Introdução aos Métodos de Crivos em Teoria dos Números

Publicações Matemáticas

Introdução aos Métodos de Crivos em Teoria dos Números

Júlio Andrade
Brown University
University of Bristol

impa 60 anos 1952 - 2012

29º Colóquio Brasileiro de Matemática

Impresso no Brasil / Printed in Brazil
Capa: Noni Geiger / Sérgio R. Vaz

29º Colóquio Brasileiro de Matemática

- Análise em Fractais – Milton Jara
- Asymptotic Models for Surface and Internal Waves – Jean-Claude Saut
- Bilhares: Aspectos Físicos e Matemáticos – Alberto Saa e Renato de Sá Teles
- Controle Ótimo: Uma Introdução na Forma de Problemas e Soluções – Alex L. de Castro
- Eigenvalues on Riemannian Manifolds – Changyu Xia
- Equações Algébricas e a Teoria de Galois – Rodrigo Gondim, Maria Eulalia de Moraes Melo e Francesco Russo
- Ergodic Optimization, Zero Temperature Limits and the Max-Plus Algebra – Alexandre Baraviera, Renaud Leplaideur e Artur Lopes
- Expansive Measures – Carlos A. Morales e Víctor F. Sirvent
- Funções de Operador e o Estudo do Espectro – Augusto Armando de Castro Júnior
- Introdução à Geometria Finsler – Umberto L. Hryniewicz e Pedro A. S. Salomão
- **Introdução aos Métodos de Crivos em Teoria dos Números – Júlio Andrade**
- Otimização de Médias sobre Grafos Orientados – Eduardo Garibaldi e João Tiago Assunção Gomes

ISBN: 978-85-244-0359-0

Distribuição: IMPA
Estrada Dona Castorina, 110
22460-320 Rio de Janeiro, RJ
E-mail: ddic@impa.br
http://www.impa.br

Conteúdo

Prefácio

A Teoria de Crivos é um conjunto de técnicas gerais em teoria dos números, dedicada a contar, ou de forma mais realista, a estimar o tamanho dos conjuntos de números inteiros que passarão por um crivo. O exemplo primordial de um conjunto que passou por um crivo é o conjunto de números primos, até um certo limite prescrito X. Do mesmo modo, o exemplo primordial de um crivo é o crivo de Eratóstenes, ou em mais geral o crivo de Legendre. O ataque direto crivando números primos usando esses métodos muito rapidamente chega a obstáculos aparentemente insuperáveis.

Uma abordagem bem sucedida é aproximar um conjunto específico de números que foi crivado (por exemplo, o conjunto de números primos) por outro conjunto mais simples (por exemplo, o conjunto dos números quase–primos), que é normalmente um pouco maior do que o conjunto original, e é mais fácil de analisar. Crivos mais sofisticados também não trabalham diretamente apenas com os conjuntos, por si só, mas também de acordo com as funções peso que são cuidadosamente escolhidas sobre esses conjuntos. Além disso, em algumas aplicações modernas, os crivos não são utilizados para estimar o tamanho de um conjunto específico, mas para produzir uma função que é grande no conjunto e pequena fora dele, sendo assim mais fácil de analisar do que a função característica do conjunto.

Crivos modernos incluem o crivo de Brun, o crivo de Selberg, o grande crivo e alguns outros que veremos nestas notas de aula. Um dos propósitos originais da teoria de crivos era tentar provar famosas e difíceis conjecturas da teoria dos números, como a conjectura dos primos gêmeos. Enquanto os grandes objetivos iniciais da teoria dos crivos ainda estão em grande parte inacabados, os crivos também

conseguiram alguns sucessos parciais, especialmente em combinação com outras ferramentas da teoria dos números.

As técnicas da teoria da crivos podem ser muito poderosas, mas elas parecem ser limitadas por um obstáculo conhecido como o problema de paridade, que a grosso modo afirma que os métodos de teoria de crivos tem extrema dificuldade em distinguir entre os números com um número ímpar de fatores primos e números com um número par de fatores primos. Este problema de paridade ainda não é muito bem compreendido.

Comparado com outros métodos na teoria dos números, a teoria dos crivos é relativamente elementar, no sentido de que tais métodos não necessitam necessariamente de conceitos sofisticados da teoria algébrica dos números ou da teoria analítica dos números. No entanto, os crivos mais avançados ainda podem ficar muito complicados e delicados (especialmente quando combinado com outras profundas técnicas em teoria dos números).

Esta notas de aulas* foram especialmente escritas para o 29º Colóquio Brasileiro de Matemática, para alunos de graduação e de pós–graduação no Brasil que irão frequentar as aulas de Teoria de Crivos. Porém é com tristeza que, apesar da grande importância da Teoria dos Números para a matemática moderna, ainda muito pouco em termos de pesquisa e ensino nesta particular área da matemática são realizados no Brasil e os alunos brasileiros interessados em Teoria dos Números ainda carecem de material e pesquisadores especialistas em Teoria dos Números. Estas notas de aulas escritas em português tem como um de seus objetivo aumentar o pequeno e restrito acervo em Teoria dos Números em língua portuguesa para alunos de universidades brasileiras.

*Estas notas de aulas são fortemente baseadas nas excelentes e modernas monografias [6, 9, 11] escritas sobre teoria dos crivos por especialistas no assunto. Estas notas também fazem parte de um livro [1], ainda em progresso, que em breve deverá ser publicado aqui no Brasil. Comentários e sugestões são bem-vindos em: j.c.andrade.math@gmail.com

Capítulo 1

Funções Aritméticas e Algumas Notações Básicas da Teoria Analítica dos Números

1.1 A notação de Bachmann–Landau

Definição 1.1.1 (notação grande 'O')**.** *Seja D um subconjunto dos números complexos $\mathbb{C}$, e $f : D \to \mathbb{C}$ uma função que assume valores complexos definida em D. Escrevemos*

$$f(x) = O(g(x))$$

se $g : D \to \mathbb{R}^+$ e se existe uma constante positiva A tal que

$$|f(x)| \leq Ag(x)$$

para todo $x \in D$.

Iremos, com alguma frequência, no decorrer destas notas usar D para denotar o conjunto dos números naturais ou $D = \mathbb{R}^+$. Uma notação alternativa para $f(x) = O(g(x))$ é

$$f(x) \ll g(x) \quad \text{ou} \quad g(x) \gg f(x).$$

Definição 1.1.2 (Ordem de Magnitude)**.** *Se $f(x) \ll g(x)$ e $g(x) \ll f(x)$, então escrevemos*

$$f(x) \asymp g(x).$$

Definição 1.1.3 (notação pequeno 'o')**.** *Se D é ilimitado, escrevemos*

$$f(x) = o(g(x))$$

se

$$\lim_{\substack{x \to \infty \\ x \in D}} \frac{f(x)}{g(x)} = 0.$$

É claro que se $f(x) = o(g(x))$, então $f(x) = O(g(x))$.

Definição 1.1.4. *Dizemos que $f(x)$ é assintótica à função $g(x)$ e escrevemos*

$$f(x) \sim g(x)$$

se

$$\lim_{\substack{x \to \infty \\ x \in D}} \frac{f(x)}{g(x)} = 1.$$

Quando escrevemos $f(x) = O(x^{\varepsilon})$, isto significa que para todo $\varepsilon > 0$ existe $C_{\varepsilon} > 0$ dependendo somente de ε tal que $|f(x)| \leq C_{\varepsilon} x^{\varepsilon}$ para todo $x \in D$. No decorrer destas notas de aula, p, q, l irão denotar números primos, n, d, k inteiros positivos, x, y, z números reais positivos. Quaisquer desvios nestes padrões estarão evidentes no contexto em que eles aparecerem. Também iremos denotar, algumas vezes, $\text{mdc}(n, d)$ como (n, d) e $\text{mmc}(n, d)$ como $[n, d]$.

1.2 A Função de Möbius

Definição 1.2.1. *A função de Möbius μ é definida da seguinte maneira:*

$$\mu(1) = 1;$$

Se $n > 1$, escrevemos $n = p_1^{a_1} \cdots p_k^{a_k}$. Então

$$\begin{aligned} \mu(n) &= (-1)^k \quad se \quad a_1 = a_2 = \cdots = a_k = 1, \\ \mu(n) &= 0 \quad \text{caso contrário.} \end{aligned}$$

Agora temos um lema básico:

Lema 1.2.2 (Propriedade Fundamental da função de Möbius)**.**

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

Demonstração. Se $n = 1$, então a fórmula é facilmente verificada. Se $n > 1$, denotemos por $n = p_1^{a_1} \cdots p_r^{a_r}$ a fatoração única de n em potências de primos distintos. Agora, seja $N = p_1 \cdots p_r$ (chamado de **radical** de n) e como $\mu(d) = 0$ a menos que d seja livre de quadrados, nós temos

$$\sum_{d|n} \mu(d) = \sum_{d|N} \mu(d).$$

A última soma contém 2^r somandos, cada um correspondendo a um subconjunto de $\{p_1, \ldots, p_r\}$, uma vez que os divisores de N estão em correspôndecia um–a–um com tais subconjuntos. O número de subconjuntos com k elementos é

$$\binom{r}{k},$$

e para cada divisor d, determinado por tal subconjunto, nós temos $\mu(d) = (-1)^k$. Assim

$$\sum_{d|N} \mu(d) = \sum_{k=0}^{r} \binom{r}{k} (-1)^k = (1-1)^r = 0.$$

Isto completa a demonstração. □

O Lema 1.2.2 é fundamental por várias razões. Primeiro, porque, em um instante, ele nos permitirá derivar uma fórmula de inversão que é útil em questões envolvendo combinatória. E em segundo, ele é a base do crivo de Eratóstenes e do crivo de Brun que iremos ver nos capítulos seguintes.

Teorema 1.2.3 (A fórmula de inversão de Möbius)**.** *Sejam f e g duas funções a valores complexos e ambas definidas nos números naturais. Se*

$$f(n) = \sum_{d|n} g(d),$$

então

$$g(n) = \sum_{d|n} \mu(d) f(n/d),$$

e reciprocamente.

Demonstração. Temos que,

$$\begin{aligned} \sum_{d|n} \mu(d) f(n/d) &= \sum_{d|n} \mu(d) \sum_{e|\frac{n}{d}} g(e) \\ &= \sum_{des=n} \mu(d) g(e) \\ &= \sum_{e|n} g(e) \sum_{d|\frac{n}{e}} \mu(d) \\ &= g(n), \end{aligned}$$

uma vez que a soma interna na pénultima linha é zero a menos que $n/e = 1$. Para estabelecermos a recíproca usamos as mesmas idéias apresentadas acima e tal exercício é deixado para o leitor. □

Teorema 1.2.4 (A fórmula dual de inversão de Möbius)**.** *Seja $\mathcal{D}$ um conjunto fechado de divisores de números naturais (isto é, se $d \in \mathcal{D}$ e $d' \mid d$, então $d' \in \mathcal{D}$). Sejam f e g duas funções que assumem valores complexos ambas definidas nos números naturais. Se*

$$f(n) = \sum_{\substack{n|d \\ d\in\mathcal{D}}} g(d),$$

então

$$g(n) = \sum_{\substack{n|d \\ d\in\mathcal{D}}} \mu\left(\frac{d}{n}\right) f(d),$$

e reciprocamente (assumindo que todas as séries convergem absolutamente).

Demonstração. Exercício. □

1.3 A Técnica de Soma Parcial.

Teorema 1.3.1. *Seja $c_1, c_2, \ldots$ uma sequência de números complexos e defina*

$$S(x) := \sum_{n\leq x} c_n.$$

Seja n_0 um inteiro positivo e fixado. Se $c_j = 0$ para $j < n_0$ e seja $f : [n_0, \infty) \to \mathbb{C}$ uma função com derivadas contínuas em $[n_0, \infty)$, então se x é um inteiro tal que $x > n_0$ nós temos que

$$\sum_{n\leq x} c_n f(n) = S(x)f(x) - \int_{n_0}^{x} S(t)f'(t)dt.$$

Demonstração.

$$\begin{aligned}\sum_{n\leq x}\{S(n)-S(n-1)\}f(n) &= \sum_{n\leq x}S(n)f(n)-\sum_{n\leq x-1}S(n)f(n+1)\\ &= S(x)f(x)-\sum_{n\leq x-1}S(n)\int_n^{n+1}f^{'}(t)dt\\ &= S(x)f(x)-\int_{n_0}^{x}S(t)f^{'}(t)dt,\end{aligned}$$

pois $S(t)$ é uma função escada que é constante em intervalos da forma $[n, n+1)$. □

Proposição 1.3.2.

$$\sum_{n\leq x}\log n = x\log x - x + O(\log x).$$

Demonstração. Usamos o teorema acima com $c_n = 1$ e $f(t) = \log t$ e deduzimos que

$$\sum_{n\leq x}\log n = [x]\log x - \int_1^x \frac{1}{x}\frac{[t]}{t}dt,$$

onde a notação $[x]$ indica o maior inteiro menor ou igual a x. E usando que $[x] = x + O(1)$ concluímos a proposição. □

Proposição 1.3.3.

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

Demonstração. Exercício. □

1.4 O Teorema de Tchebychef

Uma notação padrão em teoria dos números é que p (e algumas vezes q ou l) denotam números primos, e somas ou produtos da forma

$$\sum_{p\leq x}', \quad \sum_{p}', \quad \prod_{p\leq x}', \quad \prod_{p}'$$

indicam que as somas e os produtos são tomados sobre os números primos.

Denotemos por $\pi(x)$ o número de primos até x. Claramente, $\pi(x) = O(x)$. Em 1850, Tchebychef demonstrou, utilizando um método elementar, que

$$\pi(x) = O\left(\frac{x}{\log x}\right). \tag{1.4.1}$$

De fato, se definirmos

$$\theta(x) := \sum_{p\leq x} \log p,$$

então Tchebychef provou:

Teorema 1.4.1 (Teorema de Tchebychef)**.** *Existem constantes positivas A e B tais que*

$$Ax < \theta(x) < Bx.$$

Através do uso da técnica de soma parcial, este teorema nos dá o resultado de (1.4.1).

Podemos deduzir do Teorema 1.4.1 que sempre existe um número primo entre x e Bx/A, uma vez que

$$\theta\left(\frac{Bx}{A}\right) > A\left(\frac{Bx}{A}\right) = Bx > \theta(x).$$

Obtendo constantes A e B tal que $B/A \leq 2$, Tchebychef foi capaz de deduzir o seguinte teorema:

Teorema 1.4.2 (postualdo de Bertrand)**.** *Sempre existe um número primo entre n e $2n$, para $n \geq 1$.*

O teorema de Tchebychef representou o primeiro progresso substancial na época para uma possível solução para a famosa conjectura

de Gauss sobre o comportamento assintótico de $\pi(x)$. Com base em extensos dados numéricos e de heurísticas, Gauss previu que

$$\pi(x) \sim \frac{x}{\log x} \tag{1.4.2}$$

quando $x \to \infty$. Isto foi provado independentemente por Hadamard e de la Vallée Poussin em 1895 e é conhecido como o **teorema dos números primos**.

Demonstração de Tchebyshev do Teorema 1.4.1. O ponto chave da demonstração é notar que

$$\prod_{n<p\leq 2n} p \mid \binom{2n}{n}.$$

Uma vez que

$$\binom{2n}{n} \leq 2^{2n},$$

nós obtemos, depois de tomar logaritmos,

$$\theta(2n) - \theta(n) \leq 2n \log 2.$$

Escrevendo sucessivamente obtemos,

$$\begin{aligned} \theta(n) - \theta\left(\frac{n}{2}\right) &\leq n \log 2 \\ \theta\left(\frac{n}{2}\right) - \theta\left(\frac{n}{4}\right) &\leq \frac{n}{2} \log 2, \\ &\cdots \end{aligned}$$

e somando as desigualdades, obtemos

$$\theta(2n) \leq 4n \log 2.$$

Em outras palavras,

$$\theta(x) = O(x).$$

Logo

$$\begin{aligned} x \gg \theta(x) &\geq \sum_{\sqrt{x}<p\leq x} \log p \\ &\geq \frac{1}{2}(\log x)(\pi(x) - \pi(\sqrt{x})) \\ &\geq \frac{1}{2}(\log x)\pi(x) + O(\sqrt{x}\log x), \end{aligned}$$

e assim $\pi(x) = O(x/\log x)$. □

Usando o mesmo círculo de idéias iremos provar mais um resultado devido a Tchebychef.

Teorema 1.4.3.

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

Demonstração. Vamos estudar a fatoração em primos de $n!$. Escrevemos

$$n! = \prod_{p\leq n} p^{e_p},$$

uma vez que apenas primos $p \leq n$ podem dividir $n!$. O número de múltiplos de p que são $\leq n$ é $[n/p]$. O número de múltiplos de p^2 que são $\leq n$ é $\left[n/p^2\right]$ e assim por diante. Logo podemos ver que

$$e_p = \left[\frac{n}{p}\right] + \left[\frac{n}{p^2}\right] + \cdots,$$

onde, de fato, a soma é finita uma vez que para alguma potência p^a de p nós iremos ter $n < p^a$ e assim $[n/p^a] = 0$. Portanto deduzimos que

$$\log n! = \sum_{p\leq n} \left(\left[\frac{n}{p}\right] + \left[\frac{n}{p^2}\right] + \cdots \right) \log p.$$

E uma vez que

$$\log n! = \sum_{k\leq n} \log k = n\log n - n + O(\log n)$$

e

$$\sum_{p\le n}\left(\left[\frac{n}{p^2}\right]+\left[\frac{n}{p^3}\right]+\cdots\right)\log p \le n\sum_{p}\frac{\log p}{p(p-1)} \ll n,$$

nós temos que

$$\sum_{p\le n}\left(\left[\frac{n}{p}\right]\right)\log p = n\log n + O(n).$$

□

Teorema 1.4.4.

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

Demonstração. Colocando $c_n := (\log p)/p$ quando $n = p$ e zero caso contrário, nós aplicamos soma parcial com $f(t) := (\log t)^{-1}$ para deduzir do Teorema 1.4.3 o resultado desejado. □

1.5 Exercícios

1. Se $d(n)$ denota o número de divisores de n, mostre que $d(n) = O(\sqrt{n})$.

2. Para todo $\varepsilon > 0$, existe uma constante $C(\varepsilon)$ tal que $d(n) \le C(\varepsilon)n^{\varepsilon}$.

3. Mostre que para todo $\eta > 0$

$$d(n) < 2^{(1+\eta)\log n/\log\log n}$$

para todo n suficientemente grande.

4. Mostre que
$$\sum_{n\leq x} d(n) = x\log x + O(x).$$

5. Seja $\phi(n)$ definida como sendo a quantidade de números inteiros $1 \leq k \leq n$ tal que $(n,k) = 1$. Prove que
$$\sum_{n\leq x} \frac{\phi(n)}{n} = \frac{6}{\pi^2}x + O(\log x).$$

6. Prove que
$$\sum_{d|n} \Lambda(d) = \log n.$$
Deduza que
$$\Lambda(n) = -\sum_{d|n} \mu(d)\log d.$$
Onde
$$\Lambda(n) := \begin{cases} \log p & \text{se } n = p^a, \\ 0 & \text{caso contrário.} \end{cases}$$
$\Lambda(n)$ é chamada de função de **von Mangoldt**.

Capítulo 2

Alguns Crivos Elementares

2.1 Generalidades

Seja $\mathcal{A}$ um conjunto finito de objetos e $\mathcal{P}$ um conjunto indexado de números primos tal que para cada $p \in \mathcal{P}$ nós temos associado um subconjunto $\mathcal{A}_p$ de $\mathcal{A}$. O **problema de crivo** é estimar, superiormente e inferiormente, o tamanho do conjunto

$$\mathcal{S}(\mathcal{A}, \mathcal{P}) := \mathcal{A} \setminus \cup_{p \in \mathcal{P}} \mathcal{A}_p.$$

Esta é a formulação do problema no contexto mais geral possível. É claro, a resposta 'explícita' é dada pelo familiar princípio da combinatória de inclusão–exclusão. Mais precisamente, para cada subconjunto I de $\mathcal{P}$, denotamos por

$$\mathcal{A}_I := \cap_{p \in I} \mathcal{A}_p.$$

Então o principio de inclusão–exclusão nos fornece

$$\#\mathcal{S}(\mathcal{A}, \mathcal{P}) = \sum_{I \subseteq \mathcal{P}} (-1)^{\#I} \#\mathcal{A}_I,$$

onde para o conjunto vazio $\emptyset$ nós interpretamos $\mathcal{A}_\emptyset$ como sendo o próprio conjunto $\mathcal{A}$. Esta fórmula é a base para muitas questões em teoria de probabilidades.

Em teoria dos números nós frequentemente tomamos $\mathcal{A}$ para ser um conjunto finito de inteiros positivos e $\mathcal{A}_p$ como sendo o subconjunto de $\mathcal{A}$ consistindo de elementos que estão em certas classes de congruência módulo p. Por exemplo, se $\mathcal{A}$ é o conjunto dos números naturais $\leq x$ e $\mathcal{A}_p$ é o conjunto dos números em $\mathcal{A}$ divisíveis por p, então o tamanho de $\mathcal{S}(\mathcal{A}, \mathcal{P})$ irá ser o número de inteiros positivos $n \leq x$ que são coprimos com todos os elementos de $\mathcal{P}$.

Nós podemos também reverter esta perspectiva. Ou seja, podemos pensar de $\mathcal{S} = \mathcal{S}(\mathcal{A}, \mathcal{P})$ como sendo um conjunto dado, cujo tamanho nós queremos estimar. Procuramos fazer isso olhando a sua imagem módulo primos $p \in \mathcal{P}$ para algum conjunto de primos $\mathcal{P}$. Este será o ponto de vista para o grande crivo (Capítulo 8).

Podemos até mesmo aumentar esta perspectiva reversa da seguinte maneira: Seja $\mathcal{B}$ um conjunto finito de inteiros positivos e seja $\mathcal{T}$ um conjunto de potências de primos. Nós então podemos procurar estimar o tamanho do próprio conjunto $\mathcal{B}$. Este tratamento será discutido abaixo e é o Crivo de Gallagher.

Em algumas circunstâncias fortuitas, nós podemos ter uma família $\mathcal{F}$ de funções a valores complexos

$$f : \mathcal{B} \to \mathbb{C}$$

tal que

$$\sum_{f \in \mathcal{F}} f(n) = \begin{cases} 1 & \text{se } n \in \mathcal{B}, \\ 0 & \text{caso contrário.} \end{cases}$$

Então

$$\#\mathcal{B} = \sum_{f \in \mathcal{F}} \left(\sum_{n \in \mathcal{B}} f(n) \right) \tag{2.1.1}$$

e a soma interna pode ser investigada por outras técnicas, tais como métodos analíticos. Iremos ilustrar esta idéia na seção sobre crivos usando séries de Dirichlet.

Muitas vezes, é o caso que uma relação precisa como (2.1.1) não existe e nós só podemos ter uma aproximação para ela. Por exemplo,

$$\sum_{f\in\mathcal{F}} f(n) = \begin{cases} 1+o(1) & \text{se } n\in\mathcal{B} \\ o(1) & \text{caso contrário,} \end{cases} \tag{2.1.2}$$

quando $\#\mathcal{F}\to\infty$. Tal família de funções pode então ser usada, com grande efeito, para estimativas para $\#\mathcal{B}$.

Até mesmo o conhecimento de

$$\sum_{f\in\mathcal{F}} f(n) \geq 1 \quad \text{se} \quad n\in\mathcal{B}$$

é suficiente para nos fornecer boas cotas superiores para $\#\mathcal{B}$. De fato, nós podemos considerar

$$\#\mathcal{B} \leq \sum_{n\in\mathcal{B}} \left| \sum_{f\in\mathcal{F}} f(n) \right|^2$$

e expandir o quadrado. Depois de inverter as somas, nos deparamos

$$\sum_{n\in\mathcal{B}} f(n)\overline{f'(n)}$$

para $f, f' \in \mathcal{F}$.

Nós iremos ilustrar algumas destas idéias, primeiro com o crivo de Gallagher [10] e então com o crivo para quadrados perfeitos [12]. E na última seção iremos discutir técnicas analíticas para investigarmos crivos usando séries de Dirichlet.

2.2 O Crivo de Gallagher ou o Maior Crivo

Seja $\mathcal{B}$ um conjunto (não–vazio) finito de números inteiros e $\mathcal{T}$ um conjunto de potências de primos. Suponha que para cada $t\in\mathcal{T}$ nós temos

$$\#\mathcal{B} \pmod t \leq u(t)$$

para alguma $u(t)$. Assim $\mathcal{B}$ representa no máximo $u(t)$ classes de resíduo módulo t.

Teorema 2.2.1 (Maior Crivo de Gallagher)**.** *Nós mantemos a notação acima e seja*

$$X := \max_{b\in\mathcal{B}} |b|.$$

Se

$$\sum_{t\in\mathcal{T}} \frac{\Lambda(t)}{u(t)} - \log(2X) > 0,$$

então

$$\#\mathcal{B} \leq \frac{\sum_{t\in\mathcal{T}} \Lambda(t) - \log(2X)}{\sum_{t\in\mathcal{T}} \frac{\Lambda(t)}{u(t)} - \log(2X)},$$

onde $\Lambda(\cdot)$ *é função de von Mangoldt.*

Demonstração. Seja $t \in \mathcal{T}$ e para cada classe de resíduos $r(\mathrm{mod}\ t)$ definimos

$$Z(\mathcal{B}; t, r) := \#\{b \in \mathcal{B} : b \equiv r(\mathrm{mod}\ t)\}.$$

Então

$$\#\mathcal{B} = \sum_{r(\mathrm{\ mod\ }t)} Z(\mathcal{B}; r, t).$$

Usando a desigualdade de Cauchy–Schwarz, temos que

$$\leq u(t)^{1/2} \left(\sum_{r(\mathrm{\ mod\ }t)} Z(\mathcal{B}; r, t)^2 \right)^{1/2}.$$

Logo

$$\begin{aligned} \frac{(\#\mathcal{B})^2}{u(t)} &\leq \sum_{r(\mathrm{\ mod\ }t)} \sum_{\substack{b,b'\in\mathcal{B} \\ b,b' \equiv r(\mathrm{\ mod\ }t)}} 1 \\ &\leq \#\mathcal{B} + \sum_{\substack{b,b'\in\mathcal{B} \\ b\neq b'}} \sum_{t\mid b-b'} 1. \end{aligned}$$

Nós multiplicamos esta desigualdade por $\Lambda(t)$ e somamos sobre $t \in \mathcal{T}$. Usando

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

nós obtemos

$$\sum_{t\in\mathcal{T}} \frac{(\#\mathcal{B})^2}{u(t)}\Lambda(t) \leq (\#\mathcal{B}) \sum_{t\in\mathcal{T}} \Lambda(t) + (\log 2X)((\#\mathcal{B})^2 - \#\mathcal{B}).$$

Cancelando $\#\mathcal{B}$ e reorganizando, nós estabelecemos a desigualdade desejada. □

Este crivo deve ser comparado com o Grande Crivo discutido no Capítulo 8. A vantagem aqui é que nós podemos crivar (ou peneirar) classes de resíduo módulo potências de primos, onde no grande crivo somente classes de resíduo módulo primos são considerados. Isto explica, até certo ponto, o nome 'o maior crivo'.

Seguindo Gallagher, nós aplicamos o maior crivo para provar:

Teorema 2.2.2. *Sejam a, b inteiros tendo a propriedade que para qualquer potência de primo t existe um inteiro ν_t tal que*

$$b \equiv a^{\nu_t} (\text{mod } t).$$

Então existe um inteiro ν tal que

$$b = a^{\nu}.$$

Antes de procedermos com a demonstração deste teorema, vamos revisar algumas propriedades básicas de polinômios ciclotômicos. Lembre-se que para um inteiro positivo d, o d–ésimo polinômio ciclotômico $\Phi_d(x)$ é o polinômio minimal sobre $\mathbb{Q}$ da d-ésima raiz primitiva da unidade. Assim ele tem grau $\phi(d)$, onde $\phi(\cdot)$ é a função de Euler. Agora seja n um inteiro positivo arbitrário. Como as n-ésimas raízes da unidade podem ser particionadas de acordo com a sua ordem, nós temos a fórmula

$$x^n - 1 = \prod_{d|n} \Phi_d(x).$$

Finalmente, para um inteiro a tal que $(a, n) = 1$ definimos $f_a(n)$ a ser a ordem de a módulo n. Pelo Exercício 1 nós temos que $f_a(n) = d$ se e somente se $n \mid \Phi_d(a)$.

Demonstração do Teorema 2.2.2. Sejam a, b como no enunciado do teorema. Nós observamos que para provar o resultado nós podemos supor que a, b são positivos e $a \geq 3$ (Exercício 2). Sejam

$$\mathcal{B} := \{n \leq x : n = a^i b^j \text{ para algum } i, j\}$$

e

$$\mathcal{T} := \{t : t \text{ potência de primo, } f_a(t) \leq y\},$$

onde $y = y(x)$ é algum parâmetro a ser escolhido mais tarde. Pelo Exercício 1, $\mathcal{T}$ é um conjunto finito.

Nós mantemos a notação do Teorema 2.2.1. Se para cada potência de primo t nós temos que b é uma potência de a módulo t, então

$$u(t) \leq f_a(t).$$

Assim Teorema 2.2.1 implica que

$$\#\mathcal{B} \leq \frac{\sum_{t\in\mathcal{T}} \Lambda(t) - \log(2x)}{\sum_{t\in\mathcal{T}} \frac{\Lambda(t)}{f_a(t)} - \log(2x)}, \tag{2.2.1}$$

desde que o denominador seja positivo.

Nós temos

$$\begin{aligned}
\sum_{t\in\mathcal{T}} \Lambda(t) &= \sum_{d\leq y} \sum_{f_a(t)=d} \Lambda(t) \\
&= \sum_{d\leq y} \sum_{t|\Phi_d(a)} \Lambda(t) \\
&= \sum_{d\leq y} \log \Phi_d(a),
\end{aligned}$$

pelo Exercício 1. Claramente,

$$(a-1)^{\phi(d)} \leq |\Phi_d(a)| \leq (a+1)^{\phi(d)},$$

e assim

$$\sum_{t\in\mathcal{T}} \Lambda(t) = \sum_{d\leq y} \log|\Phi_d(a)| \asymp \sum_{d\leq y} \phi(d) \asymp y^2.$$

Nós também notamos que isto implica

$$\sum_{t\in\mathcal{T}} \frac{\Lambda(t)}{f_a(t)} \geq \frac{1}{y} \sum_{t\in\mathcal{T}} \Lambda(t) \gg y.$$

Agora escolhemos

$$y := 100\log(2x).$$

De (2.2.1) deduzimos que

$$\#\mathcal{B} \ll \log x. \tag{2.2.2}$$

Para este fim, observemos que se todas as potências de a e b forem distintas, então o conjunto $\mathcal{B}$ tem cardinalidade

$$\asymp (\log x)^2$$

(veja Exercício 3). Isto contradiz (2.2.2), e assim concluímos que para algum i_0, j_0 nós temos

$$a^{i_0} = b^{j_0}.$$

Nós podemos até mesmo supor que $(i_0, j_0) = 1$, pois caso contrário nós podemos tomar a (i_0, j_0)–ésima raiz de ambos os lados da igualdade acima.

Denotemos

$$n = \prod_p p^{\nu_p(n)}$$

para a fatoração única de um inteiro n em potências de primos. Nós deduzimos que

$$i_0\nu_p(a) = j_0\nu_p(b)$$

para todos os primos p. Como $(i_0, j_0) = 1$, isto significa que $i_0 \mid \nu_p(b)$ e $j_0 \mid \nu_p(a)$ para todos os primos p. Isto implica que a é uma j_0-ésima potência e b é uma i_0–ésima potência de algum inteiro c. As hipóteses agora implicam que para qualquer primo q existe um ν_q tal que

$$c^{j_0\nu_q} \equiv c^{i_0} (\text{mod } q),$$

que é equivalente a $f_c(q) \mid (j_0\nu_q - i_0)$ se $(q, c) = 1$.

Agora tomamos um divisor primo q de $\Phi_{j_0 t}(c)$ para todo t. Pelo Exercício 1 nós deduzimos que $f_c(q) \equiv 0 (\text{mod } j_0)$. Assim $j_0 \mid i_0$, e portanto b é uma potência de a, como desejado. □

2.3 O Crivo para Quadrados Perfeitos

O crivo para quadrados perfeitos têm a sua origem com o trabalho de Heath–Brown [12] e é usado para estimar o número de quadrados perfeitos em um dado conjunto de números inteiros. O crivo se baseia no uso de uma família de símbolos de resíduos quadráticos para crivar os quadrados perfeitos. Consequentemente, ele também é adequado para o estudo de sequências que são uniformemente distribuídas em progressões aritméticas.

Teorema 2.3.1 (O Crivo para Quadrados)**.** *Seja $\mathcal{A}$ um conjunto finito de inteiros positivos não-nulos e $\mathcal{P}$ um conjunto de primos ímpares. Definimos*

$$S(\mathcal{A}) := \#\{\alpha \in \mathcal{A} : \alpha \text{ é um quadrado}\}.$$

Então

$$S(\mathcal{A}) \leq \frac{\#\mathcal{A}}{\#\mathcal{P}} + \max_{\substack{q_1 \neq q_2 \\ q_1, q_2 \in \mathcal{P}}} \left| \sum_{a \in \mathcal{A}} \left(\frac{\alpha}{q_1 q_2} \right) \right| + E$$

onde $\left(\frac{\cdot}{q_1 q_2} \right)$ *denota o símbolo de Jacobi e*

$$E := O\left(\frac{1}{\#\mathcal{P}}\sum_{\alpha\in\mathcal{A}}\nu_{\mathcal{P}}(\alpha)+\frac{1}{(\#\mathcal{P})^2}\sum_{\alpha\in\mathcal{A}}\nu_{\mathcal{P}}(\alpha)^2\right),$$

$$\nu_{\mathcal{P}}(\alpha) := \sum_{\substack{p\in\mathcal{P}\\ p|\alpha}} 1.$$

Observação 2.3.2. *Na prática, a contribuição de E é insignificante e esperamos que as maiores contribuições para a estimativa venham dos dois primeiros termos.*

Demonstração. Começamos por observar que se $\alpha \in \mathcal{A}$ é um quadrado, então

$$\sum_{q\in\mathcal{P}}\left(\frac{\alpha}{q}\right) = \#\mathcal{P} - \nu_{\mathcal{P}}(\alpha).$$

Assim

$$S(\mathcal{A}) \le \sum_{\alpha\in\mathcal{A}}\frac{1}{(\#\mathcal{P})^2}\left(\sum_{q\in\mathcal{P}}\left(\frac{\alpha}{q}\right)+\nu_{\mathcal{P}}(\alpha)\right)^2. \tag{2.3.1}$$

Após elevarmos ao quadrado e invertermos as somas, nós obtemos que o lado direito da desigualdade (2.3.1) é

$$\sum_{\alpha\in\mathcal{A}}\frac{1}{(\#\mathcal{P})^2}\left(\sum_{q_1,q_2\in\mathcal{P}}\left(\frac{\alpha}{q_1}\right)\left(\frac{\alpha}{q_2}\right)+2\nu_{\mathcal{P}}(\alpha)\sum_{q\in\mathcal{P}}\left(\frac{\alpha}{q}\right)+\nu_{\mathcal{P}}(\alpha)^2\right).$$

A primeira soma é

$$\begin{aligned}\sum_{q_1,q_2\in\mathcal{P}}\frac{1}{(\#\mathcal{P})^2}\sum_{\alpha\in\mathcal{A}}\left(\frac{\alpha}{q_1}\right)\left(\frac{\alpha}{q_2}\right) &\le \frac{\#\mathcal{A}}{\#\mathcal{P}}+\sum_{\substack{q_1,q_2\in\mathcal{P}\\ q_1\neq q_2}}\frac{1}{(\#\mathcal{P})^2}\sum_{\alpha\in\mathcal{A}}\left(\frac{\alpha}{q_1q_2}\right)\\ &\le \frac{\#\mathcal{A}}{\#\mathcal{P}}\max_{\substack{q_1,q_2\in\mathcal{P}\\ q_1\neq q_2}}\left|\sum_{\alpha\in\mathcal{A}}\left(\frac{\alpha}{q_1q_2}\right)\right|.\end{aligned}$$

É fácil de ver que a contribuição para (2.3.1) das últimas somas é

$$E \leq \frac{2}{\#\mathcal{P}} \sum_{\alpha \in \mathcal{A}} \nu_{\mathcal{P}}(\alpha) + \frac{1}{(\#\mathcal{P})^2} \sum_{\alpha \in \mathcal{A}} \nu_{\mathcal{P}}(\alpha)^2.$$

Isto completa a demonstração. □

Corolário 2.3.3. *Seja $\mathcal{A}$ um conjunto de números inteiros não–nulos e $\mathcal{P}$ um conjunto de números primos que são coprimos com os elementos de $\mathcal{A}$. Então*

$$\mathcal{S}(\mathcal{A}) = \#\{\alpha \in \mathcal{A} : \alpha \text{ é um quadrado}\} \leq \frac{\#\mathcal{A}}{\#\mathcal{P}} + \max_{\substack{q_q, q_2 \in \mathcal{P} \\ q_1 \neq q_2}} \left| \sum_{\alpha \in \mathcal{A}} \left(\frac{\alpha}{q_1 q_2} \right) \right|.$$

Demonstração. Exercício. □

Nós iremos aplicar o crivo para quadrados para contar o número de pontos inteiros em uma curva hiper-elíptica

$$y^2 = f(x),$$

onde $f(x) \in \mathbb{Z}[x]$ é um polinômio de grau d, com discriminante não–nulo, e o qual não é o quadrado perfeito de um polinômio com coeficientes inteiros. Um famoso teorema de Siegel [20] nos diz que o número de pontos inteiros de tal curva é finito. Recentemente, estimativas efetivas deste número foram dadas por diferentes autores (veja [13]). Entretanto, estas estimativas envolvem o conhecimento do grau de Mordell-Weil do Jacobiano da curva hiper–elíptica. Nosso tratamento é elementar e pode ser adaptado para estudar o quão frequentemente um polinômio $f(x_1, \ldots, x_n)$ representa um quadrado. Como irá ser visto abaixo, a generalização irá requerer o profundo trabalho de Deligne (veja [19]).

Dado $f(x) \in \mathbb{Z}[x]$ e $k \in \mathbb{N}$, e também seja $k > 2$,

$$S_f(k) := \sum_{a \pmod k} \left(\frac{f(a)}{k} \right),$$

onde $(\cdot/k)$ é o símbolo de Jacobi.

Lema 2.3.4. *Sejam q_1, q_2 números primos distintos e seja $f \in \mathbb{Z}[x]$. Então*

$$S_f(q_1q_2) = S_{f_1}(q_1)S_{f_2}(q_2),$$

onde $f_1(x) := f(q_2x), f_2(x) := f(q_1x)$.

Demonstração. As classes de resíduos módulo q_1q_2 podem ser escritas como

$$q_1a_2 + q_2a_1,$$

com $0 \leq a_2 \leq q_2 - 1, 0 \leq a_1 \leq q_1 - 1$ (veja Exercício 4). Portanto

$$\begin{aligned} S_f(q_1q_2) &= \sum_{a_2=0}^{q_1-1}\sum_{a_1=0}^{q_2-1}\left(\frac{f(q_1a_2+q_2a_1)}{q_1q_2}\right) \\ &= \sum_{a_2=0}^{q_1-1}\sum_{a_1=0}^{q_2-1}\left(\frac{f(q_1a_2+q_2a_1)}{q_1}\right)\left(\frac{f(q_1a_2+q_2a_1)}{q_2}\right) \\ &= \sum_{a_2=0}^{q_1-1}\sum_{a_1=0}^{q_2-1}\left(\frac{f(q_2a_1)}{q_1}\right)\left(\frac{f(q_1a_2)}{q_2}\right). \end{aligned}$$

E assim o resultado segue. □

Agora seja H um número real positivo e consideremos o conjunto

$$\mathcal{A} := \{f(n) : |n| \leq H\}.$$

Usando o crivo para quadrados, o número de quadrados em $\mathcal{A}$ é, para qualquer conjunto de números primos $\mathcal{P}$ os quais não dividem o discriminante de f,

$$\leq \frac{2H+1}{\#\mathcal{P}} + \max_{\substack{q_1,q_2\in\mathcal{P}\\ q_1\neq q_2}}\left|\sum_{|n|\leq H}\left(\frac{f(n)}{q_1q_2}\right)\right| + E,$$

onde

$$E := O\left(\frac{H\log H}{\#\mathcal{P}} + \frac{H(\log H)^2}{(\#\mathcal{P})^2}\right)$$

e onde usamos a estimativa elementar $\nu_{\mathcal{P}}(\alpha) = O(\log \alpha)$.

Sejam q_1, q_2 dois primos distintos de $\mathcal{P}$. Nós temos

$$\sum_{|n|\leq H} \left(\frac{f(n)}{q_1 q_2}\right) = \sum_{a(\bmod q_1 q_2)} \left(\frac{f(a)}{q_1 q_2}\right) \sum_{\substack{|n|\leq H \\ n\equiv a(\bmod q_1 q_2)}} 1.$$

A soma interna é

$$\frac{2H}{q_1 q_2} + O(1),$$

e assim obtemos

$$\sum_{|n|\leq H} \left(\frac{f(n)}{q_1 q_2}\right) = \frac{2H}{q_1 q_2} \sum_{a(\bmod q_1 q_2)} \left(\frac{f(a)}{q_1 q_2}\right) + O(q_1 q_2).$$

Pelo lema, a soma do lado direito é o produto $S_{f_1}(q_1) S_{f_2}(q_2)$ para apropriados polinômios f_1, f_2.

Nós invocamos um célebre resultado de Weil (veja [15, p. 99]), que diz que para qualquer $g(x) \in \mathbb{Z}[x]$ com discriminante não–nulo e o qual não é quadrado perfeito de um polinômio com coeficientes inteiros, e para qualquer primo p não dividindo o discriminante de g.

$$\left| \sum_{a(\bmod p)} \left(\frac{g(a)}{p}\right) \right| \leq (\deg g - 1)\sqrt{p}.$$

Usando isto na estimativa acima nós obtemos

$$\sum_{|n|\leq H} \left(\frac{f(n)}{q_1 q_2}\right) = O\left(\frac{H}{\sqrt{q_1 q_2}} + q_1 q_2\right).$$

Vamos escolher o conjunto $\mathcal{P}$ consistindo dos primos que não dividem o discriminante de f e estão no intervalo $[z, 2z]$ para algum $z = z(H) > 0$ a ser escolhido em breve. Assim temos a estimativa final,

$$S(\mathcal{A}) \ll \frac{H \log z}{z} + \frac{H}{z} + z^2 + \frac{H(\log H)(\log z)}{z} + \frac{H(\log H)^2 (\log z)^2}{z^2}.$$

Escolhendo

$$z := H^{1/3}(\log H)^{2/3}$$

provamos o seguinte teorema:

Teorema 2.3.5. *Seja f um polinômio com discriminante não–nulo e coeficientes inteiros, o qual não é um quadrado perfeito de um polinômio com coeficientes inteiros. Seja $H > 0$. Então o número de quadrados no conjunto*

$$\{f(n) : |n| \leq H\}$$

é $O(H^{2/3}(\log H)^{4/3})$, com a constante aparecendo na estimativa para–O dependendo somente do grau de f e dos coeficientes de f.

2.4 O Crivo usando séries de Dirichlet

Algumas vezes, a sequência de números que queremos crivar exibe uma estrutura multiplicativa e as condições de crivo também exibem tal propriedade. Em tais casos, métodos analíticos usando séries de Dirichlet são bastante poderosos e diretos. Em algumas instâncias, a técnica pode até mesmo nos fornecer fórmulas assintóticas para o problema de crivos. Nós iremos ilustrar esta idéia abaixo, para mais detalhes consulte [18].

Seja $\mathcal{P}$ um conjunto de primos e $\overline{\mathcal{P}}$ o complemento no conjunto de todos os primos. Suponha que queremos contar a quantidade de números naturais $n \leq x$ que não são divisíveis por nenhum dos primos em $\mathcal{P}$. Se definirmos a série de Dirichlet

$$F(s) = \sum_{n \geq 1} \frac{a_n}{n^s} := \prod_{p \in \overline{\mathcal{P}}} \left(1 - \frac{1}{p^s}\right)^{-1},$$

nós vemos que $a_n = 1$ se n não for divisível para todo $p \in \mathcal{P}$ e $a_n = 0$ caso contrário. Assim nós queremos estudar

$$\sum_{n \leq x} a_n.$$

Pela fórmula de Perron (veja [2, pg. 245–246]), temos que

$$\sum_{n\leq x} a_n = \frac{1}{2\pi i}\int_{2-i\infty}^{2+i\infty} F(s)\frac{x^s}{s}ds.$$

O teorema abaixo é útil neste contexto.

Teorema 2.4.1 (Teorema Tauberiano)**.** *Seja* $F(s) = \sum_{n\geq 1} a_n/n^s$ *uma série de Dirichlet com coeficientes não–negativos convergindo para* $Re(s) > 1$*. Suponha que* $F(s)$ *se estende analiticamente em todos os pontos em* $\mathrm{Re}(s) = 1$ *exceto em* $s = 1$*, e que em* $s = 1$ *nós podemos escrever*

$$F(s) = \frac{H(s)}{(s-1)^{1-\alpha}}$$

para algum $\alpha \in \mathbb{R}$ *e alguma* $H(s)$ *holomórfica na região* $\mathrm{Re}(s) \geq 1$ *e não–nula lá. Então*

$$\sum_{n\leq x} a_n \sim \frac{cx}{(\log x)^{\alpha}}$$

com

$$c := \frac{H(1)}{\Gamma(1-\alpha)},$$

onde Γ *é a usual função gama.*

Nós não iremos provar este teorema aqui. Vamos agora ilustrar estas técnicas em um exemplo mais concreto. Consideremos o problema de contar o número de números naturais $n \leq x$ que podem ser escritos como a soma de dois quadrados. É bem conhecido (veja [14, p. 279]) que n pode ser escrito como uma soma de dois quadrados se e somente se para todo primo $p \equiv 3(\mathrm{mod}4)$ dividindo n, a potência de p aparecendo na fatoração única de n é par. Assim, $a_n = 1$ sempre que n puder ser escrito como uma soma de dois quadrados e é 0 caso contrário, nós então vemos que

$$F(s) := \sum_{n\geq 1} \frac{a_n}{n^s}$$
$$= \left(1-\frac{1}{2^s}\right)^{-1} \prod_{p\equiv 1(\bmod 4)} \left(1-\frac{1}{p^s}\right)^{-1} \prod_{p\equiv 3(\bmod 4)} \left(1-\frac{1}{p^{2s}}\right)^{-1}.$$

Agora nós precisamos invocar algumas propriedades básicas da função zeta de Riemann $\zeta(s)$ e de funções L de Dirichlet $L(s,\chi_4)$ associado com o caractere quadrático χ_4, definidas por

$$\zeta(s) := \sum_{n\geq 1} \frac{1}{n^s}, \qquad L(s,\chi_4) := \sum_{n\geq 1} \frac{\chi_4(n)}{n^s}$$

para $s \in \mathbb{C}$ com $\mathrm{Re}(s) > 1$. Aqui, $\chi_4(n)$ é 0 se n é par e $(-1)^{(n-1)/2}$ se n ímpar. Nós indicamos o leitor para [2] para outras propriedades destas funções. Usando o produto de Euler de $\zeta(s)$ e $L(s,\chi_4)$ nós escrevemos

$$F(s) = [\zeta(s)L(s,\chi_4)]^{1/2} H_1(s),$$

onde $H_1(s)$ é analítica e diferente de zero para $\mathrm{Re}(s) > 1/2$. Como $L(s,\chi_4)$ se estende para uma função inteira e é não–nula para $\mathrm{Re}(s) \geq 1$, nós temos

$$F(s) = \zeta(s)^{1/2} H_2(s)$$

para alguma função $H_2(s)$ holomórfica e não–nula em $\mathrm{Re}(s) \geq 1$. Assim, usando o fato que a função zeta de Riemann tem um pólo simples em $s = 1$ e que é analítica e não–nula para $\mathrm{Re}(s) = 1$, $s \neq 1$ (veja [2]), nós deduzimos que

$$F(s) = \frac{H(s)}{(s-1)^{1/2}},$$

com $H(s)$ holomórfica e não–nula na região $\mathrm{Re}(s) \geq 1$. Pelo teorema Tauberiano citado acima nós obtemos:

Teorema 2.4.2. *O número de $n \leq x$ que podem ser escritos como a soma de dois quadrados é*

$$\sim \frac{cx}{\sqrt{\log x}}$$

para algum $c > 0$*, quando* $x \to \infty$*.*

2.5 Exercícios

1. Se t é uma potência de um primo coprimo com k, mostre que $t \mid \Phi_k(a)$ se e somente se $a (\mathrm{mod}\ t)$ têm ordem k.

2. Sejam a, b inteiros. Dizemos que a e b estão relacionados se, para toda potência de primos t,

$$b \equiv a^{\nu_t} (\mathrm{mod}\ t)$$

para algum inteiro positivo ν_t. Mostre que se a e b estão relacionados, então a^2 está relacionado com b^2. Também, mostre que se $|a| \leq 2$ e $|b| \leq 2$ com a e b relacionados, então $|a| = |b|$.

3. Se a e b são números naturais ≥ 2 com $\{a^i : i \geq 1\} \cap \{b^j : j \geq 1\} = \emptyset$, mostre que o número $a^i b^j \leq x$ é

$$\asymp (\log x)^2.$$

4. Sejam t_1, t_2 inteiros positivos coprimos e $t := t_1 t_2$. Mostre que todas as classes de resíduo módulo t podem ser representadas como

$$t_1 a_2 + t_2 a_q$$

para algum $0 \leq a_2 \leq t_2 - 1$, $0 \leq a_1 \leq t_1 - 1$. Também, mostre que as classes de resíduo primas entre si módulo t podem ser representadas como acima com $(a_2, t_2) = (a_1, t_1) = 1$.

5. Um número n é chamado *quadrado completo* se se para todo primo $p \mid n$ então $p^2 \mid n$. Mostre que a quantidade de números que são quadrados completos $\leq x$ é

$$\sim c_1\sqrt{x}$$

para algum $c_1 > 0$, quando $x \to \infty$.

Capítulo 3

Um Teorema de Hardy e Ramanujam: O Método da Ordem Normal

O método da ordem normal têm a sua origem em um artigo de 1916 por G.H. Hardy e S. Ramanujan. Um tratamento mais simples e transparente deste trabalho foi dado por Paul Turán em 1934. Finalmente o método de Turán foi amplificado por P. Erdös e Mark Kac para criar o que é hoje chamado de teoria probabilística dos números. Neste capítulo estudamos como o método de Turán pode ser usado para estudar a distribuição da quantidade de fatores primos de diferentes sequências de números.

3.1 Um Teorema de Hardy e Ramanujan

Em 1916 Hardy e Ramanujan provaram que quase todos os números n são compostos de $\log \log n$ fatores primos. Para ser mais preciso eles mostraram que o número de $n \leq x$ que não satisfazem a desigualdade

$$(1-\varepsilon)\log\log n < \nu(n) < (1+\varepsilon)\log\log n$$

é $o(x)$ para todo $\varepsilon > 0$ onde $\nu(n)$ denota o número dos distintos divisores primos de n. Nós apresentamos a demonstração de Turán de tal fato.

Teorema 3.1.1. *Nós temos que*

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

e

$$\sum_{n\leq x}\nu^2(n) = x(\log\log x)^2 + O(x\log\log x).$$

Demonstração. Observemos que pelo Teorema 1.4.4

$$\begin{aligned}\sum_{n\leq x}\nu(n) &= \sum_{p\leq x}\left[\frac{x}{p}\right]\\ &= x\sum_{p\leq x}\frac{1}{p} + O(x)\\ &= x\log\log x + O(x).\end{aligned}$$

Também,

$$\begin{aligned}\sum_{n\leq x}\nu(n)^2 &= \sum_{n\leq x}\sum_{p|n}\sum_{q|n}1\\ &= \sum_{p,q\leq x}\sum_{\substack{n\leq x\\ p|n,q|n}}1\\ &= \sum_{\substack{p,q\leq x\\ p\neq q}}\left[\frac{x}{pq}\right]+\sum_{p\leq x}\left[\frac{x}{p}\right]\\ &= \sum_{pq\leq x}\left[\frac{x}{pq}\right]+O(x\log\log x)\\ &= x\sum_{pq\leq x}\frac{1}{pq}+O(x\log\log x).\end{aligned}$$

Agora,

$$\left(\sum_{p\leq\sqrt{x}}\frac{1}{p}\right)^2\leq\sum_{pq\leq x}\frac{1}{pq}\leq\left(\sum_{p\leq x}\frac{1}{p}\right)^2.$$

Uma vez que

$$\sum_{p\leq\sqrt{x}}\frac{1}{p}=\log\log\sqrt{x}+O(1)=\log\log x+O(1),$$

nós obtemos que

$$\sum_{n\leq x}\nu^2(n)=x(\log\log x)^2+O(x\log\log x).$$

Isto completa a demonstração. □

Teorema 3.1.2 (Turán)**.**

$$\sum_{n\leq x}(\nu(n)-\log\log x)^2=O(x\log\log x).$$

Demonstração. Consideremos a variância

$$\sum_{n\leq x}(\nu(n)-\log\log x)^2$$
$$=\sum_{n\leq x}\nu^2(n)-2(\log\log x)\sum_{n\leq x}\nu(n)+(\log\log x)^2\sum_{n\leq x}1,$$

e é fácil de ver que isto é $O(x\log\log x)$ pelo o que provamos acima. □

Corolário 3.1.3. *Seja $\delta>0$. O número de $n\leq x$ que não satisfazem a desigualdade*

$$|\nu(n)-\log\log x|<(\log\log x)^{\frac{1}{2}+\delta}$$

é $o(x)$.

Demonstração. Exercício. □

Relacionado aos resultados acima existe um célebre teorema de Erdös e Kac (veja [8]), que diz que, se para $\alpha\leq\beta$,

$$S(x;\alpha,\beta):=\#\left\{n\leq x:\alpha\leq\frac{\nu(n)-\log\log n}{\sqrt{\log\log n}}\leq\beta\right\},$$

então

$$\lim_{x\to\infty}\frac{S(x;\alpha,\beta)}{x}=\frac{1}{\sqrt{2\pi}}\int_{\alpha}^{\beta}e^{-t^2/2}dt.$$

A integral é a familiar integral de probabilidade associadada com a distribuição normal. Assim, o teorema de Erdös–Kac diz que a função

$$\frac{\nu(n)-\log\log n}{\sqrt{\log\log n}}$$

é 'normalmente distribuída' num certo sentido probabilístico. Recomendamos o leitor a consultar [8] para mais detalhes.

Nós dizemos que uma função $f(n)$ possui **ordem normal** $F(n)$ se, para todo $\varepsilon>0$, a desigualdade

$$(1-\varepsilon)F(n) < f(n) < (1+\varepsilon)F(n)$$

é satisfeita para quase todos os valores de n. Isto é, o número de $n \leq x$ que não satisfazem a desigualdade é $o(x)$.

3.2 O número normal de divisores primos de um polinômio

Seja $f(x)$ um polinômio irredutível com coeficientes inteiros. Nós iremos considerar aqui o problema de determinar a ordem normal de $\nu(f(n))$.

Primeiro, nós observamos que se $\nu_y(n)$ denota o número de primos dividindo n que são $\leq y$ e se $y = x^\delta$ para algum $0 < \delta < 1/2$, então para $n \leq x$ nós temos

$$\nu(n) = \nu_y(n) + (\nu(n) - \nu_y(n)) = \nu_y(n) + O(1),$$

uma vez que o número de divisores primos de n maior do que y é $O(1)$. Nós podemos portanto escrever

$$\sum_{n\leq x} \nu(f(n)) = \sum_{n\leq x} \nu_y(f(n)) + O(x). \tag{3.2.1}$$

Denotemos por $\rho_f(p)$ o número de soluções módulo p da congruência $f(x) \equiv 0 (\mathrm{mod}\ p)$. Então

$$\sum_{n\leq x} \nu_y(f(n)) = \sum_{n\leq x} \sum_{\substack{p\mid f(n)\\ p\leq y}} 1, \tag{3.2.2}$$

e assim, depois de invertemos as somas, nós devemos contar, para p fixo, o número de inteiros $n \leq x$ que pertencem a $\rho_f(p)$ classes de reíduo módulo p. Nós obtemos

$$\sum_{n\leq x} \nu(f(n)) = \sum_{p\leq y} \left(\frac{x\rho_f(p)}{p} + O(\rho_f(p)) \right) + O(x), \tag{3.2.3}$$

e uma vez que $\rho_f(p) \leq \partial(f)$, onde $\partial(f)$ denota o grau de f, nós vemos que o termo do erro vindo da soma acima é $O(y)$.

Iremos invocar agora a teoria algébrica dos números. Seja $K = \mathbb{Q}(\theta)$, onde θ é uma solução de $f(x) = 0$. O anel dos inteiros $\mathcal{O}_K$ de K é um domínio de Dedekind. É um clássico teorema de Dedekind que para todo primo p, exceto um número finito, $\rho_f(p)$ é o número de ideais primos $\mathfrak{p}$ de $\mathcal{O}_K$ tal que a norma $N_{K/\mathbb{Q}}(\mathfrak{p}) = p$. Se $\pi_K(x)$ denota o número de ideais primos cuja norma é $\leq x$, então o análogo do teorema dos números primos para corpos numéricos diz que

$$\pi_K(x) \sim \frac{x}{\log x}$$

quando $x \to \infty$. De fato, para alguma constante $c > 0$,

$$\pi_K(x) = \mathrm{li}(x) + O(xe^{-c\sqrt{\log x}}),$$

onde

$$\mathrm{li}(x) := \int_2^x \frac{dt}{\log t}$$

é a famosa **integral logarítmica**. Uma vez que a norma de qualquer ideal primo é uma potência de primo e o número de ideais primos cuja norma não é um primo não pode exceder $O(x\sqrt{\log x})$, nós deduzimos:

Teorema 3.2.1 (O Teorema de Ideais Primos)**.**

$$\sum_{p \leq x} \rho_f(p) = \mathrm{li}x + O(xe^{-c\sqrt{\log x}}),$$

para algum $c > 0$.

Corolário 3.2.2.

$$\sum_{p \leq x} \frac{\rho_f(p)}{p} = \log\log x + O(1).$$

Demonstração. Exercício. □

Nós podemos agora completar a nossa análise da ordem normal de $\nu(f(n))$. Pela nossa discussão anterior e o corolário acima,

$$\sum_{n\leq x} \nu(f(n)) = x \log\log x + O(x).$$

Também,

$$\begin{aligned}\sum_{n\leq x} \nu^2(n)(f(n)) &= \sum_{n\leq x} \nu_y^2(f(n)) + O\left(\sum_{n\leq x} \nu_y(f(n))\right) \\ &= \sum_{n\leq x} \nu_y^2(f(n)) + O(x\log\log x). \qquad (3.2.4)\end{aligned}$$

Encontramos pelo Teorema Chinês do resto que

$$\sum_{n\leq x} \nu_y^2(f(n)) = \sum_{\substack{p,q\leq y\\ p\neq q}} \left(\frac{x\rho_f(p)\rho_f(q)}{pq} + O(1)\right) + O(x\log\log x),$$

onde o último termo do erro é obtido dos termos quando $p = q$. Uma vez que

$$\sum_p \frac{\rho_f(p)}{p^2} = O(1),$$

nós temos que

$$\sum_{\substack{p,q\leq y\\ p\neq q}} \frac{\rho_f(p)\rho_f(q)}{pq} = \left(\sum_{p\leq y} \frac{\rho_f(p)}{p}\right)^2 + O(1).$$

Assim

$$\sum_{n\leq x} \nu_y^2(f(n)) = x(\log\log x)^2 + O(y^2) + O(x\log\log x).$$

Lembramos que $y = x^\delta$ com $0 < \delta < 1/2$, e assim o termo do erro $O(y^2)$ é dominado por $O(x\log\log x)$. Isto prova:

Teorema 3.2.3.

$$\sum_{n\leq x}(\nu(f(n))-\log\log x)^2 = O(x\log\log x).$$

É agora um exercício elementar deduzir que a ordem normal de $\nu(f(n))$ é $\log\log n$.

O Teorema 3.2.3 da uma estimativa de

$$\ll \frac{x}{\log\log x}$$

para o número de $n \leq x$ tal que $f(n)$ é um número primo. Uma clássica **conjectura de Buniakowski** formulada em 1854 diz que qualquer polinômio irredutível $f(x) \in \mathbb{Z}[x]$, tal que $f(\mathbb{Z}^+)$ não possui divisor comum maior do que 1, representa primos infinitas vezes. Os únicos casos conhecidos desta conjectura é o celebre teorema de Dirichlet sobre a distribuição de primos em progressões aritméticas, o qual estabelece o caso linear. Veremos mais tarde (Capítulo 6) que técnicas de crivos podem ser aplicadas para lançar algumas luzes sobre esta conjectura. Em alguns casos, os métodos chegam próximo de estabelecer a conjectura.

3.3 Estimativas sobre Números Primos

Nós podemos investigar de uma maneira similar $\nu(p-1)$ quando p varia sobre os primos. Mais precisamente, seja k um número natural e defina para $(a,k)=1$ a quantidade

$$\pi(x;k,a) := \#\{p \leq x : p \equiv a(\mathrm{mod}k)\}.$$

Então é fácil de ver que

$$\sum_{p\leq x}\nu(p-1) = \sum_{l\leq x}\pi(x;l,1),$$

onde l denota um número primo racional. Como antes, é conveniente observar que, aplicando o teorema de Chebycheff

$$\sum_{p\leq x}\nu(p-1) = \sum_{\nu_y(p-1)} + O\left(\frac{x}{\log x}\right),$$

e assim obtemos somas da forma

$$\sum_{l\le y}\pi(x;l,1)$$

para investigar com $y=x^{\delta}$ para algum $\delta>0$.

Um teorema clássico de Bombieri e Vinogradov diz que, para todo $A>0$, existe um $B=B(A)>0$ tal que

$$\sum_{k\le x^{1/2}\log^{-B}x}\max_{y\le x}\max_{(a,k)=1}\left|\pi(y;k,a)-\frac{\mathrm{li}(y)}{\phi(k)}\right|\ll\frac{x}{\log^{A}x}.$$

Este teorema será provado no Capítulo 9 e é um dos pontos altos deste curso. De fato, a maior parte do desenvolvimento do método do grande crivo (a ser discutido no Capítulo 8) culminou com esta demonstração. Nós invocamos este teorema para deduzir

$$\sum_{l\le y}\pi(x;l,1)=\sum_{l\le y}\frac{\mathrm{li}(x)}{l-1}+O\left(\frac{x}{\log^{A}x}\right),$$

se $\delta<1/2$. Uma vez que

$$\sum_{l\le y}\frac{1}{l-1}=\sum_{l\le y}\frac{1}{l}+O(1),$$

nós concluímos que

$$\sum_{p\le x}\nu(p-1)=\pi(x)\log\log x+O\left(\frac{x}{\log x}\right).$$

De uma maneira similar deduzimos que (veja Exercício 1)

$$\sum_{p\le x}\nu^{2}(p-1)=\pi(x)(\log\log x)^{2}+O(\pi(x)\log\log x).$$

Este argumento estabelece o seguinte teorema:

Teorema 3.3.1 (Teorema de Erdös)**.**

$$\sum_{p\le x}(\nu(p-1)-\log\log p)^{2}=O\left(\frac{x\log\log x}{\log x}\right).$$

Uma investigação similar pode ser feita para estudar $\nu(p-a)$ para qualquer inteiro a. O caso $a = -2$ está relacionado com a conhecida **conjectura dos primos gêmeos.** Para ser preciso, um primo p é chamado primo gêmeo se $p+2$ também é um primo. É conjecturado que existem infinitos primos gêmeos. Um interessante corolário do análogo do Teorema 3.3.1 para $\nu(p+2)$ é que o número de primos $p \leq x$ tal que $p+2$ também é primo é

$$O\left(\frac{\pi(x)}{\log\log x}\right).$$

3.4 Aplicação do método para outras sequências

O método da ordem normal pode ser formalizado da seguinte maneira.

Seja $\mathcal{A} = (a_n)$ uma sequência finita de números naturais. Seja $\mathcal{A}_1 := \mathcal{A}$ e para cada d livre de quadrados, defina

$$\mathcal{A}_d := \{a_n : a_n \equiv 0 (\text{mod } d)\}.$$

Para cada d livre de quadrados, nós escrevemos

$$\#\mathcal{A}_d = \frac{\omega(d)}{d}X + R_d,$$

ou

$$\#\mathcal{A}_d = \delta_d X + R_d,$$

onde nós pensamos X como sendo uma aproximação para a cardinalidade de $\mathcal{A}$ e R_1 como o erro em tal aproximação. A função $\delta_d = \omega(d)/d$ é para ser pensado como a 'proporção' dos elementos de $\mathcal{A}$ pertencendo a $\mathcal{A}_d$. Em particular, para primos p e q nós escrevemos

$$\#\mathcal{A}_p = \frac{\omega(p)}{p}X + R_p$$

e

$$\#\mathcal{A}_{pq} = \frac{\omega(pq)}{pq}X + R_{pq}.$$

Agora suponha que

$$a_n = O(n^C)$$

para alguma constante positiva C. Como antes, nós podemos mostrar que

$$\sum_{n\leq x} \nu(a_n) = \sum_{n\leq x} \nu_y(a_n) + O(x)$$

para algum $y = x^\delta$ e $\delta > 0$. Então obtemos

$$\sum_{n\leq x} \nu(a_n) = \sum_{p\leq y} \frac{\omega(p)}{p} X + \sum_{p\leq y} R_p + O(x)$$

e vemos que, para procedermos adiante, iremos precisar do comportamento assintótico de

$$\sum_{p\leq y} \frac{\omega(p)}{p}$$

e uma estimativa para a soma dos termos do erro

$$\sum_{p\leq y} R_p.$$

Similarmente, o estudo de

$$\sum_{n\leq x} \nu^2(a_n)$$

levaria a encontrar uma estimativa para a soma

$$\sum_{p,q\leq y} R_{pq}.$$

Iremos desenvolver um crivo geral deste método no próximo capítulo.

No exemplo da Seção 3.1, $a_n = n$, $X = x$ e $\omega(p) = 1$. Você pode determinar a_n, X e $\omega(p)$ para o exemplo da Seção 3.2?(Exercício)

3.5 Exercícios

1. Usando o teorema de Bombieri–Vinogradov prove que

$$\sum_{p\leq x}\nu^2(p-1)=\pi(x)(\log\log x)^2+O(\pi(x)\log\log x).$$

2. Prove que

$$\sum_{n\leq x}(\nu(n)-\log\log n)^2=O(x\log\log x).$$

3. Usando Teorema 3.2.1 prove que existe uma constante positiva c tal que

$$\sum_{p\leq x}\frac{\rho_f(p)}{p}=\log\log x+c+O\left(\frac{1}{\log x}\right).$$

Capítulo 4

O Crivo de Turán

Em 1934, Paul Turán (1910–1976) deu uma demonstração extremamente simples do clássico teorema de Hardy e Ramanujan sobre o número normal de fatores primos de um dado número natural. Para isto, Paul Turán desenvolveu um crivo básico que é agora chamado de o Crivo de Turán. Neste capítulo iremos estudar o Crivo de Turán e ver como ele pode ser usado para tratar outras questões sobre crivos. Por exemplo, o crivo de Turán é mais elementar do que o crivo de Eratóstenes e em alguns casos nos fornece resultados comparáveis com o anterior.

4.1 A Desigualdade Básica

Seja $\mathcal{A}$ um conjunto finito arbitrário e $\mathcal{P}$ um conjunto de números primos. Para cada primo $p \in \mathcal{P}$ nós iremos assumir um dado conjunto $\mathcal{A}_p \subseteq \mathcal{A}$. Seja $\mathcal{A}_1 := \mathcal{A}$ e para todo número livre de quadrados d composto de primos de $\mathcal{P}$ nós definimos

$$\mathcal{A}_d := \cap_{p|d} \mathcal{A}_p.$$

Fixemos um número real positivo z e definimos

$$P(z) := \prod_{\substack{p \in \mathcal{P} \\ p<z}} p.$$

Iremos estar interessado em estimar

$$S(\mathcal{A}, \mathcal{P}, z) := \#(\mathcal{A} \setminus \cup_{p|P(z)} \mathcal{A}_p).$$

Seguindo o método ilustrado na Seção 3.4, nós escrevemos para cada primo $p \in \mathcal{P}$,

$$\#\mathcal{A}_p = \delta_p X + R_p \tag{4.1.1}$$

e para primos distintos $p, q \in \mathcal{P}$,

$$\#\mathcal{A}_{pq} = \delta_p \delta_q X + R_{p,q}, \tag{4.1.2}$$

onde

$$X := \#\mathcal{A},$$

$$0 \leq \delta_p < 1.$$

Por conveniência, interpretaremos $R_{p,q}$ como sendo R_p. Heuristicamente nós podemos pensar de δ_p como a proporção de elementos de $\mathcal{A}$ que pertencem a $\mathcal{A}_p$, e de R_p como o termo do erro nesta estimativa. A mesma interpretação pode ser dada para $\delta_p \delta_q$ e $R_{p,q}$.

Teorema 4.1.1 (O Crivo de Turán)**.** *Nós mantemos a notação anterior. Seja*

$$U(z) := \sum_{p|P(z)} \delta_p.$$

Então

$$S(\mathcal{A}, \mathcal{P}, z) \leq \frac{X}{U(z)} + \frac{2}{U(z)} \sum_{p|P(z)} |R_p| + \frac{1}{U(z)^2} \sum_{p,q|P(z)} |R_{p,q}|.$$

Demonstração. Para cada elemento $a \in \mathcal{A}$, denotemos por $N(a)$ o número de primos $p \mid P(z)$ tal que $a \in \mathcal{A}_p$. Então

$$S(\mathcal{A}, \mathcal{P}, z) = \#\{a \in \mathcal{A} : N(a) = 0\} \leq \frac{1}{U(z)^2} \sum_{a \in \mathcal{A}} (N(a) - U(z))^2.$$

Assim o objetivo é derivar um limite superior para

$$\sum_{a\in\mathcal{A}}(N(a)-U(z))^2,$$

uma expressão que é reminiscente do método da ordem normal. Elevando ao quadrado o somando e expandindo temos

$$\sum_{a\in\mathcal{A}}N(a)^2-2U(z)\sum_{a\in\mathcal{A}}N(a)+XU(z)^2.$$

Para a primeira soma temos

$$\begin{aligned}
\sum_{a\in\mathcal{A}}N(a)^2 &= \sum_{a\in\mathcal{A}}\left(\sum_{\substack{p|P(z)\\ a\in\mathcal{A}_p}}1\right)^2\\
&= \sum_{p,q|P(z)}\#\mathcal{A}_p\cap\mathcal{A}_q\\
&= \sum_{\substack{p,q|P(z)\\ p\neq q}}\#\mathcal{A}_{pq}+\sum_{p|P(z)}\#\mathcal{A}_p\\
&= X\sum_{\substack{p,q|P(z)\\ p\neq q}}\delta_p\delta_q+X\sum_{p|P(z)}\delta_p+\sum_{p,q|P(z)}R_{p,q}\\
&= X\left(\sum_{p|P(z)}\delta_p\right)^2-X\sum_{p|P(z)}\delta_p^2+X\sum_{p|P(z)}\delta_p\\
&+ \sum_{p,q|P(z)}R_{p,q},
\end{aligned}$$

e, similarmente,

$$\sum_{a\in\mathcal{A}}N(a)=X\sum_{p|P(z)}\delta_p+\sum_{p|P(z)}R_p.$$

Aqui nós usamos as equações (4.1.1) e (4.1.2). Portanto

$$\sum_{a\in\mathcal{A}}(N(a)-U(z))^2 = X\sum_{p|P(z)}\delta_p(1-\delta_p)+\sum_{p,q|P(z)}R_{p,q}-2U(z)\sum_{p|P(z)}R_p.$$

Uma vez que $(1-\delta_p)\leq 1$, nós imediatamante deduzimos a cota superior enunciada no teorema. □

4.2 Contando polinômios irredutíveis em $\mathbb{F}_p[x]$

Denotemos por $\mathbb{F}_p$ o corpo finito com p elementos. Fixemos um número natural $n>1$ e seja N_n o número de polinômios mônicos irredutíveis em $\mathbb{F}_p[x]$ de grau n. Iremos obter uma fórmula exata para N_n.

Considere a série de potências

$$\sum_f T^{\partial(f)},$$

onde a soma é sobre todos os polinômios mônicos $f\in\mathbb{F}_p[x]$. Uma vez que o número total de polinômios mônicos em $\mathbb{F}_p[x]$ de grau n é p^n, nós temos que a série de potências acima é

$$\sum_{n=0}^{\infty}p^nT^n = \frac{1}{1-pT}.$$

Por outro lado, $\mathbb{F}_p[x]$ é um domínio Euclidiano e assim ele possui fatoração única. E assim podemos escrever como um produto de Euler a expressão para a série de potência acima

$$\prod_v(1-T^{\partial(v)})^{-1} = \prod_{n=1}^{\infty}(1-T^n)^{-N_n},$$

onde v percorre sobre polinômios mônicos irredutíveis em $\mathbb{F}_p[x]$. Portanto obtemos

$$(1-pT)^{-1} = \prod_{d=1}^{\infty}(1-T^d)^{-N_d}. \tag{4.2.1}$$

Usando que

$$-\log(1-pT) = \sum_{n=1}^{\infty} \frac{p^n T^n}{n}$$

e tomando logaritmos em (4.2.1) nós obtemos

$$\begin{aligned} \sum_{n=1}^{\infty} \frac{p^n T^n}{n} &= -\log(1-pT) = -\sum_{d=1}^{\infty} N_d \log(1-T^d) \\ &= \sum_{d=1}^{\infty}\sum_{e=1}^{\infty} dN_d \frac{T^{de}}{de} = \sum_{n=1}^{\infty} \frac{T^n}{n} \left(\sum_{de=n} dN_d \right). \end{aligned}$$

Isto prova:

Teorema 4.2.1. *Denotemos por N_d o número de polinômios mônicos irredutíveis em $\mathbb{F}_p[x]$ de grau d. Então*

$$\sum_{d|n} dN_d = p^n.$$

Observe que uma consequência imediata é que

$$N_n \leq \frac{p^n}{n},$$

o qual pode ser visto como o análogo para corpos de funções da cota superior de Chebycheff (1.4.1) para $\pi(x)$. De fato, é fácil deduzir que (veja Exercício 1)

$$N_n \sim \frac{p^n}{n},$$

ou mais precisamente,

$$N_n = \frac{p^n}{n} + O(p^{n/2} d(n)),$$

onde $d(n)$ denota o número de divisores de n. E podemos ver este resultado como o análogo do teorema dos números primos (1.4.2) para $\mathbb{F}_p[x]$.

Teorema 4.2.2. *Usando a mesma notação anterior. Temos que*

$$N_n = \frac{1}{n}\sum_{d|n}\mu(d)p^{n/d}.$$

Demonstração. Exercício. □

4.3 Contando polinômios irredutíveis em $\mathbb{Z}[x]$

Fixemos números naturais H e $n > 1$. Iremos aplicar o crivo de Turán para contar o número de polinômios irredutíveis

$$x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$$

com $0 \leq a_i \leq H$, $a_i \in \mathbb{Z}$. Iremos provar que este número é

$$H^n + O(H^{n-1/3}\log^{2/3} H).$$

Assim, um polinômio aleatório com coeficientes inteiros é irredutível, com probabilidade 1.

Notemos que se um polinômio é redutível em $\mathbb{Z}[x]$, então ele é redutível módulo p para algum primo p. Nossa estratégia será obter uma estimativa para o número de polinômios redutíveis.

Seja

$$\mathcal{A} := \{(a_{n-1}, a_{n-2}, \ldots, a_1, a_0) \in \mathbb{Z}^n : 0 \leq a_i < H\}.$$

Iremos pensar das n–úplas $(a_{n-1}, a_{n-2}, \ldots, a_1, a_0)$ como correspondendo ao polinômio mônico

$$x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0.$$

Nós queremos contar o número de úplas de $\mathcal{A}$ que correspondem a polinômios irredutíveis em $\mathbb{Z}[x]$. Assim, denotemos por $\mathcal{P}$ o conjunto de todos os primos e para cada primo p, denotemos por $\mathcal{A}_p$ o subconjunto de úplas correspondendo a polinômios irredutíveis módulo p. Seja $z = z(H)$ um número real positivo a ser escolhido mais tarde.

Então $S(\mathcal{A}, \mathcal{P}, z)$ representa uma cota superior para o número de polinômios redutíveis em $\mathbb{Z}[x]$, pois se um polinômio pertence a $\mathcal{A}_p$ para algum primo p, então ele é irredutível.

Observemos que $\mathcal{A}$ tem H^n elementos. Se especificarmos um polinômio mônico $g(x) \in \mathbb{F}_p[x]$, então o número de elementos de $\mathcal{A}$ que, reduzidos módulo p, são congruenntes a $g(x) (\mathrm{mod}\ p)$, é

$$\left(\frac{H}{p} + O(1)\right)^n.$$

Iremos escolher z satisfazendo $z^2 < H$ assim, para primos $p < z$, a expressão acima pode ser escrita como

$$\frac{H^n}{p^n} + O\left(\frac{H^{n-1}}{p^{n-1}}\right).$$

De nossa discussão anterior, o número de polinômios mônicos irredutíveis de grau n é

$$N_n = \frac{p^n}{n} + O(p^{n/2}),$$

onde a constante implícita depende de n. Assim o número total de polinômios em $\mathcal{A}$ correspondendo a polinômios irredutíveis em $\mathbb{F}_p[x]$ é

$$\left(\frac{H^n}{p^n} + O\left(\frac{H^{n-1}}{p^{n-1}}\right)\right)\left(\frac{p^n}{n} + O(p^{n/2})\right) = \frac{H^n}{n} + O\left(\frac{H^n}{p^{n/2}}\right) + O(H^{n-1}p).$$

Isto implica que

$$\#\mathcal{A}_p = \frac{1}{n}H^n + O(H^{n-1}p) + O(H^n/p^{n/2})$$

e, similarmente, para $p \neq q$

$$\#\mathcal{A}_p \cap \mathcal{A}_q = \frac{1}{n^2}H^n + O(H^{n-1}pq) + O(H^n/p^{n/2}) + O(H^n/q^{n/2}).$$

Podemos agora aplicar o Teorema 4.1.1 com $\delta_p = 1/n$ e

$$R_{p,q} = O(H^{n-1}pq) + O(H^n/p^{n/2}) + O(H^n/q^{n/2}).$$

Usando a cota de Chebycheff nós deduzimos:

Teorema 4.3.1. *Para $\mathcal{A}$ como acima, $n \geq 3$ e $z^2 < H$, temos*

$$S(\mathcal{A}, \mathcal{P}, z) \ll \frac{H^n \log z}{z} + H^{n-1}z^2.$$

Escolhendo

$$z := H^{1/3}(\log H)^{1/3}$$

obtemos:

Teorema 4.3.2. *Seja $n \geq 3$. O número de polinômios redutíveis*

$$x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0, \quad 0 \leq a_i \leq H, \quad a_i \in \mathbb{Z},$$

é $O(H^{n-1/3}(\log H)^{2/3})$.

4.4 Valores quadrados de polinômios

Seja $f(x) \in \mathbb{Z}[x]$ um polinômio com discriminante $\mathrm{disc}(f)$ não–nulo e o qual não é o quadrado de outro polinômio em $\mathbb{Z}[x]$. Seja $H > 0$. Nós iremos agora considerar a questão de estimar

$$\#\{|n| \leq H : f(n) \text{ é um quadrado perfeito}\}.$$

Esta questão foi discutida no Capítulo 2 no contexto do crivo para quadrados. Iremos agora aplicar o crivo de Turán. Como notado anteriormente, o teorema de Siegel sobre pontos inteiros em curvas hiper–elípticas nos fornece uma cota universal (dependendo de f). A novidade aqui é que, por um lado, nós não iremos usar o profundo trabalho de Siegel. Por outro lado, o método irá ter uma ampla aplicabilidade, tal como o caso de polinômios de várias variáveis.

Seja

$$\mathcal{A} := \{n : |n| \leq H\}$$

e para cada primo $p \nmid \operatorname{disc}(f)$, denotemos

$$\mathcal{A}_p := \{|n| \leq H : f(n) (\operatorname{mod}\ p) \text{ não é um quadrado}\}.$$

Por um resultado de Weil citado no Capítulo 2,

$$\left| \sum_{a (\operatorname{mod}\ p)} \left(\frac{f(a)}{p} \right) \right| \leq (\partial(f) - 1)\sqrt{p},$$

e assim temos que o número de $n (\operatorname{mod}\ p)$ tal que

$$y^2 \equiv f(n) (\operatorname{mod}\ p)$$

para algum inteiro y é $p/2 + O(\sqrt{p})$. Logo, o número de $n (\operatorname{mod}\ p)$ tal que $f(n)$ não é um quadrado módulo p é $p/2 + O(\sqrt{p})$. Assim

$$\begin{aligned} \#\mathcal{A}_p &= \left(\frac{2H}{p} + O(1) \right) \left(\frac{p}{2} + O(\sqrt{p}) \right) \\ &= H + O\left(\frac{H}{\sqrt{p}} + p \right). \end{aligned}$$

Similarmente, para primos distintos $p, q \nmid \operatorname{disc}(f)$,

$$\begin{aligned} \#\mathcal{A}_{pq} &= \left(\frac{2H}{pq} + O(1) \right) \left(\frac{pq}{4} + O(\sqrt{pq} + p\sqrt{q}) \right) \\ &= \frac{H}{2} + O\left(pq + \frac{H}{\sqrt{p}} + \frac{H}{\sqrt{q}} \right). \end{aligned}$$

Como $\#\mathcal{A} = 2H + O(1)$, vemos que, usando a notação da Seção 4.1, $\delta_p = 1/2$, os axiomas do crivo de Turán são satisfeitos com

$$R_p = O\left(\frac{H}{\sqrt{p}} + p \right)$$

e

$$R_{pq} = O\left(\frac{H}{\sqrt{p}} + \frac{H}{\sqrt{q}} + pq\right).$$

Aplicando Teorema 4.1.1, obtemos

$$\begin{aligned} & \#\{|n| \le H : f(n) \text{ é um quadrado}\} \\ \ll\ & \frac{H\log z}{z} + \frac{H}{\sqrt{z}} + z + \frac{1}{z^2}(Hz^{3/2} + z^4(\log z)^2). \end{aligned}$$

Escolhendo $z := \frac{H^{2/5}}{(\log H)^{4/5}}$, temos

$$\#\{|n| \le H : f(n) \text{ é um quadrado}\} = O(H^{4/5}(\log H)^{2/5}),$$

que é inferior à estimativa obtida pelo crivo para quadrados. Entretanto, se considerarmos o problema análogo para o qual $f(n)$ é um cubo ou uma potência ímpar maior, o crivo para quadrados não pode ser utilizado, porém uma estimativa similar pode ser deduzida pela técnica acima.

4.5 Exercícios

1. Com N_n como definido neste capítulo, mostre que

$$N_n = \frac{p^n}{n} + O(p^{n/2}d(n)).$$

2. Seja $\mathcal{A}$ o conjunto dos números naturais $n \le x$ e $\mathcal{P}$ o conjunto dos números primos $p < z$. Mostre que

$$\pi(x) \ll \frac{x}{\log\log x}.$$

3. Sejam $\mathcal{A}$ e $\mathcal{P}$ como na Seção 4.1. Para cada $a \in \mathcal{A}$, denotemos por $w(a)$ um número real não–negativo (chamado de peso) e denotemos por

$$X := \sum_{a \in \mathcal{A}} w(a).$$

Suponha que, para $p \in \mathcal{P}$,

$$\sum_{a \in \mathcal{A}_p} w(a) = \delta_p X + R_p$$

e, para primos distintos $p, q \in \mathcal{P}$,

$$\sum_{a \in \mathcal{A}_{pq}} w(a) = \delta_p \delta_q X + R_{p,q}.$$

Seja $N(a)$ o número de $p \in \mathcal{P}$ para o qual $a \in \mathcal{A}_p$ e

$$U(z) := \sum_{p | P(z)} \delta_p.$$

Prove que

$$\sum_{a \in \mathcal{A}} w(a)(N(a) - U(z))^2 = X \sum_{p | P(z)} \delta_p(1 - \delta_p) + \sum_{p,q | P(z)} R_{p,q} - 2U(z) \sum_{p | P(z)} R_p.$$

Capítulo 5

O Crivo de Eratóstenes

Neste capítulo iremos descrever o famoso crivo de Eratóstenes da mesma forma que foi apresentada por A. M. Legendre (1752–1833) em 1808 na segunda edição de seu livro *Théorie des Nombres*. Nós iremos tratar este assunto sob uma perspectiva moderna e vamos mostrar também que, quando combinamos o crivo de Eratóstenes com o "truque de Rankin", o crivo de Eratóstenes se torna tão poderoso quanto o crivo de Brun, sendo este último mais complicado de derivar.

5.1 O crivo de Eratóstenes

Nós iremos usar a propriedade fundamental da função de Möbius, dada no Lema 1.2.2, para estudar o número

$$\Phi(x,z) := \#\{n \leq x : n \text{ não é divisível por nenhum primo } < \text{z}\},$$

onde x, z são números reais positivos. Isto será usado como um exemplo motivador que irá nos levar mais adiante a um desenvolvimento formal do crivo de Eratóstenes.

Se denotarmos

$$P_z := \prod_{p<z} p,$$

então

$$\begin{aligned}\Phi(x,z) &= \sum_{n\le x}\sum_{d|(n,P_z)}\mu(d) = \sum_{d|P_z}\mu(d)\left[\frac{x}{d}\right] \\ &= x\sum_{d|P_z}\frac{\mu(d)}{d} + O(2^z) = x\prod_{p<z}\left(1-\frac{1}{p}\right) + O(2^z).\end{aligned}$$

Podemos usar $\Phi(x,z)$ para deduzirmos uma estimativa para $\pi(x)$ da seguinte maneira:

$$\begin{aligned}\pi(x) &= (\pi(x)-\pi(z)) + \pi(z) \\ &\le \Phi(x,z) + \pi(z) \\ &\le \Phi(x,z) + z. \qquad (5.1.1)\end{aligned}$$

Mais ainda, a desigualdade $1 - x \le e^{-x}$, válida para x positivo, implica que

$$\prod_{p<z}\left(1-\frac{1}{p}\right) \le \exp\left(-\sum_{p<z}\frac{1}{p}\right).$$

No Capítulo 1 nós estabelecemos que

$$\sum_{p<z}\frac{1}{p} \ge \log\log z + O(1).$$

Assim obtemos uma cota superior para

$$\prod_{p<z}\left(1-\frac{1}{p}\right)$$

e consequentemente para $\Phi(x,z)$. Escolhendo agora

$$z := c\log x$$

para alguma constante positiva c e usando (5.1.1), nós provamos

Proposição 5.1.1.

$$\pi(x) \ll \frac{x}{\log\log x}.$$

5.2 Teorema de Merten

Iremos derivar uma fórmula assintótica para

$$V(z) := \prod_{p<z} \left(1 - \frac{1}{p}\right),$$

e para isso iremos usar que a função zeta de Riemann $\zeta(s)$ se estende como uma função analítica para $\Re(s) > 0$ exceto por um pólo simples em $s = 1$ com resíduo 1. Nosso principal resultado nesta seção é

Teorema 5.2.1.

$$V(z) = \prod_{p<z} \left(1 - \frac{1}{p}\right) = \frac{e^{-\gamma}}{\log z} \left(1 + O\left(\frac{1}{\log z}\right)\right)$$

quando $z \to \infty$.

Demonstração. Temos que

$$-\log V(z) = \sum_{p<z} \frac{1}{p} + \sum_{k\geq 2p<z} \frac{1}{kp^k},$$

e uma vez que

$$\sum_{k\geq 2p<z} \frac{1}{kp^k} \leq \sum_{p<z} \sum_{k\geq 2} \frac{1}{p^k} = \sum_{p<z} \frac{1}{p(p-1)},$$

nós podemos escrever

$$\sum_{k\geq 2p<z} \frac{1}{kp^k} = c_0 + O\left(\frac{1}{z}\right)$$

para alguma constante positiva c_0. Assim deduzimos que

$$-\log V(z) = \sum_{p<z} \frac{1}{p} + c_0 + O\left(\frac{1}{z}\right),$$

onde, de fato,

$$c_o = -\sum_{p} \log\left(1 - \frac{1}{p}\right) + \frac{1}{p}\,.$$

Lembremos que pelo Teorema 1.4.3 do Capítulo 1 nós temos

$$R(z) := \sum_{p<z} \frac{\log p}{p} = \log z + O(1),$$

e assim por somas parciais temos que

$$\begin{aligned} \sum_{p<z} \frac{1}{p} &= \frac{R(z)}{\log z} + \int_2^z \frac{R(t)dt}{t \log^2 t} \\ &= \log\log z + c_1 + O\left(\frac{1}{\log z}\right) \end{aligned}$$

para algum $c_1 > 0$. Assim

$$-\log V(z) = \log\log z + c_0 + c_1 + O\left(\frac{1}{\log z}\right),$$

e assim

$$\prod_{p<z} \left(1 - \frac{1}{p}\right) = \frac{e^{-(c_0+c_1)}}{\log z} \left(1 + O\left(\frac{1}{\log z}\right)\right)$$

quando $z \to \infty$. Falta mostramos que $c_0 + c_1 = \gamma$, onde γ é a famosa constante de Euler.

Para $\sigma > 0$ consideremos

$$\zeta(1+\sigma) = \sum_{n=1}^{\infty} \frac{1}{n^{1+\sigma}} = \prod_p \left(1 - \frac{1}{p^{1+\sigma}}\right)^{-1}.$$

Assim

$$\begin{aligned} f(\sigma): &= \log\zeta(1+\sigma) - \sum_p \frac{1}{p^{1+\sigma}} \\ &= -\sum_p \left\{ \log\left(1 - \frac{1}{p^{1+\sigma}}\right) + \frac{1}{p^{1+\sigma}} \right\}, \end{aligned}$$

e portanto

$$c_0 = \lim_{\sigma \to 0} f(\sigma).$$

Pelo Exercício 1 nós temos que, para $\sigma > 0$,

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

Colocando

$$H(t) := \sum_{n \leq t} \frac{1}{n}$$

e

$$P(t) := \sum_{p \leq t} \frac{1}{p},$$

obtemos por somas parciais que

$$\sum_{p} \frac{1}{p^{1+\sigma}} \sigma \int_1^{\infty} \frac{P(u)}{u^{1+\sigma}} du = \sigma \int_0^{\infty} P(e^t) e^{-\sigma t} dt.$$

Similarmente,

$$\log \zeta(1+\sigma) = \sigma \int_0^{\infty} e^{-\sigma t} H(t) dt + O(\sigma).$$

Logo

$$f(\sigma) = \sigma \int_0^{\infty} e^{-\sigma t} (H(t) - P(e^t)) dt + O(\sigma).$$

Uma vez que

$$H(t) = \log t + \gamma + O\left(\frac{1}{t}\right)$$

e

$$P(e^t) = \log t + c_1 + O\left(\frac{1}{t}\right),$$

nós deduzimos que

$$\begin{aligned} f(\sigma) &= \sigma \int_0^\infty e^{-\sigma t}\left(\gamma - c_1 + O\left(\frac{1}{t+1}\right)\right) dt + O(\sigma) \\ &= \gamma - c_1 + O(\sigma). \end{aligned}$$

Tomando $\sigma \to 0$ temos

$$f(0) = c_0 = \gamma - c_1.$$

E isto conclui a demonstração. □

5.3 O truque Rankin

Consideremos a seguinte função

$$\Psi(x,z) := \#\{n \le x : \text{ se } p \mid n, \text{ então } p < z\}.$$

Através do crivo de Eratóstenes discutido na Seção 5.1 e uma análise mais fina temos que

$$\Phi(x,z) = \sum_{d \mid P_z d \le x} \mu(d)\left[\frac{x}{d}\right] = x \sum_{\substack{d \mid P_z \\ d \le x}} \frac{\mu(d)}{d} + O(\Psi(x,z)). \qquad (5.3.1)$$

Nós iremos usar agora um truque devido a Rankin para estimar $\Psi(x,z)$. Para todo $\delta > 0$, temos que

$$\Psi(x,z) = \sum_{\substack{n \le x \\ p \mid n \Rightarrow p < z}} 1 \le \sum_{\substack{n \le x \\ p \mid n \Rightarrow p < z}} \left(\frac{x}{n}\right)^\delta \le x^\delta \prod_{p<z}\left(1 - \frac{1}{p^\delta}\right)^{-1}.$$

Observemos que

$$\prod_{p<z}\left(1-\frac{1}{p^{\delta}}\right)^{-1} \ll \prod_{p<z}\left(1+\frac{1}{p^{\delta}}\right)\left(1-\frac{1}{p^{2\delta}}\right)^{-1} \ll \prod_{p<z}\left(1+\frac{1}{p^{\delta}}\right),$$

uma vez que o produto

$$\prod_{p}\left(1-\frac{1}{p^{2\delta}}\right)^{-1}$$

é convergente para $\delta > 1/2$. Portanto

$$\Psi(x,z) \ll x^{\delta}\prod_{p<z}\left(1+\frac{1}{p^{\delta}}\right).$$

Usando que $1+x \leq e^{x}$ obtemos

$$\Psi(x,z) \ll \exp\left(\delta\log x + \sum_{p<z}\frac{1}{p^{\delta}}\right).$$

Agora escolhedo $\delta := 1-\eta$ com $\eta \to 0$ quando $z \to \infty$, escrevendo

$$p^{-\delta} = p^{-1}p^{\eta} = p^{-1}e^{\eta\log p}$$

e usando a desigualdade $e^{x} \leq 1+xe^{x}$ nós deduzimos que

$$\sum_{p<z}\frac{1}{p^{\delta}} \leq \sum_{p<z}\frac{1}{p}(1+(\eta\log p)z^{\eta}),$$

uma vez que $p < z$. Assim, escolhendo

$$\eta := \frac{1}{\log z}$$

temos:

Teorema 5.3.1.

$$\Psi(x,z) \ll x(\log z)\exp\left(-\frac{\log x}{\log z}\right)$$

quando $x \to \infty$.

Teorema 5.3.2.

$$\sum_{\substack{d|P_z \\ d\leq x}} \frac{\mu(d)}{d} = \prod_{p<z}\left(1-\frac{1}{p}\right) + O\left((\log z)^2 \exp\left(-\frac{\log x}{\log z}\right)\right).$$

Demonstração. Exercício. Use Exercício 2 e o teorema acima. □

Teorema 5.3.3.

$$\Phi(x,z) = x\prod_{p<z}\left(1-\frac{1}{p}\right) + O\left(x(\log z)^2 \exp\left(-\frac{\log x}{\log z}\right)\right)$$

quando $x, z \to \infty$.

Demonstração. Imediata a partir de (5.3.1). □

Corolário 5.3.4.

$$\pi(x) \ll \frac{x}{\log x}(\log\log x).$$

Demonstração. Exercício. □

5.4 O Crivo de Eratóstenes

Seja $\mathcal{A}$ um conjunto qualquer de números naturais $\leq x$ e $\mathcal{P}$ um conjunto de primos. Para cada primo $p \in \mathcal{P}$ denotemos por $\omega(p)$ o número de classes de resíduo (distintas) módulo p. Denotemos por $\mathcal{A}_p$ o conjunto de elementos de $\mathcal{A}$ pertencendo, no mínimo, a uma destas classes de resíduo módulo p e seja $\mathcal{A}_1 := \mathcal{A}$ e para qualquer número d livre de quadrados e composto de primos em $\mathcal{P}$ definimos

$$\mathcal{A}_d := \cap_{p|d}\mathcal{A}_p$$

e

$$\omega(d) := \prod_{p|d}\omega(p).$$

Seja z um número real positivo e

$$P(z) := \prod_{p \in \mathcal{P} p < z} p.$$

Como de costume, denotamos $S(\mathcal{A}, \mathcal{P}, z)$ o número de elementos de

$$A \setminus \cup_{p|P(z)} \mathcal{A}_p.$$

Suponhamos também que exista X tal que

$$\#\mathcal{A}_d = \frac{\omega(d)}{d} X + R_d \tag{5.4.1}$$

para algum R_d.

Teorema 5.4.1 (O Crivo de Eratóstenes)**.** *Com a notação acima, suponha que as seguintes condições sejam satisfeitas:*

1. $|R_d| = O(\omega(d))$;
2. *para algum* $\kappa \geq 0$,

$$\sum_{p|P_z} \frac{\omega(p) \log p}{p} \leq \kappa \log z + O(1);$$

3. *para algum número real positivo* y, $\#\mathcal{A}_d = 0$ *para todo* $d > y$.

Então

$$S(\mathcal{A}, \mathcal{P}, z) = XW(z) + O\left(\left(X + \frac{y}{\log z}\right)(\log z)^{\kappa+1} \exp\left(-\frac{\log y}{\log z}\right)\right),$$

onde

$$W(z) := \prod_{\substack{p \in \mathcal{P} \\ p < z}} \left(1 - \frac{\omega(p)}{p}\right).$$

Para provar este teorema iremos precisar dos seguintes lemas.

Lema 5.4.2. *Com as hipóteses do Teorema 5.4.1, denotemos*

$$F(t,z) := \sum_{\substack{d\le t\\ d\mid P_z}} \omega(d).$$

Então

$$F(t,z) = O\left(t(\log z)^{\kappa}\exp\left(-\frac{\log t}{\log z}\right)\right).$$

Demonstração. Nós usamos o truque de Rankin. Para todo $\delta > 0$,

$$F(t,z) \le \sum_{d\mid P_z} \omega(d)\left(\frac{t}{d}\right)^{\delta}.$$

Uma vez que ω é multiplicativa, nós deduzimos que

$$F(t,z) \le \exp\left(\delta\log t + \sum_{p\mid P_z}\frac{\omega(p)}{p^{\delta}}\right)$$

usando a desigualdade $1+x\le e^x$. Colocando $\delta := 1-\eta$ e usando a desigualdade $e^x \le 1+xe^x$, nós obtemos

$$F(t,z) \le t\exp\left(-\eta\log t + \sum_{p\mid P_z}\frac{\omega(p)}{p} + \eta z^{\eta}\sum_{p\mid P_z}\frac{\omega(p)\log p}{p}\right).$$

Pela segunda hipótese do Teorema 5.4.1 e por somas parciais temos que

$$\sum_{p\mid P_z}\frac{\omega(p)}{p} \le \kappa\log\log z + O(1).$$

Assim

$$F(t,z) \ll t\exp(-\eta\log t + \kappa\log\log z + \kappa\eta(\log z)z^{\eta}).$$

Escolhendo $\eta := 1/\log z$ nos fornece o resultado desejado. □

Lema 5.4.3. *Com as hipóteses do Teorema 5.4.1,*

$$\sum_{\substack{d|P_z \\ d>y}} \frac{\omega(d)}{d} = O\left((\log z)^{\kappa+1} \exp\left(-\frac{\log y}{\log z}\right)\right).$$

Demonstração. Através de somas parciais, nós temos que

$$\sum_{\substack{d|P_z \\ d>y}} \frac{\omega(d)}{d} \ll \int_y^\infty \frac{F(t,z)}{t^2} dt.$$

O resultado segue do Lema anterior. □

Demonstração do Teorema 5.4.1. Pelo princípio de inclusão–exclusão e a primeira e terceira hipóteses

$$\begin{aligned} S(\mathcal{A},\mathcal{P},z) &= \sum_{\substack{d|P_z \\ d\le y}} \mu(d)\#\mathcal{A}_d \\ &= \sum_{\substack{d|P_z \\ d\le y}} \mu(d)\frac{X\omega(d)}{d} + O(F(y,z)). \qquad (5.4.2) \end{aligned}$$

Então pelo Lema 5.4.2 e 5.4.3 nós obtemos que

$$S(\mathcal{A},\mathcal{P},z) = XW(z) + O\left(\left(X + \frac{y}{\log z}\right)(\log z)^{\kappa+1} \exp\left(-\frac{\log y}{\log z}\right)\right).$$

□

Um número primo p tal que $p+2$ também é um primo é chamado de **primo gêmeo** e a famosa **conjectura dos primos gêmeos**, ainda sem solução, diz que existem infinitos primos gêmeos. Esta conjectura é uma das principais motivações de pesquisa em teoria dos crivos. De fato, o nascimento da teoria moderna dos crivos tem seu início com o artigo de Viggo Brun (veja [11]), onde ele provou que

$$\sum_{\substack{p \\ p+2 \text{ primo}}} \frac{1}{p} < \infty.$$

Antes do trabalho de Brun, ninguém sabia como atacar tal questão. Brun deduziu seu resultado a partir de uma estimativa que ele obteve através de um sofisticado método de crivo, chamado agora de Crivo de Brun, que iremos discutir no próximo capítulo. O principal interesse teórico para nós no momento é que a cota superior de Brun pode ser derivada do crivo de Eratóstenes. A cota superior derivada abaixa não é a melhor possível, mas é suficiente para deduzirmos a convergência da soma acima.

Teorema 5.4.4. *O número de primos $p \leq x$ tal que $p+2$ também é primo é*

$$\ll \frac{x(\log\log x)^2}{\log^2 x}.$$

Demonstração. Seja $\mathcal{A}$ o conjunto de números naturais $n \leq x$ e $\mathcal{P}$ o conjunto de todos os primos. Seja $z = z(x)$ um número real positivo a ser escolhido em breve. Para cada primo $p < z$ nós distinguimos as classes de resíduo 0 e -2 módulo p. Uma vez que $\mathcal{A}_p$ é vazio para $p > x+2$, nós aplicamos o Teorema 5.4.1 com $\kappa = 2$ para deduzir que

$$S(\mathcal{A}, \mathcal{P}, z) = xW(z) = O\left(x(\log z)^3 \exp\left(-\frac{\log x}{\log z}\right)\right),$$

onde

$$W(z) := \prod_{p<z}\left(1 - \frac{2}{p}\right).$$

Agora

$$W(z) = \prod_{p<z}\left(1 - \frac{2}{p}\right) \leq \exp\left(-\sum_{p<z}\frac{2}{p}\right) \ll (\log z)^{-2}.$$

Escolhemos z tal que

$$\log z = \log x / A \log\log x$$

para alguma constante positiva A grande e deduzimos que

$$S(\mathcal{A}, \mathcal{P}, z) \ll \frac{x(\log\log x)^2}{\log^2 x}.$$

Claramente, o número de primos gêmeos não pode exceder

$$\pi(z) + S(\mathcal{A}, \mathcal{P}, z) \leq z + S(\mathcal{A}, \mathcal{P}, z),$$

com z como acima. O resultado agora é imediato. □

Corolário 5.4.5 (Teorema de Brun)**.** *A soma*

$$\sum_{\substack{p \\ p+2 \text{ primo}}} \frac{1}{p}$$

é convergente.

Demonstração. Exercício. □

5.5 Exercícios

1. Usando soma parcial, mostre que

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

para $\sigma > 0$.

2. Seja $C(x) = \sum_{n \leq x} c_n$ e $f(t)$ uma função diferenciável com derivadas contínuas. Suponha que

$$\lim_{Y \to \infty} C(Y) f(Y) = 0$$

e

$$\int_1^\infty C(t) f'(t) dt < \infty.$$

Prove que

$$\sum_{n>x} c_n f(n) = -C(x)f(x) - \int_x^{\infty} C(t)f'(t)dt.$$

3. Mostre que

$$\sum_{d^2|n} \mu(d) = \begin{cases} 1 & \text{se } n \text{ é livre de quadrados,} \\ 0 & \text{caso contrário.} \end{cases}$$

Se $Q(x)$ denota o número de números livre de quadrados $\leq x$, deduza que

$$Q(x) = \frac{6}{\pi^2}x + O(\sqrt{x}).$$

Capítulo 6

O Crivo de Brun

Viggo Brun (1885–1978) introduziu o crivo que agora leva o seu nome em 1915 no artigo [5]. Neste trabalho, Brun provou que existem infinitos números inteiros n tal que n e $n+2$ possuem no máximo nove fatores primos. Ele também mostrou que se k um número par e suficientemente grande então ele é a soma de de dois inteiros, cada tendo no máximo nove fatores primos. Isto representa um tremendo avanço em direção à conjectura dos primos gêmeos e para a conjectura de Goldbach. Como consequência do seu trabalho, ele deduziu que a soma dos recíprocos da sequência de primos gêmeos converge.

Como notado anteriormente, alguns dos resultados de Brun podem ser derivados do crivo de Eratóstenes e do truque de Rankin. Entretanto, o método do crivo de Brun não pode ser deduzido dos métodos elementares do capítulo anterior. Assim o estudo do Crivo de Brun é uma essencial ferramenta na teoria dos crivos e não pode ser ignorado.

6.1 O Crivo Puro de Brun

Começamos com a simples observação de que para todos inteiros positivos ν e r tal que $0 \leq r \leq \nu - 1$,

$$\sum_{k\le r}(-1)^k(\nu/k) = (-1)^r\binom{\nu-1}{r}.$$

Esta identidade é obtida quando comparamos os coeficientes de x^r em ambos os lados de

$$(1-x)^{-1}(1-x)^{\nu} = (1-x)^{\nu-1}.$$

Seja n um inteiro positivo e N o radical de n (isto é, o produto dos divisores primos de n contados com multiplicidade 1). Nós usamos a fórmula acima com $\nu = \nu(n)$ para deduzir que, para todo $0 \le r\nu(n) - 1$,

$$\sum_{\substack{d|n\\ \nu(d)\le r}} \mu(d) = \sum_{\substack{d|N\\ \nu(d)\le r}} \mu(d) = \sum_{k\le r}(-1)^k\binom{\nu(n)}{r} = (-1)^r\binom{\nu(n)-1}{r}. \tag{6.1.1}$$

Vamos definir a **função truncada de Möbius** de d por

$$\mu_r(d) := \begin{cases} \mu(d) & \text{se } \nu(d) \le r,\\ 0 & \text{se } \nu(d) > r, \end{cases}$$

e denotemos

$$\psi_r(n) := \sum_{d|n}\mu_r(d).$$

Então (6.1.1) pode ser reescrita como

$$\psi_r(n) = (-1)^r\binom{\nu(n)-1}{r}, \tag{6.1.2}$$

que é uma fórmula que pode ser vista como uma generalização da propriedade fundamental da função de Möbius.

De (6.1.2) nós deduzimos que $\psi_r(n) \ge 0$ se r é par e $\psi_r(n) \le 0$ se r é ímpar. Isto implica que para todos inteiros positivos n e r nós temos

$$\psi_{2r+1}(n) \le \sum_{d|n}\mu(d) \le \psi_{2r}(n). \tag{6.1.3}$$

Também vemos que

$$\psi_{2r+1}(n) = \sum_{\substack{d|n \\ \nu(d)\le 2r}} \mu(d) + \sum_{\substack{d|n \\ \nu(d)=2r+1}} \mu(d)$$

$$= \psi_{2r}(n) + O\left(\sum_{\substack{d|n \\ \nu(d)=2r+1}} |\mu(d)| \right).$$

Combinando as duas fórmulas nós temos que para quaisquer inteiros positivos n e r,

$$\sum_{d|n} \mu(d) = \psi_r(n) + O\left(\sum_{\substack{d|n \\ \nu(d)=2r+1}} |\mu(d)| \right). \tag{6.1.4}$$

O ponto de partida do Crivo de Brun é usar $\psi_r(n)$, via (6.1.4), no crivo de Eratóstenes para reduzir o tamanho dos termos do erro. Para elucidar a idéia de Brun iremos aplicar suas idéias para obter um limitante superior para $\Phi(x,z)$, o número de inteiros $\le x$ que são livres de fatores primos $< z$. Como de costume, denotemos

$$P_z := \prod_{p<z} p.$$

Então por (6.1.3) nós obtemos que, para r um número par,

$$\begin{aligned} \Phi(x,z) &\le \sum_{n\le x} \sum_{d|(n,P_z)} \mu_r(d) \\ &= \sum_{d|P_z} \mu_r(d) \left[\frac{x}{d} \right] \\ &= x \sum_{d|P_z} \frac{\mu_r(d)}{d} + O(z^r), \end{aligned}$$

uma vez que $\mu_r(d) = 0$ a menos que $\nu(d) \le r$.

Agora nós mudamos a nossa atenção para

$$\sum_{d|P_z} \frac{\mu_r(d)}{d}.$$

Por inversão de Möbius,

$$\mu_r(d) = \sum_{\delta|d} \mu\left(\frac{d}{\delta}\right) \psi_r(\delta),$$

e assim

$$\begin{aligned}\sum_{d|P_z} \frac{\mu_r(d)}{d} &= \sum_{d|P_z} \frac{1}{d} \sum_{\delta|d} \mu\left(\frac{d}{\delta}\right) \psi_r(\delta) \\ &= \sum_{\delta|P_z} \frac{\psi_r(\delta)}{\delta} \sum_{d|\frac{P_z}{\delta}} \frac{\mu(d)}{d} \\ &= W(z) \sum_{\delta|P_z} \frac{\psi_r(\delta)}{\phi(\delta)},\end{aligned}$$

onde

$$W(z) := \prod_{p<z} \left(1 - \frac{1}{p}\right)$$

e ϕ denota a função de Euler. Em outras palavras,

$$\sum_{d|P_z} \frac{\mu_r(d)}{d} = W(z) + W(z) \sum_{\substack{\delta|P_z \\ \delta>1}} \frac{\psi_r(\delta)}{\phi(\delta)}. \tag{6.1.7}$$

Nosso objetivo agora é estimar

$$\sum_{\substack{\delta|P_z \\ \delta>1}} \frac{\psi_r(\delta)}{\phi(\delta)}.$$

Primeiramente observamos que, de (6.1.2),

$$\psi_r(\delta) \le \binom{\nu(\delta)-1}{r},$$

e assim a soma considerada é limitada por

$$\begin{aligned}
\sum_{\substack{\delta|P_z\\ \delta>1}} \binom{\nu(\delta)-1}{r}\frac{1}{\phi(\delta)} &\le \sum_{r+1\le m\le \pi(z)} \binom{m-1}{r} \sum_{\substack{\delta|P_z\\ \delta>1\\ \nu(\delta)=m}} \frac{1}{\phi(\delta)} \\
&\le \sum_{r+1\le m\le \pi(z)} \binom{m-1}{r} \left(\sum_{p<z}\frac{1}{p-1}\right)^m \frac{1}{m!} \\
&\le \frac{1}{r!}\sum_{m\ge r+1}\frac{1}{(m-r)!}(\log\log z + c_1)^m \\
&= \frac{(\log\log z + c_1)^r}{r!} \\
&\times \sum_{m\ge r+1}\frac{1}{(m-r)!}(\log\log z + c_1)^{m-r} \\
&\le \frac{(\log\log z + c_1)^r}{r!}\exp(\log\log z + c_1).
\end{aligned}$$

Acima nós utilizamos a estimativa elementar

$$\sum_{p<z}\frac{1}{p} < \log\log z + c_1$$

para alguma constante positiva c_1. Assim

$$\sum_{\substack{\delta|P_z\\ \delta>1}} \frac{\psi_r(\delta)}{\phi(\delta)} \le \frac{(\log\log z + c_1)}{r!}\exp(\log\log z + c_1). \tag{6.1.8}$$

Nós agora utilizamos a estimativa bem conhecida

$$\frac{1}{r!} \le \left(\frac{e}{r}\right)^r$$

para obter de (6.1.8) que

$$\sum_{\substack{\delta|P_z\\ \delta>1}} \frac{\psi_r(\delta)}{\phi(\delta)} \leq c_2 \exp(r - r\log r + r\log\Lambda)\log z, \tag{6.1.9}$$

onde

$$\Lambda := \log\log z + c_1$$

e c_2 é uma constante positiva.

Combinando (6.1.5), (6.1.7) e (6.1.9) nós obtemos que

$$\begin{aligned}\Phi(x,z) &\leq xW(z)\\ &\quad + xW(z)O(\exp(r - r\log r + r\log\Lambda)\log z) + O(z^r).\end{aligned} \tag{6.1.10}$$

O nosso objetivo agora é fazer com que o termo $r\log r$ domine e assim nos permitindo obter um termo de erro pequeno da equação (6.1.10). Nós escolhemos r a ser o inteiro par mais próximo a

$$\eta\log\log z$$

para algum $\eta = \eta(x,z)$ a ser especificado em breve. Com esta escolha de r, o termo do erro em (6.1.10) é igual a

$$x\exp(-\eta(\log\eta)(\log\log z)) + z^{\eta\log\log z},$$

verifique. Igualando os dois termos acima nós percebemos que a escolha ótima para η é

$$\eta := \frac{\alpha\log x}{(\log z)(\log\log z)} \qquad \text{para algum } \alpha < 1.$$

Em particular, para z satisfazendo

$$\log z = O((\log x)^{1-\varepsilon})$$

para qualquer $0 < \varepsilon < 1$, nós obtemos

$$\Phi(x,z) \leq xW(z) + O(x\exp(-(\log x)^{\varepsilon})).$$

Isto nos leva a estimativa

$$\pi(x) = O\left(\frac{x}{(\log x)^{1-\varepsilon}}\right).$$

Uma consideração mais fina e delicada nos mostra que

$$\log z = \frac{\alpha \log x}{\log\log x}$$

e portanto que

$$\pi(x) = O\left(\frac{x \log\log x}{\log x}\right),$$

um resultado que nós ja tínhamos pelo crivo de Eratóstenes.

Estamos agora em posiçào de formalizar o puro crivo de Brun. O ponto de partida é:

Lema 6.1.1. *Sejam n, r inteiros positivos com $r \leq \nu(n)$. Então existe $|\theta| \leq 1$ tal que*

$$\sum_{d|n} \mu(d) = \sum_{d|n\nu(d)\leq r} \mu(d) + \theta \sum_{d|n\nu(d)=r+1} \mu(d).$$

Demonstração. Exercício. □

Seja $\mathcal{A}$ um conjunto qualquer de números naturais $\leq x$ e $\mathcal{P}$ um conjunto de primos. Para cada primo $p \in \mathcal{P}$, seja $\mathcal{A}_p$ o conjunto de elementos de $\mathcal{A}$ os quais são divisíveis por p. Seja $\mathcal{A}_1 := \mathcal{A}$ e para qualquer inteiro positivo d composto de primos de $\mathcal{P}$ denotemos por $\mathcal{A}_d := \cap_{p|d}\mathcal{A}_p$. Seja z um número real positivo e $P(z) := \prod_{\substack{p\in\mathcal{P}\\ p<z}} p$. Como nos capítulos anteriores, nós queremos estimar

$$S(\mathcal{A}, \mathcal{P}, z) := \#(\mathcal{A} \setminus \cup_{p|P(z)} \mathcal{A}_p).$$

Nós iremos assumir que existe uma função multiplicativa $\omega(\cdot)$ tal que, para todo d como acima,

$$\#\mathcal{A}_d = \frac{\omega(d)}{d} X + R_d \tag{6.1.11}$$

para algum R_d, onde

$$X := \#\mathcal{A}.$$

Denotamos por

$$W(z) := \prod_{p|P(z)} \left(1 - \frac{\omega(p)}{p}\right).$$

Teorema 6.1.2 (O Crivo Puro de Brun)**.** *Nós mantemos as mesmas notações de antes e fazemos as seguintes hipóteses adicionais:*

1. $|R_d| \leq \omega(d)$ *para todo* d *livre de quadrados composto de primos de* $\mathcal{P}$*;*

2. *existe uma constante positiva* C *tal que* $\omega(p) < C$ *para todo* $p \in \mathcal{P}$*;*

3. *existem constantes positivas* C_1, C_2 *tal que*

$$\sum_{\substack{p<z \\ p\in\mathcal{P}}} \frac{\omega(p)}{p} < C_1 \log\log z + C_2.$$

Então

$$S(\mathcal{A}, \mathcal{P}, z) = XW(z)(1 + O((\log z)^{-A})) + O(z^{\eta \log\log z})$$

com $A = \eta \log \eta$*. Em particular, se* $\log z \leq c \log x / \log\log x$ *para uma apropriada constante positiva* c *suficientemente pequena, nós obtemos*

$$S(\mathcal{A}, \mathcal{P}, z) = XW(z)(1 + o(1)).$$

Demonstração. Pelo Lema 6.1.1 nós temos que para todo inteiro positivo r,

$$\begin{aligned} S(\mathcal{A},\mathcal{P},z) &= \sum_{a\in\mathcal{A}}\sum_{d|(a,P(z))}\mu(d) \\ &= \sum_{a\in\mathcal{A}}\left(\sum_{d|(a,P(z))}\mu_r(d)+\theta\sum_{\substack{d|(a,P(z))\\ \nu(d)=r+1}}\mu(d)\right) \\ &= \sum_{d|P(z)}\mu_r(d)(\#\mathcal{A}_d)+O\left(X\frac{\pi(z)^{r+1}}{(r+1)!}\right). \end{aligned}$$

De (6.1.11), da primeira hipótese e da multiplicatividade de $\omega(\dot{})$, obtemos

$$\begin{aligned} \mathcal{S}(\mathcal{A},\mathcal{P},z) &= X\sum_{d|P(z)}\frac{\mu_r(d)\omega(d)}{d}+O\left(\sum_{\substack{d|P(z)\\ \nu(d)\le r}}|R_d|\right) \\ &+ O\left(X\frac{z^{r+1}}{(r+1)!}\right) \\ &= X\sum_{d|P(z)}\frac{\mu_r(d)\omega(d)}{d}+O\left(\left(1+\sum_{p|P(z)}\omega(p)\right)^r\frac{1}{r!}\right) \\ &+ O\left(X\frac{z^{r+1}}{(r+1)!}\right). \end{aligned}$$

□

Então, aplicando a fórmula de inversão de Möbius nós deduzimos que

$$\begin{aligned} S(\mathcal{A},\mathcal{P},z) &= X\sum_{\delta|P(z)}\frac{\psi_r(\delta)\omega(\delta)}{\delta}\sum_{d|\frac{P(z)}{\delta}}\frac{\mu(d)\omega(d)}{d} \\ &+ O\left(\left(1+\sum_{p|P(z)}\omega(p)\right)^r\frac{1}{r!}\right)+O\left(X\frac{z^{r+1}}{(r+1)!}\right). \end{aligned}$$

Denotemos

$$\Omega(d) := \prod_{p|d}(p-\omega(p)).$$

Assim a primeira soma na expressão acima é

$$XW(z)\sum_{\delta|P(z)}\frac{\psi_r(\delta)\omega(\delta)}{\Omega(\delta)}.$$

Invocando agora a primeira e a segunda hipóteses do teorema nós obtemos a fórmula assintótica desejada.

6.2 O Principal Teorema de Brun

O crivo puro de Brun nos fornece resultados comparáveis com os do crivo do Eratóstenes. Nós iremos agora tentar melhorar o nosso resultado usando a função

$$f(n) = \sum_{d|n}\mu(d)g(d),$$

com $f(1) = 1$, ao invés de

$$\sum_{d|n}\mu(d),$$

para iniciar o processo de crivo.

Mais precisamente, seja $\mathcal{A}$ um conjunto finito de inteiros e $\mathcal{P}$ um conjunto de primos. Para cada primo $p \in \mathcal{P}$ suponha que seja dado um subconjunto $\mathcal{A}_p \subseteq \mathcal{A}$. Como antes, seja $\mathcal{A}_1 := \mathcal{A}$ e para um inteiro

positivo d livre de quadrados composto de primos de $\mathcal{P}$ denotemos $\mathcal{A}_d := \cap_{p|d}\mathcal{A}_p$. Seja z um número real positivo e $P(z)$ o produto dos primos em $\mathcal{P}$ que são $< z$. Usaremos a notação $\mathcal{P}^{(\delta)}$ para o conjunto de primos de $\mathcal{P}$ com os divisores primos de δ removidos. Aplicando inversão de Möbius para a soma que define f, nós obtemos

$$\mu(n)g(n) = \sum_{d|n} \mu\left(\frac{n}{d}\right) f(d),$$

assim

$$\begin{aligned}
\sum_{d|P(z)} \mu(d)g(d)\#\mathcal{A}_d &= \sum_{d|P(z)} \#\mathcal{A}_d \sum_{\delta|d} \mu\left(\frac{d}{\delta}\right) f(\delta) \\
&= \sum_{\delta|P(z)} f(\delta) \sum_{d|\frac{P(z)}{\delta}} \mu(d)\#\mathcal{A}_{d\delta} \\
&= \sum_{\delta|P(z)} f(\delta) S(\mathcal{A}_\delta, \mathcal{P}^{(\delta)}, z) \\
&= S(\mathcal{A}, \mathcal{P}, z) + \sum_{\substack{\delta|P(z) \\ \delta>1}} f(\delta) S(\mathcal{A}_\delta, \mathcal{P}^{(\delta)}, z).
\end{aligned}$$

Nós podemos interpretar a última soma como um termo de erro. Analisando este termo, Brun descobriu um escolha engenhosa para a função g que permitiu a ele obter uma cota superior e inferior para $S(\mathcal{A}, \mathcal{P}, z)$. Nós iremos fazer esta análise agora.

Primeiramente reescrevemos a soma

$$\sum_{\substack{\delta|P(z) \\ \delta>1}} f(\delta) S(\mathcal{A}_\delta, \mathcal{P}^{(\delta)}, z)$$

em uma forma mais útil. Seja $q(\delta)$ o menor divisor primo de δ, onde interpretamos $q(1)$ como infinito. Escrevemos cada δ como pt com $p = q(\delta)$. Então

$$\begin{aligned}\sum_{\substack{\delta|P(z)\\ \delta>1}} f(\delta)S(\mathcal{A}_\delta,\mathcal{P}^{(\delta)},z) &= \sum_{p|P(z)}\sum_{\substack{t|\frac{P(z)}{p}\\ p<q(t)}} S(\mathcal{A}_{pt},\mathcal{P}^{(pt)},z)f(pt)\\ &= \sum_{p|P(z)}\sum_{\substack{t|P(z)\\ p<q(t)}} S(\mathcal{A}_{pt},\mathcal{P}^{(pt)},z)\\ &\times \left(\sum_{d|t}\mu(d)(g(d)-g(pd))\right).\end{aligned}$$

Agora escrevemos $t = de$ e observemos que $p < q(t)$ se e somente se $p < q(d)$ e $p < q(e)$. Então

$$\begin{aligned}&\sum_{\substack{\delta|P(z)\\ \delta>1}} f(\delta)S(\mathcal{A}_\delta,\mathcal{P}^{(\delta)},z)\\ &\quad= \sum_{d|P(z)}\sum_{\substack{p|P(z)\\ p<q(d)}}\mu(d)(g(d)-g(pd))\sum_{\substack{e|\frac{P(z)}{d}\\ p<q(e)}} S(\mathcal{A}_{pde},\mathcal{P}^{(pde)},z).\end{aligned}$$

Se denotarmos

$$\mathcal{S}(\mathcal{A},\mathcal{P},z) := \mathcal{A}\setminus \cup_{p|P(z)}\mathcal{A}_p,$$

então um momento de reflexão nos mostra que

$$\mathcal{S}(\mathcal{A}_{pd},\mathcal{P}^{(pd)},p) = \coprod_{\substack{e|\frac{P(z)}{d}\\ p<q(e)}} \mathcal{S}(\mathcal{A}_{pde},\mathcal{P}^{(pde)},z),$$

uma vez que todo elemento do lado esquerdo deve necessariamente pertencer a um único $\mathcal{A}_{pde}$ para algum e com $q(e) > p$ (o lado direito é uma união disjunta). Assim

$$\sum_{\substack{\delta|P(z)\\ \delta>1}} f(\delta)S(\mathcal{A}_\delta, \mathcal{P}^{(\delta)}, z)$$
$$= \sum_{d|P(z)} \sum_{\substack{p|P(z)\\ p<q(d)}} \mu(d)(g(d) - g(pd))S(\mathcal{A}_{pd}, \mathcal{P}^{(pd)}, p).$$

Uma vez que $p < q(d)$,

$$S(\mathcal{A}_{pd}, \mathcal{P}^{(pd)}, p) = S(\mathcal{A}_{pd}, \mathcal{P}, p).$$

Isto prova:

Teorema 6.2.1. *Com as mesmas hipóteses anteriores e para toda função g com $g(1) = 1$, nós temos*

$$S(\mathcal{A}, \mathcal{P}, z)$$
$$= \sum_{d|P(z)} \mu(d)g(d)\#\mathcal{A}_d - \sum_{d|P(z)} \sum_{\substack{p|P(z)\\ p<q(d)}} \mu(d)(g(d)-g(pd))S(\mathcal{A}_{pd}, \mathcal{P}, p).$$

O caso quando $g(1) := 1$, $g(d) := 0$ para todo $d > 1$ nos dá o famoso resultado:

Lema 6.2.2 (Identidade de Buchstab)**.** *Nós temos que*

$$S(\mathcal{A}, \mathcal{P}, z) = \#\mathcal{A} - \sum_{p|P(z)} S(\mathcal{A}_p, \mathcal{P}, p).$$

Nós também deduzimos que:

Teorema 6.2.3. *Nós mantemos a notação do Teorema 6.2.1 e sejam g_U e g_L duas funções tal que $g_U(1) = g_L(1) = 1$, satisfazendo*

$$\mu(d)(g_U(d) - g_U(pd)) \geq 0$$

e

$$\mu(d)(g_L(d) - g_L(pd)) \leq 0$$

para todo $d \mid P(z)$, $p < q(d)$. *Então*

$$\sum_{d|P(z)} \mu(d) g_L(d) \#\mathcal{A}_d \leq S(\mathcal{A}, \mathcal{P}, z) \leq \sum_{d|P(z)} \mu(d) g_U(d) \#\mathcal{A}_d.$$

Demonstração. A condição em g_U implica que a segunda soma no Teorema 6.2.1 é não–negativo. A cota superior é agora imediata. Um argumento similar aplica-se para g_L. □

Nós iremos agora escolher as funções g_U e g_L que satisfazem a condição do Teorema 6.2.3. Nós iremos seguir o raciocínio de Brun e considerar apenas funções g tais que $g(d) = 0$ ou 1. Nós decompomos o intervalo $[1, z]$ da seguinte maneira

$$2 = z_r < z_{r-1} < \cdots < z_1 < z_0 = z$$

e definimos

$$P_{z_n,z} := \prod_{z_n \leq p < z} p.$$

Seja b um inteiro positivo fixado e denotemos

$$g_U(d) := \begin{cases} 1 & \text{se } \nu((d, P_{z_n,z})) \leq 2b + 2n - 2 \quad \forall n \leq r, \\ 0 & \text{caso contrário.} \end{cases} \tag{6.2.1}$$

Desta maneira a condição sobre o limitante superior do Teorema 6.2.3 é satisfeito. Para isto precisamos mostrar que

$$\mu(d) = 1 \quad \text{implica} \quad g_U(d) \geq g_U(pd)$$

e

$$\mu(d) = -1 \quad \text{implica} \quad g_U(d) \leq g_U(pd)$$

para todo $d \mid P(z)$ e $p < q(d)$. Suponha que $\mu(d) = 1$. O único caso a considerar é $g_U(d) = 0$ e $g_U(pd) = 1$. Neste caso temos

$$\nu((d, P_{z_m,z})) > 2b + 2m - 2 \quad \text{para algum} \quad m \leq r$$

e

$$\nu((pd, P_{z_n,z})) \leq 2b + 2n - 2 \quad \forall m \leq r.$$

Para $n = m$ a segunda desigualdade contradiz a primeira e assim esta situação não pode ocorrer. Suponha agora que $\mu(d) = -1$. O único caso a considerar é $g_U(pd) = 0$ e $g_U(d) = 1$. Em tal caso,

$$\nu((d, P_{z_n,z})) \leq 2b + 2n - 2 \quad \forall n \leq r$$

e

$$\nu((pd, P_{z_m,z})) > 2b + 2m - 2 \quad \text{para algum} \quad m \leq r.$$

Na segunda desigualdade temos

$$\nu((p, P_{z_m,z})) + \nu((d, P_{z_m,z})) > 2b + 2m - 2,$$

e isto significa que

$$z_m \leq p < z$$

e

$$2b + 2m - 3 < \nu((d, P_{z_m,z})) \leq 2b + 2m - 2.$$

Portanto

$$\nu((d, P_{z_m,z})) = 2b + 2m - 2.$$

Uma vez que $p < q(d)$ e $z_m \leq p < z$, nós deduzimos

$$\nu(d) = \nu((d, P_{z_m,z})) = 2b + 2m - 2.$$

Logo $\mu(d) = 1$, uma contradição! Observe que em todos os casos,

$$(g_U(d) - g_U(pd)) = \mu(d) g_U(d)(1 - g_U(pd)),$$

um fato que será crucial na próxima discussão.

Uma análise similar mostra que

$$g_L(d) := \begin{cases} 1 & \text{se } \nu((d, P_{z_n,z})) \le 2b + 2n - 3 \quad \forall n \le r, \\ 0 & \text{caso contrário.} \end{cases} \tag{6.2.2}$$

satisfaz a condição para g_L no Teorema 6.2.3. Também temos que

$$(g_L(d) - g_L(pd)) = -\mu(d)g_L(d)(1 - g_L(pd)).$$

Denotemos por p^+ o primo de $\mathcal{P}$ que sucede p. Escolhendo g_U e g_L como acima, temos que

Lema 6.2.4.

$$\sum_{d|P(z)} \mu(d)g_U(d)\frac{\omega(d)}{d}$$
$$= W(z)\left(1 + \sum_{p<z} \frac{\omega(p)W(p)}{pW(z)} \sum_{t|P_{p^+,z}} \frac{g_U(t)(1 - g_U(pt))}{t}\omega(t)\right),$$

$$\sum_{d|P(z)} \mu(d)g_L(d)\frac{\omega(d)}{d}$$
$$= W(z)\left(1 - \sum_{p<z} \frac{\omega(p)W(p)}{pW(z)} \sum_{t|P_{p^+,z}} \frac{g_L(t)(1 - g_L(pt))}{t}\omega(t)\right).$$

Demonstração. Iremos estabelecer a identidade para g_U e deixaremos a demonstração para g_L como um exercício para o leitor. Primeiro observamos que para $d \mid P(z)$,

$$\sum_{p|d}(g_U((d, P_{p^+,z})) - g_U((d, P_{p,z}))) = 1 - g_U(d).$$

Para ver isto, note que para $d = 1$ a afirmação é trivial. Para $d > 1$, $d = p_1 \dots p_r$ com $p_1 < \cdots < p_r$ e $p_i \in \mathcal{P}$. Então o lado esquerdo acima é

$$\sum_{i=1}^{r-1}(g_U(p_{i+1} \dots p_r) - g_U(p_i \dots p_r)) + 1 - g_U(p_r)$$

e a soma é telescópica, nos dando o resultado. Logo

$$\begin{aligned}
&\sum_{d|P(z)} \mu(d) g_U(d) \frac{\omega(d)}{d} \\
= &\sum_{d|P(z)} \mu(d) \left(1 - \sum_{p|d} (g_U(d, P_{p^+,z}) - g_U(d, P_{p,z}))\right) \frac{\omega(d)}{d} \\
= \; & W(z) + \sum_{d|P(z)} \sum_{p|d} \mu(d/p)(g_U(d, P_{p^+,z})) - g_U(d, P_{p,z}) \frac{\omega(d)}{d}.
\end{aligned}$$

Escrevemos $d = \delta pt$, onde $\delta \mid P(p)$ e $t \mid P_{p^+,z}$ para obtermos

$$\begin{aligned}
\sum_{d|P(z)} \mu(d) g_U(d) \frac{\omega(d)}{d} \; &= \; W(z) + \sum_{p<z} \frac{\omega(p)}{p} \sum_{\delta|P(p)} \frac{\mu(\delta)\omega(\delta)}{\delta} \\
&\times \sum_{t|P_{p^+,z}} \mu(t) \frac{g_U(t) - g_U(pt)}{t} \omega(t).
\end{aligned}$$

Assim temos que,

$$\begin{aligned}
&\sum_{d|P(z)} \mu(d) g_U(d) \frac{\omega(d)}{d} \\
&\qquad = W(z) + \sum_{p<z} \frac{\omega(p)}{p} W(p) \sum_{t|P_{p^+,z}} \frac{g_U(t)(1 - g_U(pt))}{t} \omega(t),
\end{aligned}$$

como desejado. Isto completa a demonstração. □

Estamos agora em posição de provar o famoso resultado de Brun:

Teorema 6.2.5 (O Crivo de Brun)**.** *Mantemos a notação da Seção 5.1. Suponha que*

1. $|R_d| \leq \omega(d)$ *para todo* d *livre de quadrados de primos de* $\mathcal{P}$*;*

2. existe uma constante $A_1 \geq 1$ *tal que*

$$0 \leq \frac{\omega(p)}{p} \leq 1 - \frac{1}{A_1};$$

3. existem constantes $\kappa > 0$ *e* $A_2 \geq 1$ *tal que*

$$\sum_{w \leq p < z} \frac{\omega(p) \log p}{p} \leq \kappa \log \frac{z}{w} + A_2, \quad \text{se} \quad 2 \leq w \leq z.$$

Sejam b *um inteiro positivo e* λ *um número real que satisfaz*

$$0 < \lambda e^{1+\lambda} < 1.$$

Então

$$S(\mathcal{A}, \mathcal{P}, z) \leq XW(z)\ 1 + 2\frac{\lambda^{2b+1}e^{2\lambda}}{1 - \lambda^2 e^{2+2\lambda}} \exp\left((2b+3)\frac{c_1}{\lambda \log z}\right) + O\left(z^{2b + \frac{2.01}{e^{2\lambda/\kappa}-1}}\right)$$

e

$$S(\mathcal{A}, \mathcal{P}, z) \geq XW(z)\ 1 - \frac{2\lambda^{2b}e^{2\lambda}}{1 - \lambda^2 e^{2+2\lambda}} \exp\left((2b+2)\frac{c_1}{\lambda \log z}\right) + O\left(z^{2b-1 + \frac{2.01}{e^{2\lambda/\kappa}-1}}\right),$$

onde

$$c_1 := \frac{A_2}{2}\ 1 + A_1\left(\kappa + \frac{A_2}{\log 2}\right).$$

Demonstração. Com as escolhas de g_U e g_L como em (6.2.1) e (6.2.2) nós aplicamos o Teorema 6.2.3 e obtemos

$$\begin{aligned} S(\mathcal{A}, \mathcal{P}, z) &\leq \sum_{d | P(z)} \mu(d) g_U(d) \#\mathcal{A}_d \\ &= \sum_{d | P(z)} \mu(d) g_U(d) \left(\frac{X\omega(d)}{d} + R_d\right). \end{aligned}$$

Agora consideremos

$$\sum_{d|P(z)} \mu(d) g_U(d) \frac{\omega(d)}{d}.$$

Pelo Lema 6.2.4, esta soma é igual a

$$W(z) \left(1 + \sum_{p<z} \frac{\omega(p)W(p)}{pW(z)} \sum_{t|P_{p^+,z}} \frac{g_U(t)(1-g_U(pt))}{t} \omega(t) \right).$$

As somas na expressão acima são estimadas por

$$\sum_{n=1}^{r} \sum_{z_n \le p < z_{n-1}} \frac{\omega(p)W(p)}{pW(z)} \sum_{t|P_{p^+,z}} \frac{g_U(t)(1-g_U(pt))}{t} \omega(t)$$

$$\le \sum_{n=1}^{r} \frac{W(z_n)}{W(z)} \sum_{z_n \le p < z_{n-1}} \frac{\omega(p)}{p} \sum_{t|P_{p^+,z}} \frac{g_U(t)(1-g_U(pt))}{t} \omega(t),$$

uma vez que $W(p) \le W(z_n)$ se $z_n \le p < z_{n-1}$. Mais ainda, para cada t que contribuiu para a soma interna na direita, necessariamente temos que $g_U(t) = 1$ e $g_U(pt) = 0$, que signifca

$$\nu((t, P_{z_n,z})) \le 2b + 2n - 2$$

e

$$\nu((pt, P_{z_n,z})) > 2b + 2n - 2,$$

e assim, quando $t \mid P_{z_n,z}$,

$$\nu(t) = \nu((t, P_{z_n,z})) = 2b + 2n - 2.$$

Assim a quantidade que queremos estimar é

$$\le \sum_{n=1}^{r} \frac{W(z_n)}{W(z)} \sum_{\substack{d|P_{z_n,z} \\ \nu(d)=2b+2n-1}} \frac{\omega(d)}{d}$$

$$\leq \sum_{n=1}^{r} \frac{W(z_n)}{W(z)} \frac{1}{(2b+2n-1)!} \left(\sum_{z_n \leq p < z} \frac{\omega(d)}{d} \right)^{2b+2n-1}.$$

Seja λ um número real satisfazendo as condições do teorema. Iremos mostrar que $z_1, \ldots, z_r$ pode ser escolhido tal que

$$\frac{W(z_n)}{W(z)} \leq e^{2(n\lambda+c)} \quad \text{para} \quad n = 1, \ldots, r, \tag{6.2.3}$$

onde

$$c := \frac{c_1}{\log z}.$$

Vamos assumir isto como verdade por um momento e continuarmos com a demonstração.

Nós obtemos

$$\sum_{z_n \leq p < z} \frac{\omega(p)}{p} \leq \sum_{z_n \leq p < z} \log \left(1 - \frac{\omega(p)}{p}\right)^{-1} = \log \frac{W(z_n)}{W(z)} < 2(n\lambda + c)$$

para $n = 1, \ldots, r$. Então segue que

$$\begin{aligned} & \sum_{n=1}^{r} \frac{W(z_n)}{W(z)} \frac{1}{(2b+2n-1)!} \left(\sum_{z_n \leq p < z} \frac{\omega(p)}{p} \right)^{2b+2n-1} \\ \leq & \sum_{n=1}^{r} e^{2n\lambda+2c} \frac{(2n\lambda+2c)^{2b+2n-1}}{(2b+2n-1)!} \\ \leq & e^{2c}(\lambda+c)^{2b-1} \sum_{n=1}^{r} \frac{(2n/e)^{2n}}{(2n)!} \left(1 + \frac{c}{n\lambda}\right)^{2n} (\lambda e^{1+\lambda})^{2n}, \end{aligned}$$

uma vez que

$$(2b+2n-1)! \geq (2n)!(2n)^{2b-1}.$$

Observemos que

$$e^{-n}\frac{n^n}{n!}$$

é decrescente e que

$$\left(1+\frac{c}{n\lambda}\right)^{2n} \leq e^{2c/\lambda},$$

e assim a soma sob consideração é no máximo

$$\begin{aligned} e^{2c}(\lambda+c)^{2b-1}(2e^{-2})e^{2c/\lambda} &= \frac{2\lambda^{2b+1}e^{2\lambda}}{1-\lambda^2e^{2+2\lambda}}\left(1+\frac{c}{\lambda}\right)^{2b-1}e^{2c(1+1/\lambda)} \\ &\leq \frac{2\lambda^{2b+1}e^{2\lambda}}{1-\lambda^2e^{2+2\lambda}}e^{(2b-1)c/\lambda+2c+2c/\lambda} \\ &\leq \frac{2\lambda^{2b+1}e^{2\lambda}}{1-\lambda^2e^{2+2\lambda}}e^{(2b+3)c/\lambda}, \end{aligned}$$

uma vez que $\lambda < 1$. Isto estabelece que

$$\sum_{d|P(z)}\mu(d)g_U(d)\frac{\omega(d)}{d} \leq W(z)\left(1+\frac{2\lambda^{2b+1}e^{2\lambda}}{1-\lambda^2e^{2+2\lambda}}e^{(2b+3)c/\lambda}\right).$$

Uma análise similar nos leva a

$$\sum_{d|P(z)}\mu(d)g_L(d)\frac{\omega(d)}{d} \geq W(z)\left(1-\frac{2\lambda^{2b}e^{2\lambda}}{1-\lambda^2e^{2+2\lambda}}e^{(2b+2)c/\lambda}\right).$$

Precisamos agora estimar

$$\sum_{d|P(z)}g_U(d)|R_d| \leq \sum_{d|P(z)}g_U(d)\omega(d).$$

A soma têm suporte nos números d que satisfazem

$$\nu((d,P_{z_n,z})) \leq 2b+2n-2 \quad \forall n \leq r.$$

Nós iremos mostrar que a soma é

$$\leq \left(1+\sum_{p<z}\omega(p)\right)^{2b}\left(1+\sum_{p<z_1}\omega(p)\right)^{2}\cdots\left(1+\sum_{p<z_r}\omega(p)\right)^{2}.$$

Para ver isto, note que d aparece em

$$\sum_{d|P(z)} g_U(d)\omega(d)$$

e $\nu(d)\leq 2b$, então ele aparece em

$$\left(1+\sum_{p<z}\omega(p)\right)^{2b}.$$

Agora suponha que $2b<\nu(d)\leq 2b+2$. Nem todos os fatores primos de d podem ser maiores do que z_1 uma vez que $\nu((d,P_{z_1,z}))\leq 2b$. Tal d portanto iria aparecer em

$$\left(1+\sum_{p<z}\omega(p)\right)^{2b}\left(1+\sum_{p<z_1}\omega(p)\right)^{2}.$$

Se $2b+2<\nu(d)\leq 2b+4$, então nem todos os fatores primos de d podem ser maiores do que z_2 uma vez que $\nu((d,P_{z_2,z}))\leq 2b+2$. Portanto tal d iria aparecer em

$$\left(1+\sum_{p<z}\omega(p)\right)^{2b}\left(1+\sum_{p<z_1}\omega(p)\right)^{2}\left(1+\sum_{p<z_2}\omega(p)\right)^{2}.$$

Procedendo desta maneira, a afirmação é estabelecida.

Pela terceira hipótese e soma parcial, nós deduzimos

$$\sum_{p<z}\omega(p)(2\mathrm{li}(z)+3),$$

onde $A:=\max(\kappa,A_2)$. Inserindo esta estimativa nos cálculos acima obtemos

$$\sum_{d|P(z)} g_U(d)\omega(d) \leq (1 + A(2\text{li } z + 3))^{2b} \prod_{n=1}^{r-1} (1 + A(2\text{li } z_n + 3))^2.$$

Nós iremos agora escolher os números z_n. Seja α um número real positivo e defina z_n por

$$\log z_n = e^{-n\alpha} \log z \qquad \text{para} \qquad n = 1, \ldots, r.$$

Uma vez que $z_r = 2$, nós temos

$$\log z_{r-1} = e^{-(r-1)\alpha} \log z > \log 2$$

e

$$\log z_r = e^{-r\alpha} \log z = \log 2,$$

e assim temos

$$e^{(r-1)\alpha} < \frac{\log z}{\log 2} = e^{r\alpha}.$$

Logo, para alguma constante absoluta B,

$$\sum_{d|P(z)} g_U(d)\omega(d) = O\left(\left(\frac{Bz}{\log z}\right)^{2b} \prod_{n=1}^{r-1} \left(\frac{Bz_n e^{n\alpha}}{\log z}\right)^2\right).$$

Observemos que, uma vez que z é suficientemente grade,

$$\prod_{n=1}^{r-1} \frac{Be^{n\alpha}}{\log z} = \frac{B^{r-1} e^{\frac{1}{2} r(r-1)\alpha}}{(\log z)^{r-1}} = \frac{B^{r-1} (\log z / \log 2)^{\frac{r-1}{2}}}{(\log z)^{r-1}} < 1,$$

$$\frac{B}{\log z} < 1,$$

e

$$\prod_{n=1}^{r-1} z_n^2 = \exp\left(2\sum_{n=1}^{r-1}\log z_n\right) = \exp\left(-2\log z\sum_{n=1}^{r-1} e^{-n\alpha}\right) \le z^{2/(e^\alpha-1)}.$$

Portanto o termo do erro é

$$O(z^{2b+2/(e^\alpha-1)}).$$

Agora temos que garantir que

$$\frac{W(z_n)}{W(z)} \le e^{2(n\lambda+c)}$$

(veja (6.2.3)). Para isto, observe que, por soma parcial,

$$\frac{W(w)}{W(z)} \le \exp\left(\kappa\log\frac{\log z}{\log w} + \frac{A_2}{\log w}\left(1 + A_1\kappa + \frac{A_1A_2}{\log 2}\right)\right)$$

para todo w, assim

$$\begin{aligned}\frac{W(z_n)}{W(z)} &\le \exp\left(\kappa n\alpha + \frac{2c_1e^{n\alpha}}{\log z}\right)\\ &= e^{2c}\exp\left(\kappa\left(n\alpha + \frac{2c_1e^\alpha}{\log z}\frac{e^{n\alpha}-1}{n}\right)\right)\\ &\le e^{2c}\exp\left(n\alpha\kappa\left(1 + \frac{2c_1e^\alpha}{\kappa\log 2}\frac{1}{\log\frac{\log z}{\log 2}}\right)\right),\end{aligned}$$

para todo $n = 1, \ldots, r$. Colocando

$$\alpha := \frac{2\lambda}{\kappa}\frac{1}{1+\epsilon}, \qquad \epsilon := \frac{1}{200e^{1/\kappa}},$$

nós obtemos

$$\frac{e^{2\lambda/\kappa}-1}{e^\alpha-1} \le \frac{2.01}{2},$$

e assim

$$\sum_{d|P(z)} g_u(d)|R_d| = O\left(z^{2b+\frac{2.01}{e^{2\lambda/\kappa}-1}}\right).$$

A análise de

$$\sum_{d|P(z)} g_L(d)|R_d|$$

é similar. Isto completa a demonstração do teorema. □

Vamos aplicar o Crivo de Brun no problema sobre primos gêmeos:

Teorema 6.2.6. *Quando $x \to \infty$,*

$$\#\{n \le x : n \text{ e } n+2 \text{ possuem no máximo sete fatores primos}\} \gg \frac{x}{(\log x)^2}.$$

Demonstração. Considere a sequência

$$\mathcal{A} := n(n+2) : \ n \le x \ .$$

Gostaríamos de contar o número de elementos nesta sequência que são livres de qualquer fator primo $< z$ para algum número z fixado. Assim, na notação do Teorema 6.2.5, $X = x$, $\mathcal{P}$ é o conjunto dos primos racionais, $\omega(2) = 1$ e $\omega(p) = 2$ para $p > 2$. As hipóteses do teorema são satisfeitas com $A_0 = \kappa = 2$, $A_1 = 3$, e assim para $b = 1$ obtemos

$$S(\mathcal{A}, \mathcal{P}, z) > \frac{1}{2}x \prod_{2<p<z}\left(1-\frac{2}{p}\right)\left(1-\frac{2\lambda^2 e^{2\lambda}}{1-\lambda^2 e^{2+2\lambda}}\exp\left(\frac{4c_1}{\lambda\log z}\right)\right) + O(z^{1+\frac{2.01}{e^{\lambda}-1}}).$$

Observemos que, para $z \to \infty$,

$$\exp\left(\frac{4c_1}{\lambda\log z}\right) = 1 + O(1/\log z).$$

Agora nós queremos escolher λ tal que

$$\frac{2\lambda^2 e^{2\lambda}}{1-\lambda^2 e^{2+2\lambda}} < 1.$$

Isto é,

$$\lambda e^{\lambda} < \frac{1}{\sqrt{2+e^2}}.$$

Usando tabela de logaritmos nós encontramos que

$$\lambda := \log(1.288)$$

satisfaz a desigualdade acima. Para esta escolha de λ, $7 > 2.01/(e^{\lambda}-1)$. Portanto

$$S(\mathcal{A},\mathcal{P},z) \gg \frac{x}{\log^2 z} + O(z^{\theta})$$

com $\theta < 8$.

Finalmente, escolhendo $z := x^{1/u}$ com $u < 8$, nós deduzimos que para no mínimo

$$\gg \frac{x}{(\log x)^2}$$

números $n \leq x$, ambos n e $n+2$ possuem no máximo sete fatores primos. □

6.3 Exercícios

1. Seja $\pi(x,z)$ o número de $n \leq x$ tais que não são divisíveis por nenhum primo $p \leq z$. Mostre que para r par,

$$\pi(x,z) \leq x \sum_{d|P_z} \frac{\mu_r(d)}{d} + O(z^r).$$

2. Usando o Crivo de Brun, deduza que

$$\pi(x) \ll \frac{x}{\log x}.$$

Capítulo 7

O Crivo de Selberg

Em 1940s Atle Selberg descobriu um novo método de crivo enquanto estudava os zeros da função zeta de Riemann. No seu estudo da função zeta de Riemann, Selberg desenvolveu o que é chamado hoje em dia de técnica de molificação, que foi a percussora do Crivo de Selberg, e ele usou tais técnicas para mostrar que uma proporção positiva dos zeros da função zeta de Riemann estão na linha crítica $\mathrm{Re}(s) = 1/2$.

O Crivo de Selberg em essência possui uma estrutura combinatorial e pode ser o caso que a sua versatilidade ainda não foi totalmente percebida e empregada.

7.1 O Teorema de Chebycheff Revisitado

Lembramos que na Seção 1.3 do Capítulo 1 nós usamos um argumento combinatorial, devido a Chebycheff, para mostrar que $\pi(x) = O(x/\log x)$. E então nos capítulos seguintes, o uso do crivo de Turán, crivo de Eratóstenes e até mesmo do crivo de Brun nos levaram a limitantes superiores mais fracos para $\pi(x)$, a saber $\pi(x) = O(x/\log\log x)$ (veja Proposição 5.1.1 do Capítulo 5) e $\pi(x) = O(x\log\log x/\log x)$ (veja Corolário 5.3.4 do Capítulo 5). A demonstração deste último resultado foi baseada no príncipio de inclusão-exclusão expresso na forma

$$\Phi(x,z) = \sum_{d|P_z} \mu(d) \sum_{\substack{n\le x \\ d|n}} 1, \tag{7.1.1}$$

e, consequentemente, na equação

$$\Phi(x,z) = \sum_{n\le x} \sum_{d|(n,P_z)} \mu(d), \tag{7.1.2}$$

onde P_z é, como usual, o produto de todos os números primos $< z$. Uma análise refinada de (7.1.2) eventualmente nos leva a fórmula

$$\Phi(x,z) = x \prod_{p<z} \left(1 - \frac{1}{p}\right) + O\left(x(\log z)^2 \exp\left(-\frac{\log x}{\log z}\right)\right) \tag{7.1.3}$$

(veja Teorema 5.3.3 do Capítulo 5) e então a estimativa superior para $\pi(x)$ apresentada acima.

Em 1947, Selberg teve a brilhante idéia de trocar a função de Möbius que aparece em (7.1.1) por uma forma quadrática, otimamente escolhida para que as estimativas resultantes sejam mínimas. Mais precisamente, a sua observação crucial foi que para qualquer sequência (λ_d) de números reais tal que

$$\lambda_1 = 1,$$

nós temos

$$\sum_{d|k} \mu(d) \le \left(\sum_{d|k} \lambda_d\right)^2 \tag{7.1.4}$$

para todo k. Usando esta observação em (7.1.2) nós temos

$$\begin{aligned}\Phi(x,z) &\leq \sum_{n\leq x}\left(\sum_{d|(n,P_z)}\lambda_d\right)^2\\ &= \sum_{n\leq x}\left(\sum_{d_1,d_2|(n,P_z)}\lambda_{d_1}\lambda_{d_2}\right)\\ &= \sum_{d_1,d_2|P_z}\lambda_{d_1}\lambda_{d_2}\sum_{\substack{n\leq x\\ [d_1,d_2]|n}}1.\end{aligned}$$

Observando que

$$\#\{n\leq x : n\equiv 0(\text{mod } d)\} = \left[\frac{x}{d}\right] = \frac{x}{d}+O(1),$$

nós obtemos a seguinte desigualdade

$$\Phi(x,z) \leq x\sum_{d_1,d_2|P_z}\frac{\lambda_{d_1}\lambda_{d_2}}{[d_1,d_2]} + O\left(\sum_{d_1,d_2|P_z}|\lambda_{d_1}||\lambda_{d_2}|\right) \qquad (7.1.5)$$

no qual a primeira soma é para ser vista como o termo principal da nossa estimativa e a O–soma como o termo do erro.

Agora, por conveniência, vamos assumir que

$$\lambda_d = 0 \quad \text{para todo} \quad d > z.$$

Isto nos dá

$$\Phi(x,z) \leq x\sum_{d_1,d_2\leq z}\frac{\lambda_{d_1}\lambda_{d_2}}{[d_1,d_2]} + O\left(\sum_{d_1,d_2\leq z}|\lambda_{d_1}||\lambda_{d_2}|\right). \qquad (7.1.6)$$

Observamos também que se tivermos $|\lambda_d| \leq 1$, então (7.1.6) nos daria um termo do erro de $O(z^2)$, o qual, para $z < x$, é menor do que o termo do erro dado pelo crivo de Eratóstenes (veja (7.1.3)). Logo é razoável esperar que o método de Selberg nos fornecerá uma melhora

para as nossas estimativas superiores de $\Phi(x,z)$, e, consequentemente, de $\pi(x)$.

Vamos estimar o termo principal em (7.1.6). A observação chave é ver a soma

$$\sum_{d_1,d_2\le z}\frac{\lambda_{d_1}\lambda_{d_2}}{[d_1,d_2]}$$

como uma forma quadrática em $(\lambda_d)_{d\le z}$, e então buscar minimizar esta forma. Usando que

$$[d_1,d_2](d_1,d_2)=d_1d_2 \tag{7.1.7}$$

e que

$$\sum_{\delta|d}\phi(\delta)=d, \tag{7.1.8}$$

nós podemos escrever

$$\begin{aligned}\sum_{d_1,d_2\le z}\frac{\lambda_{d_1}\lambda_{d_2}}{[d_1,d_2]} &= \sum_{d_1,d_2\le z}\frac{\lambda_{d_1}\lambda_{d_2}}{d_1d_2}(d_1,d_2)\\ &= \sum_{d_1,d_2\le z}\frac{\lambda_{d_1}\lambda_{d_2}}{d_{1,2}}\sum_{\delta|(d_1,d_2)}\phi(\delta)\\ &= \sum_{\delta\le z}\phi(\delta)\sum_{\substack{d_1,d_2\le z\\ \delta|(d_1,d_2)}}\frac{\lambda_{d_1}\lambda_{d_2}}{d_1d_2}\\ &= \sum_{\delta\le z}\phi(\delta)\left(\sum_{\substack{d\le z\\ \delta|d}}\frac{\lambda_d}{d}\right)^2.\end{aligned}$$

Portanto, usando a transformação

$$u_\delta:=\sum_{\substack{d\le z\\ \delta|d}}\frac{\lambda_d}{d}, \tag{7.1.9}$$

a forma quadrática inicial foi diagonalizada para

$$\sum_{\delta \leq z} \phi(\delta) u_\delta^2.$$

Nosso próximo objetivo irá ser minimizar esta nova forma diagonal, se possível.

Nós relembramos que a sequência (λ_d) foi escolhida de tal maneira que $\lambda_1 = 1$ e $\lambda_d = 0$ para $d > z$. Equação (7.1.9) nos diz que também devemos ter condições sobre (u_δ). Isto pode ser visto usando a fórmula dual de inversão de Möbius (Capítulo 1). Nós obtemos

$$\frac{\lambda_\delta}{\delta} = \sum_{\delta | d} \mu\left(\frac{d}{\delta}\right) u_d. \tag{7.1.10}$$

Assim temos as restrições

$$u_\delta = 0 \quad \text{para todo} \quad \delta > z$$

e

$$\sum_{\delta < z} \mu(\delta) u_\delta = 1. \tag{7.1.11}$$

Agora nós usamos (7.1.11) para escrever

$$\sum_{\delta \leq z} \phi(\delta) u_\delta^2 = \sum_{\delta \leq z} \phi(\delta) \left(u_\delta - \frac{\mu(\delta)}{\phi(\delta) V(z)} \right)^2 + \frac{1}{V(z)},$$

onde

$$V(z) := \sum_{d \leq z} \frac{\mu^2(d)}{\phi(d)}. \tag{7.1.12}$$

Desta expressão nós imediatamente vemos que a forma $\sum_{\delta \leq z} \phi(\delta) u_\delta^2$ tem um valor mínimo de $1/V(z)$, obtido em

$$u_\delta = \frac{\mu(\delta)}{\phi(\delta) V(z)}.$$

Com a escolha acima para u_δ, e, consequentemente, com a escolha

$$\lambda_\delta = \delta \sum_{\substack{d \le z \\ \delta | d}} \frac{\mu(d/\delta)\mu(d)}{\phi(d)V(z)} \tag{7.1.13}$$

nós obtemos

$$\Phi(x,z) \le \frac{x}{V(z)} + O\left(\sum_{d_1,d_2 \le z} |\lambda_{d_1}||\lambda_{d_2}| \right).$$

Ainda resta analisar o termo do erro O acima. Usando (7.1.13) nós obtemos

$$\begin{aligned} V(z)\lambda_\delta &= \delta \sum_{\substack{d \le z \\ \delta | d}} \frac{\mu(d/\delta)\mu(d)}{\phi(d)} = \delta \sum_{t \le \frac{z}{\delta}} \frac{\mu(t)\mu(\delta t)}{\phi(\delta t)} \\ &= \delta \sum_{\substack{t \le \frac{z}{\delta} \\ (t,\delta)=1}} \frac{\mu^2(t)\mu(\delta)}{\phi(\delta)\phi(t)} = \mu(\delta) \prod_{p|\delta} \left(1 + \frac{1}{p-1}\right) \sum_{\substack{t \le \frac{z}{\delta} \\ (t,\delta)=1}} \frac{\mu^2(t)}{\phi(t)}. \end{aligned}$$

Isto implica que

$$|V(z)||\lambda_\delta| \le |V(z)|,$$

e assim

$$|\lambda_\delta| \le 1 \quad \text{para todo} \quad \delta.$$

Combinando todos os resultados acima, temos:

Teorema 7.1.1.

$$\Phi(x,z) \le \frac{x}{V(z)} + O(z^2)$$

quando $x, z \to \infty$, *onde* $V(z) := \sum_{d \le z} \frac{\mu^2(d)}{\phi(d)}$.

Do Teorema 7.1.1 nós podemos deduzir o limitante superior de Chebycheff para $\pi(x)$:

Corolário 7.1.2.

$$\pi(x) \ll \frac{x}{\log x}.$$

Demonstração. Exercício. □

7.2 O Crivo de Selberg

Nesta seção iremos formalizar o metódo ilustrado acima. Seja $\mathcal{A}$ qualquer conjunto finito de elementos e $\mathcal{P}$ um conjunto de primos. Para cada primo $p \in \mathcal{P}$, denotemos por $\mathcal{A}_p$ um subconjunto de $\mathcal{A}$. Nós denotamos por d números que são livres de quadrados composto de primos de $\mathcal{P}$. Seja $\mathcal{A}_1 := \mathcal{A}$ e para d inteiro livre de quadrados composto de primos de $\mathcal{P}$ denotemos por $\mathcal{A}_d := \cap_{p|d}\mathcal{A}_p$. Seja z um número real positivo e denotemos por

$$P(z) := \prod_{\substack{p\in\mathcal{P}\\ p<z}} p.$$

Denote por $S(\mathcal{A}, \mathcal{P}, z)$ o número de elementos de

$$\mathcal{A}\setminus \cup_{p|P(z)} \mathcal{A}_p.$$

Teorema 7.2.1 (O Crivo de Selberg, 1947)**.** *Mantemos a notação acima e assumimos que existe $X > 0$ e uma função multiplicativa $f(\cdot)$ satisfazendo $f(p) > 1$ para todo primo $p \in \mathcal{P}$, tal que para qualquer número d livre de quadrados composto de primos de $\mathcal{P}$ nós temos*

$$\#\mathcal{A}_d = \frac{X}{f(d)} + R_d \tag{7.2.1}$$

para algum número real R_d. Nós escrevemos

$$f(n) = \sum_{d|n} f_1(d) \tag{7.2.2}$$

para alguma função multiplicativa $f_1(\cdot)$ que é únicamente determinada por f usando a fórmula de inversão de Möbius; isto é,

$$f_1(n) = \sum_{d|n} \mu(d) f\left(\frac{n}{d}\right).$$

Também, denotemos por

$$V(z) := \sum_{\substack{d\le z\\ d|P(z)}} \frac{\mu^2(d)}{f_1(d)}.$$

Então

$$S(\mathcal{A},\mathcal{P},z) \leq \frac{X}{V(z)} + O\left(\sum_{\substack{d_1,d_2\leq z\\ d_1,d_2|P(z)}} |R_{[d_1,d_2]}|\right).$$

Para provar este teorema iremos precisar da fórmula de inversão dual de Möbius e a seguine generalização:

Lema 7.2.2. *Seja f uma função multiplicativa e d_1, d_2 inteiros positivos e livre de quadrados. Então*

$$f([d_1,d_2])f((d_1,d_2)) = f(d_1)f(d_2).$$

Demonstração. Exercício. □

Demonstração do Teorema 7.2.1. Seja (λ_d) qualquer sequêcia de números reais tal que

$$\lambda_1 = 1$$

e

$$\lambda_d = 0 \quad \text{para todo } d > z.$$

Para $a \in \mathcal{A}$, denotemos

$$D(a) := \prod_{\substack{p\in\mathcal{P}\\ a\in\mathcal{A}_p}} p,$$

com a convenção de que $D(a) := 1$ se $a \notin \mathcal{A}_p$ para qualquer $p \in \mathcal{P}$. Então

$$\sum_{\substack{d|P(z)\\ a\in\mathcal{A}_d}} \mu(d) = \sum_{d|(P(z),D(a))} \mu(d) \leq \left(\sum_{d|(P(z),D(a))\lambda_d}\right)^2 = \left(\sum_{\substack{d|P(z)\\ a\in\mathcal{A}_d}} \lambda_d\right)^2. \tag{7.2.3}$$

Agora vamos olhar para a cardinalidade de $S(\mathcal{A}, \mathcal{P}, z)$. Usando (7.2.3), nós obtemos

$$\begin{aligned} S(\mathcal{A},\mathcal{P},z) &= \sum_{\substack{a\in\mathcal{A}\\ a\notin\mathcal{A}_p\ \forall p|P(z)}} 1 = \sum_{d|P(z)}\mu(d)\sum_{a\in\mathcal{A}_d}1 \\ &= \sum_{a\in\mathcal{A}}\left(\sum_{\substack{d|P(z)\\ a\in\mathcal{A}_d}}\mu(d)\right) \leq \sum_{a\in\mathcal{A}}\left(\sum_{\substack{d|P(z)\\ a\in\mathcal{A}_d}}\lambda_d\right)^2 \\ &= \sum_{a\in\mathcal{A}}\left(\sum_{\substack{d_1,d_2|P(z)\\ a\in\mathcal{A}_{[d_1,d_2]}}}\lambda_{d_1}\lambda_{d_2}\right) = \sum_{d_1,d_2\leq z}\lambda_{d_1}\lambda_{d_2}\#\mathcal{A}_{[d_1,d_2]}. \end{aligned}$$

Usando (7.2.1) a expressão acima é

$$X\sum_{\substack{d_1,d_2\leq z\\ d_1,d_2|P(z)}}\frac{\lambda_{d_1}\lambda_{d_2}}{f([d_1,d_2])} + O\left(\sum_{d_1,d_2\leq z d_1,d_2|P(z)}|\lambda_{d_1}||\lambda_{d_2}||R_{[d_1,d_2]}|\right).$$

Aqui, a primeira soma é vista como o termo principal da nossa estimativa, enquanto a soma envolvendo O é vista como o termo do erro. Como na Seção 7.1, nós também vemos o termo principal como uma forma quadrática em $(\lambda_d)_{d\leq z}$, o qual nós iremos buscar a trazer em uma forma diagonal e minimizar. Pelo Lema 7.2.2 e (7.2.1),

$$\begin{aligned}\sum_{d_1,d_2|P(z)} \frac{\lambda_{d_1}\lambda_{d_2}}{f([d_1,d_2])} &= \sum_{\substack{d_1,d_2\le z\\ d_1,d_2|P(z)}} \frac{\lambda_{d_1}\lambda_{d_2}}{f(d_1)f(d_2)} f((d_1,d_2))\\ &= \sum_{\substack{d_1,d_2\le z\\ d_1,d_2|P(z)}} \frac{\lambda_{d_1}\lambda_{d_2}}{f(d_1)f(d_2)} \sum_{\delta|(d_1,d_2)} f_1(\delta)\\ &= \sum_{\substack{\delta\le z\\ \delta|P(z)}} f_1(\delta) \sum_{\substack{d_1,d_2\le z\\ d_1,d_2|P(z)\\ \delta|(d_1,d_2)}} \frac{\lambda_{d_1}\lambda_{d_2}}{f(d_1)f(d_2)}\\ &= \sum_{\substack{\delta\le z\\ \delta|P(z)}} f_1(\delta) \left(\sum_{\substack{d\le z\\ d|P(z)\\ \delta|d}} \frac{\lambda_d}{f(d)} \right)^2.\end{aligned}$$

Assim, usando a transformação

$$u_\delta := \sum_{\substack{d\le z\\ d|P(z)\\ \delta|d}} \frac{\lambda_d}{f(d)}, \tag{7.2.4}$$

a nossa forma quadrática é reduzida à seguinte forma diagonal

$$\sum_{\substack{\delta\le z\\ \delta|P(z)}} f_1(\delta) u_\delta^2. \tag{7.2.5}$$

A fórmula dual de inversão de Möbius nos permite escrever

$$\frac{\lambda_\delta}{f(\delta)} = \sum_{\substack{d|P(z)\\ \delta|d}} \mu\left(\frac{d}{\delta}\right) u_d, \tag{7.2.6}$$

e assim, lembrando que $\lambda_d = 0$ para $d > z$ e $\lambda_1 = 1$, nós obtemos

$$u_\delta = 0 \quad \textit{para todo } \delta > z$$

e

$$\sum_{\substack{\delta \le z \\ \delta | P(z)}} \mu(\delta) u_\delta = 1.$$

Então podemos escrever, como antes,

$$\sum_{\substack{\delta \le z \\ \delta | P(z)}} f_1(\delta) u_\delta^2 = \sum_{\substack{\delta \le z \\ \delta | P(z)}} f_1(\delta) \left(u_\delta - \frac{\mu(\delta)}{f_1(\delta) V(z)} \right)^2 + \frac{1}{V(z)},$$

do qual nós vemos imediatamente que o valor mínimo da forma quadrática dada em (7.2.5) é $\frac{1}{V(z)}$ e é obtido em

$$u_\delta = \frac{\mu(\delta)}{f_1(\delta) V(z)}. \tag{7.2.7}$$

Observe que aqui nós usamos que $f_1(p) = f(p) - 1 > 0$ para todo $p \in \mathcal{P}$, logo, pela multiplicativade de $f_1(\cdot)$, vemos que os coeficientes de $f_1(d)$ aparecendo em nossa forma são positivos.

Nos resta ainda analisar o termo do erro

$$O\left(\sum_{\substack{d_1, d_2 \le z \\ d_1, d_2 | P(z)}} |\lambda_{d_1}||\lambda_{d_2}||R_{[d_1,d_2]}| \right).$$

Mais precisamente, o nosso objetivo é encontrar limitantes superior para $|\lambda_\delta|$ para todo $\delta \le z$, $\delta \mid P(z)$. De (7.2.4) e (7.2.6) nós obtemos que para tais δ, nós temos

$$\begin{aligned} V(z)\lambda_\delta &= f(\delta) \sum_{\substack{d\le z \\ d|P(z) \\ \delta|d}} \frac{\mu(d/\delta)\mu(d)}{f_1(\delta)} \\ &= f(\delta) \sum_{\substack{t\le \frac{z}{\delta} \\ t|P(z) \\ (t,\delta)=1}} \frac{\mu^2(t)\mu(\delta)}{f_1(t)f_1(\delta)} \\ &= \mu(\delta) \left(\prod_{p|\delta} \frac{f(p)}{f_1(p)} \right) \sum_{\substack{t\le \frac{z}{\delta} \\ t|P(z) \\ (t,\delta)=1}} \frac{\mu^2(t)}{f_1(t)} \\ &= \mu(\delta) \left(\prod_{p|\delta} \left(1 + \frac{1}{f_1(p)} \right) \right) \sum_{\substack{t\le \frac{z}{\delta} \\ t|P(z) \\ (t,\delta)=1}} \frac{\mu^2(t)}{f_1(t)}. \end{aligned}$$

Portanto

$$|V(z)||\lambda_\delta| \le |V(z)|,$$

e assim

$$|\lambda_\delta| \le 1.$$

Consequentemente, o termo do erro é

$$O\left(\sum_{\substack{d_1,d_2\le z \\ d_1,d_2|P(z)}} |R_{[d_1,d_2]}| \right),$$

e isto completa a demonstração do teorema. □

Para podermos usar o Teorema 7.2.1 nós precisamos de cotas inferiores para a quantidade $V(z)$.

Lema 7.2.3. *Mantemos a notação do Teorema 7.2.1. Denotamos por $\tilde{f}(\cdot)$ a ser a função completamente multiplicativa definida por $\tilde{f}(p) := f(p)$ para todos os primos p e denotamos*

$$\bar{P}(z) := \prod_{\substack{p \notin \mathcal{P} \\ p<z}} p.$$

Então

1.
$$V(z) \geq \sum_{\substack{\delta \leq z \\ p|\delta \Rightarrow p|P(z)}} \frac{1}{\tilde{f}(\delta)};$$

2.
$$f(\bar{P}(z))V(z) \geq f_1(\bar{P}(z)) \sum_{\delta \leq z} \frac{1}{\tilde{f}(\delta)}.$$

Demonstração. 1. Iremos expressar o quociente $1/f_1(d)$ aparecendo em $V(z)$ em termos da função f. Para isto, nós observamos que para d um inteiro livre de quadrados tal que $d \mid P(z)$ nós temos

$$\begin{aligned} \frac{f(d)}{f_1(d)} &= \prod_{p|d} \frac{f(p)}{f_1(p)} = \prod_{p|d} \left(1 - \frac{1}{f(p)}\right)^{-1} \\ &= \prod_{p|d} \sum_{n \geq 0} \frac{1}{f(p)^n} = \sum_{k}{}' \frac{1}{\tilde{f}(k)}, \end{aligned}$$

onde a soma $\sum_k'$ é realizda sobre inteiros k compostos de divisores primos de d. Então

$$V(z) = \sum_{\substack{d \leq z \\ d|\bar{P}(z)}} \frac{\mu^2(d)}{f(d)} \sum_{k}{}' \frac{1}{\tilde{f}(k)} \geq \sum_{\substack{\delta \leq z \\ p|\delta \Rightarrow p|P(z)}} \frac{1}{\tilde{f}(\delta)}.$$

2. Similar à parte 1, nós escrevemos

$$\begin{aligned}\frac{f(\bar{P}(z))}{f_1(\bar{P}(z))}V(z) &= \prod_{\substack{p\notin\mathcal{P}\\ p<z}}\left(1-\frac{1}{f(p)}\right)^{-1}\sum_{\substack{d\leq z\\ d|\bar{P}(z)}}\frac{\mu(d)^2}{f(d)}\sum_k{}'\frac{1}{\tilde{f}(k)}\\ &= \prod_{\substack{p\notin\mathcal{P}\\ p<z}}\left(\sum_{n\geq 0}\frac{1}{f(p)^n}\right)\sum_{\substack{d\leq z\\ d|\bar{P}(z)}}\frac{\mu(d)^2}{f(d)}\sum_k{}'\frac{1}{\tilde{f}(k)}\\ &\geq \sum_{\delta\leq z}\frac{1}{\tilde{f}(z)}. \qquad (7.2.8)\end{aligned}$$

□

7.3 O Teorema de Brun–Titchmarsh e algumas aplicações

Nesta seção iremos utilizar o Crivo de Selberg para estimar o número de primos $p \leq x$ em uma dada progressão aritmética. Mais precisamente, para quaisquer inteiros a e k primos entre si nós iremos estimar

$$\pi(x;k,a) = \#\{p \leq x : p \equiv a(\text{mod } k)\}.$$

O problema de estudar o comportamento assintótico de $\pi(x;k,a)$ tem seu ínicio em Legendre, que conjecturou que dados $(a,k) = 1$, existem infinitos primos p tal que $p \equiv a(\text{ mod } k)$. É também um caso particular da conjectura de Buniakowski relativa aos valores primos de polinômios com valores inteiros.

No final da década de 1830, Dirichlet provou a conjectura de Legendre usando propriedades importantes de certas generalizações da função zeta de Riemann (conhecida como L–funções de Dirichlet). O trabalho de Dirichlet neste problema é visto agora como o começo da teoria analítica dos números. Mais precisamente, Dirichlet mostrou que a **densidade analítica** do conjunto de primos $p \equiv a(\text{mod } k)$ é $1/\phi(k)$, isto é, que

$$\lim_{s\to 1}\frac{\sum_{p\equiv a(\bmod k)}1/p^s}{\log 1/(s-1)}=\frac{1}{\phi(k)}.$$

Na veredade, existe uma definição mais natural para a densidade de um conjunto de primos: um conjunto $\mathcal{P}$ de primos é dito ter **densidade natural** δ se o limite

$$\lim_{x\to\infty}\frac{\#\{p\le x : p\in\mathcal{P}\}}{\#\{p\le x\}}$$

existe e é igual a δ. Pode ser provado que se $\mathcal{P}$ tem densidade natural δ, então ele também possui densidade analítica δ, entretanto a recíproca não é verdadeira em geral. No caso do conjunto de primos $p\equiv a(\bmod k)$, a densidade natural existe e então, pelo teorema de Dirichlet, nós obtemos a fórmula assintótica

$$\pi(x;k,a)\sim\frac{1}{\phi(k)}\mathrm{li}(x).$$

Fórmulas assintóticas efetivas para $\pi(x;k,a)$ também são conhecidas:

1. Para todo $N>0$, existe $c(N)>0$ tal que, se

   $$k\le(\log x)^N,$$

   então

   $$\pi(x;k,a)=\frac{1}{\phi(k)}\mathrm{li}(x)+O(x\exp(-c(N)(\log x)^{1/2})),$$

   uniformemente em k (este resultado é conhecido como **teorema de Siegel–Walfisz**);

2. Assumindo uma hipótese de Riemann generalizada (para L–funções de Dirichlet) nós temos que para qualquer $k\le x$,

   $$\pi(x;k,a)=\frac{1}{\phi(k)}\mathrm{li}(x)+O(x^{1/2}\log(kx)).$$

   A constante no termo O não depende de k.

Assim, incondicionalmente, os termos do erro são conhecidos (somente) em um domínio de $k < (\log x)^N$. Finalmente, o teorema de Bombieri–Vinogradov mencionado na Seção 3.3 do Capítulo 3 nos permite controlar estes termos do erro, em 'média' e incondicionalmente, em domínios maiores de k. Este importante teorema irá ser discutito em detalhe no Capítulo 9.

Nosso objetivo nesta seção é mais modesto do que os resultados mencionados acima. Nós iremos obter um limitante superior para $\pi(x;k,a)$. Mais precisamente, nós iremos provar:

Teorema 7.3.1 (O Teorema de Brun–Titchmarsh)**.** *Sejam a e k inteiros positivos e primos entre si e seja x um númer real positivo tal que $k \leq x^\theta$ para algum $\theta < 1$. Então, para todo $\varepsilon > 0$, existe $x_0 = x_0(\varepsilon) > 0$ tal que*

$$\pi(x;k,a) \leq \frac{(2+\varepsilon)x}{\phi(k)\log(2x/k)}$$

para todo $x > x_0$.

Demonstração. Nós fixamos um número real positivo $z < x$ e observamos que

$$\begin{aligned} & \pi(x;k,a) = \pi(z;k,a) + \#\{z \leq p \leq x : p \equiv a (\mathrm{mod}\ k)\} \\ \leq\ & z + \#\{n \leq x : n \equiv a (\mathrm{mod}\ k),\ n \not\equiv 0 (\mathrm{mod}\ p) \forall p < z, (p,k) = 1\}. \end{aligned}$$

Se escolhermos

$$\mathcal{A} := \{n \leq x : n \equiv a (\mathrm{mod}\ k)\},$$

e

$$\mathcal{A}_p := \{n \leq x : n \equiv a (\mathrm{mod}\ k), n \equiv 0 (\mathrm{mod}\ p)\} \qquad \text{para todo } p \in \mathcal{P},$$

$$P(z) := \prod_{\substack{p<z \\ (p,k)=1}} p,$$

então temos

$$S(\mathcal{A}, \mathcal{P}, z) = \#\{n \leq x : n \equiv a(\text{mod } k), n \not\equiv 0(\text{mod } p) \forall p \mid P(z)\}.$$

Logo, uma vez que obtivermos uma cota superior para $S(\mathcal{A}, \mathcal{P}, z)$ nós também iremos obter um limitante superior para $\pi(x; k, a)$.

Uma vez que os primos $p \in \mathcal{P}$ sao coprimos com k, nós podemos usar o Teorema Chinês do Resto para deduzir que

$$\mathcal{A}_d = \frac{x}{kd} + O(1)$$

para todo d livre de quadrados composto de primos de $\mathcal{P}$, onde

$$\mathcal{A}_d := \cap_{p|d} \mathcal{A}_p.$$

Assim, na notação do Teorema 7.2.1, nós temos $X = x/k$, $f(d) = d$, $f_1(d) = \phi(d)$ e $R_d = O(1)$, e assim

$$S(\mathcal{A}, \mathcal{P}, z) \leq \frac{x}{kV(z)} = O(z^2),$$

onde

$$V(z) := \sum_{\substack{d \leq z \\ (d,k)=1}} \frac{\mu^2(d)}{\phi(d)}.$$

Usando parte 2 do Lema 7.2.3 junto com a fórmula (7.1.8), nós obtemos

$$\frac{k}{\phi(k)} V(z) \geq \sum_{\delta \leq z} \frac{1}{\delta} \geq \log(z) + O(1).$$

Logo

$$\pi(x; k, a) \leq z + S(\mathcal{A}, \mathcal{P}, z) \leq \frac{x}{\phi(k)(\log z + O(1))} + O(z^2).$$

Escolhendo

$$z := \left(\frac{2x}{k}\right)^{\frac{1}{2}-\varepsilon}$$

para qualquer $0 < \varepsilon < 1$ fixo, nós obtemos a estimativa superior desejada para $\pi(x; k, a)$. □

Observação 7.3.2. *Montgomery e Vaughan [16] provaram o seguinte resultado mais preciso*

$$\pi(x + y; k, a) - \pi(x; k, a) \leq \frac{2y}{\phi(k) \log(y/k)}$$

para $y > k$. Qualquer melhora na constante 2 na estimativa acima implica a não–existência de zeros de Siegel, como notado por diferente autores.

Seguindo Erdös, nós aplicamos o teorema de Brun–Titchmarsh, o crivo de Brun e o crivo de Eratóstenes ao problema de encontrar uma fórmula assintótica para o número de $n \leq x$ tal que $(n, \phi(n)) = 1$. Nós podemos provar, usando teoremas de Sylow, que esta é uma condição necessária e suficiente para todo grupo de ordem n ser cíclico.

Teorema 7.3.3 (Erdös)**.** *O número de $n \leq x$ tal que $(n, \phi(n)) = 1$ é*

$$\sim \frac{e^{-\gamma} x}{\log \log \log x}$$

quando $x \to \infty$.

Para provar este teorema, nós iremos precisar dos seguintes resultados preliminares.

Lema 7.3.4. *Seja $0 < \varepsilon < 1$ e $p < (\log \log x)^{1-\varepsilon}$. Então*

$$\sum_q{}' \frac{1}{q} > \frac{c_1 \log \log x}{p} > (\log \log x)^{\varepsilon/2},$$

onde o apóstrofo indica que a soma é feita sobre os primos $q \equiv 1 (\bmod p)$ satisfazendo $q < x^{1/(\log \log x)^2}$.

Demonstração. Exercício. □

Lema 7.3.5. *Seja p um primo qualquer. Então*

$$\sum_{\substack{q \leq x \\ q \equiv 1 (\bmod p)}} \frac{1}{q} < c_2 \left(\frac{\log \log x + \log p}{p} \right).$$

Demonstração. Este segue do Teorema 7.3.1 e somas parciais. □

Lema 7.3.6. *Seja* $0 < \varepsilon < 1$ *e* $z \leq (\log\log x)^{1+\varepsilon}$. *Então o número de* $n \leq x$ *não divisíveis por qualquer primo* $p \leq z$ *é*

$$(1 + o(1))\frac{e^{-\gamma}x}{\log z}$$

quando $x \to \infty$.

Demonstração. Isto é uma simples aplicação do crivo de Eratóstenes. □

Lema 7.3.7. *Seja* $0 < \varepsilon < 1$ *e* $z \leq (\log\log x)^{1+\varepsilon}$. *Então o número de* $n \leq x$ *possuindo o último divisor primo* p *é*

$$(1 + o(1))\frac{e^{-\gamma}x}{p\log p}$$

quando $p \to \infty$.

Demonstração. Isto é uma simples consequência do Lema 7.3.6 □

Demonstração do Teorema 7.3.3. Seja $A(x)$ o número de $n \leq x$ tal que $(n, \phi(n)) = 1$. Nós particionamos estes números em conjuntos $\mathcal{A}_p$ de acordo com o menor divisor primo p de n, e nós denotamos por $A_p(x)$ o número de elementos em $\mathcal{A}_p$. Então

$$A(x) = \sum_p A_p(x) = \sum\nolimits_1 + \sum\nolimits_2 + \sum\nolimits_3,$$

onde em $\sum_1$ nós temos $p < (\log\log x)^{1-\varepsilon}$, em $\sum_2$ nós temos $(\log\log x)^{1-\varepsilon} < p < (\log\log x)^{1+\varepsilon}$, e em $\sum_3$ nós temos $p > (\log\log x)^{1+\varepsilon}$ para qualquer $0 < \varepsilon < 1$ fixado.

Observe que para cada primo p, os números enumerados por $A_p(x)$ não possuem qualquer fator primo $q \equiv 1(\bmod\ p)$. Pelo crivo de Brun e Lema 7.3.4,

$$A_p(x) \ll \frac{x}{p} \prod_{\substack{q \equiv 1(\bmod\ p) \\ q < x^{1/(\log\log x)^2}}} \left(1 - \frac{1}{q}\right) \ll \frac{x}{p}\exp(-(\log\log x)^{\varepsilon/2}).$$

Portanto

$$\sum\nolimits_1 = o\left(\frac{x}{\log\log\log x}\right).$$

Para $\sum_2$ nós usamos Lema 7.3.7 para deduzir que

$$\sum\nolimits_2 \ll \frac{x}{\log\log\log x} \sum_p{}' \frac{1}{p},$$

onde o traço na soma indica que $(\log\log x)^{1-\varepsilon} < p < (\log\log x)^{1+\varepsilon}$.

Nós obtemos

$$\sum_p{}' \frac{1}{p} \ll \varepsilon,$$

e assim

$$\sum\nolimits_2 \ll \varepsilon \frac{x}{\log\log\log x}.$$

Finalmente, pelo crivo de Eratóstenes,

$$\sum\nolimits_3 \le (1+o(1))\frac{e^{-\gamma}x}{(1+\varepsilon)\log\log\log x},$$

uma vez que todos os divisores primos de n enumerados por $\sum_3$ são maiores do que $(\log\log x)^{1+\varepsilon}$. Por outro lado,

$$\sum\nolimits_3 \ge (1+o(1))\frac{e^{-\gamma}x}{(1+\varepsilon)\log\log\log x} - \sum_{p>y}\frac{x}{p^2} - \sum_{p>y}\sum_{\substack{q\equiv 1 (\bmod p)\\ pq\le x}}\frac{x}{pq},$$

onde $y := (\log\log x)^{1+\varepsilon}$. É fácil de ver que a penúltima soma é

$$O\left(\frac{x}{\log\log x}\right).$$

A soma final é estimada usando Lema 7.3.5 e é

$$\ll x\sum_{p>y}\frac{\log\log x + \log p}{p^2} \ll \frac{x}{(\log\log x)^{\varepsilon}},$$

por soma parcial. Isto completa a demonstração. □

7.4 Exercícios

1. Seja f uma função multiplicativa. Mostre que

$$f([d_1, d_2])f((d_1, d_2)) = f(d_1)f(d_2).$$

2. Mostre que

$$\sum_{\substack{q<z \\ q\equiv 1(\bmod\ d)}} \frac{1}{q} = \frac{\log\log z}{\phi(d)} + O(1),$$

onde a soma é tomada sobre todos os primos $q \equiv 1(\text{mod } d)$.

3. Prove que

$$\sum_{p\leq \frac{x}{2}} \frac{1}{p\log(x/p)} = O\left(\frac{\log\log x}{\log x}\right).$$

4. Seja a um inteiro positivo. Mostre que o número de primos $p \leq x$ tal que $p + a$ também é um primo é

$$\ll \frac{x}{(\log x)^2} \prod_{p|a} \left(1 + \frac{1}{p}\right).$$

Capítulo 8

O Grande Crivo

Os métodos de crivos apresentados nos capítulos anteriores são baseados no uso da função de Möbius ou de suas variações (como no caso do Crivo de Selberg). Eles são classificados como 'crivos combinatoriais'. Neste capítulo nós iremos discutir um crivo de uma natureza completamente diferente, chamado de **o grande crivo**, o qual foi introduzido por Yuri Linnik em 1941 e foi melhorado por, Rényi (1950), Roth (1965), Bombieri (1965), Davenport e Halberstam (1966), Gallagher (1967), e outros. Nós indicamos o leitor para [7, pg. 151] para uma história sobre o grande crivo. O crivo pode ser deduzido de uma bela desigualdade, conhecida como **desigualdade do grande crivo**, o qual, na primeira vista, não parece possuir o grande potencial que realmente tem.

A motivação original de Linnik foi atacar a **hipótese de Vinogradov** relativa ao tamanho do menor não–resíduo quadrático $n_p (\text{mod } p)$. Vinogradov conjecturou que

$$n_p = O(p^{\varepsilon})$$

para todo $\varepsilon > 0$. A hipótese de Riemann generalizada implica que

$$n_p = O((\log p)^2).$$

Linnik provou, usando o seu grande crivo, que o número de primos $p \leq x$ para o qual $n_p > p^{\varepsilon}$ é $O(\log \log x)$. O artigo de Linnik in-

troduziu um novo tema em teoria analítica dos números empregando idéias da teoria de probabilidade.

Como será visto neste capítulo e no próximo, o grande crivo evoluiu para uma poderosa ferramenta, sua aplicação mais significante sendo o teorema de Bombieri–Vinogradov. Este teorema tem servido frequentemente como um substituto para o uso da hipótese generalizada de Riemann em certos contextos, e irá ser discutido em detalhes no próximo capítulo.

8.1 A desigualdade do grande crivo

Nós começamos com um lema preliminar.

Lema 8.1.1. *Seja $F : [0,1] \to \mathbb{C}$ uma função diferenciável com derivadas contínuas, estendida para todo $\mathbb{R}$ por periodicidade com período 1. Seja z um inteiro positivo. Então*

$$\sum_{d\le z}\sum_{\substack{1\le a\le d\\(a,d)=1}}\left|F\left(\frac{a}{d}\right)\right| \le z^2\int_0^1 |F(\alpha)|d\alpha + \int_0^1 |F'(\alpha)|d\alpha. \tag{8.1.1}$$

Demonstração. Seja $d \le z$, $a \in [1,d] \cap \mathbb{N}$ com $(a,d) = 1$, and $\alpha \in [0,1]$. Então

$$-F\left(\frac{a}{d}\right) = -F(\alpha) + \int_{\frac{a}{d}}^{\alpha} F'(t)dt.$$

Tomando valor absoluto em ambos os lados temos

$$\left|F\left(\frac{a}{d}\right)\right| \le |F(\alpha)| + \int_{\frac{a}{d}}^{\alpha} |F'(t)|dt. \tag{8.1.2}$$

Agora fixamos $\delta > 0$ (a ser escolhido mais tarde) para que os intervalos

$$I\left(\frac{a}{d}\right) := \left(\frac{a}{d} - \delta, \frac{a}{d} + \delta\right)$$

estejam contidos em $[0,1]$. Nós integramos (8.1.2) sobre $I(a/d)$, com respeito a α, e obtemos

$$2\delta \left|F\left(\frac{a}{d}\right)\right| \leq \int_{I(\frac{a}{d})} |F(\alpha)|d\alpha + \int_{I(\frac{a}{d})} \int_{\frac{a}{d}}^{\alpha} |F^{'}(t)|dtd\alpha. \qquad (8.1.3)$$

Uma vez que $\alpha \in I(a/d)$ e $t \in [a/d, \alpha]$, nós obtemos que $t \in I(a/d)$. Logo o lado direito da desigualdade acima é

$$\begin{aligned} &\leq \int_{I(\frac{a}{d})} |F(\alpha)|d\alpha + \int_{I(\frac{a}{d})} \int_{I(\frac{a}{d})} |F^{'}(t)|dtd\alpha \\ &= \int_{I(\frac{a}{d})} |F(\alpha)|d\alpha + 2\delta \int_{I(\frac{a}{d})} |F^{'}(t)|dt \\ &= \int_{I(\frac{a}{d})} |F(\alpha)|d\alpha + 2\delta \int_{I(\frac{a}{d})} |F^{'}(\alpha)|d\alpha. \end{aligned}$$

Em outras palavras,

$$2\delta \left|F\left(\frac{a}{d}\right)\right| \leq \int_{I(\frac{a}{d})} |F(\alpha)|d\alpha + 2\delta \int_{I(\frac{a}{d})} |F^{'}(\alpha)|d\alpha. \qquad (8.1.4)$$

Agora escolhemos

$$\delta := \frac{1}{2z^2}.$$

Com esta escolha, o intervalo $I(a/d)$, $1 \leq a \leq d$, $(a, d) = 1$, $d \leq z$ não se sobrepõem (módulo 1). De fato, $x \in I(a/d) \cap I(a^{'}/d^{'})$ para $a/d \neq a^{'}/d^{'}$. Então

$$\left|x - \frac{a}{d}\right| < \delta, \quad \left|x - \frac{a^{'}}{d^{'}}\right| < \delta,$$

e assim

$$\left|\frac{a}{d} - \frac{a^{'}}{d^{'}}\right| = \frac{|ad^{'} - a^{'}d|}{dd^{'}} \neq 0, \qquad (8.1.5)$$

uma vez que se $ad^{'} = a^{'}d$, então, lembrando que $(a, d) = (a^{'}, d^{'}) = 1$, nós obtemos $d = d^{'}$, o que é falso. Assim

$$\left|\frac{a}{d}-\frac{a^{'}}{d^{'}}\right| \geq \frac{1}{dd^{'}} \geq \frac{1}{z^2}. \tag{8.1.6}$$

Combinando (8.1.5) e (8.1.6) nós somos levados a uma contradição.

Nós efetuamos a soma (8.1.4) sobre todos os intervalos $I(a/d)$ e obtemos

$$\begin{aligned}\frac{1}{z^2}\sum_{d\leq z}\sum_{\substack{1\leq a\leq d\\(a,d)=1}}\left|F\left(\frac{a}{d}\right)\right| &\leq \sum_{I(\frac{a}{d})}\int_{I(\frac{a}{d})}|F(\alpha)|d\alpha\\ &+ \frac{1}{z^2}\sum_{I(\frac{a}{d})}\int_{I(\frac{a}{d})}|F^{'}(\alpha)|d\alpha\\ &\leq \int_0^1|F(\alpha)|d\alpha+\frac{1}{z^2}\int_0^1|F^{'}(\alpha)|d\alpha.\end{aligned}$$

Isto completa a demonstração do lema. □

Agora vamos escolher

$$F(\alpha) := \left(\sum_{n\leq x}a_n e(n\alpha)\right)^2,$$

onde $(a_n)_{n\geq 1}$ é uma sequência arbitrária de números complexos, x é um inteiro positivo e para um número racional t, $e(t) := \exp(2\pi it)$.

Para simplificarmos a notação, denotemos

$$S(\alpha) := \sum_{n\leq x}a_n e(n\alpha),$$

logo

$$F(\alpha) = S(\alpha)^2, \qquad F^{'}(\alpha) = 2S(\alpha)S^{'}(\alpha).$$

Pelo Lema 8.1.1 nós obtemos

$$\sum_{d\le z}\sum_{\substack{1\le a\le d\\(a,d)=1}}\left|S\left(\frac{a}{d}\right)\right|^2 \le z^2\int_0^1|S(\alpha)|^2d\alpha + 2\int_0^1|S(\alpha)S^{'}(\alpha)|d\alpha.$$

Agora invocamos a **identidade de Parseval**, i.e.,

$$\int_0^1\left|\sum_{n\le x}a_n e(n\alpha)\right|^2 d\alpha = \sum_{n\le x}|a_n|^2,$$

e assim

$$\int_0^1|S(\alpha)|^2d\alpha = \sum_{n\le x}|a_n|^2.$$

Isto implica que

$$\sum_{d\le z}\sum_{\substack{1\le a\le d\\(a,d)=1}}\left|S\left(\frac{a}{d}\right)\right|^2 \le z^2\sum_{n\le x}|a_n|^2 + 2\int_0^1|S(\alpha)S^{'}(\alpha)|d\alpha.$$

Para o segundo termo no lado direito da desigualdade acima nós usamos a desigualdade de Cauchy–Schwarz e, novamente, a identidade de Parseval. No final obtemos

$$\begin{aligned}\int_0^1|S(\alpha)S^{'}(\alpha)|d\alpha &\le \left(\int_0^1|S(\alpha)|^2d\alpha\right)^{1/2}\left(\int_0^1|S^{'}(\alpha)|^2d\alpha\right)^{1/2}\\ &\le \left(\sum_{n\le x}|a_n|^2\right)^{1/2}\left(\sum_{n\le x}4\pi^2n^2|a_n|^2\right)^{1/2}\\ &\le 2\pi x\sum_{n\le x}|a_n|^2.\end{aligned}$$

Temos então o seguinte resultado:

Teorema 8.1.2 (A desigualdade do grande crivo)**.** *Seja $(a_n)_{n\geq 1}$ uma sequência de números complexos e x, z inteiros positivos. Então*

$$\sum_{d\leq z}\sum_{1\leq a\leq d(a,d)=1}\left|\sum_{n\leq x}a_n e\left(\frac{na}{d}\right)\right|^2 \leq (z^2+4\pi x)\sum_{n\leq x}|a_n|^2, \quad (8.1.7)$$

onde, para um número racional α, $e(\alpha) := \exp(2\pi i\alpha)$.

8.2 O grande crivo

Queremos deduzir um método de crivo obtido da desigualdade descrita no Teorema 8.1.2. Seja $\mathcal{A}$ um conjunto de inteiros positivos $n \leq x$ e $\mathcal{P}$ um conjunto de primos. Para cada $p \in \mathcal{P}$, suponha que seja dado um conjunto $w_{1,p}, \ldots, w_{\omega(p),p}$ de $\omega(p)$ classes de resíduo módulo p. Seja z um número real positivo e denote por $P(z)$ o produto dos primos $p \in \mathcal{P}$, $p < z$. Denotemos

$$\mathcal{S}(\mathcal{A},\mathcal{P},z) := \{n \in \mathcal{A} : n \not\equiv w_{i,p}(\mathrm{mod}\ p)\forall 1 \leq i \leq \omega(p),\ \forall p \mid P(z)\} \quad (8.2.1)$$

e denotamos por $S(\mathcal{A},\mathcal{P},z)$ a cardinalidade deste conjunto.

Teorema 8.2.1 (O grande crivo)**.** *Com a notação acima, nós temos*

$$S(\mathcal{A},\mathcal{P},z) \leq \frac{z^2+4\pi x}{L(z)},$$

onde

$$L(z) := \sum_{d\leq z}\mu^2(d)\prod_{p\mid d}\frac{\omega(p)}{p-\omega(p)}.$$

A idéia para a demonstração do Teorema 8.2.1 é usar somas do tipo

$$c_d(n) := \sum_{\substack{1\leq a\leq d\\(a,d)=1}} e\left(\frac{na}{d}\right), \quad (8.2.2)$$

onde $n, d \in \mathbb{N}$ e as somas acima são chamadas de **somas de Ramanujan**. Estas somas tem as seguintes propriedades:

Proposição 8.2.2. *Sejam d, $d^{'}$ inteiros positivos. Então*

1. *se $(d, d^{'}) = 1$, nós temos que $c_{dd'}(n) = c_d(n)c_{d'}(n)$;*

2. $$c_d(n) = \sum_{D|(d,n)} \mu\left(\frac{d}{D}\right) D;$$

3. *se $(d, n) = 1$, nós temos que $c_d(n) = \mu(d)$, isto é,*

$$\mu(d) = \sum_{\substack{1\leq a\leq d \\ (a,d)=1}} e\left(\frac{na}{d}\right).$$

Demonstração. A Parte 1 é deixada como exercício para o leitor. A Parte 3 é uma consequência direta da Parte 2. Provamos então a parte 2.

Seja

$$\tilde{c}_d(n) := \sum_{1\leq a\leq d} e\left(\frac{na}{d}\right).$$

Por um lado nós podemos escrever

$$\tilde{c}_d(n) = e\left(\frac{n}{d}\right) \sum_{0\leq a\leq d-1} e\left(\frac{na}{d}\right),$$

e podemos ver que isto é $e(n/d)\frac{e(n)-1}{e(n/d)-1}$ se $d \nmid n$ e $e(n/d)d$ se $d \mid n$. Em outras palavras,

$$\tilde{c}_d(n) = \begin{cases} 0 & \text{se } d \nmid n, \\ d & \text{se } d \mid n. \end{cases} \tag{8.2.3}$$

Por outro lado, nós podemos escrever

$$\begin{aligned}\tilde{c}_d(n) &= \sum_{D|d} \sum_{\substack{1\le a\le d\\(a,d)=D}} e\left(\frac{na}{d}\right)\\ &= \sum_{D|d} \sum_{\substack{1\le a_1\le \frac{d}{D}\\ \left(a_1,\frac{d}{D}\right)=1}} e\left(\frac{nDa_1}{d}\right) = \sum_{D|d} c_{\frac{d}{D}}(n).\end{aligned}$$

Usando a fórmula de inversão de Möbius nós deduzimos que

$$c_d(n) = \sum_{D|d} \mu(D)\tilde{c}_{\frac{d}{D}}(n) = \sum_{D|d} \mu\left(\frac{d}{D}\right)\tilde{c}_D(n),$$

o qual, por (8.2.3), é

$$\sum_{D|(d,n)} \mu\left(\frac{d}{D}\right) D.$$

□

Demonstração do Teorema 8.2.1. Vamos primeiro estabelecer algumas notações. Seja $d = p_1 \dots p_t$ um inteiro positivo e livre de quadrados composto de primos que dividem $P(z)$. Pelo Teorema Chinês do Resto, para todo $\mathbf{i} = (i_1, \dots, i_t)$ com $1 \le i_1 \le \omega(p_1), \dots, 1 \le i_t \le \omega(p_t)$, existe um único inteiro $w_{\mathbf{i},d}$ tal que

$$w_{\mathbf{i},d} \equiv w_{i_j,p_j} (\text{mod } p_j)\ \forall 1 \le j \le t.$$

Nós denotamos por $\omega(d)$ o número de todos os possíveis $w_{\mathbf{i},d}$ aparecendo desta maneira (isto é, quando variamos $\mathbf{i}$, mas mantemos d fixado). Claramente, $\omega(d)$ é o produto dos $\omega(p_i)$'s.

Agora seja $n \in \mathcal{S}(\mathcal{A}, \mathcal{P}, z)$. Isto implica que

$$(n - w_{\mathbf{i},d}, d) = 1$$

para todo d e $\mathbf{i}$ como acima, e assim pela parte 3 da Proposição 8.2.2 nós obtemos

$$\mu(d) = c_d(n - w_{\mathbf{i},d}) = \sum_{\substack{1\le a\le d\\(a,d)=1}} e\left(\frac{(n-w_{\mathbf{i},d})a}{d}\right). \tag{8.2.4}$$

Nós efetuamos a soma (8.2.4) sobre todos os índices $\mathbf{i}$ correspondendo a d e sobre todos os inteiros $n \in \mathcal{S}(\mathcal{A},\mathcal{P},z)$ e obtemos

$$\mu(d)\omega(d)S(\mathcal{A},\mathcal{P},z) = \sum_{\substack{1\le a\le d\\(a,d)=1}} \sum_{w_{\mathbf{i},d}} e\left(\frac{-w_{\mathbf{i},d}a}{d}\right) \sum_{n\in\mathcal{S}(\mathcal{A},\mathcal{P},z)} e\left(\frac{na}{d}\right). \tag{8.2.5}$$

Elevando ao quadrado (8.2.5) e aplicando a desigualdade de Cauchy-Schwarz nos dá

$$|\mu(d)\omega(d)S(\mathcal{A},\mathcal{P},z)|^2 \le \left(\sum_{\substack{1\le a\le d\\(a,d)=1}} \left|\sum_{w_{\mathbf{i},d}} e\left(\frac{-w_{\mathbf{i},d}a}{d}\right)\right|^2\right) \times \left(\sum_{\substack{1\le a\le d\\(a,d)=1}} \left|\sum_{n\in\mathcal{S}(\mathcal{A},\mathcal{P},z)} e\left(\frac{na}{d}\right)\right|^2\right).$$

Nós escrevemos o primeiro fator da expressão acima como

$$\sum_{\substack{1\le a\le d\\(a,d)=1}} \sum_{w_{\mathbf{i},d},w_{\mathbf{i}',d}} e\left(\frac{(w_{\mathbf{i},d}-w_{\mathbf{i}',d})a}{d}\right) = \sum_{w_{\mathbf{i},d},w_{\mathbf{i}',d}} c_d(w_{\mathbf{i},d}-w_{\mathbf{i}',d}),$$

e usando parte 2 da Proposição 8.2.2, nós temos

$$\begin{aligned}
\sum_{w_{\mathbf{i},d},w_{\mathbf{i}',d}} \sum_{D|(d,w_{\mathbf{i},d}-w_{\mathbf{i}',d}} \mu\left(\frac{d}{D}\right) D &= \sum_{D|d} \sum_{\substack{w_{\mathbf{i},d},w_{\mathbf{i}',d} \\ w_{\mathbf{i},d}\equiv w_{\mathbf{i}',d} (\bmod\ D)}} \mu\left(\frac{d}{D}\right) D \\
&= \sum_{D|d} \mu\left(\frac{d}{D}\right) D\omega(d)\omega\left(\frac{d}{D}\right) \\
&= d\omega(d) \sum_{E|d} \frac{\mu(E)\omega(E)}{E} \\
&= d\omega(d) \prod_{p|d} \left(1 - \frac{\omega(p)}{p}\right) \\
&= \omega(d) \prod_{p|d} (p - \omega(p)).
\end{aligned}$$

Isto nos fornece que

$$\begin{aligned}
&|\mu(d)\omega(d)S(\mathcal{A},\mathcal{P},z)|^2 \\
&\qquad \leq \omega(d) \prod_{p|d} (p - \omega(p)) \left(\sum_{\substack{1\leq a\leq d \\ (a,d)=1}} \left| \sum_{n\in\mathcal{S}(\mathcal{A},\mathcal{P},z)} e\left(\frac{na}{d}\right) \right|^2 \right),
\end{aligned}$$

ou, equivalentemente, que

$$\mu^2(d)S(\mathcal{A},\mathcal{P},z)^2 \prod_{p|d} \frac{\omega(p)}{p - \omega(p)} \leq \sum_{\substack{1\leq a\leq d \\ (a,d)=1}} \left| \sum_{n\in\mathcal{S}(\mathcal{A},\mathcal{P},z)} e\left(\frac{na}{d}\right) \right|^2. \tag{8.2.6}$$

Agora nós efetuamos a soma (8.2.6) sobre $d \leq z$ e usamos o Teorema 8.1.2 com a sequência $(a_n)_{n\geq 1}$ escolhida de tal maneira que a_n é 1 se $n \in \mathcal{S}(\mathcal{A},\mathcal{P},z)$ e 0 caso contrário. Nós obtemos

$$\sum_{d\leq z} \mu^2(d)S(\mathcal{A},\mathcal{P},z)^2 \prod_{p|d} \frac{\omega(p)}{p - \omega(p)} \leq (z^2 + 4\pi x)S(\mathcal{A},\mathcal{P},z),$$

a qual, depois dos cancelamentos óbvios, completa a demonstração do teorema.

□

8.3 Exercícios

1. (Identidade de Parseval) Seja $(a_n)_{n\geq 1}$ números complexos e x um inteiro positivo. Mostre que

$$\int_0^1 \left|\sum_{n\leq x} a_n e(n\alpha)\right|^2 d\alpha = \sum_{n\leq x} |a_n|^2,$$

onde $e(n\alpha) := \exp(2\pi i n\alpha)$.

2. Sejam d, d', n inteiros positivos. Mostre que , se $(d, d') = 1$, então

$$c_{dd'}(n) = c_d(n)c_{d'}(n),$$

onde $c_d(n)$ é definido em (8.2.2).

3. (A Desigualdade de Pólya–Vinogradov) Seja χ um caractere não–trivial módulo d e sejam M, N inteiros positivos. Mostre que

$$\sum_{M+1\leq n\leq M+N} \chi(n) \ll d^{1/2}\log d.$$

Capítulo 9

O Teorema de Bombieri–Vinogradov

O teorema de Bombieri–Vinogradov é considerado um dos resultados mais elegantes derivado do método do grande crivo. A principal razão está no fato de que nas muitas questões que requerem o uso da hipótese generalizada de Riemann para L–funções de Dirichlet, nós podemos substituí-la pelo teorema de Bombieri–Vinogradov. Assim, em situações envolvendo 'suficientemente várias' L–funções, o teorema pode ser reconhecido como um substituto para a hipótese generalizada de Riemann. Como veremos abaixo, uma aplicação é no problema do divisor de Titchmarsh.

Em 1976, Motohashi [17] descobriu um príncipio de indução que generaliza o teorema de Bombieri–Vinogradov e faz com que ele seja aplicável em um contexto mais amplo para funções aritméticas mais gerais. Esta perspectiva foi útil no trabalho de Bombieri *et al.* [3, 4], onde o intervalo da somatória no clássico teorema de Bombieri–Vinogradov é estendido além da barreira $\sqrt{x}$.

Neste capítulo nós iremos derivar o clássico teorema de Bombieri–Vinogradov usando um método de Vaughan, como descrito em [7]. Nós mostraremos que o método tem um amplo domínio de aplicabilidade colocando ele em um contexto mais geral. Nós então aplicaremos o resultado para abordamos o problema de Titchmarsh sobre

divisores.

9.1 Um Teorema Geral

O principal resultado desta seção é derivado como uma generalização de um método de Vaughan, o qual está no cerne de uma das demonstrações do teorema de Bombieri–Vinogradov. Antes de enunciarmos o resultado geral, vamos introduzir primeiro a classe de funções

$$\mathcal{D} := \left\{ D : \mathbb{N} \to \mathbb{C} : \sum_{n \le x} |D(n)|^2 = O(x(\log x)^\alpha) \text{ para algum } \alpha > 0 \right\}. \tag{9.1.1}$$

Nós temos as seguintes propriedades básicas:

Proposição 9.1.1. *1. Se $D \in \mathcal{D}$ e $\theta \ge 0$, então*

$$\sum_{n \le x} \frac{|D(n)|}{n^\theta} \ll x^{1-\theta}(\log x)^\alpha$$

para algum $\alpha > 0$.

2. Se $D_1, D_2 \in \mathcal{D}$, então

$$\sum_{ef \le x} |D_1(e)D_2(f)|^2 d(ef) \ll x(\log x)^\gamma$$

para algum $\gamma > 0$, onde para um inteiro positivo e, $d(e)$ denota o número de divisores de e.

Demonstração. Exercício. □

Teorema 9.1.2. *Seja x, z inteiros positivos e denotemos*

$$A(s) = \sum_{n \ge 1} \frac{a(n)}{n^s}, \quad B(s) = \sum_{n \ge 1} \frac{b(n)}{n^s}$$

as séries de Dirichlet (isto é, $a(1) = b(1) = 1$) para o qual nós escrevemos

$$\frac{A(s)}{B(s)} = \sum_{n\geq 1} \frac{c(n)}{n^s}, \qquad \frac{1}{B(s)} = \sum_{n\geq 1} \frac{\tilde{b}(n)}{n^s}$$

para algum $c(n), \tilde{b}(n) \in \mathbb{C}$. *Nós assumimos que todas estas séries são convergentes para* $\mathrm{Re}(s) > \sigma_0$, *para algum* σ_0, *e que elas satisfazem as seguintes hipóteses:*

(H1) $(a(n))_{n\geq 1}$ *é uma sequência crescente de números reais positivos;*

(H2) $b(\cdot), \tilde{b}(\cdot), c(\cdot) \in \mathcal{D}$;

(H3) existe $0 \leq \theta < 1$ *e* $0 \leq \gamma < 1$ *tal que, para todo caractere de Dirichlet não–trivial* χ *módulo* d,

$$\sum_{n\leq x} b(n)\chi(n) \ll x^{\theta}\sqrt{d}\log d + x^{\gamma}.$$

Então

1. *se* $z \leq x^{\frac{1-\theta}{3-\theta}}$,

$$\sum_{d\leq z} \frac{d}{\phi(d)} \sideset{}{^*}\sum_{\chi} \max_{y\leq x} \left| \sum_{n\leq y} c(n)\chi(n) \right|$$

$$\ll \left(z^2 x^{\frac{1}{2}} + x + z x^{\frac{5-\theta}{2(3-\theta)}} + z^2 x^{\frac{1-\theta+2\gamma}{3-\theta}} + z^{\frac{5}{2}} x^{\frac{1+\theta}{3-\theta}} a(x) \right) (\log x)^{\alpha'} \tag{9.1.2}$$

para algum $\alpha' > 0$;

2. se $z > x^{\frac{1-\theta}{3-\theta}}$,

$$\sum_{d\le z}\frac{d}{\phi(d)}\sum_{\chi}^{*}\max_{y\le x}\left|\sum_{n\le y}c(n)\chi(n)\right|$$

$$\ll\left(z^2x^{\frac{1}{2}}+x+z^{\frac{9-4\theta}{2(3-2\theta)}}x^{\frac{2-\theta}{3-2\theta}}a(x)(\log z)\right.$$
$$\left.+\,z^{\frac{3-4\theta+3\gamma}{3-2\theta}}x^{\frac{2-2\theta+\gamma}{3-2\theta}}(\log z)\right)(\log\alpha'') \quad (9.1.3)$$

para algum $\alpha'' > 0$.

Aqui, a soma $\sum_{\chi}^{*}$ *é sobre caracteres primitivos de Dirichlet* χ *módulo* d.

Demonstração. Nós denotamos

$$F(s) := \sum_{n\le U}\frac{c(n)}{n^s}, \qquad G(s) := \sum_{n\le V}\frac{\tilde{b}(n)}{n^s}$$

para algum parâmetro $U = U(x,z)$ e $V = V(x,z)$ a ser escolhido mais tarde. Nós podemos pensar F como uma aproximação para $A(s)/B(s)$ e $G(s)$ como uma aproximação para $1/G(s)$, e observamos que podemos escrever

$$\frac{A(s)}{B(s)} = F(s) - B(s)G(s)F(s) + A(s)G(s)$$
$$+\left(\frac{A(s)}{B(s)} - F(s)\right)(1 - B(s)G(s)). \quad (9.1.4)$$

Na literatura, isto é conhecido como **identidade de Vaughan** e tem seus primórdios devido a idéias de Linnik. Comparando os coeficientes de n^{-s} em ambos os lados de (9.1.4), nós deduzimos que

$$c(n) = a_1(n) + a_2(n) + a_3(n) + a_4(n),$$

onde

$$a_1(n) := \begin{cases} c(n) & \text{se } n \le U \\ 0 & \text{se } n > U \end{cases} \tag{9.1.5}$$

$$a_2(n) := -\sum_{\substack{efg=n \\ f\le V \\ g\le U}} b(e)\tilde{b}(f)c(g), \tag{9.1.6}$$

$$a_3(n) := \sum_{\substack{ef=n \\ f\le V}} a(e)\tilde{b}(f), \tag{9.1.7}$$

$$a_4(n) := -\sum_{\substack{ef=n \\ e>U \\ f>V}} c(e) \sum_{\substack{gh=f \\ h\le V}} b(g)\tilde{b}(h). \tag{9.1.8}$$

Portanto, para qualquer caractere de Dirichlet χ módulo d nós podemos escrever

$$\sum_{n\le y} c(n)\chi(n) = \sum_{1\le i\le 4} \sum_{n\le y} a_i(n)\chi(n) := \sum_{1\le i\le 4} S_i(y,\chi). \tag{9.1.9}$$

Nós provaremos o teorema estimando cada uma das somas

$$\sum_{d\le z} \frac{d}{\phi(d)} \sum_{\chi}^{*} \max_{y\le x} |S_i(y,\chi)|, \quad 1 \le i \le 4.$$

A estimativa para $S_1(y,\chi)$: Usando (9.1.5) e hipótese (H2) junto com a parte 1 da Proposição 9.1.1, nós obtemos

$$|S_1(y,\chi)| = \left| \sum_{\substack{n\le y \\ n\le U}} c(n)\chi(n) \right| \ll \sum_{n\le U} |c(n)| \ll U(\log U)^{\alpha_0}$$

para algum $\alpha_0 > 0$. O limitante acima é independente de y, χ e d, logo, relembrando que existem $\phi(d)$ caracteres de Dirichlet módulo d, nós obtemos

$$\sum_{d\leq z}\frac{d}{\phi(d)}\sum_{\chi}^{*}\max_{y\leq x}|S_1(y,\chi)|\ll z^2U(\log U)^{\alpha_0}. \tag{9.1.10}$$

A estimativa para $S_2(y,\chi)$: Usando (9.1.6) nós escrevemos

$$S_2(y,\chi)=-\sum_{\substack{efg\leq y\\ f\leq V\\ g\leq U}} b(e)\tilde{b}(f)c(g)\chi(efg),$$

o qual dividimos em duas partes de acordo com $fg\leq U$ ou $U<fg\leq UV$. A primeira soma obtida desta maneira ira ser denotada por $S_2^{'}(y,\chi)$ e a segunda soma por $S_2^{''}(y,\chi)$.

Para $S_2^{'}(y,\chi)$ nós escrevemos

$$|S_2^{'}(y,\chi)|\leq\sum_{g\leq U}|c(g)|\sum_{f\leq\min\left(V,\frac{U}{g}\right)}|\tilde{b}(f)|\left|\sum_{e\leq\frac{y}{fg}}b(e)\chi(e)\right|$$

assim podemos usar a hipótese (H3) para estimar a soma mais interna. Nós obtemos

$$\begin{aligned}
S_2^{'}(y,\chi) &\ll y^{\theta}\sqrt{d}(\log d)\sum_{g\leq U}\frac{|c(g)|}{g^{\theta}}\sum_{f\leq\min\left(V,\frac{U}{g}\right)}\frac{|\tilde{b}(f)|}{f^{\theta}}\\
&+ y^{\gamma}\sum_{g\leq U}\frac{|c(g)|}{g^{\gamma}}\sum_{f\leq\min\left(V,\frac{U}{g}\right)}\frac{|\tilde{b}(f)|}{f^{\gamma}}\\
&\ll y^{\theta}\sqrt{d}(\log d)\sum_{g\leq U}\frac{|c(g)|}{g^{\theta}}\sum_{f\leq\frac{U}{g}}\frac{|\tilde{b}(f)|}{f^{\theta}}\\
&+ y^{\gamma}\sum_{g\leq U}\frac{|c(g)|}{g^{\gamma}}\sum_{f\leq\frac{U}{g}}\frac{|\tilde{b}(f)|}{f^{\gamma}}.
\end{aligned}$$

Então, usando (H2) e parte 1 da Proposição 9.1.1, nós obtemos

$$\begin{aligned} S_2^{'}(y,\chi) &\ll y^{\theta}\sqrt{d}(\log d)\sum_{g\le U}\frac{|c(g)|}{g^{\theta}}\left(\frac{U}{g}\right)^{1-\theta}(\log U)^{\alpha_1} \\ &+ y^{\gamma}\sum_{g\le U}\frac{|c(g)|}{g^{\gamma}}\left(\frac{U}{g}\right)^{1-\gamma}(\log U)^{\alpha_2} \\ &\ll y^{\theta}\sqrt{d}(\log d)U^{1-\theta}(\log U)^{\alpha_3}+y^{\gamma}U^{1-\gamma}(\log U)^{\alpha_4} \end{aligned}$$

para alguns $\alpha_1,\alpha_2,\alpha_3,\alpha_4>0$. Isto implica que

$$\begin{aligned} &\sum_{d\le z}\frac{d}{\phi(d)}\sideset{}{^*}\sum_{\chi}\max_{y\le x}|S_2^{'}(y,\chi)| \\ &\ll x^{\theta}z^{\frac{5}{2}}(\log z)U^{1-\theta}(\log U)^{\alpha_3}+x^{\gamma}z^2(\log z)U^{1-\gamma}(\log U)^{\alpha_4}. \quad (9.1.11) \end{aligned}$$

Para $S_2^{''}(y,\chi)$ nós escrevemos

$$\begin{aligned} &\sum_{d\le z}\frac{d}{\phi(d)}\sideset{}{^*}\sum_{\chi}\max_{y\le x}|S_2^{''}(y,\chi)| \\ &=\sum_{d\le z}\frac{d}{\phi(d)}\sideset{}{^*}\sum_{\chi}\max_{y\le x}\left|\sum_{\substack{eh\le y\\ U<h\le UV}}b(e)\left(\sum_{\substack{f\le V,g\le U\\ fg=h}}\tilde{b}(f)c(g)\right)\chi(eh)\right|, \end{aligned}$$

o qual sugere usarmos a segunda desigualdade modificada do grande crivo que foi discutida no Capítulo 8. Entretanto, nós observamos que aplicando esta desigualdade diretamente ao par das sequências de números complexos

$$(b(e))_{\frac{x}{UV}<e\le\frac{x}{U}} \qquad \text{e} \qquad (\tilde{b}(f)c(g))_{fg=h,U<h\le UV},$$

nós iremos superestimar nossa expressão. Ao invés, nós devemos dividir o intervalo $(U,UV]$ em intervalos diádicos $(2^k,2^{k+1}]$ com $[\log_2 U]<$

$k < [\log_2 UV]$ e aplicar a desigualdade modificada do grande crivo para cada par de sequências

$$(b(e))_{\frac{x}{2^{k+1}}<e\le\frac{x}{2^k}} \quad \text{e} \quad (\tilde{b}(f)c(g))_{fg=h,2^k<h\le 2^{k+1}}.$$

Para cada $[\log_2 U] < k < [\log_2 UV]$ nós obtemos que

$$= \sum_{d\le z}\frac{d}{\phi(d)}\sum_{\chi}^{*}\max_{y\le x}\left|\sum_{\substack{eh\le y\\ 2^k<h\le 2^{k+1}}} b(e)\left(\sum_{\substack{f\le V,g\le U\\ fg=h}}\tilde{b}(f)c(g)\right)\chi(eh)\right|$$

$$\ll \left(z^2+\frac{x}{2^k}\right)^{1/2}(z^2+2^k)^{1/2}\left(\sum_{e}{}'\right)^{1/2}\left(\sum_{h}{}'\right)^{1/2}(\log x), \tag{9.1.12}$$

onde

$$\sum_{e}{}' := \sum_{e\le\frac{x}{2^k}}|b(e)|^2, \tag{9.1.13}$$

$$\sum_{h}{}' := \sum_{2^k<h\le 2^{k+1}}\left|\sum_{\substack{f\le V,g\le U\\ fg=h}}\tilde{b}(f)c(g)\right|^2. \tag{9.1.14}$$

Por (H2) nós temos

$$\sum_{e\le\frac{x}{2^k}}|b(e)|^2 \ll \frac{x}{2^k}\left(\log\frac{x}{2^k}\right)^{\alpha_5} \tag{9.1.15}$$

para algum $\alpha_5 > 0$, e pela desigualdade de Cauchy–Schwarz juntamente com (H2) e parte 2 da Proposição 9.1.1, nós obtemos

$$\sum_{2^k<h\le 2^{k+1}}\left|\sum_{\substack{f\le V,g\le U\\ fg=h}}\tilde{b}(f)c(g)\right|^2 \ll \sum_{2^k<h\le 2^{k+1}}d(h)\sum_{f|h}|\tilde{b}(f)|^2\left|c\left(\frac{h}{f}\right)\right|^2$$

$$\ll 2^k(\log 2^k)^{\alpha_6} \tag{9.1.16}$$

para algum $\alpha_6 > 0$. Então nós inserimos (9.1.15) e (9.1.16) em (9.1.13), (9.1.14), (9.1.12), para obtermos

$$\sum_{d\le z}\frac{d}{\phi(d)}\sideset{}{^*}\sum_{\chi}\max_{y\le x}\left|\sum_{\substack{eh\le y\\ 2^k<h\le 2^{k+1}}} b(e)\left(\sum_{\substack{f\le V, g\le U\\ fg=h}}\tilde{b}(f)c(g)\right)\chi(eh)\right|$$

$$\ll \left(z^2+\frac{zx^{1/2}}{2^{k/2}}+z2^{k/2}+x^{1/2}\right)x^{\frac{1}{2}}(\log x)^{\frac{\alpha_5}{2}+1}(\log 2^k)^{\frac{\alpha_6}{2}}.$$

Nós efetuamos a soma sobre todos os k para finalmente obtermos

$$\sum_{d\le z}\frac{d}{\phi(d)}\sideset{}{^*}\sum_{\chi}\max_{y\le x}|S_2^{''}(y,\chi)|$$

$$\ll \left(z^2+\frac{zx^{1/2}}{U^{1/2}}+z(UV)^{1/2}+x^{1/2}\right)x^{1/2}(\log x)^{\frac{\alpha_5}{2}+1}(\log UV)^{\frac{\alpha_6}{2}+1}. \tag{9.1.17}$$

Para estimar $S_3(y,\chi)$: Nós definimos uma função escada $\mathcal{A}:\mathbb{R}\to\mathbb{R}$ por $\mathcal{A}(t)=a(1)$ se $t\le 1$, $\mathcal{A}(t)=a(2)-a(1)$ se $1<t\le 2$, e em geral, $\mathcal{A}(t)=a(n)-a(n-1)$ se $n-1<t\le n$. Então nós observamos que $a(n)=\int_0^n \mathcal{A}(t)dt$ e que $\mathcal{A}(\cdot)$ é positiva, uma vez que a sequência $(a(n))_{n\ge 1}$ é crescente. Usando (9.1.7) (e assumindo, sem perda de generalidade, que y é um inteiro) nós escrevemos

$$\begin{aligned}
|S_3(y,\chi)| &= \left|\sum_{f\le V}\tilde{b}(f)\chi(f)\sum_{e\le \frac{y}{f}}a(e)\chi(e)\right| \\
&= \left|\sum_{f\le V}\tilde{b}(f)\chi(f)\sum_{e\le \frac{y}{f}}\chi(e)\int_0^e \mathcal{A}(t)dt\right| \\
&= \left|\int_0^y \mathcal{A}(t)\sum_{f\le V}\tilde{b}(f)\chi(f)\sum_{t\le e\le \frac{y}{f}}\chi(e)dt\right| \\
&\ll \int_0^y |\mathcal{A}(t)|\sum_{f\le V}|\tilde{b}(f)|\left|\sum_{t\le e\le \frac{y}{f}}\chi(e)\right|dt.
\end{aligned}$$

Observamos que podemos usar a desigualdade de Pólya–Vinogradov dada no último Exercício do Capítulo 8 para estimar a soma interna. Nós obtemos

$$S_3(y,\chi)\ll \sqrt{d}(\log d)\int_0^y |\mathcal{A}(t)|dt\sum_{f\le V}|\tilde{b}(f)|\ll \sqrt{d}(\log d)V(\log V)^{\alpha_7}a(y)$$

para algum $\alpha_7>0$. Portanto, observando que $|\mathcal{A}(t)|=\mathcal{A}(t)$ e usando (H1), nós obtemos

$$\begin{aligned}
\sum_{d\le z}\frac{d}{\phi(d)}\sideset{}{^*}\sum_{\chi}\max_{y\le x}|S_3(y,\chi)| &\ll z^{\frac{5}{2}}(\log z)V(\log V)^{\alpha_7}\max_{y\le x}|a(y)| \\
&= z^{\frac{5}{2}}(\log z)V(\log V)^{\alpha_7}a(x). \quad (9.1.18)
\end{aligned}$$

Para estimar $S_4(y,\chi)$: Usando (9.1.8) nós escrevemos

$$\sum_{d\le z}\frac{d}{\phi(d)}\sideset{}{^*}\sum_{\chi}\max_{y\le x}|S_4(y,\chi)|$$

$$=\sum_{d\le z}\frac{d}{\phi(d)}\sideset{}{^*}\sum_{\chi}\max_{y\le x}\left|\sum_{\substack{ef\le y\\ e>U\\ f>V}}c(e)\left(\sum_{\substack{gh=f\\ h\le V}}b(g)\tilde{b}(h)\right)\chi(ef)\right|,$$

o qual nos sugere usar a segunda desigualdade modificada do grande crivo ao par de sequências de números complexos

$$(c(e))_{U<e<\frac{x}{V}}\qquad\text{e}\qquad(b(g)\tilde{b}(h))_{gh=f,h\le V,V<f<\frac{x}{U}}.$$

Entretanto, para obtermos resultados ótimos, nós procedemos da mesma maneira em que lidamos com o caso para a soma $S_2''(y,\chi)$. Para ser preciso, nós dividimos o intervalo $(U,y/V]$ e intervalos diádicos $(2^k,2^{k+1}]$ com $[\log_2 U]<k\le[\log_2 y/V]$ e aplicamos a desigualdade modificada do grande crivo para cada uma das sequências abaixo

$$(c(e))_{2^k<e\le 2^{k+1}}\qquad\text{e}\qquad(b(g)\tilde{b}(h))_{gh=f,h\le V,\max(V,\frac{y}{2^{k+1}})<f<\frac{y}{2^k}}.$$

Para cada $[\log_2 U]<k\le[\log_2\frac{x}{V}]$ nós obtemos

$$\sum_{d\le z}\frac{d}{\phi(d)}\sideset{}{^*}\sum_{\chi}\max_{y\le x}\left|\sum_{\substack{ef\le y\\ e>U\\ f>V\\ 2^k<e\le 2^{k+1}}}c(e)\left(\sum_{\substack{gh=f\\ h\le V}}b(g)\tilde{b}(h)\right)\chi(ef)\right|$$

$$\ll(z^2+2^k)^{1/2}\left(z^2+\frac{x}{2^k}\right)^{1/2}\left(\sum_e{}''\right)^{1/2}\left(\sum_f{}''\right)^{1/2}\log x,\tag{9.1.19}$$

onde

$$\sum_{e}^{''} := \sum_{2^k < e \le 2^{k+1}} |c(e)|^2, \tag{9.1.20}$$

$$\sum_{f}^{''} := \sum_{V < f \le \frac{x}{2^k}} \left| \sum_{\substack{gh=f \\ h \le V}} b(g)\tilde{b}(h) \right|^2. \tag{9.1.21}$$

Nós observamos que por (H2) nós temos

$$\sum_{2^k < e \le 2^{k+1}} |c(e)|^2 \ll 2^k (\log 2^k)^{\alpha_8} \tag{9.1.22}$$

para algum $\alpha_8 > 0$, e pela desigualdade de Cauchy–Schwarz juntamente com (H2) e parte 2 da Proposição 9.1.1, nós obtemos

$$\begin{aligned} \sum_{V < f \le \frac{x}{2^k}} \left| \sum_{\substack{gh=f \\ h \le V}} b(g)\tilde{b}(h) \right|^2 &\ll \sum_{V < f \le \frac{x}{2^k}} d(f) \sum_{h \mid f} |\tilde{b}(h)|^2 \left| f\left(\frac{f}{h}\right) \right|^2 \\ &\ll \frac{x}{2^k} \left(\log \frac{x}{2^k} \right)^{\alpha_9} \end{aligned} \tag{9.1.23}$$

para algum $\alpha_9 > 0$. Nós inserimos (9.1.22) e (9.1.23) em (9.1.20), (9.1.21), (9.1.19), e obtemos

$$\begin{aligned} &\sum_{d \le z} \frac{d}{\phi(d)} \sum_{\chi}^{*} \max_{y \le x} \left| \sum_{\substack{ef \le y \\ e > U \\ f > V \\ 2^k < e \le 2^{k+1}}} c(e) \left(\sum_{\substack{gh=f \\ h \le V}} b(g)\tilde{b}(h) \right) \chi(ef) \right| \\ &\qquad \ll \left(z^2 + \frac{z x^{1/2}}{2^{k/2}} + z 2^{k/2} x^{1/2} \right) x^{1/2} (\log x)^{\frac{\alpha_9}{2}+1} (\log 2^k)^{\frac{\alpha_8}{2}}. \end{aligned}$$

Finalmente nós efetuamos a some sobre todos os k e deduzimos que

$$\sum_{d\le z}\frac{d}{\phi(d)}\sum_{\chi}^{*}\max_{y\le x}|S_4(y,\chi)| \ll \left(z^2+\frac{zx^{1/2}}{U^{1/2}}+\frac{zx^{1/2}}{V^{1/2}}+x^{1/2}\right)x^{1/2}(\log x)^{\frac{\alpha_8+\alpha_9}{2}+2}. \quad (9.1.24)$$

Nós acabamos de estimar as quatro somas $S_i(y,\chi)$, $1\le i\le 4$ (veja (9.1.10), (9.1.11), (9.1.17), (9.1.18) e (9.1.24)). Colocando todas estas estimativas juntas nós obtemos

$$\begin{aligned}
&\sum_{d\le z}\frac{d}{\phi(d)}\sum_{\chi}^{*}\max_{y\le x}\left|\sum_{n\le y}c(n)\chi(n)\right| \\
\ll\ & z^2U(\log U)^{\alpha_0} \\
+\ & x^\theta z^{5/2}(\log z)U^{1-\theta}(\log U)^{\alpha_3}+x^\gamma z^2(\log z)U^{1-\gamma}(\log U)^{\alpha_4} \\
+\ & \left(z^2+\frac{zx^{1/2}}{U^{1/2}}+z(UV)^{1/2}+x^{1/2}\right)x^{\frac12}(\log x)^{\frac{\alpha_5}{2}+1}(\log UV)^{\frac{\alpha_6}{2}+1} \\
+\ & z^{5/2}(\log z)V(\log V)^{\alpha_7}a(x) \\
+\ & \left(z^2+\frac{zx^{1/2}}{U^{1/2}}+\frac{zx^{1/2}}{V^{1/2}}+x^{1/2}\right)x^{\frac12}(\log x)^{\frac{\alpha_8+\alpha_9}{2}+2}. \qquad (9.1.25)
\end{aligned}$$

Ainda nos resta escolher apropriadamente os parâmetros U e V. Nós queremos U e V tal que $z^{5/2}x^\theta U^{1-\theta}=z^{5/2}V$, isto é, tal que

$$V=x^\theta U^{1-\theta}.$$

Agora nós analisamos a expressão

$$\begin{aligned}
E(x,z,U): &= \frac{zx}{U^{1/2}}+z^{5/2}x^\theta U^{1-\theta}+zx^{\frac12}(UV)^{1/2} \\
&= \frac{zx}{U^{1/2}}+zU\left(z^{\frac32}x^\theta U^{-\theta}+x^{\frac{1+\theta}{2}}U^{-\frac{\theta}{2}}\right).
\end{aligned}$$

Se

$$z^{\frac{3}{2}}x^{\theta}U^{-\theta} \leq x^{\frac{1+\theta}{2}}U^{-\frac{\theta}{2}},$$

isto é, se

$$z \leq x^{\frac{1-\theta}{3}}U^{\frac{\theta}{3}}, \tag{9.1.26}$$

então

$$E(x,z,U) \ll \frac{zx}{U^{1/2}} + zx^{\frac{1+\theta}{2}}U^{\frac{2-\theta}{2}}.$$

Escolhemos U tal que

$$\frac{zx}{U^{1/2}} = zx^{\frac{1+\theta}{2}}U^{\frac{2-\theta}{2}},$$

isto é,

$$U := x^{\frac{1-\theta}{3-\theta}}.$$

Voltando na equação (9.1.26), nós observamos que nossa escolha de U implica que

$$z \leq x^{\frac{1-\theta}{3-\theta}}.$$

Nós obtemos

$$\sum_{d\leq z}\frac{d}{\phi(d)}\sideset{}{^*}\sum_{\chi}\max_{y\leq x}\left|\sum_{n\leq y}c(n)\chi(n)\right|$$
$$\ll \left(z^2x^{\frac{1}{2}} + x + zx^{\frac{5-\theta}{2(3-\theta)}} + z^2x^{\frac{1-\theta+2\gamma}{3-\theta}} + z^{\frac{5}{2}}x^{\frac{1+\theta}{3-\theta}}a(x)\right)(\log x)^{\alpha'}$$

para algum $\alpha' > 0$.

Se

$$z^{\frac{3}{2}}x^{\theta}U^{-\theta} > x^{\frac{1+\theta}{2}}U^{-\frac{\theta}{2}},$$

então procedendo como fizemos acima e escolhendo U tal que

$$\frac{zx}{U^{1/2}} = z^{5/2}x^{\theta}U^{1-\theta}.$$

Assim

$$U := \frac{x^{\frac{2(1-\theta)}{3-2\theta}}}{z^{\frac{3}{3-2\theta}}}.$$

Com esta escolha de U nós obtemos que

$$z > x^{\frac{1-\theta}{3-\theta}},$$

e que

$$\sum_{d\le z}\frac{d}{\phi(d)}\sum_{\chi}^{*}\max_{y\le x}\left|\sum_{n\le y}c(n)\chi(n)\right|$$
$$\ll \left(z^2x^{\frac{1}{2}} + x + z^{\frac{9-4\theta}{2(3-2\theta)}}x^{\frac{2-\theta}{3-2\theta}}a(x)(\log z) + z^{\frac{3-4\theta+3\gamma}{3-2\theta}}x^{\frac{2-2\theta+\gamma}{3-2\theta}}(\log z)\right)(\log x)^{\alpha''}$$

para algum $\alpha'' > 0$.

Isto completa a demonstração do teorema. □

9.2 O teorema de Bombieri–Vinogradov

Nesta seção nós voltamos ao estudo do termo do erro que ocorre na fórmula assintótica para $\pi(x;d,a)$, onde a, d são primos entre si e $d \le x$. Nós já mencionamos no Capítulo 7 que se assumirmos a hipótese generalizada de Riemann, então

$$\pi(x;d,a) = \frac{\mathrm{li}(x)}{\phi(d)} + O(x^{1/2}(\log dx)).$$

Nós iremos agora provar que, em 'média' e sem usar nenhuma hipótese, o termo do erro de fato tem o tamanho $O_d(x^{1/2}\log x)$. Mais precisamente, nós iremos provar:

Teorema 9.2.1 (O teorema de Bombieri–Vinogradov)**.** *Para todo $A > 0$ existe $B = B(A) > 0$ tal que*

$$\sum_{d \le \frac{x^{1/2}}{(\log x)^B}} \max_{y \le x} \max_{(a,d)=1} \left| \pi(x; d, a) - \frac{\mathrm{li}(y)}{\phi(d)} \right| \ll \frac{x}{(\log x)^A}. \qquad (9.2.1)$$

O teorema de Bombieri–Vinogradov foi primeiramente obtido, independentemente, por E. Bombieri e A.I. Vinogradov em 1965. A demonstração deles é baseada em uma aplicação do grande crivo para estimar o número de zeros de certas L–funções em vários retângulos. Uma diferente demonstração do teorema foi obtida por Gallagher em 1968. E ainda outra demonstração foi obtida por Vaughan em 1975, e é o seu tratamento que iremos seguir abaixo (veja [7, pg. 16]).

Existem dois ingredientes principais necessários da demonstração de Vaughan do teorema de Bombieri–Vinogradov. O primeiro é um caso particular do teorema geral discutido na Seção 9.1. Nós enunciamos abaixo. O segundo é um resultado devido a Siegel e Walfisz, enunciado na Seção 7.3 do Capítulo 7, o qual nos fornece estimativas incondicionais para $\pi(x; d, a) - \mathrm{li}(x)/\phi(d)$ desde que d esteja em intervalos pequenos com respeito a x.

Teorema 9.2.2 (Vaughan)**.** *Sejam x e z inteiros positivos arbitrários. Então*

$$\sum_{d \le z} \frac{d}{\phi(d)} \sideset{}{^*}\sum_{\chi} \max_{y \le x} \left| \sum_{n \le y} \Lambda(n)\chi(n) \right| \ll (z^2 x^{1/2} + x + z x^{5/6})(\log z)(\log x)^{\alpha} \qquad (9.2.2)$$

para algum $\alpha > 0$, onde a soma $\sum_{\chi}^{*}$ é sobre caracteres primitivos de Dirichlet χ módulo d e onde $\Lambda(\cdot)$ denota a função de von Mangoldt.

Demonstração. Nós usamos a notação introduzida no Teorema 9.1.2 e denotamos

$$A(s) := -\zeta'(s) = \sum_{n \ge 1} \frac{\log n}{n^s}, \quad B(s) := \zeta(s) = \sum_{n \ge 1} \frac{1}{n^s}.$$

Assim, para $n \geq 1$,

$$a(n) = \log n,\ b(n) = 1,\ c(n) = \Lambda(n),\ \tilde{b}(n) = \mu(n),$$

e as hipóteses (H1) e (H2) do Teorema 9.1.2 estão satisfeitas. Usando a desigualdade de Pólya–Vinogradov nós vemos que, para todo d e qualquer caractere de Dirichlet não–trivial χ módulo d,

$$\sum_{n\leq x} b(n)\chi(n) = \sum_{n\leq x} \chi(n) \ll \sqrt{d}\log d.$$

Logo a hipótese (H3) é também satisfeita, com $\theta = 0 = \gamma$. Nós obtemos que se $z \leq x^{1/3}$, então

$$\sum_{d\leq z} \frac{d}{\phi(d)} \sideset{}{^*}\sum_{\chi} \max_{y\leq x} \left| \sum_{n\leq y} \Lambda(n)\chi(n) \right| \ll (z^2x^{1/2} + x + zx^{5/6} + z^{5/2}x^{1/3})(\log z)(\log x)^{\alpha'}$$

para algum $\alpha' > 0$, e se $z > x^{1/3}$, então

$$\sum_{d\leq z} \frac{d}{\phi(d)} \sideset{}{^*}\sum_{\chi} \max_{y\leq x} \left| \sum_{n\leq y} \Lambda(n)\chi(n) \right| \ll (z^2x^{1/2} + x + z^{3/2}x^{2/3})(\log z)(\log x)^{\alpha''}$$

para algum $\alpha'' > 0$. Combinando as duas estimativas nós obtemos a desigualdade desejada. □

Corolário 9.2.3. *Sejam x, z, D inteiros positivos tal que $z > D$. Então*

$$\sum_{d\leq z} \frac{d}{\phi(d)} \sideset{}{^*}\sum_{\chi} \max_{y\leq x} \left| \sum_{n\leq y} \Lambda(n)\chi(n) \right| \ll \left(zx^{1/2} + \frac{x}{z} + \frac{x}{D} + x^{5/6}\log z \right) (\log z)(\log x)^{\alpha} \quad (9.2.3)$$

para algum $\alpha > 0$.

Demonstração. Exercício. □

Demonstração do Teorema de Bombieri–Vinogradov. Primeiro observarmos que (9.2.1) é equivalente a provar que para todo $A > 0$ existe $B = B(A) > 0$ tal que

$$\sum_{d \le \frac{x^{1/2}}{(\log x)^B}} \max_{y \le x} \max_{(a,d)=1} \left| \psi(y; d, a) - \frac{y}{\phi(d)} \right| \ll \frac{x}{(\log x)^A},$$

onde

$$\psi(y; d, a) := \sum_{\substack{n \le y \\ (a,n)=1}} \Lambda(n).$$

Sejam A, y, x números reais positivos tal que $y \le x$, e seja $d \le y$. Observamos que

$$\max_{(a,d)=1} \left| \psi(y; d, a) - \frac{y}{\phi(d)} \right| \le \frac{1}{\phi(d)} \sum_{\substack{\chi (\bmod d) \\ \chi \ne \chi_0}} |\psi(y, \chi)| + \frac{\psi(y, \chi_0) - y}{\phi(d)}. \tag{9.2.4}$$

Seja o caractere $\chi \ne \chi_0$ módulo d induzido por algum caractere primitivo χ_1 módulo d_1. Assim

$$\psi(y, \chi_1) - \psi(y, \chi) \ll (\log y)(\log d)$$

(prove isto!). Usando isto em (9.2.4) nós temos

$$\max_{(a,d)=1} \left| \psi(y; d, a) - \frac{y}{\phi(d)} \right| \le \frac{1}{\phi(d)} \sum_{\substack{\chi (\bmod d) \\ \chi \ne \chi_0}} |\psi(y, \chi_1)| + \sum_{d \le z} \frac{1}{\phi(d)} \max_{y \le x} |\psi(y) - y| + (\log y)(\log d),$$

e, mais ainda, que

$$\sum_{d\le z}\max_{y\le x}\max_{(a,d)=1}\left|\psi(y;d,a)-\frac{y}{\phi(d)}\right| \ll \sum_{d\le z}\frac{1}{\phi(d)}\sum_{\substack{\chi(\bmod d)\\ \chi\neq\chi_0}}\max_{y\le x}|\psi(y,\chi_1)|$$
$$+\sum_{d\le z}\frac{1}{\phi(d)}\max_{y\le x}|\psi(y)-y|+z(\log z)(\log x),$$

onde $z=z(x)$ é um número real positivo, que depende de x, a ser especificado em breve. Para o segundo termo acima nós usamos o teorema dos números primos dado pela seguinte forma

$$\psi(x)=x+O(x\exp(-c\sqrt{\log x})),$$

obtendo que

$$\sum_{d\le z}\frac{1}{\phi(d)}\max_{y\le x}|\psi(y)-y| \ll \frac{x\log z}{(\log x)^{A+1}}. \tag{9.2.5}$$

Ainda resta estimar o primeiro termo.

Nós escrevemos todos os módulos d como $d=d_1k$ para algum inteiro positivo k e observamos que

$$\sum_{d\le z}\frac{1}{\phi(d)}\sum_{\substack{\chi(\bmod d)\\ \chi\neq\chi_0}}\max_{y\le x}|\psi(y,\chi_1)| = \sum_{d_1\le z}\sum_{k\le\frac{z}{d_1}}\frac{1}{\phi(d_1k)}\sum_{\chi_1(\bmod d_1)}\max_{y\le x}|\psi(y,\chi_1)|$$
$$\ll \sum_{d\le z}\frac{1}{\phi(d)}\log\left(\frac{2z}{d}\right)\sum_{\chi}{}^{*}|\psi(y,\chi)|,$$

onde nós usamos a estimativa

$$\sum_{k\le\frac{z}{d}}\frac{1}{dk}\ll\frac{1}{\phi(d)}\log\frac{2z}{d}.$$

Para completarmos a demonstração do teorema nós precisamos escolher z apropriadamente (i.e. da forma $z=x^{1/2}/(\log x)^B$ para alguma constante positiva $B=B(A)$) e mostrar que

$$\sum_{d\le z}\frac{1}{\phi(d)}\sum_{\chi}^{*}\max_{y\le x}|\psi(y,\chi)|\ll\frac{x}{(\log x)^A}. \tag{9.2.6}$$

Lembramos que pelo teorema de Siegel–Walfisz que diz que,

$$\psi(x,\chi)\ll x\exp(-c\sqrt{\log x}),$$

existe $B=B(A)>0$ tal que, se $d\le(\log x)^B$ e $\chi\neq\chi_0$ é um caractere módulo d, então

$$\psi(y,\chi)\ll\frac{x}{(\log x)^{A+1}}.$$

Assim

$$\sum_{d\le(\log x)^B}\frac{1}{\phi(d)}\sum_{\chi}^{*}\max_{y\le x}|\psi(y,\chi)|\ll\frac{x}{(\log x)^A}. \tag{9.2.7}$$

Agora escolhemos

$$z:=\frac{x^{1/2}}{(\log x)^B}$$

e aplicamos Corolário 9.2.3 e somas parciais para obter

$$\sum_{(\log x)^B<d\le z}\frac{1}{\phi(d)}\sum_{\chi}^{*}\max_{y\le x}|\psi(y,\chi)|\ll\frac{x}{(\log x)^A}. \tag{9.2.8}$$

Combinando (9.2.5), (9.2.7) e (9.2.8), a demonstração do teorema está completa.

□

O teorema de Bombieri–Vinogradov foi estendido por Bombieri *et al.* [4] da seguinte maneira. Seja $a\neq0$ e $x\ge y\ge3$. Então

$$\sum_{\substack{d\le x^{\frac12}y^{\frac12}\\(d,a)=1}}\left|\psi(x;d,a)-\frac{x}{\phi(d)}\right|\ll x\left(\frac{\log y}{\log x}\right)^2(\log\log x)^B.$$

Aqui B é uma constante absoluta e a constante implícita pelo símbolo $\ll$ depende somente de a.

Isto significa que na maioria das aplicações envolvendo o teorema de Bombieri–Vinogradov, nós podemos tomar

$$d \leq x^{\frac{1}{2}} \exp\left(\frac{\log x}{(\log\log x)^B}\right)$$

ao invés de $d \leq x^{1/2}(\log x)^{-B}$. Em particular, isto se aplica ao problema do divisor de Titchmarsh (a ser discutido na próxima seção) e nós podemos deduzir a existência de constantes positivas c e c_1 tal que

$$\sum_{p\leq x} d(p+a) = cx + c_1\frac{x}{\log x} + O\left(\frac{x(\log\log x)^B}{(\log x)^2}\right).$$

O problema de provar resultados da forma

$$\sum_{d\leq x^\theta} \max_{(a,d)=1} \left|\psi(x;d,a) - \frac{x}{\phi(d)}\right| \ll \frac{x}{(\log x)^A}$$

para todo $A > 0$ e $\theta > 1/2$ é extremamente difícil. Uma famosa conjectura de Elliott e Halberstam prevê que a desigualdade acima é verdadeira para todo $\theta < 1$.

Se descartarmos o valor absoluto e fixarmos a, então alguns resultados significantes foram obtidos. Sem entrar em muitos detalhes, nós mencionamos a conhecida 'funções bem-fatorável $\lambda(d)$ de nível z'. Nós temos que para todo $\varepsilon > 0$ e $z := x^{\frac{4}{7}-\varepsilon}$,

$$\sum_{\substack{(d,a)=1\\ d\leq z}} \lambda(d)\left(\psi(x;d,a) - \frac{x}{\phi(d)}\right) \ll \frac{x}{(\log x)^A}$$

para todo $A > 0$.

Este resultado possui numerosas aplicações. Por exemplo, no caso do problema do divisor de Titchmarh nós podemos mostrar que para todo $A > 0$,

$$\sum_{p\leq x} d(p+a) = cx + c_1\mathrm{li}(x) + O\left(\frac{x}{(\log x)^A}\right)$$

para constantes positivas c, c_1. Outra aplicação é vista na conjectura da raíz primitiva de Artin e também na teoria de curvas elípticas. A demonstração desta melhora na estimativa é difícil e não será vista neste curso. (veja [])

9.3 O Problema do Divisor de Titchmarsh

Seja a um inteiro fixado. Lembramos que no Capítulo 3 nós consideramos a questão de determinar o comportamento assintótico da função

$$\sum_{p \leq x} \nu(p+a)$$

para $a = -1$. Agora vamos considerar a questão mais complexa de determinar o comportamento assintótico de

$$\sum_{p \leq x} d(p+a),$$

onde $d(\cdot)$ é a funçao divisor. Este problema é conhecido na literatura como **o problema do divisor de Titchmarsh**. Foi primeiramente estudado por Titchmarsh em 1930 e está relacionado com uma famosa conjectura de Hardy e Littlewood, formulada em 1922 enunciando que todo inteiro suficientemente grande pode ser representado como a soma de um primo e dois quadrados. Já em 1930, Titchmarsh mostrou que

$$\sum_{p \leq x} d(p+a) = O(x),$$

e que, usando a hipótese generalizada de Riemann, uma fórmula assintótica explícita para $\sum_{p \leq x} d(p+a)$ (veja (9.3.1) abaixo) também é válida.

Em 1923, Hardy and Littlewood sugeriram que a sua própria conjectura era verdadeira para quase todos inteiros se assumirmos a hipótese generalizada de Riemann, e em 1928 Stanley mostrou que isto era de fato verdade. E mais tarde a hipótese do resultado de Stanley foi enfraquecida por S. Chowla (1935), A. Walfisz (1935), T.

Estermann (1936), H. Halberstam (1951), e C. Hooley (1957), sem ser completamente removida. Somente em 1960 Linnik obteve uma demonstração incondicional da conjectura de Hardy–Littlewood. Seu método, agora conhecido como 'método da dispersão', pode também ser usado para fornecer uma demonstração incondicional do problema do divisor de Titchmarsh.

Nosso objetivo nesta seção é descrever uma simples demosntração do problema do divisor de Titchmarsh. Mais precisamente, nós mostramos que:

Teorema 9.3.1. *Seja a um número inteiro fixado. Então existe uma constante positiva c tal que*

$$\sum_{p\le x} d(p+a) = cx + O\left(\frac{x\log\log x}{\log x}\right). \tag{9.3.1}$$

Demonstração. Primeiro observamos que para todo inteiro positivo n,

$$d(n) = 2\sum_{\substack{d|n\\ d\le\sqrt{n}}} 1 - \delta(n),$$

onde $\delta(n) = 1$ se n é um quadrado e 0 caso contrário. Assim

$$\begin{aligned}\sum_{p\le x} d(p+a) &= 2\sum_{p\le x}\sum_{\substack{d|p+a\\ d\le\sqrt{x}}} 1 - \sum_{p\le x}\delta(p+a)\\ &= 2\sum_{d\le\sqrt{x}} \pi(x,d,-a) + O(\sqrt{x}).\end{aligned} \tag{9.3.2}$$

Nós agora lembramos que o teorema de Bombieri–Vinogradov nos permite controlar o termo do erro na fórmula assintótica para $\pi(x;d,-a)$, desde que $d \le \sqrt{x}(\log x)^{-B}$ para alguma constante positiva B (a ser especificada mais tarde). Isto nos sugere a separar a soma no lado direito de (9.3.2) em duas partes:

$$\sum_{d\le\sqrt{x}}\pi(x,d,-a)=\sum_{d\le\frac{\sqrt{x}}{(\log x)^B}}\pi(x,d,-a)+\sum_{\frac{\sqrt{x}}{(\log x)^B}<d\le\sqrt{x}}\pi(x,d,-a). \tag{9.3.3}$$

Para a primeira soma em (9.3.3) nós escrevemos

$$\sum_{d\le\frac{\sqrt{x}}{(\log x)^B}}\pi(x,d,-a)=\sum_{d\le\frac{\sqrt{x}}{(\log x)^B}}\left(\pi(x;d,-a)-\frac{\mathrm{li}(x)}{\phi(d)}\right)+\sum_{d\le\frac{\sqrt{x}}{(\log x)^B}}\frac{\mathrm{li}(x)}{\phi(d)}$$

e usamos o teorema de Bombieri–Vinogradov para obter um limitante superior de

$$\mathrm{li}(x)\sum_{d\le\frac{\sqrt{x}}{(\log x)^B}}\frac{1}{\phi(d)}+O\left(\frac{x}{(\log x)^A}\right),$$

para qualquer $A>0$ e algum $B=B(A)>0$. Agora nós sabemos que existe uma constante positiva c_0 tal que, para todo x,

$$\sum_{d\le x}\frac{1}{\phi(d)}=c_0\log x+O(1). \tag{9.3.4}$$

(prove isto!)

Portanto

$$\sum_{d\le\frac{\sqrt{x}}{(\log x)^B}}\pi(x,d,-a)=\frac{c_0}{2}x+O(\frac{x\log\log x}{\log x}). \tag{9.3.5}$$

Para a segunda soma em (9.3.3) nós usamos o teorema de Brun–Titchmarsh e, novamente, (9.3.4). Nós obtemos

$$\sum_{\frac{\sqrt{x}}{(\log x)^B}<d\le\sqrt{x}}\pi(x;d,-a)\ll\frac{x}{\log x}\sum_{\frac{\sqrt{x}}{(\log x)^B}<d\le\sqrt{x}}\frac{1}{\phi(d)}\ll\frac{x\log\log x}{\log x}. \tag{9.3.6}$$

A demonstração do teorema está completa combinando (9.3.2), (9.3.3), (9.3.5) e (9.3.6), e colocando $c := c_0/2$. □

9.4 Exercícios

1. Use somas parciais e a desigualdade de Cauchy–Schwarz para provar a Proposição 9.1.1.

2. Prove a identidade de Vaughan (9.1.4) e (9.1.5)–(9.1.8).

3. Mostre que para a fixo e $\delta(\cdot)$ definidco como na seção anterior nós temos que

$$\sum_{p\leq x} \delta(p+a) = O(\sqrt{x}).$$

Bibliografia

[1] J.C. Andrade, *Uma Intrdução aos Métodos de Crivos em Teoria dos Números e suas Aplicações*, livro a ser publicado.

[2] T.M. Apostol, *Introduction to Analytic Number Theory*, New York: Springer–Verlag, 1976.

[3] E. Bombieri, J. Friedlander and H. Iwaniec, *Primes in arithmetic progressions to large moduli*, Acta Math., **156** (1986), 203–251.

[4] E. Bombieri, J. Friedlander and H. Iwaniec, *Primes in arithmetic progressions to large moduli II*, Math. Ann., **277**:3 (1987), 361–393.

[5] V. Brunn, *Über das Goldbachsche Gesetz und die Anzahl der Primzahlpaare*, Archiv for Matrh. og Naturvid., **B34**: 8 (1915), 19 pages.

[6] A.C. Cojocaru and M.R. Murty, *An Introduction to Sieve Methods and their Applications*, Cambridge, Cambridge University Press, 2005.

[7] H. Davenport, *Multiplicative Number Theory*, (New York: Springer–Verlag), 2000.

[8] P.D.T.A. Elliott, *Probabilistic Number Theory II: Central Limits Theorem*, New York, Berlin: Springer–Verlag, 1980.

[9] J. Friedlander and H. Iwaniec, *Opera de Cribro*, AMS Colloquium Publications Vol. 57, American Mathematical Society, Providence-RI, 2010.

[10] P.X. Gallagher, *A large sieve*, Acta Arithm. **18** (1971), 77–81.

[11] H. Halberstam e H.E. Richert, *Sieve Methods*, London Mathematical Society monographs, No 4, (London, New York: Academic Press, 1974).

[12] D.R. Heath–Brown, *The square sieve and consecutive squarefree numbers*, Math. Ann., **266** (1984), 251–259.

[13] M. Hindry e J. Silverman, *Introduction to Diophantine Geometry* (New York: Springer–Verlag), 2001.

[14] K. Ireland e M. Rosen, *A Classical Introduction to Modern Number Theory*, 2nd (New York: Springer–Verlag), 1998.

[15] W. Li, *Number Theory with Applications* (Singapore: World Scientific), 1996.

[16] H.L. Montgomery and R.C. Vaughan, *The large sieve*, Mathematika, **20** (1973), 119–134.

[17] Y. Motohashi, *An induction principle for the generalizations of Bombieri's prime number theorem*, Proc. Japan Acad., **52**:6 (1976), 273–275.

[18] M. Ram Murty, *Sieving using Dirichlet series*, In Current Trends in Number Theory, ed. S.D. Adhikari, S.A. Katre e B. Ramakrishnan, (New Delhi: Hindustan Book Agency), 2002, 111–124

[19] J.-P. Serre, *Majorations de sommes exponentielles*, Journées Arith. Caen, Astérique, **41–42** (1977), 111–126.

[20] C.L. Siegel, *The integer solutions of the equation* $y^2 = ax^n + bx^{n-1} + \cdots + k$, J. London Math. Soc., **1** (1920), 66–68.