
Marcus Vinicius Silva Nascimento

# Sobre o crivo de Eratóstenes-Legendre

**CAMPINAS**
**2015**


UNIVERSIDADE ESTADUAL DE CAMPINAS

Instituto de Matemática, Estatística
e Computação Científica

MARCUS VINICIUS SILVA NASCIMENTO

# SOBRE O CRIVO DE ERATÓSTENES-LEGENDRE

Dissertação apresentada ao Instituto de Matemática, Estatística e Computação Científica da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Mestre em matemática aplicada.

**Orientador: Jose Plinio de Oliveira Santos**

ESTE EXEMPLAR CORRESPONDE À VERSÃO FINAL DA DISSERTAÇÃO DEFENDIDA PELO ALUNO MARCUS VINICIUS SILVA NASCIMENTO, E ORIENTADA PELO PROF. DR. JOSE PLINIO DE OLIVEIRA SANTOS.

**Assinatura do Orientador**

CAMPINAS
**2015**

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca do Instituto de Matemática, Estatística e Computação Científica
Ana Regina Machado - CRB 8/5467

N17s

Nascimento, Marcus Vinicius Silva, 1980-
Sobre o crivo de Eratóstenes-Legendre / Marcus Vinicius Silva Nascimento. – Campinas, SP : [s.n.], 2015.

Orientador: José Plínio de Oliveira Santos.
Dissertação (mestrado) – Universidade Estadual de Campinas, Instituto de Matemática, Estatística e Computação Científica.

1. Método de crivo. 2. Eratóstenes, Método de crivo de. 3. Números primos. 4. Análise combinatória. I. Santos, José Plínio de Oliveira,1951-. II. Universidade Estadual de Campinas. Instituto de Matemática, Estatística e Computação Científica. III. Título.

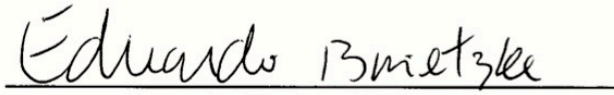
Informações para Biblioteca Digital

**Título em outro idioma:** About the Eratosthenes-Legendre sieve
**Palavras-chave em inglês:**
Sieve method
Eratosthenes sieve method
Prime numbers
Combinatorial analysis
**Área de concentração:** Matemática Aplicada
**Titulação:** Mestre em Matemática Aplicada
**Banca examinadora:**
José Plínio de Oliveira Santos [Orientador]
Eduardo Henrique de Mattos Brietzke
Robson da Silva
**Data de defesa:** 09-04-2015
**Programa de Pós-Graduação:** Matemática Aplicada

**Dissertação de Mestrado defendida em 09 de abril de 2015 e aprovada**

**Pela Banca Examinadora composta pelos Profs. Drs.**

**Prof.(a). Dr(a). JOSÉ PLÍNIO DE OLIVEIRA SANTOS**



**Prof.(a). Dr(a). EDUARDO HENRIQUE DE MATTOS BRIETZKE**

**Prof.(a). Dr(a). ROBSON DA SILVA**

## Abstract

Our aim in this work is to make a study about the *sieve method*. The motivation lies in the intent of applying this idea in a particular situation.

We splitted the study into three parts. The first part deals with definitions and basic concepts. In the second we present the *principle of inclusion-exclusion* while being something well known deserves special mention given its importance as a tool in our work. In the third and final part, we make a historical contextualization and a description of the evolution of the sieve *Eratosthenes-Legendre* ideas.

The choice of sieve, among many others, has been made taking into account two points. The first is that the *Eratosthenes-Legendre* sieve is the simplest among the sieves studied the theory of sieves. The second point is related to the fact that this sieve provide the general idea of combinatorial sieve, since the more sophisticated sieves are extensions of its basic ideas.



**Keywords**: Sieve method, Eratosthenes sieve method, Prime numbers, Combinatorial analysis.



## Resumo

Nosso objetivo, nesse trabalho, é o de fazer um estudo sobre o *método do crivo*. A motivação reside no desejo de aplicar essas ideias a uma situação particular.

Dividimos nosso trabalho em três partes. Na primeira fornecemos apenas as definições e conceitos básicos. Na segunda apresentamos *o principio da inclusão-exclusão* que embora sendo algo bastante conhecido merece destaque especial dada a sua importância como ferramenta no nosso trabalho. Na terceira e última parte, fazemos uma contextualização histórica e uma descrição da evolução das ideias do crivo de Eratóstenes-Legendre.

A escolha desse crivo, dentre tantos outros, foi feita tendo em vista dois pontos. O primeiro é que o crivo de Eratóstenes-Legendre é o mais simples dentre os crivos estudados na teoria dos crivos. O segundo ponto está relacionado com o fato deste crivo fornecer a ideia geral dos crivos combinatoriais, uma vez que os crivos mais sofisticados são extensões de suas ideias básicas.



**Palavras-chave**: Método de crivo, Eratóstenes, Método de crivo de., Números primos, Análise combinatória.

# Sumário

# Agradecimentos

Tão fundamental quanto o fazer é, criar condições para, que o fazer seja feito. Essas condições surgem, sob as mais diversas formas, por meios de pessoas e instituições. E não são poucos aqueles que criaram tais condições. Dedico portanto este espaço, àqueles que de alguma forma criaram condições para que este trabalho fosse feito.

Começo agradecendo à companheira que escolhi para vida, minha querida *Leticia*, pela paciência, carinho e atenção. Aos familiares, minha mãe *Dora*, pai *Elival*, irmã *Juliana*, sogra *Cida*, sogro *Fernando* e cunhadas *Suzi* e *Cecilia* pela ajuda rotineira e cotidiana.

Agradeço também aos amigos do instituto e os da boemia pelo convívio e por despertar em mim algo extremamente necessário, para recarregar as energias e limpar a cabeça para novas ideias, o ócio. Em particular, agradeço ao amigos Adson, Arnoldo, Carpegiane, Denise, Felipo, Jorge, Lino, Maíra, Nelson, Simone, Tatiana, Tatiane, Valter.

Agradeço imensamente, ao meu querido orientador *Prof. Plínio* pela paciência e por suas, sempre valiosas, observações. Agradeço, ainda, à banca pela leitura e por suas importantes observações.

Por fim agradeço à *Capes* pelo apoio financeiro e ao instituto *IMECC*.

# Introdução

As ideias de crivo certamente cresceram a partir de um processo chamado *crivo de Eratóstenes* proposto por volta do Século III a.C., cuja a ideia era eliminar, de uma tabela, aqueles elementos que por alguma razão eram considerados indesejaveis. Os processos atuais de crivos, no entanto, têm uma perspectiva diferente. Enquanto o *crivo de Eratóstenes*, na sua versão original, se preocupa em obter elementos atendendo a condições dadas, os crivos atuais se preocupam em contar o número de elementos atendendo a essas condições.

A teoria moderna de crivos teve início com os trabalhos de *Viggo Brun (1885-1978)* e até 30 anos atrás, três métodos de crivos eram considerados como os principais pilares da teoria, a saber, *crivo de Brun*, *crivo de Selberg* e o *grande crivo de Linnik*. Segundo os autores de [2], a versão moderna do *crivo de Eratóstenes* (tambem conhecido como *crivo de Eratóstenes-Legendre*) quando associado ao *truque de Rankin* torna-se tão poderoso quanto o *crivo de Brun*.

Desde o seu surgimento o *crivo de Eratóstenes* têm sido revisado e reinterpretado por vários matemáticos. Neste trabalho apresentamos a evolução dessas ideias desde sua origem até sua versão moderna e apresentamos, ao leitor, outras ideias comuns aos processos de crivo.

"Fundamental progress has to do with the reinterpretation of basic ideas"
*Alfred N. Whitehead*

# Capítulo 1

# Conceitos Básicos

Neste primeiro capítulo vamos revisar importantes conceitos e definições, assim como os resultados obtidos por meio deles, que serão utilizados ao longo desse trabalho. Começaremos essa revisão tratando de um assunto que permeia boa parte da matemática, a teoria de conjuntos. Não temos a intenção de um aprofundamento no tema. Nosso objetivo aqui é o de apresentar as ideias fundamentais da teoria de conjuntos que serão necessárias para o embasamento e o desenvolvimento do nosso texto. O livro adotado para essa primeira parte foi *Teoria ingênua de conjuntos* escrito por *Paul Halmos* [6].

## 1.1 Teoria de conjuntos

### 1.1.1 Conceitos elementares da teoria de conjuntos

O caminho adotado, para apresentar as ideias fundamentais da teoria de conjuntos, é o axiomático. Nesse sentido, vamos seguir o caminho trilhado por *Paul Halmos*, no livro citado, de um modo bem semelhante e as vezes até coincidente.

Como no trabalho de *Halmos* não daremos uma definição para os conceitos de

*conjuntos*
*pertinência de um elemento a um conjunto.*

A situação é análoga ao tratamento dado por *Euclides* à geometria em seu livro *Os Elementos*. *Euclides*, naquele contexto, não define pontos e retas, ele simplesmente admite a existência de tais objetos e a partir de um conjunto de regras ele deduz resultados. Portanto, no que segue, estaremos admitindo que as ideias de conjunto e de pertinência são conceitos primitivos.

Conjuntos, como intuímos, podem possuir elementos ou não. Se um conjunto $A$ não possui elementos dizemos que ele é vazio e denotamos tal conjunto por $\emptyset$. Sendo $a$ um elemento que pertence ao conjunto $A$ denotamos $a \in A$, e no caso em que $a$ não é um elemento do conjunto $A$ denotamos $a \notin A$.

O conceito primitivo *de um elemento pertencer a um conjunto* produz uma *noção* de igualdade entre conjuntos. Essa relação fundamental, entre tais conceitos, está estabelecida pelo axioma, seguinte:

**Axioma 1.1.1** (Axioma da Extensão)**.** *Dois conjuntos são iguais se, e somente se, possuem os mesmos elementos.*

A igualdade entre os conjuntos $A$ e $B$ será denotada por $A = B$ e no caso em que $A$ e $B$ não são conjuntos iguais denotamos $A \neq B$.

Vamos estabelecer, agora, uma nova relação entre conjuntos através da qual será possível reformular o Axioma (1.1.1). Sejam $A$ e $B$ conjuntos dados, dizemos que $A$ está contido em $B$, $A$ está incluído em $B$ ou ainda que $A$ é um subconjunto de $B$ se todos os elementos de $A$ são ainda elementos de $B$ e representaremos esse fato por

$$A \subset B.$$

O Axioma da Extensão em termos dessa nova relação entre conjuntos, chamada inclusão, estabelece que,

$$A = B \iff A \subset B \text{ e } B \subset A. \tag{1.1.1}$$

Essa nova caracterização fornece um método para se demonstrar a igualdade entre conjuntos. Isto é, para mostrar que $A = B$ é necessário e suficiente mostrar que,

$$A \subset B \text{ e } B \subset A.$$

Como observação é importante notar que as relações acima, de pertinência e de inclusão, são essencialmente diferentes. Para destacar uma das diferenças sabemos que a relação de inclusão é reflexiva uma vez que, por definição, todos os elementos de um dado conjunto $A$ pertencem a $A$, isto é,

$$A \subset A.$$

Por outro lado não é de todo verdade que $A \in A$. E isso é suficiente para garantir uma difença entre as relações de pertinência e de inclusão.

Um fato muito utilizado, na teoria de conjuntos, é a construção de novos conjuntos a partir de conjuntos já existentes. Deste ponto até o final desta seção nosso objetivo será o de estabelecer alguns mecanismos para a construção desses novos conjuntos. O primeiro passo nesse sentido é dado pelo axioma:

**Axioma 1.1.2** (Axioma da Especificação)**.** *A todo conjunto $A$ e a toda condição $S(x)$ podemos associar um conjunto $B$ cujos elementos são exatamente aqueles elementos $x$ de $A$ para os quais a condição $S(x)$ vale.*

O axioma acima estabelece que qualquer afirmação sensata, feita a elementos de um conjunto, especifica um novo conjunto.

Uma consequência imediata do Axioma (1.1.1) é que o Axioma da Especificação determina o conjunto $B$ de maneira única, uma vez que,

$$b \in B \iff b \in \{x \in A \mid S(x)\}$$

o que implica,

$$B = \{x \in A \mid S(x)\}$$

onde a expressão $S(x)$ denota uma afirmação feita ao elemento $x \in A$.

O próximo axioma garante que para quaisquer dois conjuntos existe um conjunto que contém ambos e nada mais.

**Axioma 1.1.3** (Axioma da Paridade)**.** *Para dois conjuntos quaisquer existe um conjunto $A$ a que ambos pertencem.*

O Axioma (1.1.3) não exclui a possibilidade de existirem mais do que dois elementos em $A$. No entanto, se $a$ e $b$ são conjuntos, e se $A$ é um conjunto tal que $a \in A$ e $b \in A$, podemos aplicar o Axioma (1.1.2) para $A$ com a sentença: $x = a \quad ou \quad x = b$. O resultado é o conjunto,

$$\{x \in A \mid x = a \text{ ou } x = b\},$$

e este conjunto, obviamente, contém apenas $a$ e $b$. A representação usual para esse último conjunto é,

$$\{a,\ b\}.$$

### 1.1.2 União e Intersecção

**Axioma 1.1.4** (Axioma da Reunião)**.** *Para toda coleção de conjuntos existe um conjunto que contém todos os elementos, de cada um dos conjuntos, da coleção dada.*

Em termos de símbolos, o axioma acima pode ser reescrito como:

$$(\forall\ \mathcal{C})\ (\exists\ U)\ (u \in U \iff u \in \{x \mid x \in X \text{ para algum } X \text{ em } \mathcal{C}\}).$$

O conjunto $U$, acima, é chamado a união (ou reunião) dos conjuntos $X$ da coleção $\mathcal{C}$ e garantimos, pelo Axioma da Extensão, a unicidade desse conjunto, isto é,

$$U = \{x \mid x \in X \text{ para algum } X \text{ em } \mathcal{C}\}. \tag{1.1.2}$$

Uma outra maneira, bastante utilizada, para representar o conjunto $U$ é dada por,

$$U = \bigcup_{X \in \mathcal{C}} X$$

e no caso em que a coleção $\mathcal{C}$ é dada por $\{A, B\}$, isto é $\mathcal{C} = \{A, B\}$, temos a representação

$$\bigcup \{X \mid X \in \mathcal{C}\} = A \cup B.$$

Assim, pela caracterização (1.1.2) temos que,

$$A \cup B = \{x \mid x \in A \text{ ou } x \in B\}.$$

Alguns fatos importantes, sobre a reunião de conjuntos, são listados abaixo:

1. $A \cup \emptyset = A$;

2. $A \cup B = B \cup A$;

3. $A \cup (B \cup C) = (A \cup B) \cup C$;

4. $A \cup A = A$;

5. $A \subset B$ se, e somente se, $A \cup B = B$.

*Demonstração.* 1. Claramente temos que $A \subset A \cup \emptyset$. Suponha no entanto que $A \cup \emptyset \not\subset A$. Então existe $x \in A \cup \emptyset$ tal que $x \notin A$. Assim como $x \in A \cup \emptyset$ temos que $x \in A$ ou $x \in \emptyset \Rightarrow x \in \emptyset$ o que é um absurdo. Portanto, segue $A \cup \emptyset \subset A$ e daí $A \cup \emptyset = A$.

2. Temos que $x \in A \cup B$ se, e somente se, $x \in A$ ou $x \in B$. Alem disso $x \in A$ ou $x \in B$ se, e somente se, $x \in B$ ou $x \in A$. Mas $x \in B$ ou $x \in A$ se, e somente se, $x \in B \cup A$.

3. Temos que $x \in A \cup (B \cup C)$ se, e somente se, $x \in A$ ou $x \in (B \cup C)$ se, e somente se, $x \in A$ ou $x \in B$ ou $x \in C$. Mas $x \in A$ ou $x \in B$ ou $x \in C$ se, e somente se, $x \in (A \cup B)$ ou $x \in C$ se, e somente se, $x \in (A \cup B) \cup C$.

4. Claramente temos que $A \subset A \cup A$. Suponha no entanto que $A \cup A \not\subset A$. Então existe $x \in A \cup A$ tal que $x \notin A$. Como $x \in A \cup A$ então $x \in A$ ou $x \in A \Rightarrow x \in A$ o que é uma contradição. Portanto segue que $A \cup A \subset A$ e daí $A \cup A = A$.

5. Suponha que $A \subset B$. Como $B \subset B$ temos que $A \cup B \subset B$, uma vez que para todo $x \in A \cup B$ temos que $x \in B$. Desde que claramente $B \subset A \cup B$ segue $A \cup B = B$. Por outro lado se $A \cup B = B$ então para todo $x \in A \cup B$ temos $x \in B$. Assim, para todo $x \in A$ ou $x \in B \Rightarrow x \in B$. Ou seja, em particular, todo $x \in A \Rightarrow x \in B$, isto é $A \subset B$.

□

Uma outra operação, que têm vários pontos em comum com a união de conjuntos, é a *intersecção* de conjuntos. Sejam os conjuntos $A$ e $B$, definimos o conjunto intersecção entre $A$ e $B$ como segue,

$$A \cap B = \{x \mid x \in A \text{ e } x \in B \}. \tag{1.1.3}$$

Alguns fatos sobre a intersecção são listados:

1. $A \cap \emptyset = \emptyset$;

2. $A \cap B = B \cap A$;

3. $A \cap (B \cap C) = (A \cap B) \cap C$;

4. $A \cap A = A$;

5. $A \subset B$ se, e somente se, $A \cap B = A$.

*Demonstração.* 1. Primeiramente temos que $\emptyset \subset A \cap \emptyset$, uma vez que, se $\emptyset \not\subset A \cap \emptyset$ existiria $x \in \emptyset$ tal que $x \notin A \cap \emptyset$ o que é um absurdo. Logo devemos ter $\emptyset \subset A \cap \emptyset$.

   Por outro lado $A \cap B \subset B$ para todo $B$, de fato, se $A \cap B \not\subset B$ existiria $x \in A \cap B$ tal que $x \notin B$. No entanto para todo $x \in A \cap B \Rightarrow x \in A$ e $x \in B$ o que é uma contradição. Portanto $A \cap B \subset B$ para todo $B$. Em particular para $B = \emptyset$ segue o resultado.

2. Temos que $x \in A \cap B$ se, e somente se, $x \in A$ e $x \in B$. Além disso $x \in A$ e $x \in B$ se, e somente se, $x \in B$ e $x \in A$. Mas $x \in B$ e $x \in A$ se, e somente se, $x \in B \cap A$.

3. Temos que $x \in A \cap (B \cap C)$ se, e somente se, $x \in A$ e $x \in (B \cap C)$ se, e somente se, $x \in A$ e $x \in B$ e $x \in C$. Mas $x \in A$ e $x \in B$ e $x \in C$ se, e somente se, $x \in (A \cap B)$ e $x \in C$ se, e somente se, $x \in (A \cap B) \cap C$.

4. Vimos em (1) que $A \cap B \subset B$ para todo $B$, portanto tomando $B = A$ segue que $A \cap A \subset A$. Por outro lado é claro que para todo $x \in A$, $x \in A \cap A$. Portanto segue que $A = A \cap A$.

5. Suponha que $A \subset B$. Como necessariamente $A \subset A$ temos que $A \subset A \cap B$, uma vez que para todo $x \in A$ temos que $x \in A$ e $x \in B$. Assim, utilizando o fato de que $A \cap B \subset A$ concluímos que $A \cap B = A$.

   Por outro lado, se $A \cap B = A$ então para todo $x \in A$ temos $x \in A \cap B \Rightarrow x \in A$ e $x \in B$. Portanto, em particular temos que para todo $x \in A \Rightarrow x \in B$ ou seja $A \subset B$.

□

A formação de um conjunto intersecção a partir de dois conjuntos é um caso particular de uma operação muito mais geral que é estabelecida pelo axioma,

**Axioma 1.1.5** (Axioma da Intersecção)**.** [1] *Para cada coleção (não vazia) de conjuntos, existe um conjunto que contém os elementos comuns a todos os conjuntos da dada coleção.*

Em termos de símbolo o Axioma da Intersecção, acima, pode ser reescrito como:

$$(\forall\, \mathcal{C})\ (\ \mathcal{C} \neq \emptyset)\ (\exists\, V)\ (v \in V \iff v \in \{x \mid x \in X \text{ para todo } X \text{ em } \mathcal{C}\}).$$

O conjunto $V$, acima, é chamado a intersecção dos conjuntos $X$ da coleção $\mathcal{C}$ e garantimos, pelo Axioma da Extensão, a unicidade desse conjunto, isto é,

$$V = \{x \mid x \in X \text{ para todo } X \text{ em } \mathcal{C}\}.$$

Uma outra maneira, bastante usual, de representar o conjunto $V$ é dada por,

$$V = \bigcap_{X \in \mathcal{C}} X.$$

Em concordância com que estabelecemos no conjunto (1.1.3), no caso particular, em que $\mathcal{C} = \{A, B\}$ temos a representação,

[1] Sob um ponto de vista completamente axiomático, temos que esse axioma é redundante uma vez que ele pode ser obtido por repetidas aplicações do Axioma da Especificação.

$$\bigcap\{X \mid X \in \mathcal{C}\} = A \cap B.$$

Os fatos estabelecidos a seguir nos dizem como relacionar as operações de união e intersecção. Sejam dados os conjuntos $A$,$B$ e $C$. Então,

1. $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$;
2. $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.

*Demonstração.* Vamos demonstrar aqui somente o item 1 uma vez que a demonstração do item 2 é análoga.

Se $x \in A \cap (B \cup C)$ então, $x \in A$ e $x \in B \cup C$. Assim, $x$ pertence a $A$ e $x$ pertence ou a $B$ ou a $C$. Se $x$ pertence a $B$, e como $x$ pertence obrigatoriamente a $A$, segue que $x$ pertence a $A \cap B$. Na outra possibilidade $x$ pertence a $C$ e novamente pela obrigatoriedade de $x$ pertencer a $A$ segue que $x$ pertence a $A \cap C$. Como não existe outra possibilidade, temos que $x$ pertence a $A \cap B$ ou $x$ pertence a $A \cap C$ o que significa dizer que $x$ pertence a $(A \cap B) \cup (A \cap C)$. Isso mostra portanto que $A \cap (B \cup C) \subset (A \cap B) \cup (A \cap C)$.

Seja agora $x \in (A \cap B) \cup (A \cap C)$. Neste caso temos que $x \in A$ e $x \in B$ ou $x \in A$ e $x \in C$. Em qualquer uma, das duas possibilidades, $x \in A$ e $x$ pertence ou a $B$ ou a $C$, isto é $x \in A \cap (B \cup C)$. Assim temos $(A \cap B) \cup (A \cap C) \subset A \cap (B \cup C)$ e isso conclui a demonstração. □

### 1.1.3 Complemento e Potência

Vamos estabelecer, agora, a noção de diferença entre conjuntos. Sejam $A$ e $B$ conjuntos dados. Definimos a diferença $B \setminus A$ por,

$$B \setminus A = \{x \in B \mid x \notin A\}.$$

e chamamos esse conjunto de o complemento relativo de $A$ em $B$. Observe que nesta definição não foi necessário o fato de $A$ ser subconjunto de $B$.

Em uma situação um pouco mais especifica, quando todos os subconjuntos $A$ a serem considerados são subconjuntos de $B$, podemos adaptar a noção de diferença entre conjuntos nesse contexto e estabelecer a noção de conjunto complementar. O *complemento* de um subconjunto $A$ de $B$, que denotaremos por $A^c$, é o conjunto de todos aqueles pontos de $B$ que não pertencem à $A$, em símbolos,

$$A^c = \{x \in B \mid x \notin A\}.$$

Os fatos básicos sobre a complementação podem ser enunciados como segue.
Sejam $A, C \subset B$, então,

1. $(A^c)^c = A$;

2. $\emptyset^c = B$ e $B^c = \emptyset$;

3. $A \cap A^c = \emptyset$ e $A \cup A^c = B$;

4. $A \subset C$ se, e somente se, $C^c \subset A^c$.

*Demonstração.* 1. Temos que $x \in (A^c)^c$ se, e somente se, $x \notin A^c$ se, e somente se, $x \in A$.

2. Por definição, temos que $\emptyset^c = \{x \in B \mid x \notin \emptyset\} = B$. Por outro lado, $B^c = \{x \in B \mid x \notin B\} = \emptyset$.

3. Por definição, temos que $A^c = \{x \in B \mid x \notin A\}$, assim $A \cap A^c = \emptyset$. Por outro lado, $A \cup A^c = B$, uma vez que $A = \{x \in B \mid x \in A\}$.

4. $A \subset C$ se, e somente se, $x \in A \Rightarrow x \in C$ se, e somente se, $x \notin C \Rightarrow x \notin A$ se, e somente se, $x \in C^c \Rightarrow x \in A^c$ se, e somente se, $C^c \subset A^c$.

□

Uma observação importante e útil, sobre complementação, conhecida como leis *De Morgan* são expressas pelas identidades abaixo,

1. $(A \cup B)^c = A^c \cap B^c$;

2. $(A \cap B)^c = A^c \cup B^c$.

*Demonstração.* Vamos demonstrar aqui somente o item 1 uma vez que a demonstração do item 2 é análoga.

De fato, dado $x \in (A \cup B)^c$ temos que $x \notin (A \cup B)$. Assim, $x \notin A$ e $x \notin B$. Logo $x \in A^c$ e $x \in B^c$ e, portanto $x \in A^c \cap B^c$, daí $(A \cup B)^c \subset A^c \cap B^c$. Por outro lado, seja $x \in A^c \cap B^c$, então $x \in A^c$ e $x \in B^c$. Assim, $x \notin A$ e $x \notin B$ e portanto $x \notin A \cup B$. Logo $x \in (A \cup B)^c$ o que conclui a demonstração.

□

De um modo muito mais geral, valem as seguintes relações,

$$\left(\bigcup_{X \in \mathcal{C}} X\right)^c = \bigcap_{X \in \mathcal{C}} X^c \quad \text{e} \quad \left(\bigcap_{X \in \mathcal{C}} X\right)^c = \bigcup_{X \in \mathcal{C}} X^c.$$

Os fatos sobre complementação implicam que, usualmente, os teoremas em teoria dos conjuntos vêm aos pares. Se em uma equação (ou inclusão) envolvendo uniões, intersecções e complementos nós trocarmos cada conjunto pelo seu complemento, mudarmos uniões por intersecções (e vice versa) e invertermos todas as inclusões, o resultado será um outro teorema. Este fato é algumas vezes chamado de **principio da dualidade** para conjuntos.

**Axioma 1.1.6** (Axioma das Potências)**.** *Para cada conjunto dado, existe uma coleção de conjuntos que contém entre seus elementos todos os subconjuntos do conjunto dado.*

Em outras palavras se $A$ é um conjunto, então existe um conjunto $\mathcal{P}$ tal que,

$$\text{Se } \ X \subset A \ \text{ então } \ X \in \mathcal{P}.$$

O conjunto $\mathcal{P}$ descrito pelo Axioma (1.1.6) pode ser maior do que o desejado, isto é, o axioma não exclui a possibilidade de existirem outros elementos que não sejam subconjuntos de $A$ em $\mathcal{P}$. Mas isto é facilmente remediado aplicando o Axioma (1.1.2) para formar o conjunto,

$$\{X \in \mathcal{P} \ \mid \ X \subset A\}. \tag{1.1.4}$$

Desde que uma condição necessária e suficiente para que $X$ pertença a (1.1.4) é que $X$ seja um subconjunto de $A$, segue que mudando a notação e denotando este conjunto ainda por $\mathcal{P}$ obtemos,

$$\mathcal{P} = \{X \ \mid \ X \subset A\}.$$

O conjunto $\mathcal{P}$ é chamado o conjunto potência de $A$ e o Axioma (1.1.1) garante sua unicidade. A dependência entre $\mathcal{P}$ e $A$ é denotada escrevendo $\mathcal{P}(A)$ ao invés de somente $\mathcal{P}$.

Vejamos uns poucos exemplos para $\mathcal{P}(A)$ em que a quantidade de elementos em $A$ é pequena.

Se $A = \emptyset$ então $\mathcal{P}(A) = \{\emptyset\}$. Se $A = \{a\}$ então $\mathcal{P}(A) = \{\emptyset, \{a\}\}$. Se $A = \{a, b\}$ então $\mathcal{P}(A) = \{\emptyset, \{a\}, \{b\}, \{a, b\}\}$.

De um modo geral, iremos mostrar, na próxima seção, que se o conjunto $A$ possui $n$ elementos então $\mathcal{P}(A)$ irá possuir $2^n$ elementos. Por enquanto, vejamos um último assunto sobre teoria de conjuntos.

### 1.1.4 Par Ordenado

Um par ordenado de $a$ e $b$, com primeira coordenada $a$ e segunda coordenada $b$, é o conjunto $(a, b)$ definido por

$$(a, b) = \{\{a\}, \{a, b\}\}.$$

Vamos provar que com essa definição a principal propriedade que se espera ter para um par ordenado se mantém. Isto é, vamos mostrar que se $(a, b)$ e $(x, y)$ são pares ordenados com $(a, b) = (x, y)$ então $a = x$ e $b = y$.

Para provar isto nós observamos que se $a$ e $b$ são iguais então o par $(a, b)$ é o conjunto $\{\{a\}\}$. Se reciprocamente $(a, b)$ é o conjunto $\{\{a\}\}$ então $\{a\} = \{a, b\}$ e assim $b \in \{a\}$ e, portanto $a = b$.

Suponha agora que $(a, b) = (x, y)$. Vimos acima que $a = b$ se, e somente se, $(a, b) = \{\{a\}\}$. Portanto supondo $a = b$ e $(a, b) = (x, y)$ devemos ter que,

$$\{\{a\}\} = \{\{x\}, \{x, y\}\},$$

e portanto

$$\{x\} = \{x, y\} \Rightarrow x = y \ \text{ e, } \ \{a\} = \{x\} \Rightarrow x = a.$$

daí segue que $a$,$b$,$x$ e $y$ são todos iguais.

Se, agora supomos $a \neq b$ e $(a,b) = (x,y)$, então ambos $(a,b)$ e $(x,y)$ contém exatamente um conjunto da forma $\{z\}$, ou seja $\{a\}$ e $\{x\}$ respectivamente, de modo que $a = x$. Desde que neste caso é também verdade que ambos $(a,b)$ e $(x,y)$ contém exatamente um par não ordenado, a saber $\{a,b\}$ e $\{x,y\}$ respectivamente, segue-se que $\{a,b\} = \{x,y\}$ e portanto em particular $b \in \{x,y\}$. Considerando que $b$ não pode ser $x$ (visto que $x = a$ e $b \neq a$) devemos ter $b = y$. Assim podemos concluir que para quaisquer pares ordenados $(a,b)$ e $(x,y)$ satisfazendo $(a,b) = (x,y)$ devemos ter $a = x$ e $b = y$.

Para finalizar vamos mostrar que, dados conjuntos $A$ e $B$, existe um conjunto que contém todos os pares ordenados $(a,b)$ com $a \in A$ e $b \in B$.

De fato, se $a \in A$ e $b \in B$, então $\{a\} \subset A$ e $\{b\} \subset B$ e portanto $\{a,b\} \subset A \cup B$. Desde que, também $\{a\} \subset A \cup B$, segue que ambos $\{a\}$ e $\{a,b\}$ são elementos do conjunto $\mathcal{P}(A \cup B)$. Isto implica que $\{\{a\},\{a,b\}\}$ é um subconjunto de $\mathcal{P}(A \cup B)$ e portanto que ele é um elemento de $\mathcal{P}(\mathcal{P}(A \cup B))$, em outras palavras $(a,b) \in \mathcal{P}(\mathcal{P}(A \cup B))$ sempre que $a \in A$ e $b \in B$. Uma vez que este conjunto é conhecido, é uma questão de rotina aplicar o Axioma da Especificação e o Axioma da Extensão para produzir o conjunto único $A \times B$ que consiste exatamente dos pares ordenados $(a,b)$ com $a \in A$ e $b \in B$. Este conjunto é chamado produto cartesiano de $A$ e $B$ e é caracterizado pelo seguinte fato

$$A \times B = \{x \mid x = (a,b) \text{ para algum } a \in A \text{ e para algum } b \in B\}.$$

## 1.2 Combinatória Enumerativa

No que segue apresentaremos um recorte de resultados básicos de Combinatória Enumerativa. E no mesmo sentido, da apresentação da teoria de conjuntos, nosso objetivo é, novamente, o de estabelecer as ideias fundamentais para o desenvolvimento do nosso trabalho. Os resultados, aqui apresentados, são básicos e podem ser encontrados em vários livros introdutórios. Nesse texto foram utilizados os livros, [1], [9] e [10].

### 1.2.1 Princípios elementares de contagem

Iniciamos esta seção com umas poucas regras que, embora obvias, estão na origem dos problemas de contagem. Para estabelecer tais regras precisamos da definição,

**Definição 1.2.1.** *Dado um conjunto finito $A$, denotaremos o número de elementos desse conjunto por $\mid A \mid$.*

Todos os conjuntos, com os quais trabalharemos nesta Seção, são finitos.

**Principio aditivo.** Se $S = \bigcup_{i=1}^{t} S_i$ é a união disjunta[2] de conjuntos $S_i$, então $\mid S \mid = \sum_{i=1}^{t} \mid S_i \mid$.

[2]Dizemos que a união $S$, definida acima, é disjunta quando $S_i \cap S_j = \emptyset$ sempre que $i \neq j$.

Uma maneira bastante comum em que o **principio aditivo** aparece é a seguinte: classificamos os elementos de $S$ de acordo com um conjunto de propriedades $e_i$ $(i = 1, \ldots, t)$ de modo que essas propriedades $e_i$ sejam mutuamente exclusivas, e definimos $S_i = \{x \in S \mid x \text{ têm a propriedade } e_i\}$. Assim para saber o número de elementos em $S$ é suficiente saber o número de elementos nos conjuntos $S_i$'s. Vejamos um exemplo de aplicação do **principio aditivo** que ilustra essa ideia.

**Exemplo 1.2.2.** *Ao longo do texto um conjunto $X$ com $n$ elementos também será chamado um $n$-conjunto. Denote por $S = \binom{X}{k}$ a família de todos $k$-subconjuntos de $X$. Assim o número de elementos de $S$, isto é $\mid S \mid$, é o usual coeficiente binomial $\binom{n}{k}$. Para o momento,*

$$\binom{n}{k}$$

*é um símbolo denotando a quantidade de $k$-subconjuntos de $X$.*

*Seja $a \in X$. Dividimos o conjunto $S$ como segue,*

$$S_1 = \{A \in S \mid a \in A\}$$

*e,*

$$S_2 = \{A \in S \mid a \notin A\}.$$

*Obtemos todos os elementos de $S_1$ através da combinação de todos $k-1$-subconjuntos de $X \setminus \{a\}$ com $a$; assim $\mid S_1 \mid = \binom{n-1}{k-1}$. Similarmente $S_2$ é a família de todos $k$-subconjuntos de $X \setminus \{a\}$ e portanto $\mid S_2 \mid = \binom{n-1}{k}$. O* **principio aditivo** *produz, portanto, a recorrência de Pascal para coeficientes binomiais,*

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k} \quad (1 \leq k \leq n)$$

*com valor inicial $\binom{n}{0} = 1$. Observe que obtivemos esta recorrência sem ter calculado o coeficiente binomial explicitamente.*

Com o propósito de enunciar um outro principio importante considere, em um caminho análogo ao trilhado por nós na Seção (1.1.4), $S$ como sendo o conjunto de todas $t$-uplas $(a_1, a_2, \ldots, a_t)$ com $a_i \in S_i$, onde $S_i$ é chamado o conjunto de coordenadas.

**Principio multiplicativo.** Se $S = \prod_{i=1}^{t} S_i$ é o produto cartesiano de conjuntos, então $\mid S \mid = \prod_{i=1}^{t} \mid S_i \mid$.

O **principio multiplicativo** estabelece que, se existem $t$ escolhas sucessivas a serem feitas, e para $1 \leq i \leq t$, a $i$-ésima escolha pode ser feita em $\mid S_i \mid$ formas, então o número total de maneiras de se fazer estas escolhas é $\prod_{i=1}^{t} \mid S_i \mid$. Vejamos um exemplo.

**Exemplo 1.2.3.** *Uma sequência finita de 0's e 1's é chamada uma palavra sobre o alfabeto $\{0, 1\}$, e a soma das quantidades de 0's e 1's fornece o tamanho da palavra.*

*Seja, $S_i = \{0,1\}$ um conjunto de coordenada para todo $1 \leq i \leq n$ e considere o conjunto*

$$S_1 \times S_2 \times \ldots \times S_n.$$

*Então o* **principio multiplicativo** *estabelece que existem $2^n$ palavras de comprimento $n$ sobre $\{0,1\}$.*

*De fato, como $S_i = \{0,1\}$ para todo $1 \leq i \leq n$ então os elementos do conjunto,*

$$S_1 \times S_2 \times \ldots \times S_n$$

*são da forma $(x_1, x_2, \ldots, x_n)$, onde $x_i \in S_i = \{0,1\}$.*

*Assim dado $x_1$, existem duas opções de escolha para $x_2$ na palavra $(x_1, x_2, \ldots, x_n)$, ou seja, dado $x_1 = 0$ temos,*

$$(0, 0, \ldots, x_n) \quad ou \quad (0, 1, \ldots, x_n)$$

*e para $x_1 = 1$ temos,*

$$(1, 0, \ldots, x_n) \quad ou \quad (1, 1, \ldots, x_n).$$

*Continuando nessa ideia, para cada uma dessas $2^2$ palavras existem duas opções de escolha para $x_3$ e, portanto, teremos um total de $2^3$ palavras ocupando as têS primeiras posições com opções de escolha em $\{0,1\}$, a saber*

$$(0,0,0,x_4,\ldots,x_n); (0,0,1,x_4,\ldots,x_n); (0,1,0,x_4,\ldots,x_n); (0,1,1,x_4,\ldots,x_n);$$
$$(1,0,0,x_4,\ldots,x_n); (1,0,1,x_4,\ldots,x_n); (1,1,0,x_4,\ldots,x_n) \text{ e } (1,1,1,x_4,\ldots,x_n).$$

*De um modo geral, se existe um total de $2^{i-1}$ palavras ocupando as $i-1$ primeiras posições com opções de escolha em $\{0,1\}$ então para as $i$ primeiras posições, com opções de escolha em $\{0,1\}$, existe um total de $2^i$ palavras.*

*Em geral, obtemos $r^n$ palavras se o alfabeto $A$ possuir $r$ elementos. Então falamos da $n$-palavra sobre o alfabeto $A$.*

Para finalizar a seção, vamos estabelecer o **principio bijetivo**. Para tanto, considere as seguintes definições e resultados.

**Definição 1.2.4.** *Dados dois conjuntos $Y$ e $Z$, dizemos que uma função $f : Y \to Z$ é bijetiva quando:*

1. *Para todo $a, b \in Y$ com $a \neq b \Rightarrow f(a) \neq f(b)$;*
2. *Dado qualquer $b \in Z$ existe $a \in Y$ tal que $f(a) = b$.*

**Definição 1.2.5.** *Sejam, $g : X \to Y$ e $f : Y \to Z$ funções dadas. Podemos definir uma função $f \circ g : X \to Z$ tal que $f \circ g(x) = f(g(x))$ para todo $x \in X$.*

**Lema 1.2.6.** *Sejam, $g : X \to Y$ e $f : Y \to Z$ funções bijetivas, então a função $f \circ g(x) = f(g(x))$ será bijetiva.*

*Demonstração.* Se $g$ e $f$ são funções bijetivas então dados $a, b \in X$ com $a \neq b$ temos $g(a) \neq g(b)$. Como $g(a), g(b) \in Y$ e $g(a) \neq g(b)$ segue que $f(g(a)) \neq f(g(b))$. Logo, sabendo que $f \circ g(x) = f(g(x))$ temos que para todo $a, b \in X$ com $a \neq b \Rightarrow f \circ g(a) \neq f \circ g(b)$ e, portanto, 1 na Definição (1.2.4) vale.

Seja agora, $z \in Z$. Como $f : Y \to Z$ é bijetiva, existe, pelo item 2 da Definição (1.2.4), um elemento $y \in Y$ tal que $f(y) = z$. Alem disso, como $g : X \to Y$ também é bijetiva existe, pelo item 2 da Definição (1.2.4), um elemento $x \in X$ tal que $g(x) = y$. Assim dado qualquer $z \in Z$ existe $x \in X$ tal que $f \circ g(x) = f(g(x)) = f(y) = z$ e, portanto, 2 na Definição (1.2.4) vale e o lema está demonstrado. □

**Principio Bijetivo.** Se existir uma bijeção entre dois conjuntos $A$ e $B$, então $\mid A \mid = \mid B \mid$.

*Demonstração.* De fato, dados conjuntos $A$ e $B$, suponha que exista uma bijeção $f : A \to B$. Se $\mid A \mid = n$, existe uma bijeção $g : I_n \to A$, onde $I_n$ é o conjunto dos naturais de 1 a $n$. Então usando o fato dado no Lema (1.2.6) temos que $f \circ g : I_n \to B$ é bijeção e portanto $\mid B \mid = n$. □

O exemplo abaixo vai fazer uso do **principio bijetivo** para estabelecer uma afirmação feita na Seção (1.1.3). Naquela situação, nos afirmamos que, se um conjunto $A$ possui $n$ elementos então o conjunto $\mathcal{P}(A)$ possui, necessariamente, $2^n$ elementos.

**Exemplo 1.2.7.** *Sejam $X$ um conjunto qualquer com $n$ elementos e $S_i = \{0, 1\}$ para todo $1 \leq i \leq n$. Indexando $X = \{x_1, x_2, \ldots, x_n\}$ de qualquer maneira e mapeando*

$$f : \mathcal{P}(X) \to S_1 \times S_2 \times \ldots \times S_n$$

*como $f(A) = (a_1, a_2, \ldots, a_n)$, onde $a_i = 1$ se $x_i \in A$ e $a_i = 0$ se $x_i \notin A$, temos uma bijeção entre $\mathcal{P}(X)$ e as palavras de comprimento $n$ sobre o alfabeto $\{0, 1\}$.*

*Sabemos, do Exemplo (1.2.3), que o conjunto das palavras com tamanho $n$ sobre o alfabeto $\{0, 1\}$ tem um total de $2^n$ elementos. Assim existindo tal bijeção, o número de elementos de $\mathcal{P}(X)$ será $2^n$.*

*Vejamos que, de fato, $f$ é uma bijeção. Sejam $A, B \in \mathcal{P}(X)$ com $A \neq B$ e suponha que $f(A) = f(B)$. Como estamos supondo $f(A) = f(B)$ segue que $a_i = b_i$ para todo $i$ e daí seguiria que $A$ e $B$ têm os mesmos elementos e portanto $A = B$, mas isso é uma contradição. Logo devemos ter $f(A) \neq f(B)$.*

*Por outro lado, dado qualquer palavra de comprimento $n$ sobre o alfabeto $\{0, 1\}$, digamos $(a_1, a_2, \ldots, a_n)$, existe um subconjunto $A \subset X$ tal que*

$$A = \{x_i \in X \mid a_i = 1\}.$$

*nós concluímos, portanto, que $\mid \mathcal{P}(X) \mid = 2^n$.*

### 1.2.2 Permutações e Combinações

Vamos estabelecer agora algumas notações e terminologias que descrevem situações especiais de aplicações dos princípios aditivo e multiplicativo.

**Definição 1.2.8.** *Seja $X$ um conjunto com $n$ elementos. Uma $k$-permutação de $X$ é uma $k$-palavra sobre o alfabeto $X$ cujas entradas são todas distintas.*

**Exemplo 1.2.9.** *Exemplos simples de 4-permutações do conjunto $\{1, 2, 3, 4, 5, 6\}$ são 1235, 1426, 5614, etc.*

Observe que, fazendo uso do **principio multiplicativo**, poderíamos calcular o número de $k$-permutações rapidamente.

**Proposição 1.2.10.** *Dado $X$ um conjunto com $n$ elementos. O número de $k$-permutações de $X$ é,*

$$n \cdot (n-1) \cdot (n-2) \ldots (n-k+1).$$

*Demonstração.* Claramente, temos $n$ possibilidades para a escolha do primeiro dígito. Uma vez que a primeira entrada foi escolhida, existem $n-1$ possíveis escolhas para a segunda entrada, e assim por diante. Depois de $k-1$ objetos terem sido escolhidos restam $n-(k-1) = n-k+1$ objetos para fazer a $k$-ésima escolha. O **principio multiplicativo** nos dá,

$$n \cdot (n-1) \cdot (n-2) \ldots (n-k+1).$$

onde $1 \leq k \leq n$. Claramente o número de $k$-permutações poderia ser escrito sob notação fatorial como

$$\frac{n!}{(n-k)!},$$

assim para $k = 0$ definimos que o número de $k$-permutações vale 1. □

**Exemplo 1.2.11.** *No nosso Exemplo (1.2.9), temos que $n = 6$ e $k = 4$. Portanto o número de 4-permutações de um conjunto com 6 elementos é $6 \cdot 5 \cdot 4 \cdot 3 = 360$.*

Um resultado que segue diretamente da Proposição (1.2.10) é,

**Corolário 1.2.12.** *Quando $k = n$ obtemos, em particular, que o número de $n$-permutações de $X$ é dado por,*

$$n! = n \cdot (n-1) \ldots 2 \cdot 1.$$

Considere agora o seguinte problema:

Dado um conjunto com $n$ elementos, denotado por $X$, qual o número de $k$-subconjuntos de $X$?

De acordo com a notação estabelecida no Exemplo (1.2.2) queremos saber quanto vale

$$\binom{n}{k},$$

e a resposta surge quando olhamos as palavras de comprimento $k$ com símbolos em $X$.

Sejam $i_1, i_2, \ldots, i_k$ elementos fixos em $X$. Claramente as $k$-palavras formadas por esses elementos estão associadas a um mesmo $k$-subconjunto $\{i_1, i_2, \ldots, i_k\}$. Portanto para cada $k$ elementos de $X$ fixado, devemos contar uma única vez as $k$-palavras formadas por esses elementos. A contagem dessas $k$-palavras é feita como segue.

**Proposição 1.2.13.** *Dado $X$ um conjunto com $n$ elementos. O número de $k$-subconjuntos de $X$ é dado por,*

$$\binom{n}{k} = \frac{n \cdot (n-1) \cdot (n-2) \ldots (n-k+1)}{k!}$$

*Demonstração.* Sabemos, da Proposição (1.2.10), que o número de maneiras em que $k$ objetos podem ser escolhidos, em ordem, de um conjunto de $n$ objetos é dado por

$$n \cdot (n-1) \cdot (n-2) \ldots (n-k+1).$$

Como cada um dos $k$-subconjuntos de $X$ pode ser permutado de $k!$ maneiras segue que

$$k! \cdot \binom{n}{k} = n \cdot (n-1) \cdot (n-2) \ldots (n-k+1)$$

portanto,

$$\binom{n}{k} = \frac{n \cdot (n-1) \cdot (n-2) \ldots (n-k+1)}{k!}.$$

Um outro caminho para escrever isso é,

$$\binom{n}{k} = \frac{n!}{k! \cdot (n-k)!} \quad (1 \leq k \leq n), \tag{1.2.1}$$

com $\binom{n}{0} = 1$. Verificamos facilmente, a partir da igualdade (1.2.1) que,

$$\binom{n}{k} = \binom{n}{n-k}.$$

□

**Proposição 1.2.14** (O Teorema Binomial)**.** *Para todo inteiro positivo $n$, temos,*

$$(x+y)^n = x^n + \binom{n}{1} x^{n-1} y + \binom{n}{2} x^{n-2} y^2 + \cdots + \binom{n}{n-1} x y^{n-1} + y^n.$$

*ou equivalentemente*

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^k.$$

*Demonstração.* Considere o produto,

$$(x+y)(x+y)\dots(x+y)$$

com $n$ parcelas.

Quando multiplicamos o produto acima, cada termo separadamente surge da escolha de $x$ ou $y$ nos parênteses. Portanto obtemos o termo $x^{n-k}y^k$ se escolhermos $y$ em $k$ desses parênteses e $x$ nos $n-k$ parênteses restante. Assim, o número de termos da forma $x^{n-k}y^k$ que nós obtemos é igual ao numero de caminhos de escolher $k$ valores $y$ nos parênteses e este número é

$$\binom{n}{k}.$$

Portanto, quando reunimos termos similares o coeficiente de $x^{n-k}y^k$ é $\binom{n}{k}$. □

Vamos estabelecer um exemplo de aplicação da Proposição (1.2.14) que será útil no próximo capitulo.

**Exemplo 1.2.15.** *Claramente,*

$$1-\binom{m}{1}+\binom{m}{2}-\binom{m}{3}+\cdots+(-1)^m\binom{m}{m}=0.$$

*Obtemos o resultado acima tomando $x=1$ e $y=-1$ na Proposição (1.2.14).*

## 1.3 Fórmulas Assintóticas

Por fim, para concluir este capítulo, vamos introduzir algumas definições e alguns resultados que serão úteis no Capítulo 3. As principais referências bibliográficas adotadas para essa seção foram [3] e [2].

### 1.3.1 Notações e Resultados

**Definição 1.3.1** (notação "O"grande)**.** *Sejam $D$ um subconjunto não vazio de números reais e $f: D \to \mathbb{R}$. Diz-se que $f = O(g)$ se existem $g: D \to \mathbb{R}_+$ e uma constante positiva $M \in \mathbb{R}$ tal que*

$$|\, f(x)\,| \leq Mg(x)$$

*para todo $x \in D$.*

Essa importante notação, é usada para ajudar a entender o comportamento de funções $f(x)$ para valores grandes de $x$. Uma definição formal para essa notação, assim como uma discussão das ideias envolvidas em tal notação, podem ser encontradas em [3]. No entanto faremos uso da Definição (1.3.1), como estabelecida em [2].

Algumas consequências da Definição (1.3.1) estão estabelecidas abaixo.

**Lema 1.3.2.** *Sejam $D$ um subconjunto não vazio de números reais e $f : D \to \mathbb{R}$. Então temos que,*

1. *$f = O(f)$ (reflexiva);*
2. *$f = O(g)$ e $g = O(h) \Rightarrow f = O(h)$ (transitiva).*

*Demonstração.* 1. Supondo $f : D \to \mathbb{R}_+$, temos que $f = O(f)$ uma vez que para todo $M \geq 1$ temos $\mid f(x) \mid \leq Mf(x)$.

2. Suponha agora que $f = O(g)$ então existe uma constante positiva $M_1$ tal que

$$\mid f(x) \mid \leq M_1 g(x).$$

Analogamente, se $g = O(h)$, então existe uma constante positiva $M_2$ tal que

$$g(x) = \mid g(x) \mid \leq M_2 h(x).$$

Portanto,

$$\mid f(x) \mid \leq M_1 g(x) \quad \text{e} \quad g(x) \leq M_2 h(x) \Rightarrow \mid f(x) \mid \leq M_1 M_2 h(x)$$

para todo $x \in D$. Daí segue que $f = O(h)$.

□

A relação $O(\cdot)$ é uma pré ordem. Apesar de cumprir o Lema (1.3.2) essa relação não cumpre a propriedade anti-simétrica. De fato, $t = O(2t)$ e $2t = O(t)$ sem que $t$ e $2t$ sejam iguais.

**Lema 1.3.3.** *Sejam $D$ um subconjunto não vazio dos números reais e $f : D \to \mathbb{R}$. Temos que $f = O(1)$ se, e somente se, $f$ é limitada.*

*Demonstração.* De fato, pela Definição (1.3.1), temos que se $f = O(1)$ então existe uma constante positiva $M$ tal que

$$\mid f(x) \mid \leq M \cdot 1$$

para todo $x \in D$, portanto $f$ é limitada. Reciprocamente, se $f$ é limitada então existe uma constante positiva $M$ tal que

$$\mid f(x) \mid \leq M$$

para todo $x \in D$. Tomando $g : D \to \mathbb{R}_+$ por,

$$g(x) = 1 \text{ para todo } x \in D$$

segue que $f = O(1)$. □

**Proposição 1.3.4.** *Seja $D$ um subconjunto não vazio de números reais. Sejam também $f, g : D \to \mathbb{R}$ e $\lambda \in \mathbb{R}$ então,*

$$f = O(h) \quad e \quad g = O(h) \Rightarrow f + \lambda g = O(h).$$

*Demonstração.* Novamente, pela Definição (1.3.1), temos que se $f = O(h)$ e $g = O(h)$ existem constantes positivas $M_1$ e $M_2$ tais que

$$| f(x) | \leq M_1 h(x) \quad \text{e} \quad | g(x) | \leq M_2 h(x)$$

para todo $x \in D$.

Seja agora, a função $[f + \lambda g](x) = f(x) + \lambda g(x)$, então

$$| f(x) + \lambda g(x) | \leq | f(x) | + |\lambda| | g(x) | \leq M_1 h(x) + |\lambda| M_2 h(x) = (M_1 + |\lambda| M_2) h(x).$$

portanto $f + \lambda g = O(h)$. □

**Proposição 1.3.5.** *Seja $D$ um subconjunto não vazio de números reais. Sejam também as funções $f_k : D \to \mathbb{R}$ com $f_k = O(1)$ para todo $k = 1, \ldots, n$ então*

$$\sum_{k=1}^{n} f_k = O(1).$$

*Demonstração.* Novamente, pela Definição (1.3.1), temos que se $f_k = O(1)$ existe uma constante positiva $c_k$ tal que

$$| f_k(x) | \leq c_k \cdot 1$$

para todo $x \in D$.

Assim,

$$\left| \sum_{k=1}^{n} f_k(x) \right| \leq \sum_{k=1}^{n} | f_k(x) | \leq (c_1 + c_2 + \cdots + c_n) \leq n \max_k \{c_k\}.$$

Como para alguma constante $c$ temos $n \max_k \{c_k\} \leq c$ segue $\sum_{k=1}^{n} f_k(x) = O(1)$. □

**Proposição 1.3.6.** *Seja $D$ um subconjunto não vazio de números reais. Sejam também $f, g : D \to \mathbb{R}$, então,*

$$f = O(f_1) \quad e \quad g = O(g_1) \Rightarrow fg = O(f_1 g_1).$$

*Demonstração.* Pela Definição (1.3.1), temos que se $f = O(f_1)$ e $g = O(g_1)$ existem constantes positivas $M_1$ e $M_2$ tais que

$$| f(x) | \leq M_1 f_1(x) \quad \text{e} \quad | g(x) | \leq M_2 g_1(x)$$

para todo $x \in D$. Portanto,

$$| f(x) | | g(x) | \leq M_1 f_1(x) | g(x) | \leq M_1 f_1(x) M_2 g_1(x)$$

logo,

$$| f(x) g(x) | \leq M_1 M_2 f_1(x) g_1(x)$$

e daí $fg = O(f_1 g_1)$. □

**Proposição 1.3.7.** *Sejam $D$ um subconjunto não vazio de números reais e $f, g : D \to \mathbb{R}_+$. Então temos que,*

1. $c \cdot O(f) = O(f)$ *onde* $c$ *é uma constante;*

2. $O(O(f)) = O(f)$*;*

*Demonstração.* 1. Sejam $c$ uma constante e $h = O(f)$. Pela Definição (1.3.1), temos que se $h = O(f)$ existe uma constante positiva $C$ tal que

$$| h(x) | \leq Cf(x)$$

para todo $x \in D$. Portanto,

$$| c \cdot h(x) | = | c | \, | h(x) | \leq | c | \, Cf(x)$$

para todo $x \in D$. Portanto, $c \cdot h = O(f)$.

2. Sejam $F = O(g)$ e $g = O(f)$. Pela Definição (1.3.1), temos que se $F = O(g)$ existe uma constante positiva $C_1$ tal que

$$| F(x) | \leq C_1 g(x)$$

para todo $x \in D$. Analogamente, se $g = O(f)$ existe uma constante positiva $C_2$ tal que

$$g(x) = | g(x) | \leq C_2 f(x)$$

para todo $x \in D$. Portanto segue que

$$| F(x) | \leq C_1 g(x) \leq C_1 C_2 f(x)$$

e daí $F = O(f)$.

□

### 1.3.2 Técnica de Soma Parcial

**Proposição 1.3.8.** *Seja* $c_1, c_2, \ldots$ *uma sequência de números complexos e defina*

$$C(x) := \sum_{n_0 \leq n \leq x} c_n$$

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

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

*Demonstração.* Primeiramente, note que para $x$ no intervalo $k \leq x < k + 1$ temos,

$$\sum_{n_0 \leq n \leq x} c_n f(n) = \sum_{n_0 \leq n \leq k} c_n f(n). \tag{1.3.1}$$

Note também que,

$$\sum_{n_0\leq n\leq k} c_n f(n) = C(k)f(k) - \sum_{n_0\leq n\leq k-1} C(n)[f(n+1)-f(n)]. \tag{1.3.2}$$

De fato,

$$C(k)f(k) - \sum_{n_0\leq n\leq k-1} C(n)[f(n+1)-f(n)] =$$

$$= C(k)f(k) + \sum_{n_0\leq r\leq k-1} C(r)f(r) - \sum_{n_0\leq s\leq k-1} C(s)f(s+1) = \sum_{n_0\leq r\leq k} C(r)f(r) - \sum_{n_0\leq s\leq k-1} C(s)f(s+1).$$

Fazendo $r = s+1$ temos

$$\sum_{n_0\leq r\leq k} C(r)f(r) - \sum_{n_0\leq s\leq k-1} C(s)f(s+1) = \sum_{n_0\leq r\leq k} C(r)f(r) - \sum_{n_0\leq r-1\leq k-1} C(r-1)f(r) =$$

$$= \sum_{n_0\leq r\leq k} C(r)f(r) - \sum_{n_0+1\leq r\leq k} C(r-1)f(r) =$$

$$= C(n_0)f(n_0) + \sum_{n_0+1\leq r\leq k} C(r)f(r) - \sum_{n_0+1\leq r\leq k} C(r-1)f(r) = C(n_0)f(n_0) + \sum_{n_0+1\leq r\leq k} [C(r)-C(r-1)]f(r)$$

e como,

$$C(n_0) = c_{n_0}\,, \quad C(r) = \sum_{n_0\leq n\leq r} c_n = \sum_{n_0\leq n\leq r-1} c_n + c_r \Rightarrow C(r) - C(r-1) = c_r$$

temos portanto das igualdades (1.3.1) e (1.3.2) que,

$$\sum_{n_0\leq n\leq x} c_n f(n) = C(k)f(k) - \sum_{n_0\leq n\leq k-1} C(n)[f(n+1)-f(n)]. \tag{1.3.3}$$

Sabemos ainda que

$$\sum_{n_0\leq n\leq k-1} C(n)[f(n+1)-f(n)] = \sum_{n_0\leq n\leq k-1} C(n)\int_n^{n+1} f'(t)dt = \sum_{n_0\leq n\leq k-1} \int_n^{n+1} C(t)f'(t)dt$$

pois, $C(t) = C(n)$ $(n \leq t < n+1)$. Portanto,

$$\sum_{n_0\leq n\leq k-1} \int_n^{n+1} C(t)f'(t)dt = \int_{n_0}^{k} C(t)f'(t)dt = \int_{n_0}^{x} C(t)f'(t)dt - \int_k^x C(t)f'(t)dt.$$

Assim,

$$\sum_{n_0\leq n\leq k-1} C(n)[f(n+1)-f(n)] = \int_{n_0}^{x} C(t)f'(t)dt - \int_k^x C(t)f'(t)dt. \tag{1.3.4}$$

Novamente, usando o fato de que, se $k \leq x < k+1$ então $C(k) = C(x)$, segue que em particular se $t \in (k, x)$ então $C(t) = C(k)$, assim

$$\int_k^x C(t)f'(t)dt = C(k)\int_k^x f'(t)dt = C(k)f(x) - C(k)f(k) = C(x)f(x) - C(k)f(k).$$

Utilizando agora as igualdades (1.3.3) e (1.3.4) obtemos a fórmula para soma parcial de *Abel*,

$$\sum_{n_0\leq n\leq x} c_n f(n) = C(k)f(k) - \sum_{n_0\leq n\leq k-1} C(n)[f(n+1)-f(n)] =$$

$$= C(k)f(k) - \left(\int_{n_0}^{x} C(t)f'(t)dt - \int_{k}^{x} C(t)f'(t)dt\right) =$$

$$= C(k)f(k) - \left(\int_{n_0}^{x} C(t)f'(t)dt - C(x)f(x) + C(k)f(k)\right) = C(x)f(x) - \int_{n_0}^{x} C(t)f'(t)dt.$$

□

Para finalizar este capítulo, apresentamos uma aplicação da Proposição (1.3.8).

**Exemplo 1.3.9.** *Suponha que queiramos provar a conhecida estimativa,*

$$\sum_{p\leq z}\frac{1}{p} = \log\log z + O(1) \quad \textit{para} \quad z \geq e$$

*onde $p$ é primo.*

*Para tanto suponha*[3]*, o resultado já provado,*

$$\sum_{p\leq z}\frac{\log p}{p} = \log z + R(z) \tag{1.3.5}$$

*onde $R(z) = O(1)$.*

*Usando somas parciais com*

$$f(t) = \frac{1}{\log t} \quad e \quad c_n = \begin{cases} \frac{\log n}{n} \text{ se } n \text{ é primo,} \\ 0 \text{ caso contrário.} \end{cases}$$

*temos que $C(x) = \sum_{p\leq x}\frac{\log p}{p}$ e,*

$$\sum_{p\leq z}\frac{1}{p} = \sum_{p\leq z}\frac{\log p}{p}\frac{1}{\log p} = \frac{C(z)}{\log z} + \int_2^z \frac{C(u)}{u(\log u)^2}du.$$

*Usando a estimativa (1.3.5) temos, portanto,*

$$\sum_{p\leq z}\frac{1}{p} = 1 + \frac{O(1)}{\log z} + \left(\int_2^z \frac{\log u}{u(\log u)^2}du + \int_2^z \frac{R(u)}{u(\log u)^2}du\right).$$

*Sabendo que*

$$\int_2^z \frac{du}{u(\log u)} = \log\log u|_2^z$$

[3]Este resultado pode ser encontrado em [2] *Teorema 1.4.3* página 9.

*segue que*

$$\sum_{p\leq z}\frac{1}{p}=\log\log z+1-\log\log 2+\frac{O(1)}{\log z}+\int_2^z\frac{R(u)}{u(\log u)^2}du.$$

*Observe agora que*

$$\int_2^z\frac{R(u)}{u(\log u)^2}du=\int_2^\infty\frac{R(u)}{u(\log u)^2}du-\int_z^\infty\frac{R(u)}{u(\log u)^2}du$$

*onde a existência das integrais improprias são asseguradas pelo fato de que* $R(u)=O(1)$.

*Além disso*

$$\int_z^\infty\frac{R(u)}{u(\log u)^2}du=O\left(\int_z^\infty\frac{du}{u(\log u)^2}\right)=O\left(\frac{1}{\log z}\right).$$

*Assim,*

$$\sum_{p\leq z}\frac{1}{p}=\log\log z+A+O\left(\frac{1}{\log z}\right)$$

*onde* $A=1-\log\log 2+\int_2^\infty\frac{R(u)}{u(\log u)^2}du$.

*Portanto, segue que*

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

*uma vez que*

$$O\left(\frac{1}{\log z}\right)\leq k\left|\frac{1}{\log z}\right|$$

*para alguma constante positiva* $k$ *e para* $z\geq e$,

$$O\left(\frac{1}{\log z}\right)\leq k\left|\frac{1}{\log z}\right|\leq k.$$

# Capítulo 2

# O Principio da Inclusão Exclusão

Neste capítulo vamos estabelecer, *o principio da inclusão-exclusão*, uma importante ferramenta de contagem. Essa ferramenta vai tornar possível obter o número de elementos em uma união finita de conjuntos.

## 2.1 O principio da inclusão-exclusão

Vamos considerar um exemplo para ilustrar nosso problema.

**Exemplo 2.1.1.** *Considere um conjunto finito $X$ e três subconjuntos $A, B$ e $C$ e suponha que queiramos contar o número de elementos no conjunto,*

$$X \setminus (A \cup B \cup C).$$

*Claramente, se soubermos contar o numero de elementos na união $A \cup B \cup C$ então respondemos ao nosso problema fazendo,*

$$| X | - | A \cup B \cup C | .$$

*Para obter o número de elementos do conjunto*

$$A \cup B \cup C$$

*tomamos a soma*

$$| A | + | B | + | C | . \tag{2.1.1}$$

*Se $A, B$ e $C$ são dois a dois disjuntos temos a resposta imediatamente pelo principio aditivo. Caso os conjuntos $A, B$ e $C$ não sejam disjuntos temos uma sobre-contagem em (2.1.1) uma vez que os elementos que pertencem a $A \cap B$, $A \cap C$ e $B \cap C$ foram contados duas vezes na soma $| A | + | B | + | C |$. Portanto devemos subtrair $| A \cap B | + | A \cap C | + | B \cap C |$ do total (2.1.1). Agora a conta está correta, exceto para os elementos na interseção $A \cap B \cap C$ que foram adicionados três vezes, mas também subtraidos três vezes. Portanto devemos ter,*

$$| A \cup B \cup C | = | A | + | B | + | C | - | A \cap B | - | A \cap C | - | B \cap C | + | A \cap B \cap C |,$$

*e portanto,*

$$| X \setminus (A \cup B \cup C) | = | X | - | A | - | B | - | C | + | A \cap B | + | A \cap C | + | B \cap C | - | A \cap B \cap C | .$$

O resultado a seguir fornece o caminho para o caso geral.

**Teorema 2.1.2.** *Sejam $A_1, A_2, \ldots, A_n$ subconjuntos de $X$. Então,*

$$\left| X \setminus \bigcup_{i=1}^{n} A_i \right| = | X | - \sum_{i=1}^{n} | A_i | + \sum_{1 \le i < j \le n} | A_i \cap A_j | - \sum_{1 \le i < j < k \le n} | A_i \cap A_j \cap A_k | + \cdots + (-1)^n | A_1 \cap \ldots \cap A_n | . \tag{2.1.2}$$

*Demonstração.* Afim de tornar clara a expressão (2.1.2) considere a observação:

obs.: A desigualdade $1 \le i_1 < i_2 < \cdots < i_t \le n$ abaixo do símbolo de soma está indicando que a soma é feita sobre todas as escolhas de inteiros $i_1, \ldots, i_t$ que satisfazem estas desigualdades.

Para a demonstração verificamos quão frequente um elemento de $x \in X$ é contado em ambos os lados. Se $x \notin \bigcup_{i=1}^{n} A_i$, então ele é contado uma vez em cada lado. Suponha então que $x \in \bigcup_{i=1}^{n} A_i$, e mais precisamente, que $x$ está em exatamente $m$ $(1 \le m \le n)$ dos conjuntos $A_i$. A conta no lado esquerdo é 0. No lado direito, usando o Exemplo (1.2.15), contamos $x$ exatamente,

$$1 - \binom{m}{1} + \binom{m}{2} - \binom{m}{3} + \cdots + (-1)^m \binom{m}{m} = 0,$$

uma vez que $\binom{m}{k}$ é o número de maneiras de se escolher $k$ conjuntos dentre os $m$ conjuntos. □

Suponha um conjunto $X$, chamado universo, e um conjunto de propriedades $E = \{e_1, e_2, \ldots, e_n\}$ cujos elementos de $X$ podem possuir ou não. Seja ainda $A_i$ o subconjunto de $X$ cujos elementos desfrutam da propriedade $e_i$ (e possivelmente de outras). Então temos:
$| X \setminus \bigcup_{i=1}^{n} A_i |$ é o número de elementos de $X$ que não possuem propriedades em $E$ e
$A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_t} \subset X$ é o conjunto dos elementos de $X$ que possuem as propriedades $e_{i_1}, e_{i_2}, \ldots, e_{i_t}$ (e talvez outras). Usando a notação,

$$\begin{aligned} N_{\supseteq T} &:= | \{x \in X \mid x \text{ possui ao menos as propriedades } T\} |; \\ N_{=T} &:= | \{x \in X \mid x \text{ possui precisamente as propriedades } T\} |, \end{aligned} \tag{2.1.3}$$

para $T \subset E$, reescrevemos o Teorema (2.1.2) como:

**Principio da Inclusão-Exclusão**. Seja $X$ um conjunto e $E = \{e_1, e_2, \ldots, e_n\}$ um conjunto de propriedades. Então,

$$N_{=\emptyset} = \sum_{T \subseteq E} (-1)^{|T|} N_{\supseteq T} = \sum_{k=0}^{n} (-1)^k \sum_{T : |T| = k} N_{\supseteq T} \tag{2.1.4}$$

onde $N_{=\emptyset} = |X \setminus \bigcup_{i=1}^{n} A_i|$, $N_{\supseteq \emptyset} = | X |$ e $N_{\supseteq T} = \left| \bigcap_{e_i \in T} A_i \right|$.

Com o propósito de tornar clara a fórmula (2.1.4) vamos ilustra-la com exemplos.

**Exemplo 2.1.3.** *Vamos supor que queiramos obter todos os números positivos, menores do que ou iguais a* 100*, que não são múltiplos de* 2*,* 3 *ou* 5*. Claramente fazendo* $X = \{1, 2, \ldots, 100\}$ *e* $E = \{e_1, e_2, e_3\}$*, onde* $e_1$ *é a propriedade* $2 \mid x$ *para* $x \in X$*,* $e_2$ *é a propriedade* $3 \mid x$ *para* $x \in X$ *e* $e_3$ *é a propriedade* $5 \mid x$ *para* $x \in X$*, o que buscamos é o conjunto* $N_{=\emptyset}$*, ou seja, o número de elementos em* $X$ *que não atende as propriedades em* $E$*.*

*Note que os elementos do conjunto,*

$$\{\emptyset, \{e_1\}, \{e_2\}, \{e_3\}, \{e_1, e_2\}, \{e_1, e_3\}, \{e_2, e_3\}, \{e_1, e_2, e_3\}\}$$

*são conjuntos de propriedades e que um elemento em* $X$ *pode admitir tais propriedades ou não.*

*Um elemento* $x \in X$ *atende a propriedade* $\{e_1, e_2\}$ *se, e somente se,* $6 \mid x$ *para* $x \in X$*, atende a propriedade* $\{e_1, e_3\}$ *se, e somente se,* $10 \mid x$ *para* $x \in X$*, atende a propriedade* $\{e_2, e_3\}$ *se, e somente se,* $15 \mid x$ *para* $x \in X$ *e atende a propriedade* $\{e_1, e_2, e_3\}$ *se, e somente se,* $30 \mid x$ *para* $x \in X$*.*

*Como* $T \in \{\emptyset, \{e_1\}, \{e_2\}, \{e_3\}, \{e_1, e_2\}, \{e_1, e_3\}, \{e_2, e_3\}, \{e_1, e_2, e_3\}\}$ *segue que*

$$\begin{aligned} N_{=\emptyset} = (-1)^{|\emptyset|} N_{\supseteq\emptyset} + (-1)^{|\{e_1\}|} N_{\supseteq\{e_1\}} + (-1)^{|\{e_2\}|} N_{\supseteq\{e_2\}} + (-1)^{|\{e_3\}|} N_{\supseteq\{e_3\}} + (-1)^{|\{e_1,e_2\}|} N_{\supseteq\{e_1,e_2\}} + \\ + (-1)^{|\{e_1,e_3\}|} N_{\supseteq\{e_1,e_3\}} + (-1)^{|\{e_2,e_3\}|} N_{\supseteq\{e_2,e_3\}} + (-1)^{|\{e_1,e_2,e_3\}|} N_{\supseteq\{e_1,e_2,e_3\}}. \end{aligned}$$

*e isso obviamente poderia ser escrito como,*

$$\sum_{k=0}^{n} (-1)^k \sum_{T:|T|=k} N_{\supseteq T}.$$

*Sabendo agora que* $\left\lfloor \frac{n}{a} \right\rfloor$[1] *é o número de inteiros no conjunto* $\{1, 2, \ldots, n\}$ *que são divisíveis por* $a$*, temos que;*

$$\begin{aligned} N_{=\emptyset} = (-1)^0 N_{\supseteq\emptyset} + (-1)^1 \left( \left\lfloor \tfrac{100}{2} \right\rfloor + \left\lfloor \tfrac{100}{3} \right\rfloor + \left\lfloor \tfrac{100}{5} \right\rfloor \right) + \\ (-1)^2 \left( \left\lfloor \tfrac{100}{6} \right\rfloor + \left\lfloor \tfrac{100}{10} \right\rfloor + \left\lfloor \tfrac{100}{15} \right\rfloor \right) + (-1)^3 \left\lfloor \tfrac{100}{30} \right\rfloor. \end{aligned}$$

*logo,*

$$N_{=\emptyset} = 100 - (50 + 33 + 20) + (16 + 10 + 6) - 3 = 26$$

*Portanto existem* 26 *números menores do que ou iguais a 100 que não são múltiplos de* 2*,* 3 *ou* 5*, a saber* $\{1, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 49, 53, 59, 61, 67, 71, 73, 77, 79, 83, 89, 91, 97\}$.

Para o próximo exemplo, vamos estabelecer o cálculo da função de Euler,

$$\varphi(n) = \mid \{d \mid 1 \leq d \leq n, (d, n) = 1\} \mid .$$

[1]Para $x \geq 0$ $\lfloor x \rfloor$ é o mesmo que,

$$\lfloor x \rfloor = \mid \{n \in \mathbb{N} : n \leq x\} \mid .$$

**Exemplo 2.1.4.** *Suponha que $n = p_1^{a_1} \dots p_t^{a_t}$ é a decomposição de $n$ em fatores primos . Sejam $X = \{1, 2, \dots, n\}$ e $e_i$ a propriedade dada por $p_i$ divide $d$ para $d \in X$. Novamente queremos calcular $N_{=\emptyset}$, ($N_{=\emptyset} = \varphi(n)$). Como $\frac{n}{\prod_{e_i \in T} p_i}$ fornece o número de inteiros menores do que ou iguais a $n$ que são divisíveis por $\prod_{e_i \in T} p_i$, temos,*

$$N_{\supseteq T} = \frac{n}{\prod_{e_i \in T} p_i}.$$

*E daí, em analogia ao exemplo anterior podemos aplicar a fórmula*

$$N_{=\emptyset} = \sum_{T \subseteq E} (-1)^{|T|} N_{\supseteq T}$$

*para obter $\varphi(n)$.*

*Note que como nesse caso, $E = \{e_1, e_2, \dots, e_t\}$ e a soma se estende para todo $T \in \mathcal{P}(E)$, a expressão acima possui a priori $2^t$ parcelas*

$$\begin{aligned} N_{=\emptyset} = (-1)^{|\emptyset|} N_{\supseteq \emptyset} + (-1)^{|\{e_1\}|} N_{\supseteq \{e_1\}} + \dots + (-1)^{|\{e_t\}|} N_{\supseteq \{e_t\}} + \dots \\ + (-1)^{|\{e_1,e_2\}|} N_{\supseteq \{e_1,e_2\}} + (-1)^{|\{e_1,e_3\}|} N_{\supseteq \{e_1,e_3\}} + \dots + (-1)^{|\{e_1,e_2,\dots,e_t\}|} N_{\supseteq \{e_1,e_2,\dots,e_t\}}. \end{aligned}$$

*e utilizando o fato de que $N_{=\emptyset}$ pode ser escrito como,*

$$N_{=\emptyset} = \sum_{k=0}^{t} (-1)^k \sum_{T:|T|=k} N_{\supseteq T}.$$

*temos,*

$$\varphi(n) = n - \sum_{i \leq t} \frac{n}{p_i} + \sum_{i<j \leq t} \frac{n}{p_i p_j} - \dots + (-1)^t \frac{n}{p_i \dots p_t} =$$

$$= n \cdot \left(1 - \sum_{i \leq t} \frac{1}{p_i} + \sum_{i<j \leq t} \frac{1}{p_i p_j} - \dots + (-1)^t \frac{1}{p_i \dots p_t}\right)$$

*ou equivalentemente*

$$\varphi(n) = n \cdot \prod_{i=1}^{t} \left(1 - \frac{1}{p_i}\right).$$

*Numericamente, se $n = 825$ temos $n = 3 \times 5^2 \times 11$. Assim*

$$\varphi(825) = 825 \cdot \left(1 - \frac{1}{3}\right)\left(1 - \frac{1}{5}\right)\left(1 - \frac{1}{11}\right) = 825 \cdot \frac{2}{3} \cdot \frac{4}{5} \cdot \frac{10}{11} = 400.$$

*Ou seja, existem 400 inteiros positivos, menores do que ou iguais a 825 que são primos com 825.*

## 2.2 Uma generalização do principio da inclusão-exclusão

Até agora expressamos $N_{=\emptyset}$ em termos do número $N_{\supseteq T}$ como na fórmula (2.1.4) e, em um caminho análogo, é natural perguntarmos por uma fórmula mais geral expressando $N_{=A}$.

A partir da notação estabelecida em (2.1.3) deduzimos que

$$N_{\supseteq A} = \sum_{T \supseteq A} N_{=T}, \tag{2.2.1}$$

uma vez que, os elementos que possuem exatamente as propriedades $T$ possuem ao menos as propriedades $A$, já que $T \supseteq A$. Agora iremos provar a reciproca,

$$N_{=A} = \sum_{T \supseteq A} (-1)^{|T|-|A|} N_{\supseteq T}, \tag{2.2.2}$$

que reduz a fórmula (2.1.4) para o caso particular $A = \emptyset$. Nós provamos as relações entre (2.2.1) e (2.2.2) para conjuntos de funções arbitrários.

**Teorema 2.2.1.** *Sejam $E$ um conjunto finito e $f, g : \mathcal{P}(E) \to K$ funções em um corpo $K$ de característica 0. Então*

$$f(A) = \sum_{T \supseteq A} g(T) \quad (\forall\ A) \iff g(A) = \sum_{T \supseteq A} (-1)^{|T|-|A|} f(T) \quad (\forall\ A)$$

*Demonstração.* Assuma que

$$f(A) = \sum_{T \supseteq A} g(T) \quad (\forall\ A).$$

Dado $A$ temos que,

$$\sum_{T \supseteq A} (-1)^{|T|-|A|} f(T) = \sum_{T \supseteq A} (-1)^{|T|-|A|} \left( \sum_{U \supseteq T} g(U) \right).$$

Como, $U \supseteq T$, $T \supseteq A$ e estamos num corpo $K$ podemos reorganizar o lado direito da última igualdade como segue

$$\sum_{T \supseteq A} (-1)^{|T|-|A|} \left( \sum_{U \supseteq T} g(U) \right) = \sum_{U \supseteq T \supseteq A} (-1)^{|T|-|A|} \left( \sum_{U \supseteq T \supseteq A} g(U) \right) =$$

$$= \sum_{U \supseteq T \supseteq A} (-1)^{|T|-|A|} \left( \sum_{U \supseteq A} g(U) \right) = \sum_{U \supseteq T \supseteq A} \sum_{U \supseteq A} (-1)^{|T|-|A|} g(U) =$$

$$= \sum_{U \supseteq A} \sum_{U \supseteq T \supseteq A} (-1)^{|T|-|A|} g(U).$$

Observe agora que, como $A$ é dado e para cada $U$ a soma ocorre em $T$, podemos colocar em evidencia $g(U)$ o que nos fornece,

$$= \sum_{U \supseteq A} \sum_{U \supseteq T \supseteq A} (-1)^{|T|-|A|} g(U) = \sum_{U \supseteq A} \left( \sum_{U \supseteq T \supseteq A} (-1)^{|T|-|A|} \right) g(U).$$

Portanto,

$$\sum_{T\supseteq A}(-1)^{|T|-|A|}f(T)=\sum_{U\supseteq A}\left(\sum_{U\supseteq T\supseteq A}(-1)^{|T|-|A|}\right)g(U).$$

Se $\mid U\setminus A\mid=m$ então,

$$\sum_{U\supseteq T\supseteq A}(-1)^{|T|-|A|}=\sum_{k=0}^{m}(-1)^k\binom{m}{k}=\begin{cases}1 & \text{se } m=0\\ 0 & \text{se } m\neq 0\end{cases}.$$

De fato, como

$$\mid A\mid\leq\mid T\mid\leq\mid U\mid$$

segue que $0\leq\mid T\mid-\mid A\mid\leq\mid U\mid-\mid A\mid=m$. Assim, escrevendo $\mid T\mid-\mid A\mid=k$, temos que para cada $k$ no conjunto $\{0,1,\ldots,m\}$ existem $\binom{m}{k}$ conjuntos $T$ com cardinalidade $\mid A\mid+k$. Portanto, para cada $k$ estaremos somando a parcela $(-1)^k$ exatamente $\binom{m}{k}$ vezes.

Assim resulta que,

$$\sum_{T\supseteq A}(-1)^{|T|-|A|}f(T)=\sum_{U\supseteq A}\left(\sum_{U\supseteq T\supseteq A}(-1)^{|T|-|A|}\right)g(U)=g(A)\iff U=A.$$

Assuma agora que,

$$g(A)=\sum_{T\supseteq A}(-1)^{|T|-|A|}f(T)\quad(\forall\ A).$$

Dado $A$ temos que,

$$\sum_{T\supseteq A}g(T)=\sum_{T\supseteq A}\sum_{U\supseteq T}(-1)^{|U|-|T|}f(U).$$

Novamente como $U\supseteq T$, $T\supseteq A$ e estamos num corpo $K$, podemos reorganizar o lado direito da última igualdade,

$$\sum_{T\supseteq A}\sum_{U\supseteq T}(-1)^{|U|-|T|}f(U)=\sum_{U\supseteq T\supseteq A}(-1)^{|U|-|T|}f(U).$$

Sabendo que $A$ é dado e para cada $U$ a soma ocorre em $T$ podemos colocar em evidencia $f(U)$ o que nos fornece,

$$\sum_{U\supseteq T\supseteq A}(-1)^{|U|-|T|}f(U)=\left(\sum_{U\supseteq T\supseteq A}(-1)^{|U|-|T|}\right)f(U).$$

Seja $\mid U\setminus A\mid=m$. Então,

$$\mid U\mid=\mid A\mid+m$$

e como $T\supseteq A$,

$$\mid T\mid=\mid A\mid+k\qquad(0\leq k\leq m)$$

segue, $\mid U\mid-\mid T\mid=m-k$ onde $0\leq k\leq m$.

Assim,

$$\sum_{U\supseteq T\supseteq A}(-1)^{|U|-|T|}=\sum_{s=0}^{m}(-1)^s\binom{m}{s}=\begin{cases}1 & \text{se } m=0\\ 0 & \text{se } m\neq 0\end{cases}.$$

onde $s=m-k$.

De fato, note que cada $k$ especifica um $s$. Note ainda que para cada $k$ no conjunto $\{0,1,\ldots,m\}$ existem $\binom{m}{k}$ conjuntos $T$ com cardinalidade $\mid A\mid +k$. Sabendo que,

$$\binom{m}{k}=\binom{m}{m-k}=\binom{m}{s}$$

temos que para cada $k$ estaremos somando a parcela $(-1)^s$ exatamente $\binom{m}{s}$ vezes.

Assim resulta que,

$$\sum_{T\supseteq A}g(T)=\left(\sum_{U\supseteq T\supseteq A}(-1)^{|U|-|T|}\right)f(U)=f(A)\iff U=A.$$

□

A fórmula (2.2.2) é agora uma consequência imediata do Teorema (2.2.1) por considerar as funções $g(A)=N_{=A}$ e $f(A)=N_{\supseteq A}$.

Uma importante generalização do *principio da inclusão-exclusão* surge quando os elementos de $X$ são ponderados. Suponha que $\omega: X\to K$ é uma função peso, que nós estendemos para $\mathcal{P}(X)$, fazendo

$$\omega(A):=\sum_{x\in A}\omega(x)\qquad \forall A\in\mathcal{P}(X)$$

com $\omega(\emptyset)=0$. Agora para um conjunto de propriedades $E=\{e_1,e_2,\ldots,e_n\}$, defina,

$$\begin{aligned}W_{\supseteq T}&:=\sum\{\omega(x)\mid x \text{ possui ao menos as propriedades } T\};\\ W_{=T}&:=\sum\{\omega(x)\mid x \text{ possui precisamente as propriedades } T\},\end{aligned}\tag{2.2.3}$$

para $T\subset E$.

Observe que para a função peso constante,

$$\omega(x)=1$$

temos $W_{\supseteq T}=N_{\supseteq T}$ e $W_{=T}=N_{=T}$.

Observe também,

$$W_{\supseteq A}=\sum_{T\supseteq A}W_{=T},$$

uma vez que, a soma dos pesos dos elementos que possuem ao menos as propriedades de $A$ é igual a soma dos pesos dos elementos que possuem precisamente as propriedades $T$ uma vez que $T\supseteq A$. Assim, aplicando o Teorema (2.2.1) com $g(A)=W_{=A}$ e $f(A)=W_{\supseteq A}$ e em seguida fazendo $A=\emptyset$,

chegamos ao *principio geral da inclusão-exclusão*,

**Principio Geral da Inclusão-Exclusão**. Sejam $X$ um conjunto, $\omega : X \to K$ uma função peso e $E = \{e_1, e_2, \ldots, e_n\}$ um conjunto de propriedades. Então,

$$W_{=\emptyset} = \sum_{T \subseteq E} (-1)^{|T|} W_{\supseteq T} = \sum_{k=0}^{n} (-1)^k \sum_{T:|T|=k} W_{\supseteq T}. \tag{2.2.4}$$

# Capítulo 3

# O Método do crivo

Nosso objetivo, nesse capítulo, é o de explicar as ideias relacionadas ao método do Crivo. Para tanto, faremos uma contextualização histórica e descreveremos a evolução das ideias de um crivo, que definiremos ao longo deste capítulo, chamado *o crivo de Eratóstenes - Legendre*. Embora este seja o mais simples crivo dentro da Teoria de crivos, o crivo de Eratóstenes - Legendre, torna-se útil para entender como funcionam os métodos de crivo uma vez que os crivos mais sofisticados[1] são extensões de suas ideias básicas.

O capítulo termina com a apresentação de uma versão moderna do crivo de Eratóstenes - Legendre associado ao *truque de Rankin*, cuja demonstração pode ser encontrada em [2]. Segundo os autores de [2] essa nova perspectiva torna o crivo de Eratóstenes - Legendre tão poderoso quanto o crivo de Brun, sendo esse último um processo de crivo que foi introduzido em um artigo de 1915 pelo matemático *Viggo Brun* e que é considerado um dos pilares da Teoria moderna de crivo.

Os livros utilizados neste capítulo foram [2], [4], [5] e [7].

## 3.1 Os crivos de Eratóstenes e Legendre

### 3.1.1 O crivo de Eratóstenes

As ideias de crivo, em teoria de números, têm sua origem rastreada por volta do Século III a.C. com o mais antigo argumento reconhecível como um crivo, devido a Eratóstenes.

De acordo com relatos sobreviventes, parece que Eratóstenes estabeleceu o seguinte processo:

> *Partindo de uma lista de números ímpares ele excluiu* $3^2$ *e cada terceiro número posterior. Em seguida, ele excluiu* $5^2$ *e cada quinto número posterior, e assim por diante.*

O processo, descrito acima, conhecido como o crivo de Eratóstenes foi primeiro descrito no trabalho de Nicomedes (280 á 210 a.C.) intitulado *Introdução à Aritmética*.

[1] Estamos nos referindo à classe dos crivos combinatoriais.

Não está claro, a partir dos relatos, que o objetivo[2] do processo era obter os números primos.

Segundo os autores *Henryk Iwaniec* e *George Greaves* a consiêcia de que uma tabela de números primos (até, por exemplo, $N$) poderia ser gerada desta forma datam de muito mais tarde. *Iwaniec*, em suas notas *Sieve Methods* de 1996, nos diz que isso ocorreu no inicio Século XIII graças a contribuição[3] de *Leonardo Pisano*.

Conforme observado por *Leonardo Pisano*, em seu *Liber Abbaci*, se um número $n$ não é primo, então ele tem necessariamente um fator primo não superior a $\sqrt{n}$. A observação de *Pisano* pode ser traduzida em termos de um teorema.

**Teorema 3.1.1.** *Se $n \neq 1$ não é um número primo, então $n$ possui necessariamente um fator primo menor do que ou igual a $\sqrt{n}$.*

*Demonstração.* Supondo que $n$ é um número composto positivo, existe ao menos um par de inteiros $n_1$ e $n_2$ tais que $n = n_1 n_2$. Sem perda de generalidade suponha $n_1 \leq n_2$. Claramente devemos ter $n_1 \leq \sqrt{n}$, uma vez que se $n_1 > \sqrt{n}$ teríamos:

$$n = n_1 n_2 \geq n_1 n_1 > \sqrt{n}\sqrt{n} = n$$

o que é um absurdo.

Usando o teorema fundamental da aritmética temos que $n_1$ é primo ou $n_1$ é produto de primos. Se $n_1$ é primo acabou. Caso contrário $n_1 = \alpha p$ onde $p$ é primo e $1 < \alpha = \frac{n_1}{p}$. Como $\max\{\alpha,\ p\} < n_1 \leq \sqrt{n}$ segue que $p < \sqrt{n}$. □

O processo hoje conhecido como o "crivo de Eratóstenes", e que já era conhecido dessa maneira nos tempos de *Legendre*, depende dessa observação de *Leonardo Pisano*. Descreveremos a seguir esse processo:

O crivo começa com a lista de todos os números inteiros entre 1 e $N$. Em seguida apagamos, dessa lista, os múltiplos de todos os números primos $p$ entre 1 e $\sqrt{N}$. Os sobreviventes, deste processo de seleção, são: o número 1 e os números primos entre $\sqrt{N}$ e $N$. Podemos ilustrar isso como segue: Para $N = 48$, temos que os primos menores do que $\sqrt{48}$ são $2, 3$ e 5, assim,

$$1, \not{2}, \not{3}, \not{4}, \not{5}, \not{6}, 7, \not{8}, \not{9}, \not{10}, 11, \not{12}, 13, \not{14}, \not{15}, \not{16}, 17, \not{18}, 19, \not{20}, \not{21}, \not{22}, 23, \not{24}, \not{25}, \not{26},$$
$$\not{27}, \not{28}, 29, \not{30}, 31, \not{32}, \not{33}, \not{34}, \not{35}, \not{36}, 37, \not{38}, \not{39}, \not{40}, 41, \not{42}, 43, \not{44}, \not{45}, \not{46}, 47, \not{48}.$$

Observe que alguns números foram cancelados mais de uma vez. Na verdade, neste processo observamos que um número $n$ é cancelado $k$ vezes, se $k$ é o número de fatores primos de $n$, que não excedem $\sqrt{N}$.

[2]Um objetivo plausível para esse processo, por exemplo, poderia ter sido a geração de uma tabela de fatores. Para este fim Eratóstenes gostaria de observar não só que número tinha sido excluído, mas de que primos era múltiplo que causou possivelmente sua eliminação múltipla.

[3]Trabalho *Liber Abbaci* de 1202 de autoria de *Leonardo Pisano*

### 3.1.2 O crivo de Legendre

Embora os problemas de obter uma tabela de primos e o de contar o número de primos menores do que um dado limite sejam diferentes, Legendre nos mostrou como obter informações sobre a quantidade de primos a partir do crivo de Eratóstenes. Em 1808, na segunda edição de seu livro *Théorie des Nombres*, *A. M. Legendre* estabeleceu uma formulação mais analítica do principio dado na Seção (3.1.1). Vamos agora estabelecer suas ideias:

Sejam $\mathscr{A}$ o conjunto dos números naturais menores do que ou igual a $x$, $\mathfrak{B}$ o conjunto dos números primos menores do que ou igual a $z$ e $\mathscr{A}_p$ o subconjunto de números em $\mathscr{A}$ que são divisíveis por $p \in \mathfrak{B}$. Seja ainda, para cada subconjunto $T \subset \mathfrak{B}$, o conjunto

$$\mathscr{A}_T := \bigcap_{p \in T} \mathscr{A}_p.$$

Temos que um elemento no conjunto $\mathscr{A}_T$ possui todas as propriedades comuns à $\mathscr{A}_p$ (com $p \in T$), ou seja, um elemento em $\mathscr{A}_T$ é divisível (ao menos) por todos os primos $p \in T$.

Temos também, denotando,

$$\mathcal{S}(\mathscr{A}, \mathfrak{B}) := \mathscr{A} \setminus \bigcup_{p \in \mathfrak{B}} \mathscr{A}_p,$$

que o conjunto $\mathcal{S}(\mathscr{A}, \mathfrak{B})$ representa os inteiros entre 1 e $x$ que não têm fatores primos menores do que $z$.

Neste contexto, usando a notação (2.1.3) introduzida no capítulo anterior, temos pelo *principio da inclusão-exclusão* que

$$| \mathcal{S}(\mathscr{A}, \mathfrak{B}) | = \sum_{T \subseteq \mathfrak{B}} (-1)^{|T|} N_{\supseteq T} \tag{3.1.1}$$

onde $N_{\supseteq T}$ denota o número de elementos em $\mathscr{A}$ que possuem ao menos as propriedades $T$. Ou seja se $T = \{p_{i_1}, p_{i_2}, \ldots, p_{i_r}\}$ então $N_{\supseteq T}$ denota o número de elementos menores do que ou igual a $x$ que são divisíveis por $p_{i_1} p_{i_2} \ldots p_{i_r}$.

Vamos agora escrever (3.1.1) de uma forma alternativa.

**Definição 3.1.2.** [4] *A função de Möbius denotada por $\mu(\cdot)$, é definida como uma função multiplicativa*[5] *satisfazendo $\mu(1) = 1$, $\mu(p) = -1$ para todo primo $p$ e $\mu(p^\alpha) = 0$ para todo inteiro $\alpha \geq 2$. Assim se $n$ não é livre de quadrados $\mu(n) = 0$ e se $n$ é um produto de $k$ primos distintos, então $\mu(n) = (-1)^k$.*

Utilizando a Definição (3.1.2) é possível ver que o lado direito da igualdade (3.1.1) pode ser reescrito como,

[4] Essa função foi introduzida em 1832 pelo matemático *A. F. Möbius* e por essa razão ela é hoje conhecida como função de Möbius. A função de Möbius definida dessa maneira pode ser encontrada em [2].

[5] Função multiplicativa é uma função definida para todos os inteiros positivos e tal que $f(mn) = f(m)f(n)$ para todo par de inteiros positivos $m$ e $n$ relativamente primos.

$$\sum_{d|P} \mu(d) \left\lfloor \frac{x}{d} \right\rfloor$$

onde $P$ denota o produto dos primos em $\mathfrak{B}$ e $\lfloor x \rfloor$ denota a parte inteira de $x$, isto é,

$$\mid \mathcal{S}(\mathscr{A}, \mathfrak{B}) \mid = \sum_{d|P} \mu(d) \left\lfloor \frac{x}{d} \right\rfloor . \tag{3.1.2}$$

De fato, sabemos que $\mathfrak{B}$ é o conjunto de todos os números primos menores do que ou igual a $z$. Desde que os divisores de $P$ estão em correspondência um a um com os subconjuntos $T \subset \mathfrak{B}$,

$$p_1 p_2 \dots p_k \iff T = \{p_1, p_2, \dots p_k\} \subset \mathfrak{B}^6 \tag{3.1.3}$$

segue que, a soma

$$\sum_{d|P} \mu(d) \left\lfloor \frac{x}{d} \right\rfloor$$

contêm $2^{|\mathfrak{B}|}$ parcelas e portanto o número de parcelas nas duas somas são iguais. Alem disso dado um divisor $d$ de $P$, associamos esse divisor ao subconjunto $T$ de $\mathfrak{B}$, como em 3.1.3, de modo que a parcela $\mu(d) \left\lfloor \frac{x}{d} \right\rfloor = (-1)^{|T|} N_{\supseteq T}$. Daí segue que

$$\sum_{d|P} \mu(d) \left\lfloor \frac{x}{d} \right\rfloor = \sum_{T \subseteq \mathfrak{B}} (-1)^{|T|} N_{\supseteq T}.$$

Como *Legendre*, baseado neste principio, publicou a fórmula

$$\mid \mathcal{S}(\mathscr{A}, \mathfrak{B}) \mid = \lfloor x \rfloor - \sum_{p_1 \leq z} \left\lfloor \frac{x}{p_1} \right\rfloor + \sum_{p_2 < p_1 \leq z} \left\lfloor \frac{x}{p_1 p_2} \right\rfloor - \sum_{p_3 < p_2 < p_1 \leq z} \left\lfloor \frac{x}{p_1 p_2 p_3} \right\rfloor + \dots \tag{3.1.4}$$

a identidade (3.1.2) é chamada de *identidade de Legendre* .

Para verificar diretamente a fórmula (3.1.4) observe que $\left\lfloor \frac{x}{p_1 p_2 \dots p_r} \right\rfloor$ é o número de inteiros positivos menores do que ou iguais a $x$ que são múltiplos do inteiro $p_1 p_2 \dots p_r$. Portanto, o lado direito da igualdade (3.1.4) representa o número de inteiros positivos menores do que ou igual a $x$ contados com multiplicidade apropriada. A saber, se $n \leq x$ tem exatamente $s \geq 0$ divisores primos menores do que ou iguais a $z$ ele é contado com multiplicidade,

$$\sum_{k=0}^{s} (-1)^k \binom{s}{k} = \begin{cases} 1, \text{ se } s = 0; \\ 0, \text{ se } s \neq 0 \end{cases} .$$

Ilustramos com um exemplo numérico a igualdade (3.1.2) e a relação disto com o crivo de Eratóstenes. A relação entre os crivos de Legendre e Eratóstenes é o motivo pelo qual o crivo de Legrendre é também chamado de *o crivo de Eratóstenes-Legendre.*

[6] O subconjunto $\emptyset$ esta em correspondencia com o divisor 1.

**Exemplo 3.1.3.** *Ao longo do texto, a função $\pi(x)$ representará o número de primos não superiores a $x$.*

*Suponha que queiramos contar o número de primos menores do que* 100.

*Primeiramente, conforme observado por Leonardo Pisano, para obter a lista de todos os primos $p$ no intervalo $\sqrt{x} < p \leq x$ somente precisamos remover, partindo dos inteiros $1 \leq n \leq x$, os múltiplos de primos menores do que ou iguais a $\sqrt{x}$. No nosso caso, para $x = 100$ devemos retirar do intervalo $1 \leq n \leq 100$ os múltiplos de primos menores do que ou igual a $\sqrt{100} = 10$, a saber os múltiplos de $\{2, 3, 5, 7\}$.*

*Assim, a quantidade de primos no intervalo $1 \leq p \leq 100$ utilizando o principio da inclusão-exclusão fornece*

$$\begin{aligned}
\pi(100) - \pi(10) + 1 &= (-1)^0 N_{\geq \emptyset} + (-1)^1 \left\lfloor \frac{100}{2} \right\rfloor + (-1)^1 \left\lfloor \frac{100}{3} \right\rfloor + (-1)^1 \left\lfloor \frac{100}{5} \right\rfloor + (-1)^1 \left\lfloor \frac{100}{7} \right\rfloor + \\
&\quad + (-1)^2 \left\lfloor \frac{100}{6} \right\rfloor + (-1)^2 \left\lfloor \frac{100}{10} \right\rfloor + (-1)^2 \left\lfloor \frac{100}{14} \right\rfloor + (-1)^2 \left\lfloor \frac{100}{15} \right\rfloor + (-1)^2 \left\lfloor \frac{100}{21} \right\rfloor \\
&\quad + (-1)^2 \left\lfloor \frac{100}{35} \right\rfloor + (-1)^3 \left\lfloor \frac{100}{30} \right\rfloor + (-1)^3 \left\lfloor \frac{100}{42} \right\rfloor + (-1)^3 \left\lfloor \frac{100}{70} \right\rfloor \\
&\quad + (-1)^3 \left\lfloor \frac{100}{105} \right\rfloor + (-1)^4 \left\lfloor \frac{100}{210} \right\rfloor
\end{aligned}$$

*ou seja,*

$$\begin{aligned}
\pi(100) - 4 + 1 &= 100 - 50 - 33 - 20 - 14 + \\
&\quad + 16 + 10 + 7 + 6 + 4 \\
&\quad + 2 - 3 - 2 - 1 \\
&\quad - 0 + 0 = 22.
\end{aligned}$$

*Portanto o número de inteiros positivos que não são múltiplos de $2, 3, 5$ ou $7$ no intervalo $1 \leq p \leq 100$ é igual a* 22. *Observe que o inteiro positivo* 1 *não é múltiplo de $2, 3, 5$ e $7$ e portanto ele foi contado acima. Além disso como estamos retirando os múltiplos dos primos menores do que* 10, *segue que o número de inteiros positivos que não são múltiplos de $2, 3, 5$ e $7$, estão no intervalo no intervalo $10 < p \leq 100$ e são primos, são em quantidade igual a* 21.

*A relação entre essa contagem e o crivo de Eratóstenes pode ser ilustrado com a ajuda do Exemplo (2.1.3), dado no Capítulo 2.*

*Para obter os números primos no intervalo $10 < p \leq 100$ devemos retirar os múltiplos dos números primos menores do que ou igual a $\sqrt{100} = 10$ do intervalo $1 \leq n \leq 100$. No Exemplo (2.1.3) retiramos os múltiplos de $2, 3$ e $5$, do intervalo $1 \leq n \leq 100$, e obtivemos o conjunto*

$$\{1, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 49, 53, 59, 61, 67, 71, 73, 77, 79, 83, 89, 91, 97\}.$$

*Assim, para obter os primos menores do que ou igual a* 100 *e maiores do que ou igual a* 10 *devemos retirar desse conjunto os múltiplos do número primo* 7*, o que resulta*

$$\{1, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 53, 59, 61, 67, 71, 73, 79, 83, 89, 97\}.$$

*Por fim devemos retirar* 1 *que não é primo para obtermos,*

$$|\ \{11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 53, 59, 61, 67, 71, 73, 79, 83, 89, 97\}\ | = 21.$$

Para finalizar a seção apresentamos um resultado, que é uma consequência da Definição (3.1.2) e que será útil futuramente.

**Lema 3.1.4** (A propriedade fundamental da função de Möbius)**.**

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

*Demonstração.* Se $n = 1$, então a afirmação do Lema é claramente verdadeira. Sejam $n > 1$ com $n = p_1^{\alpha_1} \cdot p_2^{\alpha_2} \dots p_r^{\alpha_r}$, a fatoração única de $n$ em potências de primos distintos, e $N = p_1 \cdot p_2 \dots p_r$. Como $\mu(d) = 0$, quando $d$ não é livre de quadrados, temos

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

A última soma contém $2^r$ parcelas, cada uma correspondendo a um subconjunto de $\{p_1, p_2, \dots, p_r\}$. Como os divisores de $N$ estão em correspondência um a um com tais subconjuntos, e o número de subconjuntos com $k$ elementos é

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

segue que para um divisor $d$ determinado por um subconjunto com $k$ elementos temos $\mu(d) = (-1)^k$. Assim,

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

□

### 3.1.3 Uma estimativa para $\pi(x)$ utilizando o crivo de Legendre

Definimos no Exemplo (3.1.3) a função $\pi(x)$ como sendo o número de primos menores do que ou iguais a $x$. Vamos agora discutir a questão de limitar o número $\pi(x)$ usando uma abordagem dada por *Legendre*. Fixada a notação da seção anterior e, como antes, denotando

$$\mathcal{S}(\mathscr{A}, \mathfrak{B}) := \mathscr{A} \setminus \bigcup_{p \in \mathfrak{B}} \mathscr{A}_p$$

com $z < \sqrt{x}$ segue,

$$| \mathcal{S}(\mathscr{A}, \mathfrak{B}) | \geq \pi(x) - \pi(z) + 1. \tag{3.1.5}$$

De fato, isso ocorre pois, $| \mathcal{S}(\mathscr{A}, \mathfrak{B}) |$ conta o número 1, não conta os primos $p \leq z$, conta os primos $p$ com $z < p \leq x$ e além disso, o número $| \mathcal{S}(\mathscr{A}, \mathfrak{B}) |$, conta os números compostos que são produto de primos maiores do que $z$.

Por outro lado, vimos em (3.1.2) que,

$$| \mathcal{S}(\mathscr{A}, \mathfrak{B}) | = \sum_{d|P} \mu(d) \left\lfloor \frac{x}{d} \right\rfloor.$$

Representando a parte fracionária de $x$ por $\{x\}$ temos que

$$\lfloor x \rfloor = x - \{x\},$$

assim, segue de (3.1.2) que,

$$| \mathcal{S}(\mathscr{A}, \mathfrak{B}) | = \sum_{d|P} \mu(d) \left\lfloor \frac{x}{d} \right\rfloor = \sum_{d|P} \mu(d) \left( \frac{x}{d} - \left\{ \frac{x}{d} \right\} \right) = x \sum_{d|P} \frac{\mu(d)}{d} + \sum_{d|P} \mu(d) \left( - \left\{ \frac{x}{d} \right\} \right). \tag{3.1.6}$$

Mas,

$$\left| \sum_{d|P} \mu(d) \left( - \left\{ \frac{x}{d} \right\} \right) \right| \leq \sum_{d|P} 1 = 2^{\pi(z)} \Rightarrow \sum_{d|P} \mu(d) \left( - \left\{ \frac{x}{d} \right\} \right) = O(2^{\pi(z)}).$$

Além disso, como $\mu(\cdot)$ é uma função multiplicativa

$$\sum_{d|P} \frac{\mu(d)}{d} = \prod_{p<z} \left( 1 + \frac{\mu(p)}{p} \right) = \prod_{p<z} \left( 1 - \frac{1}{p} \right).$$

Portanto de (3.1.6) devemos ter que

$$| \mathcal{S}(\mathscr{A}, \mathfrak{B}) | = x \prod_{p<z} \left( 1 - \frac{1}{p} \right) + O(2^{\pi(z)}). \tag{3.1.7}$$

Vamos agora obter uma estimativa para o produto

$$\prod_{p<z} \left( 1 - \frac{1}{p} \right). \tag{3.1.8}$$

Sabemos que a desigualdade

$$1 - x \leq e^{-x} \tag{3.1.9}$$

é valida para todo $x \in \mathbb{R}$.

De fato, a equação da reta tangente ao gráfico $g(x) = e^{-x}$ no ponto $(0, 1)$ é $y = 1 - x$. Como a função $g(x)$ tem concavidade para cima, pois $g''(x) > 0$, segue que o gráfico está sempre acima desta reta tangente. Logo

$$e^{-x} \geq 1 - x \ \forall x \in \mathbb{R},$$

com a igualdade apenas para $x = 0$.

Assim, tomando $x = \frac{1}{p}$ e multiplicando temos

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

No Exemplo (1.3.9), do Capitulo 1, obtivemos

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

Portanto um limite superior para (3.1.8) é dado por

$$\frac{1}{c \log z}$$

para uma constante $c$ suficientemente pequena. De fato,

$$\prod_{p<z} \left(1 - \frac{1}{p}\right) \leq \exp\left(-\sum_{p<z} \frac{1}{p}\right) = \exp\left(-\log\log z - O(1)\right) =$$

$$\exp\left(\log(\log z)^{-1}\right) \exp(-O(1)) = \frac{1}{\log z} \frac{1}{e^{O(1)}}$$

e para $0 < c < e^{O(1)}$ (com $\log z > 0$ )segue

$$\prod_{p<z} \left(1 - \frac{1}{p}\right) \leq \frac{1}{e^{O(1)} \log z} < \frac{1}{c \log z}. \tag{3.1.10}$$

Assim, usando (3.1.5), (3.1.7) e (3.1.10), segue que

$$\pi(x) - \pi(z) + 1 \leq \mid \mathcal{S}(\mathscr{A}, \mathfrak{B}) \mid < \frac{x}{c \log z} + O(2^{\pi(z)})$$

ou seja,

$$\pi(x) < \frac{x}{c \log z} + O(2^{\pi(z)}) + \pi(z) - 1$$

e, como $\pi(z) - 1 < z$, segue que

$$\pi(x) < \frac{x}{c \log z} + O(2^{\pi(z)}) + z.$$

Observe agora, fazendo $z = \log x$, que

$$2^{\pi(z)} < 2^{\log x} = x^{\log 2}.$$

De fato, como $\pi(z) < z$, temos $2^{\pi(z)} < 2^z$, além disso

$$2^{\log x} = \left(e^{\log 2}\right)^{\log x} = e^{\log 2 \cdot \log x} = \left(e^{\log x}\right)^{\log 2} = x^{\log 2}.$$

Assim, para alguma constante positiva $k$,

$$O(2^{\pi(z)}) < kx^{\log 2}$$

e daí,

$$\pi(x) < \frac{x}{c \log\log x} + kx^{\log 2} + \log x.$$

Por fim, utilizando o fato de que,

$$\frac{x}{c \log\log x}$$

cresce mais rapidamente do que $kx^{\log 2}$ e $\log x$, concluímos que

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

Com efeito, observe que

$$0 < \frac{kx^{\log 2}}{\frac{x}{c \log\log x}} = kc \frac{\log\log x}{x^{1-\log 2}} \qquad \text{para} \quad x > e.$$

Como a função log é monótona e para $y > 0$ temos $\log y < y$ segue,

$$0 < kc \frac{\log\log x}{x^{1-\log 2}} < kc \frac{\log x}{x^{1-\log 2}}. \tag{3.1.11}$$

Portanto, aplicando *L'Hôpital* no lado direito da desigualdade (3.1.11), temos

$$kc \frac{(\log x)'}{(x^{1-\log 2})'} = kc \frac{\frac{1}{x}}{\frac{1-\log 2}{x^{\log 2}}} = \frac{kc}{(1-\log 2)} \frac{1}{x^{1-\log 2}} \to 0 \quad \text{quando} \quad x \to \infty$$

portanto

$$\frac{kx^{\log 2}}{\frac{x}{c \log\log x}} \to 0 \quad \text{quando} \quad x \to \infty.$$

Por outro lado temos,

$$0 < \frac{\log x}{\frac{x}{c \log\log x}} = \frac{(\log x)(c \log\log x)}{x} \qquad \text{para} \quad x > e$$

e, pelo mesmo argumento acima, temos

$$0 < \frac{(\log x)(c \log\log x)}{x} < c\frac{(\log x)^2}{x}. \tag{3.1.12}$$

Portanto, aplicando *L'Hôpital*, agora no lado direito da desigualdade (3.1.12), temos

$$c\frac{((\log x)^2)'}{x'} = 2c\frac{\log x}{x}.$$

Aplicando mais uma vez *L'Hôpital*

$$2c\frac{(\log x)'}{x'} = 2c\frac{1}{x} \to 0 \quad \text{quando} \quad x \to \infty$$

e portanto

$$\frac{\log x}{\frac{x}{c \log\log x}} \to 0 \quad \text{quando} \quad x \to \infty$$

e daí segue (para $x \geq e$)

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

## 3.2 Os Problemas de Crivo

Qualquer crivo em teoria de números é baseado na seguinte ideia:

Sejam $\mathscr{A}$ uma sequência finita de inteiros, $\mathfrak{B}$ um conjunto de primos e $\Omega_p$ um conjunto de classes de resíduos indesejáveis módulo $p$ (com $p \in \mathfrak{B}$). Queremos saber quando o conjunto dos elementos da sequência $\mathscr{A}$ peneirada, não é vazio e quantos elementos sobraram nessa sequência. Em outras palavras, o problema do crivo é estimar, a quantidade de elementos do conjunto

$$(\mathscr{A}, \mathfrak{B}, \Omega) = \{n \in \mathscr{A} \mid n \mod (p) \notin \Omega_p \text{ para qualquer } p \in \mathfrak{B}\}. \tag{3.2.1}$$

Vejamos dois exemplos[7], importantes, de problemas em linguagem de crivo,

1. *conjectura dos primos gêmeos*
   Existe uma infinidade de primos $p$ tal que $p + 2$ é um primo;

2. *conjectura de Goldbach*
   Todo número par maior do que 2 é uma soma de dois primos.

**Exemplo 3.2.1** (Conjectura dos Primos Gêmeos)**.** *Considere a sequência*

$$\mathscr{A} = \{n \cdot (n+2) \mid z \leq n < z^2 - 2\},$$

[7]Existe uma grande variedade de problemas que podem ser expressos em linguagem de teoria de crivos.

*e o conjunto $\mathfrak{B}$ de números primos menores do que ou iguais a $z$. Se retirarmos os elementos da sequência $\mathscr{A}$ que deixam resto zero na divisão por $p$ ($p \in \mathfrak{B}$) temos que os elementos sobrevivem ao processo se, e somente se, $n$ e $n+2$ são ambos números primos.*

*De fato, se $n \cdot (n+2)$ têm somente fatores primos maiores do que ou iguais a $z$ então sendo $d$ um fator primo qualquer de $n \cdot (n+2)$ devemos ter que,*

$$d \mid n \cdot (n+2) \iff d \mid n \text{ ou } d \mid n+2.$$

*Portanto os divisores primos de $n$ e $n+2$ também devem ser maiores do que ou iguais a $z$. Como $z < n < z^2$ e $z < n+2 < z^2$, segue do Teorema (3.1.1) que, nem $n$ e nem $n+2$ podem ser números compostos.*

*Numericamente, seja $z = 11$, então,*

$$\mathscr{A} = \{n \cdot (n+2) \mid 11 \leq n < 119\} \quad e \quad \mathfrak{B} = \{2, 3, 5, 7, 11\}.$$

*Assim,*

$$\begin{aligned}\mathscr{A} = \{&143, 168, 195, 224, 255, 288, 323, 360, 399, 440, 483, 528, 575, 624, 675, 728, 783, 840, 899, 960,\\ &1023, 1088, 1155, 1224, 1295, 1368, 1443, 1520, 1599, 1680, 1763, 1848, 1935, 2024, 2115, 2208,\\ &2303, 2400, 2499, 2600, 2703, 2808, 2915, 3024, 3135, 3248, 3363, 3480, 3599, 3720, 3843, 3968,\\ &4095, 4224, 4355, 4488, 4623, 4760, 4899, 5040, 5183, 5328, 5475, 5624, 5775, 5928, 6083, 6240,\\ &6399, 6560, 6723, 6888, 7055, 7224, 7395, 7568, 7743, 7920, 8099, 8280, 8463, 8648, 8835, 9024,\\ &9215, 9408, 9603, 9800, 9999, 10200, 10403, 10608, 10815, 11024, 11235, 11448, 11663, 11880,\\ &12099, 12320, 12543, 12768, 12995, 13224, 13455, 13688, 13923, 14160, 14399\}.\end{aligned}$$

*Quando efetuamos o processo de crivo na sequência acima, isto é, quando retiramos os elementos da sequência $\mathscr{A}$ que deixam resto zero na divisão por $p \in \mathfrak{B}$, obtemos o conjunto*

$$\{323, 899, 1763, 3599, 5183, 10403, 11663\} = \{17 \times 19, 29 \times 31, 41 \times 43, 59 \times 61, 71 \times 73, 101 \times 103, 107 \times 109\}.$$

*Facilmente vemos que o número $(p \times (p+2))$ têm como fatores os primos gêmeos $p$ e $p+2$.*

**Exemplo 3.2.2** (Conjectura de Goldbach)**.** *Sejam $N$ um inteiro par ($N > 2$) e*

$$\mathscr{A} = \{n \cdot (N-n) \mid 2 < n < N \text{ e } (n, N) = 1\}.$$

*Seja também o conjunto $\mathfrak{B}$ formado por todos os números primos menores do que $\sqrt{N}$. Pelo mesmo argumento apresentado no Exemplo (3.2.1), se $n \cdot (N-n)$ tem somente divisores primos maiores do que $\sqrt{N}$ então necessariamente $n$ e $N-n$ têm seus divisores primos maiores do que $\sqrt{N}$. Como $\max\{n, N-n\} < N$ segue que $\sqrt{N} < n < N$ e $\sqrt{N} < N-n < N$ portanto segue do Teorema (3.1.1), juntamente com o fato de que $(n, N) = 1$, que nem $n$ e nem $N-n$ podem ser números compostos.*

*Numericamente, seja $N = 300$, então $\mathscr{A} = \{n \cdot (300-n) \mid 2 < n < 300 \text{ e } (n, 300) = 1\}$ com $\mathfrak{B} = \{2, 3, 5, 7, 11, 13, 17\}$.*

*Os elementos $n$ no intervalo $2 < n < 300$ tais que $(n, 300) = 1$ são dados pelo conjunto,*

$$\{7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 49, 53, 59, 61, 67, 71, 73, 77, 79, 83, 89, 91, 97, 101, 103, 107, 109, 113, 119, 121, 127, 131, 133, 137, 139, 143, 149, 151, 157, 161, 163, 167, 169, 173, 179, 181, 187, 191, 193, 197, 199, 203, 209, 211, 217, 221, 223, 227, 229, 233, 239, 241, 247, 251, 253, 257, 259, 263, 269, 271, 277, 281, 283, 287, 289, 293, 299\}.$$

*Veja que quando varremos os elemento do conjunto acima, geramos repetições de elementos no conjunto $\mathscr{A}$. Por exemplo, quando pegamos $n = 7$ e $n = 293$ no conjunto acima, geramos o mesmo elemento* 2051 *no conjunto $\mathscr{A}$. Portanto, retiradas as repetições, o conjunto $\mathscr{A}$ é dado por,*

$$\begin{aligned}\mathscr{A} = \{&299, 2051, 3179, 3731, 4811, 5339, 6371, 7859, 8339, 9731, 10619, 11051, 11891, 12299,\\ &13091, 14219, 14579, 15611, 16259, 16571, 17171, 17459, 18011, 18779, 19019, 19691,\\ &20099, 20291, 20651, 20819, 21131, 21539, 21659, 21971, 22139, 22211, 22331, 22379,\\ &22451, 22499\}.\end{aligned}$$

*Quando efetuamos o processo de crivo na sequência, isto é quando retiramos os elementos da sequência $\mathscr{A}$ que deixam resto zero na divisão por $p \in \mathfrak{B}$, obtemos o conjunto*

$$\begin{aligned}&\{4811, 5339, 6371, 7859, 8339, 9731, 11051, 14219, 14579, 15611, 16259, 16571, 18779, 20099,\\ &20291, 20651, 20819, 21971, 22331, 22499\} =\\ &= \{17 \times 283, 19 \times 281, 23 \times 277, 29 \times 271, 31 \times 269, 37 \times 263, 43 \times 257, 59 \times 241, 61 \times 239,\\ &67 \times 233, 71 \times 229, 73 \times 227, 89 \times 211, 101 \times 199, 103 \times 197, 107 \times 193, 109 \times 191, 127 \times 173,\\ &137 \times 163, 149 \times 151\}.\end{aligned}$$

*Facilmente vemos que o número $(p \times (N - p))$ têm como fatores os números primos $p$ e $N - p$.*

Mais uma vez reafirmamos que os problemas de obter uma tabela de números atendendo a uma condição e o problema de contar os números atendendo tal condição são problemas de natureza diferente. Vimos acima dois exemplos em que sempre é possível obter os elementos desejados, utilizando a mesma ideia, embora não saibamos como contar o número de elementos atendendo a essas condições. Os métodos modernos de crivo foram fortemente inspirados na busca de uma solução para os exemplos acima.

No que segue, daremos uma perspectiva um pouco mais geral sobre as ideias envolvidas no *crivo de Eratóstenes-Legendre*, além de fornecer outros exemplos onde as ideias podem ser aplicadas.

Sejam

$$\mathscr{A} = \{a \mid \dots\}$$

uma sequência finita de inteiros positivos menores do que ou igual a $x$, onde as propriedades que definem essa sequência estão representadas por três pontos, $\mathfrak{B}$ um conjunto de números primos menores do que ou igual a $z$ $(z < x)$. Seja também $d > 1$ um inteiro livre de quadrados e composto por primos do conjunto $\mathfrak{B}$ e defina,

$$\mathscr{A}_d := \{a \mid a \in \mathscr{A}, a \equiv 0 \mod (d)\}^8, \tag{3.2.2}$$

[8] Isto é o subconjunto de $\mathscr{A}$ cujos elementos são divisíveis por $d$.

e para $d = 1$

$$\mathscr{A}_1 := \mathscr{A}.$$

Em investigações aritméticas é comum usar uma aproximação conveniente $X$ ao invés do número exato $\mid \mathscr{A} \mid$, por exemplo para $\mid \mathscr{A} \mid = \pi(x)$ usamos $X = li$[9]. Considere então $X > 1$ uma aproximação conveniente para $\mid \mathscr{A} \mid$ e defina o erro da aproximação de $\mid \mathscr{A} \mid$ pelo número $X$ por,

$$r_1 := \mid \mathscr{A} \mid -X. \tag{3.2.3}$$

De maneira semelhante, para cada $p \in \mathfrak{B}$, podemos escolher $\omega(p)$ tal que $\frac{\omega(p)}{p}X$ seja uma aproximação[10] para $\mid \mathscr{A}_p \mid$ com erro de aproximação dado por,

$$r_p := \mid \mathscr{A}_p \mid -\frac{\omega(p)}{p} \cdot X \quad \text{para todo} \quad p \in \mathfrak{B}. \tag{3.2.4}$$

Com estas escolhas de $X$ e $\omega(p)$ agora definimos para cada inteiro $d$ livre de quadrados e composto de primos em $\mathfrak{B}$,

$$\omega(1) = 1, \qquad \omega(d) = \prod_{p \mid d} \omega(p), \qquad (\mu(d) \neq 0) \tag{3.2.5}$$

forçando, assim, $\omega(d)$ ser uma função multiplicativa.

Consistente com as equações (3.2.3) e (3.2.4), introduzimos

$$r_d := \mid \mathscr{A}_d \mid -\frac{\omega(d)}{d} \cdot X \quad (\mu(d) \neq 0). \tag{3.2.6}$$

Obviamente essas escolhas podem ser feitas de várias formas, mas como podemos esperar ela irá revelar-se melhor quanto menor for $\mid r_d \mid$.

Se estivermos interessados, como no *crivo de Eratóstenes-Legendre*, em estimar o número de elementos no conjunto

$$\mathcal{S}(\mathscr{A}, \mathfrak{B}) = \mathscr{A} \setminus \bigcup_{p \mid P(z)} \mathscr{A}_p, \tag{3.2.7}$$

onde $P(z)$ denota o produto dos primos menores do que ou igual a $z$ que pertencem a $\mathfrak{B}$, basta reescrever esse conjunto como,

$$\mathcal{S}(\mathscr{A}, \mathfrak{B}) = \{a \mid a \in \mathscr{A}, (a, P(z)) = 1\},$$

considerar a sequência finita

$$a_n = \begin{cases} 1 & \text{se } n \in \mathscr{A} \text{ e } (n, P(z)) = 1 \\ 0 & \text{caso contrário} \end{cases}.$$

[9] $li(x) = \int_2^x \frac{dt}{\log(t)}$

[10] Em termos práticos, a função $\frac{\omega(p)}{p}$, pode ser pensada como a proporção de elementos em $\mathscr{A}$ que pertencem a $\mathscr{A}_p$

e utilizar as escolhas de $X$ e $\omega(\cdot)$ como segue.

Claramente,

$$\mid \mathcal{S}(\mathscr{A}, \mathfrak{B}) \mid = \sum_{\substack{(n, P(z))=1 \\ n \leq x}} a_n$$

e pelo Lema (3.1.4)

$$\mid \mathcal{S}(\mathscr{A}, \mathfrak{B}) \mid = \sum_{\substack{(n, P(z))=1 \\ n \leq x}} a_n = \sum_{n \leq x} a_n \left( \sum_{d \mid (n, P(z))} \mu(d) \right) = \sum_{d \mid P(z)} \mu(d) \sum_{\substack{d \mid n \\ n \leq x}} a_n = \sum_{d \mid P(z)} \mu(d) \mid \mathscr{A}_d \mid .$$

Portanto,

$$\mid \mathcal{S}(\mathscr{A}, \mathfrak{B}) \mid = \sum_{d \mid P(z)} \mu(d) \mid \mathscr{A}_d \mid \tag{3.2.8}$$

e por (3.2.6) temos,

$$\mid \mathscr{A}_d \mid := \frac{\omega(d)}{d} \cdot X + r_d.$$

Segue que

$$\mid \mathcal{S}(\mathscr{A}, \mathfrak{B}) \mid = \sum_{d \mid P(z)} \mu(d) \left\{ \frac{\omega(d)}{d} \cdot X + r_d \right\} = X \cdot \sum_{d \mid P(z)} \mu(d) \frac{\omega(d)}{d} + \sum_{d \mid P(z)} \mu(d) r_d,$$

onde,

$$\sum_{d \mid P(z)} \mu(d) \frac{\omega(d)}{d} = \prod_{p \in \mathfrak{B}} \left( 1 - \frac{\omega(p)}{p} \right) \quad \text{e} \quad \sum_{d \mid P(z)} \mu(d) r_d = O\left( \sum_{d \mid P(z)} r_d \right).$$

Vejamos alguns exemplos[11].

**Exemplo 3.2.3.** *Uma discussão análoga àquela feita na Seção (3.1.3), com $\mathscr{A}$ definido abaixo, pode ser estabelecida com*

$$\mathscr{A} = \{a \in \mathbb{N} \mid Y - X \leq a \leq Y\}$$

*onde $Y \geq X > 0$.*

*Naquela situação, queríamos contar os primos menores do que $x$ e agora vamos querer contar o número de primos no intervalo $[Y - X, Y]$. Para tanto considere, como antes, o conjunto $\mathfrak{B}$ dos números primos menores do que $z$ e $P$ o produto de todos esses primos.*

*Para $\mathscr{A}_d$, como definido em (3.2.2), obtemos*

$$\mid \mathscr{A}_d \mid = \left\lfloor \frac{Y}{d} \right\rfloor - \left\lfloor \frac{Y - X}{d} \right\rfloor = \frac{X}{d} + O(1).$$

[11] Esses exemplos foram retirados de [4]

*De fato, primeiramente note que* $\left\lfloor \frac{x}{d} \right\rfloor$ *é o número de inteiros no conjunto* $\{1, 2, \ldots, x\}$ *que são divisíveis por* $d$*. Portanto, tomando*

$$\left\lfloor \frac{x}{d} \right\rfloor = \frac{x}{d} - \left\{ \frac{x}{d} \right\}$$

*onde* $\left\{ \frac{x}{d} \right\}$ *denota a parte fracionária de* $\frac{x}{d}$*, a parcela* $O(1)$ *surge como* $-\left\{ \frac{Y}{d} \right\} + \left\{ \frac{Y-X}{d} \right\}$ *e, portanto, reside no intervalo* $(-1, 1)$ *uma vez que,*

$$\frac{Y}{d} = \left\lfloor \frac{Y}{d} \right\rfloor + \left\{ \frac{Y}{d} \right\}, \quad \frac{Y-X}{d} = \left\lfloor \frac{Y-X}{d} \right\rfloor + \left\{ \frac{Y-X}{d} \right\}$$

*onde* $0 \leq \left\{ \frac{Y}{d} \right\} < 1$, $0 \leq \left\{ \frac{Y-X}{d} \right\} < 1$ *o que implica,*

$$\left\lfloor \frac{Y}{d} \right\rfloor - \left\lfloor \frac{Y-X}{d} \right\rfloor = \frac{X}{d} - \left\{ \frac{Y}{d} \right\} + \left\{ \frac{Y-X}{d} \right\}$$

*com* $-1 < -\left\{ \frac{Y}{d} \right\} + \left\{ \frac{Y-X}{d} \right\} < 1$*. Portanto, de fato*

$$\mid \mathscr{A}_d \mid = \frac{X}{d} + O(1).$$

*Se aplicarmos a identidade (3.2.8), obtemos, como na Seção (3.1.3),*

$$\mid \mathcal{S}(\mathscr{A}, \mathfrak{B}) \mid \leq \frac{X}{\log z} + O(2^{\pi(z)})$$

*e escolhendo* $z = \log X$ *encontramos*

$$\pi(Y) - \pi(Y-X) = O\left( \frac{X}{\log \log X} \right) \quad \textit{para}^{12}\ X \geq 3.$$

**Exemplo 3.2.4.** *Suponha agora que queiramos contar números dados por expressões polinomiais. Tomamos*

$$\mathscr{A} = \{ f(n) \mid Y - X < n \leq Y \}$$

*onde* $f(n)$ *é um polinômio com coeficientes inteiros, como por exemplo* $n^2 + 1$, $n \cdot (n-2)$ *ou* $n \cdot (n - C)$ [13]*. Novamente o conjunto* $\mathfrak{B}$ *denota os números primos menores do que ou igual a* $z$ *e* $P$ *o produto de todos esses primos.*

*Para estimar* $\mid \mathscr{A}_d \mid$ *vamos considerar cada classe de resíduo* $\mod (d)$ *separadamente,*

$$\mid \mathscr{A}_d \mid = \mid \{ n \mid Y - X < n \leq Y, f(n) \equiv 0 \mod (d) \} \mid .$$

*Para contar os elementos em* $\mathscr{A}_d$ *primeiro observe que desde que* $f(n)$ *é um polinômio com coeficientes inteiros,*

$$n \equiv r \mod (d) \Rightarrow f(n) \equiv f(r) \mod (d).$$

[12] O fato de $X \geq 3$ é imposto para assegurar que $\log \log X > 0$.

[13] Esse dois últimos polinômios são relevantes para investigações de *Brun* sobre as conjecturas dos *Primos Gêmeos* e de *Goldbach*

*Assim para os inteiros $l$ menores do que ou iguais a $d$ e incongruentes entre si, tal que $f(l) \equiv 0 \mod (d)$, é suficiente contar os inteiros $n$ no intervalo $Y - X < n \leq Y$ tais que $n \equiv l \mod (d)$, assim obtemos*

$$\mid \mathscr{A}_d \mid = \sum_{\substack{1 \leq l \leq d \\ f(l) \equiv 0 \mod (d)}} \sum_{\substack{Y - X < n \leq Y \\ n \equiv l \mod (d)}} 1 = \sum_{l} \left( \frac{X}{d} + O(1) \right). \tag{3.2.9}$$

*A soma mais interna sobre $n$ foi estimada de modo semelhante ao último exemplo, expressando $n$ como $l + md$. Denotando $\rho(d)$ como o número de raízes da congruência $f(l) \equiv 0 \mod (d)$, temos,*

$$\mid \mathscr{A}_d \mid = X \cdot \frac{\rho(d)}{d} + O(\rho(d)). \tag{3.2.10}$$

*A partir disto, se aplicarmos a identidade (3.2.8) obtemos,*

$$\sum_{d \mid P} \mu(d) \mid \mathscr{A}_d \mid = \sum_{d \mid P} \mu(d) \left( X \cdot \frac{\rho(d)}{d} + O(\rho(d)) \right) = X \cdot \sum_{d \mid P} \mu(d) \frac{\rho(d)}{d} + \sum_{d \mid P} \mu(d) O(\rho(d)).$$

*Em [5] vimos que $\rho(d)$ é uma função multiplicativa e que ela é não elementar (exceto em casos especiais). Portanto,*

$$\sum_{d \mid P} \mu(d) \frac{\rho(d)}{d} = \prod_{p < z} \left( 1 + \frac{\mu(p)\rho(p)}{p} \right) = \prod_{p < z} \left( 1 - \frac{\rho(p)}{p} \right)$$

*e como,*

$$\left| \sum_{d \mid P} \mu(d) O(\rho(d)) \right| \leq \sum_{d \mid P} |\mu(d) O(\rho(d))| = \sum_{d \mid P} |\mu(d)| \, |O(\rho(d))| = \sum_{d \mid P} |O(\rho(d))|$$

*uma vez que para $d$ livre de quadrados temos $\mu(d) = (-1)^k$ segue,*

$$\sum_{d \mid P} \mu(d) O(\rho(d)) = O\left( \sum_{d \mid P} \rho(d) \right).$$

*Daí,*

$$\mid \mathcal{S}(\mathscr{A}, \mathfrak{B}) \mid = X \cdot \prod_{p < z} \left( 1 - \frac{\rho(p)}{p} \right) + O\left( \sum_{d \mid P} \rho(d) \right). \tag{3.2.11}$$

## 3.3 O crivo de Eratóstenes - Legendre associado ao Truque de Rankin

Nesta seção, vamos descrever o crivo de *Eratóstenes-Legendre* sob uma perspectiva moderna[14]. No que segue, vamos considerar as notações da última seção e vamos definir a função $\omega(p)$, para

[14] A versão moderna do crivo de Eratóstenes-Legendre pode ser encontrada no livro [2] ou no artigo [8].

cada primo $p \in \mathfrak{B}$, como o número de classes de resíduos indesejáveis (distintas) módulo $p$ que estão $\mathscr{A}$.

Vamos definir também $\omega(d)$, para cada inteiro $d$ livre de quadrados e composto por primos em $\mathfrak{B}$, por

$$\omega(1) = 1 \quad \text{e} \quad \omega(d) = \prod_{p|d} \omega(p)$$

e vamos denotar $\mid \mathcal{S}(\mathscr{A}, \mathfrak{B}) \mid = \mid \mathcal{S}(\mathscr{A}, \mathfrak{B}, z) \mid$.

### 3.3.1 Uma versão moderna para o crivo de Eratóstenes-Legendre

**Teorema 3.3.1** (O Crivo de Eratóstenes)**.** *Com as notações acima, suponha que as seguintes condições sejam satisfeitas:*

1. $\mid r_d \mid = O(\omega(d))$*;*
2. *para algum* $\kappa \geq 0$*,*
$$\sum_{\substack{p \leq z \\ p \in \mathfrak{B}}} \frac{\omega(p) \log p}{p} \leq \kappa \log z + O(1);$$
3. *para algum número real positivo* $y \mid \mathscr{A}_d \mid = 0$ *para todo* $d > y$*.*

*Então,*

$$\mid \mathcal{S}(\mathscr{A}, \mathfrak{B}, z) \mid = X \cdot W(z) + O\left(X(\log z)^{\kappa+1} \exp\left(-\frac{\log X}{\log z}\right)\right), \tag{3.3.1}$$

*onde*

$$W(z) := \prod_{\substack{p \in \mathfrak{B} \\ p \leq z}} \left(1 - \frac{\omega(p)}{p}\right).$$

Para demonstrar esse teorema precisamos de alguns resultados preliminares.

**Lema 3.3.2.** *Com as hipóteses do Teorema (3.3.1), denotemos,*

$$F(t, z) := \sum_{\substack{d \leq t \\ d | P(z)}} \omega(d).$$

*Então,*

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

*Demonstração.* Primeiramente vejamos que a função $\omega(\cdot)$, definida para cada inteiro $d$ livre de quadrados e composto por primos em $\mathfrak{B}$, é uma função multiplicativa.

De fato, sejam $a, b \in \mathfrak{B}$, distintos, e $n = ab$, então

$$\omega(n) = \omega(ab) = \prod_{p|ab} \omega(p) = \prod_{p|a} \omega(p) \prod_{p|b} \omega(p) = \omega(a)\omega(b).$$

Usando o fato de que $\omega$ é uma função multiplicativa, podemos aplicar o *truque de Rankin*.

A ideia do chamado *truque de Rankin* reside no fato de que se uma função $f$ é multiplicativa e toma valores não negativos, então para $\delta > 0$ vale

$$\sum_{n\leq x} f(n) \leq \sum_{n\leq x} f(n)\left(\frac{x}{n}\right)^{\delta} \leq x^{\delta}\sum_{n=1}^{\infty}\frac{f(n)}{n^{\delta}}.$$

Portanto,

$$F(t,z) = \sum_{\substack{d\leq t\\ d|P(z)}} \omega(d) \leq \sum_{\substack{d\leq t\\ d|P(z)}} \omega(d)\left(\frac{t}{d}\right)^{\delta} \leq t^{\delta}\sum_{d|P(z)}\frac{\omega(d)}{d^{\delta}}.$$

Observe agora que, como $\omega$ é uma função multiplicativa, temos

$$\sum_{d|P(z)}\frac{\omega(d)}{d^{\delta}} = \prod_{\substack{p\leq z\\ p\in\mathfrak{B}}}\left(1+\frac{\omega(p)}{p^{\delta}}\right),$$

portanto

$$F(t,z) \leq t^{\delta}\prod_{\substack{p\leq z\\ p\in\mathfrak{B}}}\left(1+\frac{\omega(p)}{p^{\delta}}\right).$$

Usando agora a desigualdade[15] $1+x\leq e^{x}$, temos que para cada $x = \frac{\omega(p)}{p^{\delta}}$,

$$\left(1+\frac{\omega(p)}{p^{\delta}}\right) \leq \exp\left(\frac{\omega(p)}{p^{\delta}}\right)$$

portanto fazendo o produto sobre os primos $p\in\mathfrak{B}$, temos,

$$\prod_{\substack{p\leq z\\ p\in\mathfrak{B}}}\left(1+\frac{\omega(p)}{p^{\delta}}\right) \leq \exp\left(\sum_{\substack{p\leq z\\ p\in\mathfrak{B}}}\frac{\omega(p)}{p^{\delta}}\right)$$

além disso, fazendo $t^{\delta} = \exp\left(\log t^{\delta}\right)$

$$t^{\delta}\prod_{\substack{p\leq z\\ p\in\mathfrak{B}}}\left(1+\frac{\omega(p)}{p^{\delta}}\right) \leq \exp\left(\delta\log t + \sum_{\substack{p\leq z\\ p\in\mathfrak{B}}}\frac{\omega(p)}{p^{\delta}}\right)$$

e daí segue que,

$$F(t,z) \leq \exp\left(\delta\log t + \sum_{\substack{p\leq z\\ p\in\mathfrak{B}}}\frac{\omega(p)}{p^{\delta}}\right).$$

Colocando $\delta := 1-\eta$ e escrevendo

[15] Se fizermos $x$ negativo na desigualdade (3.1.9) temos essa nova desigualdade.

$$p^{-\delta} = p^{-1}p^{\eta} = p^{-1}\exp(\eta \log p)$$

temos

$$\exp\left(\delta \log t + \sum_{\substack{p\leq z\\ p\in\mathfrak{B}}} \frac{\omega(p)}{p^{\delta}}\right) = \exp\left((1-\eta)\log t + \sum_{\substack{p\leq z\\ p\in\mathfrak{B}}} \frac{\omega(p)}{p}\exp(\eta\log p)\right) =$$

$$= \exp\left(\log t - \eta\log t + \sum_{\substack{p\leq z\\ p\in\mathfrak{B}}} \frac{\omega(p)}{p}\exp(\eta\log p)\right) = \exp(\log t)\exp\left(-\eta\log t + \sum_{\substack{p\leq z\\ p\in\mathfrak{B}}} \frac{\omega(p)}{p}\exp(\eta\log p)\right) =$$

$$= t\exp\left(-\eta\log t + \sum_{\substack{p\leq z\\ p\in\mathfrak{B}}} \frac{\omega(p)}{p}\exp(\eta\log p)\right).$$

A desigualdade

$$e^x \leq 1 + xe^x \tag{3.3.2}$$

é válida para qualquer $x \in \mathbb{R}$.

De fato, a função $g(x) = (x-1)\cdot e^x$ tem como derivada $g'(x) = e^x + (x-1)\cdot e^x = xe^x$. Alem disso ela satisfaz

$$\begin{cases} g(0) = -1 \\ g'(0) = 0 \\ g'(x) > 0 \text{ para } x > 0 \\ g'(x) < 0 \text{ para } x < 0 \end{cases}$$

o que implica $x = 0$ ser ponto mínimo de $g$. Assim, $g(x) \geq -1\ \forall x \in \mathbb{R}$ e assim $e^x \leq 1 + xe^x$.

Usando, portanto, a desigualdade (3.3.2) juntamente com o fato de que a função $\exp(\cdot)$ é monótona temos,

$$t\exp\left(-\eta\log t + \sum_{\substack{p\leq z\\ p\in\mathfrak{B}}} \frac{\omega(p)}{p}\exp(\eta\log p)\right) \leq t\exp\left(-\eta\log t + \sum_{\substack{p\leq z\\ p\in\mathfrak{B}}} \frac{\omega(p)}{p}(1+\eta\log p\exp(\eta\log p)\right) =$$

$$= t\exp\left(-\eta\log t + \sum_{\substack{p\leq z\\ p\in\mathfrak{B}}} \frac{\omega(p)}{p} + \sum_{\substack{p\leq z\\ p\in\mathfrak{B}}} \frac{\omega(p)}{p}\eta\log p\exp(\eta\log p)\right) =$$

$$= t\exp\left(-\eta\log t + \sum_{\substack{p\leq z\\ p\in\mathfrak{B}}} \frac{\omega(p)}{p} + \eta\sum_{\substack{p\leq z\\ p\in\mathfrak{B}}} \frac{\omega(p)\log p}{p}\exp(\eta\log p)\right) \leq$$

$$\leq t\exp\left(-\eta\log t + \sum_{\substack{p\leq z\\ p\in\mathfrak{B}}} \frac{\omega(p)}{p} + \eta\sum_{\substack{p\leq z\\ p\in\mathfrak{B}}} \frac{\omega(p)\log p}{p}\exp(\eta\log z)\right)$$

e encontramos,

$$F(t,z) \leq t \exp\left(-\eta \log t + \sum_{\substack{p\leq z \\ p\in\mathfrak{B}}} \frac{\omega(p)}{p} + \eta z^{\eta} \sum_{\substack{p\leq z \\ p\in\mathfrak{B}}} \frac{\omega(p)\log p}{p}\right).$$

Agora pela segunda hipótese do Teorema (3.3.1) temos que para algum $\kappa \geq 0$,

$$\sum_{\substack{p\leq z \\ p\in\mathfrak{B}}} \frac{\omega(p)\log p}{p} \leq \kappa \log z + O(1).$$

Assim, fazendo

$$f(t) = \frac{1}{\log t} \quad \text{e} \quad c_n = \begin{cases} \frac{\omega(n)\log n}{n} \text{ se } n \in \mathfrak{B}, \\ 0 \text{ caso contrário.} \end{cases}$$

temos que,

$$C(z) = \sum_{\substack{p\leq z \\ p\in\mathfrak{B}}} \frac{\omega(p)\log p}{p} \leq \kappa \log z + O(1),$$

e mais, usando o fato dado no Lema (1.3.3) segue que

$$C(z) = \sum_{\substack{p\leq z \\ p\in\mathfrak{B}}} \frac{\omega(p)\log p}{p} \leq \kappa \log z + B$$

para alguma constante positiva $B$.

Assim, por somas parciais segue que,

$$\sum_{\substack{p<z \\ p\in\mathfrak{B}}} \frac{\omega(p)}{p} = \sum_{\substack{p<z \\ p\in\mathfrak{B}}} \frac{\omega(p)\log p}{p}\frac{1}{\log p} = C(z)\frac{1}{\log z} + \int_{n_0}^{z} \frac{C(u)}{u(\log u)^2}du \leq$$

$$\leq \frac{\kappa \log z + B}{\log z} + \int_{n_0}^{z} \frac{\kappa \log u + B}{u(\log u)^2}du = \kappa + \frac{B}{\log z} + \kappa \int_{n_0}^{z} \frac{du}{u \log u} + B\int_{n_0}^{z} \frac{du}{u(\log u)^2}$$

onde $n_0 = \min\{\mathfrak{B}\}$.

Como

$$\int_{n_0}^{z} \frac{du}{u\log u} = \log\log u|_{n_0}^{z} \quad \text{e} \quad \int_{n_0}^{z} \frac{du}{u(\log u)^2} = -\frac{1}{\log u}|_{n_0}^{z}.$$

Segue que

$$\sum_{\substack{p\leq z \\ p\in\mathfrak{B}}} \frac{\omega(p)}{p} \leq \kappa + \frac{B}{\log z} + \kappa\left(\log\log z - \log\log n_0\right) + B\left(-\frac{1}{\log z} + \frac{1}{\log n_0}\right),$$

assim,

$$\sum_{\substack{p\leq z \\ p\in\mathfrak{B}}} \frac{\omega(p)}{p} \leq \kappa \log\log z + O(1).$$

Novamente, usando o fato de que a função $\exp(\cdot)$ é monótona temos:

$$F(t,z) \leq t\exp\left(-\eta\log t + \sum_{\substack{p\leq z\\ p\in\mathfrak{B}}}\frac{\omega(p)}{p} + \eta z^{\eta}\sum_{\substack{p\leq z\\ p\in\mathfrak{B}}}\frac{\omega(p)\log p}{p}\right) \leq$$

$$\leq t\exp\left(-\eta\log t + \kappa\log\log z + O(1) + \eta z^{\eta}(\kappa\log z + O(1))\right) =$$

$$= t\exp\left(-\eta\log t + \kappa\log\log z + \eta z^{\eta}\kappa\log z\right)\exp(O(1))\exp(\eta z^{\eta}O(1)).$$

Escolhendo $\eta := \frac{1}{\log z}$ (onde $z > e$), temos,

$$t\exp\left(-\eta\log t + \kappa\log\log z + \eta z^{\eta}\kappa\log z\right)\exp(O(1))\exp(\eta z^{\eta}O(1)) =$$

$$= t\exp\left(-\frac{\log t}{\log z} + \log(\log z)^{\kappa} + z^{\frac{1}{\log z}}\kappa\right)\exp(O(1))\exp\left(\frac{1}{\log z}O(1)z^{\frac{1}{\log z}}\right).$$

Como

$$z^{\frac{1}{\log z}} = c \iff \log_z c = \frac{1}{\log z} \quad \text{e} \quad \log_z c = \frac{\log c}{\log z}$$

segue que

$$\log c = 1 \iff c = e.$$

Portanto,

$$t\exp\left(-\frac{\log t}{\log z} + \log(\log z)^{\kappa} + z^{\frac{1}{\log z}}\kappa\right)\exp(O(1))\exp\left(\frac{1}{\log z}O(1)z^{\frac{1}{\log z}}\right) =$$

$$= t\exp\left(-\frac{\log t}{\log z}\right)(\log z)^{\kappa}\exp(e\kappa)\exp(O(1))\exp\left(\frac{e}{\log z}O(1)\right).$$

Novamente, como $\log z > 1$ temos $\exp\left(\frac{e}{\log z}O(1)\right) \leq \exp(r)$ para alguma constante positiva $r$. Logo

$$t(\log z)^{\kappa}\exp\left(-\frac{\log t}{\log z}\right)\exp(e\kappa)\exp(O(1))\exp\left(\frac{e}{\log z}O(1)\right) \leq$$

$$\leq t(\log z)^{\kappa}\exp\left(-\frac{\log t}{\log z}\right)\exp(e\kappa)\exp(O(1))\exp(r)$$

e assim,

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

□

**Lema 3.3.3.** *Com as hipóteses do Teorema (3.3.1),*

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

*Demonstração.* Primeiramente temos,

$$\sum_{\substack{y<d\leq x\\ d|P(z)}}\frac{\omega(d)}{d}=\sum_{\substack{d\leq x\\ d|P(z)}}\frac{\omega(d)}{d}-\sum_{\substack{d\leq y\\ d|P(z)}}\frac{\omega(d)}{d}.$$

Através de somas parciais, usando,

$$f(x)=\frac{1}{x},\quad c_d=\begin{cases}\omega(d)\text{ se }\ d\in\mathfrak{B}\\ 0\text{ caso contrário}\end{cases}\quad\text{e}\quad n_0=\min\{\mathfrak{B}\}$$

temos que,

$$\sum_{\substack{y<d\leq x\\ d|P(z)}}\frac{\omega(d)}{d}=F(x,z)\frac{1}{x}+\int_{n_0}^{x}\frac{F(t,z)}{t^2}dt-\left(F(y,z)\frac{1}{y}+\int_{n_0}^{y}\frac{F(t,z)}{t^2}dt\right)=$$

$$=F(x,z)\frac{1}{x}+\int_{n_0}^{x}\frac{F(t,z)}{t^2}dt+\left(-F(y,z)\frac{1}{y}+\int_{y}^{n_0}\frac{F(t,z)}{t^2}dt\right).$$

Portanto,

$$\sum_{\substack{y<d\leq x\\ d|P(z)}}\frac{\omega(d)}{d}=F(x,z)\frac{1}{x}-F(y,z)\frac{1}{y}+\int_{y}^{x}\frac{F(t,z)}{t^2}dt.$$

Note, no entanto, que a soma $F(x,z)$ é no máximo $\sum_{d|P(z)}\omega(d)$. Assim quando $x\to\infty$ temos que $F(x,z)\frac{1}{x}\to 0$, pois $\frac{1}{x}\to 0$ quando $x\to\infty$, logo

$$\sum_{\substack{y<d\\ d|P(z)}}\frac{\omega(n)}{n}=-F(y,z)\frac{1}{y}+\int_{y}^{\infty}\frac{F(t,z)}{t^2}dt.$$

Como

$$\int_y^\infty\frac{F(t,z)}{t^2}dt=\int_y^\infty\frac{1}{t^2}O\left(t(\log z)^\kappa\exp\left(-\frac{\log t}{\log z}\right)\right)dt=\int_y^\infty O\left(\frac{1}{t}(\log z)^\kappa\exp\left(-\frac{\log t}{\log z}\right)\right)dt$$

e

$$\int_y^\infty O\left(\frac{1}{t}(\log z)^\kappa\exp\left(-\frac{\log t}{\log z}\right)\right)dt\leq\int_y^\infty A\frac{1}{t}(\log z)^\kappa\exp\left(-\frac{\log t}{\log z}\right)dt,$$

para alguma constante positiva $A$, temos

$$\int_y^\infty\frac{F(t,z)}{t^2}dt\leq A(\log z)^\kappa\int_y^\infty\frac{1}{t}\exp\left(-\frac{\log t}{\log z}\right)dt.$$

Fazendo agora uma substituição, na última integral, $u=-\frac{\log t}{\log z}$ temos que $\frac{du}{dt}=-\frac{1}{\log z}\frac{1}{t}$, portanto,

$$\int_y^\infty\frac{1}{t}\exp\left(-\frac{\log t}{\log z}\right)dt=(-\log z)\lim_{x\to\infty}\int_y^x e^u du.$$

Como

$$\int e^u du = (e^u + C) = \exp\left(-\frac{\log t}{\log z}\right) + C,$$

segue que

$$\lim_{x\to\infty}\int_y^x e^u du = \lim_{x\to\infty}\left(\exp\left(-\frac{\log t}{\log z}\right) + C\right)|_y^x = -\exp\left(-\frac{\log y}{\log z}\right).$$

Assim,

$$\int_y^\infty \frac{F(t,z)}{t^2}dt \leq A(\log z)^\kappa(-\log z)\lim_{x\to\infty}\int_y^x e^u du = A(\log z)^{\kappa+1}\exp\left(-\frac{\log y}{\log z}\right).$$

Sabemos também, pelo Lema (3.3.2), que,

$$-F(y,z)\frac{1}{y} = \frac{1}{y}O\left(y(\log z)^\kappa \exp\left(-\frac{\log y}{\log z}\right)\right) = O\left((\log z)^\kappa \exp\left(-\frac{\log y}{\log z}\right)\right)$$

assim, para alguma constante positiva $B$, temos

$$\left|-F(y,z)\frac{1}{y}\right| \leq B(\log z)^\kappa \exp\left(-\frac{\log y}{\log z}\right) \leq B(\log z)^{\kappa+1}\exp\left(-\frac{\log y}{\log z}\right).$$

Portanto, podemos concluir que

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

□

*Demonstração.* (Teorema (3.3.1))

Pelo Principio de *inclusão-exclusão* temos,

$$|\,\mathcal{S}(\mathscr{A},\mathfrak{B},z)\,| = \sum_{\substack{d|P(z)\\ d\leq y}}\mu(d)\,|\,\mathscr{A}_d\,| = \sum_{\substack{d|P(z)\\ d\leq y}}\mu(d)\left(X\frac{\omega(d)}{d} + r_d\right) = \sum_{\substack{d|P(z)\\ d\leq y}}\mu(d)X\frac{\omega(d)}{d} + \sum_{\substack{d|P(z)\\ d\leq y}}\mu(d)r_d.$$

Como

$$\left|\sum_{\substack{d|P(z)\\ d\leq y}}\mu(d)r_d\right| \leq \sum_{\substack{d|P(z)\\ d\leq y}}|\mu(d)r_d| = \sum_{\substack{d|P(z)\\ d\leq y}}|\mu(d)|\,|r_d| = \sum_{\substack{d|P(z)\\ d\leq y}}|r_d|$$

uma vez que para $d$ livre de quadrados temos $\mu(d) = (-1)^k$ segue

$$\left|\sum_{\substack{d|P(z)\\ d\leq y}}\mu(d)r_d\right| \leq \sum_{\substack{d|P(z)\\ d\leq y}}|r_d|.$$

Pela primeira hipótese temos que $\mid r_d \mid = O(\omega(d))$, portanto existe uma menor constante $c_d > 0$ tal que $\mid r_d \mid \leq c_d\omega(d)$. Fazendo $C = \max_d\{c_d\}$ temos que

$$\sum_{\substack{d|P(z)\\ d\leq y}} \mid r_d \mid \leq C \sum_{\substack{d|P(z)\\ d\leq y}} \omega(d)$$

portanto,

$$\mid \mathcal{S}(\mathscr{A}, \mathfrak{B}, z) \mid = \sum_{\substack{d|P(z)\\ d\leq y}} \mu(d) X \frac{\omega(d)}{d} + O(F(y, z)). \tag{3.3.3}$$

Novamente, utilizando o fato de que $\omega$ é uma função multiplicativa temos que

$$\sum_{d|P(z)} \mu(d)\frac{\omega(d)}{d} = \prod_{\substack{p<z\\ p\in\mathfrak{B}}} \left(1 + \frac{\mu(p)\omega(p)}{p}\right) = \prod_{\substack{p<z\\ p\in\mathfrak{B}}} \left(1 - \frac{\omega(p)}{p}\right).$$

Sabendo também que,

$$\sum_{d|P(z)} \mu(d)\frac{\omega(d)}{d} = \sum_{\substack{d|P(z)\\ d\leq y}} \mu(d)\frac{\omega(d)}{d} + \sum_{\substack{d|P(z)\\ y<d}} \mu(d)\frac{\omega(d)}{d}$$

segue que a equação (3.3.3) pode ser escrita como

$$\mid \mathcal{S}(\mathscr{A}, \mathfrak{B}, z) \mid = X\left(\sum_{d|P(z)} \mu(d)\frac{\omega(d)}{d} - \sum_{\substack{d|P(z)\\ y<d}} \mu(d)\frac{\omega(d)}{d}\right) + O(F(y, z))$$

e, portanto,

$$\mid \mathcal{S}(\mathscr{A}, \mathfrak{B}, z) \mid = XW(z) - X \sum_{\substack{d|P(z)\\ y<d}} \mu(d)\frac{\omega(d)}{d} + O(F(y, z)). \tag{3.3.4}$$

Como

$$\left|\sum_{\substack{d|P(z)\\ y<d}} \mu(d)\frac{\omega(d)}{d}\right| \leq \sum_{\substack{d|P(z)\\ y<d}} \left|\mu(d)\frac{\omega(d)}{d}\right| = \sum_{\substack{d|P(z)\\ y<d}} |\mu(d)| \left|\frac{\omega(d)}{d}\right|,$$

$d$ é um inteiro positivo livre de quadrados e $\omega(d) > 0$, segue que

$$\left|\sum_{\substack{d|P(z)\\ y<d}} \mu(d)\frac{\omega(d)}{d}\right| \leq \sum_{\substack{d|P(z)\\ y<d}} \frac{\omega(d)}{d}.$$

Portanto,

$$\mid \mathcal{S}(\mathscr{A}, \mathfrak{B}, z) \mid = XW(z) - XO\left(\sum_{\substack{d|P(z)\\ y<d}} \frac{\omega(d)}{d}\right) + O(F(y, z)). \tag{3.3.5}$$

Pelos Lemas (3.3.2) e (3.3.3) obtemos que

$$\mid \mathcal{S}(\mathscr{A}, \mathfrak{B}, z) \mid = XW(z) - XO\left(O\left((\log z)^{\kappa+1} \exp\left(-\frac{\log y}{\log z}\right)\right)\right) + O\left(O\left(y(\log z)^{\kappa} \exp\left(-\frac{\log y}{\log z}\right)\right)\right)$$

e pela parte 2 da proposição (1.3.7), temos

$$\mid \mathcal{S}(\mathscr{A}, \mathfrak{B}, z) \mid = XW(z) - XO\left((\log z)^{\kappa+1} \exp\left(-\frac{\log y}{\log z}\right)\right) + O\left(y(\log z)^{\kappa} \exp\left(-\frac{\log y}{\log z}\right)\right).$$

Fazendo $y = CX$ para alguma constante positiva $C$, temos que

$$y(\log z)^{\kappa} \exp\left(-\frac{\log y}{\log z}\right) = O\left(X(\log z)^{\kappa+1} \exp\left(-\frac{\log y}{\log z}\right)\right).$$

De fato,

$$\left|CX(\log z)^{\kappa} \exp\left(-\frac{\log y}{\log z}\right)\right| \leq \mid C \mid \left|X(\log z)^{\kappa+1} \exp\left(-\frac{\log y}{\log z}\right)\right|.$$

Portanto,

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

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

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

Por fim vamos mostrar que

$$X(\log z)^{\kappa+1} \exp\left(-\frac{\log y}{\log z}\right) = O\left(X(\log z)^{\kappa+1} \exp\left(-\frac{\log X}{\log z}\right)\right).$$

De fato, como

$$\exp\left(-\frac{\log C}{\log z}\right) \leq \exp(r)$$

para alguma constante positiva $r$, temos,

$$\left|X(\log z)^{\kappa+1} \exp\left(-\frac{\log CX}{\log z}\right)\right| = \left|X(\log z)^{\kappa+1} \exp\left(-\frac{\log C}{\log z} - \frac{\log X}{\log z}\right)\right| =$$

$$= \left|X(\log z)^{\kappa+1} \exp\left(-\frac{\log C}{\log z}\right) \exp\left(-\frac{\log X}{\log z}\right)\right| \leq \exp(r) \left|X(\log z)^{\kappa+1} \exp\left(-\frac{\log X}{\log z}\right)\right|$$

Portanto,

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

$$= XW(z) + O\left(X(\log z)^{\kappa+1} \exp\left(-\frac{\log X}{\log z}\right)\right)$$

e isso conclui a demonstração. □

Uma consequência imediata do último resultado é o corolário.

**Corolário 3.3.4.** *Se $\pi(X)$ denota o número de primos menores do que ou iguais a $X$. Então,*

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

*Demonstração.* De fato, sejam $\mathscr{A}$ o conjunto dos inteiros positivos menores do que $X$, $\mathfrak{B}$ o conjunto de números primos menores do que ou iguais a $z$, para algum $z$ a ser escolhido futuramente.

Quando tomamos $\omega(d) = 1$, estamos contando uma classe de resíduo módulo $d$, no nosso caso queremos os inteiros na classe de resíduo 0 módulo $d$. Além disso, a função $\omega(\cdot)$ definida desta forma, para todos os inteiros $d$ livres de quadrados e compostos por primos em $\mathfrak{B}$, é multiplicativa.
Com efeito, sejam $a, b \in \mathfrak{B}$, distintos, então,

$$\omega(ab) = 1 = 1 \cdot 1 = \omega(a)\omega(b).$$

Com essa escolha para a função $\omega(\cdot)$ vemos que as hipóteses do Teorema (3.3.1) satisfazem:

1. $| r_d | = O(1)$;
2. No Exemplo (1.3.9), temos que para $\kappa = 1$,

$$\sum_{p \leq z} \frac{\log p}{p} = \log z + R(z);$$

   com $R(z) = O(1)$

3. para $C > 0$, $y = CX$ $| \mathscr{A}_d | = 0$ para todo $d > y$.

Então pelo Teorema (3.3.1) temos

$$| \mathcal{S}(\mathscr{A}, \mathfrak{B}, z) | = XW(z) + O\left(X(\log z)^2 \exp\left(-\frac{\log X}{\log z}\right)\right),$$

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

Vimos também, na seção (3.1.3), que para alguma constante positiva

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

portanto,

$$\mid \mathcal{S}(\mathscr{A},\mathfrak{B},z)\mid < X\frac{1}{c\log z}+O\left(X(\log z)^2\exp\left(-\frac{\log X}{\log z}\right)\right).$$

Escolhendo $z=\exp\left(\frac{\log X}{3\log\log X}\right)$, segue que

$$\log z=\frac{\log X}{3\log\log X}$$

e portanto temos

$$\begin{aligned}\mid \mathcal{S}(\mathscr{A},\mathfrak{B},z)\mid &< X\frac{1}{c\frac{\log X}{3\log\log X}}+O\left(X\left(\frac{\log X}{3\log\log X}\right)^2\exp\left(-\frac{\log X}{\frac{\log X}{3\log\log X}}\right)\right)=\\ &=X\frac{3\log\log X}{c\log X}+O\left(X\left(\frac{\log X}{3\log\log X}\right)^2\exp\left(-3\log\log X\right)\right)=\\ &=X\frac{3\log\log X}{c\log X}+O\left(X\left(\frac{\log X}{3\log\log X}\right)^2(\log X)^{-3}\right)=\\ &=X\frac{3\log\log X}{c\log X}+O\left(\frac{X}{(3\log\log X)^2(\log X)}\right).\end{aligned}$$

Portanto temos que

$$\mid \mathcal{S}(\mathscr{A},\mathfrak{B},z)\mid < X\frac{3\log\log X}{c\log X}+f_1(X),$$

onde $\mid f_1(X)\mid \leq A\left|\frac{X}{(3\log\log X)^2(\log X)}\right|$ para alguma constante positiva $A$.

Observe que

$$\begin{aligned}\left|X\frac{3\log\log X}{c\log X}+f_1(X)\right| &\leq \left|X\frac{3\log\log X}{c\log X}\right|+|f_1(X)|\leq\\ &\leq\left|X\frac{3\log\log X}{c\log X}\right|+A\left|\frac{X}{(3\log\log X)^2(\log X)}\right|,\end{aligned}$$

observe também que a função

$$\left|X\frac{3\log\log X}{c\log X}\right|$$

cresce mais rapidamente do que a função

$$A\left|\frac{X}{(3\log\log X)^2(\log X)}\right|.$$

De fato,

$$\lim_{X\to\infty} \frac{A\left|\frac{X}{(3\log\log X)^2(\log X)}\right|}{\left|X\frac{3\log\log X}{c\log X}\right|} = \lim_{X\to\infty} A\left|\frac{X}{(3\log\log X)^2(\log X)}\right|\left|\frac{c\log X}{3X\log\log X}\right| = \lim_{X\to\infty}\left|\frac{Ac}{(3\log\log X)^3}\right|$$

portanto

$$\lim_{X\to\infty}\left|\frac{Ac}{(3\log\log X)^3}\right| \to 0.$$

Portanto podemos concluir que

$$\mid \mathcal{S}(\mathscr{A},\mathfrak{B},z) \mid = O\left(\frac{X}{\log X}\log\log X\right).$$

Observe no entanto que

$$\pi(X) - \pi(z) \leq \mid \mathcal{S}(\mathscr{A},\mathfrak{B},z) \mid,$$

uma vez que no conjunto $\mid \mathcal{S}(\mathscr{A},\mathfrak{B},z) \mid$ estão: o número 1, todos os primos do intervalo $z < p < X$ e os números compostos que são produto de primos maiores do que $z$. Assim, segue que

$$\pi(X) \leq \mid \mathcal{S}(\mathscr{A},\mathfrak{B},z) \mid + \pi(z) \leq \mid \mathcal{S}(\mathscr{A},\mathfrak{B},z) \mid + z.$$

Como $z = \exp\left(\frac{\log X}{3\log\log X}\right)$, temos

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

De fato, para qualquer constante positiva $C$ a função $C\frac{X}{\log X}\log\log X$ cresce mais rapidamente do que $\exp\left(\frac{\log X}{3\log\log X}\right)$, i.e.,

$$\lim_{X\to\infty}\left|\frac{\exp\left(\frac{\log X}{3\log\log X}\right)}{C\frac{X}{\log X}\log\log X}\right| \to 0.$$

□

## 3.4 Algumas Considerações

Tivemos, nesse trabalho, a intenção de estudar o crivo de *Eratóstenes - Legendre* com o objetivo de entender as ideias envolvidas nos processos de crivo. A escolha desse crivo, dentre todos os outros, foi feita tendo em vista dois pontos. O primeiro é que o crivo de *Eratóstenes - Legendre* é o mais simples dentre os crivos estudados na teoria dos crivos. O segundo ponto reside no fato

de que esse crivo fornece a ideia geral dos crivos uma vez que os crivos[16] mais sofisticados são extensões de suas ideias básicas.

Uma motivação desse trabalho reside no desejo de aplicar as ideias de crivo a uma situação particular, o que será feito em algum trabalho futuro.

Ao leitor interessado em obter mais ideias e resultados, um bom texto introdutório é [4]. Neste texto as ideias são apresentadas de maneira objetiva e clara. Um livro, sobre o assunto, também introdutório e mais completo é sem dúvida [5]. Embora essa última bibliografia seja rica em conteúdo, ela leva em consideração muitos detalhes técnicos, o que sob meu ponto de vista faz com que o leitor se perca nos detalhes em prejuízo as ideias. Um texto mais recente, sobre teoria de crivos, e que também foi utilizado neste trabalho é [2]. Neste último são discutidos e apresentados muitos outros tipos de crivos combinatoriais e não combinatoriais.

[16] Crivos combinatoriais.

# Referências


[1] M. Aigner. *A Course in Enumeration*. Berlin: Springer, 2007.

[2] A. C. Cojocaru e M.R. Murty. *An Introduction to Sieve Methods and their Applications*. Cambridge: Cambridge University Press, 2005.

[3] R. L. Graham, D. E. Knuth e O. Patashnik. *Matemática Concreta*. Rio de Janeiro: LTC, 1995.

[4] G. Greaves. *Sieves in Number Theory*. New York: Springer, 2001.

[5] H. Halberstam e Richert H. E. *Sieve Methods*. London: Academic Press, 1974.

[6] P. R. Halmos. *Teoria ingênua dos conjuntos*. São Paulo: Polígono, 1970.

[7] H. Iwaniec. *Notas de aula: Sieve Methods*. Piscataway: Não Publicada, 1996.

[8] M. R. Murty e N. Saradha. "On the sieve of Eratosthenes." Em: *Canadian Journal of Mathematics* 39.5 (1987), pp. 1107–1122.

[9] J. P. O. Santos, M. P. Mello e I. T. C. Murari. *Introdução à Análise Combinatória*. Rio de Janeiro: Ciência Moderna, 2007.

[10] A. Slomson. *An Introduction to Combinatorics*. Great Britain: Chapman e Hall, 1991.