Universidade de Brasília
Instituto de Ciências Exatas
Departamento de Matemática

Ortogonalidade da Função de Möbius

por

Josimar Joao Ramirez Aguirre

Orientador: Leandro Martins Cioletti

Brasília
2014

AGRADECIMENTOS

Desejo dedicar meu trabalho em primeira instância a Deus que me deu força, fé, saúde, sempre me apoiou, dotou-me de grandes dons e talentos e que me deu esperança para alcançar meus objetivos.

Gostaria de agradecer especialmente, aos meus pais Clemente e Francisca, que sempre me apoiaram com grande espírito, mesmo estando longe de mim, contribuindo imensamente para alcançar minhas metas.

Agradeço da mesma forma a meus tios Samuel e Salomé, que são como segundos pais pra mim. Estarei sempre grato por seus conselhos e porque me tomaram como um filho na sua casa quando precisei deles. Também agradeço a meus irmãos, que sempre me apoiaram para continuar estudando, assim como ao resto da minha família e a meus amigos que me deram força e apoio nas horas difíceis.

Agradeço de forma particular a meu orientador Dr. Leandro Martins Cioletti, que teve a paciência necessária para orientar-me, apoiar-me e cujas contribuições ajudaram a melhorar este trabalho assim como a me tornar uma grande pessoa e profissional. Da mesma forma agradeço a meu primeiro orientador Dr. Nigel John Edward Pitt por mostrar-me o caminho a seguir neste trabalho e pela ajuda dispensada.

Aos professores que estiveram comigo durante este longo caminho, sempre me dando orientação profissional, ética na aquisição de conhecimentos e fortalecendo a minha formação como estudante.

Obrigado.

Resumo

Nesta dissertação de Mestrado apresentamos uma nova prova do Teorema de Davenport (1937), e a prova de Terence Tao que a conjectura de Chowla implica a conjectura de Sarnak.

Na primeira parte do trabalho apresentamos a teoria básica das L-funções bem como uma variação método de Vinogradov, usando as identidades de Vaughan. Em seguida, usamos estas ferramentas para mostrar o Teorema de Davenport. A principal referência desta parte são os capítulos 5 e 13 do livro *Analitic Number Theory* de Henryk Iwaniec e Emmanuel Kowalski, [9].

A prova que a Conjectura de Chowla implica na Conjectura de Sarnak é baseada em princípio de grandes desvios, obtido por uma variação do método do segundo momento. A exposição é inspirada na primeira parte do artigo de Peter Sarnak, intitulado *Three Lectures on the Mobius Function Randomness and Dynamics*, [16].

Palavras-chave: Ortogonalidade da função de Möbius, Método de Vinogradov, Identidades de Vaughan, Formas Bilineares, Desigualdades diofântinas, L-funções, Região livre de zeros, caracteres de Dirichlet.

Abstract

In this Master's thesis we present a new proof of Davenport's Theorem (1937), and the Terence Tao's proof that Chowla conjecture implies Sarnak's conjecture.

In the first part of this work we present the basic theory of L-functions and a variation of the Vinogradov's method using the Vaughan's identities. Then we use these tools to prove Davenport's Theorem. This section is based on chapters 5 and 13 of the reference *Analytic Number Theory* by Henryk Iwaniec and Emmanuel Kowalski, [9].

The Chowla's Conjecture implies Sarnak's Conjecture is based on a principle of large deviations obtained by variation of the second moment method. The exposition is inspired on the first part of Peter Sarnak's article entitled *Three Lectures on the Mobius Function Randomness and Dynamics*, [16].

Keywords: Möbius orthogonality, Vinogradov's Method , Vaughan's Identities, Bilinear Forms, Diophantine inequalities, L-functions, Region free of zeros, Dirichlet's Characters.

A Matemática é a rainha das
ciências e a Teoria dos Números
é a rainha da Matemática.

Carl Friedrich Gauss

Sumário

Introdução

Existe uma lenda na Alemanha sobre Barbarrosa (o imperador Federico I): ele era muito querido pelo povo alemão e dado que ele morreu numa guerra, seu corpo foi enterrado num túmulo distante. Originou-se uma lenda de que ele ainda estava vivo, dormindo numa caverna da montanha Kyffhäuser, mas acordaria e voltaria, depois de centenas de anos, quando o povo alemão precisasse dele.

Dizem que alguém perguntou a Hilbert:

> *Se você voltasse a vida, como Barbarrosa, depois de quinhentos anos, o que você faria?*
> *Respondeu Hilbert:*
> *- Perguntaria se alguém já provou a hipótese de Riemann.*

A hipótese de Riemann, formulada pela primeira vez por Bernhard Riemann em 1859, é para muito matemáticos o problema em aberto mais importante na matemática contemporânea. Ela é uma conjectura sobre a distribuição dos zeros da função zeta de Riemann. Esta função é definida como segue: para $\mathrm{Re}(s) > 1$,

$$\zeta(s) = \sum_{n=1}^{\infty} n^{-s} = \prod_{p} (1 - p^{-s})^{-1},$$

onde o produto acima percorre sob os números primos p. Em seguida, estendemos esta função meromorficamente a todo plano complexo, para maiores detalhes vide Apêndice 5.3. Esta conjectura afirma que os zeros não triviais da função zeta estão localizados na reta $\mathrm{Re}(s) = 1/2$.

Muitos matemáticos de grande reputação tentaram provar ou refutar essa afirmação, mas até hoje este problema continua em aberto. Em matemática, muitas vezes para tentar resolver um problema procura-se alguma equivalência que permita por em jogo novas ferramentas; vejamos agora uma forma distinta de olhar este problema.

Considere os números naturais, $1, 2, 3, 4, 5, \ldots$ e em seguida, descarte aqueles que sejam divisíveis pelo quadrado de um número maior que 1, isto é, apagamos da lista os números $4, 8, 9, 16, 18, \ldots$ desta forma obtemos todos os naturais livres de quadrados

$$1, 2, 3, 5, 6, 7, 10, 11, 13, 14, 15, \ldots$$

Cada um dos números na lista anterior, exceto o um, tem uma fatoração única como produto de primos distintos, alguns destes naturais, livres de quadrados, são o produto de um número par de primos distintos e outros de um número impar de primos distintos. Chamaremos um número natural de "bom" se ele for o número um ou produto de um número par de primos distintos, caso contrario, chamaremos de "mau". Assim $6 = 2.3$ é bom e $30 = 2.3.5$ é mau. A hipótese de Riemann é equivalente a dizer que, para qualquer natural n suficientemente grande, a diferença numérica entre os números bons e os maus no intervalo $[1, n]$ não é muito grande, veja Teorema abaixo. Para tornar esta discussão mais precisa, considere a função de Möbius, $\mu : \mathbb{N} \to \mathbb{C}$ definida por

$$\mu(n) := \begin{cases} 1, & \text{se } n = 1; \\ (-1)^r, & \text{se } n \text{ é produto de } r \text{ primos distintos;} \\ 0, & \text{se } n \text{ não é livre de quadrados.} \end{cases}$$

A função de Möbius claramente codifica a estrutura multiplicativa de um número livre de quadrados. Por suas diversas propriedades esta função é uma das mais importantes em teoria dos números. Veremos algumas destas propriedades interessantes no capítulo seguinte. Usando esta função podemos formalizar a discussão feita acima através do seguinte teorema

Teorema. *A hipótese de Riemann é equivalente à afirmação que* $\forall \epsilon > 0$ *tem-se que*

$$M(x) = \sum_{n \leqslant x} \mu(n) = O(x^{1/2+\epsilon}).$$

Demonstração. Veja [18], Teorema 14.25. □

É bem conhecido que a hipótese de Riemann é equivalente a vários outros problemas de matemática. Uma característica importante sobre a equivalência acima é que podemos dar um argumento heurístico, usando a teoria de probabilidades, que sugere que $M(x) = O(x^{1/2+\epsilon})$ [1] e portanto que a hipótese de Riemann é verdadeira. Este argumento é explicado abaixo:

Vamos lançar N vezes uma moeda "honesta", com N suficientemente grande. Seja X a variável aleatória que conta o número de caras obtidas nestes N lançamentos. Supondo que os lançamentos são independentes então podemos afirmar que X tem distribuição binomial, media $\mu = N/2$ e variância $\sigma^2 = N/4$. Pela desigualdade de Chebyshev temos que

$$\mathbb{P}\Big(\big|X - \frac{N}{2}\big| < KN^{1/2}\Big) = 1 - \mathbb{P}\Big(\big|X - \frac{N}{2}\big| \geqslant 2K\frac{N^{1/2}}{2}\Big) \geqslant 1 - \frac{1}{4K^2}.$$

[1] Os símbolos O(g) e o(g) representam as notações de Landau, veja o Apêndice

Seja Y a variável aleatória que denota o número de caras menos o número de coroas, isto é, $Y = X - (N - X) = 2(X - N/2)$. Usando esta igualdade na expressão obtemos a seguinte desigualdade

$$\mathbb{P}\Big(|Y| < 2KN^{1/2}\Big) \geqslant 1 - \frac{1}{4K^2}.$$

Tomando $K = N^{\epsilon}/2$, temos que

$$\mathbb{P}\Big(|Y| < N^{1/2+\epsilon}\Big) \geqslant 1 - \frac{1}{4(N^{\epsilon}/2)^2} = 1 - \frac{1}{N^{2\epsilon}} \underset{N\to\infty}{\to} 1, \tag{1}$$

isto é, com probabilidade próximo de um o número de caras menos o número de coroas tem crescimento limitado por $N^{1/2+\epsilon}$.

Assumindo que $\mu(n)$, para todo n livre de quadrados, é uma variável aleatória que toma valores em $\{-1, 1\}$ com probabilidade uniforme e supondo que $\mu(n)$'s são independentes podemos pensar em $M(x)$ como sendo a variável aleatória Y. Assim concluímos de (1) que a igualdade $M(x) = O(x^{1/2+\epsilon})$ e a Hipótese de Riemann (pelo teorema acima) são verdadeiras com probabilidade próximo de um.

Este argumento mostra a forte relação que existe entre a Hipótese de Riemann e a importância de estudar o comportamento da função acumulativa de Möbius (Função de Mertens).

Outro resultado que desperta nosso interesse em estudar a função de Möbius vem do fato de que o comportamento desta função está intimamente relacionado ao Teorema dos Números Primos (TNP), que diz que o número, $\pi(x)$, de números primos menores ou iguais que x é assintoticamente igual a $x/\log x$, isto é,

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

que é um dos principais resultados em Teoria Analítica dos Números. Veremos mais adiante que o TNP é equivalente a $\mu(n)$ ter assintoticamente valor médio nulo, isto é,

$$\sum_{n\leqslant x} \mu(n) = o(x).$$

Em [16], Peter Sarnak, expressa a igualdade acima simplesmente dizendo que a função de Möbius é ortogonal a função constante igual a 1. Caso geral, nos dizemos que $\mu(n)$ é (assintoticamente) ortogonal a $\xi(n)$ se a igualdade abaixo é satisfeita

$$\sum_{n\leqslant N} \mu(n)\xi(n) = o(x), \tag{2}$$

para funções aritméticas $\xi(n)$.

No Capítulo 1 apresentaremos as definições e propriedades básicas de algumas funções multiplicativas e alguns resultados clássicos da Teoria Analítica dos

Números. Por último veremos algumas equivalências do TNP ressaltando aquela que envolve a função de Möbius.

No Capítulo 2 expomos a teoria básica das L-funções, nos concentrando nos problemas de como estão distribuídos seus zeros no plano complexo. Usando esta ferramenta vamos provar o Teorema dos Números Primos Generalizado, que nos diz qual é o comportamento assintótico da função acumulativa dos coeficientes de Dirichlet da derivada logarítmica de uma L-função. Resultado chave na abordagem adotada no capítulo seguinte.

No Capítulo 3, iremos estudar o comportamento assintótico de (2) no caso em que $\xi(n) = e(\alpha n)$, que se trata do resultado provado pela primeira vez por H. Davenport. Mas aqui não seguiremos as ideias originais dele, vamos demonstrar esse resultado fazendo uso do Capítulo 2, do método de Vinogradov, das identidades de Vaughan e de alguns fatos elementares sobre formas bilineares, seguindo de perto a abordagem de H. Iwaniec e Kowalski [9].

Por último, no Capítulo 4 abordamos o comportamento dinâmico do fluxo de Möbius seguindo as notas de uma palestra apresentada recentemente por Peter Sarnak no *Institute for Advanced Studies (IAS)* e também a primeira parte do artigo de Peter Sarnak, intitulado *Three Lectures on the Mobius Function Randomness and Dynamics*.

Também no Capítulo 4, apresentamos uma conjectura devida a Sarnak e alguns resultados sugerindo que ela seja verdadeira.

Capítulo 1

Preliminares

1.1 Resultados Básicos

Nesta seção apresentaremos as propriedades básicas satisfeitas por algumas funções aritméticas que serão necessárias a frente. A propriedade fundamental da função de Möbius é dada no seguinte teorema

Teorema 1.1 (Identidade de Möbius). *Para todo* $n \in \mathbb{N}$

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

Demonstração. Se $n = 1$ então a fórmula é facilmente verificada.

Para todo $n > 1$, denotamos por $n = p_1^{a_1} \dots p_r^{a_r}$ a fatoração única de n em potências de primos distintos. Chamamos de $N = p_1 \dots p_r$ o radical de n. Já que $\mu(d) = 0$ a menos que d seja livre de quadrados, temos

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

A última soma contém 2^r parcelas, cada uma correspondendo ao um subconjunto de $\{p_1, \dots, p_r\}$, uma vez que os divisores de N estão em correspondência um-a-um com tais subconjuntos. O número de subconjuntos com k elementos é o coeficiente binomial de r tomado k e para cada divisor d, determinado por tal subconjunto, temos que $\mu(d) = (-1)^k$. Assim

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

e isto completa a demonstração. $\square$

A identidade dada no teorema anterior é a principal razão para a peculiar definição da função de Möbius, que pode parecer um pouco artificial. Em particular, a definição de $\mu(n)$ como 0 quando n não é livre de quadrados parece ser desmotivador. Por outro lado, a função de Liouville $\lambda(n)$, que é idêntica à função de Möbius nos inteiros livre de quadrados, se estende aos inteiros não livre de quadrados e por isto parece ser muito mais natural trabalhar com ela. Porém, esta função não satisfaz a identidade obtida no teorema acima, e este é o motivo de darmos atenção especial a função de Möbius.

Uma das razões pelas quais o Teorema 1.1 é fundamental, é porque ele nos permite deduzir uma fórmula de inversão que é útil em questões envolvendo combinatória.

TEOREMA 1.2 (A Fórmula de Inversão de Möbius). *Sejam f e g duas funções aritméticas. Se*

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

então

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

e reciprocamente.

Demonstração. Trocando a ordem de somas temos as seguintes igualdades

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

onde na última igualdade usamos a Identidade de Möbius. A recíproca se prova de forma análoga. □

Outra identidade que podemos obter a partir do Teorema 1.1 é que o crescimento médio da função acumulativa de Möbius é limitado, como mostra o teorema abaixo.

TEOREMA 1.3. *Para todo $n \geqslant 1$, temos*

$$\left| \sum_{n \leqslant x} \frac{\mu(n)}{n} \right| \leqslant 1$$

Demonstração. Seja x um número real positivo. Note que

$$\sum_{n\leqslant x}\mu(n)\left[\frac{x}{n}\right]=1,$$

onde $[x]$ denota a parte inteira de x. De fato,

$$\sum_{n\leqslant x}\mu(n)\left[\frac{x}{n}\right]=\sum_{n\leqslant x}\sum_{d|n}\mu(d)=1.$$

Usando que $[x]=x-\{x\}$ e a identidade que acabamos de mostrar segue que

$$1=\sum_{n\leqslant x}\mu(n)\left[\frac{x}{n}\right]=\sum_{n\leqslant x}\mu(n)\frac{x}{n}-\sum_{n\leqslant x}\mu(n)\left\{\frac{x}{n}\right\}$$

somando nos extremos da igualdade acima, a segunda parcela à direita e a desigualdade triangular obtemos

$$\begin{aligned} x\Big|\sum_{n\leqslant x}\frac{\mu(n)}{n}\Big| &= \Big|1+\sum_{n\leqslant x}\mu(n)\left\{\frac{x}{n}\right\}\Big|\leqslant 1+\sum_{n\leqslant x}\left\{\frac{x}{n}\right\}\leqslant 1+\{x\}+\sum_{2\leqslant n\leqslant x}\left\{\frac{x}{n}\right\}\\ &< 1+\{x\}+[x]-2+1=x, \end{aligned}$$

onde na última desigualdade usamos que $\{x/n\}\leqslant 1$. Finalmente, cancelando o fator x obtemos o resultado desejado. □

Além disso, uma relação muito usada em Teoria Analítica dos Números que necessitaremos mais adiante é dada no seguinte teorema

Teorema 1.4 (Identidade de Abel). *Sejam $a(n)$ uma função aritmética e*

$$A(x)=\sum_{n\leqslant x}a(n),$$

onde $A(x)=0$ se $x<1$. Seja $f:\mathbb{R}\to\mathbb{R}$ uma função com derivada contínua no intervalo $[y,x]$, onde $0<y<x$. Então é válida a seguinte igualdade

$$\sum_{y<n\leqslant x}a(n)f(n)=A(x)f(x)-A(y)f(y)-\int_y^x A(t)f'(t)dt \tag{1.1}$$

Demonstração. Já que $A(x)$ é uma função degrau com salto $f(n)$ nos inteiros, a soma em (1.1) pode ser expressa como uma integral de Riemann-Stieltjes

$$\sum_{y<n\leqslant x}a(n)f(n)=\int_y^x f(t)\,dA(t).$$

Usando a fórmula de integração por partes segue que

$$\begin{aligned}\sum_{y<n\leqslant x} a(n)f(n) &= f(x)A(x) - f(y)A(y) - \int_y^x A(t)df(t) \\ &= f(x)A(x) - f(y)A(y) - \int_y^x A(t)f'(t)dt\end{aligned}$$

□

Como mencionamos anteriormente, a função Zeta e a função de Möbius estão relacionadas, a primeira relação é dada por

TEOREMA 1.5. *Seja $s = \sigma + it$ um número complexo, se $\sigma > 1$ então*

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

Demonstração. Usando a definição de Zeta, a fórmula do produto de Series de Dirichlet e a Identidade de Möbius temos que

$$\begin{aligned}\zeta(s)\sum_{n=1}^{\infty}\frac{\mu(n)}{n^s} &= \sum_{m=1}^{\infty}\frac{1}{m^s}\sum_{n=1}^{\infty}\frac{\mu(n)}{n^s} \\ &= \sum_{l=1}^{\infty}\sum_{nm=l}\frac{\mu(n)}{l^s} \\ &= \sum_{l=1}^{\infty}\frac{1}{l^s}\sum_{n\mid l}\mu(n) = 1.\end{aligned}$$

Usando produto de Euler e a desigualdade triangular podemos verificar que se $\sigma > 1$ então $\zeta(s) \neq 0$ e isto completa a prova do teorema. □

A função de Möbius frequentemente aparece relacionada com outras funções aritméticas. Já vimos sua relação com a função Zeta, e seguir vamos mostrar que ela também está relacionada com a função de von Mangoldt, definida por

$$\Lambda(n) := \begin{cases} \log p, & \text{se } n = p^k \text{ para algum primo } p \text{ e } k > 0; \\ 0, & \text{caso contrário.} \end{cases}$$

Esta função possui, entre outras, a seguinte propriedade

LEMA 1.1. *Se $n \geqslant 1$, temos*

$$\log n = \sum_{d\mid n} \Lambda(d) \tag{1.2}$$

Demonstração. Se $n = 1$ o lema é certamente verdadeiro pois, ambos lados são zero. Caso $n > 1$, escrevemos $n = \prod_{k=1}^{r} p_k^{a_k}$ assim

$$\log n = \sum_{k=1}^{r} a_k \log(p_k).$$

Considere agora a soma do lado direito de (1.2). Os termos não nulos são os divisores d da forma p_k^m, com $m = 1, 2, \ldots, a_k$ e $k = 1, 2, \ldots, r$. Portanto

$$\sum_{d|n} \Lambda(d) = \sum_{k=1}^{r} \sum_{m=1}^{a_k} \Lambda(p_k^m) = \sum_{k=1}^{r} \sum_{m=1}^{a_k} \log p_k = \sum_{k=1}^{r} a_k \log(p_k) = \log n.$$

□

O próximo lema mostra como estão relacionadas as funções de Möbius e von Malgoldt

Lema 1.2. *Se $n \geqslant 1$ temos*

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

Demonstração. Segue do Lema 1.1 e da Fórmula de Inversão de Möbius que

$$\begin{aligned} \Lambda(n) &= \sum_{d|n} \mu(d) \log \frac{n}{d} = \log n \sum_{d|n} \mu(d) - \sum_{d|n} \mu(d) \log d \\ &= -\sum_{d|n} \mu(d) \log d \end{aligned}$$

□

Agora relacionamos a função de Von Mangoldt com a função Zeta.

Lema 1.3. *Para $Re(s) > 1$ temos*

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

Demonstração. Como $\mathrm{Re}(s) > 1$ então $\zeta(s)$ é absolutamente convergente então usando a definição de Zeta, produto de Series de Dirichlet e o Lema 1.1 temos

$$\begin{aligned} \zeta(s) \sum_{n \geqslant 1} \frac{\Lambda(n)}{n^s} &= \sum_{n \geqslant 1} \frac{1}{n^s} \sum_{m \geqslant 1} \frac{\Lambda(m)}{m^s} = \sum_{l \geqslant 1} \sum_{nm = l} \frac{\Lambda(m)}{l^s} \\ &= \sum_{l \geqslant 1} \frac{1}{l^s} \sum_{n|l} \Lambda(n) = \sum_{l \geqslant 1} \frac{\log\, l}{l^s} = -\zeta'(s). \end{aligned}$$

□

1.2 O Teorema dos Números Primos

Seja $\pi(x)$ o número de primos $p \leqslant x$. Euclides provou que existem infinitos primos. Em particular $\pi(x) \to \infty$ quando $x \to \infty$, e o estudo do comportamento de $\pi(x)$, como função de x se intensificou bastante a partir do século dezoito.

Inspecionando suas tabelas, Gauss (1792) e Legendre (1798) conjeturaram que $\pi(x)$ se comportaria assintoticamente como $x/\log x$, isto é

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

Esta conjectura foi provada em 1896, independentemente, por Jacques Salomon Hadamard e Charles-Jean de La Vallée Poussin e este resultado ficou conhecido na literatura como o Teorema dos Números Primos. Atualmente existem muitas provas deste resultado (Veja [2], para uma delas). Nesta seção estamos interessados em mostrar algumas equivalências deste teorema, sendo para nós a mais importante aquela existente com a função de Möbius. Para isto definimos antes as funções de Chebyshev

DEFINIÇÃO 1.1. *Para $x > 0$ definimos a função Ψ de Chebyshev por*

$$\Psi(x) = \sum_{n \leqslant x} \Lambda(n).$$

Já que $\Lambda(n) = 0$ a menos que n seja potência de um primo, podemos escrever $\Psi(x)$ como segue

$$\Psi(x) = \sum_{n \leqslant x} \Lambda(n) = \sum_{m \geqslant 1} \sum_{p \leqslant x^{1/m}} \log p.$$

Vamos verificar que a soma sobre m é finita. De fato, a soma é vazia se $x^{1/m} < 2$, isto é, se $m > \log x/\log 2$, portanto

$$\Psi(x) = \sum_{m \leqslant \log_2 x} \sum_{p \leqslant x^{1/m}} \log p.$$

Podemos expressar a função Ψ de forma ligeiramente diferente, introduzindo outra função de Chebyshev.

DEFINIÇÃO 1.2. *Para $x > 0$ definimos a função ϑ de Chebyshev pela equação*

$$\vartheta(x) = \sum_{p \leqslant x} \log p.$$

Note que $\Psi(x)$ pode ser reescrita como

$$\Psi(x) = \sum_{m \leqslant \log_2 x} \vartheta(x^{1/m}). \tag{1.3}$$

O seguinte teorema relaciona os dois quocientes $\Psi(x)/x$ e $\vartheta(x)/x$

TEOREMA 1.6. *Para $x > 0$ temos*

$$0 \leqslant \frac{\Psi(x)}{x} - \frac{\vartheta(x)}{x} \leqslant \frac{(\log x)^2}{2\sqrt{x}\log 2}.$$

E portanto

$$\lim_{x\to\infty}\Big(\frac{\Psi(x)}{x} - \frac{\vartheta(x)}{x}\Big) = 0$$

Demonstração. Da equação (1.3) temos que

$$0 \leqslant \Psi(x) - \vartheta(x) = \sum_{2\leqslant m\leqslant \log_2 x} \vartheta(x^{1/m}).$$

Segue da definição de $\vartheta(x)$ a desigualdade trivial $\vartheta(x) \leqslant x\log x$, assim

$$\begin{aligned} 0 \leqslant \Psi(x) - \vartheta(x) &\leqslant \sum_{2\leqslant m\leqslant \log_2 x} x^{1/m}\log(x^{1/m}) \leqslant (\log_2 x)\sqrt{x}\log\sqrt{x} \\ &= \frac{\log x}{\log 2}.\frac{\sqrt{x}}{2}\log x = \frac{\sqrt{x}(\log x)^2}{2\log 2} \end{aligned}$$

dividindo por x obtemos o resultado. □

Agora usaremos a identidade de Abel para expressar a funções $\vartheta(x)$ e $\pi(x)$ em termos de integrais

TEOREMA 1.7. *Para $x > 2$ temos*

$$\vartheta(x) = \pi(x)\log x - \int_2^x \frac{\pi(t)}{t}dt$$

e

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

Demonstração. Seja $a(n)$ a função característica dos primos, isto é $a(n) = 1$ se n for primo e $a(n) = 0$ caso contrário, temos

$$\pi(x) = \sum_{p\leqslant x} 1 = \sum_{1<n\leqslant x} a(n), \quad \vartheta(x) = \sum_{p\leqslant x}\log p = \sum_{1<n\leqslant x} a(n)\log n.$$

Tome $f(x) = \log x$ em (1.1) com $y = 1$ obtemos

$$\vartheta(x) = \sum_{p\leqslant x}\log p = \sum_{1<n\leqslant x} a(n)\log n = \pi(x)\log x - \pi(1)\log 1 - \int_1^x \frac{\pi(t)}{t}dt,$$

o que prova a primeira identidade pois $\pi(t) = 0$ para $t < 2$. Para a segunda identidade, definimos $b(n) = a(n)\log n$, assim podemos escrever

$$\pi(x) = \sum_{\frac{3}{2}<n\leqslant x} \frac{b(n)}{\log n}, \qquad \vartheta(x) = \sum_{n\leqslant x} b(n).$$

Tomando $f(x) = 1/\log x$ em (1.1) com $y = 3/2$ obtemos

$$\pi(x) = \frac{\vartheta(x)}{\log x} - \frac{\vartheta(3/2)}{\log \frac{3}{2}} + \int_{\frac{3}{2}}^{x} \frac{\vartheta(t)}{t\log^2 t}dt$$

o que prova a segunda afirmação pois $\vartheta(t) = 0$ se $t < 2$. $\square$

Apresentamos, em seguida, a primeira equivalência do TNP.

TEOREMA 1.8. *O Teorema dos Números Primos é equivalente à afirmação*

$$\lim_{x\to\infty} \frac{\Psi(x)}{x} = 1 \tag{1.4}$$

Demonstração. Pelo Teorema 1.6 basta provar a equivalência acima com respeito a função $\vartheta(x)$. Do Teorema 1.7

$$\frac{\vartheta(x)}{x} = \frac{\pi(x)\log x}{x} - \frac{1}{x}\int_2^x \frac{\pi(t)}{t}dt$$

e

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

então para mostrar que o TNP implica (1.4) é suficiente mostrar que

$$\lim_{x\to\infty} \frac{1}{x}\int_2^x \frac{\pi(t)}{t}dt = 0.$$

Pelo TNP, $\frac{\pi(t)}{t} = O\Big(\frac{1}{\log t}\Big)$ para $t \geqslant 2$. Assim

$$\frac{1}{x}\int_2^x \frac{\pi(t)}{t}dt = O\Big(\frac{1}{x}\int_2^x \frac{dt}{\log t}\Big). \tag{1.5}$$

Decompondo o intervalo de integração e usando majorantes triviais temos

$$\int_2^x \frac{dt}{\log t} = \int_2^{\sqrt{x}} \frac{dt}{\log t} + \int_{\sqrt{x}}^x \frac{dt}{\log t} \leqslant \frac{\sqrt{x}}{\log 2} + \frac{x-\sqrt{x}}{\log\sqrt{x}}.$$

Dividindo por x ambos membros da desigualdade acima segue de (1.5) que

$$\lim_{x\to\infty} \frac{1}{x}\int_2^x \frac{\pi(t)}{t}dt = 0,$$

mostrando que a igualdade (1.4) é verdadeira. Reciprocamente, suponha que vale (1.4), portanto $\vartheta(t) = O(t)$ e

$$\frac{\log x}{x}\int_2^x \frac{\vartheta(t)}{t\log^2 t}dt + O\Big(\frac{\log x}{x}\int_2^x \frac{dt}{\log^2 t}\Big)$$

usando novamente as cotas triviais temos

$$\int_2^x \frac{dt}{\log^2 t} = \int_2^{\sqrt{x}} \frac{dt}{\log^2 t} + \int_{\sqrt{x}}^x \frac{dt}{\log^2 t} \leqslant \frac{\sqrt{x}}{\log^2 2} + \frac{x-\sqrt{x}}{\log^2\sqrt{x}}.$$

Daí segue imediatamente que

$$\frac{\log x}{x}\int_2^x \frac{\vartheta(t)}{t\log^2 t}\,dt \to 0, \quad \text{quando } x \to \infty$$

o que conclui a prova. □

Para finalizar este capítulo provaremos que o TNP está fortemente relacionado com a ortogonalidade da função de Möbius. Mais explicitamente temos o

Teorema 1.9. *O Teorema dos Números Primos é equivalente à identidade*

$$\frac{M(x)}{x} = \frac{1}{x}\sum_{n\leqslant x}\mu(n) = o(1)$$

Para demonstrar o teorema acima precisamos da seguinte definição.

Definição 1.3. *Para todo $x \geqslant 1$ definimos*

$$H(x) = \sum_{n\leqslant x}\mu(n)\log n.$$

O próximo teorema nos dá informações sobre o comportamento assintótico de $M(x)/x$, mostrando que ele pode ser obtido através do comportamento assintótico da função $H(x)/(x\log x)$.

Teorema 1.10. *Temos*

$$\lim_{x\to\infty}\Big(\frac{M(x)}{x} - \frac{H(x)}{x\log x}\Big) = 0 \tag{1.6}$$

Demonstração. Tomando $f(t) = \log t$ no Teorema 1.4 obtemos

$$H(x) = \sum_{n\leqslant x}\mu(n)\log n = M(x)\log x - \int_1^x \frac{M(t)}{t}dt.$$

Portanto se $x > 1$ temos

$$\frac{M(x)}{x} - \frac{H(x)}{x\log x} = \frac{1}{x\log x}\int_1^x \frac{M(t)}{t}dt.$$

Daí para provar o teorema basta mostrar que

$$\lim_{x\to\infty} \frac{1}{x\log x}\int_1^x \frac{M(t)}{t}dt = 0. \tag{1.7}$$

Usando a estimativa trivial $M(x) = O(x)$ obtemos

$$\int_1^x \frac{M(t)}{t}dt = O\Big(\int_1^x dt\Big) = O(x)$$

daí segue que (1.7) é verdadeira e logo (1.6). □

Vamos agora provar a condição suficiente do Teorema 1.9, isto é

TEOREMA 1.11. *O Teorema dos Números Primos implica que*

$$\lim_{x\to\infty} \frac{M(x)}{x} = 0$$

Demonstração. A ideia é usar o TNP na forma $\Psi(x) \backsim x$ para provar, pelo teorema anterior, que $H(x)/(x\log x) \to 0$, quando $x \to \infty$. Para isto necessitamos da seguinte identidade

$$-H(x) = -\sum_{n\leqslant x} \mu(n)\log n = \sum_{n\leqslant x} \mu(n)\Psi\left(\frac{x}{n}\right). \tag{1.8}$$

(Ver [2], pag 92). Como $\Psi(x) \sim x$, dado $\epsilon > 0$ existe uma constante $A > 0$ tal que se $x \geqslant A$ então

$$\Big|\frac{\Psi(x)}{x} - 1\Big| < \epsilon.$$

Em outras palavras, tem-se

$$|\Psi(x) - x| < \epsilon x, \ \text{ quando } \ x \geqslant A. \tag{1.9}$$

Escolhemos $x > A$ e decompomos a soma do lado direito de (1.8) em duas parcelas do tipo $\sum_{n\leqslant y} + \sum_{y<n\leqslant x}$, onde $y = [x/A]$. Observe que na primeira soma $n \leqslant x/A$, e portanto $x/n \geqslant A$. Daí podemos usar (1.9) para escrever

$$\left|\Psi\left(\frac{x}{n}\right) - \frac{x}{n}\right| < \epsilon\frac{x}{n}, \quad \text{se } \ n \leqslant y. \tag{1.10}$$

Somando e subtraindo x/n temos

$$\begin{aligned}\sum_{n\leqslant y}\mu(n)\Psi\left(\frac{x}{n}\right) &= \sum_{n\leqslant y}\mu(n)\Big(\frac{x}{n}+\Psi\left(\frac{x}{n}\right)-\frac{x}{n}\Big)\\ &= x\sum_{n\leqslant y}\frac{\mu(n)}{n}+\sum_{n\leqslant y}\mu(n)\Big(\Psi\left(\frac{x}{n}\right)-\frac{x}{n}\Big),\end{aligned}$$

tomando módulo, usando a desigualdade triangular, a estimativa (1.10) e o Teorema 1.3 obtemos

$$\begin{aligned}\Big|\sum_{n\leqslant y}\mu(n)\Psi\left(\frac{x}{n}\right)\Big| &\leqslant x\Big|\sum_{n\leqslant y}\frac{\mu(n)}{n}\Big|+\sum_{n\leqslant y}\Big|\Psi\left(\frac{x}{n}\right)-\frac{x}{n}\Big|\\ &< x+\epsilon\sum_{n\leqslant y}\frac{x}{n} < x+\epsilon x(1+\log y)\\ &< x+\epsilon x+\epsilon x\log x.\end{aligned}$$

Isto conclui a estimativa da primeira soma. Passamos agora para a segunda soma. Agora temos $y < n \leqslant x$ assim $n \geqslant y+1$. Portanto

$$\frac{x}{n}\leqslant\frac{x}{y+1}<A$$

pois $y \leqslant x/A < y+1$. A desigualdade $x/n < A$, implica que $\Psi(\frac{x}{n}) \leqslant \Psi(A)$. Por tanto a segunda soma está limitada por $x\Psi(A)$, logo a soma inteira em (1.7) está limitada por

$$(1+\epsilon)x+\epsilon x\log x+x\Psi(A) < (2+\Psi(A))x+\epsilon x\log x,$$

se $\epsilon < 1$. Em outras palavras, dado algum ϵ tal que $0<\epsilon<1$ temos

$$|H(x)| < (2+\Psi(A))x+\epsilon x\log x, \quad \text{se } x > A$$

ou equivalentemente

$$\frac{|H(x)|}{x\log x} < \frac{2+\Psi(A)}{\log x}+\epsilon$$

Agora escolha $B > A$ tal que $x > B$ implica $(2+\Psi(A))/\log x < \epsilon$. Logo para $x > B$ temos

$$\frac{|H(x)|}{x\log x} < 2\epsilon$$

o que mostra que $H(x)/(x\log x) \to 0$ quando $x \to \infty$. □

A condição necessária do Teorema 1.9, é dada no seguinte teorema.

TEOREMA 1.12. *A relação*

$$M(x) = o(x), \quad \textit{quando } x \to \infty \tag{1.11}$$

implica $\Psi(x) \backsim x$*, quando* $x \to \infty$*.*

Demonstração. Primeiro vamos mostrar que podemos expressar $\Psi(x)$ pela fórmula

$$\Psi(x) = x - \sum_{\substack{q,d \\ qd \leqslant x}} \mu(d) f(q) + O(1), \tag{1.12}$$

onde f é definida no parágrafo seguinte. Em seguida, usando que $M(x) = o(x)$ mostramos que a soma no lado direito de (1.12) é uma função $o(x)$ quando $x \to \infty$.

A função f mencionada acima é dada por $f(n) = \sigma_0(n) - \log n - 2\gamma$, onde γ é a constante de Euler-Mascheroni [1] e $\sigma_0(n) = d(n)$ é o número de divisores de n. Para obter (1.12) consideramos as seguintes identidades

$$[x] = \sum_{n \leqslant x} 1, \quad \Psi(x) = \sum_{n \leqslant x} \Lambda(n), \quad 1 = \sum_{n \leqslant x} \left[\frac{1}{n}\right].$$

Da Fórmula de Inversão de Möbius obtemos as três igualdades

$$1 = \sum_{d|n} \mu(d) \sigma_0\left(\frac{n}{d}\right), \quad \Lambda(n) = \sum_{d|n} \mu(d) \log \frac{n}{d}, \quad \left[\frac{1}{n}\right] = \sum_{d|n} \mu(d).$$

Usando estas seis igualdades mostramos que

$$\begin{aligned}
[x] - \Psi(x) - 2\gamma &= \sum_{n \leqslant x} \left(1 - \Lambda(n) - 2\gamma \left[\frac{1}{n}\right]\right) \\
&= \sum_{n \leqslant x} \sum_{d|n} \mu(d) \left(\sigma_0\left(\frac{n}{d}\right) - \log \frac{n}{d} - 2\gamma\right) \\
&= \sum_{\substack{qd \leqslant x \\ q,d}} \mu(d) \big(\sigma_0(q) - \log q - 2\gamma\big) = \sum_{\substack{qd \leqslant x \\ q,d}} \mu(d) f(q).
\end{aligned}$$

Desta igualdade segue (1.12). Portanto a prova do teorema estará completa se mostramos que

$$\sum_{\substack{qd \leqslant x \\ q,d}} \mu(d) f(q) = o(x), \quad \text{quando } x \to \infty. \tag{1.13}$$

---

[1] A constante de Euler-Mascheroni γ é definida por

$$\gamma = \lim_{n \to \infty} \left[\sum_{m \leqslant n} \frac{1}{m} - \log n\right].$$

Para isto, vamos usar a seguinte identidade: (que pode ser obtida através de um argumento geométrico simples, para maiores detalhes [2], Teorema 3.17.)

$$\sum_{\substack{qd\leqslant x\\ q,d}}\mu(d)f(q)=\sum_{n\leqslant b}\mu(n)F\Big(\frac{x}{n}\Big)+\sum_{n\leqslant a}f(n)M\Big(\frac{x}{n}\Big)-F(a)M(b), \tag{1.14}$$

onde a e b são números reais positivos tais que $ab=x$ e $F(x)=\sum_{n\leqslant x}f(n)$.

Vamos mostrar que $F(x)=O(\sqrt{x})$. De fato, é suficiente usar a fórmula (veja [2], Teorema 3.3)

$$\sum_{n\leqslant x}\sigma_0(n)=x\log x+(2\gamma-1)x+O(\sqrt{x}),$$

e a relação

$$\sum_{n\leqslant x}\log n=\log[x]!=x\log x-x+O(\log x)$$

para concluir que

$$\begin{aligned}F(x)&=\sum_{n\leqslant x}\sigma_0(n)-\sum_{n\leqslant x}\log n-2\gamma\sum_{n\leqslant x}1\\ &=x\log x+(2\gamma-1)x+O(\sqrt{x})-(x\log x-x+O(\log x))-2\gamma x\\ &=O(\sqrt{x})+O(\log x)\\ &=O(\sqrt{x}).\end{aligned}$$

Logo existe uma constante $B>0$ tal que $|F(x)|\leqslant B\sqrt{x}$ para todo $x\geqslant 1$. Usando esta estimativa na primeira soma de (1.14) obtemos

$$\Big|\sum_{n\leqslant b}\mu(n)F\left(\frac{x}{n}\right)\Big|\leqslant B\sum_{n\leqslant b}\sqrt{\frac{x}{n}}\leqslant A\sqrt{xb}=\frac{Ax}{\sqrt{a}}, \tag{1.15}$$

para alguma constante $A>B>0$. Agora seja $\epsilon>0$ e escolha $a>1$ tal que $A/\sqrt{a}<\epsilon$. Usando a desigualdade (1.15) temos que

$$\Big|\sum_{n\leqslant b}\mu(n)F\left(\frac{x}{n}\right)\Big|<\epsilon x,$$

para todo $x\geqslant 1$. Vamos lidar agora com segundo somatório de (1.14). Note que a depende de ϵ e não de x. Já que $M(x)=o(x)$, quando $x\to\infty$, para o mesmo ϵ escolhida acima, podemos afirmar que existe $c>0$ tal que

$$x>c\Rightarrow\frac{|M(x)|}{x}<\frac{\epsilon}{K},$$

onde K é uma constante positiva (que será calculada em breve). Usando estas desigualdades, a segunda soma de (1.14) é limitada por

$$\Big|\sum_{n\leqslant a} f(n)M\left(\frac{x}{n}\right)\Big| \leqslant \sum_{n\leqslant a}|f(n)|\frac{\epsilon}{K}\frac{x}{n} = \frac{\epsilon x}{K}\sum_{n\leqslant a}\frac{|f(n)|}{n},$$

sempre que $x/n > c$, para todo $n \leqslant a$. Portanto a estimativa acima permanece verdadeira para $x > ac$. Tomando

$$K = \sum_{n\leqslant a}\frac{|f(n)|}{n}$$

e substituindo na última desigualdade ficamos com

$$\Big|\sum_{n\leqslant a} f(n)M\big(\frac{x}{n}\big)\Big| < \epsilon x, \quad \text{se } x > ac.$$

O último termo de (1.14) pode ser estimado como segue

$$|F(a)M(b)| \leqslant A\sqrt{a}|M(b)| < A\sqrt{a}b < \epsilon\sqrt{b}\sqrt{a}b = \epsilon\sqrt{x}b < \epsilon x,$$

desde que $\sqrt{x} > a$, ou $x > a^2$. Juntando todas as estimativas concluímos finalmente que

$$\Big|\sum_{\substack{qd\leqslant x\\ q,d}}\mu(d)f(q)\Big| < 3\epsilon x,$$

se $x > a^2$ e $x > ac$, onde a e c dependem de ϵ. Isto finaliza a prova do teorema. $\square$

Em resumo, vimos neste capítulo que o TNP nos garante que $M(x) = o(x)$, ou seja,

$$\sum_{n\leqslant x}\mu(n) = o(x).$$

Este fato é o primeiro resultado de ortogonalidade da função de Möbius, neste caso com respeito a função constante igual a um. No Capítulo 3, veremos outro caso de ortogonalidade da função de Möbius.

Capítulo 2

Teoria de L-Funções

2.1 Definição e Preliminares

É muito difícil atualmente mover-se pela Teoria dos Números sem encontrar a função Zeta ou uma L-função, mas por que elas são tão importantes? Uma boa analogia pode ser feita com matrizes. Em Álgebra Linear uma matriz é a maneira mais canônica de codificar todas as informações de uma transformação linear. Sua importância está no fato que as propriedades das transformações lineares estão relacionadas as propriedades das matrizes e as propriedades destas são fáceis de serem identificadas.

De maneira similar, vários dos objetos de interesse na Teoria dos Números (Corpo dos Números Algébricos, Representações de Galois e Formas Automórficas) podem ter algumas de suas informações codificadas em objetos analíticos chamados de L-funções. Como no caso de matrizes, podemos re-interpretar propriedades dos objetos citados acima em termos das propriedades das L-funções. Além do mais podemos mostrar que as L-funções associadas a uma classe de objetos coincide com aquelas associadas a outros e estas conexões estão relacionadas a vários teoremas profundos do século 20 em Teoria dos Números, incluindo entre eles o Teorema de Modularidade para Curvas Elípticas.

Como já mencionamos, a teoria de L-funções é uma ferramenta da Teoria Analítica dos Números contemporânea que tornou-se muito importante, mesmo que em grande parte sua estrutura teórica ainda seja de caráter conjectural.

Nesta teoria são construídas, entre outras coisas, generalizações da função Zeta de Riemann e as L-séries para carácteres de Dirichlet. Neste texto estes serão os exemplos que usaremos em geral, para ilustrar as propriedades deste tipo de função.

Vamos começar nossa exposição deste assunto pela definição abstrata das L-funções.

Definição 2.1. *Uma L-função é uma função complexa denotada por $L(f,s)$ satisfazendo as seguintes condições 1) e 2)*

1) *Tanto a série de Dirichlet como o produto de Euler de grau $d \geqslant 1$*

$$L(f,s) = \sum_{n\geqslant 1} \lambda_f(n) n^{-s} = \prod_p (1-\alpha_1(p)p^{-s})^{-1} \dots (1-\alpha_d(p)p^{-s})^{-1}, \quad (2.1)$$

com $\lambda_f(1) = 1$, $\lambda_f(n) \in \mathbb{C}$, $\alpha_i(p) \in \mathbb{C}$; devem ser absolutamente convergentes para $Re(s) > 1$. Os $\alpha_i(p)$, $1 \leqslant i \leqslant d$ são chamados de raízes locais ou parâmetros locais de $L(f,s)$ em p, e satisfazem

$$|\alpha_i(p)| < p \text{ para todo } \quad p.$$

2) *Deve existir um inteiro $q(f) \geqslant 1$, chamado o condutor de $L(f,s)$, tal que $\alpha_i(p) \neq 0$ para todo $1 \leqslant i \leqslant d$, sempre que $p \nmid q(f)$. Um primo $p \nmid q(f)$ é chamado de não ramificado.*

Além das condições anteriores deve existir uma função, chamada de Fator Gama, definida por

$$\gamma(f,s) = \pi^{\frac{-ds}{2}} \prod_{j=1}^{d} \Gamma\left(\frac{s+k_j}{2}\right), \quad (2.2)$$

onde $k_j \in \mathbb{C}$ (chamados de parâmetros locais de $L(f,s)$ no infinito) devem ser reais ou aparecer junto com seu conjugado. Além do mais, para todo $j = 1, \dots, d$, devem satisfazer $Re(k_j) > -1$.

Esta última condição nos diz que $\gamma(f,s)$ não tem zeros em $\mathbb{C}$ e não tem polos para $Re(s) \geqslant 1$.

Em seguida definimos o conceito de L-função completa.

Definição 2.2. *Uma L-função completa $\Lambda(f,s)$ associada a $L(f,s)$ é uma função da forma*

$$\Lambda(f,s) = q(f)^{\frac{s}{2}} \gamma(f,s) L(f,s), \quad (2.3)$$

onde $L(f,s)$ é uma L-função, $\gamma(f,s)$, $q(f)$ são o fator gama e o condutor de $L(f,s)$, respectivamente. Além disto exigimos que $\Lambda(f,s)$ admite uma extensão meromorfa, de ordem 1 (veja definição 5.8*) a todo plano cujos polos estão no máximo em $s = 0$ e $s = 1$, e também que satisfaça a equação funcional*

$$\Lambda(f,s) = \varepsilon(f)\Lambda(\overline{f}, 1-s),$$

onde $\overline{f}$ é uma função associada a f (dual de f) mediante as igualdades $\lambda_{\overline{f}}(n) = \overline{\lambda}_f(n)$, $\gamma(\overline{f},s) = \gamma(f,s)$, $q(\overline{f}) = q(f)$ e $\varepsilon(f)$ é um número complexo de valor absoluto 1, chamado de "número principal" de $L(f,s)$.

EXEMPLO 2.1. *A função Zeta de Riemann*

$$\zeta(s) = \sum_{n \geqslant 1} n^{-s} = \prod_p (1 - p^{-s})^{-1}$$

é o exemplo mais clássico de uma L-função de ordem 1, onde $\lambda_\zeta(n) = 1$ *para todo* $n \in \mathbb{N}$ *e* $\alpha_1(p) = 1$ *para todo* p *primo. Além disso, basta tomar* $q(\zeta) = 1$ *e todos os primos são não ramificados. O fator gama é dado por*

$$\gamma(\zeta, s) = \pi^{-s/2}\Gamma\left(\frac{s}{2}\right).$$

Observe que $\varepsilon(\zeta) = 1$ *e a equação funcional da L-função completa neste caso é a conhecida equação funcional da função Zeta de Riemann*

$$\pi^{-\frac{s}{2}}\Gamma\left(\frac{s}{2}\right)\zeta(s) = \pi^{\frac{-(1-s)}{2}}\,\Gamma\left(\frac{1-s}{2}\right)\zeta(1-s).$$

EXEMPLO 2.2. *Uma L-série de Dirichlet é uma função da forma*

$$L(\chi, s) = \sum_{n \geqslant 1} \frac{\chi(n)}{n^s},$$

onde χ *é um caráter de Dirichlet módulo* q *(veja definição* 5.4 *do Apêndice), dado que o caráter é completamente multiplicativo, tem-se que*

$$L(\chi, s) = \sum_{n \geqslant 1} \frac{\chi(n)}{n^s} = \prod_p (1 + \frac{\chi(p)}{p^s} + \frac{\chi(p)^2}{p^{2s}} + \ldots) = \prod_p (1 - \chi(p)p^{-s})^{-1}.$$

Pela definição de caráter temos $\chi(1) = 1$, $|\chi(p)| \leqslant 1 < p$ *e o condutor* $q(\chi) = q$, *pois se* $p \nmid q(\chi)$ *então* $(p, q(\chi)) = 1$ *e* $\chi(p) \neq 0$. *Podemos provar que o fator gama é dado por*

$$\gamma(s) = \pi^{-\frac{s}{2}}\Gamma\left(\frac{s+\delta}{2}\right),$$

onde $\delta = 0$ *se* $\chi(-1) = 1$, *e* $\delta = 1$ *se* $\chi(-1) = -1$ *e o número principal é dado por* $\varepsilon(\chi) = i^{-k}\tau(\chi)/\sqrt{q}$, *onde* $\tau(\chi)$ *é definida na Seção* 5.1, *equação* (5.2). *Para maiores detalhes veja [9], Capítulo 3. Também podemos provar que a equação funcional é dada por*

$$\Lambda(s, \chi) = \varepsilon(\chi)\Lambda(1 - s, \overline{\chi}),$$

veja [9], Teorema 4.15.

Procuramos estimativas para várias quantidades relacionadas com $L(f, s)$. Para uma L-função fixada, em geral, o único parâmetro de interesse é $s \in \mathbb{C}$. Mas quando estamos interessando em estimativas que sejam válidas para um conjunto de L-funções é importante levar em conta o grau, o condutor e os parâmetros locais de cada uma delas.

A maioria dos resultados para $L(f,s)$ são expressos convenientemente em termos do seu condutor analítico dado por

$$\mathfrak{q}(f,s) = q(f)\mathfrak{q}_\infty(s) := q(f)\prod_{j=1}^{d}(|s+k_j|+3). \tag{2.4}$$

Usamos a notação

$$\mathfrak{q}(f) \equiv \mathfrak{q}(f,0) = q(f)\prod_{j=1}^{d}(|k_j|+3).$$

Observamos que $\mathfrak{q}(f) \geqslant q(f)\prod_{j=1}^{d} 3 = 3^d q(f)$ e portanto $d < \log \mathfrak{q}(f)$. Além disso vamos usar com frequência a estimativa

$$\mathfrak{q}(f,s) \leqslant \mathfrak{q}(f)(|s|+3)^d. \tag{2.5}$$

Se $L(f,s)$ e $L(g,s)$ são L-funções segue diretamente da definição que $L(f,s)L(g,s)$ também é uma L-função com condutor $q(f)q(g)$ e condutor analítico $\mathfrak{q}(f,s)\mathfrak{q}(g,s)$, fator gama $\gamma(f,s)\gamma(g,s)$ e número principal $\varepsilon(f)\varepsilon(g)$.

Além disso, $L(\overline{f},s)$ é outra L-função por construção, com o mesmo grau, condutor, fator gama e $\overline{\varepsilon}(f) = \varepsilon(\overline{f})$. Quando $L(f,s)$ é inteira, fixado $t \in \mathbb{R}$ a L-função deslocada dada por $L(g,s) = L(f,s+it)L(\overline{f},s-it)$ tem fator gamma $\gamma(g,s) = \gamma(f,s+it)\gamma(f,s-it)$.

Se $f = \overline{f}$ dizemos que f é auto-dual. Neste caso a serie de Dirichlet da L-função têm coeficientes reais e o número principal é real, e portanto $\varepsilon(f) = \pm 1$. O número $\varepsilon(f)$ é chamado de sinal da equação funcional. Por abuso de notação vamos dizer que $L(f,s)$ é auto-dual quando f o for.

A proposição abaixo é originalmente devido a Shimura e de grande importância em diversas aplicações.

PROPOSIÇÃO 2.1. *Se $L(f,s)$ é auto-dual com $\varepsilon(f) = -1$, então $L(f,\frac{1}{2}) = 0$.*

Demonstração. Tomando $s = 1/2$ na equação funcional e observando que o fator gama é não nulo temos a identidade $L(f,\frac{1}{2}) = -L(f,\frac{1}{2})$, o que implica imediatamente que $L(f,\frac{1}{2}) = 0$. $\square$

Sejam $L(f,s)$ e $L(g,s)$ duas L-funções de grau d e e, com componentes locais no infinito k_i e ν_j, e raízes locais $\alpha_i(p)$ e $\beta_j(p)$, respectivamente. Para cada $p \nmid q(f)q(g)$ fixado, defina

$$L_p(f \otimes g, s) = \prod_{i,j}(1-\alpha_i(p)\beta_j(p)p^{-s})^{-1}.$$

Dizemos que o par de funções f e g tem uma **Convolução de Rankin-Selberg** se existe uma L-função de grau $d \cdot e$, tal que

$$L(f \otimes g, s) = \prod_{p \nmid q(f)q(g)} L_p(f \otimes g, s) \prod_{p \mid q(f)q(g)} H_p(p^{-s}),$$

onde

$$H_p(p^{-s}) = \prod_{j=1}^{de} (1 - \gamma_j(p)p^{-s})^{-1}, \quad \text{com} \quad |\gamma_j(p)| < p,$$

com seu fator gama dado por

$$\gamma(f \otimes g, s) = \pi^{-\frac{des}{2}} \prod_{i,j} \Gamma\left(\frac{s + \mu_{i,j}}{2}\right),$$

onde $\mathrm{Re}(\mu_{i,j}) \leqslant \mathrm{Re}(k_i + \nu_j)$ e $|\mu_{i,j}| \leqslant |k_i| + |\nu_j|$. Além disso, o condutor divide $q(f)^e q(g)^d$ e $L(f \otimes g, s)$ possui um polo em $s = 1$ se $g = \overline{f}$.

Chamamos $L(f \otimes g, s)$ de convolução de Rankin-Selberg de f e g, ou quadrado de Rankin-Selberg se $g = \overline{f}$.

Note que se existem $L(f \otimes f, s)$ ou $L(f \otimes \overline{f}, s)$ as estimativas das raízes locais k_j podem ser melhoradas para $|\alpha_i(p)| < \sqrt{p}$ e $Re(k_j) > -\frac{1}{2}$.

Observando que $q(f \otimes g) | q(f)^e q(g)^d$ segue de (2.5) que

$$\mathfrak{q}(f \otimes g, s) \leqslant \mathfrak{q}(f)^e \mathfrak{q}(g)^d (|s| + 3)^{de}. \tag{2.6}$$

Exemplo 2.3. *Note que a função $\zeta(s)$ é autodual e o quadrado de Rankin-Selberg de $\zeta(s)$ é a própria função $\zeta(s)$. Mais geralmente, dois caráteres de Dirichlet quaisquer χ_1 e χ_2 admitem uma convolução de Rankin-Selberg dada por $L(\chi_1 \otimes \chi_2, s) = L(\chi_3, s)$, onde χ_3 é um caráter primitivo induzido pelo produto $\chi_1\chi_2$. Observamos que temos $\chi_3 = \chi_1\chi_2$ se, por exemplo, os condutores q_i de χ_i são coprimos (Lembre-se que $\mathbb{Z}_{mn} \cong \mathbb{Z}_m \times \mathbb{Z}_n$ se $(m, n) = 1$). Em particular, o quadrado de Rankin-Selberg de $L(s, \chi)$ é $\zeta(s)$.*

Vamos estabelecer agora uma fórmula exata, conhecida como "equação funcional aproximada"que fornece expressões analíticas convenientes para $L(f, s)$ na faixa crítica, onde a série não é absolutamente convergente.

Antes de prosseguir introduzimos mais algumas notações. Seja $f : \mathbb{C} \to \mathbb{C}$ é uma função, denotamos por (caso existam)

$$\int_{(k)} f(z)dz \quad \text{e} \quad \int_{<k>} f(z)dz$$

as integrais complexas ao longo das retas $k = \mathrm{Re}(s)$ e $k = \mathrm{Im}(s)$, respectivamente.

TEOREMA 2.1. *Sejam $L(f,s)$ uma L-função e G uma função par que é holomorfa e limitada na faixa $-4 < Re(u) < 4$ e $G(0) = 1$. Fixado $X > 0$ para todo s pertencente a faixa $0 \leqslant \sigma \leqslant 1$, temos*

$$L(f,s) = \sum_n \frac{\lambda_f(n)}{n^s} V_s\left(\frac{n}{X\sqrt{q}}\right) + \varepsilon(f,s) \sum_n \frac{\overline{\lambda_f}(n)}{n^{1-s}} V_{1-s}\left(\frac{nX}{\sqrt{q}}\right) - R, \tag{2.7}$$

onde $V_s(y)$ é uma função suave definida por

$$V_s(y) = \frac{1}{2\pi i} \int_{(3)} y^{-u} G(u) \frac{\gamma(f,s+u)}{\gamma(f,s)} \frac{du}{u}$$

e

$$\varepsilon(f,s) = \varepsilon(f) q(f)^{\frac{1}{2}-s} \frac{\gamma(f,1-s)}{\gamma(f,s)}.$$

A última parcela é nula, isto é, $R = 0$ se $\Lambda(f,s)$ é inteira, caso contrário

$$R = \left(\underset{u=1-s}{res} + \underset{u=-s}{res} \right) \frac{\Lambda(f,s+u)}{q^{s/2}\gamma(f,s)} \frac{G(u)}{u} X^u.$$

Demonstração. Considere a integral de Lebesgue

$$I(X,f,s) = \frac{1}{2\pi i} \int_{(3)} X^u \Lambda(f,s+u) \frac{G(u)}{u}\, du,$$

Afirmação: a integral $I(X,f,s)$ existe e é finita. De fato, como $\mathrm{Re}(u) = 3$ tem-se que $G(u)$ é limitada, pela definição de L-função completa temos

$$\Lambda(f,s+u) = q(f)^{s/2} \gamma(f,s+u) L(f,s+u),$$

sendo $L(f,s)$ é uniformemente convergente para $Re(s) \geqslant 1+\delta$, $\delta > 0$. Resta estudar o comportamento de $\gamma(f,s)$ para σ fixada e $t \to \infty$. Pela definição de $\gamma(f,s)$ em (2.2) basta analisar o crescimento de $\Gamma(x+iy)$ quando $y \to \infty$. Usando a Fórmula de Stirling [1]

$$\Gamma(x+iy) = \sqrt{\frac{2\pi}{x+iy}} \Big(\frac{x+iy}{e}\Big)^{x+iy} \cdot O(1),$$

para todo y suficientemente grande. Usando o simbolo $f \ll g$ de forma equivalente a $f = O(g)$ temos

$$\begin{aligned}
(x+iy)^{x+iy} &= \exp((x+iy)\log(x+iy)) \\
&= \exp\big((x+iy)[\log(x^2+y^2) + i\arg(x+iy)]\big) \\
&\ll \exp\big(x\log(x^2+y^2) - y\arg(x+iy)\big) \\
&= \exp\left(-y\arctan(y/x)\right)\exp\big(x\log(x^2+y^2)\big) \\
&\ll \exp\left(-y\arctan(y/x)\right) y^A, \quad A = A(x).
\end{aligned}$$

[1] Veja apêndice, proposição 5.1

Como $\arctan\,(y/x) \ll \frac{\pi}{2}$ quando y cresce, logo $\Gamma(x+iy) \to 0$ quando $y \to \infty$ daí pelo teorema da Convergência Dominada de Lebesgue temos que a integral existe.

Para todo $T > 0$, podemos usar o Teorema dos Resíduos para transladar a integração da reta $\mathrm{Re}(s) = 3$ para a reta $\mathrm{Re}(s) = -3$, como segue

$$\int\limits_{(3)} D + \int\limits_{<T>} D + \int\limits_{(-3)^-} D + \int\limits_{<-T>^-} D = 2\pi i \left(\underset{u=0}{\mathrm{res}} + \underset{u=1-s}{\mathrm{res}} + \underset{u=-s}{\mathrm{res}} \right) D, \tag{2.8}$$

onde $(a)^-$ indica a integração sobre a reta $\mathrm{Re}(s) = a$ percorrida no sentido de cima para baixo (analogamente para $< a >^-$) e

$$D := D(u) = X^u \Lambda(f, s+u) \frac{G(u)}{u}.$$

Nosso próximo passo é obter estimativas para

$$M = \int\limits_{<T>} D(u)\ du.$$

Seja $u = v + iw$, $s = \sigma + it$, como $G(u)$ é limitada em $-4 \leqslant Re(u) \leqslant 4$ então

$$\begin{aligned} M &= \int_{-3}^{3} X^{v+iT} \Lambda(f, (\sigma+v) + i(t+T)) G(v+it) \frac{dv + idT}{v+iT} \\ &\ll \int_{-3}^{3} X^{v} \Lambda(f, (\sigma+v) + i(t+T)) \frac{dv}{v+iT} \\ &= \int_{-3}^{3} \frac{v-iT}{v^2+T^2} X^{v} \Lambda(f, (\sigma+v) + i(t+T)) dv. \end{aligned}$$

Já que a L-função completa $\Lambda(f,s)$, para todo $T \neq 0$ fixado, é uniformemente limitada nos segmentos de reta $v \pm iT$, com $-3 \leqslant v \leqslant 3$, segue que

$$M \ll \int_{-3}^{3} \frac{v-iT}{v^2+T^2} dv = \frac{i}{T} \arctan \frac{3}{T}.$$

Desta forma $M \to 0$ quando $T \to \infty$. O mesmo acontece com a última integral no lado esquerdo de (2.8). Assim

$$\int\limits_{(3)} D + \int\limits_{(-3)^-} D = 2\pi i \left(\underset{u=0}{\mathrm{res}} + \underset{u=1-s}{\mathrm{res}} + \underset{u=-s}{\mathrm{res}} \right) D. \tag{2.9}$$

Passamos agora ao cálculo dos resíduos. Como $s = 0$ é um polo simples de D temos que
$$\operatorname*{res}_{u=0} D = \lim_{u\to 0} X^u \Lambda(f, s+u) G(u) = \Lambda(f, s).$$

Além disso, se $\Lambda(f,s)$ é inteira, os dois últimos resíduos são zero e logo $R = 0$. Caso $\Lambda(f,s)$ não seja inteira, usando (2.9), a equação acima e a definição de R, temos que

$$I(X,f,s) + \frac{1}{2\pi i} \int\limits_{(-3)^-} X^u \Lambda(f, s+u) \frac{G(u)}{u}\, du = \Lambda(f,s) + Rq(f)^{s/2}\gamma(f,s).$$

Fazendo a mudança de variáveis $u = -v$, obtemos

$$I(X,f,s) + \frac{1}{2\pi i} \int\limits_{(3)} X^{-u} \Lambda(f, s-u) \frac{G(u)}{u}\, du = \Lambda(f,s) + Rq(f)^{s/2}\gamma(f,s) \qquad (2.10)$$

usando a equação funcional para $\Lambda(f, s-u)$, segue que a integral acima satisfaz a seguinte identidade

$$\begin{aligned}\frac{1}{2\pi i} \int\limits_{(3)} X^{-u} \Lambda(f, s-u) \frac{G(u)}{u}\, du &= \frac{\varepsilon(f)}{2\pi i} \int\limits_{(3)} X^{-u} \Lambda(\overline{f}, 1-s+u) \frac{G(u)}{u}\, du \\ &= \varepsilon(f) I(X^{-1}, \overline{f}, 1-s).\end{aligned}$$

Substituindo esta última igualdade em (2.10) ficamos com

$$\Lambda(f,s) = I(X,f,s) + \varepsilon(f) I(X^{-1}, \overline{f}, 1-s) - Rq(f)^{s/2}\gamma(f,s). \qquad (2.11)$$

Expandindo em séries de Dirichlet (absolutamente convergentes), e usando a notação $q = q(f)$, temos

$$\begin{aligned}I(X,f,s) &= \frac{1}{2\pi i} \int\limits_{(3)} X^u \Lambda(f, s+u) G(u) \frac{du}{u} \\ &= \frac{1}{2\pi i} \int\limits_{(3)} X^u q^{\frac{s}{2}+\frac{u}{2}} \gamma(f, s+u) \sum_{n\geqslant 1} \lambda_f(n) n^{-s-u} G(u) \frac{du}{u} \\ &= q^{\frac{s}{2}} \sum_{n\geqslant 1} \lambda_f(n) n^{-s} \frac{1}{2\pi i} \int\limits_{(3)} X^u \gamma(f, s+u) \left(\frac{X\sqrt{q}}{n}\right)^u G(u) \frac{du}{u} \\ &= q^{\frac{s}{2}} \sum_{n\geqslant 1} \lambda_f(n) n^{-s} \gamma(f,s) \frac{1}{2\pi i} \int\limits_{(3)} X^u \frac{\gamma(f, s+u)}{\gamma(f,s)} \frac{G(u)}{u} \left(\frac{n}{X\sqrt{q}}\right)^{-u} du.\end{aligned}$$

Fazendo uso da definição de V_s na última expressão da igualdade acima obtemos a igualdade

$$I(X,f,s) = q^{\frac{s}{2}} \sum_{n\geqslant 1} \lambda_f(n) n^{-s} \gamma(f,s) V_s\left(\frac{n}{X\sqrt{q}}\right).$$

Analogamente, para $I(X^{-1}, \overline{f}, 1-s)$ temos

$$\begin{aligned}
I(X^{-1},\overline{f},1-s) &= \frac{1}{2\pi i}\int_{(3)} X^{-u}\Lambda(f,1-s+u)G(u)\frac{du}{u} \\
&= \frac{1}{2\pi i}\int_{(3)} X^{u} q^{\frac{1-s}{2}+\frac{u}{2}}\gamma(f,1-s+u)\sum_{n\geqslant 1}\frac{\overline{\lambda_f}(n)}{n^{1-s+u}}G(u)\frac{du}{u} \\
&= q^{\frac{1-s}{2}}\sum_{n\geqslant 1}\frac{\overline{\lambda_f}(n)}{n^{1-s}}\gamma(f,1-s)V_{1-s}\left(\frac{nX}{\sqrt{q}}\right) \\
&= \frac{\varepsilon(f,s)}{\varepsilon(f)} q^{s/2}\gamma(f,s)\sum_{n\geqslant 1}\frac{\overline{\lambda_f}(n)}{n^{1-s}}.
\end{aligned}$$

Portanto da igualdade (2.11) obtemos

$$\Lambda(f,s) = q^{\frac{s}{2}}\gamma(f,s)\left[\sum_n \frac{\lambda_f(n)}{n^s}V_s\left(\frac{n}{X\sqrt{q}}\right) + \varepsilon(f,s)\sum_n \frac{\overline{\lambda_f}(n)}{n^{1-s}}V_{1-s}\left(\frac{nX}{\sqrt{q}}\right) - R\right].$$

Dividindo ambos os membros da equação acima por $q^{s/2}\gamma(f,s)$, finalmente obtemos (2.7). $\square$

2.2 Contando zeros de L-funções

Um dos assuntos mais profundos da teoria de L-funções é a distribuição dos zeros de $L(f,s)$.

Vamos iniciar esta seção com um teorema que dá informações sobre a localização dos zeros de uma L-função e que também mostra que a sequência dos recíprocos tem norma $\ell^{1+\varepsilon}(\mathbb{N})$ finita.

Lema 2.1. *Seja $L(f,s)$ seja uma L-função. Todos os zeros ρ's de $\Lambda(f,s)$ estão na faixa crítica $0 \leqslant \sigma \leqslant 1$. Além do mais para todo $\epsilon > 0$ temos*

$$\sum_{\rho\neq 0,1} |\rho|^{-1-\epsilon} < +\infty.$$

Demonstração. Dado que a expansão em produto de Euler de $L(f,s)$ é absolutamente convergente e $\gamma(f,s)$ é não nulo para $\mathrm{Re}(s) > 1$, então pela definição de

L-função completa não existem zeros de $\Lambda(f,s)$ nesta região. Invocando a equação funcional, podemos observar que o mesmo ocorre para $\mathrm{Re}(s) < 0$.

Para provar a segunda afirmação basta mostrar que o expoente de convergência κ do conjunto de zeros ρ's da nossa L-função, definido por

$$\kappa = \inf\left\{ a \in (0,\infty) : \sum_{\rho\neq 0} |\rho|^{-a} < \infty \right\},$$

é menor ou igual do que um.

Para isto consideramos a função inteira $h(s) = \Lambda(f,s)(s)^r(s-1)^r$, onde r é a ordem do polo em $s=1$ de $\Lambda(f,s)$ (que é a mesma em $s=0$ pela equação funcional).

Note que h tem ordem um e seus zeros diferentes de 0 e 1 são também zeros de $\Lambda(f,s)$. Seguindo as mesmas ideias da prova do Teorema de Fatoração de Hadamard (veja Apêndice, Teorema (5.3)), concluímos que $\kappa \leq 1$. □

Note que se $\rho = \beta + i\gamma$ é zero de $\Lambda(f,s)$, então $\overline{\rho}$ é um zero do dual $\Lambda(\overline{f},s)$. Portanto, pela equação funcional, os pontos refletidos na linha crítica $\rho^* = 1-\overline{\rho} = 1-\beta+i\gamma$ é um zero de $\Lambda(f,s)$ (com sua correspondente multiplicidade).

É claro que não existe simetria entre os zeros triviais de $L(f,s)$ que vêm dos polos de $\Gamma(\frac{s+k_j}{2})$.

Daqui em diante usamos a notação log para denotar o ramo principal do logaritmo complexo. As próximas proposições (elementares em análise complexa) nos fornecem dois resultados que serão úteis no restante de todo este capítulo.

PROPOSIÇÃO 2.2. *Se $z_1,\dots,z_k$ são números complexos tais que $Re(z_j) > 0$ então*

$$\log(z_1\cdot\ldots\cdot z_k) = \log(z_1) + \ldots + \log(z_k).$$

PROPOSIÇÃO 2.3. *Sejam $U \subset \mathbb{C}$ aberto, $f_1,\dots,f_n : U \to \mathbb{C}$ funções holomorfas tais que $f_j(z) \neq 0$ para todo $z \in U$ e $j = 1,\dots,n$. Se $f = f_1\cdot f_2\cdot\ldots\cdot f_n$ então*

$$\frac{f'}{f} = \sum_{j=1}^{n} \frac{f_j'}{f_j}.$$

TEOREMA 2.2. *Seja $L(f,s)$ seja uma L-função. Existem constantes $a = a(f)$ e $b = b(f)$ tais que*

$$(s(1-s))^r\Lambda(f,s) = e^{a+bs}\prod_{\rho\neq 0,1}\big(1-\frac{s}{\rho}\big)e^{s/\rho}. \tag{2.12}$$

Portanto

$$-\frac{L'}{L}(f,s) = \frac{1}{2}\log q + \frac{\gamma'}{\gamma}(f,s) - b + \frac{r}{s} + \frac{r}{s-1} - \sum_{\rho\neq 0,1}\big(\frac{1}{s-\rho} + \frac{1}{\rho}\big). \tag{2.13}$$

Ambas expressões são absolutamente convergentes em subconjuntos compactos que não têm zeros ou polos (lembre-se que r é a ordem do polo de $L(f,s)$ em $s=1$).

Demonstração. Dado que a função $\Lambda(f,s)$ tem no máximo polos em 0 e 1 então $(s(1-s))^r\Lambda(f,s)$ é uma função inteira, daí pelo Teorema de Fatoração de Hadamard temos

$$h(s) \equiv (s(1-s))^r\Lambda(f,s) = e^{g(s)}P(s),$$

pois $h(s)$ não tem zeros em $s=0$. Por definição de L-função completa, temos que $\Lambda(f,s)$ é de ordem 1 o que implica que h é de ordem 1. Logo o gênero de h satifaz $\mu \leqslant 1$, grau$(g) \leqslant 1$ e o posto de h satisfaz $\delta \leqslant 1$; donde concluímos (2.12).

Vamos provar agora que (2.12) implica (2.13). Considere s tal que $0 < \mathrm{Re}(s) < 1$ e $|s| < |\rho_{min}|$, onde $0 < |\rho_{min}| \leqslant |\rho|$ para toda raiz de $L(f,s)$. Então temos de (2.12) que

$$\frac{\Lambda(f,s)}{e^{a+bs}} = \frac{\prod_{\rho\neq 0,1}\left(1-\frac{s}{\rho}\right)e^{s/\rho}}{[s(1-s)]^r}.$$

Observe que todos os fatores do lado direito da igualdade acima têm parte real positiva. Logo podemos tomar o logaritmo em ambos lados e usar a Proposição 2.2 para concluir que

$$\log\left(\frac{\Lambda(f,s)}{e^{a+bs}}\right) = \sum_{\rho\neq 0,1}\left[\log\left(1-\frac{s}{\rho}\right)+\frac{s}{\rho}\right] - r(\log s + \log(1-s)). \tag{2.14}$$

Derivando os dois lados da equação acima com respeito à variável s, e aplicando a Proposição 2.3 ficamos com

$$-\frac{L'}{L}(f,s) = \frac{1}{2}\log q + \frac{\gamma'}{\gamma}(f,s) - b + \frac{r}{s} + \frac{r}{s-1} - \frac{d}{ds}\sum_{\rho\neq 0,1}\left[\log\left(1-\frac{s}{\rho}\right)+\frac{s}{\rho}\right]. \tag{2.15}$$

Próximo passo é mostrar que podemos trocar a ordem da soma com a derivada na igualdade acima. Para isto vamos usamos a seguinte desigualdade de interpolação:

$$\left|\log\left(1-\frac{s}{\rho}\right)+\frac{s}{\rho}\right| = \left|\int_\beta \frac{1}{z-\rho}+\frac{1}{\rho}\,dz\right| \leqslant \int_\beta \left|\frac{1}{z-\rho}+\frac{1}{\rho}\right|\,|dz|,$$

onde β é um caminho unindo os pontos 0 e s, suave por partes e que não intercepta os zeros ρ's. Agora observamos que se $s\in K$, onde K é um compacto no interior da faixa crítica que não possui os zeros de $L(f,s)$, então podemos afirmar que o integrando ao longo de β satisfaz a seguinte estimativa:

$$\left|\frac{1}{z-\rho}+\frac{1}{\rho}\right| = \frac{|z|}{|\rho|^2\,|(z/\rho)-1|} \ll \frac{1}{|\rho|^2}. \tag{2.16}$$

O que mostra que

$$\sum_{\rho\neq 0,1}\left|\log\left(1-\frac{s}{\rho}\right)+\frac{s}{\rho}\right| \ll \sum_{\rho\neq 0,1}\frac{1}{|\rho|^2}$$

Do Lema 2.1 segue que o lado direito da estimativa acima é finito e portanto podemos concluir do Teste-M de Weirstrass que o lado esquerdo converge uniforme e absolutamente nos compactos no interior da faixa crítica. E assim concluímos que podemos trocar a ordem da soma com a derivada em (2.15). Obtendo desta forma igualdade (2.13) nestes compactos. Para finalizar a prova basta usar a estimativa (2.16) para mostrar que o lado direito de (2.13) pode ser estendido analiticamente a todo o plano menos os zeros e polos de $L(f,s)$. Com a série que aparece em (2.13) convergindo absolutamente nas partes compactas do domínio acima. $\square$

Supondo que a derivada logarítmica de $L(f,s)$ possui expansão em série de Dirichlet para $s \neq \rho$ podemos escrever

$$-\frac{L'}{L}(f,s) = \sum_{n\geqslant 1} \Lambda_f(n) n^{-s}, \tag{2.17}$$

para alguma sequência $(\Lambda_f(n))$. Esta notação é inspirada no Lema 1.3 e será reservada neste texto para o caso em que os coeficientes $\Lambda_f(n)$ são suportados nas potências de primos. Neste caso os coeficientes da série de Dirichlet, em termos das raízes locais $\alpha_i(p)$, é dado por

$$\Lambda_f(p^k) = \sum_{j=1}^{d} \alpha_j(p)^k \log p. \tag{2.18}$$

De fato, segue da definição de L-função que

$$L(f,s) = \prod_p \prod_{j=1}^{d} (1-\alpha_j(p)p^{-s})^{-1}.$$

Já que $|\alpha_j(p)| < p$ temos que $(1-\alpha_j(p)p^{-s})^{-1}$ tem parte real positiva. Portanto podemos tomar o logaritmo na igualdade acima e usar a Proposição 2.2 para ver que

$$\log L(f,s) = \sum_p \sum_{j=1}^{d} -\log(1-\alpha_j(p)p^{-s}).$$

Pela definição de L-função podemos concluir que a convergência acima é absoluta e uniforme nas partes compactas de $\{s \in \mathbb{C} : \mathrm{Re}(s) > 1\}$. Derivando com respeito a variável s e usando que $|\alpha_j(p)p^{-s}| < p^{1-\sigma} < 1$, temos

$$\begin{aligned} -\frac{L'}{L}(f,s) &= \sum_p \sum_{j=1}^{d} \frac{\alpha_j(p)p^{-s}\log p}{1-\alpha_j(p)p^{-s}} = \sum_p \sum_{j=1}^{d} \sum_{k=1}^{\infty} \alpha_j^k(p)(p^k)^{-s} \log p \\ &= \sum_{k=1}^{\infty} \sum_p \left(\sum_{j=1}^{d} \alpha_j^k(p) \log p \right) (p^k)^{-s}. \end{aligned} \tag{2.19}$$

o que queríamos provar.

Para caráteres de Dirichlet, vamos mostrar que $\Lambda_\chi(n) = \chi(n)\Lambda(n)$. De fato, observe que $|\chi(n)\Lambda(n)| \leqslant \log n$, garante a convergência absoluta da série de Dirichlet com coeficientes $\chi(n)\Lambda(n)$, para $Re(s) > 1$ e assim podemos usar o produto de Series de Dirichlet para ver que

$$\begin{aligned} L(\chi,s)\sum_{n\geqslant 1}\frac{\chi(n)\Lambda(n)}{n^s} &= \sum_{m\geqslant 1}\frac{\chi(m)}{m^s}\sum_{n\geqslant 1}\frac{\chi(n)\Lambda(n)}{n^s} = \sum_{l\geqslant 1}\frac{\chi(l)}{l^s}\sum_{n|l}\Lambda(n) \\ &= \sum_{l\geqslant 1}\frac{\chi(l)\log l}{l^s} \\ &= -L'(\chi,s). \end{aligned}$$

e portanto

$$-\frac{L'(\chi,s)}{L(\chi,s)} = \sum_{n\geqslant 1}\frac{\chi(n)\Lambda(n)}{n^s}.$$

Em particular, para a função ζ temos $\chi(n) \equiv 1$ e usando o resultado acima podemos ver que $\Lambda_\zeta(n) = \Lambda(n)$.

Agora passamos a análise da distribuição dos zeros das L-funções.

Proposição 2.4. *Seja $L(f,s)$ uma L-função de grau $d \geqslant 1$. Denote por ρ's os zeros de $\Lambda(f,s)$ diferente de $0, 1$.*

(1) O número de zeros $\rho = \beta + i\gamma$ tais que $|\gamma - T| \leqslant 1$, é chamado de $m(T,f)$ e satisfaz

$$m(T,f) \ll \log \mathfrak{q}(f,iT), \tag{2.20}$$

com uma constante absoluta.

(2) Para todo s na faixa $\frac{1}{2} \leqslant \sigma \leqslant 2$, temos

$$\frac{L'}{L}(f,s) + \frac{r}{s} + \frac{r}{s-1} - \sum_{|s+k_j|<1}\frac{1}{s+k_j} - \sum_{|s-\rho|<1}\frac{1}{s-\rho} \ll \log \mathfrak{q}(f,s), \tag{2.21}$$

com uma constante absoluta.

(3) A constante $b(f)$ em (2.12) satisfaz

$$Re(b(f)) = -\sum_\rho Re(\rho^{-1}). \tag{2.22}$$

Demonstração. Vamos provar, em primeiro lugar, a equação (2.22). Multiplicando a equação funcional por $(s(1-s))^r$ temos

$$(s(1-s))^r\,\Lambda(f,s) = \varepsilon(f)\,(s(1-s))^r\,\Lambda(\overline{f},1-s).$$

Pela primeira parte do Teorema 2.2, podemos reescrever a equação acima como segue

$$e^{a(f)+b(f)s}\prod_{\rho\neq 0,1}\left(1-\frac{s}{\rho}\right)e^{s/\rho}=\varepsilon(f)e^{a(\overline{f})+b(\overline{f})(1-s)}\prod_{\rho\neq 0,1}\left(1-\frac{1-s}{\overline{\rho}}\right)e^{(1-s)/\overline{\rho}}.$$

Aplicando logaritmo, seguindo as mesmas ideias da demonstração do Teorema anterior, temos que o lado esquerdo da igualdade acima é dado por

$$a+bs+\sum_{\rho\neq 0,1}\left[\log\left(1-\frac{s}{\rho}\right)+\frac{s}{\rho}\right] \tag{2.23}$$

enquanto que o lado direito é dado por

$$\log\varepsilon+a(\overline{f})+\overline{b}(1-s)+\sum_{\rho\neq 0,1}\left[\log\left(1-\frac{1-s}{\overline{\rho}}\right)+\frac{1-s}{\overline{\rho}}\right]. \tag{2.24}$$

De fato, na última equação usamos que $\overline{b}=b(\overline{f})$. Para mostrar isso basta pegar a segunda igualdade no Teorema 2.2 para f e $\overline{f}$, fazer o conjugado na segunda equação e subtrair da primeira.

Lembrando que (2.23) é igual (2.24) e derivando estas duas expressões com respeito de s ficamos com

$$b+\sum_{\rho\neq 0,1}\left[\frac{\rho}{\rho-s}\cdot\frac{-1}{\rho}+\frac{1}{\rho}\right]=-\overline{b}+\sum_{\rho\neq 0,1}\left[\frac{\overline{\rho}}{\overline{\rho}-1+s}\cdot\frac{1}{\overline{\rho}}-\frac{1}{\overline{\rho}}\right].$$

Agrupando a somas, obtemos

$$2\ \mathrm{Re}(b(f))=b(f)+b(\overline{f})=-\sum_{\rho\neq 0,1}\left[\frac{1}{s-\rho}+\frac{1}{1-s-\overline{\rho}}+\frac{1}{\rho}+\frac{1}{\overline{\rho}}\right]. \tag{2.25}$$

Usando a desigualdade triangular, a cota $0\leqslant Re(\rho)\leqslant 1$ e o fato que a soma acima é absolutamente convergente para $s\in K$, com K compacto

$$\left|\frac{1}{s-\rho}+\frac{1}{1-s-\overline{\rho}}\right|=\frac{|1-2\ \mathrm{Re}(\rho)|}{|s-\rho||1-s-\overline{\rho}|}\leqslant\frac{3}{|s-\rho||1-s-\overline{\rho}|}\ll|\rho|^{-2},$$

assim $(s-\rho)^{-1}+(1-s-\overline{\rho})^{-1}\ll|\rho|^{-2}$ para $s\neq\rho$, e similarmente $\rho^{-1}+\overline{\rho}^{-1}\ll|\rho|^{-2}$. Desta forma as series

$$\sum_{\rho\neq 0,1}\left[\frac{1}{s-\rho}+\frac{1}{1-s-\overline{\rho}}\right]\quad\text{e}\quad\sum_{\rho\neq 0,1}\left[\frac{1}{\rho}+\frac{1}{\overline{\rho}}\right]$$

são, usando Teste-M de Weierstrass, absolutamente convergentes pelo Lema 2.1.

Desta forma podemos decompor o somatório em (2.25) e a pela observação sobre os zeros feita antes da Proposição 2.2 podemos ver que o primeiro somatório acima é nulo e juntando todos fatos temos

$$2\mathrm{Re}(b(f)) = -\sum_{\rho\neq 0,1}\left[\frac{1}{\rho}+\frac{1}{\bar{\rho}}\right] = -2\sum_{\rho\neq 0,1}\mathrm{Re}\left(\frac{1}{\rho}\right).$$

Em seguida, vamos mostrar (2.20). Sejam $T \geqslant 2$ e $s = 3 + iT$. Da equação (2.18), temos que

$$\Lambda_f(n) := \begin{cases} \sum_{j=1}^{d} \alpha_j(p)^k \log p, & \text{se } n = p^k; \\ 0, & \text{caso contrário.} \end{cases}$$

Usando a desigualdade triangular junto com $|\alpha_j(p)| < p$ e que $n = p^k$ temos

$$|\Lambda_f(n)| = \left|\sum_{j=1}^{d} \alpha_j(p)^k \log p\right| \leqslant \sum_{j=1}^{d} p^k \log p = d\ n \log p \leqslant d\ n \log n.$$

Portanto segue da desigualdade triangular, da estimativa acima, o fato $s = 3 + iT$ e desigualdade que aparece acima da equação (2.5) temos que

$$\begin{aligned} \left|\frac{L'}{L}(f,s)\right| &= \left|\sum_{n\geqslant 1} \Lambda_f(n) n^{-s}\right| \leqslant \sum_{n\geqslant 1} d\ n \log n |n^{-s}| = \sum_{n\geqslant 1} d \log n\ n^{1-3} \\ &= \sum_{n\geqslant 1} n^{-2}\ d \log n = d\zeta'(2) \ll \log \mathfrak{q}(f). \end{aligned} \tag{2.26}$$

Pela fórmula de Stirling, para $-1/2 \leqslant \sigma \leqslant 2$, temos que

$$\frac{1}{2}\log q + \frac{\gamma'}{\gamma}(f,s) \ll \log \mathfrak{q}(f,s),$$

(veja [9] pag. 151). Observe que para todo zero da forma $\rho = \beta + i\gamma$ temos

$$\mathrm{Re}\left(\frac{1}{s-\rho}\right) = \frac{3-\beta}{(3-\beta)^2 + (T-\gamma)^2} \quad \text{e} \quad 0 < \beta < 1$$

então

$$\frac{2}{9+(T-\gamma)^2} < \mathrm{Re}\left(\frac{1}{s-\rho}\right) < \frac{3}{4+(T-\gamma)^2}.$$

Tomando a parte real de 2.13 e usando 2.22 obtemos

$$\begin{aligned} -\mathrm{Re}\left(\frac{L'}{L}(f,s)\right) &= \frac{1}{2}\log q + \mathrm{Re}\left(\frac{\gamma'}{\gamma}(f,s)\right) + \sum_{\rho}\mathrm{Re}\left(\rho^{-1}\right) + \frac{3r}{9+T^2} + \frac{2r}{4+T^2} \\ &\quad - \sum_{\rho}\mathrm{Re}\left(\frac{1}{s-\rho}+\frac{1}{\rho}\right). \end{aligned}$$

Então das estimativas acima

$$\sum_{\rho} \operatorname{Re}\left(\frac{1}{s-\rho}\right) \ll \log \mathfrak{q}(f, iT) + \log \mathfrak{q}(f) \ll \mathfrak{q}(f, iT).$$

Portanto

$$\sum_{\rho} \frac{1}{1+(T-\gamma)^2} \ll \log \mathfrak{q}(f, iT), \tag{2.27}$$

dai, se $|\gamma - T| \leqslant 1$, temos

$$\frac{1}{2} m(T, f) \leqslant \sum_{\rho} \frac{1}{1+(T-\gamma)^2} \ll \log \mathfrak{q}(f, iT)$$

donde segue (2.20). Para provar (2.21), escrevemos $s = \sigma + it$ e de (2.26) tem-se

$$-\frac{L'}{L}(f, s) = -\frac{L'}{L}(f, s) + \frac{L'}{L}(f, 3+it) + O(\log \mathfrak{q}(f, s)).$$

Em seguida, aplicando (2.13)

$$\begin{aligned} -\frac{L'}{L}(f, s) = \frac{1}{2}\log q + \frac{\gamma'}{\gamma}(f, s) - b + \frac{r}{s} + \frac{r}{s-1} - \sum_{\rho}\left(\frac{1}{s-\rho} + \frac{1}{\rho}\right) - \frac{1}{2}\log q + b \\ - \frac{\gamma'}{\gamma}(f, 3+it) - \frac{r}{3+it} - \frac{r}{2+it} + \sum_{\rho}\left(\frac{1}{3+it-\rho} + \frac{1}{\rho}\right) + O(\log \mathfrak{q}(f, s)). \end{aligned}$$

Dai podemos obter a seguinte igualdade

$$-\frac{L'}{L}(f, s) = \frac{\gamma'}{\gamma}(f, s) + \frac{r}{s} + \frac{r}{s-1} - \sum_{\rho}\left(\frac{1}{s-\rho} - \frac{1}{3+it-\rho}\right) + O(\log \mathfrak{q}(f, s)). \tag{2.28}$$

No somatório separamos os zeros com $|s-\rho| < 1$ e estimamos o restante por $\log \mathfrak{q}(f, s)$, já que para $|s-\rho| \geqslant 1$

$$\left|\frac{1}{s-\rho} - \frac{1}{3+it-\rho}\right| \leqslant 1 + \frac{1}{|3+it-\rho|} \leqslant \frac{4}{1+(T-\gamma)^2}$$

assim podemos usar a estimativa em (2.27).

Além do mais, no outro caso podemos reduzir os termos na soma, isto é

$$\sum_{|s-\rho|<1}\left(\frac{1}{s-\rho} - \frac{1}{3+it-\rho}\right) = \sum_{|s-\rho|<1}\left(\frac{1}{s-\rho}\right)$$

pois o segundo termo na soma não satisfaz a condição $|s-\rho| < 1$. Substituindo em (2.28)

$$-\frac{L'}{L}(f, s) = \frac{\gamma'}{\gamma}(f, s) + \frac{r}{s} + \frac{r}{s-1} - \sum_{|s-\rho|<1} \frac{1}{s-\rho} + O(\log \mathfrak{q}(f, s)).$$

Agora por definição de $\gamma(f,s)$ e o fato que $\Gamma(s+1) = s\Gamma(s)$ segue que

$$\gamma(f,s) = \pi^{-ds/2} \prod_{j=1}^{d} \frac{2\Gamma\left(1+\frac{s+k_j}{2}\right)}{s+k_j}$$

aplicando derivada logarítmica

$$\begin{aligned}
\frac{\gamma'}{\gamma}(f,s) &= -\frac{d}{2}\log\pi - \sum_j \frac{1}{s+k_j} + \sum_j \frac{\Gamma'}{\Gamma}\left(1+\frac{s+k_j}{2}\right) \\
&= -\frac{d}{2}\log\pi - \sum_{|s+k_j|<1} \frac{1}{s+k_j} - d + \sum_j \frac{\Gamma'}{\Gamma}\left(1+\frac{s+k_j}{2}\right) \\
&= -\sum_{|s+k_j|<1} \frac{1}{s+k_j} + O(\log \mathfrak{q}_\infty(s)).
\end{aligned}$$

Para ver que a última igualdade é verdadeira basta usar a Fórmula de Stirling, para obter assintoticamente a seguinte igualdade

$$\frac{\Gamma'}{\Gamma}(z) = \log z - \frac{1}{2z}$$

o que implica que

$$\sum_j \frac{\Gamma'}{\Gamma}\left(1+\frac{s+k_j}{2}\right) \ll \sum_j \log(|s+k_j|+3) = \log \mathfrak{q}_\infty(s).$$

Portanto (2.21) segue. □

2.3 A Região livre de zeros

No estudo das L-funções precisamos com frequência fazer cálculos ou estimativas de integrais complexas numa região livre de zeros de $L(f,s)$ para várias f. Sabemos que todos os zeros não-triviais das L-funções estão localizados na faixa crítica $0 \leqslant \mathrm{Re}(s) \leqslant 1$. Nesta seção sob certas condições, determinamos uma sub-região da faixa crítica que é livre de zeros.

Na próxima seção usamos o nosso conhecimento sobre ausência de zeros nesta região para obter aproximações assintóticas para a função acumulativa dos coeficientes da série de Dirichlet para a derivada logarítmica de $L(f,s)$ (Teorema de Primo Generalizado).

No que segue apresentamos dois resultados que nos dão a localização de alguns zeros reais de uma L-função e também nos mostra uma nova região do plano complexo livre de zeros.

LEMA 2.2. *Seja $L(f,s)$ uma L-função de grau d com $Re(\Lambda_f(n)) \geqslant 0$ para $(n, q(f)) = 1$. Suponha que em primos ramificados $|\alpha_j(p)| \leqslant p/2$. Então $L(f,1) \neq 0$. Além disso, se $r \geqslant 0$ é a ordem do polo de $L(f,s)$ em $s=1$, então existe uma constante $c > 0$ tal que $L(f,s)$ tem no máximo r zeros reais no intervalo*

$$\sigma > 1 - \frac{c}{d(r+1)\log \mathfrak{q}(f)}.$$

Demonstração. Sejam β_j os zeros reais de $L(f,s)$ tais que $\frac{1}{2} \leqslant \beta_j \leqslant 1$. Tomando a parte real de (2.21) para $s \in \mathbb{R}$ com $s = \sigma > 1$, temos

$$-\mathrm{Re}\frac{L'}{L}(f,s) - \mathrm{Re}\frac{r}{s} - \mathrm{Re}\frac{r}{s-1} + \sum_{|s+k_j|<1} \mathrm{Re}\frac{1}{s+k_j} + \sum_{|s-\rho|<1} \mathrm{Re}\frac{1}{s-\rho}$$

é igual a $O(\log \mathfrak{q}(f,s))$. Já que $s = \sigma > 1$ então

$$\mathrm{Re}\ \frac{r}{s} = \frac{r}{\sigma} < r \quad \text{e} \quad \mathrm{Re}\ \frac{r}{s-1} = \frac{r}{\sigma-1}.$$

Destas estimativas temos que

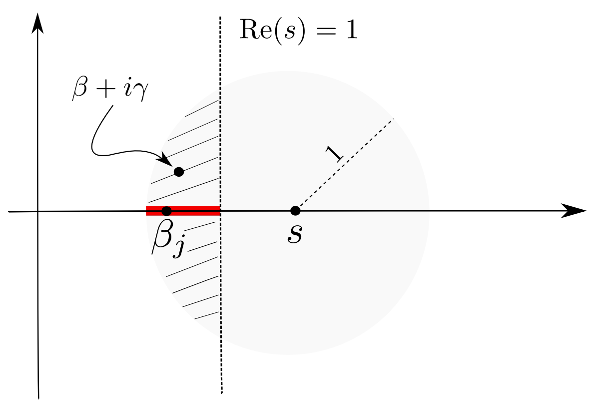
$$\sum_{|s+k_j|<1} \mathrm{Re}\frac{1}{s+k_j} + \sum_{|s-\rho|<1} \mathrm{Re}\frac{1}{s-\rho} < \mathrm{Re}\frac{L'}{L}(f,s) + \frac{r}{\sigma-1} + O\left(\log \mathfrak{q}(f,s)\right).$$

Dado que $s = \sigma > 1$ e $\mathrm{Re}(k_j) > -1$, segue que $\mathrm{Re}(1/(s+k_j)) > 0$ para todo j. Assim

$$\sum_{|s-\rho|<1} \mathrm{Re}\frac{1}{s-\rho} < \mathrm{Re}\frac{L'}{L}(f,s) + \frac{r}{\sigma-1} + O\left(\log \mathfrak{q}(f,s)\right).$$

Note que podemos ignorar no lado esquerdo da desigualdade acima os zeros cuja parte imaginária é não nula e os zeros com $\beta < 1/2$, pois para estes zeros $\rho = \beta + i\gamma$ temos

$$\mathrm{Re}\frac{1}{\sigma-\rho} = \frac{\sigma-\beta}{|\sigma-\rho|^2} \geqslant 0,$$



Além disso, como uma condição para a validade da equação (2.21) é que $\sigma \leqslant 2$, temos da desigualdade (2.5) que

$$\log \mathfrak{q}(f,s) \leqslant \log \mathfrak{q}(f) + d\log(\sigma+3) \leqslant \log \mathfrak{q}(f) + d\log(5),$$

isto é, $\log \mathfrak{q}(f,s) \ll \log \mathfrak{q}(f)$. Portanto

$$\sum_j \frac{1}{\sigma-\beta_j} < Re\frac{L'}{L}(f,\sigma) + \frac{r}{\sigma-1} + O(\log \mathfrak{q}(f)). \tag{2.29}$$

Como, por hipótese, $Re(\Lambda_f(n)) \geqslant 0$ para $(n, q(f)) = 1$ temos

$$Re\frac{L'_{nr}}{L_{nr}}(f,\sigma) \leqslant 0, \tag{2.30}$$

onde L_{nr} é o produto de Euler de $L(f,s)$ restrito a primos não ramificados (veja a identidade (2.17)).

Em seguida, vamos estimar a contribuição de primos ramificados, para a primeira parcela do direito da desigualdade (2.29), usando o fato que $|\alpha_j(p)| \leqslant p/2$. Neste caso, segue de (2.19) que (observando que agora a soma é feita apenas sobre os divisores primos de $q(f)$)

$$\left|\sum_{p|q(f)}\sum_{j=1}^{d} \frac{\alpha_j(p)p^{-\sigma}\log p}{1-\alpha_j(p)p^{-\sigma}}\right| \leqslant \sum_{p|q(f)}\sum_{j=1}^{d} \frac{p^{1-\sigma}\log p}{2-p^{1-\sigma}} \leqslant d\sum_{p|q(f)} \log p \leqslant d\log q(f).$$

Usando as estimativas obtidas acima para os termos na soma do lado direito de 2.29 segue que

$$\sum_j \frac{1}{\sigma-\beta_j} < \frac{r}{\sigma-1} + O(d\log \mathfrak{q}(f)). \tag{2.31}$$

Suponha que $\beta_j = 1$ para algum j. Então (interpretando um zero como um polo de ordem negativa) temos que $r < 0$, assim da estimativa acima temos

$$\frac{1-r}{\sigma-1} + \sum_{\substack{j \\ \beta_j \neq 1}} \frac{1}{\sigma-\beta_j} < O(d\log \mathfrak{q}(f)),$$

como os termos da soma no lado esquerdo são positivos e $1-r>0$ podemos tomar σ tão perto de 1 quanto se queira. Assim o termo do lado esquerdo da desigualdade acima alguma hora será maior que $O(d\log \mathfrak{q}(f))$ que é uma contradição. Portanto $\beta_j = 1$ não é possível, daí concluímos que $L(f,1) \neq 0$.

Suponha que $\beta_j > 1 - c\,(d(r+1)\log \mathfrak{q}(f))^{-1}$ para $1 \leqslant j \leqslant n$. Escolhemos convenientemente $\sigma = 1 + 2c\,(d\log \mathfrak{q}(f))^{-1}$. Tomando c suficientemente pequeno temos que $|\sigma-\beta_j| \leqslant 1$. Além disso, fazendo manipulações algébricas simples temos que

$$\left[\frac{d\log\mathfrak{q}(f)}{2c+c/(r+1)}\right]^{-1} = 1 + \frac{2c}{d\log\mathfrak{q}(f)} + \frac{c}{d(r+1)\log\mathfrak{q}(f)} - 1 > \sigma - \beta_j.$$

Somando agora sobre os zeros e usando a estimativa (2.31) temos

$$n \cdot \frac{d\log\mathfrak{q}(f)}{2c+c/(r+1)} < \sum_j \frac{1}{\sigma-\beta_j} < \frac{r}{\sigma-1} + O(d\log\mathfrak{q}(f)).$$

Pela escolha de σ temos $1/(\sigma-1) = d\log\mathfrak{q}(f)/2c$, logo

$$\frac{nd\log\mathfrak{q}(f)}{2c+c/(r+1)} < \left(\frac{r}{2c} + O(1)\right) d\log\mathfrak{q}(f).$$

Isto implica que

$$n < r + \frac{r}{2(r+1)} + O(c)$$

e $n \leqslant r$ se c é suficientemente pequeno. O que conclui a prova. □

TEOREMA 2.3. *Seja $L(f,s)$ uma L-função de grau d tal que as convoluções de Rankin-Selberg $L(f\otimes f,s)$ e $L(f\otimes\overline{f},s)$ existem, a última tem um polo simples em $s=1$ e a primeira é uma função inteira se $f\neq\overline{f}$. Suponha que para todo primo ramificado, $|\alpha_j(p)|^2 \leq p/2$. Então existe uma constante $c>0$ tal que $L(f,s)$ não têm zeros na seguinte região*

$$\sigma \geqslant 1 - \frac{c}{d^4\log(\mathfrak{q}(f)(|t|+3|)}, \tag{2.32}$$

exceto possivelmente um zero real simples $\beta_f < 1$. Se tal zero existir então f é necessariamente auto-dual.

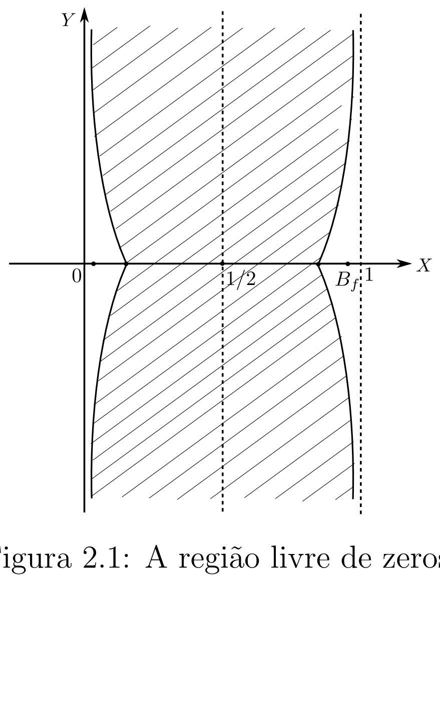


Figura 2.1: A região livre de zeros.

Demonstração. Para $t \in \mathbb{R}$, seja $L(g,s)$ uma L-função definida por

$$L(g,s) = \zeta(s)L(f,s+it)^2 L(\overline{f},s-it)^2 L(f\otimes f,s+2it)L(\overline{f}\otimes\overline{f},s-2it)L(f\otimes\overline{f},s)^2$$

de grau $(1+2d)^2$. Da desigualdade (2.6)

$$\begin{aligned}\mathfrak{q}(g) &= 3\mathfrak{q}(f,it)^4\mathfrak{q}(f\otimes f,2it)\mathfrak{q}(\overline{f}\otimes\overline{f},2it)\mathfrak{q}(f\otimes\overline{f},t)^2\\ &\leqslant \mathfrak{q}(f,it)^4\mathfrak{q}(f\otimes f,2it)^4\mathfrak{q}(f\otimes\bar{f})^2\\ &\leqslant \mathfrak{q}(f)^{4+12d}(|t|+3)^{6d^2}.\end{aligned}$$

Então o seu condutor analítico satisfaz

$$\mathfrak{q}(g) \leqslant \mathfrak{q}(f)^{4+12d}(|t|+3)^{6d^2} \tag{2.33}$$

A definição de $L(g,s)$ foi feita de modo que os coeficientes, em primos não ramificados, da serie de Dirichlet de menos a derivada logarítmica sejam reais não negativos.

De fato, para um primo p não ramificado, o fator local em p para $L(g,s)$ posto na forma 2.1 terá "raízes" 1 com multiplicidade um, $\alpha_j p^{it}$ e $\overline{\alpha}_j p^{-it}$ com multiplicidade dois, $\alpha_j\overline{\alpha}_k$ com multiplicidade dois, $\alpha_j\alpha_k p^{2it}$ e $\overline{\alpha}_j\overline{\alpha}_k p^{-2it}$ com multiplicidade um, onde α_j,α_k percorre as raízes para $L(f,s)$. Portanto para todo $k \geqslant 1$, a soma das k-ésimas potências destas "raízes" é dada por

$$\left|1+\sum_j \alpha_j^k p^{kit} + \sum_j \overline{\alpha}_j^k p^{-kit}\right|^2 \geqslant 0,$$

assim $\Lambda_g(n) \geqslant 0$ para algum n coprimo com $q(f)$.

Seja $\rho = \beta + i\gamma$ um zero de $L(f,s)$ com $\beta \geqslant \frac{1}{2}$ e $\gamma \neq 0$ (zero não real). Vamos fixar agora $t=\gamma$ na definição de $L(g,s)$. Afirmamos que $L(g,s)$ tem um polo em $s=1$ de ordem menor ou igual a 3. De fato, sabemos que a função zeta tem um polo simples em $s=1$ (veja o Apêndice) e por hipótese $L(f\otimes\bar{f})$ tem polo simples em $s=1$. Além disso, nenhuma L-função pode ter polos em $s=1+i\gamma$, $\gamma\neq 0$ pois uma L-função completa $\Lambda(f,s)$ tem polos no máximo em $s=1$ e $s=0$ então se $L(f,s)$ tivesse um polo em $s=1+it$, $\gamma(f,s)$ teria que ter um zero no mesmo ponto o que não acontece.

Dado $\rho=\beta+i\gamma$ um zero de $L(f,s)$, segue que β é um zero de $L(g,s)$ de ordem maior ou igual que 4. Portanto pelo Lema 2.2 temos

$$\beta < 1 - \frac{c}{(1+2d)^2(r+1)\log\mathfrak{q}(g)}.$$

Usando a equação (2.33) podemos ver que

$$d^2\log\mathfrak{q}(g) \ll d^4\log\mathfrak{q}(f)(|t|+3).$$

Portanto, já que $r \leqslant 3$, obtemos a seguintes desigualdades

$$\beta < 1 - \frac{\tilde{c}}{d^2 \log \mathfrak{q}(g)} < 1 - \frac{c'}{d^4 \log(\mathfrak{q}(f)(|t|+3))}, \tag{2.34}$$

para algumas constantes $\tilde{c} > 0$ e $c' > 0$. Isto prova o teorema para os zeros não reais de $L(f,s)$.

Passamos agora à análise do caso dos zeros reais de $L(f,s)$. Agora será conveniente tomar $t = 0$ na definição de $L(g,s)$. Note qualquer zero de $L(f,s)$ é zero de $L(g,s)$ com multiplicidade pelo menos 4. Por outro lado, o ponto $s = 1$ é um polo de $L(g,s)$ de ordem 3 se $f \neq \overline{f}$ e de ordem 5 se $f = \overline{f}$, pois por definição para uma L função deslocada precisamos que $L(f,s)$ seja inteira.

Por tanto se f fosse auto dual, pelo Lema 2.2 a função $L(g,s)$ teria no máximo 5 raízes reais β_j satisfazendo

$$\beta_j \geqslant 1 - \frac{c}{(1+2d)^2(r+1)\log \mathfrak{q}(g)},$$

mas como cada β_j tem multiplicidade pelo menos 4, segue que só pode acontecer um zero real nessa região. De fato, se β_j não é um zero simples de $L(f,s)$ então β_j é um zero de ordem pelo menos 8 para $L(g,s)$ o que é uma contradição. Este único zero, se existir, é conhecido como zero excepcional ou zero de Siegel. Isto conclui a prova, a menos de mostrar que o possível zero excepcional de $L(f,s)$, caso f seja auto dual satisfaz $\beta_f < 1$.

Considere

$$L(h,s) = \zeta(s)L(f,s)^2 L(f \otimes \bar{f}, s).$$

Assumindo $L(f,1) = 0$, segue-se que $L(h,s)$ é inteira, pois o zero duplo se cancela com os polos de $\zeta(s)$ e $L(f \otimes \bar{f}, s)$. Usando raciocínio análogo ao caso anterior podemos ver que os coeficientes da serie de Dirichlet de menos a derivada logarítmica de $L(h,s)$, quando suportados em potências de primos, são não negativos. Mais precisamente, para primos não ramificados p e $k \geqslant 1$ temos que os fatores locais de $L(h,s)$ é da forma (2.1) com raízes 1 com multiplicidade um, $\alpha_j(p)$ com multiplicidade dois e $\alpha_i(p)\overline{\alpha_j}(p)$ com multiplicidade um, assim

$$\Lambda_h(p^k) = \left|1 + \sum_j \alpha_j^k\right|^2 \geqslant 0. \tag{2.35}$$

Usando a equação (2.17), temos que

$$-\frac{L'}{L}(h,s) = \sum_p \sum_{k \geqslant 1} \Lambda_h(p^k) p^{-ks}.$$

Como $\mathrm{Re}(s) > 1$ é simplesmente conexo, $L(h,s)$ é analítica nesta região e $L(h,s) \neq 0$, então existe um ramo do logaritmo tal que $\log(L(h,s))$ é uma função analítica em $\mathrm{Re}(s) > 1$ e além do mais $\log(L(h,s))$ é uma primitiva de $L'/L(h,s)$.

Levando em conta que as somas do lado direito da série acima convergem em subconjuntos compactos da região $Re(s) > 1$, obtemos a seguinte igualdade,

$$\log L(h,s) = \sum_p \sum_{k>0} \frac{1}{k \log p} \Lambda_h(p^k) p^{-ks} \tag{2.36}$$

onde todos os coeficientes são não negativos.

Seja $\sigma_0 \leqslant 1$ o maior zero real de $L(h,s)$, o que corresponde a primeira singularidade de $\log L(h,s)$ quando s decresce. Note que σ_0 existe pois $\zeta(-2) = 0$.

Pelo Lema de Landau (veja Lema 5.1 do Apêndice) a serie (2.36) converge para todo real $\sigma > \sigma_0$, assim por (2.36) tem-se $\log L(h,\sigma) \geqslant 0$ e $|L(h,\sigma)| \geqslant 1$. Tomando $\sigma \to \sigma_0$, obtemos uma contradição. Desta forma $L(f,1) \neq 0$ e $\beta_f < 1$. □

2.4 O Teorema dos Números Primos Generalizado

Como vimos no capítulo anterior, o TNP possui muitas equivalências, uma delas era a seguinte:

$$\Psi(x) = \sum_{n \leqslant x} \Lambda(n) \sim x.$$

Motivados por este resultado, vamos a estudar agora o comportamento assintótico de somas do tipo

$$\Psi(f,x) = \sum_{n \leqslant x} \Lambda_f(n). \tag{2.37}$$

A expansão assintótica obtida será chamada de Teorema dos Números Primos para L-funções, ou Teorema dos Números Primos Generalizado. Por abuso de notação, usaremos simplesmente TNP para nos referir a estes resultados assintóticos.

Uma parte importante destas expansões se refere ao controle rigoroso do erro.

A estratégia adotada aqui privilegia resultados explícitos nos parâmetros em detrimento a otimalidade das estimativas.

Como poderá perceber o leitor os resultados são cada vez mais forte a medida que aumentamos nosso conhecimento sobre a região livre de zeros de $L(f,s)$. Um dos melhores resultados conhecidos atualmente é que $L(f,s)$ não tem zeros na região (2.32) exceto possivelmente o zero excepcional β_f no caso em f é auto-dual.

Nesta seção vamos considerar que o zero excepcional está no intervalo abaixo:

$$1 - \frac{c}{d^4 \log(3\mathfrak{q}(f))} \leqslant \beta_f < 1 \tag{2.38}$$

Note que esta informação sobre a localização de β_f é obtida diretamente do Teorema 2.3, tomando $t = 0$.

Em geral, ainda não temos uma cota forte para os coeficientes de $L(f,s)$. Mas como veremos cotas para $\Lambda_f(n)$ em média serão suficientes para nossos propósitos.

Antes de prosseguir destacamos que vamos provar o teorema sob a seguinte hipótese:

$$\sum_{n\leqslant x} |\Lambda_f(n)|^2 \ll xd^2 \log^2(x\mathfrak{q}(f)) \tag{2.39}$$

Em geral, não se sabe ainda qual é precisamente a classe de L-funções que satisfaz esta exigência, mas queremos observar que no entanto se trata de uma suposição razoável, já que ela é válida para distintas classes de f, por exemplo, a função Zeta de Riemann. De fato, pelo TNP $\Psi(x) \sim x$ então

$$\sum_{n\leqslant x} |\Lambda(n)|^2 \ll x\log^2(3x).$$

Assumindo a estimativa (2.39) e que $y \in [1,x]$ temos, usando a desigualdade de Cauchy, as seguintes estimativas

$$\begin{aligned}
\left(\sum_{x<n\leqslant x+y} |\Lambda_f(n)|\right)^2 &\leqslant \sum_{x<n\leqslant x+y} |\Lambda_f(n)|^2 \sum_{x<n\leqslant x+y} 1 = y \sum_{x<n\leqslant x+y} |\Lambda_f(n)|^2 \\
&= y\left(\sum_{n\leqslant x+y} |\Lambda_f(n)|^2 - \sum_{n\leqslant x} |\Lambda_f(n)|^2\right) \\
&\ll y\big((x+y)d^2\log^2((x+y)\mathfrak{q}(f)) + xd^2\log^2(x\mathfrak{q}(f))\big) \\
&\leqslant y\big(2xd^2\log^2(2x\mathfrak{q}(f)) + xd^2\log^2(x\mathfrak{q}(f))\big) \\
&\ll xyd^2\log^2(x\mathfrak{q}(f)),
\end{aligned}$$

desta forma obtemos a estimativa

$$\sum_{x<n\leqslant x+y} |\Lambda_f(n)| \ll d\sqrt{xy}\log(x\mathfrak{q}(f)). \tag{2.40}$$

Agora estamos pronto para enunciar e provar o Teorema do Número Primo Generalizado.

TEOREMA 2.4 (Teorema do Número Primo Generalizado). *Seja $L(f,s)$ uma L-função na qual* (2.32) *é a região livre de zeros com no máximo um zero excepcional real β_f no intervalo* (2.38)*, onde c é uma constante positiva. Supondo que vale a cota* (2.39) *temos*

$$\Psi(f,x) = \sum_{n\leqslant x} \Lambda_f(n) = rx - \frac{x^{\beta_f}}{\beta_f} + O\left(x\exp\left(\frac{-cd^{-4}\log x}{\sqrt{\log x} + 3\log\mathfrak{q}(f)}\right)(d\log x\mathfrak{q}(f))^4\right), \tag{2.41}$$

para $x \geqslant 1$. *Por convenção o termo* x^{β_f}/β_f *não está em* (2.41) *se o zero excepcional não existe.*

Demonstração. Seja $\phi : \mathbb{R} \to \mathbb{C}$ uma função tal que ϕ tem suporte em $[0, x+y]$, $\phi(z) = 1$ se $1 \leqslant z \leqslant x$ e $|\phi(z)| \leqslant 1$ caso contrário. Os parâmetro x, y são fixos porém arbitrários e y será escolhido no intervalo $1 \leqslant y \leqslant x$. Desta forma temos que

$$\left|\Psi(f,x) - \sum_n \Lambda_f(n)\phi(n)\right| = \left|\sum_{x<n\leqslant x+y} \Lambda_f(n)\phi(n)\right| \leqslant \sum_{x<n\leqslant x+y} |\Lambda_f(n)|$$
$$\ll d\sqrt{xy}\log(x\mathfrak{q}(f)).$$

Portanto obtemos a seguinte igualdade assintótica

$$\Psi(f,x) = \sum_n \Lambda_f(n)\phi(n) + O\left(d\sqrt{xy}\log(x\mathfrak{q}(f))\right).$$

Considere a seguinte função

$$\phi(z) := \begin{cases} \min\left(1, 1 + \dfrac{x-z}{y}\right), & 0 \leqslant z \leqslant x+y; \\ 0, & z > x+y. \end{cases}$$

A função assim definida satisfaz as condições mencionadas acima. De fato, temos que se $1 \leqslant z \leqslant x$ então $1 \leqslant 1 + \frac{x-z}{y}$. Daí $\phi(z) = 1$ se $1 \leqslant z \leqslant x$. Agora se $x < z \leqslant x+y$ temos que $1 > 1 + \frac{x-z}{y}$ e portanto $\phi(z) = 1 + \frac{x-z}{y}$. Além disso como $x < z \leqslant x+y$ segue que $-1 \leqslant \frac{x-z}{y} < 0$, desta forma $|\phi(z)| \leqslant 1$ para $x < z \leqslant x+y$.

Integrando por partes repetidamente (levando em conta que a função ϕ é contínua e $\phi(x+y) = 0$) obtêm-se

$$\int_0^{x+y} \phi(z)z^{s-1}\, dz = \frac{(-1)^\upsilon}{s(s+1)\ldots(s+\upsilon-1)} \int_0^{x+y} \phi^{(\upsilon)}(z)z^{s+\upsilon-1}\, dz.$$

Analisemos agora o comportamento assintótico, para casos particulares do lado direito na última igualdade

Para $\upsilon = 1$

$$\frac{-1}{s}\int_0^{x+y} \phi'(z)z^s dz \ll \frac{x^\sigma}{|s|}.$$

Para $\upsilon = 2$ (já que $\phi''(z) \equiv 0$)

$$\frac{1}{s(s+1)}\int_0^{x+y} \phi''(z)z^{s+1} dz \ll \frac{x^{\sigma+1}}{|s|^2 y}.$$

Portanto a Transformada de Mellin de $\phi(z)$ (veja o Apêndice, definição 5.3) satisfaz

$$\widehat{\phi}(s) = \int_0^{x+y} \phi(z) z^{s-1} dz \ll \frac{x^\sigma}{|s|} \min\left(1, \frac{x}{|s|y}\right), \tag{2.42}$$

para $s = \sigma + it$, $\frac{1}{2} \leqslant \sigma \leqslant 2$.

Pela Fórmula da Transformada Inversa de Mellin

$$\phi(z) = \frac{1}{2\pi i} \int_{(2)} \widehat{\phi}(s) z^{-s} ds. \tag{2.43}$$

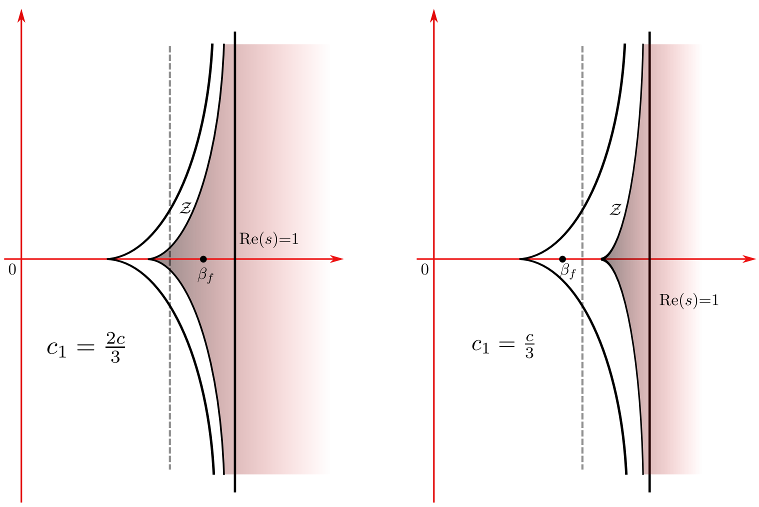
Dado que esta última integral converge absolutamente, temos que

$$\sum_n \Lambda_f(n)\phi(n) = \sum_{n \geqslant 1} \Lambda_f(n) \frac{1}{2\pi i} \int_{(2)} \widehat{\phi}(s) n^{-s} ds = \frac{1}{2\pi i} \int_{(2)} -\frac{L'}{L}(f,s)\widehat{\phi}(s) ds. \tag{2.44}$$

Definimos a seguinte região, que será necessário para nosso propósito

$$A = \left\{ s = \sigma + it \mid \sigma \geqslant 1 - \frac{c_1}{d^4 \log(\mathfrak{q}(f)(|t|+3))} \right\},$$

onde $c_1 = 2c/3$, ou $c_1 = c/3$, dependendo se o zero excepcional está na metade direita do segmento (2.38) ou não (com a possibilidade da não existência). Seja $\mathcal{Z}$ a fronteira desta região. Observe que todos os zeros de $L(f,s)$ estão a uma distancia de $\mathcal{Z}$ maior ou igual que $c/6d^4 \log(\mathfrak{q}(f)(|t|+3))$.



Nosso próximo passo é estimar para $s \in \mathcal{Z}$ cada parcela do lado esquerdo de (2.21), isto é,

$$-\frac{L'}{L}(f,s) - \frac{r}{s} - \frac{r}{s-1} + \sum_{|s+k_j|<1} \frac{1}{s+k_j} + \sum_{|s-\rho|<1} \frac{1}{s-\rho} \ll \log \mathfrak{q}(f,s). \tag{2.45}$$

Como os zeros estão a uma distância maior ou igual que $c/6d^4\log(\mathfrak{q}(f)(|t|+3))$ de $\mathcal{Z}$, podemos usar a estimativa dada em (2.20) com $\rho=\sigma+i\gamma$ para verificar que

$$\begin{aligned}\sum_{|s-\rho|<1}\frac{1}{s-\rho} &\ll \sum_{|t-\gamma|<1}\frac{6d^4\log(\mathfrak{q}(f)(|t|+3))}{c}=\frac{6d^4\log(\mathfrak{q}(f)(|t|+3))}{c}m(t,f)\\ &\ll \frac{6d^4\log(\mathfrak{q}(f)(|t|+3))}{c}\log(\mathfrak{q}(f,it)).\end{aligned}$$

Usando a definição de condutor analítico obtemos

$$\begin{aligned}\log\mathfrak{q}(f,it) &= \log\left(q(f)\prod_{j=1}^{d}(|it+k_j|+3)\right)\leqslant \log\left(q(f)(|t|+3)^d\prod_{j=1}^{d}(|k_j|+3)\right)\\ &= \log(\mathfrak{q}(f))+d\log(|t|+3)\ll \log\left(\mathfrak{q}(f)(|t|+3)\right).\end{aligned}$$

Usando as duas estimativas acima temos

$$\sum_{|s-\rho|<1}\frac{1}{s-\rho}\ll d^4\log^2\left(\mathfrak{q}(f)(|t|+3)\right).$$

Agora vamos estimar a primeira soma do lado esquerdo de (2.45). Pela desigualdade triangular e o fato que $|\mathrm{Re}(s)|\leqslant |s|$ segue que

$$\left|\sum_{|s+k_j|<1}\frac{1}{s+k_j}\right|\leqslant \sum_{|\sigma+\mathrm{Re}(k_j)|<1}\frac{1}{|\sigma+\mathrm{Re}(k_j)|}.$$

Como $s\in\mathcal{Z}$ temos que

$$\sigma+\mathrm{Re}(k_j)=\frac{[1+\mathrm{Re}(k_j)]d^4\log\left(\mathfrak{q}(f)(|t|+3)\right)-c_1}{d^4\log\left(\mathfrak{q}(f)(|t|+3)\right)}.$$

Então usando a regra de L'Hospital e o fato que $\mathrm{Re}(k_j)>-1$

$$\lim_{t\to^+\infty}\frac{1}{\sigma+\mathrm{Re}(k_j)}=\frac{1}{1+\mathrm{Re}(k_j)}$$

Se denotamos por $K=\min_j\{\mathrm{Re}(k_j)\}$, obtemos a seguinte estimativa

$$\begin{aligned}\sum_{|\sigma+Re(k_j)|<1}\frac{1}{\sigma+Re(k_j)} &\leqslant \sum_{|\sigma+\mathrm{Re}(k_j)|<1}\frac{1}{\sigma+K}\\ &\ll \frac{d}{1+K}\ll d.\end{aligned}$$

Além disso, para $s\in\mathcal{Z}$

$$\left|\frac{r}{s}\right| \leqslant \frac{r}{|\sigma|} = \left|\frac{rd^4 \log\left(\mathfrak{q}(f)(|t|+3)\right)}{d^4 \log\left(\mathfrak{q}(f)(|t|+3)\right) - c_1}\right| = O(1),$$

e

$$\left|\frac{r}{s-1}\right| \leqslant \frac{r}{|\sigma - 1|} = \frac{rd^4 \log\left(\mathfrak{q}(f)(|t|+3)\right)}{c_1}.$$

Juntando as estimativas acima na equação (2.45) temos que para $s \in \mathcal{Z}$

$$-\frac{L'}{L}(f,s) \ll d^4 \log^2\left(\mathfrak{q}(f)(|t|+3)\right). \tag{2.46}$$

Na sequência vamos mostrar, usando o Teorema dos Resíduos, que podemos mover a integração do contorno $Re(s) = 2$ para $\mathcal{Z}$ na equação (2.43). Antes de aplicar o Teorema dos Resíduos, vamos analisar o que acontece com a integral nas linhas horizontais. Para isso tome $t \in \mathbb{R}$ e mostremos que

$$\lim_{t\to\infty} \int\limits_{<t>} \frac{L'}{L}(f,s)\widehat{\phi}(s)ds = 0.$$

De fato, para $\sigma_0 \leqslant Re(s) \leqslant 2$, onde σ_0 é a parte real (associado a t) de um ponto que pertence à fronteira da região livre de zeros, temos

$$\int\limits_{<t>} \frac{L'}{L}(f,s)\widehat{\phi}(s)ds = \int_{\sigma_0}^{2} \frac{L'}{L}(f,s)\widehat{\phi}(s)d\sigma \ll \int_{\sigma_0}^{2} \frac{L'}{L}(f,s)\frac{x^2}{|t|}d\sigma.$$

Aplicando o Teorema da Convergência Dominada de Lebesgue e observando que, se $t \to \infty$ então $\sigma \to 1$, ficamos com

$$\lim_{t\to\infty} \int_{-1}^{2} 1_{[\sigma_0,2]}(\sigma)\frac{L'}{L}(f,s)\frac{x^2}{|t|}d\sigma = \int_{1}^{2} \lim_{t\to\infty} \frac{L'}{L}(f,s)\frac{x^2}{|t|}d\sigma = 0.$$

Agora, aplicando o Teorema dos Resíduos na equação (2.44) obtemos

$$\sum_n \Lambda_f(n)\phi(n) = r\widehat{\phi}(1) - \widehat{\phi}(\beta_f) + \frac{1}{2\pi i}\int\limits_{\mathcal{Z}} -\frac{L'}{L}(f,s)\widehat{\phi}(s)ds.$$

Observamos que a presença do resíduo envolvendo o zero excepcional depende apenas de sua existência e localização com respeito ao contorno $\mathcal{Z}$.

Para $s = 1$ ou $s = \beta_f$ temos

$$\widehat{\phi}(s) = \int_0^x z^{s-1}dz + O(y) = \frac{x^s}{s} + O(y),$$

enquanto que a integral de linha sobre $\mathcal{Z}$ pode ser estimada, usando (2.46) e (2.42), pela seguinte expressão

$$d^4 \int\limits_{\mathcal{Z}} \frac{x^\sigma}{|s|}\min\left(1, \frac{x}{|s|y}\right) \log^2\left(\mathfrak{q}(f)(|t|+3)\right)|ds| \ll d^4 x^{\sigma(T)} \log^3\left(\mathfrak{q}(f)T\right),$$

onde $T = x/y$ e $\sigma(T) = 1 - c_1/d^4 \log(\mathfrak{q}(f)T)$. De fato, suponha primeiro que $T \geqslant |t| + 3$, então $|s| \ll T\mathfrak{q}(f)$ e

$$d^4 \int\limits_{\mathcal{Z}\cap\{|s|\ll T\mathfrak{q}(f)\}} \frac{x^\sigma}{|s|} \min\left(1, \frac{x}{|s|y}\right) \log^2\left(\mathfrak{q}(f)(|t|+3)\right)|ds| \leqslant d^4 x^{\sigma(T)} \log^2\left(\mathfrak{q}(f)T\right) \int\limits_{\mathcal{Z}\cap\{|s|\ll T\mathfrak{q}(f)\}} \frac{|ds|}{|s|}$$

$$\ll d^4 x^{\sigma(T)} \log^3\left(\mathfrak{q}(f)T\right).$$

Por outro lado, se $T < |t| + 3$, o integrando vai para zero quando $|s| \to \infty$ e assim a integral acima é finita.

Agrupando os resultados acima obtemos

$$\Psi(f,x) = rx - \frac{x^{\beta_f}}{\beta_f} + O\left(d^4(xT^{-\frac{1}{2}} + x^{\sigma(T)}) \log^3(x\mathfrak{q}(f))\right), \tag{2.47}$$

onde T esta sujeito apenas a restrição $1 \leqslant T \leqslant x$. Escolhendo $T = \exp(\frac{1}{3}\sqrt{\log x})$ temos que

$$\begin{aligned} x^{\sigma(T)} &= \exp(\sigma(T)\log x) = \exp\left(\left(1 - \frac{c_1}{d^4 \log\left(\mathfrak{q}(f)\exp\left(\frac{1}{3}\sqrt{\log x}\right)\right)}\right)\log x\right) \\ &= x\exp\left(\frac{-3c_1 d^{-4}\log x}{3\log(\mathfrak{q}(f)) + \sqrt{\log x}}\right). \end{aligned}$$

Além do mais, tem-se que para dito valor de T

$$T^{-\frac{1}{2}} = O\left(\exp\left(\frac{-3c_1 d^{-4}\log x}{3\log(\mathfrak{q}(f)) + \sqrt{\log x}}\right)\right).$$

Portanto, substituindo as duas últimas igualdades em (2.47) obtemos o resultado desejado. □

Observação 2.1. *É possível simplificar o erro comprometendo levemente a "força" do resultado. Por exemplo, temos*

$$\psi(f,x) = rx - \frac{x^{\beta_f}}{\beta_f} + O\left(\sqrt{\mathfrak{q}(f)}x\exp\left(-\frac{c}{2d^4}\sqrt{\log x}\right)\right)$$

De fato isto é verdade, porque este erro é menor do que 2.41 *e portanto esta afirmação é trivial dado o resultado obtido acima.*

Capítulo 3

Somas Exponenciais - Teorema de H. Davenport

Os métodos de crivos (veja o Apêndice) é uma ferramenta de grande utilidade em numerosos problemas da Teoria Analítica dos Números. Eles apareceram ao estudar problemas aditivos como a Conjectura dos Primos Gêmeos e a Conjectura de Goldbach. No estudo destes métodos, o matemático norueguês Viggo Brun foi um dos primeiros a introduzir o procedimento de exclusão-inclusão para atacar os problemas que neles apareciam. Por exemplo, o crivo de Eratóstenes segue essa linha.

Nos 50 anos que sucederam os trabalhos de Viggo Brun muitas ramificações do procedimento de exclusão-inclusão foram desenvolvidas, e este procedimento ainda se constitui uma técnica fundamental em diversos tipos de crivos. Embora as ideias principais sejam elementares, é necessário incorporar argumentos analíticos para obter estimativas mais finas. Desafortunadamente os métodos de crivos não são poderosos o bastante para fornecer, por exemplo, números primos em sequências gerais.

Nesse contexto, Vinogradov (1891-1983), considerado um dos criadores da Teoria Analítica dos Números Moderna, inventou um novo método para aproximar somas exponenciais

$$S_f(N) = \sum_{1 \leqslant n \leqslant N} e(f(n)),$$

onde, como em todo o restante deste capítulo, $e(\alpha) := \exp(2\pi i \alpha)$, f uma função suave tomando valores complexos que é chamada de função de amplitude e N é um inteiro positivo chamado de comprimento da soma.

Neste capítulo estudaremos uma variante do Método de Vinogradov usando as Identidades de Vaughan e Formas Bilineares para estudar o comportamento assintótico da soma

$$\sum_{n \leqslant x} \mu(n) e(\alpha n) \tag{3.1}$$

Vamos começar o capítulo, lembrando uma estimativa para Formas Bilineares sobre segmentos diádicos usando a Desigualdade de Cauchy.

3.1 Formas Bilineares

O núcleo da Teoria dos Números Primos consiste na distribuição dos números primos em segmentos curtos, em progressões aritméticas, em polinômios homogêneos e em outras sequências esparsas semelhantes. Este território se expande para os corpos numéricos, onde os ideais primos desempenham o papel de primos.

Agora gostaria de passar por um território diferente, aquele de problemas envolvendo números primos que dificilmente podem ser tratados pelos métodos das L-funções (a teoria de L-funções normalmente é adequada para estudos envolvendo funções aritméticas multiplicativas). Estes métodos que vamos apresentar em seguida são baseados no conceito de Formas Bilineares. Considere a soma

$$S_f(x) = \sum_{p \leqslant x} f(p),$$

onde f é uma função aritmética definida em primos, não necessariamente em todos os números inteiros (certamente pode-se estender f a todos os inteiros arbitrariamente, porém uma extensão natural pode não ser fácil de ser feita). Ao estudar a soma acima, para uma função f particular, é comum fazermos uso de outras somas relacionadas a ela. Mas muitas vezes, nem a série de Dirichlet associada

$$\sum_n f(n) n^{-s},$$

nem a serie de Fourier

$$\sum_n f(n) e(nz),$$

têm propriedades boas o suficiente para tirarmos as conclusões que desejamos. Então procuramos relacioná-las com outras somas ou função dando lugar, de alguma maneira, às Formas Bilineares.

Para estimar Formas Bilineares começamos olhando para somas duplas do tipo

$$\Psi(\alpha, \beta) = \sum_m \sum_n \alpha_m \beta_n \phi(m, n), \tag{3.2}$$

onde $\alpha = (\alpha_m)$, $\beta = (\beta_n)$ são vetores em $\mathbb{C}^m$ e $\mathbb{C}^n$, respectivamente e

$$\Phi = (\phi(m, n)) \tag{3.3}$$

é uma matriz de tamanho $m \times n$ com entradas complexas.

Dada Φ o objetivo é dar estimativas não triviais para $\Psi(\alpha, \beta)$ válidas para vetores α e β arbitrários, isto é, estimativas para a norma do operador representado pela matriz Φ. Muitas vezes, uma tentativa de usar a estrutura específica desses vetores traz de volta uma imagem espelhada da situação original. Assim é natural a pergunta, qual é a vantagem de organizar as quantidades originais em Formas Bilineares? Uma das chaves para esta resposta é que as sequências α e β são essencialmente independentes uma da outra e portanto o cancelamentos em $\Psi(\alpha,\beta)$ é possível, via às propriedades da matriz de coeficientes $\phi(m,n)$.

Uma das estimativas, nem sempre a melhor possível, é dada por

$$|\Psi(\alpha,\beta)|^2 \leqslant \Delta^2||\alpha||^2||\beta||^2, \tag{3.4}$$

onde $\Delta = \Delta(\Phi)$ é a norma do correspondente operador linear e

$$||\alpha||^2 = \sum |\alpha_m|^2, \quad ||\beta||^2 = \sum |\beta_n|^2.$$

A forma mais transparente de estimar $\Psi(\alpha,\beta)$ é usando a desigualdade de Cauchy, para isto escolha primeiro m estar na soma externa e n na soma interna obtendo

$$|\Psi(\alpha,\beta)|^2 \leqslant ||\alpha||^2 \sum_m \left| \sum_n \beta_n \phi(m,n) \right|^2. \tag{3.5}$$

Trocando a ordem do somatório obtemos

$$\sum_m \left| \sum_n \beta_n \phi(m,n) \right|^2 = \sum_{n_1}\sum_{n_2} \beta_{n_1}\overline{\beta}_{n_2} \sum_m \phi(m,n_1)\overline{\phi}(m,n_2). \tag{3.6}$$

Os termos onde $n_1 = n_2$, são chamados de termos da diagonal, e note que eles não colaboram de forma determinante nos cancelamentos na soma interna pois além de serem não negativos

$$\sum_n |\phi(m,n)|^2,$$

sua a quantidade de parcelas é da ordem de da raíz do número de entradas da matriz. Por outro lado, para a maioria de n_1, n_2 fora da diagonal poderíamos encontrar consideráveis cancelamentos na soma

$$\sum_m \phi(m,n_1)\overline{\phi}(m,n_2),$$

por causa das variações independentes de sinal de $\phi(m,n_1), \phi(m,n_2)$. Como se organizar esta última, para explorar estes cancelamentos depende muito das características particulares de Φ. A continuação vamos explorar algumas destas técnicas.

3.2 Somas Tipo I e Tipo II

Um dos grandes problemas em Teoria Analítica dos Números é, dada uma função $f : \mathbb{N} \to \mathbb{C}$, encontrar estimativas da soma

$$V = \sum_{p} f(p),$$

onde p percorre sobre os números primos. Esta soma está também intimamente relacionada a

$$S = \sum_{n} \Lambda(n) f(n),$$

onde Λ é a função de Von Mangoldt. Note que S pode ser obtido a partir V tomando $f(n)$ igual a $f(n)\log n$, quando as contribuições dos termos $n = p^{\nu}$ são pequenas. Mas, às vezes, na prática é mais fácil trabalhar com a soma S.

Se f é uma função suave com suporte compacto em $\mathbb{R}_+$ podemos expressar S em termos dos zeros de $\zeta(s)$, integrando $-\widehat{f}\zeta'(s)/\zeta(s)$, onde $\widehat{f}$ é a Transformada de Mellin de f. Assim obtemos

$$S = \widehat{f}(1) - \sum_{n>0} \widehat{f}(-2n) - \sum_{\rho} \widehat{f}(\rho),$$

para ver isto basta usar (2.44) e uma generalização do Princípio do Argumento que é dada pelo Teorema 3.6 da referência [5] .

Muitas funções aritméticas interessantes $f : \mathbb{N} \to \mathbb{Z}$ não são definidas em todos os números reais, assim a fórmula anterior não pode ser aplicada, em tais casos, preferimos de usar a notação sequencial, $\mathcal{A} = (a_n)$ em lugar da função $f(n) = a_n$. Neste contexto, podemos apelar ao seguinte principio.

Princípio de Aleatoriedade de Möbius

A função de Möbius $\mu(n)$ muda de sinal "aleatoriamente", desta forma para uma sequência "razoável"(isto será discutido no próximo capítulo) de números complexos $\mathcal{A} = (a_n)$, a soma

$$M(\mathcal{A}, x) = \sum_{n \leqslant x} \mu(n) a_n$$

é relativamente pequena devido ao cancelamento de seus termos.

Como mencionamos acima não vamos dar ainda uma definição precisa de sequências razoáveis, mas para tentar tornar intuitiva a discussão que segue sugerimos que o leitor pense neste momento sobre sequências "razoáveis"como sendo aquelas que não são obtidas por construções muito rebuscadas. Na prática, grande parte das sequências que trabalhamos são razoáveis, mas uma exceção importante acontece quando se joga com pesos de crivos que conspiram com a função de Möbius.

Veremos a seguir como o problema de estimar as somas citada acima pode ser transformado no estudo de dois tipos de somas, e como estas podem ser analisadas com argumentos conhecidos em Teoria Analítica dos Números. Esta decomposição em duas somas é parte essencial do Método de Vinogradov, que veremos na próxima secção.

Lembramos que alguns dos argumentos ao longo desta seção não serão dados de maneira formal, já que nosso objetivo aqui é expor as ideias essenciais usadas no tratamento das somas do tipo S e dar uma noção do Método de Vinogradov.

Seja $\mathcal{A} = (a_n)$ uma sequência de números reais, não negativos. Defina

$$S(\mathcal{A}, x) = \sum_{n \leqslant x} \Lambda(n) a_n \tag{3.7}$$

Sabemos que, da Fórmula de Inversão de Möbius

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

Substituindo esta identidade em 3.7 temos

$$S(\mathcal{A}, x) = -\sum_{d} \mu(d)(\log d) A_d(x), \tag{3.8}$$

onde

$$A_d(x) = \sum_{\substack{n \leqslant x \\ n \equiv 0 (mod\ d)}} a_n. \tag{3.9}$$

As quantidades $A_d(x)$, bem como $A_d(x) \log d$ são "razoáveis" se a sequência a_n é também "razoável". Portanto, de acordo com o Princípio de Aleatoriedade de Möbius, esperamos que a contribuição das parcelas com d's muito grande em (3.8) seja desprezível, e assim a soma $S(\mathcal{A}, x)$ deva ser bem aproximada por

$$S(\mathcal{A}, D, x) = -\sum_{d \leqslant D} \mu(d)(\log d) A_d(x) \tag{3.10}$$

para D relativamente pequeno. Agora para $d \leqslant D$ as quantidades $A_d(x)$ pode satisfazer (em média)

$$A_d(x) = g(d)A(x) + r_d(x),$$

onde $g(d)$ é uma função multiplicativa com $0 \leqslant g(p) < 1$ para todo p, $A(x) \equiv A_1(x)$, e $r_d(x)$ o termo de erro, que é relativamente pequeno. Aqui $g(d)$ representa a densidade da subsequência de elementos a_n com n sendo divisível por d. Naturalmente assumimos que $g(d)$ é multiplicativa, para usar a ideia intuitiva de que a divisibilidade por primos distintos são "eventos independentes" (isto não é exatamente certo em modelos mais fortes de crivos).

Escrevemos

$$H(D) = -\sum_{d \leqslant D} \mu(d)(\log d)g(d), \quad R(D) = -\sum_{d \leqslant D} \mu(d)(\log d)r_d(D)$$

Substituindo estas definições em (3.10) obtemos

$$S(\mathcal{A}, D, x) = H(D)A(x) + R(D)$$

A função de densidade g usualmente satisfaz condições de regularidade (distribuição sobre primos), assim existe o limite de $H(D)$ quando D tende a ∞, i.e. $H(D) \to H(\infty) =: H$ com

$$H = -\sum_{d} \mu(d)(\log d)g(d) \tag{3.11}$$

Portanto, a discussão acima nos conduz a seguinte fórmula assintótica

$$S(\mathcal{A}, x) \sim HA(x), \quad \text{quando } x \to \infty. \tag{3.12}$$

Embora não se possa ignorar o termo de erro arbitrariamente, conjectura-se que a fórmula assintótica acima seja verdadeira para sequências $A = (a_n)$ que aparecem naturalmente. Esta fórmula já havia sido obtida por vários argumentos heurísticos, por exemplo, pelo Método do Círculo (Veja [9], cap 20).

Claro, que a prova rigorosa desta expansão assintótica é feita de maneira um pouco diferente da apresentada aqui. O que deve ser resaltado é que a fórmula heurística que aparece em (3.12), em vários casos naturais em que S pode ser avaliada rigorosamente, nunca deixou de ser válida.

Vamos ver, usando as notações acima, o Método de Vinogradov de maneira abstrata. Começamos aplicando repetidamente o processo de inclusão-exclusão para reordenar $S(\mathcal{A}, x)$ em somas de dois tipos

$$S_1 = \sum_{d \leqslant D} \sum_{dn \leqslant x} \alpha_d a_{dn} \tag{3.13}$$

e

$$S_2 = \sum_{\substack{mn \leqslant x \\ M < m \leqslant N}} \sum \beta_m \gamma_n a_{mn}, \tag{3.14}$$

que são as chamadas somas do Tipo I e somas do Tipo II respectivamente, onde α_d, β_m, γ_n são números reais independentes (essencialmente limitados) e D, M, N são parâmetros adequados (nem pequenos nem grandes em relação a x).

Observe que a soma do Tipo S_1 pode ser escrita como

$$S_1 = \sum_{d \leqslant D} \alpha_d A_d(x), \tag{3.15}$$

onde $\mathcal{A}_d$ é definido em 3.9.

Assumindo alguma regularidade na variação da sequência a_n pode-se mostrar que $A_d(x)$ é pequeno para qualquer $d \leqslant D$ devido à cancelamentos de seus termos.

A soma do tipo S_2 pode ser vista como uma Forma Bilinear com coeficientes β_m, γ_n. Para obter estimativas é conveniente separar S_2 em Formas Bilineares sobre segmentos diádicos

$$\mathcal{B}(M,N) = \sum_{M<m\leqslant 2M} \beta_m \sum_{N<n\leqslant 2N} \gamma_n a_{mn}.$$

Como na seção anterior, podemos usar a Desigualdade de Cauchy para encontrar limitantes como segue

$$|\mathcal{B}(M,N)|^2 \leqslant \left(\sum_m |\beta_m|^2\right) \sum_{n_1} \sum_{n_2} |\gamma_{n_1}\gamma_{n_2}| \sum_m a_{mn_1}\overline{a}_{mn_2}.$$

Combinando os resultados para somas tipo S_1 e S_2 podemos deduzir um limitante para $S(\mathcal{A},x)$ que é frequentemente forte. Na próxima seção fornecemos detalhes do Método de Vinogradov em contextos particulares.

3.3 Método de Vinogradov

Nesta seção, desenvolvemos uma fórmula para somas sobre primos usando modificações das ideias originais de Vinogradov. Os argumentos são de natureza combinatória e têm muito em comum com métodos de crivos. Defina $\beta_n(x) = (1+\omega(n,\sqrt{x}))^{-1}$ para $1 < n \leqslant x$, onde $\omega(n,\sqrt{x})$ é o número de primos $p|n$ com $p \leqslant \sqrt{x}$. Se n é livre de quadrados, temos

$$\sum_{p|n, p\leqslant\sqrt{x}} \beta_{n/p}(x) = \begin{cases} 1 & \text{se } n \text{ tem um fator primo} \leqslant \sqrt{x}; \\ 0 & \text{se } n \text{ é primo com } \sqrt{x} < n \leqslant x. \end{cases}$$

De fato, para o segundo caso não existe nenhum termo na somatória, dai o resultado é trivial. Para o primeiro caso, observamos que um inteiro positivo geral, livre de quadrados, tem sua decomposição em fatores primos dada da seguinte forma $n = p_1p_2\dots p_k$ com $p_i \neq p_j$ para $i \neq j$. Suponha que $p_1, p_2, \dots, p_r \leqslant \sqrt{x}$ com $r \leqslant k$, dai

$$\sum_{p|n, p\leqslant\sqrt{x}} \beta_{n/p}(x) = \sum_{i=1}^{r} \beta_{p_1\dots p_{i-1}p_{i+1}\dots p_r} = \sum_{i=1}^{r} (1+(r-1))^{-1} = \sum_{i=1}^{r} \frac{1}{r} = 1.$$

Portanto para qualquer sequência de números complexos $\mathcal{A} = (a_n)$, denotando por $\sum^{\flat}$ a soma restrita a números livres de quadrados, temos

$$\begin{aligned}
\sideset{}{^\flat}\sum_{1<n\leqslant x} a_n &= \sideset{}{^\flat}\sum_{1<n\leqslant \sqrt{x}} a_n + \sideset{}{^\flat}\sum_{\sqrt{x}<n\leqslant x} a_n = \sideset{}{^\flat}\sum_{1<n\leqslant \sqrt{x}} a_n + \sum_{\sqrt{x}<p\leqslant x} a_p + \sideset{}{^\flat}\sum_{\substack{\sqrt{x}<n\leqslant x\\ n\neq p}} a_n\\
&= \sum_{\sqrt{x}<p\leqslant x} a_p + \sum_{\substack{1<n\leqslant \sqrt{x}\\ (n,p^2)=1}} a_n \sum_{\substack{p\leqslant\sqrt{x}\\ p|n}} \beta_{n/p}(x) + \sum_{\substack{\sqrt{x}<n\leqslant x\\ (n,p^2)=1}} a_n \sum_{\substack{p\leqslant\sqrt{x}\\ p|n}} \beta_{n/p}(x)\\
&= \sum_{\sqrt{x}<p\leqslant x} a_p + \sum_{\substack{p\leqslant\sqrt{x}, mp\leqslant x\\ (m,p)=1}} \sideset{}{^\flat}\sum \beta_m(x).a_{mp}
\end{aligned}$$

Desta forma obtemos a seguinte igualdade

$$\sum_{\sqrt{x}<p\leqslant x} a_p = \sideset{}{^\flat}\sum_{1<n\leqslant x} a_n - \sum_{\substack{p\leqslant\sqrt{x}, mp\leqslant x\\ (m,p)=1}} \sideset{}{^\flat}\sum \beta_m(x) a_{mp}.$$

Seja $P(z)$ o produto de todos os primos $p \leqslant z$ com $2 \leqslant z \leqslant \sqrt{x}$. Aplicando a identidade acima a subsequência de números a_n com $(n, P(z)) = 1$, obtemos

$$\sum_{\sqrt{x}<p\leqslant x} a_p = \sideset{}{^\flat}\sum_{\substack{1<n\leqslant x\\ (n,P(z))=1}} a_n - \sum_{\substack{z<p\leqslant\sqrt{x},\ mp\leqslant x\\ p\nmid m,(m,P(z))=1}} \sideset{}{^\flat}\sum \beta_m(x) a_{mp} \tag{3.16}$$

A fim de chegar ao resultado que dá o conceito do Método de Vinogradov, vamos estimar o erro que obtém-se ao retirar certas restrições nas somas sobre os elementos da sequência.

Nosso intuito é retirar a restrição a números livres de quadrados e a condição $p \nmid m$. O erro cometido ao negligenciar estes termos pode ser estimado usando a Desigualdade de Cauchy por

$$\begin{aligned}
\sum_{\substack{n\leqslant x\\ \forall p>z, p^2|n}} |a_n| &\leqslant \left(\sum_{n\leqslant x} |a_n|^2\right)^{\frac{1}{2}} \left(\sum_{z<p} xp^{-2}\right)^{\frac{1}{2}}\\
&\leqslant ||\mathcal{A}(x)|| x^{\frac{1}{2}} \left(\sum_{z<p} p^{-2}\right)^{\frac{1}{2}}\\
&\leqslant ||\mathcal{A}(x)|| x^{\frac{1}{2}} z^{-\frac{1}{2}},
\end{aligned} \tag{3.17}$$

onde $||\mathcal{A}(x)||$ é definido por

$$||\mathcal{A}(x)|| = \left(\sum_{n\leqslant x} |a_n|^2\right)^{\frac{1}{2}}.$$

A isto acrescentarmos os termos a_1 e a_p com $p \leqslant \sqrt{x}$ e estimamos sua contribuição, usando novamente a Desigualdade de Cauchy, por

$$\sum_{n\leqslant\sqrt{x}} |a_n| \leqslant ||\mathcal{A}(x)||x^{\frac{1}{4}}. \tag{3.18}$$

Em seguida, tiramos a condição $(n, P(z)) = 1$. Para isso primeiro fazemos uso da Fórmula de Legendre

$$\sum_{\substack{n\leqslant x\\(n,P(z))=1}} a_n = \sum_{d|P(z)} \mu(d)A_d(x). \tag{3.19}$$

Observe que este fato pode ser provado, lembrando a propriedade fundamental da função de Möbius

$$\sum_{\substack{n\leqslant x\\(n,P(z))=1}} a_n = \sum_{n\leqslant x} a_n \left(\sum_{d|(n,P(z))} \mu(d)\right) = \sum_{d|P(z)} \mu(d) \sum_{\substack{n\leqslant x\\d|n}} a_n = \sum_{d|P(z)} \mu(d)A_d(x).$$

Separamos os termos $A_d(x)$ com $d \leqslant \sqrt{x}$ e estimamos os demais. Primeiro pela Desigualdade de Cauchy, denotado por $\tau(n)$ a função divisor que é definido como o número de divisores positivos de n, segue que

$$\begin{aligned}
\sum_{\substack{d|P(z)\\d>\sqrt{x}}} |A_d(x)| &\leqslant \sum_{\substack{dn\leqslant x\\d|P(z),d>\sqrt{x}}}\sum |a_{dn}| \\
&\leqslant \Big(\sum_{\substack{dn\leqslant x\\d|P(z),d>\sqrt{x}}}\sum |a_{dn}|^2\Big)^{\frac{1}{2}} \Big(\sum_{\substack{dn\leqslant x\\d|P(z),d>\sqrt{x}}}\sum 1\Big)^{\frac{1}{2}} \\
&\leqslant ||\mathcal{A}(x)|| \Big(\sum_{\substack{dn\leqslant x\\d|P(z),d>\sqrt{x}}}\sum \tau(dn)\Big)^{\frac{1}{2}} \\
&\leqslant ||\mathcal{A}(x)||(x\log 3x)^{\frac{1}{2}} \left(\sum_{d|P(z),d>\sqrt{x}} \tau(d)d^{-1}\right)^{\frac{1}{2}}.
\end{aligned}$$

Vamos estimar agora a última soma que aparece no lado esquerdo da desigualdade acima. Aplicamos o Truque de Rankin [1] para $\alpha = 2\varepsilon > 0$ como segue

$$\begin{aligned}\sum_{d|P(z),d>\sqrt{x}} \tau(d)d^{-1} &\leqslant x^{-\varepsilon} \sum_{d|P(z)} \tau(d)d^{2\varepsilon-1} \leqslant x^{-\varepsilon} \prod_{p\leqslant z}(1+2p^{2\varepsilon-1}) \\ &\leqslant x^{-\varepsilon} \prod_{p}(1+p^{-1-\varepsilon})^{2z^{3\varepsilon}} \leqslant x^{-\varepsilon}\zeta(1+\varepsilon)^{2z^{3\varepsilon}} \\ &\leqslant x^{-\varepsilon}\left(1+\frac{1}{\varepsilon}\right)^{2z^{3\varepsilon}} = \exp\left(-\frac{\log x}{\log z}\right)(1+\log z)^{2e^3},\end{aligned}$$

onde a última desigualdade vem do fato que $\zeta(s) \leqslant \frac{1}{s-1}$ para valores de s perto de 1 (veja Lema 5.4 no Apêndice), e a última igualdade vem da escolha $\varepsilon = 1/\log z$.

Usaremos estas estimativas poder demonstrar o próximo teorema, que nos dará uma ideia clara do Método de Vinogradov. Este método normalmente é usado para estudar as somas exponenciais tipo Weyl, para maiores os detalhes veja [9], Seção 8.5, mas a variante que estamos interessados aqui, como veremos, conserva as ideias originais deste método.

TEOREMA 3.1. *Seja $2 \leqslant z \leqslant \sqrt{x}$. Para qualquer sequência de números complexos $\mathcal{A} = (a_n)$ temos*

$$\sum_{p\leqslant x} a_p = \sum_{\substack{d|P(z)\\ d\leqslant\sqrt{x}}} \mu(d)A_d(x) - \sideset{}{^\flat}\sum_{(m,P(z))=1} \beta_m(x) \sum_{\substack{mp\leqslant x\\ z<p\leqslant\sqrt{x}}} a_{mp} + \varepsilon(x,z)\sqrt{x}||\mathcal{A}(x)||, \tag{3.20}$$

onde $P(z)$ é o produto de todos os primos $p \leqslant z$, $\sum^\flat$ é a soma restrita aos números livres de quadrados, $A_d(x)$ é definido por 3.9, *$|\beta_m(x)| \leqslant 1$ e $\varepsilon(x,z)$ (que também depende de $\mathcal{A}$) satisfaz*

$$|\varepsilon(x,z)| \leqslant \frac{2}{\sqrt{z}} + exp\left(-\frac{\log x}{2\log z}\right)(\log 3x)^{21}. \tag{3.21}$$

Demonstração. Defina

$$\varepsilon(x,z) = \frac{1}{\sqrt{x}||\mathcal{A}||}\left(\sum_{p\leqslant x} a_p - \sum_{\substack{d|P(z)\\ d\leqslant\sqrt{x}}} \mu(d)A_d(x) + \sideset{}{^\flat}\sum_{(m,P(z))=1} \beta_m(x) \sum_{\substack{mp\leqslant x\\ z<p\leqslant\sqrt{x}}} a_{mp}\right)$$

[1]Seja (a_n) uma sequência de números reais não negativos, logo para todo $\alpha > 0$ temos

$$\sum_{n>x} a_n \leqslant x^{-\alpha} \sum_{n>x} a_n n^\alpha \leqslant x^{-\alpha} \sum_{n\geqslant 1} a_n n^\alpha$$

esta observação é conhecida como “Truque de Rankin”.

Vamos mostrar que $\varepsilon(x,z)$ assim definida satisfaz a condição (3.21) o que vai concluir nossa prova. De fato, primeiro tomamos o módulo na igualdade acima

$$|\varepsilon(x,z)| = \frac{1}{\sqrt{x}||\mathcal{A}||}\left|\sum_{p\leqslant x} a_p - \sum_{\substack{d|P(z)\\ d\leqslant\sqrt{x}}} \mu(d)A_d(x) + \sideset{}{^\flat}\sum_{(m,P(z))=1} \beta_m(x) \sum_{\substack{mp\leqslant x\\ z<p\leqslant\sqrt{x}}} a_{mp}\right|$$

em seguida, separamos a soma dupla considerando os casos $p|m$ e $p\nmid m$ e usando a equação (3.16), obtemos que a expressão do lado direito acima é igual a

$$\frac{1}{\sqrt{x}||\mathcal{A}||}\left|\sum_{p\leqslant\sqrt{x}} a_p + \sideset{}{^\flat}\sum_{\substack{1<n\leqslant x\\ (n,P(z))=1}} a_n + \sum_{\substack{z<p\leqslant\sqrt{x},\ mp\leqslant x\\ p|m,(m,P(z))=1}} \sideset{}{^\flat}\sum \beta_m(x)a_{mp} - \sum_{\substack{d|P(z)\\ d\leqslant\sqrt{x}}} \mu(d)A_d(x)\right|$$

Dado que a soma sobre os números livres de quadrados é igual a soma sobre todos os termos menos a soma sobre os termos não livre de quadrados, temos que expressão acima pode ser reescrita como

$$\frac{1}{\sqrt{x}||\mathcal{A}||}\left|\sum_{p\leqslant\sqrt{x}} a_p + \sum_{\substack{1<n\leqslant x\\ (n,P(z))=1}} a_n - \sum_{\substack{1<n\leqslant x\\ \forall p>z,p^2|n}} a_n + \sum_{\substack{z<p\leqslant\sqrt{x},\ mp\leqslant x\\ p|m,(m,P(z))=1}} \sideset{}{^\flat}\sum \beta_m(x)a_{mp} - \sum_{\substack{d|P(z)\\ d\leqslant\sqrt{x}}} \mu(d)A_d(x)\right|.$$

Usando a fórmula de Legendre (3.19), na expressão acima, obtemos que $|\varepsilon(x,z)|$ é igual a

$$\frac{1}{\sqrt{x}||\mathcal{A}||}\left|\sum_{p\leqslant\sqrt{x}} a_p + \sum_{\substack{d|P(z)\\ d>\sqrt{x}}} \mu(d)A_d(x) - \sum_{\substack{1<n\leqslant x\\ \forall p>z,p^2|n}} a_n + \sum_{\substack{z<p\leqslant\sqrt{x},\ mp\leqslant x\\ p|m,(m,P(z))=1}} \sideset{}{^\flat}\sum \beta_m(x)a_{mp}\right|$$

Agora mostramos como estimar cada parcela desta última expressão. Por exemplo, usando a desigualdade (3.18)

$$\left|\sum_{p\leqslant\sqrt{x}} a_p\right| \leqslant \sum_{n\leqslant\sqrt{x}} |a_n| \leqslant ||\mathcal{A}(x)||x^{\frac{1}{4}}.$$

Além disso, segue da Desigualdade de Cauchy e da estimativa (3.17) que

$$\left|\sum_{\substack{n\leqslant x\\ \forall p>z,p^2|n}} a_n\right| \leqslant \sum_{\substack{n\leqslant x\\ \forall p>z,p^2|n}} |a_n| \leqslant ||\mathcal{A}(x)||x^{\frac{1}{2}}z^{-\frac{1}{2}},$$

Por último, usando a duas estimativas que aparecem acima do enunciado do Teorema 3.1 temos

$$\left|\sum_{\substack{d|P(z)\\ d>\sqrt{x}}} \mu(d)A_d(x)\right| \leqslant ||\mathcal{A}(x)||(x\log 3x)^{\frac{1}{2}}\left(\exp\left(-\frac{\log x}{\log z}\right)(1+\log z)^{2e^3}\right)^{\frac{1}{2}}$$

$$= ||\mathcal{A}(x)||(x\log 3x)^{\frac{1}{2}}\left(\exp\left(-\frac{\log x}{2\log z}\right)\right)(1+\log z)^{e^3}$$

$$\leqslant ||\mathcal{A}(x)||x^{\frac{1}{2}}(\log 3x)^{21}\left(\exp\left(-\frac{\log x}{2\log z}\right)\right).$$

Observe que na soma dupla temos que, se $p|m$ e $n = mp$ então n não é livre de quadrados, daí a soma é nula. Aplicando a Desigualdade Triangular juntamente com as estimativas acima obtemos o resultado desejado. □

OBSERVAÇÃO 3.1. *Mais precisamente temos* 3.20 *com* $\beta_m(x) = (1+\omega(m,\sqrt{x}))^{-1}$, *onde* $\omega(m,\sqrt{x})$ *é o número de primos* $p|m$ *com* $p \leqslant \sqrt{x}$, *mas propositalmente não fizemos uso explicito desta expressão para enfatizar que isto não é necessário. A primeira soma no lado direito de* 3.20 *é do tipo* S_1 *(veja* 3.15 *com* $D = \sqrt{x}$*), enquanto à soma dupla do lado direito de* 3.20 *é da forma* 3.14 *com* $M = z$ *e* $N = \sqrt{x}$.

Portanto, vimos que o método de Vinogradov aplicado a somas sobre primos permite expressá-la em função de somas do tipo S_1 e S_2, as quais podem ser estudadas pelas ideias descritas na seção anterior.

Na próxima seção veremos as Identidades de Vaughan que junto ao método estudado aqui permitirá chegar ao objetivo deste capítulo, isto é, o Teorema de Davenport.

3.4 Identidades de Vaughan

As Identidades de Vaughan são algumas identidades encontradas por R. C. Vaughan (1977) de muita utilidade em Teoria dos Números, por exemplo, ajudam a simplificar o Método de Vinogradov para somas trigonométricas. Vale ressaltar também que elas foram muito importante na demostração do Teorema de Bombieri-Vinogradov, que é considerado o teorema mais importante da teoria multiplicativa de números. A primeira destas identidades é dada na seguinte proposição.

PROPOSIÇÃO 3.1. *Sejam* $y, z \geqslant 1$. *Logo para tudo* $n > z$ *temos*

$$\Lambda(n) = \sum_{\substack{b|n\\ b\leqslant y}} \mu(b)\log\frac{n}{b} - \sum\sum_{\substack{bc|n\\ b\leqslant y, c\leqslant z}} \mu(b)\Lambda(c) + \sum\sum_{\substack{bc|n\\ b>y, c>z}} \mu(b)\Lambda(c). \tag{3.22}$$

Demonstração. Sabemos pela fórmula de inversão de Möbius que

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

Portanto basta separar a soma acima nas parcelas onde $b \leqslant y$ e $b > y$, onde $1 \leqslant y \leqslant n$ é arbitrário. Para obter a fórmula desejada basta trabalhar na parcela onde $b > y$. Para esta soma podemos usar a identidade 1.2 e ficar com

$$\sum_{\substack{b|n \\ b>y}} \mu(b) \log \frac{n}{b} = \sum_{\substack{b|n \\ b>y}} \mu(b) \sum_{c|\frac{n}{b}} \Lambda(c) = \sum_{\substack{bc|n \\ b>y}} \sum \mu(b)\Lambda(c)$$

Portanto

$$\begin{aligned}
\Lambda(n) = \sum_{b|n} \mu(b) \log \frac{n}{b} &= \sum_{\substack{b|n \\ b>y}} \mu(b) \log \frac{n}{b} + \sum_{\substack{b|n \\ b\leqslant y}} \mu(b) \log \frac{n}{b} \\
&= \sum_{\substack{b|n \\ b\leqslant y}} \mu(b) \log \frac{n}{b} + \sum_{\substack{bc|n \\ b>y}} \sum \mu(b)\Lambda(c) \\
&= \sum_{\substack{b|n \\ b\leqslant y}} \mu(b) \log \frac{n}{b} + \sum_{\substack{bc|n \\ b>y, c>z}} \sum \mu(b)\Lambda(c) + \sum_{\substack{bc|n \\ b>y, c\leqslant z}} \sum \mu(b)\Lambda(c) \\
&= \sum_{\substack{b|n \\ b\leqslant y}} \mu(b) \log \frac{n}{b} + \sum_{\substack{bc|n \\ b>y, c>z}} \sum \mu(b)\Lambda(c) - \sum_{\substack{bc|n \\ b\leqslant y, c\leqslant z}} \sum \mu(b)\Lambda(c).
\end{aligned}$$

A última igualdade vêm do fato que para $n > z$ temos

$$\sum_{\substack{bc|n \\ c\leqslant z}} \sum \mu(b)\Lambda(c) = \sum_{c\leqslant z} \Lambda(c) \sum_{b|\frac{n}{c}} \mu(b) = \sum_{n\leqslant z} \Lambda(n) = 0.$$

$\square$

Similarmente, existe uma identidade envolvendo a função de Möbius, está é dada na seguinte proposição.

Proposição 3.2. *Sejam $y, z \geqslant 1$. Se $m > \max(y,z)$ então*

$$\mu(m) = -\sum_{\substack{bc|m \\ b\leqslant y, c\leqslant z}} \sum \mu(b)\mu(c) + \sum_{\substack{bc|m \\ b>y, c>z}} \sum \mu(b)\mu(c) \tag{3.23}$$

Demonstração. Pela identidade de Möbius

$$\sum_{bc|m}\sum \mu(b)\mu(c) = \sum_{b|m}\mu(b)\sum_{c|\frac{m}{b}}\mu(c) = \mu(m). \tag{3.24}$$

Decompondo a soma do lado esquerdo de (3.24) em quatro somas onde $b \leqslant y$, $b > y$, $c \leqslant z$, $c > z$, temos

$$\mu(m) = \sum_{\substack{bc|m \\ b\leqslant y, c\leqslant z}}\sum \mu(b)\mu(c) + \sum_{\substack{bc|m \\ b\leqslant y, c>z}}\sum \mu(b)\mu(c) + \sum_{\substack{bc|m \\ b>y, c\leqslant z}}\sum \mu(b)\mu(c) + \sum_{\substack{bc|m \\ b>y, c>z}}\sum \mu(b)\mu(c). \tag{3.25}$$

Levando em conta que por hipótese $m > \max(y, z)$ temos

$$\sum_{\substack{bc|m \\ b\leqslant y}}\sum \mu(b)\mu(c) = \sum_{\substack{b|m \\ b\leqslant y}}\mu(b)\sum_{c|\frac{m}{b}}\mu(c) = \sum_{m\leqslant y}\mu(m) = 0.$$

Assim, separando a soma do lado esquerdo da equação acima, com as restrições $c \leqslant z$ e $c > z$ temos

$$\sum_{\substack{bc|m \\ b\leqslant y, c>z}}\sum \mu(b)\mu(c) = -\sum_{\substack{bc|m \\ b\leqslant y, c\leqslant z}}\sum \mu(b)\mu(c).$$

Analogamente podemos mostrar que

$$\sum_{\substack{bc|m \\ b>y, c\leqslant z}}\sum \mu(b)\mu(c) = -\sum_{\substack{bc|m \\ b\leqslant y, c\leqslant z}}\sum \mu(b)\mu(c).$$

Substituindo estas igualdades em (3.25), obtemos o resultado desejado. □

3.5 Teorema de H. Davenport

As somas exponenciais desempenham um papel importante no desenvolvimento da Teoria de Números desde Gauss, quando foram utilizados para provar a lei da reciprocidade quadrática. Como obtemos uma soma exponencial? Consideramos o seguinte problema: temos uma função f definida num conjunto finito $\mathcal{A}$ e que assume valores no conjunto de inteiros não negativos menores que um certo N. Queremos medir, de alguma forma, que tão bem distribuídos estão os valores de dita função. Diremos que os valores estão bem distribuídos se a função assume cada possível valor, aproximadamente o mesmo número de vezes. A chave está em que entendemos por aproximadamente, e como medir isto. Vamos ver um exemplo. Na seguinte tabela, tomamos $N = 12$

n	$\#\{f(x) = n\}$
0	12
1	8
2	9
3	10
4	5
5	12
6	7
7	4
8	11
9	12
10	14
11	9

Tabela 3.1

Para medir a distribuição de uma forma quantitativa, dividimos a circunferência unitária em 12 partes iguais, e atribuímos a cada subdivisão um "peso" correspondente ao número de vezes que a função assume um dado valor, como vemos na figura abaixo

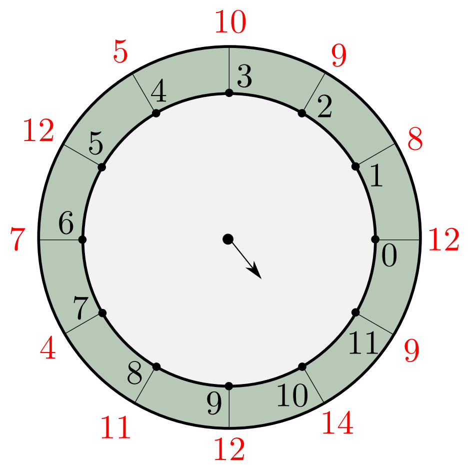


Figura 3.1: No exemplo acima o valor 2 é assumido 9 vezes.

O vetor mostrado nesta figura é a média de todos os vetores que ligam o origem com cada uma das subdivisões da circunferência, tomado com pesos indicado em vermelho. Intuitivamente, quanto mais distribuídos os valores da função, menor seria o módulo do vetor média. Porém, o fato que o vetor média seja pequeno não é suficiente para que os valores da função estejam bem distribuídos, pois se os valores se concentram ao redor de dois pontos opostos, eles podem se cancelar, dando como resultado uma média pequena. Por outro lado, se distribuímos os valores da função contando os pontos de dois em dois, a média obtida pode ser grande. Por isso, para obter um conceito razoável de boa distribuição, precisamos que o vetor média seja pequeno se distribuirmos os valores contando os pontos em blocos de vários tamanhos.

Como calcular o módulo do vetor média algebricamente? Para isso uma boa ideia é usar os números complexos. Os pontos marcados na circunferência unitária são dados por ξ^i, $i = 0, 1, \ldots, N-1$ onde

$$\xi = \exp\left(\frac{2\pi i}{N}\right) = \cos\left(\frac{2\pi}{N}\right) + i\text{sen}\left(\frac{2\pi}{N}\right)$$

é uma raiz N-ésima da unidade em $\mathbb{C}$. Se denotamos por $\psi(k) = \xi^k$ para todo k inteiro e distribuímos os valores contando os pontos de c em c, onde $c = 1, 2, \ldots, N-1$ o vetor média é dado por

$$\frac{1}{\#\mathcal{A}} \sum_{k=0}^{N-1} \#\{x \in \mathcal{A} : f(x) = k\} \cdot \psi(ck) = \frac{1}{\#\mathcal{A}} \sum_{x \in \mathcal{A}} \psi(cf(x)).$$

Este tipo de somas são chamadas somas exponenciais. Se os valores da função estão bem distribuídos, esta soma seria pequena em valor absoluto para todo c. Por isso é bastante relevante encontrar cotas para tais somas.

Agora como um exemplo das numerosas aplicações do método descrito nas seções anteriores provamos a estimativa de H. Davenport para somas exponenciais.

Seja $\alpha \in \mathbb{R}$. Pelo Teorema de Aproximação Diofantina de Dirichlet (veja o Apêndice, Teorema 5.1) existem inteiros a, q tais que

$$\left|\alpha - \frac{a}{q}\right| < \frac{1}{q^2} \quad \text{com} \quad (a, q) = 1. \tag{3.26}$$

Usando esta aproximação, vamos deduzir uma cota para a soma

$$S(\alpha; x) = \sum_{n \leqslant x} e(\alpha n)\mu(n). \tag{3.27}$$

TEOREMA 3.2. *Suponha que* α *satisfaz* (3.26). *Então para* $x \geqslant 2$ *temos*

$$S(\alpha, x) \ll (q^{\frac{1}{2}} x^{\frac{1}{2}} + q^{-\frac{1}{2}} x + x^{\frac{4}{5}})(\log x)^3, \tag{3.28}$$

onde a constante implicada é absoluta.

Antes de apresentar a prova deste teorema, vamos mostrar alguns fatos sobre somas exponenciais. Com o intuito de provar o Teorema 3.2, vamos estimar a soma exponencial abaixo, como segue: primeiro passo usamos a cota trivial, isto é,

$$\left|\sum_{1 \leqslant n \leqslant N} e(\alpha n)\right| \leqslant \sum_{1 \leqslant n \leqslant N} |e(\alpha n)| = \sum_{1 \leqslant n \leqslant N} 1 = N.$$

Em seguida, observamos que se $\alpha = 0$ a soma acima é exatamente N. Por outro lado, se $\alpha \neq 0$

$$\sum_{1\leqslant n\leqslant N} e(\alpha n) = \frac{e(\alpha(N+1)) - e(\alpha)}{e(\alpha)-1} = \frac{e(\frac{\alpha N}{2}+\alpha)\left[e(\frac{\alpha N}{2}) - e(-\frac{\alpha N}{2})\right]}{e(\frac{\alpha}{2})\left[e(\frac{\alpha}{2}) - e(-\frac{\alpha}{2})\right]}$$
$$= \frac{\operatorname{sen}(\pi\alpha N)}{\operatorname{sen}(\pi\alpha)} e\left(\frac{\alpha}{2}(N+1)\right).$$

Definimos a função distância ao inteiro mais próximo $\|\cdot\| : \mathbb{R} \to [0, 1/2]$, como segue

$$\|\alpha\| = \frac{1}{2} - \left|\{\alpha\} - \frac{1}{2}\right|.$$

Uma vez que a função inteiro mais próximo é periódica de período 1 em $\mathbb{R}$ basta entendê-la no intervalo $[0, 1]$. Note que $|sen(\pi\alpha)| \geqslant 2\|\alpha\|$, obtemos

$$\left|\sum_{1\leqslant n\leqslant N} e(\alpha n)\right| = \left|\frac{sen(\pi\alpha N)}{sen(\pi\alpha)}\right| \leqslant \frac{1}{|sen(\pi\alpha)|} \leqslant \frac{1}{2\|\alpha\|}, \quad \alpha \notin \mathbb{Z}.$$

Portanto segue da estimativa trivial e da anterior que

$$\left|\sum_{1\leqslant n\leqslant N} e(\alpha n)\right| \leqslant \min\left(N, \frac{1}{2\|\alpha\|}\right).$$

Tomemos, em seguida, um número real $x(m) > 0$, pela estimativa acima temos que

$$\left|\sum_{1\leqslant n\leqslant N} e(\alpha m n)\right| \leqslant \min\left(N, \frac{1}{2\|\alpha m\|}\right).$$

Agora, dado que a exponencial complexa tem módulo um, segue que

$$\sum_{|m|\leqslant M} \left|\sum_{\substack{1\leqslant n\leqslant N\\ n\leqslant x(m)}} e(\alpha m n)\right| \leqslant \sum_{|m|\leqslant M} \min\left(N, x(m), \frac{1}{2\|\alpha m\|}\right). \tag{3.29}$$

Nosso próximo passo é estimar a soma

$$\sum_{|m|\leqslant M} \min\left(N, \frac{1}{\|\alpha m\|}\right).$$

Para isso, dividimos a soma sobre m em blocos de q termos consecutivos (com talvez um bloco incompleto). Observe que o número de tais blocos é menor ou igual que $1 + 2Mq^{-1}$. Vamos mostrar primeiro como obter a estimativa em um dos blocos. A soma restrita a um bloco é dada por

$$\sum_{m=0}^{q-1} \min\left(N, \frac{1}{\|\alpha(m_1+m)\|}\right),$$

onde m_1 é o primeiro inteiro pertencente ao bloco. Da condição 3.26 e do fato que $0 \leqslant m < q$ segue que

$$\alpha(m_1+m) = \alpha m_1 + \frac{am}{q} + O\left(\frac{1}{q}\right).$$

Como m percorre de 0 a $q-1$ e $(a,q)=1$, o número am atravessa o conjunto completo de resíduos $\mod q$. Pondo $am \equiv r(\mod q)$ a soma sobre o bloco pode ser reescrita como

$$\sum_{r=0}^{q-1} \min\left(N, \frac{1}{||(r+b)/q + O(1/q)||}\right),$$

onde b é o inteiro mais próximo de $q\alpha m_1$. Existe uma quantidade finita de valores de r na soma para os quais a segunda expressão no mínimo é maior ou igual a N, aqueles para os quais o mínimo resíduo absoluto é pequeno. Para estes, devemos tomar N como o valor de aquele mínimo. Para os outros valores de r, se l denota o mínimo resíduo absoluto de $(r+b)(\mod q)$ temos

$$\left\|\frac{r+b}{q} + O\left(\frac{1}{q}\right)\right\| \gg \frac{l}{q}.$$

Portanto a soma acima satisfaz

$$\sum_{r=0}^{q-1} \min\left(N, \frac{1}{\|(r+b)/q + O(1/q)\|}\right) \ll N + \sum_{l=1}^{q/2} \frac{q}{l}.$$

Usando as cotas estabelecidas acima, encontramos uma estimativa para a soma sobre todos os blocos, que é dada por

$$\begin{aligned}\sum_{|m|\leqslant M} \min\left(N, \frac{1}{2\|\alpha m\|}\right) &\ll (1+2Mq^{-1})\left(N + \sum_{1\leqslant l\leqslant q} ql^{-1}\right)\\ &= N + 2MNq^{-1} + (2M+q)\sum_{1\leqslant l\leqslant q} l^{-1}\\ &\leqslant N + 2MNq^{-1} + (2M+q)\log q.\end{aligned}$$

A última desigualdade devido ao fato que

$$\sum_{1\leqslant l\leqslant q} l^{-1} \leqslant \int_1^q \frac{dt}{t} = \log q.$$

Portanto

$$\sum_{|m|\leqslant M}\min\left(N,\frac{1}{2||\alpha m||}\right)\ll(M+N+MNq^{-1}+q)\log 2q. \tag{3.30}$$

Similarmente para $x(m)=x/m$ encontramos uma estimativa seguindo os mesmos passos dados acima (considerando separadamente o caso em que $1\leqslant m\leqslant q/2$ e seu complementar no intervalo $[1,M]$), isto é

$$\sum_{1\leqslant m\leqslant M}\min\left(\frac{x}{m},\frac{1}{2||\alpha m||}\right)\ll(M+xq^{-1}+q)\log 2qx. \tag{3.31}$$

Juntando todas estas cotas obtemos o seguinte lema.

Lema 3.1. *Para qualquer número $x(m)>0$ temos*

$$\sum_{|m|\leqslant M}\left|\sum_{\substack{1\leqslant n\leqslant N\\ n\leqslant x(m)}}e(\alpha mn)\right|\ll(M+N+MNq^{-1}+q)\log 2q, \tag{3.32}$$

$$\sum_{1\leqslant m\leqslant M}\left|\sum_{mn\leqslant x}e(\alpha mn)\right|\ll(M+xq^{-1}+q)\log 2qx. \tag{3.33}$$

Demonstração. Para provar (3.32), simplesmente basta usar as estimativas (3.29) e (3.30). Enquanto a prova de (3.33) segue diretamente de (3.29) e (3.31). □

Agora, vamos procurar cotas para formas bilineares gerais, como

$$B(x;N)=\sum_{N\leqslant n\leqslant 2N}\left|\sum_{mn\leqslant x}\gamma_m e(\alpha mn)\right|$$

onde γ_m são números complexos com $|\gamma_m|\leqslant 1$. Usando a Desigualdade de Cauchy

$$\begin{aligned}
B^2(x;N)&\leqslant\left(\sum_{N<n\leqslant 2N}1\right)\left(\sum_{N<n\leqslant 2N}\left(\sum_{mn\leqslant x}\gamma_m e(\alpha mn)\right)^2\right)\\
&\leqslant 2N\sum\sum_{1\leqslant m_1\leqslant m_2\leqslant x/N}|\gamma_{m_1}\gamma_{m_2}|\left|\sum_{\substack{N<n\leqslant 2N\\ nm_2\leqslant x}}e(\alpha m_1 n)\overline{e(\alpha m_2 n)}\right|\\
&\leqslant 2N\sum\sum_{1\leqslant m_1\leqslant m_2\leqslant x/N}\left|\sum_{\substack{N<n\leqslant 2N\\ nm_2\leqslant x}}e(\alpha(m_1-m_2)n)\right|.
\end{aligned}$$

Vamos aplicar (3.32) com $M = \frac{x}{N}$, a soma do lado direito da desigualdade acima para obter

$$2N \sum_{1 \leqslant m_1 \leqslant m_2 \leqslant x/N} \sum \left| \sum_{\substack{N \leqslant n \leqslant 2N \\ nm_2 \leqslant x}} e(\alpha(m_1 - m_2)n) \right| \ll (\frac{x}{N} + N + \frac{x}{q} + q) x \log 2q.$$

desta cota deduzimos o seguinte resultado.

LEMA 3.2. *Para quaisquer números complexos α_m, β_n com $|\alpha_m| \leqslant 1$, $|\beta_n| \leqslant 1$, temos*

$$\sum_{\substack{mn \leqslant x \\ m > M, n > N}} \sum \alpha_m \beta_n e(\alpha mn) \ll (\frac{x}{M} + \frac{x}{N} + \frac{x}{q} + q)^{\frac{1}{2}} x^{\frac{1}{2}} (\log x)^2. \tag{3.34}$$

Demonstração. Pela desigualdade de Cauchy e pela estimativa anterior temos

$$\begin{aligned} \Big| \sum_{\substack{mn \leqslant x \\ m > M, n > N}} \sum \alpha_m \beta_n e(\alpha mn) \Big|^2 &\leqslant \Big(\sum_{N < n < x/M} |\beta_n|^2 \Big) \sum_{M \leqslant m_1 \leqslant m_2 \leqslant x/N} \sum |\alpha_{m_1} \alpha_{m_2}| \times \\ &\quad \times \left| \sum_{\substack{N < n < x/M \\ nm_2 \leqslant x}} e(\alpha(m_1 - m_2)n) \right| \\ &\leqslant x \sum_{M \leqslant m_1 \leqslant m_2 \leqslant x/N} \sum \left| \sum_{\substack{N < n < x/M \\ nm_2 \leqslant x}} e(\alpha(m_1 - m_2)n) \right| \\ &\ll (\frac{x}{M} + \frac{x}{N} + \frac{x}{q} + q) x \log 2q. \end{aligned}$$

Notando que $2q \leqslant x$ para x suficientemente grande, basta tomar a raiz quadrada para obter o resultado desejado. □

Passamos agora à demonstração do Teorema 3.2

Demonstração. Pela Identidade de Vaughan

$$\Lambda(n) = \sum_{\substack{b|n \\ b \leqslant y}} \mu(b) \log \frac{n}{b} - \sum_{\substack{bc|n \\ b \leqslant y, c \leqslant z}} \sum \mu(b) \Lambda(c) + \sum_{\substack{bc|n \\ b > y, c > z}} \sum \mu(b) \Lambda(c).$$

Substituindo esta identidade na soma $S(\alpha;x)$ temos

$$\begin{aligned}
S(\alpha;x) &= \sum_{n\leqslant x} e(\alpha n)\Lambda(n)\\
&= \sum_{n\leqslant x} e(\alpha n)\Big[\sum_{\substack{b|n\\ b\leqslant y}} \mu(b)\log\frac{n}{b} - \sum\sum_{\substack{bc|n\\ b\leqslant y, c\leqslant z}} \mu(b)\Lambda(c) + \sum\sum_{\substack{bc|n\\ b>y, c>z}} \mu(b)\Lambda(c)\Big]\\
&= \sum_{n\leqslant x}\sum_{\substack{b|n\\ b\leqslant y}} e(\alpha n)\mu(b)\log\frac{n}{b} - \sum_{n\leqslant x}\sum\sum_{\substack{bc|n\\ b\leqslant y, c\leqslant z}} e(\alpha n)\mu(b)\Lambda(c)\\
&\quad + \sum_{n\leqslant x}\sum\sum_{\substack{bc|n\\ b>y, c>z}} e(\alpha n)\mu(b)\Lambda(c)\\
&= \sum\sum_{\substack{lm\leqslant x\\ m\leqslant M}} e(\alpha lm)\mu(m)\log l - \sum\sum\sum_{\substack{lmn\leqslant x\\ m\leqslant M, n\leqslant N}} e(\alpha lmn)\Lambda(n)\mu(m)\\
&\quad + \sum\sum\sum_{\substack{lmn\leqslant x\\ m>M, n>N}} e(\alpha lmn)\Lambda(n)\mu(m) + O(N)
\end{aligned}$$

Escolhemos $M = N = x^{\frac{2}{5}}$ e estimamos a s duas primeiras somas da última igualdade acima aplicando (3.33), o que nos fornece

$$\begin{aligned}
\Bigg|\sum\sum_{\substack{lm\leqslant x\\ m\leqslant M}} e(\alpha lm)\mu(m)\log l\Bigg| &\leqslant \sum_{m\leqslant M}\Bigg|\sum_{lm\leqslant x} e(\alpha lm)\log l\Bigg| \leqslant \log x\sum_{m\leqslant M}\Bigg|\sum_{lm\leqslant x} e(\alpha lm)\Bigg|\\
&\ll \log x(x^{\frac{2}{5}} + xq^{-1} + q)\log 2qx\\
&\ll (x^{\frac{4}{5}} + xq^{-1} + q)(\log x)^2.
\end{aligned}$$

Da mesma maneira a segunda soma podemos estimar como segue

$$\begin{aligned}
\Bigg|\sum\sum\sum_{\substack{lmn\leqslant x\\ m\leqslant M, n\leqslant N}} e(\alpha lmn)\Lambda(n)\mu(m)\Bigg| &\leqslant \sum_{m\leqslant M}\Bigg|\sum_{\substack{lmn\leqslant x\\ n\leqslant N}} e(\alpha lmn)\Lambda(n)\Bigg|\\
&\leqslant \log x\sum_{m\leqslant M}\Bigg|\sum_{mn\leqslant x} e(\alpha lmn)\Bigg|\\
&\ll \log x(x^{\frac{2}{5}} + xq^{-1} + q)\log 2qx\\
&\ll (x^{\frac{4}{5}} + xq^{-1} + q)(\log x)^2.
\end{aligned}$$

Para estimar a última soma de $S(\alpha; x)$, vamos transformá-la numa Forma Bilinear e assim aplicar o Lema 3.2. Para isso fazemos a mudança de variável $ln = k$ e definimos $c(k)$ como segue

$$c(k) = \sum_{\substack{n|k, n>N \\ k>N}} \Lambda(n) \leqslant \sum_{n|k} \Lambda(n) = \log k.$$

Desta forma $c(k)/\log k \leqslant 1$, e aplicando 3.34, temos

$$\begin{aligned}
\left| \sum\sum\sum_{\substack{lmn\leqslant x \\ m>M, n>N}} e(\alpha lmn)\Lambda(n)\mu(m) \right| &= \left| \sum\sum\sum_{\substack{mk\leqslant x, k>N \\ m>M, n>N, n|k}} e(\alpha mk)\Lambda(n)\mu(m) \right| \\
&= \left| \sum\sum_{\substack{mk\leqslant x \\ m>M, k>N}} \left(\sum_{n>N, n|k} \Lambda(n) \right) e(\alpha mk)\mu(m) \right| \\
&= \log x \left| \sum\sum_{\substack{mk\leqslant x \\ m>M, k>N}} \frac{c(k)}{\log k} e(\alpha mk)\mu(m) \right| \\
&\ll (x^{\frac{3}{5}} + xq^{-1} + q)^{\frac{1}{2}} x^{\frac{1}{2}} (\log x)^3.
\end{aligned}$$

Juntando as estimativas para as três somas acima obtemos

$$\begin{aligned}
S(\alpha; x) &\ll \left(x^{\frac{4}{5}} + xq^{-1} + q \right) (\log x)^2 + \left(x^{\frac{3}{5}} + xq^{-1} + q \right)^{\frac{1}{2}} x^{\frac{1}{2}} (\log x)^3 \\
&\ll \left(q^{\frac{1}{2}} x^{\frac{1}{2}} + q^{-\frac{1}{2}} x + x^{\frac{4}{5}} \right) (\log x)^3.
\end{aligned}$$

Isto conclui a prova. $\square$

Levando em conta as mesmas ideias da demostração acima, mas usando a Identidade de Vaughan para a função de Möbius em lugar da função de Von Mangoldt, obtemos o seguinte resultado.

TEOREMA 3.3. *Suponha que α satisfaz* (3.26). *Logo para $x \geqslant 2$ temos*

$$\sum_{m\leqslant x} \mu(m)e(\alpha m) \ll \left(q^{\frac{1}{2}} x^{\frac{1}{2}} + q^{-\frac{1}{2}} + x^{\frac{4}{5}} \right)^{\frac{1}{2}} x^{\frac{1}{2}} (\log x)^4. \tag{3.35}$$

Neste parte do trabalho vamos lembrar a resultado obtido no capítulo anterior. Pela observação 2.1 do Teorema do Número Primo Generalizado temos

$$\Psi(f,x) = \sum_{n \leqslant x} \Lambda_f(n) = rx - \frac{x^{\beta_f}}{\beta_f} + O\left(\sqrt{\mathfrak{q}(f)}x \exp\left(-\frac{c}{2d^4}\sqrt{\log x}\right)\right).$$

Assim para um caráter primitivo de Dirichlet χ módulo q, sabemos que sua L-função é de ordem 1 (χ é completamente multiplicativo). Além disso, é bem conhecido que se χ é não trivial então $L(\chi, s)$ é inteira e caso contrario ela tem um polo simples em $s = 1$ com resíduo 1 (veja o Apêndice, Teorema 5.2) Portanto temos que

$$\sum_{n \leqslant x} \chi(n)\Lambda(n) = \delta_\chi x - \frac{x^{\beta_\chi}}{\beta_\chi} + O\left(\sqrt{q}x \exp\left(-\frac{c}{2}\sqrt{\log x}\right)\right), \tag{3.36}$$

onde $\delta_\chi = 1$ se χ é o carácter trivial e, $\delta_\chi = 0$ caso contrario.

Na sequência vamos ver um teorema que nos permitirá tirar a parte principal da estimativa acima. Não apresentaremos aqui a demonstração deste importante resultado devido a tecnicalidade envolvida em sua prova (para maiores detalhes, veja [9], pag. 122 e 123).

Teorema 3.4 (Siegel). *Para qualquer caráter de Dirichlet real χ módulo q, temos*

$$\beta_\chi \geqslant 1 - \frac{c(\varepsilon)}{q^\varepsilon},$$

para qualquer $\varepsilon > 0$, com uma constante $c(\varepsilon) > 0$ dependendo só de ε. Se $\varepsilon < \frac{1}{2}$ o valor explícito desta constante não é conhecido.

Um resultado de muito interesse para nosso estudo se deduz diretamente do teorema anterior, mais explicitamente, temos a seguinte estimativa.

Corolário 3.1. *Para um carácter primitivo de Dirichlet χ modulo $q > 2$, temos*

$$\sum_{p \leqslant x} \chi(p) \ll \sqrt{q}x(\log x)^{-A}, \tag{3.37}$$

$$\sum_{n \leqslant x} \chi(n)\mu(n) \ll \sqrt{q}x(\log x)^{-A} \tag{3.38}$$

para todo $A > 0$. A constante implicada depende só de A.

Demonstração. Dado que o carácter é não trivial temos em (3.36) que $\delta_\chi = 0$. Além do mais, do Teorema de Siegel acima tem-se que o zero excepcional é uniformemente limitado então o segundo termo do lado direito de (3.36) é limitado por $x/(\log x)^A$ com $A > 0$, desta forma temos

$$\sum_{n \leqslant x} \chi(n)\Lambda(n) \ll \sqrt{q}x(\log x)^{-A}.$$

Agora observe que

$$\sum_{p\leqslant x}\chi(p)\log p+\sum_{p^k\leqslant x,k\geqslant 2}\chi(p^k)\log p=\sum_{n\leqslant x}\chi(n)\Lambda(n).$$

Podemos estimar a segunda soma do lado esquerdo da equação acima, como segue

$$\left|\sum_{p^k\leqslant x,k\geqslant 2}\chi(p^k)\log p\right|\leqslant\sum_{p^k\leqslant x,k\geqslant 2}\log p\leqslant\sum_{p\leqslant\sqrt{x}}\log p\leqslant\sqrt{x}\log x.$$

Juntando as estimativas acima temos

$$\sum_{p\leqslant x}\chi(p)\log p\ll\frac{\sqrt{q}x}{(\log x)^A}+\sqrt{x}\log x.$$

De esta forma obtemos a seguinte cota

$$\sum_{p\leqslant x}\chi(p)\log p\ll\sqrt{q}x(\log x)^{-A}. \tag{3.39}$$

Finalmente fazendo uso da Fórmula de Abel, temos

$$\sum_{p\leqslant x}\chi(p)\log p=\log x\sum_{p\leqslant x}\chi(p)-\chi(2)\log 2-\int_2^x\sum_{p\leqslant t}\chi(p)\frac{dt}{t}.$$

Desta última igualdade e da equação 3.39 segue que

$$\sum_{p\leqslant x}\chi(p)\log x\ll\sqrt{q}x(\log x)^{-A}+x.$$

Se dividimos por $\log x$ e redefinimos a constante A, obtemos 3.37.

Agora vamos provar 3.38. Para isso escrevemos

$$\sum_{p\leqslant x}\chi(p)+\sum_{n\leqslant x}\chi(n)\mu(n)=1-\sum_{1<m\leqslant x}\chi(m)\mu(m)\sum_{p_m<p\leqslant x/m}\chi(p),$$

onde p_m representa o maior divisor primo de m.

Seja $r\geqslant 0$ o número de divisores primos de m. Pelas restrições que aparecem na soma interna do lado direito temos que $m^{1/r}\leqslant p_m<x/m$, assim $m<x^{r/(r+1)}$. Portanto usando a estimativa 3.37, a soma interna sobre p satisfaz

$$\sum_{p\leqslant\frac{x}{m}}\chi(p)\ll\frac{\sqrt{q}x}{m}\left(\log\frac{x}{m}\right)^{-A}.$$

Observando que $\log m < \frac{r}{r+1}\log x$ segue que $\log x < (r+1)\log\frac{x}{m}$ e dado que $(r+1)^A \ll 2^r = \sigma_0(m)$ concluímos que

$$\sum_p \chi(p) \ll \frac{\sqrt{q}x}{m}\left(\frac{r+1}{\log x}\right)^A \ll \frac{\sigma_0(m)\sqrt{q}x}{m(\log x)^A}.$$

Finalmente, somando sobre m obtemos a estimativa (3.38) com um fator extra $(\log x)^2$, reescrevendo A obtemos dito resultado. $\square$

Usando as somas de Gauss podemos obter estimativas para a correlação da função de Möbius com caracteres no grupo aditivo $(\mathbb{Z}/q\mathbb{Z})^*$ de classes residuais módulo m. Usando (5.3) e as relações de ortogonalidade dos caráteres temos

$$\begin{aligned}\sum_{m\leqslant x}\mu(m)e\left(\frac{am}{q}\right) &= \sum_{m\leqslant x}\mu(m)\frac{1}{\varphi(q)}\sum_{\overline{\chi}(\bmod q)}\chi(am)\tau(\overline{\chi})\\ &= \frac{1}{\varphi(q)}\sum_{m\leqslant x}\mu(m)\chi(m)\sum_{\overline{\chi}(\bmod q)}\chi(a)\sum_{b(\bmod q)}\overline{\chi}(b)e\left(\frac{b}{q}\right)\\ &= \frac{1}{\varphi(q)}\sum_{m\leqslant x}\mu(m)\chi(m)\sum_{b(\bmod q)}e\left(\frac{b}{q}\right)\sum_{\overline{\chi}(\bmod q)}\chi(a)\overline{\chi}(b)\\ &= \frac{1}{\varphi(q)}\sum_{m\leqslant x}\mu(m)\chi(m)\varphi(q)e\left(\frac{a}{q}\right)\\ &= \sum_{m\leqslant x}\mu(m)\chi(m)e\left(\frac{a}{q}\right).\end{aligned}$$

Portanto usando a estimativa (3.38) obtemos

$$\sum_{m\leqslant x}\mu(m)e\left(\frac{am}{q}\right) \ll qx(\log x)^{-A}. \tag{3.40}$$

Agora seja α um número real. Pelo Teorema da Aproximação Diofantina, dado $Q \geqslant 1$ existe um número racional a/q com $(a,q)=1$, $1\leqslant q\leqslant Q$ tais que

$$\left|\alpha - \frac{a}{q}\right| \leqslant \frac{1}{qQ}. \tag{3.41}$$

Usando novamente a Fórmula de Abel, temos

$$\begin{aligned}\sum_{m\leqslant x}\mu(m)e(\alpha m) &= \sum_{m\leqslant x}\mu(m)e\left(\frac{am}{q}\right)e\left(\alpha m-\frac{am}{q}\right)\\ &= e\left(\alpha x-\frac{ax}{q}\right)\sum_{m\leqslant x}\mu(m)e\left(\frac{am}{q}\right)-e\left(\alpha-\frac{a}{q}\right)e\left(\frac{a}{q}\right)\\ &\quad -2\pi i\int_1^x\left(\sum_{m\leqslant t}\mu(m)e\left(\frac{am}{q}\right)\right)\left(\alpha-\frac{a}{q}\right)e\left(\alpha t-\frac{at}{q}\right)dt.\end{aligned}$$

Portanto usando (3.41) obtemos a seguinte desigualdade

$$\left|\sum_{m\leqslant x}\mu(m)e(\alpha m)\right|\leqslant\left(1+\frac{2\pi x}{qQ}\right)\left|\sum_{m\leqslant y}\mu(m)e\left(\frac{am}{q}\right)\right|,$$

para algum $1\leqslant y\leqslant x$. Usando a estimativa (3.40), na desigualdade anterior obtemos

$$\begin{aligned}\sum_{m\leqslant x}\mu(m)e(\alpha m) &\ll \left(1+\frac{2\pi x}{qQ}\right)\left(q\ y\ (\log y)^{-A}\right)\\ &\ll \left(q+\frac{x}{Q}\right)x(\log x)^{-5A}.\end{aligned}\tag{3.42}$$

para $x\geqslant 2$ e $A\geqslant 0$.

Finalmente, usando o fato que $q\leqslant Q$, podemos estimar a soma (3.1) como segue: Primeiro separemos a soma em questão em dois casos.

Primeiro caso: se $q\leqslant xQ^{-1}$ da estimativa (3.42) temos

$$\begin{aligned}\sum_{m\leqslant x}\mu(m)e(\alpha m) &\ll \left(q+\frac{x}{Q}\right)x(\log x)^{-5A}\\ &\leqslant Q^{-1}x^2(\log x)^{-5A}.\end{aligned}$$

Segundo caso: se $xQ^{-1}<q\leqslant Q$ do Teorema 3.3 segue que

$$\begin{aligned}\sum_{m\leqslant x}\mu(m)e(\alpha m) &\ll (q^{\frac{1}{2}}x^{\frac{1}{2}}+q^{-\frac{1}{2}}+x^{\frac{4}{5}})^{\frac{1}{2}}x^{\frac{1}{2}}(\log x)^4\\ &\leqslant Q^{\frac{1}{4}}x^{\frac{3}{4}}(\log x)^4+x^{\frac{9}{10}}(\log x)^4.\end{aligned}$$

Portanto somando estas estimativas obtemos

$$\sum_{m\leqslant x}\mu(m)e(\alpha m)\ll Q^{-1}x^2(\log x)^{-5A}+Q^{\frac{1}{4}}x^{\frac{3}{4}}(\log x)^4+x^{\frac{9}{10}}(\log x)^4.$$

Fazendo a escolha $Q=x(\log x)^{-4A}$, temos o seguinte resultado, que torna-se o teorema principal deste capítulo.

Teorema 3.5. *Para qualquer número real α e $x \geqslant 2$ temos*

$$\sum_{m \leqslant x} \mu(m) e(\alpha m) \ll x (\log x)^{-A}. \tag{3.43}$$

para todo $A > 0$, onde a constante na expansão assintótica só depende de A.

Como falamos anteriormente, esta estimativa foi provada pela primeira vez em 1937 por H. Davenport ([6]), mas fazendo uso de outras ferramentas da Teoria dos Números, já que a demonstração apresentada aqui faz uso de métodos não existentes naquela época.

No próximo capítulo, vamos estudar a ortogonalidade de Möbius de um ponto de vista diferente. Com os resultado obtidos nela a ortogonalidade da função com a função constante igual a um (equivalente ao TNP) e o Teorema de Davenport se tornarão casos particulares.

Capítulo 4

Aspecto Dinâmico da Função de Möbius

No capítulo anterior, apresentamos o Principio de Aleatoriedade de Möbius. De uma maneira informal, ele diz que a função de Möbius não está correlacionada com nenhuma sequência $\xi(n)$ chamada de, naquele momento, "razoável". Por não correlacionada entenderemos

$$\sum_{n \leqslant x} \mu(n)\xi(n) = o\left(\sum_{n \leqslant x} |\xi(n)|\right). \tag{4.1}$$

O objetivo deste capítulo é ver alguns aspectos do Principio de Aleatoriedade identificando o que entendemos por sequências "razoáveis". Com objetivo de definir este conceito, poderíamos usar as ideias e notação de complexidade computacional. Isto é, ξ é "razoável" ou de baixa complexidade se é de classe polinomial P, ou seja, cada valor de $\xi(n)$ pode ser calculado em $O((\log n)^B)$ passos para algum $B = B(\xi)$.

Isto nos leva a questionar se $\mu \in P$. Esta é uma pergunta interessante, mas sua resposta ainda permanece em aberto. Supondo que $\mu \in P$, temos uma função $\xi(n) \in P$ para o qual o Principio de Aleatoriedade (4.1) não é satisfeito. Para ver isto, basta definir a função dada por $\xi(n) = -1$ se n é primo e zero caso contrário. Então é um resultado recente que $\xi(n) \in P$, veja [1]. Além disso,

$$\frac{1}{x}\sum_{n \leqslant x} \xi(n)\mu(n) = \frac{1}{x}\sum_{n \leqslant x} |\xi(n)|1 \sim \frac{1}{\log x}.$$

Esta observação, mostra que o ponto de vista de complexidade computacional não é a noção correta para formular-se o Princípio da Aleatoriedade de Möbius ou $\mu -$ aleatoriedade.

Portanto temos que procurar outro ponto de vista para estudar este conceito de aleatoriedade. Será mais conveniente para nossos propósitos trabalhar com sequências $\xi(n)$ limitadas e estudar um caso particular desta aleatoriedade, a chamada de ortogonalidade, lembramos que isto é

$$\sum_{n\leqslant x}\mu(n)\xi(n)=o(x) \tag{4.2}$$

Neste contexto vimos nos capítulos anteriores que as funções $\xi(n)=1$ (equivalência do TNP) e $\xi(n)=e(\alpha n)$ (Teorema de Davenport) satisfazem a ortogonalidade com a função de Möbius. Do mesmo modo, existe um resultado devido a Wirsing que diz que se $\xi(n)$ é real, multiplicativa e $-1\leqslant\xi(n)\leqslant 1$, então ξ e ortogonal a μ se, e somente se, a série de coeficientes não negativos sobre primos

$$\sum_{p}\frac{1+\xi(p)}{p}$$

diverge, veja [21]. Por outro lado, se substituímos $\xi(n)$ por $\xi(n+a)$ ou mais geralmente por $\xi(n+a_1)\xi(n+a_2)\ldots\xi(n+a_t)$ com $0<a_1<a_2<\ldots<a_t$, então a ortogonalidade com μ ainda está em aberto. Por exemplo, as correlações locais de μ consigo mesma, isto é, $\sum_{n\leqslant x}\mu(n+a_1)\mu(n+a_2)\cdots\mu(n+a_t)$ com $a_1,\ldots a_t$ como acima. Este fato inspirou a seguinte conjectura devido a Chowla.

CONJECTURA 4.1 (Chowla). *Sejam $0\leqslant a_1<a_2<\ldots<a_t$ e $k_1,k_2,\ldots k_t$ em $\{1,2\}$ não todos pares, então quando $N\to\infty$ temos*

$$\sum_{n\leqslant N}\mu^{k_1}(n+a_1)\mu^{k_2}(n+a_2)\ldots\mu^{k_t}(n+a_t)=o(N).$$

A conjectura de Chowla parece uma afirmação de que, para n escolhido aleatoriamente entre 1 e x. As "variáveis aleatórias" $\mu(n+1),\ldots,\mu(n+k)$ são assintoticamente independentes entre si, quando $x\to\infty$. Mas isto não é bem formulado, já que por exemplo, os "eventos" $\mu(n)=0$ e $\mu(n+4)=0$ tem uma forte correlação entre si, basicamente devido a ambos serem fortemente relacionados ao "evento" de n ser divisível por 4. Uma interpretação mais precisa da conjectura é que as "variáveis aleatórias" $\mu(n+1),\ldots,\mu(n+k)$ são, assintoticamente, condicionalmente independentes par a par.

Veremos mais a frente que existe uma forte relação entre esta conjectura e o Principio de Aleatoriedade, mas especificamente o caso da ortogonalidade.

Passamos a apresentar agora o ponto de vista dinâmico introduzido por Peter Sarnak, que ajudará na procura do formalismo para o Principio de Aleatoriedade. Por último veremos também a conhecida conjectura de Sarnak.

4.1 Fluxos - Entropia Topológica

DEFINIÇÃO 4.1 (Fluxo). *Um fluxo F, é um par $F=(X,T)$, donde X é um espaço topológico compacto, e $T:X\to X$ uma aplicação contínua. Se existe um $x\in X$*

e uma função $f \in C(X)$ *tal que* $\xi(n) = f(T^n x)_{n \geqslant 1}$*, dizemos que* $\xi(n)$ *e realizada pelo fluxo* (X, T)*.*

Note que uma sequência dada pode ser realizada por muitos fluxos. Vejamos isto no seguinte exemplo.

EXEMPLO 4.1. *Seja* $\xi(n)$ *uma função aritmética que assume valores num conjunto finito* $\Lambda \subset \mathbb{R}$*. Denotemos por* $\Omega = \Lambda^{\mathbb{N}}$ *o espaço de configurações de* ξ *com a topologia produto. Como* Λ *é finito (logo compacto) segue do Teorema de Tychonoff que* Ω *é compacto. Seja* $T : \Omega \to \Omega$ *a aplicação deslocamento definido por*

$$T(x_1, x_2, \ldots) = (x_2, x_3, \ldots).$$

Sejam $\xi = (\xi(1), \xi(2), \ldots) \in \Omega$*,* $f = \pi_1$ *(projeção na primeira coordenada) e para* $k \in \Lambda$ *arbitrário seja* $x = (k, \xi) \in \Omega$*. Então*

$$T(x) = \xi, \quad T^n(x) = (\xi(n), \xi(n+1), \ldots) \Rightarrow f(T^n(x)) = \xi(n).$$

Assim ξ *é realizada pelo fluxo* (Ω, T)*. Também existe outra realização de* ξ *mediante o fluxo* $F_\xi = (X_\xi, T)$*, onde* X_ξ *é o fecho em* Ω *da órbita de* ξ *determinada por* T*, isto é,* $X_\xi = \overline{\{T^j \xi\}}_{j=1,2,\ldots}$*, e aplicação contínua é a restrição de* T *ao fecho da órbita.*

Note que se $\xi(n)$ é realizada pelo fluxo F e $\eta(n)$ por G logo $\xi(n+a)\eta(n+b)$ é realizada pelo fluxo produto $F \times G = (X_F \times X_G, T_F \times T_G)$.

Para medir a complexidade de um fluxo F, consideramos a **Entropia Topológica** $h(F)$, introduzida por Adler (1965). Para nosso propósito será suficiente considerar este conceito em espaços métricos devido a Bowen.

Sejam (X, d) um espaço métrico compacto e $T : X \to X$ uma função contínua. Definimos para cada n natural uma nova métrica sobre X da seguinte maneira:

$$d_n(x, y) = \max\left\{d(T^i(x), T^i(y)) : 0 \leqslant i < n\right\}.$$

Usando esta métrica faremos algumas definições que vão ser importantes para a definição de Entropia Topológica. Observe que se (X, d) é um espaço métrico então a topologia induzida por d_n em X é mais grossa que a topologia induzida por d. Logo se (X, d) é um espaço métrico compacto então (X, d_n) também é um espaço métrico compacto.

DEFINIÇÃO 4.2. *Sejam* n *um número natural,* $\epsilon > 0$ *e* X *um espaço métrico compacto. Um subconjunto* Y *de* X *de diz* (n, ϵ)*-separado com respeito à função* T *se para qualquer* $x, y \in Y$*,* $x \neq y$ *tem-se* $d_n(x, y) > \epsilon$*. Denotemos por*

$$N(n, \epsilon) = \max\{\#Y : Y \textit{ é } (n, \epsilon)\textit{-separado com respeito a } T\}.$$

DEFINIÇÃO 4.3. *Seja (X, d) um espaço topológico compacto. $F = (X, T)$ um fluxo, então a Entropia Topológica é definida por*

$$h(F) = \lim_{\epsilon \to 0} \limsup_{n \to \infty} \frac{\log N(n, \epsilon)}{n}.$$

Observe que X é compacto, logo $N(n, \epsilon)$ é finito, assim o limite que aparece acima expressa o crescimento exponencial médio de subconjuntos "distinguíveis" por T.

Existem outras definições de Entropia Topológica, devido a Adler e Dinanburg. No que segue apresentamos estas definições e mostramos que todas elas são equivalentes.

A definição dada por Adler é a seguinte: dados (X, d) um espaço métrico compacto, $\epsilon > 0$ e $n \in \mathbb{N}$ considere uma cobertura de X:

$$\mathcal{Y} = \{A \subset X : \operatorname{diam}_{d_n}(A) = \epsilon \ \text{ e } \ A \text{ é aberto}\}.$$

Como (X, d_n) é compacto podemos encontrar uma sub-coleção finita $A_1, \dots, A_r$ que cobre X. Denote por $S(n, \epsilon)$ o menor $r \in \mathbb{N}$ tal que $A_1, \dots, A_r$ é uma cobertura de X. Defina $h(F, \epsilon)$ por

$$h(F, \epsilon) := \limsup_{n \to \infty} \frac{1}{n} \log S(n, \epsilon).$$

Observando que a função $\epsilon \mapsto h(F, \epsilon)$ é monótona não-crescente podemos garantir que existe o limite

$$h_{cob}(F) := \lim_{\epsilon \to 0} h(F, \epsilon).$$

A definição dada por Dinanburg é como segue. Um subconjunto $Z \subset X$ é chamado de conjunto (n, ϵ)-gerador se para todo $x \in X$ existe $z \in Z$ tal que $d_n(x, z) \leqslant \epsilon$. Em outras palavras, $X = \cup_{z \in Z} B_{d_n}(z, \epsilon)$. Seja

$$M(n, \epsilon) = \min\{\# Z : Z \text{ é } (n, \epsilon)\text{-gerador}\}$$

então por argumentos semelhantes dados acima existe

$$h_{ger}(F) = \lim_{\epsilon \to 0} \limsup_{n \to \infty} \frac{1}{n} \log M(n, \epsilon).$$

OBSERVAÇÃO 4.1. *As definições acima podem ser dadas de forma semelhante para fluxos contínuos $\phi = \phi_t$.*

Note que as três definições apresentadas acima estão relacionadas a mesma ideia. A Entropia Topológica indica se o crescimento do número de órbitas que se distinguem ao longo de um tempo finito (arbitrário) cresce exponencialmente. (caso $h(F) > 0$).

Vamos mostrar agora que as três definições de entropia coincidem. Se um conjunto Y um conjunto (n,ϵ)-separado cuja cardinalidade é exatamente $N(n,\epsilon)$ dizemos que Y é um subconjunto (n,ϵ)-separador maximal. Note que tais conjuntos são (n,ϵ)-geradores. De fato, suponha por absurdo que existe um $x \in X$ tal que, para todo $y \in Y$ temos $d_n(x,y) > \epsilon$. Então $Y \cup \{x\}$ é um conjunto (n,ϵ) separador e isto contradiz a maximalidade de Y. Portanto $M(n,\epsilon) \leqslant N(n,\epsilon)$, logo $h_{ger}(F) \leqslant h(F)$. Por outro lado, para todo $\epsilon > 0$ temos $N(n,\epsilon) \leqslant M(n,\epsilon/2)$. Portanto

$$h(F) = \lim_{\epsilon\to 0} \limsup_{n\to\infty} \frac{1}{n} \log N(n,\epsilon) \leqslant \lim_{\epsilon\to 0} \limsup_{n\to\infty} \frac{1}{n} \log M(n,\epsilon/2) = h_{ger}(F).$$

Observamos que $S(n,2\epsilon) \leqslant M(n,\epsilon)$, pois o diâmetro de uma bola de raio ϵ é menor ou igual que 2ε e assim qualquer cobertura por bolas de raio ϵ é uma cobertura por conjuntos de diâmetro menor ou igual a 2ϵ. Note também que $S(n,2\epsilon) \leqslant M(n,\epsilon) \leqslant S(n,\epsilon/2)$, desta forma

$$h_{cob}(F) = \lim_{\epsilon\to 0} \limsup_{n\to\infty} \frac{1}{n} \log S(n,2\epsilon) \leqslant \lim_{\epsilon\to 0} \limsup_{n\to\infty} \frac{1}{n} \log M(n,\epsilon) = h_{ger}(F).$$

$$h_{ger}(F) = \lim_{\epsilon\to 0} \limsup_{n\to\infty} \frac{1}{n} \log M(n,\epsilon) \leqslant \lim_{\epsilon\to 0} \limsup_{n\to\infty} \frac{1}{n} \log S(n,\epsilon/2) = h_{cob}(F).$$

Isto prova a afirmação. Vamos na sequência dar um exemplo simples do cálculo da Entropia. Advertimos porém que mas na maioria dos casos o cálculo da Entropia não é fácil de ser realizado.

Exemplo 4.2. *Consideremos o fluxo definido por T, sendo T uma isometria e um espaço métrico (X,d) compacto. Então temos que $d_n = d$ e assim $N(n,\epsilon) = N(1,\epsilon)$, isto é, $N(n,\epsilon)$ não depende de n. Portanto*

$$h(F) = \lim_{\epsilon\to 0} \limsup_{n\to\infty} \frac{\log N(1,\epsilon)}{n} = 0.$$

A próxima proposição estabelece a razão do nome de Entropia Topológica. Isto é, porque a entropia depende unicamente da topologia, apesar de ter sido definido usando a estrutura métrica do espaço.

Proposição 4.1. *A Entropia Topológica não depende da métrica escolhida. Se duas métricas induzem a mesma topologia então a entropia do fluxo é invariante.*

Demonstração. Considere duas métricas d e d' que induzem a mesma topologia. Para $\epsilon > 0$, defina

$$\delta(\epsilon) = \sup_{x,y\in X} \{d'(x,y) : d(x,y) \leqslant \epsilon\}.$$

Já que d e d' induzem a mesma topologia em X tem-se que $\delta(\epsilon) \to 0$ quando $\epsilon \to 0$. Se U é um conjunto tal que $d_n(x,y) < \varepsilon$ para todo x, y em U, então $\sup_{x,y\in U} (d'_n(x,y))$

é menor ou igual que $\delta(\epsilon)$. Logo, se Y é um conjunto (n,ϵ)-gerador minimal, na métrica d_n, então Y é um conjunto $(n,\delta(\epsilon))$-gerador com a métrica d'_n. Portanto temos que

$$\lim_{\delta(\epsilon)\to 0}\limsup_{n\to\infty}\frac{1}{n}\log M(n,\delta(\epsilon)) \leqslant \lim_{\epsilon\to 0}\limsup_{n\to\infty}\frac{1}{n}\log M(n,\epsilon).$$

O mesmo argumento, trocando as métricas d e d' nos fornece a desigualdade oposta. Isto prova o resultado. □

Em seguida, vamos ver um resultado que relaciona a Entropia Topológica de um fluxo composto com uma função contínua sobrejetora e a Entropia do fluxo original, devido a Bowen.

TEOREMA 4.1. *Seja* (a_n) *uma sequencia limitada, realizada pelo fluxo* (X,T)*, onde* $X=\overline{\{T^k(a_n),k\geqslant 0\}}$ *e* $f(T^k(a_n))=a_{k+1}$*,* $k\geqslant 0$ *(como no exemplo* 4.1*). Se* (a_n) *é realizada por outro fluxo* (Y,σ)*, então* $h(Y,\sigma)\geqslant h(X,T)$*.*

Demonstração. Dado que (a_n) é realizada por (Y,σ), existem uma função contínua $g:Y\to\mathbb{C}$ e um ponto $y\in Y$ tal que $a_n=g(\sigma^n(y))$. Denote $\widetilde{Y}=\overline{\{\sigma^n(y),n\geqslant 0\}}$ o fecho da órbita por σ para o ponto $y\in Y$. Logo (a_n) é realizada também pelo fluxo $(\widetilde{Y},\sigma)$. Considere o seguinte diagrama comutativo

$$\begin{tikzcd} \widetilde{Y} \arrow[r, "\sigma"] \arrow[d, "h"'] & \widetilde{Y} \arrow[d, "h"] \\ X \arrow[r, "T"'] & X \end{tikzcd}$$

onde h é uma função contínua sobrejetora definida como segue

$$h(z)=(g(z),g(\sigma(z)),g(\sigma^2(z)),\ldots).$$

Temos que $h(y)=(a_n)_{n\geqslant 0}\in X$ e $h(z)\approx(a_{n+k})_{n\geqslant 0}\in X$, dado que X é compacto segue que $h(z)\in X$.

Portanto $h(Y,\sigma)\geqslant h(\widetilde{Y},\sigma)\geqslant h(X,T)$. □

Furstenberg teve a ideia de estudar $\xi(n)$ como o tempo de retorno de um sistema dinâmico. Assim podemos entender o conceito, que até agora era informal, de sequência "razoável" usando a ideia de Entropia Topológica. Na próxima seção apresentaremos esta ideia, assim como uma nova forma de enunciar o Principio de Aleatoriedade de Möbius, mais precisamente o caso da ortogonalidade.

4.2 Conjectura de Sarnak

Definição 4.4. *Um fluxo F é determinístico se $h(F) = 0$. A sequência $\xi(n)$ é determinística se ela é realizada por algum fluxo determinístico.*

Agora vamos ver um resultado que mostra que a função de Möbius é em algum sentido aleatória.

Teorema 4.2. *A função de Möbius $\mu(n)$ é não determinística.*

Demonstração. Sejam $X = \{-1, 0, 1\}^{\mathbb{N}}$, T a aplicação deslocamento em X, $w = (\mu(1), \mu(2), \ldots) \in X$. Defina

$$X_M = \overline{\{T^j(w)\}}_{j \geqslant 0} \subseteq X.$$

Então $M = (X_M, T_M)$ é um fluxo de Möbius, onde T_M é a aplicação T restrita a X_M.

Agora considere $\eta = (\mu^2(1), \mu^2(2), \ldots) \in Y = \{0, 1\}^{\mathbb{N}}$. Se definimos $Y_S = \overline{\{T^j(\eta)\}}_{j \geqslant 0}$, obtemos o fluxo $S = (Y_S, T_S)$, conhecido como o fluxo livre de quadrados.

Podemos munir aos conjuntos $\{0, 1, -1\}$ e $\{0, 1\}$ da topologia discreta na qual são compactos e usar o Teorema de Tychonoff para garantir que os conjuntos X e Y são compactos.

O fluxo S é denominado livre de quadrados, pois as componentes de η são zeros e uns, dependendo se o valor na posição n é ou não livre de quadrados e, o ponto importante e que S é um fator do fluxo M, no seguinte sentido.

Considere a função $\pi : X_M \to Y_S$, dada por $\pi(x_1, x_2, \ldots) = (x_1^2, x_2^2, \ldots)$. Note que $\pi(w) = \eta$, π é contínua e sobrejetora, e além do mais o diagrama abaixo é comutativo

$$\begin{array}{ccc} X_M & \xrightarrow{T_M} & X_M \\ \pi \downarrow & & \downarrow \pi \\ Y_s & \xrightarrow[T_s]{} & Y_s \end{array}$$

Para calcular a entropia do fluxo S. Usamos em Y_S a métrica dada pela restrição da métrica $d : \{0, 1\}^{\mathbb{N}} \times \{0, 1\}^{\mathbb{N}} \to \mathbb{R}$ definida por $d(x, x) = 0$ para todo $x \in \{0, 1\}^{\mathbb{N}}$ e para $x \neq y \in \{0, 1\}^{\mathbb{N}}$ por

$$d(x, y) = \frac{1}{2^n}, \quad \text{onde } n = \min\{i \in \mathbb{N} : x_i \neq y_i\}.$$

Note que a métrica definida desta maneira induz a topologia produto em Y_S. Podemos provar que a Entropia Topológica de S é dada por

$$h(S) = \log 2 \frac{6}{\pi^2},$$

usando a ideia de conjuntos admissíveis (veja [13]).

Agora como S é um fator de M temos que $h(M) \geqslant h(S) = \log 2\frac{6}{\pi^2}$, logo $h(M) > 0$ e $\mu(n)$ é não determinística. Além disso, observe que qualquer outro fluxo que realize $\mu(n)$ (pelo teorema 4.1) também é não determinístico. $\square$

Com objetivo de apresentar a Conjectura de Sarnak, vamos dar a seguinte definição.

DEFINIÇÃO 4.5. *A função de Möbius μ é ortogonal a um fluxo F se é ortogonal a todas as sequências $\xi(n)$ realizadas por F, isto é,*

$$\sum_{n \leqslant x} \mu(n)\xi(n) = o(x).$$

Finalmente temos a conjectura de Peter Sarnak que tambem é conhecida como Lei da Aleatoriedade de Möbius.

CONJECTURA 4.2 (Lei da Aleatoriedade de Möbius). *A função de Möbius μ é ortogonal a qualquer fluxo determinístico limitado, isto é, μ é ortogonal a qualquer sequencia ξ determinística limitada.*

A Lei da Aleatoriedade de Möbius é satisfeita para um grande número de fluxos, por exemplo, podemos mencionar alguns deles:

1. Se F é o fluxo trivial (isto é, só um ponto) a lei de aleatoriedade é equivalente ao Teorema dos Números Primos (veja capitulo 2).

2. Se F é finito (X_F finito) a lei de aleatoriedade é essencialmente o Teorema dos Números Primos para Progressões Aritméticas.

3. Se F é a rotação do círculo ($X_F = \mathbb{R}/\mathbb{Z}$ e $T_\alpha x = x + \alpha$) a lei de aleatoriedade pode ser provado usando o Teorema de Weierstrass [1] e o Teorema de H. Davenport (veja capitulo 4).

A distribuição das órbitas em um fluxo determinístico geral pode ser bastante irregular e o leitor pode ser um pouco cético quanto a validade da Conjectura de Sarnak. No entanto, o próximo resultado mostra que ela é pelo menos tão realista quanto a Conjectura de Chowla.

TEOREMA 4.3. *A conjectura de Chowla implica a conjectura de Sarnak*

Vejamos primeiro um resultado que nós vai ajudar na prova do teorema acima.

---

[1] O teorema de Weierstrass afirma que se f é uma função contínua em X então existem $m \in \mathbb{N}$ e constantes $c_i \in X$ tal que

$$f(x) \approx \sum_{k=-m}^{m} c_k e(kx)$$

Proposição 4.2. *Assuma a conjectura de Chowla. Logo para algum $m \geqslant 1$, algum $\epsilon > 0$, e coeficientes $c_1, \ldots, c_m$ limitados em módulo por 1, temos*

$$\mathbb{P}\left(\left|\frac{1}{m}\sum_{i=1}^{m} c_i\mu(n+i)\right| \geqslant \epsilon\right) \leqslant C\exp(-\epsilon^2 m/C) + o_{x\to\infty;m,\epsilon}(1), \tag{4.3}$$

onde C é uma constante e $o_{x\to\infty;m,\epsilon}(1)$ vai para zero quando $x \to \infty$ para m, ϵ fixos (uniformemente na escolha de coeficientes $c_1, \ldots, c_m$), e n é escolhido uniformemente e de maneira aleatória de 1 a x.

Demonstração. Seja k um inteiro par, suficientemente grande a ser otimizado adiante. Pela desigualdade de Chebyshev

$$\mathbb{P}\left(\left|\frac{1}{m}\sum_{i=1}^{m} c_i\mu(n+i)\right| \geqslant \epsilon\right) \leqslant \frac{1}{(\epsilon m)^k}\mathbb{E}\left|\sum_{i=1}^{m} c_i\mu(n+i)\right|^k.$$

Expandindo a k-ésima potência e usando a desigualdade triangular, podemos limitar o lado direito da desigualdade acima por

$$\frac{1}{(\epsilon m)^k}\sum_{1\leqslant i_1,\ldots,i_k\leqslant m} |\mathbb{E}\mu(n+i_1)\ldots\mu(n+i_k)|. \tag{4.4}$$

Os termos da soma são sempre limitados por 1. Pela conjectura de Chowla, podem de fato ser limitados por $o_{n\to\infty,m}(1)$ a menos que nenhum dos índices $i_1, \ldots, i_k$ ocorra um número ímpar de vezes, de modo que estes k índices são, de fato, agrupados em no máximo $k/2$ classes. Então nós estamos enfrentando agora um problema de combinatória enumerativa, que consiste em contar quantas k-úplas $(i_1, \ldots, i_k)$ são desta forma. Pode-se estimar essa contagem como segue: olhando para $i_1, \ldots, i_k$ de esquerda para a direita, há dois casos: cada j é ou "novo"(se i_j não é igual a nenhum dos $i_1, \ldots, i_{j-1}$), ou "repetido". No primeiro caso, existem no máximo m escolhas para i_j. No segundo caso, há no máximo k. Além disso, no máximo $k/2$ dos casos são "novos", portanto uma vez que forem decididos quais índices j são novos ou repetidos, há no máximo $m^{k/2}k^{k/2}$ escolhas restantes a serem feitas (assumindo que $m \geqslant k$). Assim a contagem total é limitada por $2^k m^{k/2}k^{k/2}$. Usando esta cota, podemos majorar 4.4 por

$$\frac{1}{(\epsilon m)^k}\left(2^k m^{k/2}k^{k/2} + o_{n\to\infty,\epsilon,m,k}(1)\right) = \left(\frac{4k}{\epsilon^2 m}\right)^{k/2} + o_{n\to\infty,\epsilon,m,k}(1).$$

Se escolhemos k como sendo o maior inteiro par menor que $\epsilon^2 m/10$, obtemos

$$\begin{aligned}\mathbb{P}\left(\left|\frac{1}{m}\sum_{i=1}^{m} c_i\mu(n+i)\right| \geqslant \epsilon\right) &\leqslant \left(\frac{4k}{\epsilon^2 m}\right)^{k/2} + o_{n\to\infty,\epsilon,m,k}(1)\\ &\leqslant \left(\frac{4\epsilon^2 m}{10\epsilon^2 m}\right)^{\epsilon^2 m/20} + o_{n\to\infty,\epsilon,m,k}(1)\\ &= \exp\left(\frac{\epsilon^2 m}{20}\log\left(\frac{2}{5}\right)\right) + o_{x\to\infty,\epsilon,m}(1).\end{aligned}$$

tomando $1 < C < 21$ obtemos o resultado desejado. □

Apresentamos agora a prova do Teorema 4.3.

Demonstração. Seja f uma sequência determinística, sem perda de generalidade podemos supor que f toma valores reais e é limitada por 1. Sejam $\epsilon > 0$, $m \geqslant 1$ e f_ϵ o arredondamento de f para o múltiplo mais próximo de ϵ menor que um. Assim, a partir da hipótese de entropia zero, há no máximo $\exp(o_{m\to\infty;\epsilon,f}(m))$ diferentes valores para a m-úpla $(f_\epsilon(n+1),\ldots,f_\epsilon(n+m))$, com n variando sobre os números naturais. Seja S_m o conjunto de todas essas m-úplas. Da proposição anterior e da cota para a união, temos que

$$\begin{aligned}&\mathbb{P}\left(\sup_{(c_1,\ldots,c_m)\in S_m}\left|\frac{1}{m}\sum_{i=1}^{m} c_i\mu(n+i)\right| \geqslant \epsilon\right)\\ &\qquad\qquad \leqslant \big(C\exp(-\epsilon^2 m/C) + o_{x\to\infty;m,\epsilon}(1)\big)\exp(o_{m\to\infty;\epsilon,f}(m)).\end{aligned}$$

Portanto

$$\begin{aligned}&\mathbb{P}\left(\left|\frac{1}{m}\sum_{i=1}^{m} f_\epsilon(n+i)\mu(n+i)\right| \geqslant \epsilon\right)\\ &\qquad\qquad \leqslant \big(C\exp(-\epsilon^2 m/C) + o_{x\to\infty;m,\epsilon}(1)\big)\exp(o_{m\to\infty;\epsilon,f}(m)).\end{aligned}$$

O lado direito da última desigualdade pode ser simplificado como mostramos abaixo

$$\mathbb{P}\left(\left|\frac{1}{m}\sum_{i=1}^{m} f_\epsilon(n+i)\mu(n+i)\right| \geqslant \epsilon\right) \leqslant \epsilon + o_{x\to\infty;m,\epsilon}(1),$$

se m é suficientemente grande dependendo de ϵ, C. Agora, da caracterização da esperança $EX = \int_0^\infty (1-F(x))dx - \int_{-\infty}^0 F(x)dx$, e da limitação do integrando por um temos

$$\begin{aligned}\left|\mathbb{E}\frac{1}{m}\sum_{i=1}^{m} f_\epsilon(n+i)\mu(n+i)\right| &= \int_0^1 \mathbb{P}\left(\left|\frac{1}{m}\sum_{i=1}^{m} f_\epsilon(n+i)\mu(n+i)\right| > \epsilon\right) d\epsilon \\ &\quad - \int_{-1}^{0}\left(1-\mathbb{P}\left(\left|\frac{1}{m}\sum_{i=1}^{m} f_\epsilon(n+i)\mu(n+i)\right| > \epsilon\right)\right) d\epsilon \\ &\leqslant \epsilon + o_{x\to\infty;m,\epsilon}(1) - 1 + \epsilon + o_{x\to\infty;m,\epsilon}(1).\end{aligned}$$

Portanto

$$\left|\mathbb{E}\frac{1}{m}\sum_{i=1}^{m} f_\epsilon(n+i)\mu(n+i)\right| \leqslant o_{x\to\infty;m,\epsilon}(1) + 2\epsilon.$$

Mas, pela invariância por translação da função f_ϵ temos

$$\mathbb{E}f_\epsilon(n+i)\mu(n+i) = \frac{1}{x}\sum_{n<x} f_\epsilon(n)\mu(n) + o_{x\to\infty;m}(1),$$

para cada $1 \leqslant i \leqslant m$. Isto segue dado que pela proposição anterior n é escolhido de forma uniforme em $[1, x]$, então para $1 \leqslant i \leqslant m$ fixo temos

$$\left|\mathbb{E}f_\epsilon(n+i)\mu(n+i) - \frac{1}{x}\sum_{n<x} f_\epsilon(n)\mu(n)\right| \leqslant \frac{1}{x}|f_\epsilon(1)\mu(1) + \ldots + f_\epsilon(i)\mu(i)| \underset{x\to\infty}{\longrightarrow} 0.$$

Desta forma

$$\mathbb{E}\frac{1}{m}\sum_{i=1}^{m} f_\epsilon(n+i)\mu(n+i) = \frac{1}{x}\sum_{n<x} f_\epsilon(n)\mu(n) + o_{x\to\infty;m}(1).$$

Isto nos fornece (para uma escolha apropriada de m)

$$\left|\sum_{n<x} f_\epsilon(n)\mu(n)\right| \leqslant (o_{x\to\infty;\epsilon,f}(1) + 2\epsilon)x.$$

e assim

$$\left|\sum_{n<x} f(n)\mu(n)\right| \leqslant (o_{x\to\infty;\epsilon,f}(1) + 3\epsilon)x.$$

deixando ϵ ir a zero suficientemente devagar em função de x, obtemos a Conjectura de Sarnak. □

Assim temos descrito as ideias de Sarnak para definir de alguma maneira a aleatoriedade da função de Möbius. Além do mais vimos também suas razões para acreditar neste fato. Isto torna interessante e relevante o ponto de vista dinâmico, pois o que está por traz da aleatoriedade de Möbius é a Hipótese de Riemann.

Capítulo 5

Apêndice

Definição 5.1 (Métodos de Crivos). *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 'explicita' 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 $\mathcal{A}$ como sendo um conjunto finito de inteiros positivos e $\mathcal{A}_p$ como sendo o subconjunto de $\mathcal{A}$ consistindo de elementos de que estão em certas classes de congruência módulo p. Por exemplo, se $\mathcal{A}$ é o conjunto dos números naturais menores ou iguais que x e $\mathcal{A}_p$ é o conjuntos 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 $n \leqslant x$ que são coprimos com todos os elementos de $\mathcal{P}$.

Definição 5.2. *[Símbolos de Landau]*

Em matemática, la notação de Landau, também conhecida como notação de ordem é uma notação para á comparação assintótica de funções. Se f e g são funções complexas num entorno do x_0, então

1. *$f = O(g)$ quando $x \to x_0$ se e somente se existe $\epsilon > 0$ tal que $|f(x)| < \epsilon |g(x)|$ para tudo x num entorno do x_0.*

2. *$f = o(g)$ quando $x \to x_0$ se e somente se para tudo $\epsilon > 0$ temos que $|f(x)| < \epsilon |g(x)|$ para tudo x num entorno do x_0.*

   Além do mais, se $f(x)$ e $g(x)$ duas funções definidas para $x > x_0$ e $g(x) \neq 0$, $\forall x > x_0$. Os símbolos

   $$f = o(g), \quad f = O(g)$$

   significam respectivamente que $\frac{f(x)}{g(x)} \to 0$ quando $x \to \infty$ e que $\frac{f(x)}{g(x)}$ está limitado para x suficientemente grande. A mesma notação é valida para x tendendo para $-\infty$.

3. *$f \sim g$, com f e g definidas numa vizinhança de um ponto x_0 (finito ou infinito) significa que $\frac{f(x)}{g(x)} \to 1$ quando $x \to x_0$ e são chamadas assintoticamente iguais.*

LEMA 5.1 (Lema de Landau). *Seja $a_n \geqslant 0$ para todo $n \geqslant 1$. Suponha que a serie*

$$f(s) = \sum_{n \geqslant 1} a_n n^{-s}$$

converge absolutamente para todo s com $Re(s) > \sigma_0$, mas não para algum s com $Re(s) < \sigma_0$; em outras palavras, σ_0 é a abscisa de convergência absoluta de $f(s)$. Então $s = \sigma_0$ é uma singularidade de $f(s)$, i.e, $f(s)$ não possui continuação analítica a uma vizinhança de σ_0.

Equivalentemente, se σ é finito, $\sum_{n \geqslant 1} \frac{a_n}{n^s}$ converge em $Re(s) > \sigma$ e a função f admite continuação analítica em torno de $s = \sigma$, então a serie $\sum_{n \geqslant 1} \frac{a_n}{n^s}$ converge em $Re(s) > \sigma - \epsilon$ para algum $\epsilon > 0$.

Demonstração. Mostraremos a segunda forma do lema. Seja $a = 1 + \sigma$, daí f é analítica em a e portanto

$$f(s) = \sum_{k \geqslant 0} \frac{f^{(k)}(a)}{k!}(s - a)^k,$$

sendo esta série absolutamente convergente num disco aberto de centro em a e raio maior que 1, pois f é analítica em σ.

Temos que

$$f^{(k)}(a) = (-1)^k \sum_{n \geqslant 1} \frac{a_n (\log n)^k}{n^a},$$

logo

$$f(s) = \sum_{k \geqslant 0} \sum_{n \geqslant 1} \frac{a_n (\log n)^k}{n^a k!}(a - s)^k$$

neste disco. Como o raio de convergência é maior que 1, para algum $\epsilon > 0$ temos que a expressão acima válida com $s = \sigma - \epsilon$, logo $a - s = 1 + \epsilon > 0$, e desta forma os termos na série dupla são não negativos, permitindo a troca dos somatórios. Assim temos que

$$\begin{aligned} f(\sigma - \epsilon) &= \sum_{n \geqslant 1} \frac{a_n}{n^a} \sum_{k \geqslant 0} \frac{((1+\epsilon)\log n)^k}{k!} \\ &= \sum_{n \geqslant 1} \frac{a_n}{n^a} e^{(1+\epsilon)\log n} = \sum_{n \geqslant 1} \frac{a_n}{n^{\sigma - \epsilon}}, \end{aligned}$$

isto é, a série $\sum_{n \geqslant 1} \frac{a_n}{n^s}$ converge para $s = \sigma - \epsilon$, logo também converge em $Re(s) > \sigma - \epsilon$. □

Definição 5.3 (Transformada de Mellin). *Seja* $\phi : (0, +\infty) \to \mathbb{C}$. *A transformada de Mellin de* ϕ *é definida por*

$$\widehat{\phi}(s) = \int_0^\infty \phi(x) x^{s-1} dx$$

e sua transformada inversa é dada por

$$\phi(x) = \frac{1}{2\pi i} \int_{c - i\infty}^{c + i\infty} x^{-s} \widehat{\phi}(s) ds.$$

A transformada $\widehat{\phi}(s)$ *existe se a integral*

$$\int_0^\infty |\phi(x)| x^{k-1} dx$$

é limitada para algum $k > 0$*, em tal caso a transformada inversa existe se* $c > k$.

Proposição 5.1 (Fórmula de Stirling). *A fórmula assintótica de Stirling*

$$\Gamma(s) = \left(\frac{2\pi}{s}\right)^{\frac{1}{2}} \left(\frac{s}{e}\right)^s \left[1 + O\left(\frac{1}{|s|}\right)\right] \tag{5.1}$$

é válida para os ângulos cujo $|arg(s)| \leqslant \pi - \varepsilon$ *com a constante dependendo de* ε.

Demonstração. Veja [10], pag. 321. □

Teorema 5.1 (Aproximação Diofantina de Dirichlet). *Seja* α *um número real e* n *um inteiro. Então existem inteiros* k *e* b *com* $0 < k < n$*, tais que*

$$\frac{-1}{n} < k\alpha - b < \frac{1}{n}.$$

Demonstração. Note que cada um dos $n+1$ números da forma $a_i = i\alpha - [i\alpha]$, $i = 0, 1, \ldots, n$ pertencem ao intervalo $0 \leqslant a_i < 1$. Assim pelo princípio da casa de pombos existem no mínimo um intervalo da forma $[\frac{r}{n}, \frac{r+1}{n})$, $0 \leqslant r < n$, que contém dois destes elementos, que são denotados por a_m e a_j. Para tais m e j temos que $0 \leqslant a_m - a_j < \frac{1}{n}$ ou $0 \leqslant a_j - a_m < \frac{1}{n}$, daí

$$\frac{-1}{n} < a_m - a_j < \frac{1}{n};$$

$$\frac{-1}{n} < m\alpha - [m\alpha] - (j\alpha - [j\alpha]) < \frac{1}{n};$$

$$\frac{-1}{n} < (m-j)\alpha - \big([m\alpha] - [j\alpha]\big) < \frac{1}{n};$$

tomando $B = [m\alpha] - [j\alpha]$ temos

$$\Big|(m-j)\alpha - B\Big| < \frac{1}{n}.$$

Para completar a demonstração basta observar que, se $m > j$ é suficiente tomar $k = m - j$ e $b = B$. Por outro lado, se $m < j$ tomamos $k = j - m$ e $b = -B$. Isto conclui a prova. □

5.1 Caráteres

DEFINIÇÃO 5.4 (Caráter de Dirichlet). *Um caráter de Dirichlet é uma função χ definida sobre os inteiros tomando valores complexos com as seguintes propriedades:*

- *Existe um número $k \in \mathbb{N}$ tal que $\chi(n) = \chi(n+k)$ para tudo n.*
- *Se $(n,k) > 1$ então $\chi(n) = 0$; se $(n,k) = 1$ então $\chi(n) \neq 0$.*
- *$\chi(mn) = \chi(n)\chi(m)$ para tudo $m, n \in \mathbb{N}$.*

OBSERVAÇÃO 5.1. *Se $(a,k) = 1$, o Teorema de Euler diz que $a^{\varphi(k)} \equiv 1 (mod\ k)$, daí*

$$\chi(a)^{\varphi(k)} = \chi(a^{\varphi(k)}) = \chi(1) = 1.$$

Portanto para todo a primo relativo com k, $\chi(a)$ é uma $\varphi(k)$-ésima raiz da unidade.

No caso geral, seja G um grupo abeliano finito. Um homomorfismo $\chi : G \to \mathbb{C}^*$ é chamado caráter de G. Portanto (na notação multiplicativa) χ tem a propriedade

$$\chi(xy) = \chi(x)\chi(y) \quad \text{para todo } x, y \in G,$$

$\chi(1) = 1$ e $\chi(x)^m = 1$, onde $m = |G|$ é a ordem de G. Os caráteres de G formam um grupo $\widehat{G}$ com a multiplicação dado por

$$(\chi_1\chi_2)(x) = \chi_1(x)\chi_2(x) \quad \text{para todo } x \in G.$$

$\widehat{G}$ é chamado o grupo dual de G, o elemento identidade é o caráter trivial $\chi_0(x) = 1$ para todo $x \in G$. Note que se G é um grupo cíclico de ordem m e g um gerador de G, cada caráter de G é do tipo

$$\chi_a(x) = e^{2\pi iay/m} \quad \text{se } x = g^y.$$

para alguma classe residual $a(\text{mod } m)$. O inverso do caráter χ é o caráter $\overline{\chi} = \chi^{-1}$ definido por $\overline{\chi}(x) = \overline{\chi(x)}$.

Observação 5.2. *Note que um caráter de Dirichlet mod q é um caráter no grupo $(\mathbb{Z}/q\mathbb{Z})^*$. Para cada q, existe um carácter de Dirichlet tomando valor 1 em todos os inteiros n coprimos com q, este carácter é chamado de carácter principal módulo q.*

Associado a cada caráter χ, além de seu módulo m, está um número natural m^*, seu condutor. O condutor é o menor divisor de m tal que χ pode ser escrito como $\chi = \chi_0\chi^*$, onde χ_0 é o caráter principal módulo m e χ^* é um caráter a modulo m^*. Para alguns caráteres o condutor é igual ao módulo, tais caráteres recebem o nome de primitivos.

Proposição 5.2 (Relações de Ortogonalidade). *Sejam χ_1 e χ_2 dois caráteres de G. Então*

$$\sum_{x\in G} \chi_1(x)\overline{\chi_2}(x) = \begin{cases} |G| & se \quad \chi_1 = \chi_2; \\ 0 & se \quad \chi_1 \neq \chi_2, \end{cases}$$

e sua relação dual. Se $x, y \in G$ então

$$\sum_{\chi\in\widehat{G}} \chi(x)\chi(y^{-1}) := \begin{cases} |G| & se \quad x = y; \\ 0 & se \quad x \neq y. \end{cases}$$

Demonstração. Para a primeira afirmação, se $\chi_1 = \chi_2$ logo $\chi_1(x)\overline{\chi_2}(x) = 1$ para todo $x \in G$ e dado que existem $|G|$ homomorfismos distintos tem-se a primeira igualdade, agora se $\chi_1 \neq \chi_2$ existe um $y \in G$ tal que $\chi_1(y)\overline{\chi_2}(y) \neq 1$, desta forma temos

$$\begin{aligned}(1 - \chi_1(y)\overline{\chi_2}(y)) \sum_{x\in G} \chi_1(x)\overline{\chi_2}(x) &= \sum_{x\in G} \chi_1(x)\overline{\chi_2}(x) - \sum_{x\in G} \chi_1(y)\overline{\chi_2}(y)\chi_1(x)\overline{\chi_2}(x) \\ &= \sum_{x\in G} \chi_1(x)\overline{\chi_2}(x) - \sum_{x\in G} \chi_1(xy)\overline{\chi_2}(xy) \\ &= \sum_{x\in G} \chi_1(x)\overline{\chi_2}(x) - \sum_{x\in G} \chi_1(x)\overline{\chi_2}(x) = 0.\end{aligned}$$

Como $\chi_1(y)\overline{\chi_2}(y) \neq 1$ tem-se

$$\sum_{x\in G} \chi_1(x)\overline{\chi_2}(x) = 0.$$

A segunda afirmação segue do fato que os caráteres são distintos, dai o grupo dual de $\widehat{G}$ é um grupo cíclico de ordem m e portanto $\widehat{G}$ e isomorfo a G. $\square$

Os caráteres de G formam um sistema ortogonal completo, assim uma função $f : G \to \mathbb{C}$ tem sua expansão de Fourier

$$f = \frac{1}{|G|} \sum_{\psi\in\widehat{G}} \langle f, \psi\rangle \psi$$

com coeficientes

$$\langle f, \psi\rangle = \sum_{g\in G} f(g)\overline{\psi}(g).$$

No caso de funções em classes residuais, e sobre corpos finitos, os caráteres aditivos e multiplicativos estão presentes e são necessários para transformar a expansão de Fourier de um sistemas em outro, ao fazê-lo aparecem as somas de Gauss.

Consideremos as somas de Gauss associadas com caráteres em classes residuais, digamos mod m, para um caráter multiplicativo $\chi(\text{mod } m)$ defina

$$\tau(\chi) = \sum_{b(mod\ m)} \chi(b)e(\frac{b}{m}). \tag{5.2}$$

multiplicando por $\overline{\chi}(a)$ e somando sobre χ obtemos, usando as relações de ortogonalidade acima, que

$$e(\frac{a}{m}) = \frac{1}{\varphi(m)} \sum_{\chi(mod\ m)} \overline{\chi}(a)\tau(\chi) \quad \text{se } (a, m) = 1. \tag{5.3}$$

Esta é a expansão de Fourier de caráteres aditivos em função de caráteres multiplicativos. Similarmente

$$\chi(a)\tau(\overline{\chi}) = \sum_{b(mod\ m)} \overline{\chi}(b)e(\frac{ab}{m}) \quad \text{se } (a, m) = 1. \tag{5.4}$$

Um resultado importante a respeito da existência de polos das L-series de Dirichlet é o seguinte teorema

TEOREMA 5.2. *Seja χ um caráter de Dirichlet de módulo q. Então $L(\chi, s)$ estende-se a uma função holomorfa em $Re(s) > 0$ com um polo simples em $s = 1$ se χ é principal. Caso contrario $L(\chi, s)$ é holomorfa também em $s = 1$.*

Demonstração. Se χ é principal, temos que

$$L(\chi, s) = \zeta(s) \prod_{p|q} (1 - p^{-s}).$$

Então o resultado segue do polo simples de ζ.

Agora se χ não é principal, por somação parcial podemos escrever

$$L(\chi, s) = \sum_{n \geqslant 1} (\chi(1) + \ldots + \chi(n))(n^{-s} - (n+1)^{-s}).$$

Das relações de ortogonalidade, sabemos que $\chi(1) + \ldots + \chi(q) = 0$, assim $\chi(1) + \ldots + \chi(n)$ é limitado para todo n. Enquanto,

$$n^{-s} - (n+1)^{-s} = n^{-s}(1 - (1 + 1/n)^{-s}) = sn^{-s-1} + O(n^{-s-2}),$$

onde a constante implicada pode ser tomada uniforme sobre s em conjuntos compactos. Consequentemente, a representação em soma de $L(\chi, s)$ dada acima converge uniformemente para $Re(s) \geqslant \varepsilon$ para todo $\varepsilon > 0$. Isto conclui a prova. $\square$

5.2 Funções Inteiras

Definição 5.5. *Seja f uma função inteira com zeros $\{a_1, a_2, \ldots\}$ tomados com suas multiplicidades e ordenados tal que $|a_1| \leqslant |a_2| \leqslant \ldots$. A função f é chamada de posto finito se existe um inteiro ρ tal que*

$$\sum_{n \geqslant 1} |a_n|^{-(\rho+1)} < \infty.$$

Se ρ é o menor inteiro que satisfaz a desigualdade acima, então dizemos que f tem posto ρ. Uma função com só um número finito de zeros é dito ser de posto zero.

Definição 5.6. *Seja f uma função inteira de rango ρ com zeros $\{a_1, a_2, \ldots\}$. A chamada forma estândar de f é dado por o seguinte produto*

$$P(z) = \prod_{n \geqslant 1} E_p(z/a_n),$$

onde

$$E_p(z) = (1 - z) \exp\left(z + \frac{z^2}{2} + \ldots \frac{z^p}{p}\right).$$

Definição 5.7 (Funções de gênero finito). *Uma função f tem gênero finito se f tem posto finito e f pode ser expresso como segue*

$$f(z) = z^m e^{g(z)} P(z),$$

onde $P(z)$ está na forma estândar, g é um polinômio e m é a multiplicidade do zero em $z = 0$. Se p é o posto de f e q é o grau do polinômio g, logo $\mu = max\{p, q\}$ é chamado o gênero de f.

DEFINIÇÃO 5.8. *[Funções inteiras de ordem finito]*

Uma função inteira f é de ordem finito se existem uma constante positiva a e um $r_0 > 0$ tal que $|f(z)| < \exp(|z|^a)$ para $|z| > r_0$. Se f não é de ordem finito se diz que têm ordem infinito.

Se f tem ordem finito logo o número $\lambda = inf\{a : |f(z)| < exp(|z|^a)\}$ para $|z|$ suficientemente grande é chamado de ordem de f.

5.2.1 Teorema de Fatoração de Hadamard

Nesta seção vamos provar o fato que cada função de ordem finito têm gênero finito. Já que uma função de gênero finito pode ser factorado de uma forma particular agradável, isto fornece um teorema de fatoração.

LEMA 5.2. *Sejam f uma função inteira não constante de ordem λ com $f(0) = 1$ e $\{a_1, a_2, \ldots\}$ os zeros de f levando em conta sua multiplicidade e ordenados tal que $|a_1| \leqslant |a_2| \leqslant \ldots$. Se um inteiro $p > \lambda - 1$ então*

$$\frac{d^p}{dz^p}\left[\frac{f'(z)}{f(z)}\right] = -p!\sum_{n=1}^{\infty}\frac{1}{(a_n - z)^{p+1}}$$

para $z \neq a_1, a_2, \ldots$.

Demonstração. (Veja [5], pag 289). □

TEOREMA 5.3 (Teorema de fatorizarão de Hadamard). *Se f é uma função inteira de ordem finito de ordem λ, então f tem gênero $\mu \leqslant \lambda$.*

Demonstração. Seja p o inteiro maior menor ou igual a λ, assim $p \leqslant \lambda < p + 1$. Primeiro vamos provar que f tem posto finito e que o posto não pode ser maior que p. Sejam $\{a_1, a_2, \ldots\}$ os zeros de f contados com sua multiplicidade e ordenados tal que $|a_1| \leqslant |a_2| \leqslant \ldots$. Devemos provar que

$$\sum_{n=1}^{\infty}|a_n|^{-(p+1)} < \infty. \tag{5.5}$$

Sem perda de generalidade podemos assumir que $f(0) = 1$. De fato, se f tem um zero na origem de multiplicidade m e $M(r) = \max\{|f(z)| : |z| = r\}$. Então para algum $\epsilon > 0$ e $|z| = r$ temos que

$$\begin{aligned}\log|f(z)z^{-m}| \leqslant \log[M(r)r^{-m}] &\leqslant \log\left(\exp(r^{\lambda})\right) - m\log r \\ &\leqslant r^{\lambda+\epsilon} - m\log r \leqslant r^{\lambda+2\epsilon},\end{aligned}$$

se r é suficientemente grande. Daí $f(z)z^{-m}$ é uma função inteira de ordem λ sem zero na origem. Já que a multiplicação por um escalar não afeta ao ordem, a suposição $f(0) = 1$ está justificado.

Seja $n(r)$ = número de zeros de f em $B(0;r)$ então $n(r)\log 2 \leqslant \log M(r)$ (Veja [5], pag. 284). Desde que f tem ordem λ, $\log M(r) \leqslant r^{\lambda+\frac{1}{2}\epsilon}$ para algum $\epsilon > 0$, disto segue que $\lim\limits_{r\to\infty} n(r)r^{-(\lambda+\epsilon)} = 0$. Por tanto $n(r) \leqslant r^{\lambda+\epsilon}$ para r suficientemente grande.

Como $|a_1| \leqslant |a_2| \leqslant \dots$, $k \leqslant n(|a_k|) \leqslant |a_k|^{\lambda+\epsilon}$ para todo k maior que algum inteiro k_0 temos que

$$|a_k|^{-(p+1)} \leqslant k^{-(p+1)/(\lambda+\epsilon)},$$

para $k > k_0$. Desta forma se ϵ é escolhido tal que $\lambda + \epsilon < p + 1$ (lembre-se que $\lambda < p+1$) logo $\sum |a_k|^{-(p+1)}$ é dominado por uma serie convergente e (5.5) segue.

Seja $f(z) = P(z)\exp(g(z))$ onde P é um produto canônico na forma estândar. Temos que para $z \neq a_k$

$$\frac{f'(z)}{f(z)} = g'(z) + \frac{P'(z)}{P(z)}.$$

Usando o Lema 5.2 ficamos com

$$-p!\sum_{n=1}^{\infty}(a_n - z)^{-(p+1)} = g^{(p+1)}(z) + \frac{d^p}{dz^p}\left[\frac{P'(z)}{P(z)}\right].$$

Porém da definição da forma estândar de $P(z)$ podemos notar que

$$-p!\sum_{n=1}^{\infty}(a_n - z)^{-(p+1)} = \frac{d^p}{dz^p}\left[\frac{P'(z)}{P(z)}\right],$$

para $z \neq a_1, a_2, \dots$.

Por tanto $g^{(p+1)} \equiv 0$ e g tem que ser um polinômio de grau menor ou igual que p. Assim o gênero de $f \leqslant p \leqslant \lambda$ $\square$

5.3 A Função Zeta de Riemann

A função Zeta de Riemann $\zeta(s)$ tem seu origem na identidade expressa pelas duas fórmulas

$$\zeta(s) = \sum_{n\geqslant 1} \frac{1}{n^{-s}},$$

onde n percorre os inteiros positivos, e

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

onde p percorre os números primos. Esta função é uma das mais estudadas em Teoria Analítica dos Números pelo grande número de propriedades que possui. Nesta seção só mencionaremos algumas propriedades que serão de utilidade em nosso estudo. Para um estudo mais profundo desta função sugerimos ao leitor o livro [18]. Vejamos o seguinte lema conhecido como a Fórmula de Soma Parcial de Euler.

LEMA 5.3. *Seja $\phi(x)$ alguma função com derivada contínua no intervalo (a,b). Logo se $[x]$ denota o maior inteiro que não excede x então*

$$\sum_{a<n\leqslant b}\phi(n)=\int_a^b\phi(x)dx+\int_a^b\big(x-[x]-\frac{1}{2}\big)\phi'(x)dx+ \\ +\big(a-[a]-\frac{1}{2}\big)\phi(a)-\big(b-[b]-\frac{1}{2}\big)\phi(b).$$

Demonstração.

$$\begin{aligned}\int_a^b[x]\phi'(x)dx&=\int_{[a]}^{[b]}[x]\phi'(x)dx-\int_{[a]}^{a}[x]\phi'(x)dx+\int_{[b]}^{b}[x]\phi'(x)dx\\&=\sum_{n=[a]}^{[b]-1}n\int_n^{n+1}\phi'(x)dx-[a]\int_{[a]}^{a}\phi'(x)dx+[b]\int_{[b]}^{b}\phi'(x)dx\\&=\sum_{n=[a]}^{[b]-1}n\{\phi(n+1)-\phi(n)\}-[a]\{\phi(a)-\phi([a])\}+[b]\{\phi(b)-\phi([b])\}\\&=\sum_{n=[a]+1}^{[b]}\phi(n)-[a]\phi(a)+[b]\phi(b).\end{aligned}$$

Além do mais

$$\int_a^b\left(x-\frac{1}{2}\right)\phi'(x)dx=\left(b-\frac{1}{2}\right)\phi(b)-\left(a-\frac{1}{2}\right)\phi(a)-\int_a^b\phi(x)dx$$

Portanto

$$\int_a^b\left(x-[x]-\frac{1}{2}\right)\phi'(x)dx=\sum_{n=[a]+1}^{[b]}\phi(n)-\int_a^b\phi(x)dx+\left([a]-a+\frac{1}{2}\right)\phi(a)+ \\ +\left([b]-b+\frac{1}{2}\right)\phi(b)$$

o que prova o resultado. □

Agora, como um casso particular, seja $\phi(n)=n^{-s}$, onde $s\neq 1$ e sejam a e b inteiros positivos. Logo

$$\sum_{n=a+1}^{b}\frac{1}{n^s}=\frac{b^{1-s}-a^{1-s}}{1-s}-s\int_a^b\frac{x-[x]-1/2}{x^{s+1}}dx+\frac{1}{2}(b^{-s}-a^{-s}).$$

Tomando $\sigma>1$, $a=1$, fazer $b\to\infty$ e somando 1 a ambos lados, obtemos

$$\zeta(s)=s\int_1^\infty\frac{x-[x]-1/2}{x^{s+1}}dx+\frac{1}{s-1}+\frac{1}{2}. \tag{5.6}$$

Já que $[x]-x+\frac{1}{2}$ e limitada, esta integral é convergente para $\sigma > 0$, e convergente uniformemente em toda região finita na direita de $\sigma = 0$. Assim define uma função analítica em s, para $\sigma > 0$. Portanto o lado direito fornece uma continuação analítica de $\zeta(s)$ até $\sigma = 0$, e podemos ver claramente um polo simples em $s = 1$ com resíduo 1. Agora vejamos uma aproximação da função Zeta para s perto de um.

Lema 5.4. *Para s tendendo a 1, temos*

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

onde γ é a constante de Euler Mascheroni.

Demonstração. Da equação (5.6), temos que

$$\begin{aligned}
\lim_{s\to 1}\left[\zeta(s) - \frac{1}{s-1}\right] &= \lim_{s\to 1}\left[s\int_1^\infty \frac{[x]-x+\frac{1}{2}}{x^{s+1}}dx + \frac{1}{2}\right] = \int_1^\infty \frac{[x]-x+\frac{1}{2}}{x^2}dx + \frac{1}{2} \\
&= \lim_{n\to\infty}\int_1^n \frac{[x]-x}{x^2}dx + \int_1^\infty \frac{1}{2x^2}dx + \frac{1}{2} \\
&= \lim_{n\to\infty}\left[\sum_{m=1}^{n-1} m\int_m^{m+1}\frac{dx}{x^2} - \int_1^n \frac{dx}{x} + 1\right] \\
&= \lim_{n\to\infty}\left[\sum_{m=1}^{n-1}\frac{1}{m+1} + 1 - \log n\right] \\
&= \lim_{n\to\infty}\left[\sum_{m=1}^{n-1}\frac{1}{m} - \log n\right] = \gamma
\end{aligned}$$

□

Referências Bibliográficas

[1] M. Agrawal, N. kayal e N. Sexana. *Primes is in P*, Annals of Math. (**160**), 781-793, 2004.

[2] Tom M. Apostol. *Introduction to the Analytic Number Theory*, Springer, California Institute of Tecnology, 1976.

[3] R.C. Baker. *Diophantine Inequalitities*, London Mathematical Society, Oxford University Press, 1986.

[4] R. Bowen. *Entropy for Group Endomorphisms and Homogeneous Spaces*, Trans. Amer. Math. Soc., **153**, 401-414, (1971).

[5] John B. Conway.*Functions of One Complex Variable*, Springer-Verlag, New York, 1978.

[6] H. Davenport. *On some infinite series involving arithmetical functions*, Math. Oxf **8** (1937), pag. 313-320.

[7] H.M. Edwards. *Riemann's Zeta Functions*, Dover Publications, New York, 1974.

[8] G.H. Hardy and E.M. Wright. *An Introduction to the Theory of Numbers*, Fourth Edition, Oxford, 1960.

[9] H. Iwaniec and E. Kowalski. *Analytic Number Theory*, American Mathematical Society, USA, 2004.

[10] A.I. Markushevich. *Theory of Functions of a Complex Variable*, Volume II, Prentice-Hall, Inc. 1965.

[11] F. E. Brochero Martinez, C. G. Tamm de A. Moreira, N. C. Saldanha, E. Tengan. *Teoria dos Números: Um passeio com primos e outros números familiares pelo mundo inteiro*, IMPA, Rio de Janeiro, 2011.

[12] K. Oliveira, M. Viana. *Teoria Ergódica: Um Curso Introdutório*, IMPA, 2010.

[13] RYAN PECKNER. *Uniqueness of the Measure of Maximal Entropy for the Squarefree flow*, arxiv:1205.2905v4, 2013.

[14] NIGEL J.E. PITT. *Two cases of the Möbius Randomness Principle via Vaughan's Identity and Bilinear Forms*, Universiy of Maine, 2013.

[15] J. J. RODRÍGUEZ HERRERA. *La Función de Moebius y un Análisis Probabilístico de la Hipótesis de Riemann* Universidad Nacional de Colombia, 2011.

[16] PETER SARNAK. *Three Lectures on the Möbius Function Randomness and Dynamics*, Institute for Advanced Study, New Yersey, 2011.

[17] TERENCE TAO. (14 Outubro 2012) *The Chowla conjecture and the Sarnak conjecture*, Disponível em: http://terrytao.wordpress.com/2012/10/14/the-chowla-conjecture-and-the-sarnak-conjecture/.

[18] E.C. TITCHMARSH. *The Theory of the Riemann Zeta-Function*, University of Oxford, 1951.

[19] I. M. VINOGRADOV. *Representation of an odd number as a sum of three primes*, Dokl.Akad. Nausk SSSR, 1965.

[20] B. WEISS. *Single Orbits Dynamics*, American Mathematical Society, CBMS 95, 2000.

[21] E. WIRSING. *Das asymptotische Verhalten von Summen über multiplikative Funktionen*, Acta Math. Acad. Sci. Hung, 1967.

[22] B. YEMINI. *Entropia Topológica del Flujo Geodésico*, Universidad de la República del Uruguay, 2013.