,2]$ e não se anula aí (em $(1,2]$ pelo produto de Euler, e em $s=1$ por hipótese); logo ela tem um logaritmo contínuo em todo o compacto $[1,2]$, que é limitado. Dois logaritmos contínuos da mesma função em $(1,2]$ diferem por uma constante $2\pi ik$, e portanto $\log L(\sigma,\chi)$ é limitado para $1<\sigma\leq 2$.

Consequentemente,

$$\sum_{p\equiv a\,(q)}\frac{1}{p^s}=\frac{1}{\phi(q)}\log\zeta(s)+O(1),$$

provando o Teorema de Dirichlet (Teorema 6.1).

Mais precisamente, vale o seguinte resultado mais geral, cuja parte sobre convergência e extensão já foi obtida acima; a não-anulação em $s=1$, que é o ingrediente principal da prova do Teorema de Dirichlet, será demonstrada na Seção 6.6.

Teorema 6.31 (Extensão meromorfa e não-anulação de $L(s,\chi)$)**.** *Seja χ um caractere de Dirichlet. Então a série definindo $L(s,\chi)$ converge absolutamente para $\mathrm{Re}(s)>1$ e possui extensão meromorfa à região $\{s\in\mathbb{C};\mathrm{Re}(s)>0\}$ com único (possível) polo em $s=1$, que só ocorre no caso em que $\chi=\chi_0$. Além disso, quando $\chi\neq\chi_0$,*

$$L(1,\chi)\neq 0.$$

6.6 A não-anulação de $L(1,\chi)$

Vamos demonstrar que, para χ não principal, $L(1,\chi)\neq 0$ (última afirmação do Teorema 6.31). Dividimos a prova em dois casos, a depender de χ ser real ou não-real. Lembre que χ é real se todos os valores de $\chi(n)$ são números reais. Note que isso significa $\chi(n)\in\{-1,0,1\}$ ou, alternativamente, que $\chi^2(n)=\chi_0(n)$. O caso não-real é o mais fácil.

6.6.1 Caso χ não-real

Demonstração para χ não-real. Na prova de que $\zeta(1+it)\neq 0$, $t\neq 0$ (Seção 5.2), utilizamos a seguinte expressão:

$$\log\zeta(\sigma+it)=\sum_p\sum_{m\geq 1}\frac{1}{mp^{m\sigma}}p^{-imt},$$

junto com

$$\log\zeta(\sigma)=\sum_p\sum_{m\geq 1}\frac{1}{mp^{m\sigma}}\cdot 1,\qquad \log\zeta(\sigma+2it)=\sum_p\sum_{m\geq 1}\frac{1}{mp^{m\sigma}}\,e^{-2imt\log p}.$$

De maneira análoga, para $\sigma>1$ temos:

$$\begin{aligned}\log L(\sigma,\chi)&=\log\prod_p\Big(1-\frac{\chi(p)}{p^\sigma}\Big)^{-1}\\&=\sum_p\log\Big[\Big(1-\frac{\chi(p)}{p^\sigma}\Big)^{-1}\Big]\\&=\sum_p\sum_{m\geq 1}\frac{1}{mp^{m\sigma}}\chi(p)^m.\end{aligned}$$

Note que

$$\chi(p)=\begin{cases}0 & \text{se } p\mid q,\\ e^{i\theta_p}\text{ para algum }\theta_p\in\mathbb{R} & \text{se } p\nmid q.\end{cases}$$

Substituindo acima, temos que

$$\log L(\sigma,\chi) = \sum_{p\nmid q}\sum_{m=1}^{\infty}\frac{1}{mp^{m\sigma}}e^{im\theta_p}.$$

De maneira análoga:

$$\log L(\sigma,\chi^2) = \sum_{p\nmid q}\sum_{m=1}^{\infty}\frac{1}{mp^{m\sigma}}e^{2im\theta_p}, \quad \log L(\sigma,\chi_0) = \sum_{p\nmid q}\sum_{m=1}^{\infty}\frac{1}{mp^{m\sigma}}\cdot 1.$$

Assim, definindo

$$R(\sigma) = 3\log L(\sigma,\chi_0) + 4\log L(\sigma,\chi) + \log L(\sigma,\chi^2),$$

tem-se

$$\operatorname{Re}(R(\sigma)) = \sum_{p\nmid q}\sum_{m=1}^{\infty}\frac{1}{mp^{m\sigma}}\operatorname{Re}(3+4e^{im\theta_p}+e^{2im\theta_p}).$$

Agora,

$$\operatorname{Re}(3+4e^{im\theta_p}+e^{2im\theta_p}) = 3+4\cos(m\theta_p)+\cos(2m\theta_p) \geq 0$$

pela identidade trigonométrica 3-4-1: $3+4\cos u+\cos 2u = 2(1+\cos u)^2$.

Consequentemente,

$$\operatorname{Re}(R(\sigma)) = \log|L(\sigma,\chi_0)^3L(\sigma,\chi)^4L(\sigma,\chi^2)| \geq 0.$$

Donde segue que

$$|L(\sigma,\chi_0)^3L(\sigma,\chi)^4L(\sigma,\chi^2)| \geq 1. \tag{6.6}$$

O resto do argumento segue exatamente como no Teorema dos Números Primos. Suponha que $L(s,\chi)$ tenha um zero em $s=1$. Isso significaria que $L(s,\chi) = (s-1)F(s)$, com $F(s)$ holomorfa em uma vizinhança de $s=1$. Por outro lado, sabemos que:

- $L(s,\chi_0) = \frac{G(s)}{s-1}$ (com G holomorfa),
- $L(s,\chi^2) = H(s)$ (com H holomorfa, pois $\chi^2 \neq \chi_0$).

Substituindo em (6.6), obtemos

$$|(\sigma-1)F(\sigma)^4G(\sigma)^3H(\sigma)| \geq 1,$$

com F, G, H holomorfas em $s=1$. Fazendo $\sigma \to 1^+$, temos uma contradição. □

6.6.2 Caso χ real

Considere a função de convolução $f = 1 * \chi$. Vamos precisar dos seguintes resultados:

Lema 6.32. *Seja χ um caractere real e f como acima. Então*

$$f(n) \geq \begin{cases} 1 & \text{se } n \text{ é um quadrado,} \\ 0 & \text{caso contrário.} \end{cases}$$

Demonstração. Como χ é multiplicativa, f também o é. Logo, basta verificar a desigualdade acima para potências de primos $n = p^k$. Escreva

$$f(p^k) = \sum_{j=0}^{k}\chi(p^j).$$

Como χ é real, $\chi(p) \in \{-1,0,1\}$. Vamos provar a desigualdade em cada um dos casos.

1. Se $\chi(p) = -1$: $f(p^k) = 1 + (-1) + 1 + (-1) + \cdots = \begin{cases} 0 & \text{se } k \text{ é ímpar,} \\ 1 & \text{se } k \text{ é par.} \end{cases}$
2. Se $\chi(p) = 0$: $f(p^k) = 1 + 0 + \cdots + 0 = 1$.
3. Se $\chi(p) = 1$: $f(p^k) = 1 + 1 + \cdots + 1 = k + 1$.

Veja que em todos os casos temos $f(p^k) \geq 0$ e, se k é par, vale que $f(p^k) \geq 1$, concluindo a prova do lema. □

O lema abaixo segue, por exemplo, somando por partes (Seção 2.3).

Lema 6.33. *Existe uma constante A tal que*

$$\sum_{n \leq x} \frac{1}{\sqrt{n}} = 2\sqrt{x} + A + O\Big(\frac{1}{\sqrt{x}}\Big),$$

a saber,

$$A = -1 - \frac{1}{2}\int_1^{+\infty} \{t\}\, t^{-3/2}\, dt.$$

O próximo resultado já foi visto no caso da função zeta (Seção 5.3) e a prova é análoga. Vamos escrevê-la aqui por conveniência do leitor.

Lema 6.34. *Seja $\chi \neq \chi_0$ e $\mathrm{Re}(s) = \sigma > 0$. Então*

$$\sum_{n \leq x} \frac{\chi(n)}{n^s} = L(s,\chi) + O(x^{-\sigma}).$$

Demonstração. Seja $S_\chi(t) = \sum_{n \leq t} \chi(n)$. Usando a soma por partes (Seção 2.3),

$$\sum_{n \leq x} \frac{\chi(n)}{n^s} = \frac{S_\chi(x)}{x^s} + s\int_1^x S_\chi(t) t^{-s-1}\, dt.$$

Tomando $\mathrm{Re}(s) > 1$ e fazendo $x \to \infty$, temos

$$L(s,\chi) = s\int_1^{+\infty} S_\chi(t) t^{-s-1}\, dt.$$

Além disso, como $S_\chi(t) = O(1)$ (Lema 6.28), a expressão do lado direito fornece uma extensão de $L(s,\chi)$ para $\mathrm{Re}(s) > 0$. Subtraindo as expressões:

$$\begin{aligned} \sum_{n \leq x} \frac{\chi(n)}{n^s} &= L(s,\chi) + \frac{S_\chi(x)}{x^s} - s\int_x^{+\infty} S_\chi(t) t^{-s-1}\, dt \\ &= L(s,\chi) + O(x^{-\sigma}) + O\Big(\int_x^{\infty} t^{-\sigma-1}\, dt\Big) \\ &= L(s,\chi) + O(x^{-\sigma}). \end{aligned}$$

□

Demonstração para χ real não principal. A ideia é mostrar que a função $f = 1 * \chi$ é "grande demais" para que D_f seja holomorfa em $s = 1$.

Considere a soma

$$S(x) = \sum_{n \leq x} \frac{f(n)}{\sqrt{n}}.$$

Pela minoração do Lema 6.32, temos que

$$S(x) \geq \sum_{m\leq\sqrt{x}} \frac{1}{m} \gg \log x. \tag{6.7}$$

Por outro lado, pela definição de f:

$$S(x) = \sum_{n\leq x} \frac{1}{\sqrt{n}} \sum_{n_1 n_2 = n} \chi(n_2) = \sum_{n_1 n_2 \leq x} \frac{1}{\sqrt{n_1 n_2}} \chi(n_2).$$

Pelo método da hipérbole (Seção 2.5), com $y = \sqrt{x}$:

$$S(x) = S_I + S_{II} - S_{III},$$

onde

$$\begin{cases} S_I = \displaystyle\sum_{n_1\leq\sqrt{x}} \frac{1}{\sqrt{n_1}} \sum_{n_2\leq x/n_1} \frac{\chi(n_2)}{\sqrt{n_2}}, \\ S_{II} = \displaystyle\sum_{n_2\leq\sqrt{x}} \frac{\chi(n_2)}{\sqrt{n_2}} \sum_{n_1\leq x/n_2} \frac{1}{\sqrt{n_1}}, \\ S_{III} = \Big(\displaystyle\sum_{n_1\leq\sqrt{x}} \frac{1}{\sqrt{n_1}}\Big)\Big(\sum_{n_2\leq\sqrt{x}} \frac{\chi(n_2)}{\sqrt{n_2}}\Big). \end{cases}$$

Usando os Lemas 6.33 e 6.34, temos que

$$\begin{aligned} S_I &= \sum_{n_1\leq\sqrt{x}} \frac{1}{\sqrt{n_1}} \Big(L(\tfrac{1}{2},\chi) + O\big((x/n_1)^{-1/2}\big)\Big) \\ &= L(\tfrac{1}{2},\chi)(2x^{1/4} + A + O(x^{-1/4})) + O\Big(\frac{1}{\sqrt{x}} \sum_{n_1\leq\sqrt{x}} 1\Big) \\ &= 2L(\tfrac{1}{2},\chi)x^{1/4} + O(1), \end{aligned}$$

$$\begin{aligned} S_{II} &= \sum_{n_2\leq\sqrt{x}} \frac{\chi(n_2)}{\sqrt{n_2}} \Big(2\sqrt{\frac{x}{n_2}} + A + O\Big(\Big(\frac{x}{n_2}\Big)^{-1/2}\Big)\Big) \\ &= 2\sqrt{x} \sum_{n_2\leq\sqrt{x}} \frac{\chi(n_2)}{n_2} + A \sum_{n_2\leq\sqrt{x}} \frac{\chi(n_2)}{\sqrt{n_2}} + O(1) \\ &= 2\sqrt{x}(L(1,\chi) + O(x^{-1/2})) + A(L(\tfrac{1}{2},\chi) + O(x^{-1/4})) + O(1) \\ &= 2L(1,\chi)\sqrt{x} + O(1), \end{aligned}$$

$$\begin{aligned} S_{III} &= (2x^{1/4} + A + O(x^{-1/4}))(L(\tfrac{1}{2},\chi) + O(x^{-1/4})) \\ &= 2x^{1/4}L(\tfrac{1}{2},\chi) + O(1). \end{aligned}$$

De modo que

$$S(x) = S_I + S_{II} - S_{III} = 2L(1,\chi)\sqrt{x} + O(1).$$

Segue que, se $L(1,\chi) = 0$, teríamos $S(x) = O(1)$, o que contradiz a estimativa (6.7).

Logo, temos que $L(1,\chi) \neq 0$ também no caso em que χ é real, concluindo a prova do Teorema 6.31 e, com ela, a do Teorema de Dirichlet. □

Referências Bibliográficas

[Dav00] H. Davenport, *Multiplicative Number Theory*, 3ª ed., revista por H. L. Montgomery, Graduate Texts in Mathematics 74, Springer, 2000.

[Hil05] A. J. Hildebrand, *Introduction to Analytic Number Theory*, Math 531 Lecture Notes, Fall 2005 (versão 2013.01.07), University of Illinois at Urbana-Champaign.

[Ten15] G. Tenenbaum, *Introduction to Analytic and Probabilistic Number Theory*, 3ª ed., Graduate Studies in Mathematics 163, American Mathematical Society, 2015.