

Quando Frações Encontram Círculos: geometria por trás dos números

Jones Colombo

Universidade Federal Fluminense

XI Bienal de Matemática

Agosto de 2026

Sumário

0.1	Aproximação de Números Reais por Números Racionais	1
0.1.1	O Princípio da Casa dos Pombos de Dirichlet	1
1	Frações Decimais, Sequências de Farey e Círculos de Ford	5
1.1	Introdução	5
1.2	A Função Maior Inteiro	6
1.3	Representação Decimal de Números Racionais	7
1.4	A Função ϕ de Euler e o Princípio de Inclusão-Exclusão	9
1.4.1	A Função ϕ de Euler	10
1.4.2	Abordagem Alternativa: Dedução por Somas de Divisores	11
1.5	Séries Infinitas e a Representação Decimal dos Números Reais	12
1.6	Periodicidade das Expansões Decimais	15
1.6.1	Decimais Periódicos Puros	16
1.7	Grafos de Divisibilidade	18
2	Sequências de Farey	21
2.1	Qual é o Comprimento de uma Série de Farey?	23
2.2	Outra demonstração da propriedade fundamental do MDC	24
3	Círculos de Ford	25
3.1	Frações Adjacentes	27
3.2	Os Círculos de Ford nos Complexos	29
4	Geometria Hiperbólica	31
4.1	Construções Euclidianas Úteis	31
4.2	Construções com Régua e Compasso na Geometria Hiperbólica	35
4.3	O Modelo do Disco de Poincaré	35
4.4	Primeiras Construções	36
4.5	Conexão Analítica: O Modelo do Semi-Plano Superior, os Círculos de Ford e as Ternas Pitagóricas	39
4.5.1	Para onde vão os pontos de Farey	42
4.5.2	Os Círculos de Ford dentro do disco	43
4.5.3	Conexão com as Ternas Pitagóricas	43
5	Melhor Aproximação Racional de Números Irracionais; Círculos de Ford e o Teorema de Hurwitz	45
5.1	Qualidade da Aproximação	45
5.2	Triângulos Circulares	47

6	Frações Contínuas	53
6.1	Definição - Notação	53
6.2	Convergentes	55
6.3	Aproximações Sucessivas	58
6.4	Propriedades dos Convergentes	61
6.5	Problemas Resolvidos	69
6.6	Problemas Propostos	70
	Referências Bibliográficas	71

Introdução

0.1 Aproximação de Números Reais por Números Racionais

O problema de aproximar um número real arbitrário por um número racional está na base da teoria das aproximações diofantinas. A ferramenta mais fundamental para este fim é um argumento surpreendentemente simples, porém poderoso, introduzido por P. G. Lejeune Dirichlet, conhecido modernamente como o *Princípio da Casa dos Pombos* (ou Princípio das Gavetas de Dirichlet).

0.1.1 O Princípio da Casa dos Pombos de Dirichlet

Lema 0.1 (Princípio da Casa dos Pombos) *Se mais de n objetos forem distribuídos em n caixas, então pelo menos uma caixa conterá dois ou mais objetos.*

Utilizando este princípio combinatório elementar, Dirichlet provou um teorema fundamental sobre a aproximação de números reais.

Teorema 0.2 (Teorema de Aproximação de Dirichlet) *Para qualquer número real α e qualquer número inteiro positivo Q , existem números inteiros a e q tais que $1 \leq q \leq Q$ e*

$$\left| \alpha - \frac{a}{q} \right| < \frac{1}{qQ}.$$

Demonstração: Consideremos a parte fracionária de um número real, denotada por $\{x\} = x - \lfloor x \rfloor$, onde $\lfloor x \rfloor$ representa a parte inteira de x . Claramente, para todo x , temos $0 \leq \{x\} < 1$.

Analisemos os seguintes $Q + 1$ números reais localizados no intervalo $[0, 1)$:

$$0, \{\alpha\}, \{2\alpha\}, \{3\alpha\}, \dots, \{Q\alpha\}.$$

Subdividamos o intervalo semiaberto $[0, 1)$ em Q subintervalos disjuntos da seguinte forma:

$$\left[0, \frac{1}{Q}\right), \left[\frac{1}{Q}, \frac{2}{Q}\right), \left[\frac{2}{Q}, \frac{3}{Q}\right), \dots, \left[\frac{Q-1}{Q}, 1\right).$$

Como temos $Q + 1$ números e apenas Q subintervalos (as “caixas”), pelo princípio da casa dos pombos, pelo menos dois desses números devem pertencer ao mesmo subintervalo.

Suponhamos que estes dois números sejam $\{r\alpha\}$ e $\{s\alpha\}$, com $0 \leq r < s \leq Q$. Como eles estão no mesmo intervalo de comprimento $1/Q$, a diferença absoluta entre eles deve ser estritamente menor que $1/Q$:

$$|\{s\alpha\} - \{r\alpha\}| < \frac{1}{Q}.$$

Substituindo a definição de parte fracionária, obtemos:

$$|(s\alpha - \lfloor s\alpha \rfloor) - (r\alpha - \lfloor r\alpha \rfloor)| < \frac{1}{Q},$$

o que equivale a

$$|(s - r)\alpha - (\lfloor s\alpha \rfloor - \lfloor r\alpha \rfloor)| < \frac{1}{Q}.$$

Definamos agora os inteiros q e a como:

$$q = s - r, \quad a = \lfloor s\alpha \rfloor - \lfloor r\alpha \rfloor.$$

Dado que $0 \leq r < s \leq Q$, temos que q é um número inteiro que satisfaz $1 \leq q \leq Q$. A desigualdade torna-se:

$$|q\alpha - a| < \frac{1}{Q}.$$

Dividindo ambos os lados por q (já que $q > 0$), concluímos que:

$$\left| \alpha - \frac{a}{q} \right| < \frac{1}{qQ}.$$

Isso encerra a demonstração. □

Como $q \leq Q$, temos que $\frac{1}{qQ} \leq \frac{1}{q^2}$. Portanto, um corolário imediato e muito célebre deste teorema é que, se α for um número irracional, existem infinitas frações racionais $\frac{a}{q}$ que satisfazem a desigualdade:

$$\left| \alpha - \frac{a}{q} \right| < \frac{1}{q^2}.$$

Neste trabalho vamos expandir e refinar a ideia de aproximação diofantina introduzida acima. Em vez de nos contentarmos com uma prova de existência, como a de Dirichlet, vamos construir um aparato geométrico e combinatório explícito para **encontrar e ordenar** essas aproximações racionais de um número real de forma sistemática. O fio condutor de todo o texto é, portanto, sempre o mesmo problema — aproximar números reais por números racionais —, abordado por três caminhos complementares: o aritmético (frações decimais e congruências), o geométrico (sequências de Farey e círculos de Ford, inclusive vistos sob a ótica da geometria hiperbólica) e o algébrico (frações contínuas).

O objetivo central deste trabalho pode ser dividido em três frentes principais:

- Introduzir a Série de Farey de ordem n (denotada por $\mathcal{F}_n$), que é o conjunto ordenado de todas as frações irredutíveis no intervalo $[0, 1]$ cujos denominadores não excedem n . Queremos mostrar como os números reais são “cercados” por essas frações. Por exemplo, para $\mathcal{F}_4$, temos:

$$\mathcal{F}_4 = \left\{ \frac{0}{1}, \frac{1}{4}, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \frac{1}{1} \right\}$$

- Provar rigorosamente duas propriedades matemáticas profundas e quase mágicas que regem essas séries:

- **A Propriedade do Determinante:** Se $\frac{a}{b}$ e $\frac{c}{d}$ são duas frações consecutivas em $\mathcal{F}_n$, então:

$$bc - ad = 1$$

Isso significa que a área do paralelogramo formado pelos vetores (b, a) e (d, c) no plano linear é exatamente 1, uma conexão direta com a geometria dos números de Minkowski.

- **A Propriedade da Mediana (Interpolação):** A primeira fração que surge entre duas frações consecutivas $\frac{a}{b}$ e $\frac{c}{d}$ ao passarmos para uma ordem superior $\mathcal{F}_{b+d}$ é a sua *mediana* (ou *mediant*):

$$\frac{a+c}{b+d}$$

- Concluir que as Séries de Farey oferecem uma aproximação **mais precisa** do que o Teorema de Dirichlet.

Enquanto Dirichlet garante que, para qualquer α , existe *pelo menos* uma fração tal que $|\alpha - \frac{a}{q}| < \frac{1}{q^2}$, seguiremos Rademacher no uso das propriedades das frações vizinhas de Farey para provar que, para qualquer número real α , existem duas frações consecutivas de $\mathcal{F}_n$ que o comprimem, de tal forma que:

$$\left| \alpha - \frac{a}{b} \right| \leq \frac{1}{b(b+d)} \quad \text{e} \quad \left| \alpha - \frac{c}{d} \right| \leq \frac{1}{d(b+d)}$$

Em resumo: Vamos transformar uma prova de existência abstrata (o método de Dirichlet) em um **algoritmo geométrico e construtivo**. As Séries de Farey funcionam como uma “peneira” perfeitamente calibrada que permite aproximar qualquer número irracional com precisão cirúrgica, pavimentando o caminho para o estudo posterior dos Círculos de Ford e das Frações Contínuas.

Roteiro do Texto

Para levar a cabo este programa, o texto está organizado em seis capítulos, que percorrem o problema da aproximação racional por três vias distintas — porém intimamente relacionadas.

No **Capítulo 1** partimos de um objeto bem mais elementar, a representação decimal de um número racional, e investigamos por que algumas dízimas são periódicas e outras não. Essa investigação nos conduz, de forma natural, à linguagem das congruências e à função ϕ de Euler, culminando no Teorema de Euler-Fermat e em um método explícito para calcular o resto da divisão de n por d . O capítulo se encerra introduzindo, ainda de modo informal, as sequências de Farey e os círculos de Ford, que serão o objeto dos capítulos seguintes.

No **Capítulo 2** construímos rigorosamente as sequências de Farey $\mathcal{F}_n$ e demonstramos suas duas propriedades fundamentais, mencionadas acima: a propriedade do determinante e a propriedade da mediana. Mostramos também como obter $\mathcal{F}_n$ a partir de $\mathcal{F}_{n-1}$, e usamos as sequências de Farey para dar uma nova demonstração da identidade de Bézout.

No **Capítulo 3** associamos a cada fração de Farey p/q uma circunferência tangente ao eixo real — o círculo de Ford $C(p/q)$ — e mostramos que dois destes círculos são sempre tangentes ou totalmente exteriores um ao outro, e que essa tangência ocorre exatamente quando as frações correspondentes são vizinhas de Farey. Essa tradução geométrica da vizinhança de Farey é o que permitirá, mais adiante, medir a qualidade de uma aproximação racional através do raio do círculo correspondente.

No **Capítulo 4** damos um passo lateral, necessário para o que segue: apresentamos os fundamentos da geometria hiperbólica no modelo do disco de Poincaré, incluindo inversões, transformações de Möbius e construções com régua e compasso. Isso nos permite, ao final do capítulo, reinterpretar os círculos de Ford como horociclos de um grupo fuchsiano agindo sobre o disco hiperbólico — uma mudança de perspectiva que ilumina por que os círculos de Ford se comportam do jeito que se comportam.

No **Capítulo 5** usamos a geometria dos círculos de Ford para atacar diretamente o problema da melhor aproximação racional: mostramos que todo número irracional admite infinitas aproximações h/k com $|\gamma - h/k| < 1/(2k^2)$, e provamos o Teorema de Hurwitz, que identifica $\sqrt{5}$ como a melhor constante possível nesse tipo de desigualdade.

Por fim, no **Capítulo 6** revisitamos o mesmo problema por um caminho inteiramente diferente, algébrico: as frações contínuas. Construímos os convergentes de uma fração contínua, provamos que eles fornecem as melhores aproximações racionais possíveis (Teorema de Lagrange) e reencontramos, por essa via puramente aritmética, o mesmo limiar $1/(2b^2)$ que havia surgido no Capítulo 5 através da geometria dos círculos de Ford — um encontro que ilustra bem o espírito deste trabalho: um mesmo fenômeno de aproximação, visto de ângulos diferentes.

Capítulo 1

Frações Decimais, Sequências de Farey e Círculos de Ford

Este capítulo baseia-se em conceitos fundamentais sobre a expansão decimal dos números reais e a teoria aritmética elementar. A estruturação das propriedades da função maior inteiro, a decomposição de frações decimais e a análise da função ϕ de Euler utilizam a fundamentação teórica rigorosa presente em SANTOS [10], complementada pelas conexões clássicas combinatórias entre sequências ordenadas e partições geométricas descritas por CONWAY e GUY [5].

1.1 Introdução

Sabemos que existem outros números para lá dos inteiros, como $2/3$, $4/7$, ..., que também se comportam de maneiras muito interessantes. Chamar-lhes-emos simplesmente frações, embora os matemáticos se refiram usualmente a eles como *números racionais*. Algo curioso surge do fato de que um número racional é definido como a razão entre dois números inteiros.

Esperamos que você esteja familiarizado com os usos mais mundanos das frações; por isso, o verdadeiro tema deste capítulo é como as frações podem ser usadas para lançar luz sobre algumas propriedades subtis dos números inteiros — e, ao final, como elas se organizam em duas construções notáveis: as *sequências de Farey* e os *círculos de Ford*.

$$\frac{4}{6} = \frac{2}{3} = \frac{6}{9} = \frac{202}{303} = \dots$$

A regra de ouro para as frações é que se pode multiplicar o numerador e o denominador pelo mesmo número não alteramos o valor da fração. Uma fração sem tal fator comum está na sua *forma irredutível* ($2/3$ no exemplo acima). A regra de ouro permite somar, subtrair, multiplicar e dividir frações:

$$\frac{2}{3} + \frac{1}{4} = \frac{8}{12} + \frac{3}{12} = \frac{11}{12}; \quad \frac{2}{3} \times \frac{1}{4} = \frac{4}{6} \times \frac{1}{4} = \frac{1}{6};$$

$$\frac{2}{3} \div \frac{1}{4} = \frac{2/3}{1/4} = \frac{2}{3} \times \frac{4}{1} = \frac{8}{3}.$$

É claro que as frações surgem por toda a matemática e ciências. De fato, foi esta ubiquidade que convenceu a irmandade pitagórica de que os números governavam o mundo. Uma das descobertas que mais os impressionou foi uma aplicação pouco provável das frações à música: as notas de duas cordas vibrantes semelhantes soam harmoniosas precisamente quando a razão dos seus comprimentos é uma fração simples.

Antes de chegarmos às sequências de Farey e aos círculos de Ford — que são um de nossos destinos neste texto — precisamos construir algumas ferramentas: a função maior inteiro, a representação decimal dos racionais e dos reais, a função ϕ de Euler e a teoria da periodicidade das dízimas. Cada uma dessas ferramentas reaparecerá, mais adiante, na explicação de por que as frações de Farey se comportam tão bem.

1.2 A Função Maior Inteiro

Antes de falarmos em casas decimais, precisamos de uma noção precisa de “parte inteira” de um número, que usaremos repetidamente ao longo do capítulo.

Seja a um número racional. Denotamos por $\lfloor a \rfloor$ o maior inteiro que não ultrapassa a , isto é:

$$\lfloor a \rfloor = \max\{m \in \mathbb{Z} \mid m \leq a\}.$$

Definição 1.1 A função de $\mathbb{Q}$ em $\mathbb{Z}$ definida por $x \rightarrow \lfloor x \rfloor$ chama-se **função maior inteiro** (ou **função piso**) sobre $\mathbb{Q}$.

Proposição 1.2 Sejam m, n inteiros, $n > 0$. Se q é o quociente da divisão euclidiana de m por n , então $q = \left\lfloor \frac{m}{n} \right\rfloor$.

Demonstração: Escrevamos $m = nq + r$, com $0 \leq r < n$. Então

$$\frac{m}{n} = \frac{nq + r}{n} = q + \frac{r}{n}.$$

Como $0 \leq r < n$, temos

$$0 = \frac{0}{n} \leq \frac{r}{n} < \frac{n}{n} = 1,$$

e portanto $\left\lfloor \frac{r}{n} \right\rfloor = 0$. Logo

$$\left\lfloor \frac{m}{n} \right\rfloor = \left\lfloor q + \frac{r}{n} \right\rfloor = q + \left\lfloor \frac{r}{n} \right\rfloor = q.$$

■

Exemplo 1.3 Mostremos que o expoente com que um número primo $p > 0$ aparece como fator de $n!$, para todo $n \geq 1$, é

$$\left\lfloor \frac{n}{p} \right\rfloor + \left\lfloor \frac{n}{p^2} \right\rfloor + \dots$$

Se $p > n$, então p não é fator primo de nenhum dos fatores de $n!$, e portanto o expoente de p em $n!$ é zero. Como, neste caso, $n < p^r$ para todo $r \geq 1$, também

$$\left\lfloor \frac{n}{p} \right\rfloor + \left\lfloor \frac{n}{p^2} \right\rfloor + \dots = 0.$$

Se $p \leq n$, pela Proposição 1.2 o quociente da divisão de n por p é $\left\lfloor \frac{n}{p} \right\rfloor$; logo p é divisor exatamente dos fatores $p, 2p, \dots, \left\lfloor \frac{n}{p} \right\rfloor p$ de $n!$ (e apenas destes), já que $\left\lfloor \frac{n}{p} \right\rfloor p$ é o último múltiplo de p que não supera n . Assim, p divide $\left\lfloor \frac{n}{p} \right\rfloor$ fatores de $n!$. O mesmo argumento aplicado a esses $\left\lfloor \frac{n}{p} \right\rfloor$ fatores mostra que os múltiplos de p^2 entre eles totalizam $\left\lfloor \frac{n}{p^2} \right\rfloor$, e assim por diante. Logo o expoente de p em $n!$ é, efetivamente, a soma dada no enunciado.

Por exemplo, o expoente de 3 em $20!$ é

$$\left\lfloor \frac{20}{3} \right\rfloor + \left\lfloor \frac{20}{9} \right\rfloor + \left\lfloor \frac{20}{27} \right\rfloor + \dots = 6 + 2 + 0 + \dots = 8.$$

1.3 Representação Decimal de Números Racionais

Todo número racional que puder ser escrito sob a forma

$$\frac{r}{10^n}$$

onde $r \in \mathbb{Z}$ e $n \in \mathbb{N}$, chama-se número racional decimal.

Exemplo 1.4

$$\frac{3}{10}, \quad \frac{-19}{100}, \quad \frac{1}{250}, \quad \frac{4}{1000}.$$

Proposição 1.5 Um elemento $a \in \mathbb{Q}$ é um número racional decimal se, e somente se, existem $r \in \mathbb{Z}$ e $\alpha, \beta \in \mathbb{N}$ tais que

$$a = \frac{r}{2^\alpha \cdot 5^\beta}.$$

Demonstração: Suponhamos a conforme o enunciado e, sem perda de generalidade, $\alpha \geq \beta$. Então

$$a = \frac{r}{2^\alpha \cdot 5^\beta} = \frac{r \cdot 5^{\alpha-\beta}}{2^\alpha \cdot 5^\beta \cdot 5^{\alpha-\beta}} = \frac{5^{\alpha-\beta} \cdot r}{10^\alpha},$$

e portanto a é um número racional decimal.

Reciprocamente, se a é racional decimal, existem $r \in \mathbb{Z}$ e $n \in \mathbb{N}^*$ tais que

$$a = \frac{r}{10^n} = \frac{r}{2^n \cdot 5^n},$$

o que encerra a demonstração. ■

Observemos que

$$\frac{12345}{1000} = \frac{10000 + 2000 + 300 + 40 + 5}{1000} = 10 + 2 + \frac{3}{10} + \frac{4}{100} + \frac{5}{1000} = 12 + \frac{3}{10} + \frac{4}{100} + \frac{5}{1000}.$$

Generalizando, seja

$$a = \frac{n}{10^r} \quad (n, r \geq 0)$$

e suponhamos que a representação decimal de n seja $b_1 b_2 \dots b_s a_1 a_2 \dots a_r$ (isto é, $n = a_r + a_{r-1} \cdot 10 + \dots + a_1 \cdot 10^{r-1} + m \cdot 10^r$, onde $m = b_s + \dots + b_1 \cdot 10^{s-1}$ e $a_r, \dots, a_1, b_s, \dots, b_1$ são os algarismos de n). Então:

$$a = \frac{n}{10^r} = m + \frac{a_1}{10} + \frac{a_2}{10^2} + \dots + \frac{a_{r-1}}{10^{r-1}} + \frac{a_r}{10^r} \quad (*)$$

Nessa decomposição, m é natural (sua representação decimal é $b_1 b_2 \dots b_s$), ao passo que as demais parcelas são números racionais decimais não inteiros, pois $0 \leq a_i \leq 9$ ($i = 1, \dots, r$). Fica evidenciada em (*) a importância das frações decimais

$$\frac{1}{10}, \frac{1}{100}, \frac{1}{1000}, \dots$$

em nosso sistema de numeração — daí os nomes especiais que recebem, respectivamente: décimo, centésimo, milésimo, ...

A notação para expressar a é sugestiva em face da decomposição obtida:

$$a = m, a_1 a_2 \dots a_r.$$

Esta é a chamada **representação** (ou **forma**) **decimal** de a . Nela, m é a **parte inteira**, a_1 é a **primeira casa decimal**, a_2 é a **segunda casa decimal**, etc. A vírgula separa a parte inteira da parte decimal. Por exemplo:

$$\frac{12\,345}{1\,000} = 12,345,$$

uma vez que, como vimos, $\frac{12\,345}{1\,000} = 12 + \frac{3}{10} + \frac{4}{100} + \frac{5}{1\,000}$. Lê-se o numeral 12,345 assim: “doze inteiros e trezentos e quarenta e cinco milésimos”. Ou seja, após a leitura do numeral à direita da vírgula lê-se também a fração decimal (de numerador igual a 1) correspondente ao seu último algarismo.

Na representação decimal de a pode ocorrer (mantendo as notações anteriores) que

$$b_1 = \dots = b_s = a_1 = \dots = a_{k-1} = 0 \text{ e } a_k \neq 0, \quad 1 \leq k \leq r.$$

Neste caso, $m = 0$ e $a = 0,0 \dots 0 a_k \dots a_r$. Por exemplo, se

$$a = \frac{123}{10\,000} = \frac{123}{10^4},$$

como 123 tem 3 algarismos e $r = 4$, então $a_4 = 3$, $a_3 = 2$, $a_2 = 1$, $a_1 = 0$, $b_1 = 0$, e portanto $a = 0,0123$.

Se $a < 0$ e $|a| = m, a_1 a_2 \dots a_r$, a representação decimal de a é, por definição, $a = -m, a_1 a_2 \dots a_r = -(m, a_1 a_2 \dots a_r)$. Por exemplo:

$$\frac{-12\,345}{1\,000} = -12,345, \quad \frac{-123}{10\,000} = -0,0123.$$

Em geral, se $a/(2^\alpha 5^\beta)$ é uma fração própria reduzida (a inteiro), ela tem uma expansão decimal finita, porque pode sempre ser transformada numa fração com denominador que é uma potência de 10. Por exemplo, se $\alpha \geq \beta$,

$$\frac{a}{2^\alpha 5^\beta} = \frac{a \cdot 5^{\alpha-\beta}}{2^\alpha \cdot 5^\alpha} = \frac{a \cdot 5^{\alpha-\beta}}{10^\alpha}.$$

Se $\beta > \alpha$, multiplicamos por $2^{\beta-\alpha}$, e há β casas decimais. Por exemplo,

$$\frac{147}{200} = \frac{735}{1000} = 0,735.$$

Reciprocamente, se uma fração tem uma expansão decimal finita, a multiplicação por uma potência adequada de 10 dá um inteiro; na forma reduzida, portanto, o seu denominador contém apenas 2 e 5 como fatores primos. Fica assim provado (juntamente com a Proposição 1.5) que:

Observação 1.6 Uma fração irredutível $\frac{a}{m}$ tem expansão decimal finita se, e somente se, $m = 2^\alpha 5^\beta$ para alguns $\alpha, \beta \geq 0$.

1.4 A Função ϕ de Euler e o Princípio de Inclusão-Exclusão

Antes de tratarmos das frações que *não* têm expansão decimal finita, vale a pena introduzir uma ferramenta combinatória — a função ϕ de Euler — que vai reaparecer repetidas vezes ao longo deste texto: primeiro na periodicidade das dízimas, depois nas sequências de Farey.

O Princípio de Inclusão-Exclusão é uma técnica combinatória útil para contar o número de elementos na união de vários conjuntos finitos. Para dois conjuntos A e B , a cardinalidade da união é

$$|A \cup B| = |A| + |B| - |A \cap B|.$$

Para três conjuntos A , B e C , expandimos a ideia somando as cardinalidades individuais, subtraindo as interseções duplas e readicionando a interseção tripla:

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

De forma geral, para r conjuntos, alternamos a soma e a subtração das interseções de ordem superior. Vamos aplicar este princípio para resolver dois problemas preliminares.

Problema 1.7 Dentre os números de 1 até 3600 inclusive, quantos são divisíveis por 3 ou por 7?

Demonstração: Seja $S = \{1, 2, \dots, 3600\}$. Definimos os subconjuntos:

- $A = \{x \in S \mid 3 \text{ divide } x\}$, com $|A| = \frac{3600}{3} = 1200$;
- $B = \{x \in S \mid 7 \text{ divide } x\}$, com $|B| = \left\lfloor \frac{3600}{7} \right\rfloor = 514$.

O conjunto dos números divisíveis por 3 e por 7 simultaneamente é $A \cap B$, correspondente aos múltiplos de $\text{mmc}(3, 7) = 21$:

$$|A \cap B| = \left\lfloor \frac{3600}{21} \right\rfloor = 171.$$

Pelo Princípio de Inclusão-Exclusão, a quantidade de números divisíveis por 3 ou por 7 é

$$|A \cup B| = |A| + |B| - |A \cap B| = 1200 + 514 - 171 = 1543.$$

■

Problema 1.8 Dado um inteiro positivo m e números $p_1, \dots, p_r$ menores do que ou iguais a m , primos entre si dois a dois, encontrar uma fórmula para o número de inteiros positivos menores do que ou iguais a m que não são divisíveis por nenhum dos $p_1, \dots, p_r$.

Demonstração: Seja $S = \{1, 2, \dots, m\}$. Para cada $i \in \{1, \dots, r\}$, seja A_i o conjunto dos elementos de S divisíveis por p_i . Como os p_i são primos entre si, $|A_i| = \frac{m}{p_i}$, $|A_i \cap A_j| = \frac{m}{p_i p_j}$, e assim por diante.

Queremos o número de elementos que não pertencem a nenhum dos A_i , isto é, a cardinalidade do complementar da união:

$$\left| \bigcap_{i=1}^r A_i^c \right| = |S| - |A_1 \cup \dots \cup A_r|.$$

Pelo Princípio de Inclusão-Exclusão,

$$|S| - \sum_{i=1}^r |A_i| + \sum_{1 \leq i < j \leq r} |A_i \cap A_j| - \dots + (-1)^r |A_1 \cap \dots \cap A_r|,$$

que, substituindo as cardinalidades, é

$$m - \sum_{i=1}^r \frac{m}{p_i} + \sum_{1 \leq i < j \leq r} \frac{m}{p_i p_j} - \dots + (-1)^r \frac{m}{p_1 \dots p_r}.$$

Colocando m em evidência, a expressão fatora-se elegantemente como

$$m \left(1 - \frac{1}{p_1}\right) \left(1 - \frac{1}{p_2}\right) \dots \left(1 - \frac{1}{p_r}\right).$$

■

1.4.1 A Função ϕ de Euler

Definição 1.9 A função ϕ de Euler (ou função indicatriz de Euler) associa a cada inteiro positivo m a quantidade de inteiros positivos menores ou iguais a m que são relativamente primos com m (isto é, $\text{MDC}(x, m) = 1$). Equivalentemente, $\phi(n)$ conta, dentre as frações

$$\frac{0}{n}, \frac{1}{n}, \frac{2}{n}, \dots, \frac{n-1}{n},$$

quantas têm n como menor denominador possível, isto é, quantas já estão em forma irredutível.

Utilizando o resultado do Problema 1.8, podemos enunciar e provar o teorema fundamental para o cálculo de $\phi(m)$.

Teorema 1.10 *Seja $m > 1$ um inteiro cuja decomposição em fatores primos é $m = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_r^{\alpha_r}$. Então*

$$\phi(m) = m \left(1 - \frac{1}{p_1}\right) \left(1 - \frac{1}{p_2}\right) \dots \left(1 - \frac{1}{p_r}\right).$$

Demonstração: Um número $x \leq m$ é relativamente primo com m se, e somente se, não compartilha nenhum fator primo com m , ou seja, não é divisível por nenhum dos $p_1, \dots, p_r$. Como os fatores primos de um número são primos entre si dois a dois, isto recai exatamente sobre as hipóteses do Problema 1.8; a aplicação direta da fórmula lá deduzida encerra a demonstração. ■

Exemplo 1.11 (Cálculo de $\phi(100)$) *A fatoração prima de 100 é $100 = 2^2 \times 5^2$; os fatores primos distintos são $p_1 = 2$ e $p_2 = 5$. Pelo Teorema 1.10,*

$$\phi(100) = 100 \left(1 - \frac{1}{2}\right) \left(1 - \frac{1}{5}\right) = 100 \times \frac{1}{2} \times \frac{4}{5} = 40.$$

Portanto, existem 40 números menores ou iguais a 100 que são primos com 100. Equivalentemente: numa fração com denominador 100, ela só se simplifica se o numerador for divisível por 2 ou por 5. Metade dos numeradores são ímpares, e 4 em cada 5 do restante não são divisíveis por 5; logo $\phi(100) = 100 \times \frac{1}{2} \times \frac{4}{5} = 40$, confirmando o cálculo.

1.4.2 Abordagem Alternativa: Dedução por Somas de Divisores

Outra forma elegante de calcular os valores de $\phi(m)$ baseia-se numa identidade aritmética fundamental sobre os divisores de um número:

$$\sum_{d|m} \phi(d) = m,$$

onde a soma é feita sobre todos os divisores positivos d de m — ou seja, *qualquer número inteiro é a soma dos totientes dos seus divisores*. Essa relação permite obter os valores de $\phi(m)$ de forma indutiva a partir dos valores conhecidos dos divisores menores. Partindo de $\phi(1) = 1$:

- $m = 2$, divisores 1, 2: $\phi(1) + \phi(2) = 2 \implies \phi(2) = 1$.
- $m = 3$, divisores 1, 3: $\phi(1) + \phi(3) = 3 \implies \phi(3) = 2$.
- $m = 4$, divisores 1, 2, 4: $\phi(1) + \phi(2) + \phi(4) = 4 \implies \phi(4) = 2$.
- $m = 6$, divisores 1, 2, 3, 6: $\phi(1) + \phi(2) + \phi(3) + \phi(6) = 6 \implies \phi(6) = 2$.
- $m = 12$, divisores 1, 2, 3, 4, 6, 12: $\phi(1) + \phi(2) + \phi(3) + \phi(4) + \phi(6) + \phi(12) = 12 \implies \phi(12) = 4$, o que também confere com o Teorema 1.10: $\phi(12) = 12 \left(1 - \frac{1}{2}\right) \left(1 - \frac{1}{3}\right) = 4$.

Essa identidade dá uma forma alternativa, sem recorrer diretamente à fatoração prima ou ao Princípio de Inclusão-Exclusão, para determinar os valores de ϕ passo a passo — e reaparecerá, na Subseção ??, ligada ao comprimento das sequências de Farey.

1.5 Séries Infinitas e a Representação Decimal dos Números Reais

Até aqui tratamos apenas de frações com expansão decimal *finita*. Para tratar frações como $1/3$ ou $1/7$, e mais geralmente qualquer número real, precisamos de séries infinitas.

Uma série infinita

$$\sum_{n=1}^{\infty} aq^{n-1} = a + aq + aq^2 + \dots$$

onde a e q são números reais dados, com $a \neq 0$ e $q \neq 0$, chama-se série geométrica.

Proposição 1.12 *Seja $q \in \mathbb{R}$ tal que $0 < q < 1$. Então a série geométrica $a + aq + aq^2 + \dots$ converge, qualquer que seja $a \in \mathbb{R}$, e sua soma é*

$$s = \frac{a}{1-q}.$$

Demonstração: Para todo índice n , a n -ésima soma parcial é

$$s_n = a + aq + \dots + aq^{n-1}.$$

Daí $qs_n = aq + aq^2 + \dots + aq^n$, e subtraindo:

$$s_n - qs_n = a - aq^n = a(1 - q^n),$$

donde

$$s_n = \frac{a(1 - q^n)}{1 - q} = \frac{a}{1 - q} + \frac{-a}{1 - q} q^n.$$

Como $0 < q < 1$, a sequência $\left(\frac{-a}{1-q} q^n\right)$ converge para 0; logo (s_n) converge, e

$$\lim s_n = \frac{a}{1 - q}.$$

■

Seja m um inteiro positivo qualquer e consideremos inteiros $a_1, a_2, \dots$ tais que $0 \leq a_i \leq 9$ ($i = 1, 2, \dots$). Pela Proposição 1.12 (aplicada termo a termo, com razão $1/10$), a série

$$m + \frac{a_1}{10} + \frac{a_2}{10^2} + \dots$$

converge em $\mathbb{R}$. A soma dessa série é um número real α , denotado

$$\alpha = m, a_1 a_2 a_3 \dots,$$

chamado representação (ou forma) decimal de α .

Se para todo índice r existe um índice $s \geq r$ tal que $a_s \neq 0$, a representação diz-se **infinita**. Por exemplo, a representação decimal da soma de

$$1 + \frac{0}{10} + \frac{2}{10^2} + \frac{0}{10^3} + \frac{2}{10^4} + \dots$$

é infinita: 1,020202...

Caso contrário, existe um índice r tal que $a_r = a_{r+1} = \dots = 0$; a representação é então **finita**, coincidindo com o caso já estudado na seção anterior:

$$\alpha = m \quad \text{ou} \quad \alpha = m, a_1 a_2 \dots a_{r-1}.$$

Teorema 1.13 *Se a é um número real positivo, então existem inteiros $m \geq 0$ e $a_1, a_2, \dots$, com $0 \leq a_i \leq 9$, tais que*

$$a = m, a_1 a_2 a_3 \dots$$

Demonstração: Seja $m = \lfloor a \rfloor$. Então $m \leq a < m + 1$. Se $a = m$, a demonstração termina. Caso contrário, $m < a < m + 1$; consideremos o maior dos números

$$m, m + \frac{1}{10}, m + \frac{2}{10}, \dots, m + \frac{9}{10}$$

que não supera a . Se $m + \frac{a_1}{10}$ é esse número, então

$$m + \frac{a_1}{10} \leq a < m + \frac{a_1 + 1}{10} =: \alpha'_1.$$

Se $a = \alpha'_1$, então $a = m, a_1$ e terminamos. Caso contrário, repetimos o raciocínio: tomamos o maior dos números $m + \frac{a_1}{10} + \frac{k}{100}$ ($k = 0, \dots, 9$) que não supera a , obtendo a_2 , e assim sucessivamente.

Se, para algum r , $a = m, a_1 a_2 \dots a_r$, a demonstração termina. Se isto nunca ocorre, consideremos os conjuntos

$$S = \{\alpha_1, \alpha_2, \dots\} \quad \text{e} \quad T = \{\alpha'_1, \alpha'_2, \dots\},$$

onde $\alpha_r = m + \frac{a_1}{10} + \dots + \frac{a_r}{10^r}$ e $\alpha'_r = \alpha_r + \frac{1}{10^r}$. Todo elemento de S é menor que todo elemento de T , e dado $\varepsilon > 0$, tomando r tal que $\frac{1}{10^r} < \varepsilon$, temos

$$\alpha'_r - \alpha_r = \frac{1}{10^r} < \varepsilon.$$

Como $\alpha_r \leq a < \alpha'_r$, segue $a - \alpha_r < \varepsilon$; e como para todo $n \geq r$ vale $\alpha_r \leq \alpha_n \leq a$, obtemos

$$|a - \alpha_n| = a - \alpha_n \leq a - \alpha_r < \varepsilon, \quad \text{para todo } n \geq r.$$

Logo $\alpha_n \rightarrow a$, isto é,

$$m + \frac{a_1}{10} + \frac{a_2}{10^2} + \dots = a,$$

o que permite escrever $a = m, a_1 a_2 a_3 \dots$ ■

Na representação decimal de $\alpha \geq 0$, m é sua parte inteira, a_1 a primeira casa decimal, a_2 a segunda, e assim por diante. Se $\alpha < 0$ e $|\alpha| = m, a_1 a_2 \dots$ (finita ou infinita), define-se $\alpha = -m, a_1 a_2 \dots = -(m, a_1 a_2 \dots)$. Assim, todo número real α admite uma representação decimal

$$m, a_1 a_2 a_3 \dots \quad \text{ou} \quad -m, a_1 a_2 a_3 \dots,$$

finita ou infinita, com $m, a_1, a_2, \dots \in \mathbb{N}$ e $0 \leq a_i \leq 9$.

Exemplo 1.14 (A expansão decimal de $\sqrt{3}$) Vamos obter a representação decimal de $\alpha = \sqrt{3}$, até a terceira casa decimal, usando o procedimento da demonstração do Teorema 1.13.

Como $m = \lfloor \sqrt{3} \rfloor = 1$, procuramos, na sequência

$$1; 1,1; 1,2; \dots; 1,9,$$

o maior número que não supera $\sqrt{3}$. Como $(1,7)^2 = 2,89$ e $(1,8)^2 = 3,24$, esse número é 1,7: $a_1 = 7$.

Em seguida, na sequência $1,7; 1,71; \dots; 1,79$, procuramos o maior número que não supera $\sqrt{3}$. Como $(1,73)^2 = 2,9929$ e $(1,74)^2 = 3,0276$, esse número é 1,73: $a_2 = 3$.

Por fim, em $1,731; 1,732; \dots; 1,739$, observando que $(1,732)^2 = 2,999824$ e $(1,733)^2 = 3,003289$, a resposta é 1,732: $a_3 = 2$. Logo

$$\sqrt{3} = 1,732\dots$$

Geometricamente: dividimos $[1, 2]$ em dez subintervalos iguais, numerados de 0 a 9:

	0	1	2	3	4	5	6	7	8	9	
	1	1,1	1,2	1,3	1,4	1,5	1,6	1,7	1,8	1,9	(2)

Como $1,7 < \sqrt{3} < 1,8$, $\sqrt{3}$ pertence ao subintervalo 7, e $\sqrt{3} = 1,7\dots$. Dividindo agora $[1,7; 1,8]$ em dez subintervalos iguais:

	0	1	2	3	4	5	6	7	8	9	
	1,7	1,71	1,72	1,73	1,74	1,75	1,76	1,77	1,78	1,79	(1,8)

Como $1,73 < \sqrt{3} < 1,74$, $\sqrt{3}$ pertence ao subintervalo 3, e $\sqrt{3} = 1,73\dots$, e assim por diante.

Exemplo 1.15 (A expansão de $\sqrt{3}$ na base 2) Assim como chegamos à representação decimal de $\alpha \in \mathbb{R}$, poderíamos chegar à sua representação binária, subdividindo intervalos em dois em vez de dez. Vejamos isso para $\sqrt{3} = \sqrt{(11)_2}$.

A parte inteira de $\sqrt{(11)_2}$ é 1. Dividimos $[1, 2]$ em dois subintervalos de mesma amplitude, atribuindo os dígitos 0 e 1 em ordem. Como $\sqrt{3}$ pertence ao segundo subintervalo (o de $[1,5; 2]$, já que $(1,5)^2 = 2,25 < 3$), então

$$\sqrt{(11)_2} = 1,1\dots$$

Agora subdividimos $\left[1 + \frac{1}{2}, 2\right]$ em dois, com numerais 0 e 1. Como $\sqrt{3}$ pertence ao primeiro desses subintervalos (pois $(1,75)^2 = 3,0625 > 3$), então

$$\sqrt{(11)_2} = 1,10\dots$$

Continuando este processo bit a bit — em vez de dígito a dígito — obtemos a expansão binária de $\sqrt{3}$, da mesma forma que obtivemos sua expansão decimal.

1.6 Periodicidade das Expansões Decimais

De maneira mais formal, qualquer número racional $\frac{a}{b}$ pode ser convertido em fração decimal através do algoritmo clássico da divisão. Nosso objetivo agora é estudar a estrutura aritmética dessas expansões e compreender a periodicidade das dízimas a partir da teoria das congruências — e, com isso, obter uma nova demonstração do teorema de Euler-Fermat.

Antes de prosseguirmos, vale a pena fixar a linguagem das *congruências*, que usaremos repetidamente a partir daqui: ela nada mais é do que uma forma cômoda de falar sobre restos de divisão.

Definição 1.16 *Dados dois inteiros a e b , e um inteiro positivo m , dizemos que a é **côngruo a b módulo m** , e escrevemos*

$$a \equiv b \pmod{m},$$

quando a e b deixam o mesmo resto na divisão euclidiana por m — ou, equivalentemente, quando m divide a diferença $a - b$.

Por exemplo, $17 \equiv 2 \pmod{5}$, pois tanto 17 quanto 2 deixam resto 2 na divisão por 5 (e, de fato, $5 \mid 15$). Assim como uma igualdade, uma congruência pode ser somada, subtraída e multiplicada, em ambos os lados, por um mesmo inteiro, ou por outra congruência de mesmo módulo, sem que a relação deixe de valer — é essa flexibilidade algébrica, aliada à sua leitura direta em termos de restos, que torna a notação $\pmod{m}$ tão conveniente para o que segue.

Suponhamos, sem perda de generalidade, que trabalhamos com uma fração própria irredutível $\frac{a}{m}$ em $(0, 1)$, de modo que $0 < a < m$ e $\text{MDC}(a, m) = 1$.

Tomemos a fração $\frac{1}{7}$. O processo de divisão euclidiana se desenvolve assim:

1,0000000 ...	7
−7	0,1428571 ...
30	
−28	
20	
−14	
60	
−56	
40	
−35	
50	
−49	
10	

No último passo reaparece o resto original 1, o que implica que a sequência de restos parciais $\{3, 2, 6, 4, 5, 1\}$ começa a se repetir indefinidamente:

$$\frac{1}{7} = 0,142857142857 \dots = 0,\overline{142857}.$$

É um fato familiar que qualquer decimal terminante ou periódico é a expansão decimal de um número racional; é o problema inverso que nos interessa agora. Já

vimos exemplos de decimais finitos, como $0,2 = \frac{1}{5}$ e $0,025 = \frac{1}{40}$. Mas $\frac{1}{3} = 0,333 \dots$ é diferente:

1,0000000...	3
-9	0,3333...
10	
-9	
10	
-9	
10	

Dizemos que estes decimais são **periódicos**: $\frac{1}{3}$ tem período de comprimento 1, e $\frac{1}{7}$ tem período de comprimento 6. Em vez de escrever uma infinidade de algarismos, desenhamos uma linha sobre o período:

$$\frac{1}{3} = 0, \overline{3}, \quad \frac{1}{7} = 0, \overline{142857}.$$

Já $\frac{1}{6} = 0,1\overline{6}$ não inicia o seu período logo na vírgula. Os decimais periódicos que iniciam o período na vírgula, como $\frac{1}{3}$ e $\frac{1}{7}$, chamam-se **periódicos puros**; o caso de $\frac{1}{6}$ é dito *misto*, eles aparecem pois o denominador possui em sua fatoração os primos 2 e 5. Não teremos muito interesse neste caso — concentramo-nos nos dois casos anteriores (finito e periódico puro).

1.6.1 Decimais Periódicos Puros

Mostraremos que, se A/m é uma fração própria reduzida com $\text{MDC}(10, m) = 1$, então A/m tem expansão decimal periódica pura.

Quando dividimos A por m , $\frac{A}{m} = 0.q_1q_2q_3 \dots$, e em geral

$$r_{j+1} = 10r_j - q_j m, \quad 0 \leq q_j < 10, \quad 0 < r_j < m.$$

Todos os restos r_j são primos com m : por indução, $r_1 = A$ é primo com m ; e como $r_{j+1} \equiv 10r_j \pmod{m}$, se $\text{MDC}(r_j, m) = 1$ e $\text{MDC}(10, m) = 1$, então $\text{MDC}(r_{j+1}, m) = 1$.

Como há infinitos r 's, todos entre 1 e $m-1$, dois deles devem coincidir. Sejam r_j e r_{j+l} os primeiros iguais. Se $j > 1$, então r_{j-1} está definido, e

$$10r_{j-1} \equiv r_j = r_{j+l} \equiv 10r_{j+l-1} \pmod{m} \implies r_{j-1} \equiv r_{j+l-1} \pmod{m}$$

(usando $\text{MDC}(m, 10) = 1$); como r_{j-1} e r_{j+l-1} estão entre 1 e $m-1$, são pequenos demais para diferir por um múltiplo de m sem serem iguais — contradizendo a minimalidade de j . Logo $j = 1$: a sequência de restos, e portanto a sequência de dígitos q_j , é puramente periódica, com período l .

Seja $\lambda(m)$ o comprimento do período da expansão decimal de $1/m$ (mostraremos adiante que toda fração própria reduzida A/m tem o mesmo período $\lambda(m)$). Temos $\lambda(m) \leq \phi(m)$, pois há apenas $\phi(m)$ classes de resíduos primas com m , e pelo menos dois de $r_1, \dots, r_{\phi(m)+1}$ devem coincidir.

Às vezes $\lambda(m) = \phi(m)$: por exemplo, $\lambda(7) = \phi(7) = 6$ e $\lambda(17) = \phi(17) = 16$, como já vimos: $\frac{1}{17} = 0, \overline{0588235294117647}$. Mas $\frac{1}{3} = 0, \overline{3}$ dá $\lambda(3) = 1$, enquanto $\phi(3) = 2$; e

$\frac{1}{21} = 0, \overline{047619}$ dá $\lambda(21) = 6$, enquanto

$$\phi(21) = 21 \left(1 - \frac{1}{3}\right) \left(1 - \frac{1}{7}\right) = 12.$$

Mostra-se que, se m tem dois ou mais divisores primos distintos, então $\lambda(m) \leq \frac{1}{2}\phi(m)$; assim, para atingir o período máximo $\phi(m)$ é necessário (mas não suficiente, como vimos para $p = 3$) que m seja potência de um primo. Não se sabe sequer se há infinitos primos p para os quais o período máximo $\phi(p) = p - 1$ é atingido (voltaremos a isto na Seção ??, sobre primos longos).

Se $\frac{A}{m} = \frac{r_1}{m} = 0, \overline{q_1 q_2 \dots q_l}$, então

$$\frac{r_j}{m} = 0, \overline{q_j q_{j+1} \dots q_l q_1 \dots q_{j-1}},$$

pois começar com $A = r_j$ em vez de $A = r_1$ segue o mesmo procedimento, obtendo os mesmos q 's e r 's em ordem cíclica, apenas começando o período num lugar diferente. Por exemplo:

$$\begin{array}{ll} 1/7 = 0, \overline{142857} & 3/7 = 0, \overline{428571} \\ 2/7 = 0, \overline{285714} & 6/7 = 0, \overline{857142} \\ 4/7 = 0, \overline{571428} & 5/7 = 0, \overline{714285} \end{array}$$

todos gerados pelo mesmo ciclo de restos 1, 3, 2, 6, 4, 5.

Vejamos agora $1/41$:

$$1/41 = 0, \overline{02439},$$

com restos 1, 10, 18, 16, 37. Note que

$$10 \equiv 10^1, \quad 18 \equiv 10^2, \quad 16 \equiv 10^3, \quad 37 \equiv 10^4, \quad 1 \equiv 10^5 \pmod{41},$$

isto é, todos esses restos são congruentes a potências de 10 módulo 41.

Teorema 1.17 *As classes de resíduos (mod 41) contendo 1, 10, 18, 16, 37 formam um grupo multiplicativo.*

Demonstração: Se $r \equiv 10^k$ e $s \equiv 10^l \pmod{41}$, então $rs \equiv 10^{k+l} \pmod{41}$, também uma potência de 10. Se $r \equiv 10^k \pmod{41}$, $k = 1, \dots, 5$, então $s \equiv 10^{6-k} \pmod{41}$ é o inverso de r , pois $rs \equiv 10^6 \equiv 1 \pmod{41}$. Logo as classes de 1, 10, 18, 16, 37 formam um subgrupo do grupo de classes primas com 41. ■

Não há magia no número 41: em geral, se os restos na expansão de $1/m$ são $r_1, \dots, r_\lambda$, então, como $r_1 = 1$ e $r_{j+1} \equiv 10r_j \pmod{m}$, segue que

$$r_j \equiv 10^{j-1} \pmod{m}, \quad j = 1, 2, \dots$$

Como os r_j ($j = 1, \dots, \lambda$) são incongruentes módulo m , as potências 10^j também o são; além disso $r_{\lambda+1} = r_1 = 1$, logo $10^\lambda \equiv 1 \pmod{m}$ é a menor potência de 10 congruente a 1 módulo m . As classes contendo $r_1, \dots, r_\lambda$ coincidem com as contendo $10^0, \dots, 10^{\lambda-1}$, formando um subgrupo do grupo de classes primas com m — de ordem $\phi(m)$. Pelo teorema de Lagrange (elementos da teoria de grupos), a ordem desse subgrupo divide $\phi(m)$:

Teorema 1.18 $\lambda(m)$ divide $\phi(m)$.

Prova alternativa, sem recorrer à teoria de grupos: Consideremos $m = 41$. Como 1, 10, 18, 16, 37 são os restos de $1/41$, os restos de $2/41$ seriam 2, 20, 36, 32, 74, exceto que $74 > 41$ deve ser reduzido módulo 41, dando 2, 20, 36, 32, 33 — precisamente as classes $\equiv 2 \cdot 10^k \pmod{41}$. Analogamente, 3, 30, 13, 7, 29 são os restos de $3/41$. Assim, qualquer fração própria reduzida $A/41$ gera apenas cinco restos, e o período de qualquer $A/41$ é o mesmo de $1/41$. Os $\phi(41)$ números primos com 41 dividem-se, desta forma, em conjuntos disjuntos de $\lambda(41)$ elementos; se há k conjuntos, $k \cdot \lambda(41) = \phi(41)$, e portanto $\lambda(41) \mid \phi(41)$. O argumento geral é idêntico. ■

Como $10^{\lambda(m)} \equiv 1 \pmod{m}$, temos $m \mid (10^{\lambda(m)} - 1)$. Se $\phi(m) = k\lambda(m)$, usando $x^k - 1 = (x - 1)(x^{k-1} + \dots + 1)$ com $x = 10^{\lambda(m)}$, vemos que $m \mid (10^{\phi(m)} - 1)$, isto é, $10^{\phi(m)} \equiv 1 \pmod{m}$. Provamos assim um caso especial do teorema de Fermat-Euler:

Se $\text{MDC}(m, 10) = 1$, então $10^{\phi(m)} \equiv 1 \pmod{m}$.

Em particular, se m é primo $p \neq 2, 5$, $10^{p-1} \equiv 1 \pmod{p}$ — um caso particular do Pequeno Teorema de Fermat, ao qual voltaremos na Seção ???. A importância de 10 nestes enunciados vem apenas de trabalharmos na base 10; trabalhando numa base q qualquer, obtemos o teorema completo:

Teorema 1.19 (Fermat-Euler) Se $\text{MDC}(q, m) = 1$, então $q^{\phi(m)} \equiv 1 \pmod{m}$.

1.7 Grafos de Divisibilidade

A periodicidade que acabamos de estudar admite uma bela visualização. Fixado um divisor m , podemos construir um grafo cujos nós representam não os números em si, mas os *restos* da divisão por m , e cujas movimentações determinam o resto da divisão de um número n por m .

Para construir o grafo, distribuimos os restos $0, 1, \dots, m-1$ em um círculo. As arestas orientadas são determinadas da seguinte forma:

Criamos as setas internas representando a operação “multiplicar o resto acumulado por 10 e escolher o próximo resto módulo m ”, seguindo $r_{j+1} \equiv 10r_j \pmod{m}$. Como vimos, em algum momento esse ciclo se fechará.

Exemplo 1.20 Se $m = 7$, os seus restos são: 0, 1, 2, 3, 4, 5, 6.

- $0 \equiv 10 \cdot 0 \pmod{7}$;
- $3 \equiv 10 \cdot 1 \pmod{7}$; $2 \equiv 10 \cdot 3 \pmod{7}$; $6 \equiv 10 \cdot 2 \pmod{7}$;
- $4 \equiv 10 \cdot 6 \pmod{7}$; $5 \equiv 10 \cdot 4 \pmod{7}$; $1 \equiv 10 \cdot 5 \pmod{7}$;

Podemos criar o grafo seguindo o caminho $0 \rightarrow 0$ e $1 \rightarrow 3 \rightarrow 2 \rightarrow 6 \rightarrow 4 \rightarrow 5 \rightarrow 1$. Veja a Figura 1.1.

Este grafo é, portanto, apenas a mesma recorrência $r_{j+1} = 10r_j - q_j m$ da seção anterior, desta vez lida ao contrário: em vez de extrair dígitos de um resto, construímos o resto a partir dos dígitos.

Agora vamos utilizar este grafo para dividir um número por m .

Números de mais dígitos. A animação de percorrer caminhos só ocorre de fato se n tiver mais de um dígito (por exemplo: 10, 45, 132). A regra de movimentação, para cada novo dígito lido da esquerda para a direita, segue os dois passos que já vimos algebricamente na seção anterior:

1. Andar pelas setas internas: “multiplicar o resto acumulado por 10” — a mesma operação $r_{j+1} \equiv 10r_j \pmod{m}$ de antes, ao avançarmos uma casa decimal.
2. Andar pelo círculo (bordas): “somar o valor do próximo dígito”.

Exemplo 1.21 Se $m = 7$ e $n = 4523$, seguimos os seguintes passos:

- **Passo 1:** Comece no dígito 4. Como o resto de 4 por 7 é 4, a bolinha começa no vértice 4.
- **Passo 2:** Multiplicamos o resto acumulado por 10: $10 \times 4 \equiv 5 \pmod{7}$. A bolinha segue a seta interna correspondente até o vértice 5.
- **Passo 3:** Agora use o segundo dígito, que é o 5, e ande sobre o círculo no sentido horário 5 unidades, terminando no vértice 3.
- **Passo 4:** Ande pelo grafo internamente até o vértice 2 (pois $10 \times 3 \equiv 2 \pmod{7}$).
- **Passo 5:** O próximo dígito é 2. Fazendo 2 saltos pela borda do círculo, chegamos ao vértice 4.
- **Passo 6:** Depois de ler 452, o resto acumulado é 4.
- **Passo 7:** Ande internamente até o vértice 5 (pois $10 \times 4 \equiv 5 \pmod{7}$).
- **Passo 8:** O próximo dígito é 3. Andando 3 unidades pela borda, chegamos ao vértice 1.

Portanto, o resto da divisão de 4523 por 7 é 1. Veja que $4523 = 7 \times 646 + 1$.

Para uma animação do processo explicado acima, acesse a página: <https://www.im-uff.mat.br/javascript/gd.html>

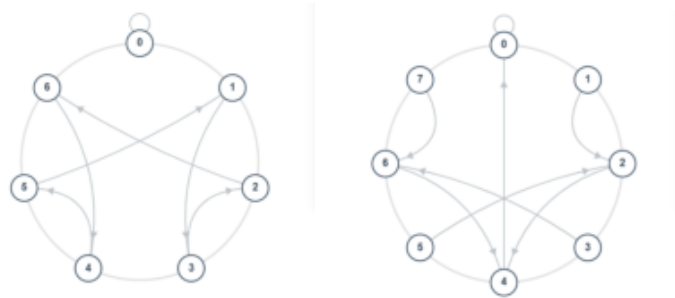


Figura 1.1: Grafo de divisibilidade do 7 e do 8

Capítulo 2

Sequências de Farey

Dizemos que os inteiros p e q são primos entre si, se $\text{MDC}(p, q) = 1$.

Vamos construir o conjunto de todas as frações p/q não negativas com $\text{MDC}(p, q) = 1$, chamado de sequência de Farey. Para isso vamos definir uma nova operação entre números racionais p/q e r/s , por

$$\frac{m}{n} \oplus \frac{r}{s} = \frac{m+r}{n+s}$$

Observe que

$$\frac{m}{n} < \frac{m+r}{n+s} < \frac{r}{s}.$$

Vamos inserir $\frac{m+r}{n+s}$ entre as duas frações adjacentes $\frac{m}{n}$ e $\frac{r}{s}$. Iniciaremos com as frações $\frac{0}{1}$ e $\frac{1}{1}$,

$$\frac{0}{1}, \frac{1}{2}, \frac{1}{1};$$

e o próximo será:

$$\frac{0}{1}, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{1}{1};$$

e nos próximos devemos introduzir 4 termos e depois 8... e assim por diante. Mas vamos adicionar uma imposição: Na próxima etapa o denominador máximo só pode ser 4. Portanto,

$$\frac{0}{1}, \frac{1}{4}, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \frac{1}{1};$$

e as frações $\frac{2}{5}$ e $\frac{3}{5}$ não são acrescentadas nesta etapa, mas claramente serão na próxima.

Por que essa construção funciona? Por que, por exemplo, cada mediant fração $(r+r')/(s+s')$ acaba aparecendo nos termos mais baixos? (Se r, r', s e s' fossem todos ímpares, obteríamos par/par; de alguma forma a construção garante que frações com numeradores e denominadores ímpares nunca aparecem.) E por que todas as frações possíveis r/s ocorrem exatamente uma vez? Por que uma fração específica não pode ocorrer duas vezes, ou não ocorrer? Todas essas perguntas têm respostas incrivelmente simples, baseadas no seguinte fato fundamental: Se r/s e r'/s' são frações consecutivas em qualquer fase de construção, temos

$$r's - rs' = 1, \tag{1}$$

ou escrito de outra forma

$$\det \begin{bmatrix} s & s' \\ r & r' \end{bmatrix} = 1.$$

No início esta relação é trivialmente verdadeira no $(1 \cdot 1 - 0 \cdot 0 = 1)$ e quando introduzimos uma nova mediana $(r + r')/(s + s')$, precisamos checar que

$$(r + r')s - r(s + s') = 1 \text{ e } r'(s + s') - (r + r')s' = 1$$

Mas ambas as equações são equivalentes a condição inicial. Portanto 1 é invariante em todos os estágios da construção.

Uma questão ainda permanece em aberto. É possível que alguma fração a/b com $\text{MDC}(a, b) = 1$ não apareça? A resposta é não, porque nós podemos nos confinar a construção às vizinhanças imediatas de a/b , e nesta região isso é fácil de se analisar: Inicialmente temos

$$\frac{0}{1} = \frac{r}{s} < \left\{ \frac{a}{b} \right\} < \frac{r'}{s'} = \frac{1}{0'}$$

as chaves ao redor $\frac{a}{b}$ indica que isto ainda não foi verificado. Então em algum estágio

$$\frac{r}{s} < \left\{ \frac{a}{b} \right\} < \frac{r'}{s'},$$

e posso construir $(r + r')/(s + s')$ e temos 3 casos para considerar. Ou $(r + r')/(s + s') = a/b$ e já consegui, ou $(r + r')/(s + s') < a/b$ e faço $r \leftarrow r + r', s \leftarrow s + s'$; ou $(r + r')/(s + s') > a/b$ e, neste caso, faço $r' \leftarrow r + r', s' \leftarrow s + s'$. Este processo não pode continuar indefinidamente por causa da condição

$$\frac{a}{b} - \frac{r}{s} > 0 \text{ e } \frac{r'}{s'} - \frac{a}{b} > 0$$

implica que

$$as - br \geq 1 \text{ e } br' - as' \geq 1,$$

daí

$$(r' + s')(as - br) + (r + s)(br' - as') \geq r' + s' + r + s$$

e daí, é o mesmo que $a + b \geq r' + s' + r + s$ por 1. Desde que ou r e s ou r' e s' aumenta em cada uma das etapas, então deveremos alcançar o objetivo em no máximo $a + b$ passos.

A partir desta construção vamos construir a série de Farey, denotado por F_n , é o conjunto de todas as frações reduzidas entre 0 e 1 cujos denominadores são $\leq N$, arranjos de forma crescente. Por exemplo,

$$\begin{array}{c} 0 \quad 1 \\ \frac{1}{1} \quad \frac{1}{1} \\ 0 \quad 1 \quad 1 \\ \frac{1}{1} \quad \frac{1}{2} \quad \frac{1}{1} \\ 0 \quad 1 \quad 1 \quad 2 \quad 1 \\ \frac{1}{1} \quad \frac{1}{3} \quad \frac{1}{2} \quad \frac{1}{3} \quad \frac{1}{1} \\ 0 \quad 1 \quad 1 \quad 1 \quad 2 \quad 3 \quad 1 \\ \frac{1}{1} \quad \frac{1}{4} \quad \frac{1}{3} \quad \frac{1}{2} \quad \frac{1}{3} \quad \frac{1}{4} \quad \frac{1}{1} \\ 0 \quad 1 \quad 1 \quad 1 \quad 2 \quad 1 \quad 3 \quad 2 \quad 3 \quad 4 \quad 1 \\ \frac{1}{1} \quad \frac{1}{5} \quad \frac{1}{4} \quad \frac{1}{3} \quad \frac{1}{5} \quad \frac{1}{2} \quad \frac{1}{5} \quad \frac{1}{3} \quad \frac{1}{4} \quad \frac{1}{5} \quad \frac{1}{1} \end{array}$$

e portanto, o F_6 se torna

$$\frac{0}{1}, \frac{1}{6}, \frac{1}{5}, \frac{1}{4}, \frac{1}{3}, \frac{2}{5}, \frac{1}{2}, \frac{3}{5}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}, \frac{5}{6}, \frac{1}{1}.$$

Em geral, posso obter F_N , começando F_1 e, em seguida, inserindo medianas alcançar F_N . Posso fazer isso desde que o denominador não seja muito grande. Não se perde nenhuma fração dessa maneira, porque sabemos que a construção não perde, e porque uma mediana com denominador N nunca é formado a partir de uma fração cujo denominador é $> N$.

O método descrito revela que F_N pode ser obtido de maneira simples de F_{N-1} ; simplesmente introduzimos as frações $(r + r')/N$ entre frações consecutivas r/s e r'/s' de F_{N-1} cujas somas dos denominadores dão N . Por exemplo, para obter F_7 de F_6 , basta inserir, $\frac{1}{7}, \frac{1}{7}, \dots, \frac{6}{7}$ de acordo com a regra obtemos

$$F_7 = \frac{0}{1}, \frac{1}{7}, \frac{1}{6}, \frac{1}{5}, \frac{1}{4}, \frac{2}{7}, \frac{1}{3}, \frac{2}{5}, \frac{3}{7}, \frac{4}{7}, \frac{1}{2}, \frac{4}{7}, \frac{3}{5}, \frac{2}{3}, \frac{5}{7}, \frac{3}{4}, \frac{4}{5}, \frac{5}{6}, \frac{1}{7}, \frac{1}{1}.$$

Quando N é primo, $N - 1$ novas frações são inseridas, mas quando N não for primo muito menos frações são inseridas, pois este processo só entra numeradores que são relativamente primos com N .

Vamos dar um salto e tentar prever quais as frações que aparecem em F_{12} . Considere as frações $0/12, 1/12, \dots, 11/12$ de acordo com o seu denominador após simplificação:

Denominador após simplificação	Frações originais
1	0/12
2	6/12
3	4/12, 8/12
4	3/12, 9/12
6	2/12, 10/12
12	1/12, 5/12, 7/12, 11/12

Veja que somente serão acrescentados $4 = \phi(12)$ termos a F_{11} .

2.1 Qual é o Comprimento de uma Série de Farey?

A resposta é obviamente que a n -ésima série de Farey tem comprimento

$$1 + \phi(1) + \phi(2) + \phi(3) + \dots + \phi(n-1) + \phi(n)$$

(o 1 inicial vem do fato de contarmos tanto $0/1$ como $1/1$).

Não há uma fórmula simples para esta soma particular de números de ϕ , mas sabe-se que a resposta é aproximadamente

$$3 \left(\frac{n}{\pi} \right)^2 \approx 0.3039635509 \times n^2$$

e que a aproximação fica proporcionalmente melhor à medida que n aumenta. Por exemplo, a décima série de Farey tem comprimento 33, em comparação com $3 \times 100/\pi^2 \approx 30.4$, e a centésima tem comprimento 3045, em comparação com $(3 \times 100^2)/\pi^2 \approx 3039.6$.

2.2 Outra demonstração da propriedade fundamental do MDC

Já sabemos que $\text{MDC}(r, s) = 1$ e $0 < r \leq s$ sempre podemos encontrar a e b tais que

$$ar - sb = 1. \quad (2)$$

(Na verdade sabemos que vale $ar + sb = \text{MDC}(a, b)$). A série de Farey nos fornece outra prova 2, porque podemos considerar o b/a sendo a fração que precede r/s em F_n . Por exemplo, uma solução de $3a - 7b = 1$ é $r = 5$ e $s = 2$, desde que $\frac{2}{5}$ precede $\frac{3}{7}$ em F_7 . As séries de Farey nos permitem concluir que sempre podemos encontrar solução para 2 com $0 \leq r < s \leq b$, se $0 < a \leq b$. E reciprocamente se $0 < r \leq s$ e $\text{MDC}(r, s) = 1$ nós podemos resolver 2 com $0 < a \leq b \geq s$ por colocar r/s e encontrar a fração r/s que a precede em F_n .

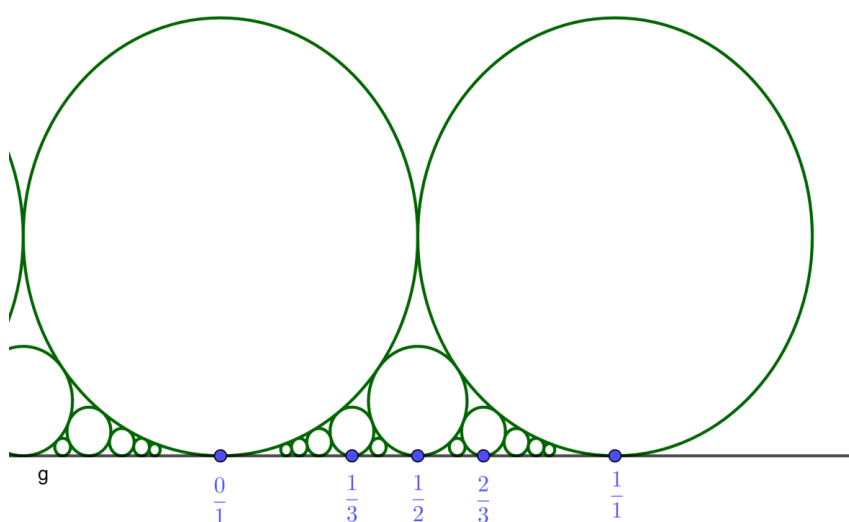
Capítulo 3

Círculos de Ford

Este capítulo segue de perto a exposição de [7], com contribuições também de [8] e [4].

Começamos com frações reais. Elas são geralmente representadas, juntamente com os irracionais reais, por pontos em uma reta. Seja esta reta o eixo x em um plano xy de coordenadas retangulares.

Através de cada ponto racional $x = p/q$, onde p e q são inteiros com $\text{MDC}(p, q) = 1$, construímos um círculo de raio $1/(2q^2)$ tangente ao eixo x e situado no semiplano superior. É este círculo, que toca o eixo x no ponto geralmente tomado para representar a fração, que será a imagem geométrica da fração. Os inteiros são representados por



círculos de raio $1/2$; as frações $1/3$, $2/3$, $4/3$, etc., por círculos de raio $1/18$; e assim por diante. Cada pequeno intervalo do eixo x contém pontos de tangência de infinitos desses círculos. Sejam p/q e P/Q duas frações diferentes em seus termos mais baixos. Considere a distância entre os centros de seus círculos representativos. Na figura, a distância horizontal AC é $|(P/Q) - (p/q)|$ e a distância vertical CB é a diferença dos

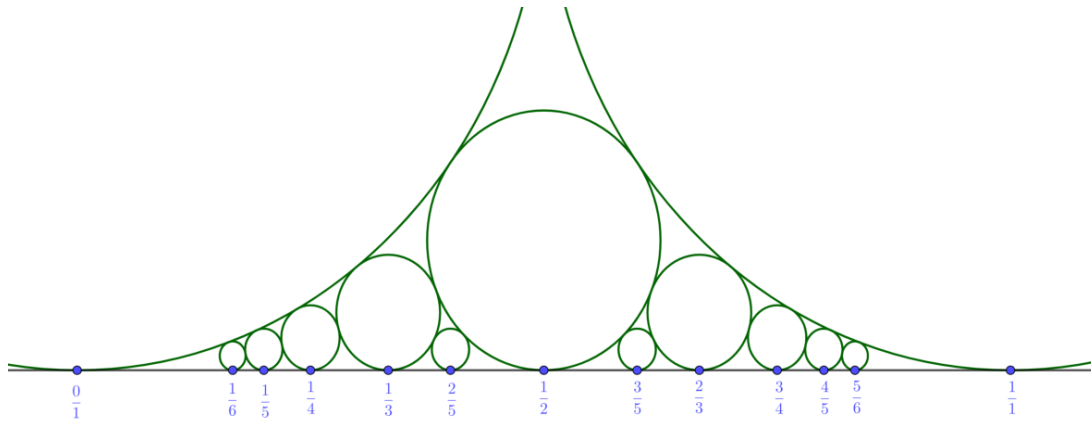
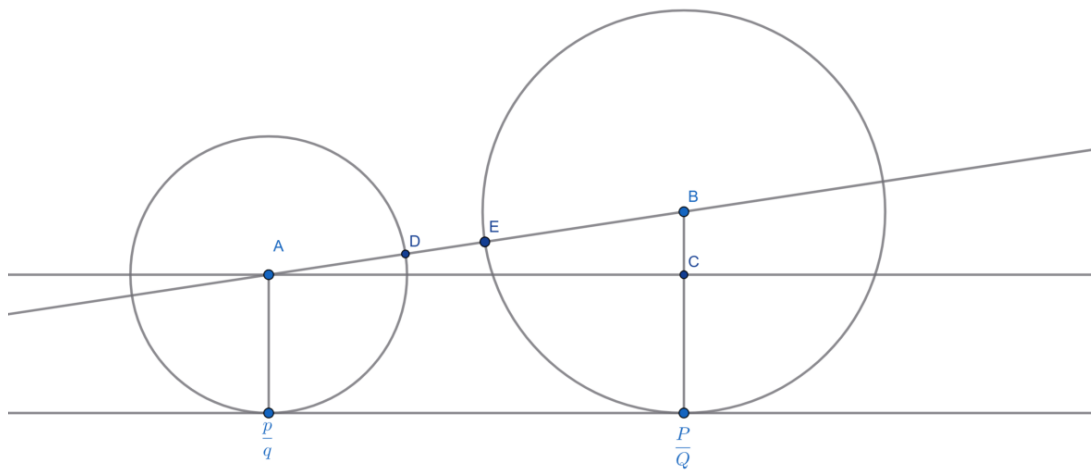


Figura 3.1: Círculos de Ford

raios, $|(1/2Q^2) - (1/2q^2)|$. Temos



$$\begin{aligned}
 AB^2 &= \left(\frac{P}{Q} - \frac{p}{q}\right)^2 + \left(\frac{1}{2Q^2} - \frac{1}{2q^2}\right)^2 \\
 &= \left(\frac{1}{2Q^2} + \frac{1}{2q^2}\right)^2 + \frac{(Pq - pQ)^2 - 1}{Q^2q^2} \\
 &= (AD + EB)^2 + \frac{(Pq - pQ)^2 - 1}{Q^2q^2}.
 \end{aligned} \tag{1}$$

Se $|Pq - pQ| > 1$, então $AB > AD + EB$, e os dois círculos são totalmente externos um ao outro.

Se $|Pq - pQ| = 1$, então $AB = AD + EB$, e os dois círculos são tangentes.

Se $|Pq - pQ| < 1$, temos, uma vez que $Pq - pQ$ é um inteiro, que $Pq - pQ = 0$. Disto, $P/Q = p/q$, o que é contrário à suposição de que as frações são diferentes. Portanto, esta última desigualdade não é possível. Podemos afirmar o seguinte resultado:

Teorema 3.1 *Os círculos representativos de duas frações distintas são tangentes ou totalmente externos um ao outro.*

3.1 Frações Adjacentes

Chamaremos duas frações p/q e P/Q de adjacentes se seus círculos representativos forem tangentes. A condição para isso é que $|Pq - pQ| = |\det \begin{bmatrix} P & p \\ Q & q \end{bmatrix}| = 1$. (Assumimos daqui em diante que as frações aparecem em seus termos mais baixos.) Provaremos

Teorema 3.2 *Cada fração p/q possui uma fração adjacente.*

Que inteiros P e Q existem satisfazendo a equação $|Pq - pQ| = 1$ (ou, arranjando os sinais adequadamente, satisfazendo $Pq - pQ = 1$) é uma proposição fundamental da teoria dos números. Muitas provas foram dadas. Por uma questão de completeza, provamos o teorema aqui. A prova dada é arranjada de modo a cobrir o caso de frações complexas que aparecerão subsequentemente neste texto.

A prova é por indução. Claramente, o teorema vale para $|q| = 1$, uma vez que $p/1$ tem então a fração adjacente $(p+1)/1$. Para provar o teorema em geral, assumimos que todas as frações cujos denominadores são menores em valor absoluto que $|q|$ possuem frações adjacentes e o provamos para a fração p/q .

Seja n o inteiro mais próximo de p/q , donde podemos escrever, com m inteiro,

$$\frac{p}{q} = n + \frac{m}{q} = \frac{nq + m}{q}, \quad 0 < |m| < |q|.$$

Como $|m| < |q|$, então q/m tem uma fração adjacente r/s , então $|sq - rm| = 1$. Então a fração

$$\frac{P}{Q} = n + \frac{s}{r} = \frac{nr + s}{r}$$

é adjacente a p/q ; pois

$$|Pq - pQ| = |(nr + s)q - (nq + m)r| = |sq - rm| = 1.$$

Isso estabelece o teorema.

Podemos agora dar uma fórmula para todas as frações adjacentes a p/q .

Teorema 3.3 *Se P/Q é adjacente a p/q , então todas as frações adjacentes a p/q são*

$$\frac{P_n}{Q_n} = \frac{P + np}{Q + nq},$$

onde n assume todos os valores inteiros.

Encontramos facilmente que as frações dadas são adjacentes a p/q , pois

$$|(P + np)q - p(Q + nq)| = |Pq - pQ| = 1.$$

Encontramos que P_n/Q_n e P_{n+1}/Q_{n+1} são adjacentes entre si, uma vez que

$$|[P + np][Q + (n + 1)q] - [P + (n + 1)p][Q + nq]| = |Pq - pQ| = 1.$$

Os círculos correspondentes a essas frações formam um anel em torno do círculo de p/q , todos tangentes ao círculo de p/q , e cada um tangente aos círculos que o precedem e o seguem na sequência (Figura 2). Que este anel envolve completamente o círculo de p/q , vemos de

$$\frac{P_n}{Q_n} = \frac{p}{q} + \frac{Pq - pQ}{q(Q + nq)} = \frac{p}{q} \pm \frac{1}{q^2(n + Q/q)}. \quad (2)$$

Quando $n \rightarrow +\infty$, P_n/Q_n aproxima-se de p/q por um lado; quando $n \rightarrow -\infty$, P_n/Q_n aproxima-se de p/q pelo outro.

É óbvio a partir da imagem geométrica que não é possível desenhar um círculo situado no semiplano superior, tocando o eixo x , e tangente ao círculo de p/q sem intersectar os círculos do anel que envolve o círculo de p/q . Segue-se que não há outras frações adjacentes a p/q .

Teorema 3.4 *Das frações adjacentes a p/q ($|q| > 1$), exatamente duas têm denominadores numericamente menores que q .*

Que dois círculos do anel precedente em torno do círculo de p/q são maiores que esse círculo é bastante evidente a partir da imagem geométrica. Vemos também que

$$|Q + nq| < |q|, \quad \text{ou} \quad |n + Q/q| < 1$$

para exatamente dois valores de n ; ou seja, aqueles inteiros entre os quais $-Q/q$ se encontra. Para um desses dois valores de n , $n + Q/q$ é positivo, para o outro, negativo. Vemos então de (2) que uma das frações do Teorema 3.4 é maior que p/q , a outra é menor.

Adicionaremos mais um círculo à nossa coleção. A fração $1/0$ é formalmente adjacente aos inteiros $p/1$, uma vez que $1 \cdot 1 - p \cdot 0 = 1$. Tomamos como seu círculo a reta $y = 1$, que toca os círculos de todos os inteiros. Consideramos como o interior do círculo a parte do plano acima da reta.

3.2 Os Círculos de Ford nos Complexos

Adiamos por um momento a demonstração da primeira metade do teorema de Hurwitz (Teorema 5.2, a ser provado no Capítulo 5) para preparar uma nova ferramenta. Esta é uma figura geométrica introduzida por L. R. Ford, consistindo em certos círculos que têm algo a ver com as sequências de Farey. É útil pensar nestes "círculos de Ford" como estando no plano complexo $z = x + iy$.

Seja $C(h/k)$ o círculo com centro em $h/k + i/(2k^2)$ e raio $1/(2k^2)$. Assim, $C(h/k)$ é o círculo

$$\left| z - \left(\frac{h}{k} + \frac{i}{2k^2} \right) \right| = \frac{1}{2k^2},$$

que está no semiplano superior e é tangente ao eixo x em $x = h/k$. Estes círculos têm uma propriedade importante.

Capítulo 4

Geometria Hiperbólica

Este capítulo segue de perto a exposição de [1] e [2].

Em geral, quando se fala em Geometria Hiperbólica sempre esperamos uma abordagem analítica. Portanto, trazer uma abordagem sintética apresenta um certo ar de novidade. Esta forma de abordar a Geometria Hiperbólica nos permite alcançar alguns resultados de maneira muito mais rápida que na construção analítica.

4.1 Construções Euclidianas Úteis

Neste capítulo só trataremos de propriedades euclidianas. Mas será a boa compreensão destas construções que nos habilitará para as construções hiperbólicas que se seguirão.

A primeira coisa que precisamos ter em mente é o conceito de reflexão, vamos iniciar com a reflexão em torno de uma reta. Uma reflexão em torno de uma reta ℓ é uma aplicação do plano no plano, que pega um ponto A e o leva em A' de tal forma que a reta ℓ se torna a mediatriz do segmento AA' (veja a Figura 4.1). Uma outra aplicação é a inversão na circunferência $\mathcal{C}$ de centro O e raio r . A inversão é uma aplicação que a cada ponto A associa um ponto A' , com a seguinte propriedade: o segmento AA' está sobre o raio OA da circunferência e $OA \times OA' = r^2$ (veja a Figura 4.1).

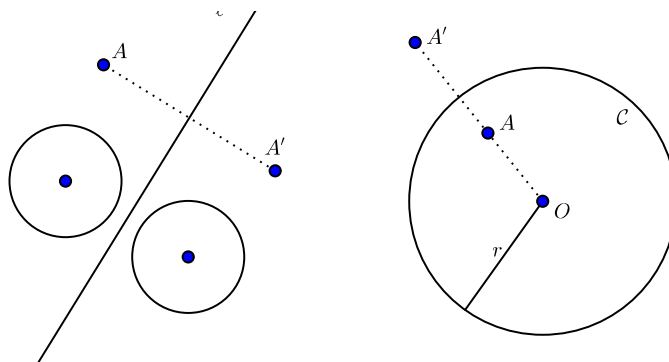


Figura 4.1: Reflexão em torno da reta ℓ e inversão em $\mathcal{C}$

Vamos iniciar a primeira construção.

Construção 4.1 *Construa uma circunferência que passe por três pontos não colineares A , B e C .*

Construa os segmentos AB e BC . Sejam ℓ_1, ℓ_2 as mediatrizes correspondentes. Chame de O o ponto de encontro de ℓ_1 e ℓ_2 . Então a circunferência desejada tem centro O e passa, digamos, por A . ■

Construção 4.2 *Inverta um ponto com respeito a uma circunferência C com centro O .*

Existem várias formas de se fazer isso. Aqui está uma:

Digamos que o ponto B está fora da circunferência, construa o segmento OB (veja a Figura 4.2 abaixo), e depois uma outra circunferência C' , com centro no ponto médio do segmento OB e que passe por O . Seja P um dos pontos de interseção de C e C' (observe que o triângulo OPB é retângulo, uma vez que OB é o diâmetro de C'). Considere a reta ℓ perpendicular a OB que passa por P . Seja A o ponto de interseção de ℓ com OB . Então A é o inverso de B com respeito a C . De fato, observando o triângulo OBP e o triângulo OAP , vemos que ambos são semelhantes, uma vez que compartilham o ângulo $B\hat{O}P$. Pela semelhança temos que

$$\frac{OP}{OB} = \frac{OA}{OP} \Rightarrow OA \cdot OB = OP^2, \text{ mas } OP = r.$$

Se queremos inverter um ponto A que está no interior da circunferência (diferente de O), basta reverter esta construção. Considere a reta ℓ perpendicular ao segmento OA . Seja P o ponto de interseção de ℓ com a circunferência. Considere ℓ' a perpendicular a OP passando por P . Então o ponto B será determinado pela interseção da semirreta OA com ℓ' . ■

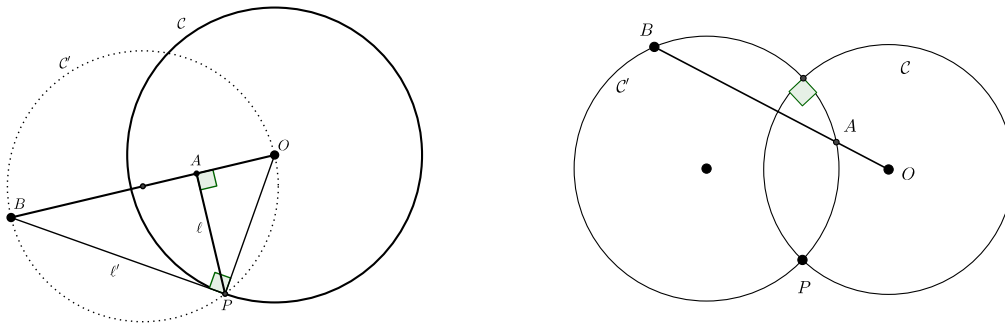


Figura 4.2: Construções 4.1 e 4.2

Existe um método mais rápido se tivermos uma condição adicional.

Construção 4.3 *Dado um par de circunferências ortogonais C e C' com centro O, O' , respectivamente, e A um ponto sobre C' , construa o inverso de A com respeito a C .*

Simplesmente construa a semirreta OA , como na Figura 4.2 à direita; um dos pontos de interseção da semirreta OA com C' é o A ; o outro, B , será o inverso de A com respeito a C . Para vermos isso, veja a Figura 4.3 à esquerda, marque o ponto P de interseção de C e C' e trace os segmentos OP, PB e PA . Agora veja que os triângulos OPA e OBP são semelhantes. ■

Para concluir a semelhança você necessita dos dois seguintes resultados:

Resultado 1: Um ângulo APB inscrito em uma circunferência é metade do ângulo central AOB .

Resultado 2: A medida de um ângulo formado por uma secante e uma tangente APB é igual à metade do arco correspondente POB (veja a figura à direita em 4.3).

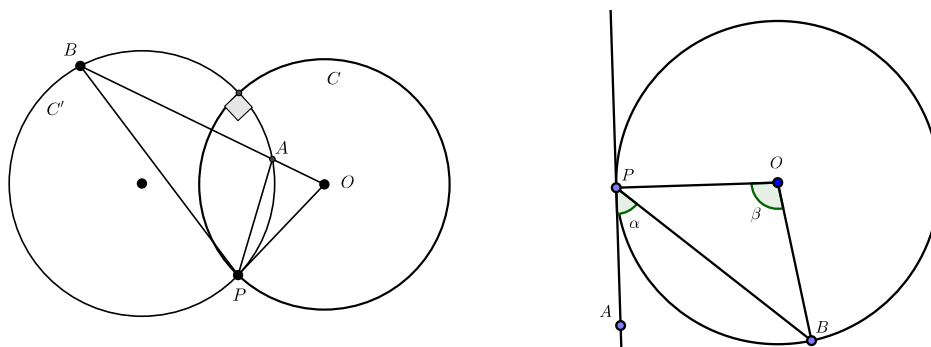


Figura 4.3: Demonstração da construção 4.3

Uma consequência disso é o seguinte:

Lema 4.4 *Sejam C, C' circunferências ortogonais e A um ponto sobre C' . Então o inverso de A com respeito a C está sobre C' . Similarmente, se A e B são inversos com respeito a C e estão sobre uma circunferência C' , então C e C' são ortogonais.*

Lema 4.5 *Seja C uma circunferência com centro O . Seja A um ponto diferente de O . Então o lugar geométrico de todas as circunferências passando por A e ortogonais a C é uma reta. Se A está no interior de (sobre, no exterior de) C , esta linha está no exterior da (tangente a, no interior de) C . Além disso, qualquer linha no exterior de, ou tangente a, C é tal lugar geométrico.*

Demonstração: Como já vimos, toda circunferência que passa por A e é ortogonal a C passa também pelo inverso B de A com respeito à circunferência C . Então o lugar geométrico dos centros destas circunferências deve ser a mediatriz do segmento AB , os quais estão no exterior de C . Reciprocamente, qualquer circunferência sobre a mediatriz passando sobre os dois pontos A e B deve ser ortogonal a C .

Se A está sobre C , então este lugar geométrico deve ser uma linha tangente a C por A .

Também é simples mostrar que qualquer linha no exterior de, ou tangente a C , é um tal lugar geométrico. $\square$

Construção 4.6 *Dada uma circunferência C com centro O , e um ponto A no exterior de C , construa uma única circunferência C' com centro A , ortogonal a C .*

Seja C'' uma circunferência com centro no ponto médio do segmento AO e que passa por A e O (veja o lado direito da Figura 4.4); seja P um dos pontos de interseção das circunferências C e C'' . Então a circunferência desejada tem centro A e passa por P . $\blacksquare$

No lema 4.4 e na construção 4.6 estabelecemos um fato muito importante: com respeito à circunferência C dada, as circunferências ortogonais C' estão em

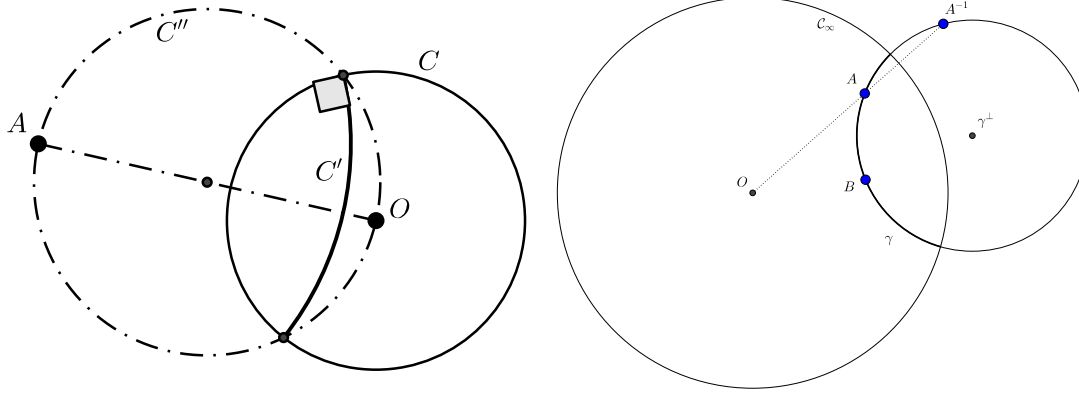


Figura 4.4: Construções 4.6 e 4.7

correspondência com pontos no exterior de $\mathcal{C}$ (os centros de $\mathcal{C}'$ são chamados seus polos) e cada ponto no interior de $\mathcal{C}$ corresponde a uma reta (chamada de polar) no exterior de $\mathcal{C}$.

Fixando a circunferência $\mathcal{C}$, denote o polo da circunferência γ por $\gamma^\perp$; reciprocamente, cada ponto B no exterior de $\mathcal{C}$ é um polo de alguma circunferência, denotada por $B^\perp$, como ilustrado na Figura 4.5. Denote o polar de cada ponto A no interior de $\mathcal{C}$ por $A^\perp$; reciprocamente, cada reta ℓ que não intercepta $\mathcal{C}$ é o polar de algum ponto, denotado por $\ell^\perp$.

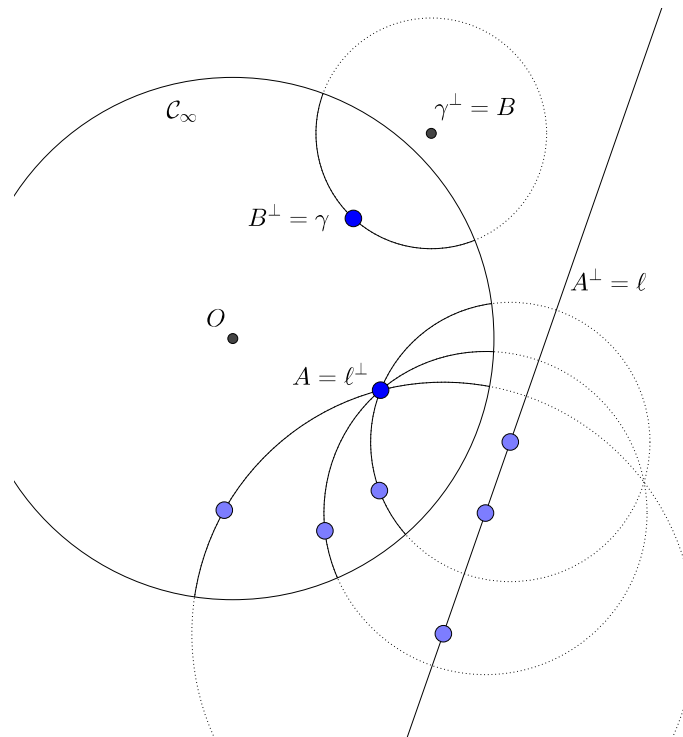


Figura 4.5: Polos e Polares

Falando em um nível mais sofisticado, os polos de nossas geodésicas estão no plano projetivo real $\mathbb{RP}^2$ - então ao juntar os pontos no infinito obtemos os polos que por sua vez são diâmetros de $\mathcal{C}$. Isso, por sua vez, nos fornece uma evidência de que o espaço das retas no plano hiperbólico é um aberto da faixa de Möbius, isto é, $\mathbb{RP}^2$ menos um disco fechado limitado por $\mathcal{C}$.

4.2 Construções com Régua e Compasso na Geometria Hiperbólica

O Espaço hiperbólico é um local onde a geometria (hiperbólica) não-euclidiana pode ser construída. Este espaço deve possuir um atributo incomum que é: à medida que você se afasta de um ponto, o espaço ao seu redor expande exponencialmente. É quase impossível criar modelos físicos precisos deste espaço usando qualquer um dos métodos usuais. Mas em 1997 a matemática Daina Taimina da Universidade de Cornell inventou um modelo usando o crochê! Depois disso, todos começaram a ver a geometria hiperbólica em diversos locais - folhas de alface, corais, etc. Quem estiver interessado veja o site <http://www.theiff.org/main.html> do Institute for Figuring.

4.3 O Modelo do Disco de Poincaré

Nesta seção vamos levar a cabo algumas construções utilizando régua e compasso na geometria hiperbólica. Vamos considerar o modelo do disco de Poincaré. Sempre nos referiremos à borda deste disco como $\mathcal{C}_\infty$ e ao seu centro por O . Chamaremos



Figura 4.6: Modelo do Espaço Hiperbólico por Daina Taimina

o disco aberto de $\mathbb{D}$. Quando nos referirmos ao inverso, a menos de claro aviso, significará o inverso com respeito a $\mathcal{C}_\infty$; $X^\perp$ significa o polar, polo, etc., de algum objeto X com respeito a $\mathcal{C}_\infty$, sempre no sentido do capítulo anterior.

$\mathbb{D}$ é conhecido como Disco de Poincaré e será o nosso plano hiperbólico. Os pontos desta região são os pontos do plano hiperbólico, até aqui sem muitas novidades. As retas hiperbólicas são os arcos de circunferências ortogonais a $\mathcal{C}_\infty$ ou diâmetros de $\mathcal{C}_\infty$.

A geometria hiperbólica satisfaz a maioria dos axiomas de Euclides, mas é preciso alterar o quinto axioma pelo seguinte: Dados uma reta do plano hiperbólico e um ponto fora da mesma, existe um número infinito de retas hiperbólicas que contêm o dado ponto e não encontram a reta dada. Outras semelhanças são a possibilidade de definir distância, assim como a noção de área no plano hiperbólico. Existe ainda outra diferença importante: a área de um triângulo hiperbólico depende somente de seus ângulos e não do comprimento de seus lados, assim o caso AAA é um caso de congruência na geometria hiperbólica.

4.4 Primeiras Construções

Construção 4.7 *Dados pontos A e $B \in \mathbb{D}$, construa a reta hiperbólica AB . De maneira equivalente, dados dois pontos A, B e a circunferência $\mathcal{C}_\infty$, construa a única circunferência passando por A, B e que seja ortogonal a $\mathcal{C}_\infty$.*

Invertendo A (com respeito a $\mathcal{C}_\infty$) obtemos A^{-1} . A circunferência procurada é obtida fazendo a circunferência que passa pelos pontos A, B e A^{-1} . ■

Construção 4.8 *Dados pontos A e $B \in \mathbb{D}$, e seus polares $A^\perp$ e $B^\perp$, construa a reta hiperbólica AB .*

Se conhecemos os polares $A^\perp$ e $B^\perp$, obtenha P , o ponto de interseção, e faça a circunferência com centro em P e raio $PA = PB$ (veja a Figura 4.7 à direita). ■

Na prática, se A e B estão relativamente próximos em $\mathbb{D}$, $A^\perp$ e $B^\perp$ são quase paralelas e P pode ser difícil de encontrar. Por sorte, a mediatriz do segmento (euclidiano) AB também passa por P , o que permite aumentar a precisão da construção.

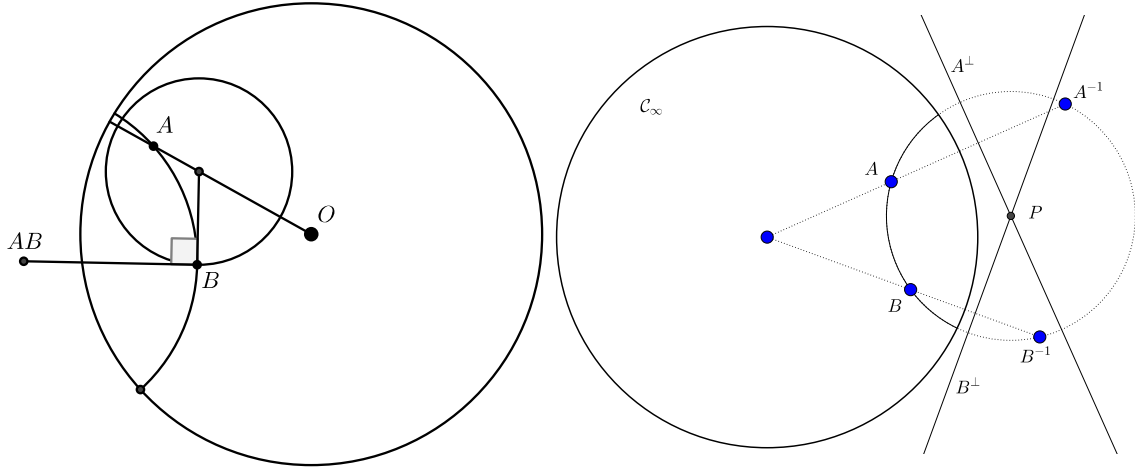


Figura 4.7: Circunferência e Retas Hiperbólicas

Construção 4.9 Dados pontos A e $B \in \mathbb{D}$, construa uma circunferência hiperbólica com centro A e passando por B .

Vamos admitir que sabemos que a curva procurada $\mathcal{C}$ é a única circunferência que é ortogonal a toda geodésica que passa por A . Em particular, $\mathcal{C}$ deve ser ortogonal a OA , visto que este faz parte de um diâmetro de $\mathcal{C}_\infty$. Além disso, $\mathcal{C}$ também deve ser perpendicular à reta hiperbólica AB . Então $\mathcal{C}$ deve ser a circunferência (euclidiana) com centro na interseção destas duas linhas e que passa por B . ■

Construção 4.10 Dados pontos A e $B \in \mathbb{D}$, construa a mediatriz do segmento hiperbólico AB . Similarmente, construa o ponto médio do segmento hiperbólico AB .

O procedimento é o mesmo que o realizado no plano euclidiano. Desenhe duas circunferências hiperbólicas $\mathcal{C}_1$ e $\mathcal{C}_2$, com centros em A e B , respectivamente, e passando por B e A . Sejam P, Q os pontos de $\mathcal{C}_1 \cap \mathcal{C}_2$. Então a reta hiperbólica PQ é a mediatriz e o ponto médio de AB é a interseção de PQ com AB . ■

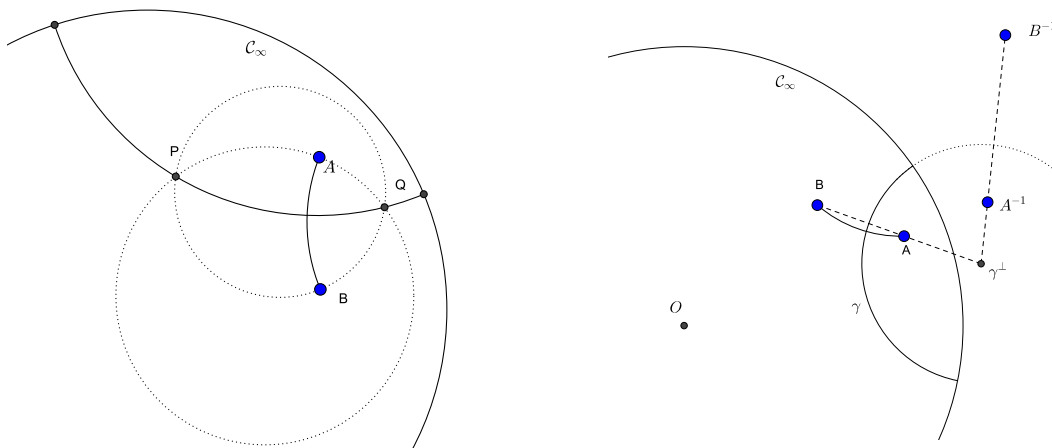


Figura 4.8: Mediatriz Hiperbólica

Construção 4.11 Dado o ponto $A \in \mathbb{D}$ e uma reta hiperbólica γ , construa a reflexão de A por γ .

Imitando a construção no caso euclidiano, construa uma circunferência hiperbólica com centro em A e com raio suficientemente grande para cortar γ em dois pontos: B e C . Construa circunferências hiperbólicas centradas nos pontos B e C que passem por A . O ponto procurado A' será o outro ponto de interseção dessas duas circunferências. Observe que, se traçarmos o segmento AA' , obteremos a mediatriz hiperbólica do segmento hiperbólico BC . ■

Na construção acima é importante observar que o ponto A' é a inversão de A através de γ ; desta forma, os conceitos de inversão em uma circunferência e a reflexão em uma reta hiperbólica se mostram equivalentes. Então, ao introduzir uma inversão em uma circunferência, não estamos introduzindo um conceito novo. Além disso, com esta propriedade em mente, podemos fazer a construção 4.10 da seguinte forma: Considere que já conhecemos a mediatriz hiperbólica γ de AB . Então a inversão através de γ permuta A e B . Por outro lado, γ deve ser ortogonal a C_∞ , então a inversão (através de γ) também deve intercambiar A^{-1} e B^{-1} . Daí $\gamma^\perp$ está sobre a interseção de AB e $A^{-1}B^{-1}$ (veja a Figura 4.8 à direita).

Construção 4.12 Sejam $A, B, C \in \mathbb{D}$, construa uma circunferência hiperbólica com centro em A e com raio congruente ao segmento hiperbólico BC .

Inicie construindo a mediatriz hiperbólica γ do segmento AB . Inverta C através de γ e obtenha P . A circunferência desejada tem centro A e passa por P . ■

Construção 4.13 Sejam γ uma reta hiperbólica, A um ponto de γ e um ângulo euclidiano α . Construa uma outra reta hiperbólica que encontra γ em A formando com esta o ângulo α .

Veja a Figura 4.9. Considere em $A^\perp$ o ponto $\gamma^\perp$. Seja ℓ a reta que passa por A e $\gamma^\perp$. Marque ℓ' a reta que passa por A e forma com ℓ o ângulo α . Seja P a interseção de ℓ com ℓ' . Afirmação: A reta procurada é dada por $P^\perp$. ■

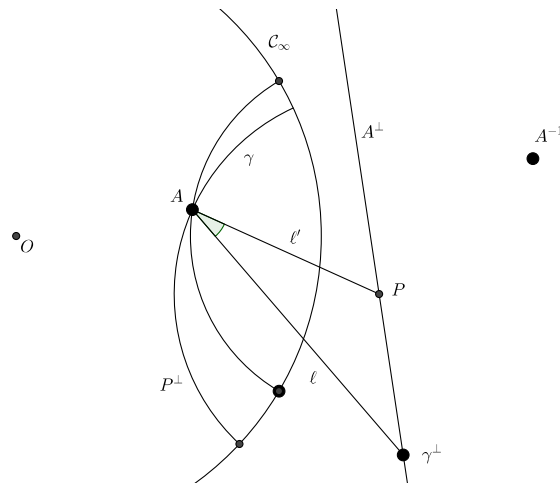


Figura 4.9: Construção do ângulo hiperbólico

Exemplo: invertendo os Círculos de Ford

Temos, como sempre, o empacotamento dos Círculos de Ford: circunferências tangentes à reta f em cada ponto racional, tangentes entre si segundo o padrão das frações de Farey. Fizemos a inversão dessa figura na circunferência auxiliar de centro

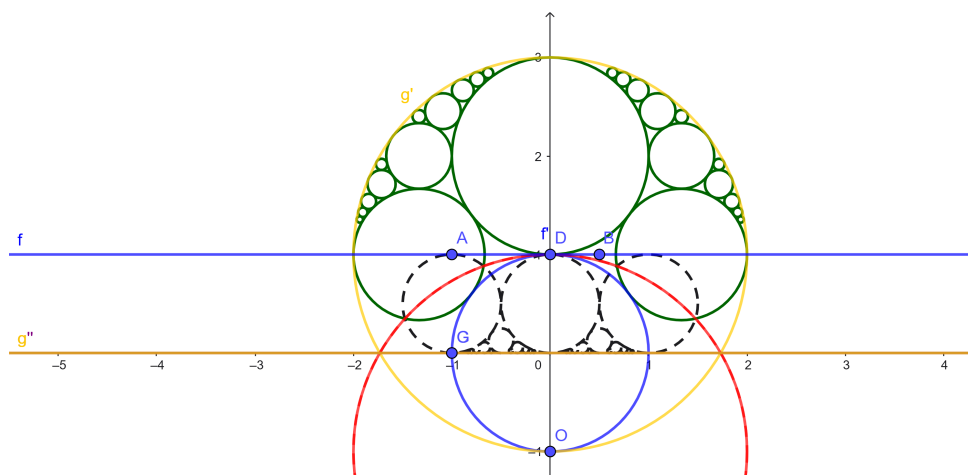


Figura 4.10: Resultado da Inversão dos Círculos de Ford

$(-1,0)$ e raio 2. Já sabíamos disso, pois construímos com régua e compasso no modelo do disco de Poincaré: inverter numa circunferência é o mesmo que refletir numa reta. A inversão troca o interior pelo exterior da circunferência de inversão, preserva ângulos e preserva tangências — exatamente como faz uma reflexão comum.

O resultado é a nuvem de circunferências verdes dentro da circunferência amarela g' (centro $(1,0)$, raio 2): é a mesma família de Círculos de Ford, apenas refletida. Cada tangência da figura original vira uma tangência na figura nova — nenhuma circunferência deixa de tocar suas vizinhas. E a própria reta f , que era a "base" onde tudo era tangente, também se transforma: sua imagem é a circunferência g' , que passa a ser a nova "base" da figura refletida.

Vamos pensar o interior de g' como um modelo de disco de Poincaré, com essa circunferência fazendo o papel de horizonte (os pontos infinitamente distantes). Nesse modelo, inverter numa circunferência ortogonal ao horizonte é uma reflexão hiperbólica — o análogo, na geometria hiperbólica, de refletir uma figura num espelho reto.

Como essa reflexão preserva tangências e ângulos, a nuvem verde tem, dentro do disco g' , exatamente as mesmas propriedades hiperbólicas que os Círculos de Ford originais: continuam sendo horociclos (circunferências tangentes ao horizonte), e o padrão de quem toca quem reproduz fielmente as frações de Farey. Só a aparência euclidiana mudou; a geometria hiperbólica da figura permanece a mesma.

A justificativa cuidadosa de por que uma inversão se comporta, de fato, como uma reflexão hiperbólica é mais natural na via analítica: com fórmulas explícitas para as transformações de Möbius — das quais as inversões são a versão antiholomorfa — essa equivalência fica imediata. É para lá que seguimos agora.

4.5 Conexão Analítica: O Modelo do Semi-Plano Superior, os Círculos de Ford e as Ternas Pitagóricas

A transição entre abordagens sintéticas e analíticas revela conexões profundas entre a geometria hiperbólica, a teoria dos números e a Álgebra. Vamos agora retomar os Círculos de Ford $C(h/k)$ do Capítulo 3 – ali estudados como objetos do plano complexo $z = x + iy$ – e situá-los dentro do modelo do disco de Poincaré $\mathbb{D}$ que construímos neste capítulo.

Para isso, precisamos primeiro reconhecer que o semi-plano superior onde vivem os Círculos de Ford é, ele próprio, um modelo da geometria hiperbólica: o *Modelo do Semi-Plano Superior de Poincaré*, que denotaremos por $\mathbb{H}^2$ (não confundir com o uso puramente informal da expressão “plano hiperbólico” feito antes neste capítulo):

$$\mathbb{H}^2 = \{z \in \mathbb{C} \mid \text{Im}(z) > 0\}.$$

A fronteira geométrica deste modelo no infinito é a reta real estendida $\partial\mathbb{H}^2 = \mathbb{R} \cup \{\infty\}$ – exatamente onde moram as frações de Farey h/k e onde cada Círculo de Ford $C(h/k)$ é tangente ao eixo real.

Existe um mapeamento analítico bijetivo e isométrico entre $\mathbb{H}^2$ e o disco unitário $\mathbb{D} = \{w \in \mathbb{C} \mid |w| < 1\}$: a célebre *Transformada de Cayley*, uma transformação de Möbius da forma

$$w = f(z) = \frac{z - i}{z + i}. \quad (1)$$

Esta função leva os pontos de $\mathbb{H}^2$ no interior de $\mathbb{D}$ e, de modo crucial, estende-se à fronteira enviando a reta real $\mathbb{R}$ sobre a circunferência unitária $\partial\mathbb{D} = \mathcal{C}_\infty$. Vamos verificar isso!

Seja $z = x + iy$ com $y > 0$. Calculamos:

$$|f(z)| = \left| \frac{z - i}{z + i} \right| = \frac{|z - i|}{|z + i|}.$$

Mas

$$|z - i|^2 = x^2 + (y - 1)^2, \quad |z + i|^2 = x^2 + (y + 1)^2.$$

Assim,

$$|f(z)| < 1 \iff |z - i| < |z + i| \iff x^2 + (y - 1)^2 < x^2 + (y + 1)^2.$$

Simplificando:

$$(y - 1)^2 < (y + 1)^2 \iff y^2 - 2y + 1 < y^2 + 2y + 1 \iff -2y < 2y \iff y > 0.$$

Portanto, se $z \in \mathbb{H}^2$, então $|f(z)| < 1$. Logo,

$$f(\mathbb{H}^2) \subseteq \mathbb{D}.$$

Devemos mostrar que para todo $w \in \mathbb{D}$ existe $z \in \mathbb{H}^2$ tal que $f(z) = w$. Invertamos a transformação:

$$w = \frac{z - i}{z + i} \implies w(z + i) = z - i \implies wz + wi = z - i \implies z(w - 1) = -i(1 + w).$$

Assim,

$$z = i \frac{1+w}{1-w}, \quad (w \neq 1).$$

Escrevendo $w = u + iv$ com $u^2 + v^2 < 1$, obtemos:

$$z = i \frac{1+u+iv}{1-u-iv}.$$

Multiplicando numerador e denominador pelo conjugado do denominador:

$$z = i \frac{(1+u+iv)(1-u+iv)}{(1-u)^2 + v^2}.$$

Desenvolvendo o numerador (sem o fator i):

$$(1+u+iv)(1-u+iv) = (1-u^2-v^2) + i(2v).$$

Logo,

$$z = \frac{i(1-u^2-v^2) - 2v}{(1-u)^2 + v^2}.$$

Portanto,

$$\operatorname{Im}(z) = \frac{1-u^2-v^2}{(1-u)^2 + v^2}.$$

Como $|w| < 1$, temos $u^2 + v^2 < 1$, logo o numerador é positivo e o denominador é positivo. Assim,

$$\operatorname{Im}(z) > 0,$$

ou seja, $z \in \mathbb{H}$. Logo, para cada $w \in \mathbb{D}$ existe $z \in \mathbb{H}^2$ com $f(z) = w$. Portanto,

$$\mathbb{D} \subseteq f(\mathbb{H}^2).$$

Para entender melhor como funciona a transformação veja o próximo exemplo

Exemplo 4.14 Tomei três circunferências no semiplano superior, cada uma determinada por três pontos: (A_1, B_1, C_1) , (A_2, B_2, C_2) e (A_3, B_3, C_3) . Apliquei a cada um desses nove pontos a transformação

$$f(z) = \frac{z-i}{z+i},$$

obtendo $A'_1 = f(A_1)$, $B'_1 = f(B_1)$, $C'_1 = f(C_1)$, e assim por diante. Em seguida, usando a ferramenta que traça a circunferência por três pontos, desenhei a circunferência determinada por cada trinca de imagens.

O que a figura evidencia:

- cada trinca de pontos-imagem ainda determina uma circunferência — f nunca "quebra" um círculo em outra coisa;
- os pontos que estavam sobre a reta h (o eixo real) têm imagens que também caem todas sobre uma mesma curva — a fronteira da figura-imagem;
- circunferências que cruzavam h em dois pontos têm imagens que cruzam essa curva-fronteira exatamente nas imagens desses dois pontos;
- relações de estar "dentro" ou "fora" de outra circunferência são preservadas: nada se inverte ou se embaralha, f é uma correspondência contínua e biunívoca.

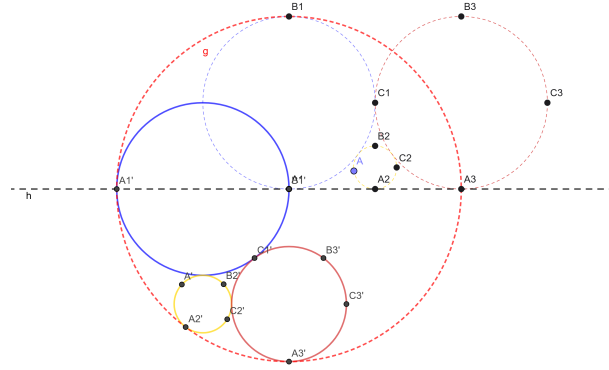


Figura 4.11: Ação da transformada de Cayley sobre três circunferências do semiplano superior, cada uma determinada por três pontos

Para quem já conhece o assunto

Os valores particulares

$$f(i) = 0, \quad f(0) = -1, \quad f(-1) = i, \quad f(1) = -i, \quad f(\infty) = 1$$

ilustram concretamente o que já provamos acima: i , o "centro" de $\mathbb{H}^2$, vai para o centro do disco, e os pontos reais $0, \pm 1, \infty$ — que determinam o eixo h da figura — caem todos sobre C_∞ , confirmando que h é enviada, ponto a ponto, sobre a fronteira do disco.

Duas propriedades gerais de Möbius explicam tudo o que se observa:

1. *toda transformação de Möbius leva o conjunto das circunferências generalizadas (círculos e retas, indistintamente) nele mesmo — por isso cada trinca de pontos, esteja ela sobre um círculo ou sobre uma reta, sempre produz uma imagem concíclica ou colinear; como f não tem polo em nenhuma das retas/círculos considerados aqui, as imagens são sempre círculos genuínos;*
2. *f preserva a razão cruzada de quatro pontos, e é isso que garante, de forma mais forte, que qualquer outro ponto de um dos círculos originais tem imagem sobre o círculo-imagem correspondente — não é coincidência que só testamos três pontos por curva.*

Vale um contraste com o exemplo anterior: lá usamos uma *inversão* numa circunferência, um mapa antiholomorfo (envolve $\bar{z}$), que funciona como uma *reflexão* — reverte orientação, mas permanece dentro do mesmo modelo (o semiplano). Aqui f é holomorfa, portanto conforme e *preserva* orientação: não é uma reflexão hiperbólica, mas sim a mudança de coordenadas que troca de modelo, levando o semiplano superior exatamente sobre o disco de Poincaré usual.

4.5.1 Para onde vão os pontos de Farey

Seja h/k uma fração de Farey ($h, k \in \mathbb{Z}$, $k > 0$, $\text{mdc}(h, k) = 1$, como definido no Capítulo anterior), vista como ponto da fronteira $\partial\mathbb{H}^2$. Aplicando (1) a $z = h/k$:

$$w = f\left(\frac{h}{k}\right) = \frac{\frac{h}{k} - i}{\frac{h}{k} + i} = \frac{h - ki}{h + ki}.$$

Multiplicando numerador e denominador pelo conjugado do denominador, $(h - ki)$:

$$w = \frac{(h - ki)^2}{(h + ki)(h - ki)} = \frac{h^2 - k^2 - 2hki}{h^2 + k^2}.$$

Logo $w = x + iy$ com

$$x = \frac{h^2 - k^2}{h^2 + k^2}, \quad y = \frac{-2hk}{h^2 + k^2}. \quad (2)$$

Como $h, k \in \mathbb{Z}$, as coordenadas (x, y) são **racionais**: cada fração de Farey h/k vai parar em um ponto racional de $\mathcal{C}_\infty$. (O sinal de y é apenas uma escolha de orientação em (1); com $f(z) = \frac{i-z}{i+z}$ obtém-se $y = +\frac{2hk}{h^2+k^2}$, sem mudar nada do que segue.)

4.5.2 Os Círculos de Ford dentro do disco

Vimos no Capítulo 3 que dois Círculos de Ford $C(h/k)$ e $C(h'/k')$ são tangentes exatamente quando h/k e h'/k' são vizinhos de Farey, isto é, $|hk' - h'k| = 1$ (Teorema 3.1). Essa é uma propriedade de *tangência*, e as transformações de Möbius têm exatamente a propriedade que precisamos para transportá-la ao disco: elas levam circunferências e retas em circunferências e retas, e preservam ângulos – em particular, preservam tangência.

Assim, ao aplicarmos a Transformada de Cayley (1) a cada Círculo de Ford $C(h/k)$ (que é tangente ao eixo real em h/k , ponto de $\partial\mathbb{H}^2$), obtemos uma circunferência dentro de $\mathbb{D}$, tangente internamente a $\mathcal{C}_\infty$ precisamente no ponto racional (x, y) de (2) – a imagem de $C(h/k)$ pelo disco. Como a Transformada de Cayley preserva tangência, dois desses círculos-imagem se tocam se, e somente se, os Círculos de Ford originais se tocavam, isto é, se, e somente se, h/k e h'/k' eram vizinhos de Farey. Ou seja: *a figura dos Círculos de Ford do Capítulo 3, com toda a sua combinatória de tangências, transporta-se intacta para dentro do disco de Poincaré, sem exigir nenhum cálculo novo além do que já tínhamos.*

4.5.3 Conexão com as Ternas Pitagóricas

A expressão (2) tem ainda um significado aritmético clássico. Como $(x, y) \in \mathcal{C}_\infty$, vale $x^2 + y^2 = 1$; de fato,

$$x^2 + y^2 = \frac{(h^2 - k^2)^2 + 4h^2k^2}{(h^2 + k^2)^2} = \frac{(h^2 + k^2)^2}{(h^2 + k^2)^2} = 1.$$

Escrevendo essas coordenadas em termos de um triângulo retângulo de lados inteiros, obtemos

$$A = h^2 - k^2, \quad B = 2hk, \quad C = h^2 + k^2, \quad A^2 + B^2 = C^2,$$

que é exatamente a *Parametrização de Euclides* das ternas pitagóricas. Uma ressalva importante: essa fórmula só produz uma terna **primitiva** (com $\text{mdc}(A, B, C) = 1$) quando h, k , além de coprimos, têm paridades opostas. Se h e k forem ambos ímpares – o que é perfeitamente possível entre frações de Farey reduzidas –, A, B, C saem todos pares, e a terna obtida é o dobro de uma terna primitiva. Assim, os pontos racionais de

$\mathcal{C}_\infty$ correspondem, via esta parametrização, a *todas* as ternas pitagóricas – primitivas ou não –, e não apenas às primitivas.

Fecha-se assim o elo entre as construções sintéticas deste capítulo, os Círculos de Ford e as frações de Farey do Capítulo 3, e as ternas pitagóricas – preparando o terreno para o Teorema de Hurwitz no Capítulo 5.

Capítulo 5

Melhor Aproximação Racional de Números Irracionais; Círculos de Ford e o Teorema de Hurwitz

Este capítulo segue de perto a exposição de [9].

5.1 Qualidade da Aproximação

Voltemos à nossa teoria da aproximação. Se γ é irracional e N é um inteiro positivo dado, mostramos a existência de uma aproximação racional h/k , com $k \leq N$, tal que

$$\left| \gamma - \frac{h}{k} \right| < \frac{1}{k(N+1)}.$$

Em particular, temos então

$$\left| \gamma - \frac{h}{k} \right| < \frac{1}{k^2}.$$

Estes argumentos elementares podem ser usados para mostrar a existência de infinitas frações h/k tais que

$$\left| \gamma - \frac{h}{k} \right| < \frac{1}{k^2}.$$

No entanto, usando sequências de Farey, podemos provar um teorema muito mais forte.

Teorema 5.1 *Se γ é irracional, existem infinitas frações h/k tais que*

$$\left| \gamma - \frac{h}{k} \right| < \frac{1}{2k^2}.$$

Demonstração: Suponha que na sequência de Farey de ordem N temos

$$\frac{a}{b} < \gamma < \frac{c}{d}.$$

Desejamos mostrar que ou

$$\gamma - \frac{a}{b} < \frac{1}{2b^2} \quad \text{ou} \quad \frac{c}{d} - \gamma < \frac{1}{2d^2}. \quad (1)$$

Suponha o contrário,

$$\gamma - \frac{a}{b} \geq \frac{1}{2b^2} \quad \text{e} \quad \frac{c}{d} - \gamma \geq \frac{1}{2d^2}. \quad (2)$$

Então

$$\frac{c}{d} - \frac{a}{b} \geq \frac{1}{2b^2} + \frac{1}{2d^2} = \frac{b^2 + d^2}{2b^2d^2}.$$

Agora sabemos pela teoria das sequências de Farey que

$$\frac{c}{d} - \frac{a}{b} = \frac{1}{bd},$$

de modo que

$$0 \geq \frac{b^2 + d^2}{2b^2d^2} - \frac{1}{bd} = \frac{(b - d)^2}{2b^2d^2}.$$

Mas isto só é possível se $b = d$. Além disso, isto implica $b = d = 1$, uma vez que $ad - bc = -1$. Portanto, para uma sequência de Farey de ordem $N > 1$, (2) deve ser falsa e (1) verdadeira, o que prova o teorema.

Assim, somos levados à questão de saber se existem infinitas frações h/k com uma aproximação ainda melhor

$$\left| \gamma - \frac{h}{k} \right| < \frac{1}{ck^2} \quad (3)$$

com $c > 2$, e qual é o maior valor de c possível.

Esta questão foi completamente respondida por A. Hurwitz.

Teorema 5.2 (Hurwitz) *Para qualquer $c \leq \sqrt{5}$ positivo, a desigualdade (3) tem infinitas soluções. Existe, no entanto, um irracional γ para o qual (3) tem apenas finitas soluções no caso $c > \sqrt{5}$.*

A segunda metade deste teorema é facilmente resolvida. Tomamos $\gamma = \frac{1+\sqrt{5}}{2}$ e $0 < \alpha < 1$ e perguntamos por frações h/k tais que

$$\left| \frac{h}{k} - \frac{1+\sqrt{5}}{2} \right| < \frac{\alpha}{\sqrt{5}k^2}. \quad (4)$$

Se escrevermos

$$\frac{h}{k} - \frac{1+\sqrt{5}}{2} = \frac{\theta}{\sqrt{5}k^2},$$

a desigualdade anterior significa $|\theta| < \alpha < 1$. Temos

$$h - \frac{k}{2} = \frac{\sqrt{5}k}{2} + \frac{\theta}{\sqrt{5}k}$$

e, após elevar ao quadrado e reorganizar,

$$h^2 - hk - k^2 = \theta + \frac{\theta^2}{5k^2}.$$

O inteiro no lado esquerdo não pode ser 0 para inteiros h, k não ambos nulos, portanto temos

$$1 \leq \left| \theta + \frac{\theta^2}{5k^2} \right| < \alpha + \frac{\alpha^2}{5k^2}$$

ou

$$k^2 < \frac{\alpha^2}{5(1-\alpha)}.$$

Isto restringe os denominadores k , e (4) então também permite apenas finitos h para cada um dos finitos k . Portanto, (3) tem de fato apenas finitas soluções h/k para

$$c = \frac{\sqrt{5}}{\alpha} > \sqrt{5}.$$

5.2 Triângulos Circulares

Recordemos que dois círculos de Ford distintos nunca se intersectam. Eles são tangentes se e somente se as suas frações são adjacentes em alguma sequência de Farey.

Vamos precisar determinar a ordenada x do ponto de tangência de dois círculos de Ford, quando eles são tangentes. Sejam $C(h/k)$ e $C(l/m)$ dois círculos de Ford tangentes entre si, com centros

$$A = \left(\frac{h}{k}, \frac{1}{2k^2} \right), \quad B = \left(\frac{l}{m}, \frac{1}{2m^2} \right)$$

e raios

$$r_1 = \frac{1}{2k^2}, \quad r_2 = \frac{1}{2m^2}.$$

Quando duas circunferências são tangentes *externamente*, o ponto de tangência $w = (u, v)$ está sobre o segmento que une os dois centros, a uma distância r_1 de A e r_2 de B (pois $AB = r_1 + r_2$ é precisamente a condição de tangência externa). Logo w divide o segmento AB internamente na razão

$$AP : PB = r_1 : r_2.$$

Se um ponto P divide o segmento AB na razão $AP : PB = r_1 : r_2$, então

$$P = \frac{r_2 A + r_1 B}{r_1 + r_2}.$$

Note que o peso de A é r_2 e o peso de B é r_1 (os pesos ficam "trocados" em relação à razão); é por isso a razão fica como

$$\frac{1}{k^2} : \frac{1}{m^2} = m^2 : k^2.$$

Logo a coordenada u :

$$u = \frac{r_2 \cdot \frac{h}{k} + r_1 \cdot \frac{l}{m}}{r_1 + r_2} = \frac{\frac{1}{2m^2} \cdot \frac{h}{k} + \frac{1}{2k^2} \cdot \frac{l}{m}}{\frac{1}{2k^2} + \frac{1}{2m^2}}.$$

Portanto

$$u = \frac{hk + lm}{k^2 + m^2} \quad (5)$$

Agora, a configuração de todos os círculos de Ford mostra triângulos circulares que são formados por arcos de círculos mutuamente tangentes (ver Figuras 5.1 e 5.2). Sejam os círculos $C(H/K)$, $C(h/k)$, $C(h_1/k_1)$ com $0 < K < k < k_1$. As frações H/K e h/k aparecem como adjacentes na sequência de Farey de ordem k . A fração h_1/k_1 não está nesta sequência de Farey. No entanto, uma vez que h_1/k_1 é adjacente a H/K bem como a h/k (por causa da tangência dos seus círculos de Ford), deve ser a mediana

$$\frac{h_1}{k_1} = \frac{H + h}{K + k},$$

numa sequência de Farey de ordem superior

$$h_1 = H + h, \quad k_1 = K + k. \quad (6)$$

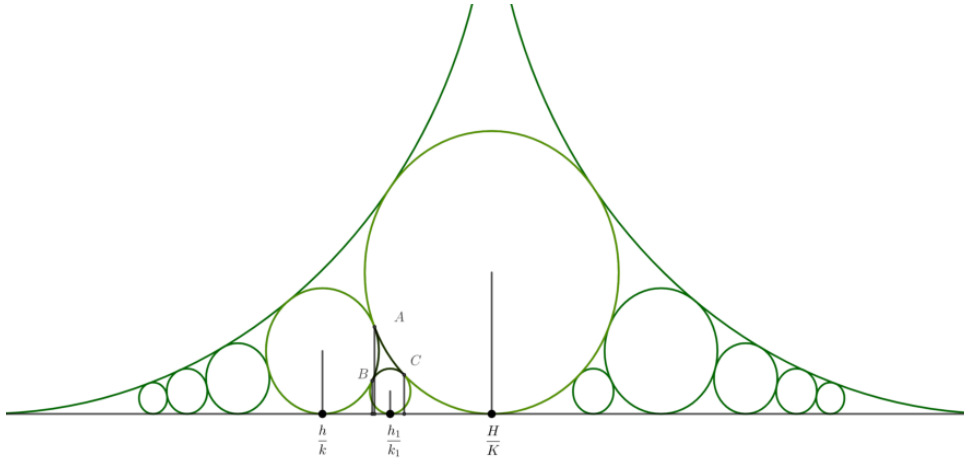


Figura 5.1: Configuração de círculos de Ford.

Voltamos agora à demonstração do Teorema 5.2 (de Hurwitz).

Seja γ irracional. A linha vertical $x = \gamma$ não pode atingir qualquer ponto de tangência dos círculos de Ford. Deve passar, portanto, através do interior de infinitos dos triângulos circulares.

Seja um tal triângulo formado por arcos dos círculos $C(h/k)$, $C(H/K)$ e $C(h_1/k_1)$. Uma vez que a configuração dos círculos de Farey é simétrica em relação à linha $x = \frac{1}{2}$, podemos substituir γ por $1 - \gamma$ se necessário e então assumir, sem perda de generalidade,

$$\frac{h}{k} < \gamma < \frac{H}{K} \quad (7)$$

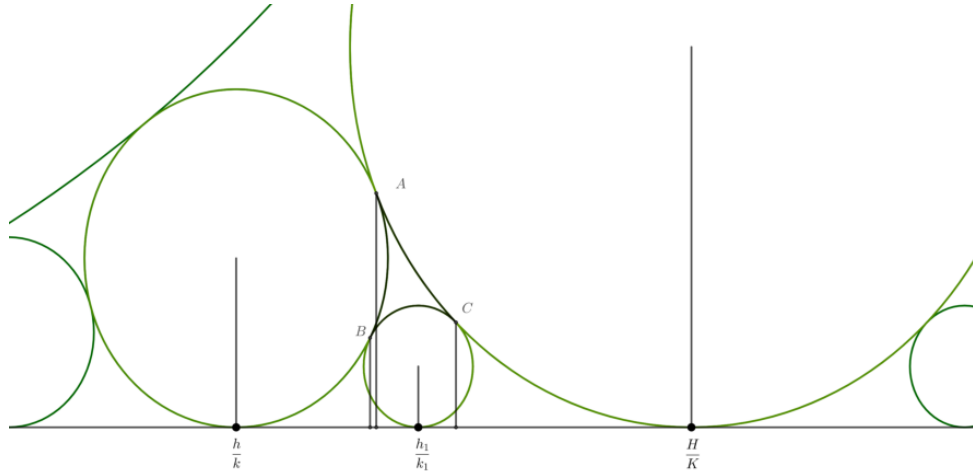


Figura 5.2: Um triângulo circular formado por três círculos de Ford mutuamente tangentes.

e também, claro,

$$\frac{h}{k} < \frac{h_1}{k_1} < \frac{H}{K}. \quad (8)$$

Seja A o ponto de tangência de $C(h/k)$ e $C(H/K)$; B o de $C(h/k)$ e $C(h_1/k_1)$; e C o de $C(h_1/k_1)$ e $C(H/K)$. Sejam a, b, c as coordenadas x de A, B, C . Então, em vista de (5), temos

$$\begin{aligned} a &= \frac{hk + HK}{k^2 + K^2}, \\ b &= \frac{hk + h_1k_1}{k^2 + k_1^2}, \\ c &= \frac{h_1k_1 + HK}{k_1^2 + K^2}. \end{aligned} \quad (9)$$

Assim

$$\begin{aligned} c - a &= \frac{HK(k^2 - k_1^2) + h_1k_1(k^2 + K^2) - hk(k_1^2 + K^2)}{(k^2 + K^2)(k_1^2 + K^2)} \\ &= \frac{kK(Hk - hK) + k_1k(h_1k - hk_1) - Kk_1(Hk_1 - h_1K)}{(k^2 + K^2)(k_1^2 + K^2)} \\ &= \frac{kK + k_1k - Kk_1}{(k^2 + K^2)(k_1^2 + K^2)} \end{aligned}$$

por causa do Teorema 3.1 e da ordenação (8). Finalmente, com (6), obtemos

$$c - a = \frac{kK + k^2 - K^2}{(k^2 + K^2)(k_1^2 + K^2)}.$$

Seja

$$s = \frac{k}{K} > 1.$$

Então temos

$$c - a = \frac{s^2 + s - 1}{K^2(s^2 + 1)((s + 1)^2 + 1)},$$

e, uma vez que

$$s^2 + s - 1 > 1 + 1 - 1 = 1,$$

concluimos

$$c - a > 0.$$

Similarmente, obtemos

$$\begin{aligned} c - b &= \frac{kk_1 + Kk + Kk_1}{(k^2 + k_1^2)(k_1^2 + K^2)} = \frac{K^2 + 3Kk + k^2}{(k^2 + k_1^2)(K^2 + k_1^2)} \\ &= \frac{s^2 + 3s + 1}{K^2((s + 1)^2 + 1)(s^2 + (s + 1)^2)} > 0. \end{aligned}$$

No entanto, $b - a$ pode ser positivo ou negativo. De fato, um cálculo análogo produz

$$b - a = \frac{k^2 - kK - K^2}{(k^2 + K^2)(k^2 + k_1^2)} = \frac{s^2 - s - 1}{K^2(s^2 + 1)(s^2 + (s + 1)^2)}.$$

Aqui

$$s^2 - s - 1 = \left(s - \frac{1 + \sqrt{5}}{2}\right) \left(s - \frac{1 - \sqrt{5}}{2}\right),$$

e, uma vez que

$$s - \frac{1 - \sqrt{5}}{2} > \frac{1}{2} - \frac{1 - \sqrt{5}}{2} = \frac{\sqrt{5}}{2} > 0,$$

vemos que o sinal de $b - a$ é o mesmo que o sinal de

$$s - \frac{1 + \sqrt{5}}{2}.$$

Consideramos agora separadamente os casos $b > a$ e $b < a$ ($b = a$ é impossível, uma vez que o polinômio em (5.2) não pode anular-se para valores racionais de s).

Caso I: $b > a$ ou $s > \frac{1 + \sqrt{5}}{2}$.

Pretendemos mostrar que neste caso

$$\left| \gamma - \frac{H}{K} \right| < \frac{1}{\sqrt{5}K^2}.$$

De fato, porque neste caso a e c são as abscissas extremas do triângulo circular que é atingido por $x = \gamma$, temos

$$a < \gamma < c < \frac{H}{K}$$

e

$$0 < \frac{H}{K} - \gamma < \frac{H}{K} - a = \frac{H}{K} - \frac{hk + HK}{k^2 + K^2}$$

$$= \frac{k}{K(k^2 + K^2)} = \frac{s}{K^2(s^2 + 1)}.$$

Agora

$$\left(s - \frac{1 + \sqrt{5}}{2}\right) \left(s + \frac{1 - \sqrt{5}}{2}\right) > 0,$$

uma vez que o primeiro fator e, a fortiori, o segundo fator são positivos sob a nossa suposição. Assim,

$$s^2 - \sqrt{5}s + 1 > 0$$

ou

$$s^2 + 1 > \sqrt{5}s,$$

$$\frac{s}{s^2 + 1} < \frac{1}{\sqrt{5}},$$

de modo que temos neste caso:

$$0 < \frac{H}{K} - \gamma < \frac{1}{\sqrt{5}K^2},$$

como anunciado.

Caso II: $b < a$ ou $1 < s < \frac{1+\sqrt{5}}{2}$.

Desta vez pretendemos mostrar

$$\left| \frac{h_1}{k_1} - \gamma \right| < \frac{1}{\sqrt{5}k_1^2}.$$

Sob as nossas condições

$$b < a < c$$

e, portanto,

$$b < \gamma < c.$$

Derivamos primeiro que

$$\frac{h_1}{k_1} - b > c - \frac{h_1}{k_1}.$$

Isto é bastante claro geometricamente, uma vez que C está mais alto no círculo $C(h_1/k_1)$ do que B , tendo $C(H/K)$ um raio maior do que o de $C(h/k)$. Explicitamente, temos

$$\frac{h_1}{k_1} - b = \frac{h_1}{k_1} - \frac{hk + h_1k_1}{k^2 + k_1^2} = \frac{k}{k_1(k^2 + k_1^2)}$$

$$c - \frac{h_1}{k_1} = \frac{h_1k_1 + HK}{k_1^2 + K^2} - \frac{h_1}{k_1} = \frac{K}{k_1(K^2 + k_1^2)}$$

e, de fato,

$$\frac{k}{k_1(k^2 + k_1^2)} - \frac{K}{k_1(K^2 + k_1^2)} = \frac{(k - K)(k_1^2 - kK)}{k_1(k^2 + k_1^2)(K^2 + k_1^2)}$$

$$= \frac{k^2 - K^2}{k_1(k^2 + k_1^2)(K^2 + k_1^2)} > 0.$$

Assim, por causa de (5.2), temos

$$\frac{h_1}{k_1} - c < \frac{h_1}{k_1} - \gamma < \frac{h_1}{k_1} - b$$

e, por causa de (5.2),

$$\left| \frac{h_1}{k_1} - \gamma \right| < \frac{h_1}{k_1} - b = \frac{k}{k_1(k^2 + k_1^2)} = \frac{s}{k_1(s^2 + (s+1)^2)}.$$

Aqui, novamente, no entanto,

$$\frac{s}{s^2 + (s+1)^2} < \frac{1}{\sqrt{5}}.$$

De fato, o contrário

$$\frac{s}{s^2 + (s+1)^2} > \frac{1}{\sqrt{5}}$$

implicaria

$$s\sqrt{5} > 2s^2 + 2s + 1$$

ou

$$s^2(\sqrt{5} - 2) + s(\sqrt{5} - 2) - 1 > 0$$

ou

$$\begin{aligned} 0 &< s^2 + s - \frac{1}{\sqrt{5} - 2} = s^2 + s - (\sqrt{5} + 2) \\ &= \left(s + \frac{3}{2} + \frac{\sqrt{5}}{2} \right) \left(s - \frac{1}{2} - \frac{\sqrt{5}}{2} \right), \end{aligned}$$

o que é errado, uma vez que o Caso II pressupõe $s - \frac{1}{2} - \frac{\sqrt{5}}{2} < 0$.

Portanto, em ambos os casos encontramos uma fração l/m tal que

$$\left| \gamma - \frac{l}{m} \right| < \frac{1}{\sqrt{5}m^2}.$$

Aqui l/m foi determinado pelo triângulo ABC atravessado pela linha $x = \gamma$. Como esta linha corta infinitos tais triângulos, há infinitas frações l/m satisfazendo (5.2), o que prova o teorema de Hurwitz.

Capítulo 6

Frações Contínuas

Este capítulo segue de perto a exposição de [10].

6.1 Definição - Notação

O método que se encontra em *Os Elementos de Euclides* serve para se encontrar o máximo divisor comum de dois números.

Basicamente, como veremos a seguir, este método é usado para se converter uma fração em fração contínua. Iniciamos, pois, com um exemplo. Vamos encontrar o máximo divisor comum de 79 e 28 pelo processo das divisões sucessivas.

$$79 = 2 \times 28 + 23$$

$$28 = 1 \times 23 + 5$$

$$23 = 4 \times 5 + 3$$

$$5 = 1 \times 3 + 2$$

$$3 = 1 \times 2 + 1$$

$$2 = 2 \times 1 + 0$$

Logo, $\text{MDC}(79, 28) = 1$, uma vez que 1 é o último resto não-nulo nesta sequência de divisões sucessivas.

Como consequência imediata destas igualdades, podemos expressar o número racional $79/28$ da seguinte forma:

$$\begin{aligned}
\frac{79}{28} &= 2 + \frac{23}{28} = 2 + \frac{1}{\frac{28}{23}} \\
&= 2 + \frac{1}{1 + \frac{5}{23}} = 2 + \frac{1}{1 + \frac{1}{\frac{23}{5}}} \\
&= 2 + \frac{1}{1 + \frac{1}{4 + \frac{3}{5}}} = 2 + \frac{1}{1 + \frac{1}{4 + \frac{1}{\frac{5}{3}}}} \\
&= 2 + \frac{1}{1 + \frac{1}{4 + \frac{1}{1 + \frac{2}{3}}}} = 2 + \frac{1}{1 + \frac{1}{4 + \frac{1}{1 + \frac{1}{\frac{3}{2}}}}} \\
&= 2 + \frac{1}{1 + \frac{1}{4 + \frac{1}{1 + \frac{1}{1 + \frac{1}{2}}}}}
\end{aligned}$$

Uma expressão como esta é chamada de **fração contínua**. Uma maneira mais compacta de escrever esta mesma fração contínua é $[2, 1, 4, 1, 1, 2]$ onde os números $2, 1, 4, 1, 1, 2$ são chamados de **quocientes parciais**. Estes números são os quocientes que aparecem na sequência de divisões sucessivas para o cálculo do máximo divisor comum de 79 e 28. Chamamos **fração contínua simples** a uma expressão onde todos os numeradores são iguais a 1.

1 Seja $\frac{19}{11}$. Temos,

$$\begin{aligned}
\frac{19}{11} &= 1 + \frac{8}{11} = 1 + \frac{1}{\frac{11}{8}} = 1 + \frac{1}{1 + \frac{3}{8}} = 1 + \frac{1}{1 + \frac{1}{\frac{8}{3}}} = 1 + \frac{1}{1 + \frac{1}{2 + \frac{2}{3}}} \\
&= 1 + \frac{1}{1 + \frac{1}{2 + \frac{1}{\frac{3}{2}}}} = 1 + \frac{1}{1 + \frac{1}{2 + \frac{1}{1 + \frac{1}{2}}}}.
\end{aligned}$$

Logo, $19/11 = [1, 1, 2, 1, 2]$.

É de imediata verificação que toda fração contínua finita representa um número racional. Embora tenhamos afirmado que todo racional pode ser expresso sob a forma de uma fração contínua finita, esta representação não é única. Isto porque

$$[a_1, a_2, \dots, a_{n-1}, a_n] = [a_1, a_2, \dots, a_{n-1}, a_n - 1, 1],$$

se $a_n > 1$.

Por exemplo,

$$[1, 2, 3] = 1 + \frac{1}{2 + \frac{1}{3}} = 1 + \frac{1}{\frac{7}{3}} = 1 + \frac{3}{7} = \frac{10}{7}.$$

Por outro lado,

$$[1, 2, 2, 1] = 1 + \frac{1}{2 + \frac{1}{2 + \frac{1}{1}}} = 1 + \frac{1}{2 + \frac{1}{3}} = \frac{10}{7}.$$

Portanto, $[1, 2, 3] = [1, 2, 2, 1]$.

e, portanto, $[2, 1, 4, 5, 3] = [2, 1, 4, 5, 2, 1]$.

De modo geral, se $a_n > 1$, temos

$$[a_1, a_2, \dots, a_n] = [a_1, a_2, \dots, a_n - 1, 1].$$

As observações que acabamos de fazer podem ser resumidas no seguinte teorema.

Teorema 6.1 *Todo número racional pode ser representado de duas maneiras distintas sob a forma de fração contínua e toda fração contínua (simples) finita representa um número racional.*

Chamamos a atenção do leitor para o fato de que a unicidade da representação de um número racional em fração contínua (a menos da modificação do último termo a_n) é garantida pelo Teorema 6.1.

A representação de $19/11$ em fração contínua é dada por $[1, 1, 2, 1, 2]$ e como

$$\frac{11}{19} = 0 + \frac{1}{\frac{19}{11}} = 0 + \frac{1}{1 + \frac{1}{\frac{1}{2 + \frac{1}{1 + \frac{1}{2}}}}}$$

temos que $11/19 = [0, 1, 1, 2, 1, 2]$.

Na realidade temos, em geral, que se a representação em fração contínua do racional $\frac{p}{q}$ ($p > q$) é dada por $[a_1, a_2, \dots, a_n]$, então a representação de $\frac{q}{p}$ é dada por $[0, a_1, a_2, \dots, a_n]$.

Isto é consequência imediata do fato de

$$\frac{q}{p} = 0 + \frac{1}{\frac{p}{q}}.$$

6.2 Convergentes

Vimos que qualquer número racional pode ser representado sob a forma de uma fração contínua (simples)

$$\frac{p}{q} = [a_1, a_2, \dots, a_{n-1}, a_n]$$

onde a_1 é um inteiro positivo, negativo ou zero, e $a_2, a_3, \dots, a_n$ são inteiros positivos.

Consideremos as frações

$$c_1 = \frac{a_1}{1}, \quad c_2 = a_1 + \frac{1}{a_2}, \quad c_3 = a_1 + \frac{1}{a_2 + \frac{1}{a_3}}, \dots$$

obtidas pelas expansões das frações contínuas

$$[a_1], [a_1, a_2], [a_1, a_2, a_3], \dots$$

Estas frações são chamadas de primeiro, segundo, terceiro, ... convergentes, respectivamente, da fração contínua $[a_1, a_2, a_3, \dots, a_{n-1}, a_n]$.

É claro que o n -ésimo convergente é igual à própria fração contínua. Nos teoremas seguintes mostramos algumas propriedades satisfeitas pelos convergentes de uma fração contínua.

Se considerarmos

$$c_1 = \frac{a_1}{1} = \frac{p_1}{q_1} \quad \text{onde} \quad p_1 = a_1 \quad \text{e} \quad q_1 = 1$$

teremos,

$$c_2 = a_1 + \frac{1}{a_2} = \frac{a_1 a_2 + 1}{a_2} = \frac{p_2}{q_2}$$

onde $p_2 = a_1 a_2 + 1$ e $q_2 = a_2$.

Se calcularmos c_3, c_4, c_5 , obtemos, respectivamente:

$$c_3 = \frac{a_3 p_2 + p_1}{a_3 q_2 + q_1} = \frac{p_3}{q_3}$$

$$c_4 = \frac{a_4 p_3 + p_2}{a_4 q_3 + q_2} = \frac{p_4}{q_4}$$

$$c_5 = \frac{a_5 p_4 + p_3}{a_5 q_4 + q_3} = \frac{p_5}{q_5}$$

Observando estes resultados podemos conjecturar que os numeradores p_i 's e os denominadores q_i 's dos convergentes c_i 's satisfazem as seguintes relações:

$$p_i = a_i p_{i-1} + p_{i-2}$$

$$q_i = a_i q_{i-1} + q_{i-2}$$

No teorema abaixo provamos, por indução, que estas relações se verificam para $i = 3, 4, 5, \dots, n$.

Teorema 6.2 *Seja $c_i = p_i/q_i$ o i -ésimo convergente da fração contínua $[a_1, a_2, \dots, a_n]$. Então o numerador p_i e o denominador q_i de c_i satisfazem as seguintes relações:*

$$p_i = a_i p_{i-1} + p_{i-2}$$

$$q_i = a_i q_{i-1} + q_{i-2}$$

para $i = 3, 4, 5, \dots, n$; onde

$$p_1 = a_1, \quad p_2 = a_2 a_1 + 1, \quad q_1 = 1, \quad q_2 = a_2.$$

Demonstração: Como já vimos o teorema é válido para $i = 3$, i.e.,

$$c_3 = \frac{p_3}{q_3} = \frac{a_3 p_2 + p_1}{a_3 q_2 + q_1}.$$

Assumimos, agora, que a relação do Teorema 6.2 se verifica para todo $j \leq i$ com $3 \leq i < n$. Isto significa que

$$c_i = [a_1, a_2, \dots, a_i] = \frac{p_i}{q_i} = \frac{a_i p_{i-1} + p_{i-2}}{a_i q_{i-1} + q_{i-2}}$$

Observando que,

$$c_i = a_1 + \frac{1}{a_2 + \frac{1}{a_3 + \cdots + \frac{1}{a_{i-1} + \frac{1}{a_i}}}}$$

e que,

$$c_{i+1} = a_1 + \frac{1}{a_2 + \frac{1}{a_3 + \cdots + \frac{1}{a_{i-1} + \frac{1}{a_i + \frac{1}{a_{i+1}}}}}}$$

vemos que c_{i+1} pode ser obtido de c_i simplesmente pela substituição de a_i por $a_i + \frac{1}{a_{i+1}}$. Isto nos diz que se pudermos mostrar que os números p_{i-1} , p_{i-2} , q_{i-1} e q_{i-2} dependem somente dos quocientes parciais $a_1, a_2, \dots, a_{i-1}$, poderemos usar (6.2) para a obtenção de c_{i+1} pois estamos assumindo, como hipótese de indução, a validade de (6.2) para todo $j \leq i$. Como

$$\frac{p_{i-1}}{q_{i-1}} = \frac{a_{i-1}p_{i-2} + p_{i-3}}{a_{i-1}q_{i-2} + q_{i-3}}$$

os números p_{i-1} e q_{i-1} dependem somente dos números a_{i-1} e dos números $p_{i-2}, q_{i-2}, p_{i-3}, q_{i-3}$ os quais dependem dos precedentes a 's, p 's e q 's. Desta forma $p_{i-2}, q_{i-2}, p_{i-3}$ e q_{i-3} dependem somente dos primeiros $i-1$ quocientes parciais $a_1, a_2, \dots, a_{i-1}$ sendo independentes de a_i . Portanto eles não serão alterados com a substituição de a_i por $a_i + \frac{1}{a_{i+1}}$. Podemos, portanto, utilizar (6.2) para a obtenção de c_{i+1} bastando, para isto, substituir a_i por $a_i + \frac{1}{a_{i+1}}$

$$\begin{aligned} c_{i+1} &= \frac{\left(a_i + \frac{1}{a_{i+1}}\right) p_{i-1} + p_{i-2}}{\left(a_i + \frac{1}{a_{i+1}}\right) q_{i-1} + q_{i-2}} \\ &= \frac{(a_{i+1}a_i + 1)p_{i-1} + a_{i+1}p_{i-2}}{(a_{i+1}a_i + 1)q_{i-1} + a_{i+1}q_{i-2}} \\ &= \frac{a_{i+1}(a_i p_{i-1} + p_{i-2}) + p_{i-1}}{a_{i+1}(a_i q_{i-1} + q_{i-2}) + q_{i-1}} \\ &= \frac{a_{i+1}p_i + p_{i-1}}{a_{i+1}q_i + q_{i-1}} \\ &= \frac{p_{i+1}}{q_{i+1}} \end{aligned}$$

o que conclui a demonstração por indução. $\square$

Por uma questão de conveniência vamos definir $p_0 = 1$, $p_{-1} = 0$, $q_0 = 0$, $q_{-1} = 1$. Com estas definições as equações do Teorema 6.2

$$p_i = a_i p_{i-1} + p_{i-2}$$

$$q_i = a_i q_{i-1} + q_{i-2}$$

passam a ser verdadeiras para todo $i = 1, 2, 3, \dots$

A relação obtida no próximo teorema nos permitirá deduzir, facilmente, que para todo convergente $c_i = p_i/q_i$ temos que $\text{mdc}(p_i, q_i) = 1$.

Teorema 6.3 *A relação*

$$p_i q_{i-1} - p_{i-1} q_i = (-1)^i$$

se verifica para todo $i \geq 0$, onde p_i e q_i são, respectivamente, o numerador e o denominador do i -ésimo convergente.

Demonstração: Para $i = 0$ temos $p_0 q_{-1} - p_{-1} q_0 = 1 = (-1)^0$ uma vez que $p_0 = q_{-1} = 1$ e $p_{-1} = q_0 = 0$.

Vamos assumir, como hipótese de indução, a validade da relação para i e mostrar que a mesma relação também se verifica quando substituirmos i por $i + 1$.

Sabemos, do teorema anterior, que

$$p_{i+1} = a_{i+1} p_i + p_{i-1} \quad \text{e} \quad q_{i+1} = a_{i+1} q_i + q_{i-1}.$$

Logo,

$$\begin{aligned} p_{i+1} q_i - p_i q_{i+1} &= (a_{i+1} p_i + p_{i-1}) q_i - p_i (a_{i+1} q_i + q_{i-1}) \\ &= a_{i+1} p_i q_i + p_{i-1} q_i - a_{i+1} p_i q_i - p_i q_{i-1} \\ &= -(p_i q_{i-1} - p_{i-1} q_i) \end{aligned}$$

Utilizando, pois, a hipótese de indução, obtemos

$$p_{i+1} q_i - p_i q_{i+1} = -(-1)^i = (-1)^{i+1}$$

o que conclui a demonstração. $\square$

Corolário 6.4 *Para todo convergente $c_i = \frac{p_i}{q_i}$ temos que $\text{mdc}(p_i, q_i) = 1$.*

Demonstração: Pelo teorema temos que $p_i q_{i-1} - p_{i-1} q_i = (-1)^i$. Isto nos diz que qualquer divisor comum de p_i e q_i deve ser um divisor de 1 ou -1. Logo o máximo divisor comum de p_i e q_i deve ser igual a 1. $\square$

6.3 Aproximações Sucessivas

Nesta seção descrevemos um processo de obtenção de aproximações sucessivas, por racionais, para um número irracional.

Seja α um irracional e seja $a_1 = \lfloor \alpha \rfloor$, isto é, a_1 é o maior inteiro menor do que α . Logo,

$$\alpha = a_1 + \frac{1}{x_1},$$

e, claramente, $x_1 = \frac{1}{\alpha - a_1}$ é irracional e $x_1 > 1$. Podemos, pois, escrever x_1 na forma

$$x_1 = a_2 + \frac{1}{x_2}$$

onde $a_2 = \lfloor x_1 \rfloor$, x_2 irracional e $x_2 > 1$. Podemos repetir este processo, obtendo:

$$\alpha = a_1 + \frac{1}{x_1}$$

$$\begin{aligned}
x_1 &= a_2 + \frac{1}{x_2} \\
x_2 &= a_3 + \frac{1}{x_3} \\
&\vdots \\
x_n &= a_{n+1} + \frac{1}{x_{n+1}}
\end{aligned}$$

onde todos os a_i 's ($i > 1$) são inteiros maiores ou iguais a 1 e todos os x_i 's são irracionais maiores do que 1. O fato de cada x_i ser irracional nos garante que este processo pode ser repetido um número qualquer de vezes. Utilizando as equações (6.3) vemos que

$$\begin{aligned}
\alpha &= a_1 + \frac{1}{x_1} = a_1 + \frac{1}{a_2 + \frac{1}{x_2}} = a_1 + \frac{1}{a_2 + \frac{1}{a_3 + \frac{1}{x_3}}} \\
&= a_1 + \frac{1}{a_2 + \frac{1}{a_3 + \frac{1}{a_4 + \frac{1}{x_4}}}} \\
&= \dots
\end{aligned}$$

Definimos $[a_1, a_2, a_3, \dots] = \lim_{n \rightarrow \infty} [a_1, a_2, \dots, a_n]$. Vamos ilustrar este processo obtendo a expansão de $\sqrt{3}$. Sendo $a_1 = \lfloor \sqrt{3} \rfloor = 1$ e

$$\sqrt{3} = a_1 + \frac{1}{x_1} = 1 + \frac{1}{x_1}$$

temos

$$x_1 = \frac{1}{\sqrt{3} - 1} = \frac{1}{(\sqrt{3} - 1)(\sqrt{3} + 1)} = \frac{\sqrt{3} + 1}{2}.$$

Consequentemente

$$\sqrt{3} = 1 + \frac{1}{x_1} = 1 + \frac{1}{\frac{\sqrt{3}+1}{2}}.$$

Como $a_2 = \lfloor \frac{\sqrt{3}+1}{2} \rfloor = 1$ temos

$$\frac{\sqrt{3}+1}{2} = 1 + \frac{1}{x_2}$$

donde obtemos

$$x_2 = \frac{1}{\frac{\sqrt{3}+1}{2} - 1} = \sqrt{3} + 1.$$

Logo,

$$\begin{aligned}
\sqrt{3} &= 1 + \frac{1}{x_1} = 1 + \frac{1}{1 + \frac{1}{\sqrt{3}+1}} \\
&= 1 + \frac{1}{1 + \frac{1}{\sqrt{3}-1}}.
\end{aligned}$$

Como $a_3 = \lfloor \sqrt{3} + 1 \rfloor = 2$, temos

$$\sqrt{3} + 1 = x_2 = a_3 + \frac{1}{x_3} = 2 + \frac{1}{x_3}.$$

Resolvendo esta última equação para x_3 obtemos:

$$x_3 = \frac{1}{(\sqrt{3} + 1 - 2)} = \frac{\sqrt{3} + 1}{2}.$$

Sendo $x_3 = x_1$ concluímos que x_4 será igual a x_2 e desta forma, continuando com este processo, iremos obter para a sequência $a_1, a_2, a_3, \dots$ os valores $1, 1, 2, 1, 2, 1, 2, \dots$. Logo a fração contínua infinita representando $\sqrt{3}$ será dada por:

$$\sqrt{3} = [1, 1, 2, 1, 2, 1, 2, \dots] = [1, \overline{1, 2}].$$

Chamamos **fração contínua periódica** a uma representação como esta em que uma sequência de números se repete periodicamente. Colocamos uma barra sobre a parte que se repete que é chamada de **período** da fração contínua.

Para $\alpha = \sqrt{6}$ obtemos a seguinte sequência

$$a_1 = \lfloor \sqrt{6} \rfloor = 2$$

$$x_1 = \frac{1}{\sqrt{6} - 2} = \frac{\sqrt{6} + 2}{2}$$

$$a_2 = \lfloor \frac{\sqrt{6} + 2}{2} \rfloor = 2$$

$$x_2 = \frac{1}{\frac{\sqrt{6} + 2}{2} - 2} = \sqrt{6} + 2$$

$$a_3 = \lfloor \sqrt{6} + 2 \rfloor = 4$$

$$x_3 = \frac{1}{(\sqrt{6} + 2) - 4} = \frac{\sqrt{6} + 2}{2} = x_1$$

Como $x_3 = x_1$, vemos que $a_4 = a_2, a_5 = a_3, a_6 = a_2, a_7 = a_3, \dots$. Logo,

$$\sqrt{6} = [2, 2, 4, 2, 4, \dots] = [2, \overline{2, 4}].$$

Dada uma fração contínua periódica podemos reverter o processo acima para a obtenção do número irracional representado por ela.

$$[2, \overline{2, 4}] = 2 + \frac{1}{2 + \frac{1}{4 + \frac{1}{2 + \frac{1}{4 + \frac{1}{2 + \frac{1}{4 + \ddots}}}}}}$$

onde

$$y = 2 + \frac{1}{4 + \frac{1}{2 + \frac{1}{4 + \frac{1}{2 + \frac{1}{4 + \ddots}}}}}$$

Desta última igualdade observamos que,

$$y = 2 + \frac{1}{4 + \frac{1}{y}}$$

da qual obtemos a equação $2y^2 - 4y - 1 = 0$, ou seja, $y = \frac{2+\sqrt{6}}{2}$, uma vez que y é positivo. Logo,

$$\begin{aligned} [2, \overline{2, 4}] &= 2 + \frac{1}{y} = 2 + \frac{1}{\left(\frac{2+\sqrt{6}}{2}\right)} \\ &= 2 + \frac{2}{2 + \sqrt{6}} = 2 + \frac{2(2 - \sqrt{6})}{-2} \\ &= \frac{-4 + 4 - 2\sqrt{6}}{-2} = \sqrt{6}. \end{aligned}$$

Até o presente momento já vimos que toda fração contínua finita representa um racional, que todo racional é representado por uma fração contínua finita e que um irracional é representado por uma fração contínua infinita.

Vale mencionar que nem todo irracional possui uma representação periódica quando representado sob a forma de fração contínua. O número π possui a seguinte representação

$$\pi = [3, 7, 15, 1, 292, 1, 1, 1, 2, 1, 3, \dots]$$

onde não há nenhuma sequência (período) que se repete. **Lagrange**, em 1770, caracterizou todos os irracionais que possuem representação periódica quando expressos sob a forma de fração contínua. Ele mostrou que a fração contínua infinita que representa um irracional é periódica se, e somente se, este irracional for raiz de um polinômio da forma $ax^2 + bx + c = 0$ onde a, b e c são inteiros. Este resultado nos diz, em particular, que somente irracionais algébricos podem ter representação periódica.

6.4 Propriedades dos Convergentes

No teorema seguinte mostramos algumas importantes propriedades da sequência $c_1, c_2, c_3, \dots$ dos convergentes de uma fração contínua.

Teorema 6.5 *A sequência $c_1, c_2, c_3, \dots$ dos convergentes de uma fração contínua satisfaz as seguintes propriedades:*

- (i) $c_1 < c_3 < c_5 < c_7 < \dots < c_{2n+1}$
- (ii) $c_2 > c_4 > c_6 > \dots > c_{2n}$
- (iii) $c_{2n+1} < c_{2n+2} < c_{2n+3}$

Demonstração: Pelo Teorema 6.3 temos que

$$p_i q_{i-1} - p_{i-1} q_i = (-1)^i$$

a qual é válida independentemente da fração contínua ser finita ou não. Dividindo ambos os lados desta igualdade por $q_i q_{i-1}$ obtemos:

$$\frac{p_i}{q_i} - \frac{p_{i-1}}{q_{i-1}} = \frac{(-1)^i}{q_i q_{i-1}}.$$

Como $c_i = \frac{p_i}{q_i}$ temos que

$$c_i - c_{i-1} = \frac{(-1)^i}{q_i q_{i-1}} \quad (1)$$

Mas sendo

$$c_i - c_{i-2} = \frac{p_i}{q_i} - \frac{p_{i-2}}{q_{i-2}} = \frac{p_i q_{i-2} - p_{i-2} q_i}{q_i q_{i-2}} \quad (2)$$

obtemos, usando o fato de $p_i = a_i p_{i-1} + p_{i-2}$ e $q_i = a_i q_{i-1} + q_{i-2}$, que

$$\begin{aligned} c_i - c_{i-2} &= \frac{(a_i p_{i-1} + p_{i-2}) q_{i-2} - p_{i-2} (a_i q_{i-1} + q_{i-2})}{q_i q_{i-2}} \\ &= \frac{a_i (p_{i-1} q_{i-2} - p_{i-2} q_{i-1})}{q_i q_{i-2}} \\ &= \frac{a_i (-1)^{i-1}}{q_i q_{i-2}} \end{aligned}$$

isto é,

$$c_i - c_{i-2} = \frac{a_i (-1)^{i-1}}{q_i q_{i-2}} \quad (3)$$

Tomando $i = 2$ e $i = 3$ em (1) obtemos, respectivamente,

$$c_2 - c_1 = \frac{1}{q_2 q_1} > 0 \quad \text{e} \quad c_3 - c_2 = \frac{-1}{q_3 q_2} < 0$$

uma vez que todos os q_i 's são positivos.

A equação (3) nos diz que para $i = 3$ temos

$$c_3 - c_1 = \frac{a_3}{q_3 q_1} > 0$$

pois a_3, q_3 e q_1 são, todos, positivos. Sendo, pois, $c_1 < c_3, c_1 < c_2$ e $c_2 > c_3$ temos que $c_1 < c_3 < c_2$. Usando, agora, $i = 3$ e $i = 4$ em (1) e $i = 4$ em (3) obtemos $c_3 < c_4 < c_2$. Repetindo este processo obtemos a sequência de desigualdades:

$$\begin{aligned} c_5 &< c_6 < c_4 \\ c_7 &< c_8 < c_6 \\ &\vdots \end{aligned}$$

Combinando estas desigualdades obtemos

$$c_1 < c_3 < c_5 < c_7 < \cdots < c_{2i+1} < \cdots < c_{2n} < \cdots < c_6 < c_4 < c_2$$

o que conclui a demonstração. ■

O que acabamos de mostrar é que a sequência dos convergentes de índice ímpar forma uma sequência crescente que é limitada superiormente e que a sequência dos convergentes de índice par forma uma sequência decrescente e limitada inferiormente.

Há um resultado fundamental em Análise que diz que toda sequência crescente e limitada superiormente converge e que toda sequência decrescente e limitada inferiormente também converge.

Sejam, pois, ℓ_i o limite da sequência $c_1, c_3, c_5, \dots, c_{2i+1}, \dots$ e ℓ_p o limite da sequência $c_2, c_4, c_6, \dots, c_{2i}, \dots$

Como os números q_i 's são calculados através da relação $q_i = a_i q_{i-1} + q_{i-2}$ e os números $a_i (i \geq 2)$ e $q_i (i \geq 1)$ são todos positivos concluímos que a sequência dos q_i 's cresce indefinidamente. Isto implica que se tomarmos $i = 2j$ em (1), teremos

$$c_{2j} - c_{2j-1} = \frac{1}{q_{2j} q_{2j-1}}$$

o que nos permite concluir que $\lim_{j \rightarrow \infty} (c_{2j} - c_{2j-1}) = 0$.

Sendo $\lim_{j \rightarrow \infty} c_{2j-1} = \ell_i$ e $\lim_{j \rightarrow \infty} c_{2j} = \ell_p$ concluímos que $\ell_p = \ell_i$.

Mencionamos, anteriormente, a possibilidade de se obter um processo de aproximações sucessivas, por racionais, para um número irracional. Já vimos como obter a representação sob a forma de fração contínua para um irracional e, também, que a sequência $c_1, c_2, c_3, \dots$ dos convergentes de uma fração contínua converge.

Provamos, a seguir, que o limite ℓ para o qual a sequência dos convergentes converge é, na realidade, o número irracional que deu origem à fração contínua. Embora este fato pareça óbvio ele precisa ser provado. Para isto necessitamos do seguinte resultado.

Teorema 6.6 *Para qualquer número real α temos:*

$$[a_1, a_2, \dots, a_{i-1}, \alpha] = \frac{\alpha p_{i-1} + p_{i-2}}{\alpha q_{i-1} + q_{i-2}} \quad (4)$$

onde $a_1, a_2, a_3, \dots$ é uma sequência infinita de inteiros positivos com a possível exceção de a_1 e as sequências dos p_i 's e q_i 's são dadas pelo Teorema 6.2, isto é,

$$p_0 = 1, p_{-1} = 0, q_0 = 0, q_{-1} = 1$$

$$p_i = a_i p_{i-1} + p_{i-2}$$

$$q_i = a_i q_{i-1} + q_{i-2}, \quad i \geq 1.$$

Demonstração: Para $n = 1$ o resultado deve ser visto como

$$\alpha = \frac{\alpha p_0 - p_{-1}}{\alpha q_0 + q_{-1}}$$

o qual é verdadeiro pelas condições iniciais.

Para $n = 2$ temos

$$[a_1, \alpha] = \frac{\alpha p_1 + p_0}{\alpha q_1 + q_0} = \frac{\alpha a_1 + 1}{\alpha}$$

o que é verdadeiro uma vez que $[a_1, \alpha] = a_1 + \frac{1}{\alpha}$.

Estabelecemos, agora, o resultado por indução. Considerando verdadeiro o resultado para $[a_1, a_2, \dots, a_{i-1}, \alpha]$ temos

$$\begin{aligned} [a_1, a_2, \dots, a_{i-1}, a_i + \frac{1}{\alpha}] &= \frac{(a_i + \frac{1}{\alpha})p_{i-1} + p_{i-2}}{(a_i + \frac{1}{\alpha})q_{i-1} + q_{i-2}} \\ &= \frac{\alpha(a_i p_{i-1} + p_{i-2}) + p_{i-1}}{\alpha(a_i q_{i-1} + q_{i-2}) + q_{i-1}} \\ &= \frac{\alpha p_i + p_{i-1}}{\alpha q_i + q_{i-1}} \end{aligned}$$

o que prova (4). ■

Consideramos, agora, a sequência $a_1, a_2, a_3, \dots$ dada por (6.3) e a sequência dos convergentes $c_i = p_i/q_i$. Sabemos que,

$$\begin{aligned} \alpha &= a_1 + \frac{1}{x_1} = [a_1, x_1] = [a_1, a_2 + \frac{1}{x_2}] \\ &= [a_1, a_2, x_2] = [a_1, a_2, a_3 + \frac{1}{x_3}] \\ &= [a_1, a_2, a_3, x_3] = [a_1, a_2, \dots, a_{i-1} + \frac{1}{x_{i-1}}] \\ &= [a_1, a_2, a_3, \dots, a_{i-1}, x_{i-1}] \end{aligned}$$

e pelo Teorema 6.6 temos

$$\alpha = [a_1, a_2, a_3, \dots, a_{i-1}, x_{i-1}] = \frac{x_{i-1}p_{i-1} + p_{i-2}}{x_{i-1}q_{i-1} + q_{i-2}}.$$

Tomando a diferença

$$\begin{aligned} \alpha - c_{i-1} &= \frac{x_{i-1}p_{i-1} + p_{i-2}}{x_{i-1}q_{i-1} + q_{i-2}} - \frac{p_{i-1}}{q_{i-1}} \\ &= \frac{-(p_{i-1}q_{i-2} - p_{i-2}q_{i-1})}{q_{i-1}(x_{i-1}q_{i-1} + q_{i-2})} \end{aligned} \tag{5}$$

podemos concluir que $\lim_{i \rightarrow \infty} (\alpha - c_{i-1}) = 0$ uma vez que a sequência dos q_i é crescente e os números x_i 's são positivos.

Portanto $\alpha = \lim_{i \rightarrow \infty} c_i = \lim_{i \rightarrow \infty} [a_1, a_2, \dots, a_i] = [a_1, a_2, a_3, \dots]$ ou seja, o limite da sequência dos convergentes da representação do irracional α sob a forma de fração contínua é igual ao próprio α .

O Teorema 6.6 nos permite provar que toda fração contínua simples infinita representa um irracional. Isto é o que mostramos no próximo teorema.

Teorema 6.7 *Toda fração contínua simples infinita $[a_1, a_2, a_3, \dots]$ representa um irracional.*

Demonstração: Denotando $[a_1, a_2, a_3, \dots]$ por α nós observamos, pelo Teorema 6.5, que α está entre c_i e c_{i+1} e, portanto, $0 < |\alpha - c_i| < |c_{i+1} - c_i|$.

Multiplicando por q_i esta desigualdade e utilizando (1) temos

$$0 < |\alpha q_i - p_i| < |c_{i+1} q_i - c_i q_i| < \frac{1}{q_{i+1}}.$$

Supondo α racional, isto é, $\alpha = \frac{a}{b}$, a e b inteiros com $b > 0$, a desigualdade acima, após multiplicação por b nos fornece

$$|a q_i - b p_i| < \frac{b}{q_{i+1}}.$$

Como a sequência dos q_i é crescente podemos escolher i suficientemente grande de forma que $b < q_{i+1}$.

Desta forma o inteiro $a q_i - b p_i$ estaria entre 0 e 1, o que é impossível. ■

Nos próximos teoremas são provados dois importantes resultados. Mostramos que todo convergente $\frac{p_n}{q_n}$ de α satisfaz

$$\left| \alpha - \frac{p_n}{q_n} \right| < \frac{1}{q_n^2}$$

e que se o racional a/b , $b > 0$ verificar

$$\left| \alpha - \frac{a}{b} \right| < \frac{1}{2b^2}$$

então a/b é um dos convergentes da representação de α em fração contínua.

Teorema 6.8 *Todo convergente $c_i = \frac{p_i}{q_i}$ de α satisfaz*

$$\left| \alpha - \frac{p_i}{q_i} \right| < \frac{1}{q_i^2}.$$

Demonstração: Por (6.3) temos que $a_{n+1} < x_n$. Este fato, juntamente com (5), nos fornece

$$\left| \alpha - \frac{p_i}{q_i} \right| = \frac{1}{q_i(x_i q_i + q_{i-1})} < \frac{1}{q_i(a_{i+1} q_i + q_{i-1})}.$$

Usando o Teorema 6.2 substituímos $a_{i+1} q_i + q_{i-1}$ por q_{i+1} obtendo

$$\left| \alpha - \frac{p_i}{q_i} \right| < \frac{1}{q_i q_{i+1}} < \frac{1}{q_i^2}$$

onde usamos o fato de $q_i < q_{i+1}$ uma vez que a sequência dos q_i é estritamente crescente. ■

Teorema 6.9 *Seja α um irracional e p_n/q_n os convergentes da expansão de α em fração contínua. Se a/b for um racional com $b > 0$ tal que*

$$\left| \alpha - \frac{a}{b} \right| < \left| \alpha - \frac{p_n}{q_n} \right|$$

para algum $n \geq 1$, então $b > q_n$. Mais ainda, se $|\alpha b - a| < |\alpha q_n - p_n|$ para algum $n \geq 0$, então $b \geq q_{n+1}$.

Demonstração: Inicialmente mostramos que a segunda parte do teorema implica a primeira. Supondo falsa a primeira parte teremos a existência de um racional a/b com

$$\left| \alpha - \frac{a}{b} \right| < \left| \alpha - \frac{p_n}{q_n} \right| \quad \text{e} \quad b \leq q_n.$$

Destas desigualdades obtemos

$$|\alpha b - a| < |\alpha q_n - p_n|.$$

Mas a segunda parte do teorema nos diz que isto implica $b \geq q_{n+1}$. Logo temos uma contradição pois $q_n < q_{n+1}$ para $n \geq 0$.

Para provarmos a segunda parte procedemos, novamente, por contradição. Suponhamos $|\alpha b - a| < |\alpha q_n - p_n|$ e $b < q_{n+1}$.

Consideremos o sistema linear em x e y

$$\begin{aligned} p_n x + p_{n+1} y &= a, \\ q_n x + q_{n+1} y &= b, \\ p_n x + p_{n+1} y &= a. \end{aligned} \tag{6}$$

Pelo Teorema 6.3 o determinante principal deste sistema é ± 1 e, consequentemente, este sistema possui uma solução em inteiros x e y . Na realidade, ambos x e y , são diferentes de zero. Isto porque se $x = 0$ então $b = y q_{n+1}$ o que implica $y > 0$ e $b \geq q_{n+1}$, em contradição com $b < q_{n+1}$.

Se $y = 0$ então $a = x p_n, b = x q_n$ e

$$\begin{aligned} |\alpha b - a| &= |\alpha x q_n - x p_n| \\ &= |x| |\alpha q_n - p_n| \geq |\alpha q_n - p_n| \end{aligned}$$

uma vez que $|x| \geq 1$ o que nos dá, novamente, uma contradição.

A seguir mostramos que x e y possuem sinais opostos. Se $y < 0$, então $x q_n = b - y q_{n+1}$, isto é, $x > 0$.

Se $y > 0$ então $b < y q_{n+1}$ pois $b < q_{n+1}$. Portanto $x q_n$ é negativo o que nos diz que $x < 0$.

Concluimos, agora, de (5) que $\alpha q_n - p_n$ e $\alpha q_{n+1} - p_{n+1}$ possuem sinais opostos e, portanto, $x(\alpha q_n - p_n)$ e $y(\alpha q_{n+1} - p_{n+1})$ possuem o mesmo sinal.

Do sistema (6) obtemos $\alpha b - a = x(\alpha q_n - p_n) + y(\alpha q_{n+1} - p_{n+1})$.

Desta forma, como os dois termos da direita da igualdade acima possuem o mesmo sinal, temos

$$|\alpha b - a| = |x(\alpha q_n - p_n) + y(\alpha q_{n+1} - p_{n+1})|$$

$$\begin{aligned}
&= |x(\alpha q_n - p_n)| + |y(\alpha q_{n+1} - p_{n+1})| \\
&> |x(\alpha q_n - p_n)| = |x| |\alpha q_n - p_n| \\
&\geq |\alpha q_n - p_n|.
\end{aligned}$$

Como isto nos dá uma contradição, o teorema está provado. ■

Teorema 6.10 (Lagrange) *Seja α um número irracional. Se existir um racional $a/b, b \geq 1$, tal que*

$$\left| \alpha - \frac{a}{b} \right| < \frac{1}{2b^2}$$

então a/b é um dos convergentes da expansão de α em fração contínua.

Demonstração: Suponhamos que exista a/b um racional satisfazendo a hipótese do teorema e que a/b não seja um convergente da fração contínua de α . Sem perda de generalidade podemos supor $\text{mdc}(a, b) = 1$. Seja n o inteiro tal que $q_n \leq b < q_{n+1}$. Para este inteiro n a desigualdade $|\alpha b - a| < |\alpha q_n - p_n|$ é impossível pelo Teorema 6.9. Logo

$$|\alpha q_n - p_n| \leq |\alpha b - a| < \frac{1}{2b},$$

$$\left| \alpha - \frac{p_n}{q_n} \right| < \frac{1}{2bq_n}.$$

Usando o fato de que $a/b \neq p_n/q_n$ e que $bp_n - aq_n$ é um inteiro não nulo obtemos

$$\begin{aligned}
\frac{1}{bq_n} &\leq \frac{|bp_n - aq_n|}{bq_n} = \frac{\left| \frac{p_n}{q_n} - \frac{a}{b} \right|}{bq_n} \\
&\leq \frac{\left| \alpha - \frac{a}{b} \right| + \left| \alpha - \frac{p_n}{q_n} \right|}{bq_n} \\
&\leq \frac{1}{2b^2} + \frac{1}{2bq_n}
\end{aligned}$$

e isto implica $b < q_n$ o que, sendo uma contradição, completa a prova do teorema. ■

É interessante notar que o mesmo limiar $\frac{1}{2b^2}$ já havia aparecido no Capítulo 5 (Teorema 5.1), onde se mostrou, por meio da geometria dos círculos de Ford, que todo irracional admite infinitas aproximações racionais dentro desse limiar. O Teorema de Lagrange acima complementa esse resultado por outro caminho, inteiramente algébrico: mostra que qualquer aproximação racional que já esteja dentro desse limiar é necessariamente um dos convergentes da fração contínua de α . As frações contínuas e os círculos de Ford descrevem, assim, por vias diferentes, o mesmo fenômeno de aproximação racional.

Exemplo 6.11 *Considere o número $\frac{\sqrt{5}+1}{2}$. Definimos o valor inicial como:*

$$x_0 = \phi = \frac{\sqrt{5}+1}{2} \approx 1,6180339 \dots$$

Passo 1 • **Parte inteira:** Como $1 < \frac{\sqrt{5}+1}{2} < 2$, temos:

$$a_0 = \lfloor x_0 \rfloor = 1$$

• **Resto:** O resto r_0 é a diferença entre o número e sua parte inteira:

$$r_0 = x_0 - a_0 = \frac{\sqrt{5}+1}{2} - 1 = \frac{\sqrt{5}-1}{2}$$

Passo 2 • **Novo valor:** Invertamos o resto obtido no passo anterior:

$$x_1 = \frac{1}{r_0} = \frac{2}{\sqrt{5}-1}$$

Racionalizando o denominador multiplicando pelo conjugado $(\sqrt{5}+1)$:

$$x_1 = \frac{2(\sqrt{5}+1)}{(\sqrt{5}-1)(\sqrt{5}+1)} = \frac{2(\sqrt{5}+1)}{5-1} = \frac{2(\sqrt{5}+1)}{4} = \frac{\sqrt{5}+1}{2}$$

• **Parte inteira e Resto:** Note que $x_1 = x_0$. Portanto, o processo se repete identicamente:

$$a_1 = \lfloor x_1 \rfloor = 1$$

$$r_1 = x_1 - a_1 = \frac{\sqrt{5}-1}{2}$$

Como $x_n = x_0$ para todo $n \geq 0$, todas as partes inteiras subsequentes serão iguais a 1. Assim, a sequência de coeficientes é $[1; 1, 1, 1, \dots]$, o que nos dá a expansão:

$$\phi = 1 + \frac{1}{1 + \frac{1}{1 + \frac{1}{1 + \ddots}}}$$

Vamos assumir que essa estrutura infinita converge para um valor real positivo x :

$$x = 1 + \frac{1}{1 + \frac{1}{1 + \frac{1}{1 + \ddots}}}$$

Como a estrutura de frações se estende infinitamente para baixo, o denominador da primeira fração possui exatamente a mesma forma e comportamento que toda a expressão original x . Substituindo a expressão em **negrito** pelo próprio x , reduzimos a estrutura infinita a uma equação algébrica simples:

$$x = 1 + \frac{1}{x}$$

Multipliquemos todos os termos por x :

$$x^2 = x + 1 \Leftrightarrow x^2 - x - 1 = 0$$

Resolvendo

$$x = \frac{-(-1) \pm \sqrt{(-1)^2 - 4(1)(-1)}}{2(1)} = \frac{1 \pm \sqrt{5}}{2}$$

Dado que a fração contínua original é uma soma de termos estritamente positivos iniciando em 1, o valor final de x deve ser necessariamente positivo ($x > 0$). Descartamos a raiz negativa $\frac{1-\sqrt{5}}{2}$.

6.5 Problemas Resolvidos

Problema 6.12 Mostrar que para $\alpha \in \mathbb{R}$ a seguinte relação se verifica

$$q_n |q_{n-1}\alpha - p_{n-1}| + q_{n-1} |q_n\alpha - p_n| = 1.$$

Solução 6.13 Como α está entre p_n/q_n e p_{n-1}/q_{n-1} , temos

$$\left| \alpha - \frac{p_{n-1}}{q_{n-1}} \right| + \left| \alpha - \frac{p_n}{q_n} \right| = \left| \frac{p_{n-1}}{q_{n-1}} - \frac{p_n}{q_n} \right| = \frac{1}{q_n q_{n-1}}.$$

Basta multiplicar, ambos os membros, por $q_n q_{n-1}$, que o resultado segue.

Problema 6.14 Mostrar que se $a_1 > 0$ e $\frac{p_n}{q_n} = [a_1, a_2, \dots, a_n]$ então,

$$\frac{p_n}{p_{n-1}} = [a_n, a_{n-1}, \dots, a_2, a_1].$$

Solução 6.15 Para $n = 1$ temos $[a_1] = a_1 = \frac{p_1}{p_0}$. Portanto o resultado é verdadeiro para $n = 1$.

Assumimos que,

$$\frac{p_k}{p_{k-1}} = [a_k, a_{k-1}, \dots, a_2, a_1]$$

onde $1 \leq k < n$. Então,

$$\begin{aligned} [a_{k+1}, a_k, a_{k-1}, \dots, a_2, a_1] &= a_{k+1} + \frac{1}{[a_k, a_{k-1}, \dots, a_2, a_1]} \\ &= a_{k+1} + \frac{1}{p_k/p_{k-1}} \\ &= \frac{a_{k+1}p_k + p_{k-1}}{p_k} = \frac{p_{k+1}}{p_k} \end{aligned}$$

pelo Teorema 6.2. Portanto o resultado é verdadeiro por indução.

Problema 6.16 Uma fração contínua simples é chamada **simétrica** no caso em que $a_i = a_{n-i}$ para $0 \leq i \leq n$. Por exemplo $[3, 2, 1, 2, 3]$ é simétrica. Mostrar que se o racional r/s , com $\text{mdc}(r, s) = 1$, possui representação simétrica então,

$$r | (s^2 + (-1)^n).$$

Solução 6.17 Sendo $\text{mdc}(r, s) = 1$ temos que $\frac{r}{s} = \frac{p_n}{q_n} \Rightarrow r = p_n$ e $s = q_n$ uma vez que $\text{mdc}(p_n, q_n) = 1$.

Pelo problema anterior e pelo fato de $\frac{r}{s}$ possuir representação simétrica temos que,

$$\frac{p_n}{p_{n-1}} = [a_n, a_{n-1}, \dots, a_2, a_1] = [a_1, a_2, \dots, a_n] = \frac{r}{s}$$

o que nos diz que $q_n = p_{n-1} = s$.

Como pelo Teorema 6.3 $p_n q_{n-1} - p_{n-1} q_n = (-1)^n$ concluímos que,

$$r q_{n-1} - s^2 = (-1)^n.$$

Logo, $r q_{n-1} = s^2 + (-1)^n$ e, portanto, $r | (s^2 + (-1)^n)$.

6.6 Problemas Propostos

1. Encontrar o número racional representado sob a forma de fração contínua em cada item abaixo:

- (a) $[3, 1]$
- (b) $[1, 1, 1]$
- (c) $[0, 6, 5]$
- (d) $[1, 2, 3, 4]$

2. A representação sob a forma de fração contínua simples infinita (não-periódica) do número e é dada por:

$$e = [2, 1, 2, 1, 1, 4, 1, 1, 6, 1, 1, 8, \dots]$$

Encontrar os primeiros 6 convergentes desta fração contínua.

3. Encontrar o irracional representado pela fração contínua $[1, \overline{1, 1, 1, \dots}]$.
4. Expressar os seguintes racionais sob a forma de fração contínua
 - (a) $11/7$
 - (b) $-51/23$
 - (c) $114/235$
 - (d) $34/21$
5. Mostrar que se $p_n/q_n = [a_1, a_2, \dots, a_n]$ então $q_n/q_{n-1} = [a_n, a_{n-1}, \dots, a_3, a_2]$.
6. Suponha que $r > s > 0$, $\text{mdc}(r, s) = 1$ e que $\frac{r}{s} = [a_1, a_2, \dots, a_n]$. Mostrar que se $p_n | (q_n^2 + (-1)^n)$ então $a_i = a_{n-i}$ para $1 \leq i < n$, isto é, $[a_1, a_2, \dots, a_n]$ é simétrica.

Referências Bibliográficas

- [1] Goodmann-Strauss, C., *Compass and Straightedge in the Poincaré disk*, American Mathematical Society, January, 2001.
- [2] Katok, S., *Fuchsian Groups*, Chicago: Library of Congress Cataloging, 1992.
- [3] APOSTOL, T. M. Ford Circles. In: **Modular Functions and Dirichlet Series in Number Theory**, 2nd ed. New York: Springer-Verlag, p. 99-102, 1997.
- [4] CONWAY, J. H.; GUY, R. K. Farey Fractions and Ford Circles. In: **The Book of Numbers**. New York: Springer-Verlag, p. 152-154, 1996.
- [5] CONWAY, J. H.; GUY, R. K. **The Book of Numbers**. New York: Springer-Verlag, 1996.
- [6] COXETER, H. S. M. Circles, horocycles, and equidistant curves. In: **Introduction to Geometry**, 2nd ed. New York: John Wiley & Sons, p. 287-290, 1961.
- [7] FORD, L. R. Fractions. **American Mathematical Monthly**, v. 45, p. 586-601, 1938.
- [8] HARDY, G. H.; WRIGHT, E. M. Farey Series and a Theorem of Minkowski. In: **An Introduction to the Theory of Numbers**, 5th ed. Oxford: Clarendon Press, p. 23-37, 1979.
- [9] RADEMACHER, H. **Higher Mathematics from an Elementary Point of View**. Boston: Birkhäuser, 1983.
- [10] SANTOS, J. P. O. **Introdução à Teoria dos Números**. Rio de Janeiro: IMPA, 1998. (Coleção Matemática Universitária).