# Capítulo 6

# Funções Geratrizes. Produtórias Complexas. Algumas Identidades Combinatórias

*Conteúdo*



O curto capítulo que ora iniciamos é dedicado a duas classes de objetos matemáticos úteis, as chamadas funções geratrizes e as produtórias complexas infinitas. Trata-se de dois temas não necessariamente próximos, mas ligados a diversas identidades e relações úteis. Faremos uso de ambos nos capítulos que se seguem. Neste capítulo, como em outros adiante, pressuporemos que o estudante possua noções básicas sobre a teoria das funções de variável complexa.

Funções geratrizes, introduzidas na Seção 6.1, página 313, desempenham um elegante papel em Análise Combinatória. O leitor poderá encontrar na bela referência [155] uma vasta coleção de identidades combinatórias interessantes que podem ser engenhosamente demonstradas com o uso de funções geratrizes de sequências, assim como outras referências à literatura pertinente. Funções geratrizes são também relevantes no estudo de certas sequências de funções e no estudo de propriedades de sequências numéricas. Ilustraremos essa última afirmação estudando através de funções geratrizes a chamada sequência de Fibonacci e os chamados números de Bernoulli. No Capítulo 16 faremos uso de funções geratrizes para estudar e demonstrar algumas propriedades úteis de algumas das soluções especiais que encontramos no Capítulo 15, como os polinômios de Legendre, de Hermite, de Laguerre, de Tchebychev e as funções de Bessel.

Produtórias infinitas, assim como séries, são encontradas amiúde no estudo de propriedades de certas funções de interesse. Faremos uso das mesmas, por exemplo, no Capítulo 7, dedicado à função gama de Euler. Na Seção 6.2, página 325, discutimos a noção de convergência de produtórias infinitas e estudamos alguns exemplos úteis.

## 6.1 Funções Geratrizes

• **Funções geratrizes**

Seja $\{a_n,\ n \in \mathbb{N}_0\}$ uma sequência de números reais ou complexos. Define-se a *função geratriz* da sequência $\{a_n,\ n \in \mathbb{N}_0\}$ como sendo a função dada por

$$G_{\{a_n\}}(t) \ := \ \sum_{n=0}^{\infty} a_n\, t^n \ .$$

Essa definição pressupõe que a série de potências em $t$ do lado direito seja convergente em alguma região aberta do plano complexo, digamos $|t| < T$, para algum $T > 0$. Isso nem sempre é o caso. Por exemplo, se $a_n = n!$ a série acima tem raio de convergência nulo.





• **Funções geratrizes exponenciais**

A *função geratriz exponencial* da sequência $\{a_n,\ n \in \mathbb{N}_0\}$ é definida por

$$E_{\{a_n\}}(t) \ := \ \sum_{n=0}^{\infty} \frac{a_n}{n!}\, t^n \ .$$

Novamente, essa definição pressupõe que a série de potências em $t$ do lado direito seja convergente em alguma região aberta do plano complexo, digamos $|t| < T$. Observe que o lado direito, caso seja convergente, coincide com a série de Taylor centrada em 0 da função $E_{\{a_n\}}(t)$ e, portanto, para cada $k$ vale $a_k = \frac{d^k E_{\{a_n\}}}{dt^k}(0)$.

• **Funções geratrizes de Dirichlet**

Para certos tipos de sequências é conveniente definir outro tipo de função geratriz, substituindo os monômios $t^n$ por outras funções de $t$: $\sum_{n=0}^{\infty} a_n\, S_n(t)$. O exemplo mais importante desse tipo de função geratriz é aquele no qual se toma $S_n(t) = 1/n^t$, $n \geq 1$. Isso nos conduz à próxima definição.

A *função geratriz de Dirichlet*[1] da sequência $\{a_n,\ n \in \mathbb{N}\}$ é definida por

$$D_{\{a_n\}}(t) \ := \ \sum_{n=1}^{\infty} \frac{a_n}{n^t} \ ,$$

desde que a série do lado direito convirja com a variável $t$ em alguma região aberta do plano complexo.

A mais famosa das funções geratrizes de Dirichlet é a *função zeta de Riemann*[2], que é a função geratriz de Dirichlet da sequência constante $a_n = 1$, $n \geq 1$:

$$\zeta(s) \ := \ \sum_{n=1}^{\infty} \frac{1}{n^s} \ . \tag{6.1}$$

Como facilmente se vê, a série do lado direito converge na região do plano complexo definida por $\mathrm{Re}(s) > 1$. A função zeta de Riemann desempenha um papel de grande importância na teoria das funções de variável complexa e na teoria de números, pois várias de suas propriedades estão relacionadas a propriedades do conjunto dos números primos. A ela dedicamos o Capítulo 8, página 386. Vide também [171], [419], [420], [111] e outras referências citadas no Capítulo 8.

• **Funções geratrizes de Lambert**

A *função geratriz de Lambert*[3] da sequência $\{a_n,\ n \in \mathbb{N}\}$ é definida por

$$L_{\{a_n\}}(t) \ := \ \sum_{n=1}^{\infty} a_n \frac{t^n}{1 - t^n} \ ,$$

desde que a série do lado direito convirja com a variável $t$ em alguma região aberta do plano complexo. As funções geratrizes de Lambert são também denominadas *séries de Lambert*.

• **Algumas propriedades de funções geratrizes**

As funções geratrizes definidas acima têm várias propriedades algébricas interessantes, como mostrado nos exercícios que seguem.

**E. 6.1** *Exercício*. Se $\{a_n\}$ e $\{b_n\}$ são duas sequências cujas funções geratrizes $G_{\{a_n\}}(t)$ e $G_{\{b_n\}}(t)$ têm uma região aberta de convergência comum, mostre que

$$G_{\{a_n\}}(t)\, G_{\{b_n\}}(t) \ = \ G_{\{c_n\}}(t) \ ,$$

onde

$$c_n \ = \ \sum_{p=0}^{n} a_{n-p}\, b_p \ .$$

♣

[1] Johann Peter Gustav Lejeune Dirichlet (1805–1859).
[2] Georg Friedrich Bernhard Riemann (1826–1866).
[3] Johan Heinrich Lambert (1728–1777).

**E. 6.2** *Exercício*. Se $\{a_n\}$ e $\{b_n\}$ são duas sequências cujas funções geratrizes exponenciais $E_{\{a_n\}}(t)$ e $E_{\{b_n\}}(t)$ têm uma região aberta de convergência comum, mostre que

$$E_{\{a_n\}}(t)\,E_{\{b_n\}}(t) \;=\; E_{\{c_n\}}(t)\;,$$

onde

$$c_n \;=\; \sum_{p=0}^{n}\binom{n}{p}\,a_{n-p}\,b_p\;.$$

✠

**E. 6.3** *Exercício*. Se $\{a_n\}$ e $\{b_n\}$ são duas sequências cujas funções geratrizes de Dirichlet $D_{\{a_n\}}(t)$ e $D_{\{b_n\}}(t)$ têm uma região aberta de convergência comum, mostre que

$$D_{\{a_n\}}(t)\,D_{\{b_n\}}(t) \;=\; D_{\{c_n\}}(t)\;,$$

onde

$$c_n \;=\; \sum_{\substack{p=1\\ n/p\ inteiro}}^{n} a_{n/p}\,b_p\;.$$

✠

**E. 6.4** *Exercício*. Se $\{a_n\}$ é uma sequência cuja função geratriz de Lambert é $L_{\{a_n\}}(t)$, mostre que

$$L_{\{a_n\}}(t) \;=\; \sum_{m=1} b_m t^m \;=\; G_{\{b_n\}}(t)\;,$$

onde $b_0 := 0$ e, para $m > 0$,

$$b_m \;:=\; \sum_{\substack{n=1\\ m/n\ inteiro}}^{m} a_n\;.$$

✠

Passemos a discutir algumas aplicações das funções geratrizes.

### 6.1.1 Algumas Identidades Combinatórias

A sequência de exercícios dirigidos que segue apresenta-nos uma série de identidades combinatórias de interesse (usaremos algumas no Capítulo 18, página 900). A primeira obtém-se através de uma função geratriz.

**E. 6.5** *Exercício dirigido*. Para $n \in \mathbb{N}$, $m \in \mathbb{N}_0$, defina-se o conjunto

$$\mathbf{N}^n_m \;:=\; \Big\{(a_1,\,\ldots,\,a_n)\in\mathbb{N}_0^n,\ a_1+\cdots+a_n = m\Big\}\;.$$

Seja $|\mathbf{N}^n_m|$ o número de elementos de $\mathbf{N}^n_m$. $|\mathbf{N}^n_m|$ representa o número de maneiras de colocar exatamente $m$ objetos indistinguíveis em $n$ posições distintas, eventualmente permitindo sobreposições (e.g., colocar $m$ bolas indistinguíveis em $n$ caixas distintas). Mostre que

$$|\mathbf{N}^n_m| \;=\; \binom{n+m-1}{m} \;=\; \frac{(n+m-1)!}{(n-1)!\,m!}\;.$$

*Sugestão*. Mostre primeiramente que, para cada $n$, a função geratriz da sequência $|\mathbf{N}^n_m|$, $m = 0,\,1,\,2,\,\ldots$, é

$$\sum_{m=0}^{\infty}|\mathbf{N}^n_m|\,t^m \;=\; \left(\frac{1}{1-t}\right)^n\;. \tag{6.2}$$

Para isso, mostre que, para $|t| < 1$,

$$\left(\frac{1}{1-t}\right)^n \;=\; \left(\sum_{a=0}^{\infty}t^a\right)^n \;=\; \sum_{a_1,\,\ldots,\,a_n=0}^{\infty} t^{a_1+\cdots+a_n} \;=\; \sum_{m=0}^{\infty}|\mathbf{N}^n_m|\,t^m.$$

De (6.2), obtenha

$$|\mathbf{N}^n_m| \;=\; \frac{1}{m!}\frac{d^m}{dt^m}\left(\frac{1}{1-t}\right)^n\bigg|_{t=0} \;=\; \frac{n\cdots(m+m-1)}{m!} \;=\; \frac{(n+m-1)!}{(n-1)!\,m!} \;=\; \binom{n+m-1}{m}\;.$$

✠



**E. 6.6** *Exercício dirigido*. Para $n \in \mathbb{N}$, $m \in \mathbb{N}_0$, defina-se o conjunto

$$\mathbf{M}^n_m \;:=\; \Big\{(a_1,\,\ldots,\,a_n)\in\mathbb{N}_0^n,\ 0\leq a_1+\cdots+a_n \leq m\Big\}\;.$$

Seja $|\mathbf{M}^n_m|$ o número de elementos de $\mathbf{M}^n_m$. $|\mathbf{M}^n_m|$ representa o número de maneiras de colocar de zero a no máximo $m$ objetos indistinguíveis em $n$ posições distintas, eventualmente permitindo sobreposições (e.g., colocar de $0$ a $m$ bolas indistinguíveis em $n$ caixas distintas). Mostre que

$$|\mathbf{M}^n_m| \;=\; \binom{n+m}{m} \;=\; \frac{(n+m)!}{n!\,m!}\;. \tag{6.3}$$

*Sugestão*. Convença-se que

$$|\mathbf{M}^n_m| \;=\; \sum_{k=0}^{m}|\mathbf{N}^n_k| \;=\; \sum_{k=0}^{m}\binom{n+k-1}{k}\;. \tag{6.4}$$

Usando a bem conhecida *identidade de Pascal*[4]

$$\binom{a}{b} \;=\; \binom{a-1}{b}+\binom{a-1}{b-1} \tag{6.5}$$

conclua que

$$\begin{aligned}
\sum_{k=0}^{m}\binom{n+k}{k} \;=\; 1+\sum_{k=1}^{m}\binom{n+k}{k} &\stackrel{(6.5)}{=} 1+\sum_{k=1}^{m}\binom{n+k-1}{k}+\sum_{k=1}^{m}\binom{n+k-1}{k-1}\\
&= 1+\sum_{k=1}^{m}\binom{n+k-1}{k}+\sum_{l=1}^{m}\binom{n+l-1}{l-1}\\
&\stackrel{l'=l-1}{=} 1+\sum_{k=1}^{m}\binom{n+k-1}{k}+\sum_{l'=0}^{m-1}\binom{n+l'}{l'}
\end{aligned}$$

e, assim, conclua que

$$\sum_{k=0}^{m}\binom{n+k-1}{k} \;=\; 1+\sum_{k=1}^{m}\binom{n+k-1}{k} \;=\; \sum_{k=0}^{m}\binom{n+k}{k}-\sum_{l'=0}^{m-1}\binom{n+l'}{l'} \;=\; \binom{n+m}{m}\;.$$

Por (6.4), isso prova (6.3). ✠

A identidade $\displaystyle\sum_{k=0}^{m}\binom{n+k-1}{k} = \binom{n+m}{m}$, provada acima, é conhecida como *segunda identidade de Pascal* ou *identidade da soma paralela*. Para outras identidades combinatórias úteis, vide [155].

**E. 6.7** *Exercício*. Seguindo passos análogos aos do último exercício, demonstre a *identidade da soma vertical*:

$$\binom{n+1}{m+1} \;=\; \sum_{j=0}^{n}\binom{j}{m}\;.$$

✠

**E. 6.8** *Exercício*. As denominações *identidade da soma paralela* e *identidade da soma vertical* provêm da relação dos coeficientes binomiais $\binom{a}{b}$ com o triângulo de Pascal. Desenhe um triângulo de Pascal e entenda o significado dessas identidades e a razão de seus nomes. ✠

**E. 6.9** *Exercício dirigido*. As identidades $|\mathbf{N}^n_m| = \binom{n+m-1}{m}$ e $|\mathbf{M}^n_m| = \binom{n+m}{m}$ podem ser obtidas de uma forma talvez mais direta e simples, dependendo do gosto do leitor. Suponha que se tenha $m$ bolas pretas e $n$ bolas brancas. Convença-se que há $\binom{n+m}{m} = \binom{n+m}{n}$ arranjos possíveis dessas bolas (supondo que as bolas pretas são indistinguíveis entre si, e que o mesmo valha para as brancas). Uma maneira de fazer esse raciocínio é imaginar as $n+m$ bolas enfileiradas e contar de quantas maneiras distintas essas fileiras podem ser formadas. Há $(n+m)!$ permutações das $n+m$ bolas, das quais devem ser fatoradas $m!$ permutações envolvendo apenas bolas pretas

[4] Blaise Pascal (1623–1662).

e $n!$ permutações envolvendo apenas bolas brancas, fornecendo assim $\binom{n+m}{m}$ arranjos. Convença-se também que, pela definição, esse número de arranjos é igual a $|\mathbf{M}_m^n|$. Isso provou que $|\mathbf{M}_m^n| = \binom{n+m}{m}$. Convença-se que, pela definição, $|\mathbf{N}_m^n| = |\mathbf{M}_m^n| - |\mathbf{M}_{m-1}^n|$. Tem-se, então $|\mathbf{N}_m^n| = \binom{n+m}{m} - \binom{n+m-1}{m-1} = \binom{n+m-1}{m}$, onde a última igualdade segue da identidade de Pascal (6.5). ✠

### 6.1.2 Números de Fibonacci

Seja $a_n$, $n = 0, 1, 2, 3, 4 \ldots$, a sequência definida recursivamente da seguinte forma:

$$a_0 = 1, \quad a_1 = 1, \quad a_{n+2} = a_{n+1} + a_n, \quad \forall\, n \geq 0 .$$

Essa sequência é denominada *sequência de Fibonacci*[5]. Cada elemento da sequência de Fibonacci é a soma de seus dois antecessores. Os primeiros elementos da sequência de Fibonacci são

$$1,\ 1,\ 2,\ 3,\ 5,\ 8,\ 13,\ 21,\ 34,\ 55,\ 89,\ 144,\ 233,\ 377,\ 610,\ 987,\ 1597,\ 2584,\ 4181,\ 6765,\ 10946,\ 17711,\ 28657,\ \ldots .$$

Fibonacci introduziu a sequência que leva seu nome em um problema de seu livro *Liber abbaci*, de 1202 (livro esse que introduziu o sistema decimal arábico na Europa, em substituição ao sistema de algarismos romanos, usado até então): *"Um certo homem coloca um casal de coelhos em um local cercado de muros por todos os lados. Quantos pares de coelhos podem ser produzidos a partir daquele casal em um ano se for suposto que a cada mês cada casal gera um novo casal, o qual se torna fértil em um mês"*. A resposta (supondo que nenhum coelho morre) é que, após $n$ meses, tem-se $a_n$ pares de coelhos, sendo $a_n$ dado acima. Trata-se provavelmente do primeiro modelo de evolução de populações. A sequência de Fibonacci é surpreendentemente rica em propriedades, sendo possivelmente uma das mais pesquisadas, existindo até mesmo uma publicação periódica ("Fibonacci Quarterly") dedicada a seu estudo.

Um fato que confere aos números de Fibonacci um sabor especial é que os mesmos aparecem frequentemente na Natureza. Há, por exemplo, uma forte probabilidade de os números de pétalas em flores de determinadas espécies de plantas serem números de Fibonacci. O mesmo se dá com o número de voltas espirais na casca de abacaxis e de pinhas, com o número de ramos de plantas e árvores, com o número de padrões de um determinado tipo nas conchas de caramujos etc[6]. A razão do surgimento de números de Fibonacci em contextos biológicos está relacionado à formação e reprodução de padrões, mas é apenas parcialmente entendida atualmente.

No intuito de ilustrar a utilidade de funções geratrizes de sequências, vamos demonstrar a seguinte identidade para os elementos da sequência de Fibonacci:

$$a_n = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^{n+1} - \left(\frac{1-\sqrt{5}}{2}\right)^{n+1}\right], \tag{6.6}$$

válida para todo $n \geq 0$. Essa expressão permite obter cada $a_n$ diretamente em termos de $n$ e é denominada *fórmula de Binet*[7], tendo sido obtida por esse autor (usando métodos matriciais) em 1843. Essa fórmula, porém, foi obtida pela primeira vez por de Moivre[8] cerca de cem anos antes e foi encontrada também por Euler[9] e por Daniel Bernoulli[10]. A constante $\frac{1+\sqrt{5}}{2}$ ($\approx 1,6180339887\ldots$) é a célebre *razão áurea*, muitas vezes denotada pelo símbolo $\varphi$, e cujas propriedades são estudadas desde a Antiguidade. Note-se que $\frac{1-\sqrt{5}}{2} = 1 - \varphi = -1/\varphi$, de modo que (6.6) pode ser também expressa como

$$a_n = \frac{1}{\sqrt{5}}\left[\varphi^{n+1} - (1-\varphi)^{n+1}\right] = \frac{1}{\sqrt{5}}\left[\varphi^{n+1} - \left(-\frac{1}{\varphi}\right)^{n+1}\right]. \tag{6.7}$$

[5] Leonardo Pisano, cognominado "Fibonacci" (1170–1250).

[6] Para algumas referências:

- S. L. Basin, "The Fibonacci Sequence as it appears in Nature", The Fibonacci Quarterly, **1**, (1963), 53–57.
- A. Brousseau, "Fibonacci Statistics in Conifers", The Fibonacci Quarterly, **7** (1969), 525–532.
- P. B. Onderdonk, "Pineapples and Fibonacci Numbers", The Fibonacci Quarterly, **8** (1970), 507–508.

Um livro clássico sobre o assunto é [416]. A área da Biologia e da Matemática que se dedica ao estudo da formação e evolução de padrões é denominada Filotaxia.

[7] Jacques Philippe Marie Binet (1786–1856).

[8] Abraham de Moivre (1677–1754).

[9] Leonhard Euler (1707–1783).

[10] Daniel Bernoulli (1700–1782).



A função geratriz da sequência de Fibonacci é

$$F(t) = \sum_{n=0}^{\infty} a_n\, t^n . \tag{6.8}$$

Mostremos primeiramente que a série de potências do lado direito tem um raio de convergência não nulo. Pelo teste da razão vale, para $n > 0$,

$$\left|\frac{a_{n+1}t^{n+1}}{a_n t^n}\right| = \frac{a_{n+1}}{a_n}\,|t| = \frac{a_n + a_{n-1}}{a_n}\,|t| = \left(1 + \frac{a_{n-1}}{a_n}\right)|t| \leq 2|t| ,$$

pois $\frac{a_{n-1}}{a_n} \leq 1$, já que a sequência de Fibonacci é crescente. Logo, a série converge absolutamente pelo menos na região $|t| < 1/2$. A verdadeira região de convergência é um pouco maior (como veremos adiante), mas não precisaremos desse fato por ora, pois tudo o que necessitamos é da existência de um raio de convergência não nulo, o que justifica as manipulações que faremos.

Façamos uso da definição da sequência de Fibonacci para obter uma fórmula explícita para $F(t)$. Temos que

$$\begin{aligned}
F(t) &= 1 + t + \sum_{n=2}^{\infty} a_n\, t^n \\
&= 1 + t + \sum_{n=2}^{\infty}(a_{n-1} + a_{n-2})\, t^n = 1 + t + \sum_{n=2}^{\infty} a_{n-1}\, t^n + \sum_{n=2}^{\infty} a_{n-2}\, t^n \\
&= 1 + t + t\sum_{n=1}^{\infty} a_n\, t^n + t^2 \sum_{n=0}^{\infty} a_n\, t^n \\
&= 1 + t + t(F(t) - 1) + t^2 F(t) .
\end{aligned}$$

Assim, $(1 - t - t^2)F(t) = 1$ e, portanto,

$$F(t) = \frac{1}{1 - t - t^2} .$$

A ideia agora é obter a expansão em série de Taylor de $F(t)$ em torno de $t = 0$ e compará-la a (6.8), para assim obter uma expressão explícita para os $a_n$'s. Para isso, ao invés de calcularmos as derivadas de $F$ em $t = 0$, é mais fácil proceder da seguinte forma. Escrevemos $1 - t - t^2 = -(t - \gamma_1)(t - \gamma_2)$ onde

$$\gamma_1 = \frac{\sqrt{5} - 1}{2}, \qquad \gamma_2 = -\frac{\sqrt{5} + 1}{2} .$$

Assim,

$$\begin{aligned}
F(t) &= \frac{1}{1 - t - t^2} = -\frac{1}{(t - \gamma_1)(t - \gamma_2)} = \frac{1}{\gamma_1 - \gamma_2}\left[\frac{1}{\gamma_1 - t} - \frac{1}{\gamma_2 - t}\right] \\
&= \frac{1}{\sqrt{5}}\left[\frac{1}{\gamma_1}\left(\frac{1}{1 - \frac{t}{\gamma_1}}\right) - \frac{1}{\gamma_2}\left(\frac{1}{1 - \frac{t}{\gamma_2}}\right)\right] \\
&= \frac{1}{\sqrt{5}} \sum_{n=0}^{\infty}\left[\frac{1}{\gamma_1^{n+1}} - \frac{1}{\gamma_2^{n+1}}\right] t^n \\
&= \frac{1}{\sqrt{5}} \sum_{n=0}^{\infty}\left[(-\gamma_2)^{n+1} - (-\gamma_1)^{n+1}\right] t^n \\
&= \sum_{n=0}^{\infty} \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^{n+1} - \left(\frac{1-\sqrt{5}}{2}\right)^{n+1}\right] t^n ,
\end{aligned}$$

onde usamos que $1/\gamma_1 = -\gamma_2$. Comparando com (6.8) obtemos (6.6), como queríamos. Da última expressão, vê-se também que o raio de convergência da série de potências que define $F$ é $(\sqrt{5}-1)/2 \approx 0,6180339887\ldots$.

No Exercício E. 10.58, página 583, reobtemos (6.6) usando técnicas matriciais, as quais permitem obter os valores de $a_n$ mesmo para condições iniciais outras que não $a_0 = a_1 = 1$. Essas mesmas ideias são usadas no Exercício E. 10.59, página 584, para tratar da chamada *sequência de Fibonacci generalizada*, definida pela relação de recorrência $a_{n+2} = \alpha a_{n+1} + \beta a_n$, onde $\alpha$ e $\beta$ são constantes (reais ou complexas).

### 6.1.3 Números de Bernoulli

A sequência de números racionais denominados *números de Bernoulli*[11] tem importância destacada na Teoria dos Números, especialmente devido à sua relação com a função zeta de Riemann, definida acima. Os números de Bernoulli também aparecem na expansão em série de Taylor da função tangente e na chamada fórmula de Euler-Maclaurin. Os chamados números de Bernoulli, denotados por $B_n$, com $n \in \mathbb{N}_0$, são definidos de sorte que sua função geratriz exponencial $E_{\{B_n\}}(z)$ seja a função $z/(e^z-1)$, ou seja, são definidos por

$$\frac{z}{e^z-1} \;=:\; \sum_{n=0}^{\infty} \frac{B_n}{n!}\, z^n \,. \tag{6.9}$$

Devido ao fato de $z/(e^z-1)$ ter um polo em $z = \pm 2\pi i$ sendo, porém, analítica em $|z| < 2\pi$, concluímos *a priori* que a série de potências do lado direito é convergente para $|z| < 2\pi$. Multiplicando (6.9) por $e^z - 1 = \sum_{m=1}^{\infty} z^m/m!$, obtemos

$$z \;=\; \left(\sum_{n=0}^{\infty} \frac{B_n}{n!}\, z^n\right)\left(\sum_{m=1}^{\infty} \frac{z^m}{m!}\right) \;=\; \sum_{p=0}^{\infty}\left(\sum_{q=0}^{p} \frac{B_q}{q!(p+1-q)!}\right) z^{p+1} \,,$$

de onde concluímos que

$$B_0 \;=\; 1 \qquad \text{e} \qquad \sum_{q=0}^{p} \frac{B_q}{q!(p+1-q)!} \;=\; 0 \quad \text{para todo } p \geq 1 \,.$$

Multiplicando a segunda relação acima por $(p+1)!$ a mesma torna-se

$$\sum_{q=0}^{p} B_q \binom{p+1}{q} \;=\; 0 \quad \text{para todo } p \geq 1 \,,$$

forma essa mais frequentemente encontrada na literatura. Essa relação acima permite obter recursivamente os coeficientes $B_n$ a partir de $B_0 = 1$. De fato, isolando o termo com $q = p$, temos

$$B_p \;=\; \frac{-1}{p+1} \sum_{q=0}^{p-1} B_q \binom{p+1}{q} \quad \text{para todo } p \geq 1 \,, \tag{6.10}$$

ou seja,

$$B_p \;=\; -p! \sum_{q=0}^{p-1} \frac{B_q}{q!(p+1-q)!} \quad \text{para todo } p \geq 1 \,. \tag{6.11}$$

Usando a fórmula (6.10) é possível obter os primeiros números de Bernoulli, vide Tabela 6.1, página 320.

A contemplação da Tabela 6.1 permite conjecturar que, exceto $B_1$, todos os $B_n$ com $n$ ímpar sejam nulos. Veremos abaixo que essa conjectura é verdadeira. A impressão, porém, que os $B_n$'s não nulos crescem lentamente, obtida da observação dos primeiros elementos da sequência, é falsa. Devido ao fato de a série de potências em não convergir para $|z| = 2\pi$ concluímos que os $|B_n|$ não nulos devem assintoticamente ser maiores que, ou da ordem de, $n!/(2\pi)^n$ para $n$ grande. Isso de fato é correto ($B_{98}$, por exemplo, é da ordem de $1,13 \times 10^{76}$) e a expressão precisa será apresentada em (6.18). Outra conjectura que se pode levantar da observação da Tabela 6.1 é que os sinais dos números de Bernoulli com índice par (exceto $B_0$) são alternados. Esse fato é também correto e será provado mais adiante.

[11] Jakob Bernoulli (1654–1705).



| $n$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $B_n$ | 1 | $-\frac{1}{2}$ | $\frac{1}{6}$ | 0 | $-\frac{1}{30}$ | 0 | $\frac{1}{42}$ | 0 | $-\frac{1}{30}$ | 0 | $\frac{5}{66}$ | 0 | $-\frac{691}{2730}$ | 0 | $\frac{7}{6}$ | 0 | $-\frac{3617}{510}$ | 0 |

Tabela 6.1: Números de Bernoulli $B_n$ para $n = 0, \ldots, 17$.

Separando o termo com $B_1$ de (6.9), que é $-z/2$, e passando-o para o lado esquerdo, obtemos

$$\frac{z}{e^z-1} + \frac{z}{2} \;=\; 1 + \sum_{n=2}^{\infty} \frac{B_n}{n!}\, z^n \,.$$

O lado esquerdo vale $\frac{z}{2}\coth\left(\frac{z}{2}\right)$, como facilmente se constata. Concluímos assim que

$$z\coth(z) \;=\; 1 + \sum_{n=2}^{\infty} \frac{2^n B_n}{n!}\, z^n \,, \tag{6.12}$$

para $|z| < 2\pi$. Como $z\coth(z)$ é uma função par, vemos de (6.12) que, exceto $B_1$, todos os demais $B_n$'s com $n$ ímpar são nulos. Com esse conhecimento podemos escrever

$$z\coth(z) \;=\; \sum_{n=0}^{\infty} \frac{4^n B_{2n}}{(2n)!}\, z^{2n} \,. \tag{6.13}$$

Como $z\,\mathrm{cotg}\,(z) = iz\coth(iz)$, obtemos também

$$z\,\mathrm{cotg}\,(z) \;=\; \sum_{n=0}^{\infty} \frac{(-1)^n 4^n B_{2n}}{(2n)!}\, z^{2n} \,, \tag{6.14}$$

para $|z| < 2\pi$.

Há uma conclusão importante a se obter de (6.14). A função $z\,\mathrm{cotg}\,(z)$ satisfaz também a igualdade, obtida primeiramente por Euler em 1748,

$$z\,\mathrm{cotg}\,(z) \;=\; 1 - 2z^2 \sum_{k=1}^{\infty} \frac{1}{(k\pi)^2 - z^2} \,. \tag{6.15}$$

Demonstrações dessa importante relação podem ser encontradas nestas notas no Exercício E. 7.5, página 359, no Exercício-dirigido E. 6.17, página 342, ou no Exercício-dirigido E. 36.23, página 1856 (para outras referências, vide esses Exercícios-dirigidos). Agora, para $|z| < \pi$, podemos escrever, já que $k \geq 1$,

$$\frac{1}{(k\pi)^2 - z^2} \;=\; \frac{1}{(k\pi)^2}\left(\frac{1}{1-\left(\frac{z}{k\pi}\right)^2}\right) \;=\; \frac{1}{(k\pi)^2}\sum_{n=0}^{\infty}\left(\frac{z}{k\pi}\right)^{2n} \,,$$

e reinserindo isso em (6.15), obtemos

$$z\,\mathrm{cotg}\,(z) \;=\; 1 - \sum_{n=1}^{\infty}\left(2\sum_{k=1}^{\infty}\frac{1}{(k\pi)^{2n}}\right) z^{2n} \,.$$

Comparando a (6.14), obtemos finalmente

$$\sum_{k=1}^{\infty} \frac{1}{k^{2n}} \;=\; (-1)^{n+1}\frac{4^n\, B_{2n}}{2(2n)!}\,\pi^{2n} \;=\; (-1)^{n+1}\frac{(2\pi)^{2n}\, B_{2n}}{2(2n)!} \,, \tag{6.16}$$

válida para todo inteiro $n > 0$. Note que o lado esquerdo é igual a $\zeta(2n)$, onde $\zeta$ é a função zeta de Riemann, definida em (6.1), página 314. A célebre expressão (6.16) foi obtida pela primeira vez em 1735, por Euler, resolvendo assim parcialmente um problema, denominado *problema de Basel*, levantado por Mengoli[12] em 1644, que consistia em encontrar uma fórmula fechada para as somas $\sum_{k=1}^{\infty} \frac{1}{k^m}$, $m \in \mathbb{N}$, $m > 1$, as quais envolvem potências de inversas de números inteiros. Os primeiros resultados obtidos de (6.16) são

$$\sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6}, \quad \sum_{k=1}^{\infty} \frac{1}{k^4} = \frac{\pi^4}{90}, \quad \sum_{k=1}^{\infty} \frac{1}{k^6} = \frac{\pi^6}{945}, \quad \sum_{k=1}^{\infty} \frac{1}{k^8} = \frac{\pi^8}{9450}, \quad \sum_{k=1}^{\infty} \frac{1}{k^{10}} = \frac{\pi^{10}}{93555}. \tag{6.17}$$

Como o lado esquerdo de (6.16) é sempre positivo e não nulo concluímos daquela identidade que os sinais da sequência $B_{2n}$, $n \geq 1$, são alternados e que os $B_{2n}$'s nunca se anulam. Como o lado esquerdo de (6.16) converge a 1 quando $n \to \infty$ (por quê?), obtemos a expressão assintótica

$$B_{2n} \overset{n\to\infty}{\approx} (-1)^{n+1} \frac{2(2n)!}{(2\pi)^{2n}}. \tag{6.18}$$

Diversos textos tratam de outras propriedades elementares dos números de Bernoulli. Recomendamos, em particular, [155]. Vide também [390]. Para uma prova de (6.16) em alguns casos particulares usando séries de Fourier, vide os exercícios da Seção 36.6, página 1853. Para uma prova geral de (6.16) usando séries de Fourier, vide [128]. Para uma discussão aparentada, vide Seção 24.B, página 1305, destas Notas.

O estudante deve interessar-se em saber que é até hoje um problema aberto determinar fórmulas exatas para as séries $\sum_{k=1}^{\infty} \frac{1}{k^m} = \zeta(m)$ quando $m$ é um número ímpar maior que 1. Além de não haver tais fórmulas exatas, sabe-se muito pouco sobre $\zeta(m)$ com $m$ ímpar. Apenas em 1979 foi demonstrado, por R. Apéry[13], que $\zeta(3)$ é um número irracional[14]. Em 2001, Wadim Zudilin demonstrou que ao menos um dentre os números $\zeta(5)$, $\zeta(7)$, $\zeta(9)$, $\zeta(11)$ é irracional[15]. Em 2000, Tanguy Rivoal demonstrou que há infinitos $\zeta(m)$, com $m$ ímpar, que são irracionais[16] sem, no entanto, poder especificar quais são.

### 6.1.4 Números de Bell

Os chamados *números de Bell* são relevantes na contagem do número de partições de um conjunto finito[17] e surgem em identidades combinatórias. Há diversas formas de os apresentarmos e escolhemos aqui fazê-lo no contexto de uma dessas identidades. Afirmamos que para todo $n \in \mathbb{N}_0$ vale a seguinte relação:

$$\sum_{k=0}^{\infty} \frac{k^n}{k!} = e\, T_n, \tag{6.19}$$

onde $T_n$ são números naturais que satisfazem as relações de recorrência

$$T_0 = 1, \qquad T_1 = 1 \qquad \text{e} \qquad T_{n+1} = \sum_{p=0}^{n} \binom{n}{p} T_p, \quad n \geq 1. \tag{6.20}$$

Assim,

$$T_0 = 1, \quad T_1 = 1, \quad T_2 = 2, \quad T_3 = 5, \quad T_4 = 15, \quad T_5 = 52, \quad T_6 = 203, \quad T_7 = 877, \quad T_8 = 4140 \text{ etc.}$$

Como veremos mais adiante, o número $T_n$ conta o número de partições de um conjunto com $n$ elementos e também possui uma interpretação em termos de probabilidades, a qual apresentamos brevemente adiante ($T_n$ é o $n$-ésimo momento da distribuição de Poisson com média 1).

[12] Pietro Mengoli (1626–1686).
[13] Roger Apéry (1916–1994).
[14] Para o trabalho original: Roger Apéry, "Irrationalité de $\zeta(2)$ et $\zeta(3)$, Astérisque, **61** (1979), 11–13. Vide também Alfred van der Poorten, "A proof that Euler missed. Apéry's proof of the irrationality of $\zeta(3)$. An informal report", Math. Intell., **1** (1979), 195–203.
[15] Wadim Zudilin. "One of the numbers $\zeta(5)$, $\zeta(7)$, $\zeta(9)$, $\zeta(11)$ is irrational". Russ. Math. Surv. 56 (4): 774–776 (2001).
[16] Tanguy Rivoal, "La fonction Zêta de Riemann prend une infinit'e de valeurs irrationnelles aux entiers impairs". Comptes rendus de l'Académie des Sciences, Paris, **331** (2000), 267–270.
[17] A noção de partição de um conjunto é introduzida na página 44.



Os números $T_n$ são denominados *números de Bell*[18] e a relação (6.19) é denominada *fórmula de Dobiński*[19], e foi obtida em 1877. A denominação em honra a Bell deve-se a um trabalho seu sobre esses números, datado de 1934[20], mas, em verdade, o primeiro estudo sistemático dos números de Bell fora feito por Ramanujan[21], mais de 25 anos antes, no segundo de seus célebres *Notebooks*. Por isso, a designação dos números $T_n$ em honra a Bell é controversa.

- **Prova da fórmula de Dobiński**

Para $n \in \mathbb{N}_0$, defina $F_n := \sum_{k=0}^{\infty} \frac{k^n}{k!}$. Para $n = 0$ temos

$$F_0 = \sum_{k=0}^{\infty} \frac{1}{k!} = e. \tag{6.21}$$

Para $n > 0$, podemos começar a somatória em $k = 1$ e temos $F_n = \sum_{k=1}^{\infty} \frac{k^n}{k!}$. É claro disso que

$$F_1 = \sum_{k=1}^{\infty} \frac{1}{(k-1)!} = \sum_{k=0}^{\infty} \frac{1}{k!} = e. \tag{6.22}$$

Constate que, para $n \in \mathbb{N}$,

$$F_{n+1} = \sum_{k=1}^{\infty} \frac{k^{n+1}}{k!} = \sum_{k=1}^{\infty} \frac{k^n}{(k-1)!} = \sum_{k=0}^{\infty} \frac{(k+1)^n}{k!} = 1 + \sum_{k=1}^{\infty} \frac{(k+1)^n}{k!}.$$

Escreva agora $(k+1)^n = \sum_{p=0}^{n} \binom{n}{p} k^p$, insira isso na igualdade acima e obtenha, após trocar a ordem das somas (justifique isso!),

$$F_{n+1} = 1 + \sum_{k=1}^{\infty} \frac{1}{k!} \left( \sum_{p=0}^{n} \binom{n}{p} k^p \right) = 1 + \sum_{p=0}^{n} \binom{n}{p} \left( \sum_{k=1}^{\infty} \frac{k^p}{k!} \right).$$

O termo com $p = 0$ é $\sum_{k=1}^{\infty} \frac{1}{k!} = e - 1$. Assim, obtenha

$$F_{n+1} = e + \sum_{p=1}^{n} \binom{n}{p} \left( \sum_{k=1}^{\infty} \frac{k^p}{k!} \right) = e + \sum_{p=1}^{n} \binom{n}{p} F_p. \tag{6.23}$$

Definido $T_n := e^{-1} F_n$, $n \in \mathbb{N}_0$, ou seja,

$$T_n := \frac{1}{e} \sum_{k=0}^{\infty} \frac{k^n}{k!},$$

teremos, pelo que vimos acima em (6.21), (6.22) e (6.23),

$$T_0 = 1, \qquad T_1 = 1 \qquad \text{e} \qquad T_{n+1} = \sum_{p=0}^{n} \binom{n}{p} T_p, \quad n \geq 1,$$

provando o que desejávamos estabelecer, a fórmula de Dobiński.

- **Distribuição de Poisson de média 1**

A relação (6.19) tem uma interpretação em termos de probabilidades. Consideremos $\mathbb{N}_0 = \{0, 1, 2, \ldots\}$ como um espaço amostral com a distribuição de probabilidades $P\big(\{k\}\big) = \frac{1}{ek!}$ para todo $k \in \mathbb{N}_0$. Naturalmente, $P\big(\{k\}\big) > 0$ e $\sum_{k=0}^{\infty} P\big(\{k\}\big) = 1$. A distribuição $P$ é a distribuição de Poisson[22] com média igual a 1[23].

[18] Eric Temple Bell (1883–1960).
[19] Donald Gabriel Dobiński (–).
[20] E. T. Bell, "Exponential polynomials". Annals of Mathematics. **35:** 258–277 (1934). doi:10.2307/1968431. JSTOR 1968431.
[21] Srinivasa Aiyangar Ramanujan (1887–1920).
[22] Siméon Denis Poisson (1781–1840).
[23] De modo geral, a distribuição de Poisson de média $\lambda > 0$ é dada por $P\big(\{k\}\big) = e^{-\lambda} \frac{\lambda^k}{k!}$.

Para $n \in \mathbb{N}_0$, a expressão

$$T_n \;:=\; \sum_{k=0}^{\infty} k^n P\big(\{k\}\big) \;=\; \frac{1}{e}\sum_{k=0}^{\infty}\frac{k^n}{k!}$$

representa o $n$-ésimo *momento* da distribuição de probabilidades $P$. As relações (6.20) fornecem expressões recursivas para esses momentos. Assim, por exemplo, o valor médio da distribuição é $T_1 = 1$ e sua variância é $T_2 - (T_1)^2 = 1$.

**• Números de Bell e partições**

A noção de partição de um conjunto foi introduzida à página 44. Seja $X$ um conjunto e seja $\mathcal{C} = \{C_\lambda,\ \lambda \in \Lambda\}$ uma coleção de subconjuntos de $X$ (que indexamos por um conjunto de índices $\Lambda$). Dizemos que $\mathcal{C}$ é uma *partição* de $X$ se $C_\lambda \cap C_{\lambda'} = \emptyset$ sempre que $\lambda \neq \lambda'$ e se $\bigcup_{\lambda\in\Lambda} C_\lambda = X$. É evidente que uma coleção $\mathcal{C}$ de subconjuntos de $X$ é uma partição de $X$ se e somente se cada $x \in X$ pertence a um e somente um conjunto $C_\lambda$.

Coloquemos a questão de saber quantas partições podem existir em um conjunto finito.

Seja $P_n$ o número de partições de um conjunto de $n$ elementos distintos, sendo que honorificamente declaramos que $P_0 = 1$ (ou seja, consideramos que o conjunto vazio tem apenas uma partição). Afirmamos que $P_n = T_n$ para todo $n \in \mathbb{N}_0$, sendo $T_n$ os números de Bell, definidos acima. A argumentação é a seguinte.

Seja $C = \{c_1,\ \ldots,\ c_{n+1}\}$ um conjunto com $n+1$ elementos distintos. Seja $\mathcal{P}$ uma partição de $C$ e, dentro dessa partição, considere-se o subconjunto $D$ de $C$ que contém o elemento $c_1$. O conjunto $C \setminus D$ pode ter $k$ elementos, sendo $k$ um número de 0 a $n$, correspondentes aos casos extremos em que $D = \{c_1,\ \ldots,\ c_{n+1}\}$ e $D = \{c_1\}$. Se eliminarmos $D$ da partição $\mathcal{P}$ o que sobra é uma partição de $k$ elementos e, por definição, há $P_k$ de tais partições. Esses $k$ elementos comporão um subconjunto de $\{c_2,\ \ldots,\ c_{n+1}\}$, que é um conjunto com $n$ elementos. Há $\binom{n}{k}$ possíveis subconjuntos de $k$ elementos em $\{c_2,\ \ldots,\ c_{n+1}\}$. Assim, o número de partições associadas a um dado $k$ é $\binom{n}{k} P_k$. Portanto, como $k$ assume todos os valores de 0 a $n$, concluímos que $P_{n+1} = \sum_{k=0}^{n}\binom{n}{k} P_k$. Trata-se das mesmas relações de recorrência dos números de Bell. Como $P_0 = 1$ e $P_1 = 1$, os mesmos valores assumidos por $T_0$ e $T_1$, concluímos que $P_n = T_n$ para todo $n \in \mathbb{N}_0$.

Essa constatação do significado de $T_n$ em termos de partições torna ainda mais notável fórmula de Dobiński (6.19).

**• Uma majoração grosseira para os números $T_n$**

Com uso das relações de recorrência (6.20) podemos encontrar uma estimativa grosseira para um majorante dos números de Bell $T_n$ que é suficiente para os propósitos que teremos logo adiante[24]. Essa estimativa é

$$T_n \;<\; 2n!\, e^n\,, \qquad n \in \mathbb{N}_0\,. \tag{6.24}$$

São conhecidas estimativas muito melhores na literatura – duas serão apresentada mais adiante – mas não precisaremos delas por enquanto.

Vamos demonstrá-la por indução. Ela é claramente válida para $T_0$ e $T_1$. Vamos admitir que o seja para todo $p = 1,\ \ldots,\ n$. Por (6.20), teríamos

$$T_{n+1} \;<\; 2\sum_{p=0}^{n}\binom{n}{p} p!\, e^p \;=\; 2n!\, e^n \sum_{p=0}^{n}\frac{e^{p-n}}{(n-p)!} \;\overset{q=n-p}{=}\; 2n!\, e^n \sum_{q=0}^{n}\frac{e^{-q}}{q!} \;<\; 2n!\, e^n \sum_{q=0}^{\infty}\frac{e^{-q}}{q!} \;=\; 2n!\, e^n e^{e^{-1}}\,.$$

Como $e^{-1} < 1$ e $n! < (n+1)!$, obtemos $T_{n+1} < 2(n+1)!\, e^{n+1}$, provando (6.24) para todo $n$.

**• A função geratriz exponencial dos números de Bell**

A função geratriz exponencial dos números de Bell $T(x) = \sum_{n=0}^{\infty}\frac{x^n}{n!}T_n$ pode ser calculada facilmente. Poderemos usar sua fórmula explícita para obter melhores estimativas para $T_n$ quando $n$ é "grande".

Antes, observemos, porém, que devido à estimativa grosseira (6.24), podemos afirmar que a série que define $T(x)$ é absolutamente convergente ao menos na região $|x| < e^{-1}$, $x \in \mathbb{C}$, e define uma função analítica nessa região. Os cômputos

[24] A saber, para demonstrar convergência da função geratriz exponencial associada aos números de Bell em um certo aberto de $\mathbb{C}$.



a seguir são autoexplicativos.

$$T(x) \;=\; \sum_{n=0}^{\infty}\frac{x^n}{n!}T_n \;=\; 1+\sum_{n=1}^{\infty}\frac{x^n}{n!}T_n \;\overset{n\to n+1}{=}\; 1+\sum_{n=0}^{\infty}\frac{x^{n+1}}{(n+1)!}T_{n+1} \;\overset{(6.20)}{=}\; 1+\sum_{n=0}^{\infty}\sum_{p=0}^{n}\frac{x^{n+1}}{(n+1)!}\binom{n}{p}T_p$$

$$= 1+\sum_{p=0}^{\infty}\sum_{n=p}^{\infty}\frac{x^{n+1}}{(n+1)!}\binom{n}{p}T_p \;=\; 1+\sum_{p=0}^{\infty}\frac{T_p}{p!}\left(\sum_{n=p}^{\infty}\frac{x^{n+1}}{(n+1)(n-p)!}\right)\,.$$

Com a mudança $n \to n+p$ temos, finalmente,

$$T(x) \;=\; 1+\sum_{p=0}^{\infty}\frac{T_p}{p!}\left(\sum_{n=0}^{\infty}\frac{x^{n+p+1}}{(n+p+1)n!}\right)\,.$$

Diferenciando ambos os lados dessa expressão em relação a $x$, e diferenciando a série de potências termo a termo, o que é permitido na sua região de convergência, obtemos

$$T'(x) \;=\; \sum_{p=0}^{\infty}\frac{T_p}{p!}\left(\sum_{n=0}^{\infty}\frac{x^{n+p}}{n!}\right) \;=\; \sum_{p=0}^{\infty}\frac{T_p}{p!}x^p\left(\sum_{n=0}^{\infty}\frac{x^n}{n!}\right) \;=\; e^x\sum_{p=0}^{\infty}\frac{T_p}{p!}x^p \;=\; e^x T(x)\,.$$

Assim, $T(x)$ satisfaz a equação diferencial $T'(x) = e^x T(x)$ com a condição inicial $T(0) = T_0 = 1$. Como facilmente se vê, a equação diferencial tem por solução $T(x) = \exp\big(e^x + c\big)$, com $c$ sendo uma constante, e a imposição da condição inicial implica $c = -1$. Logo, concluímos que

$$T(x) \;=\; \exp\big(e^x - 1\big) \;=\; e^{-1}e^{e^x}\,. \tag{6.25}$$

Observe-se que essa função é inteira, *i.e.*, analítica em toda a parte como função de $x \in \mathbb{C}$ e, portanto, a série de potências que a define é também convergente em toda parte, não apenas na região $|x| < e^{-1}$.

Dessa expressão, podemos obter mais uma prova da fórmula de Dobiński. De fato, temos por (6.25),

$$T(x) \;=\; e^{-1}\sum_{k=0}^{\infty}\frac{e^{kx}}{k!}\,. \qquad \text{Assim, para a } n\text{-ésima derivada, temos} \qquad T^{(n)}(x) \;=\; e^{-1}\sum_{k=0}^{\infty}\frac{k^n}{k!}e^{kx}\,.$$

Pela definição da função geratriz exponencial, temos $T^{(n)}(0) = T_n$, e tomando-se $x = 0$ na expressão acima, temos $T_n = T^{(n)}(0) = e^{-1}\sum_{k=0}^{\infty}\frac{k^n}{k!}$, que é novamente a fórmula de Dobiński.

**• Uma majoração melhor para os números $T_n$**

A expressão para a função geratriz $T$ permite aperfeiçoar as estimativas para majorantes de $T_n$. Como $T_n = T^{(n)}(0)$, a fórmula integral de Cauchy nos informa que

$$T_n \;=\; \frac{n!}{2\pi i e}\oint_C \frac{e^{e^z}}{z^{n+1}}\,dz\,, \tag{6.26}$$

onde $C$ é um circuito fechado dando uma volta em torno da origem em sentido anti-horário. Usando que $\left|e^{e^z}\right| \leq e^{|e^z|} \leq e^{e^{|z|}}$ e tomando $C$ como um círculo de raio $R$ centrado na origem, obtemos facilmente a majoração

$$T_n \;\leq\; \frac{e^{e^R}}{e}\frac{n!}{R^n}\,, \tag{6.27}$$

a qual, para $R > 1$, já fornece uma estimativa assintótica melhor que a de (6.24). Nada nos impede de escolher $R$ como função de $n$ na expressão (6.27). Se, por exemplo, adotarmos $R = \ln(n+1)$ para $n > 0$ e usarmos a estimativa $n! < \sqrt{2\pi}\, n^{n+\frac{1}{2}}e^{-n}e^{\frac{1}{12n}}$, obtida em (7.105), página 376, teremos,

$$T_n \;<\; \sqrt{2\pi n}\,\left(\frac{n}{\ln(n+1)}\right)^n e^{\frac{1}{12n}}\,, \qquad n \;>\; 0\,. \tag{6.28}$$

Verifique!

Usando o método conhecido como "*steepest descent*" para estimar a integral complexa em (6.26), é possível obter estimativas assintóticas ainda melhores para os números $T_n$, mas não discutiremos mais essa questão aqui e remetemos o leitor interessado à literatura pertinente. Estimativas assintóticas para os números $T_n$ têm relevância na Teoria de Números e na Análise Combinatória, sendo ainda hoje objeto de pesquisa.

## 6.2 Notas Sobre Convergência de Produtórias

Seja $\{a_n,\ n\in\mathbb{N}\}$ uma sequência de números complexos. Definimos a produtória infinita $\prod_{k=1}^{\infty} a_k$ como o limite $\lim_{M\to\infty}\prod_{k=1}^{M} a_k$ se este existir. Podemos escrever $\prod_{k=1}^{M} a_k = \exp\left(\sum_{k=1}^{M}\log a_k\right)$ e, assim, tentar associar a convergência da sequência $\prod_{k=1}^{M} a_k$ à convergência da sequência $\sum_{k=1}^{M}\log a_k$ quando $M\to\infty$. Com isso poderíamos reduzir a questão da convergência de produtórias infinitas à questão da convergência de séries, tema sobre o qual muito é conhecido. Há, porém, dois problemas aqui: em primeiro lugar, alguns $a_k$'s podem ser nulos (em cujo caso desejamos que $\prod_{k=1}^{\infty} a_k$ seja nula) e para tais $a_k$'s o logaritmo $\log a_k$ não está definido. Em segundo lugar, a questão da convergência de $\sum_{k=1}^{M}\log a_k$ pode depender da folha onde os logaritmos complexos estão definidos, ponto que é particularmente importante no caso se que $\{a_n,\ n\in\mathbb{N}\}$ é uma sequência de funções em $\mathbb{C}$. A proposição a seguir fornece condições suficientes sob as quais esses problemas podem ser contornados.

No que segue, log refere-se sempre à primeira folha do logaritmo, isto é, para $a=|a|e^{-i\phi}\in\mathbb{C}$ com $\phi\in(-\pi,\ \pi]$ e $|a|>0$, o logaritmo $\log a$ é definido por $\log a := \ln|a| + i\phi$.

**Proposição 6.1** *Seja $\{a_n,\ n\in\mathbb{N}\}$ uma sequência de números complexos. Uma condição suficiente para que a produtória $\prod_{k=1}^{\infty} a_k$ seja convergente é que valha $\sum_{n=1}^{\infty}|a_n-1|<\infty$, em cujo caso tem-se $\prod_{k=1}^{\infty} a_k = \left[\prod_{k=1}^{N} a_k\right]\exp\left(\sum_{n=N+1}^{\infty}\log a_n\right)$ para todo $N\in\mathbb{N}$ suficientemente grande. Sob as hipóteses, a série $\sum_{n=N+1}^{\infty}\log a_n$ é absolutamente convergente.* □

*Comentários.* **1.** A condição $\sum_{n=1}^{\infty}|a_n-1|<\infty$ implica que pode haver no máximo um número finito de $a_n$'s nulos, os quais, se ocorrerem, estarão no fator $\prod_{k=1}^{N} a_k$ o qual será, portanto, nulo, implicando a nulidade de $\prod_{k=1}^{\infty} a_k$. **2.** A convergência absoluta de $\sum_{n=N+1}^{\infty}\log a_n$ assegura que a produtória $\prod_{k=1}^{\infty} a_k$ é invariante por permutações dos índices. ♣

Prova da Proposição 6.1. Seja $z\in\mathbb{C}$ com $|z|<1$. Sabemos que $\log(1+z)=\sum_{m=1}^{\infty}\frac{(-1)^{m+1}z^m}{m}$. Logo, $|\log(1+z)|\leq\sum_{m=1}^{\infty}|z|^m=\frac{|z|}{1-|z|}$. Assim, para $|z|<1/2$, teremos $|\log(1+z)|<2|z|$.

Se $\sum_{n=1}^{\infty}|a_n-1|$ converge, então existe $N\in\mathbb{N}$ tal que $|a_n-1|<1/2$ para todo $n\geq N$. Escrevendo $\log a_n=\log\big(1+(a_n-1)\big)$ concluímos que $\big|\log a_n\big|<2|a_n-1|$ para todo $n\geq N$. Essa desigualdade, junto à hipótese que $\sum_{n=1}^{\infty}|a_n-1|<\infty$, implica que $\sum_{n=N+1}^{\infty}\log a_n$ converge absolutamente. Para $M>N$ temos $\prod_{k=1}^{M} a_k = \left[\prod_{k=1}^{N} a_k\right]\exp\left(\sum_{n=N+1}^{M}\log a_n\right)$ e a convergência absoluta de $\sum_{n=N+1}^{\infty}\log a_n$ implica a existência do limite $\lim_{M\to\infty}\prod_{k=1}^{M} a_k$. ■

A Proposição 6.1 tem um corolário extremamente útil que diz respeito a produtórias de funções analíticas em um domínio compacto comum.

**Corolário 6.1** *Seja $\{a_n(z),\ n\in\mathbb{N}\}$ uma sequência de funções analíticas em um domínio compacto comum $K\subset\mathbb{C}$ e suponhamos que $\sum_{n=1}^{\infty}\sup_{z\in K}|a_n(z)-1|<\infty$. Então, a produtória infinita $\prod_{k=1}^{\infty} a_k(z)$ converge uniformemente em $K$ e define nesse domínio uma função analítica. Essa função pode ser escrita como $\prod_{k=1}^{\infty} a_k(z) = \left[\prod_{k=1}^{N} a_k(z)\right]\exp\left(\sum_{n=N+1}^{\infty}\log a_n(z)\right)$ para todo $N\in\mathbb{N}$ suficientemente grande, sendo que $\sum_{n=N+1}^{\infty}\log a_n(z)$ converge absoluta e uniformemente em $K$ e define uma função analítica nesse domínio.* □

Prova. A hipótese $\sum_{n=1}^{\infty}\sup_{z\in K}|a_n(z)-1|<\infty$ implica que existe $N\in\mathbb{N}$ tal que $\sup_{z\in K}|a_n(z)-1|<1/2$ para todo $n\geq N$ e, portanto, valem para todo $n\geq N$ as afirmações que $a_n$ não se anula em $K$ e que $\log\big(a_n(z)\big)$ é analítica em $K$. Pela Proposição 6.1, vale também que $\prod_{k=1}^{\infty} a_k(z)$ converge para cada $z\in K$ e vale $\prod_{k=1}^{\infty} a_k(z) =$



$\left[\prod_{k=1}^{N} a_k(z)\right]\exp\left(\sum_{n=N+1}^{\infty}\log a_n(z)\right)$, sendo que $\sum_{n=N+1}^{\infty}\log a_n(z)$ converge absolutamente para cada $z\in K$. Como vimos na demonstração da Proposição 6.1, vale $|\log a_n(z)|<2|a_n(z)-1|\leq 2\sup_{z\in K}|a_n(z)-1|$. Com isso, vemos pela hipótese $\sum_{n=1}^{\infty}\sup_{z\in K}|a_n(z)-1|<\infty$ que $\sum_{n=N+1}^{\infty}\log a_n(z)$ converge uniformemente em $K$ e, portanto, define uma função analítica nesse domínio. Isso implica também a convergência uniforme em $K$ da sequência de produtórias finitas $\prod_{k=1}^{M} a_k(z)$. Note-se que cada função $\prod_{k=1}^{M} a_k(z)$ é analítica em $K$ (por ser um produto finito de funções analíticas). Portanto, o limite $\prod_{k=1}^{\infty} a_k(z)$ define igualmente uma função analítica em $K$, por ser o limite uniforme de uma sequência de funções analíticas em um compacto. ■

### 6.2.1 Uma Dedução Elementar do Produto de Wallis

Como aplicação de nossa discussão anterior, vamos aqui apresentar uma dedução direta da chamada *fórmula de Wallis*, ou *fórmula do produto de Wallis*. Uma segunda demonstração independente será também apresentada em nossa discussão de propriedades da Função Gama de Euler, no Capítulo 7, página 344. Vide relação (7.46), página 359.

Para $n\geq 2$, vale a identidade

$$\big(\operatorname{sen} x\big)^n \;=\; -\frac{1}{n}\frac{d}{dx}\Big(\cos(x)\big(\operatorname{sen} x\big)^{n-1}\Big)+\frac{n-1}{n}\big(\operatorname{sen} x\big)^{n-2}, \tag{6.29}$$

cuja prova segue das linhas seguintes:

$$\begin{aligned}\big(\operatorname{sen} x\big)^n \;&=\; -\left(\frac{d}{dx}\cos x\right)\big(\operatorname{sen} x\big)^{n-1} \;=\; -\frac{d}{dx}\Big(\cos(x)\big(\operatorname{sen} x\big)^{n-1}\Big)+\cos x\frac{d}{dx}\Big(\big(\operatorname{sen} x\big)^{n-1}\Big)\\ &=\; -\frac{d}{dx}\Big(\cos(x)\big(\operatorname{sen} x\big)^{n-1}\Big)+(n-1)\big(\cos x\big)^2\big(\operatorname{sen} x\big)^{n-2}\\ &=\; -\frac{d}{dx}\Big(\cos(x)\big(\operatorname{sen} x\big)^{n-1}\Big)+(n-1)\big(\operatorname{sen} x\big)^{n-2}-(n-1)\big(\operatorname{sen} x\big)^{n}.\end{aligned}$$

Com o uso de (6.29) pode-se calcular recursivamente integrais envolvendo potências da função seno. Definindo para todo $n\in\mathbb{N}_0$

$$J_n \;:=\; \int_0^{\pi/2}\big(\operatorname{sen} x\big)^n\,dx$$

teremos por (6.29), para $n\geq 2$, as relações recursivas

$$J_n \;=\; \frac{n-1}{n}J_{n-2}. \tag{6.30}$$

Verifique! Observe-se agora que

$$J_0 \;=\; \frac{\pi}{2} \qquad \text{e} \qquad J_1 \;=\; 1.$$

Logo, por (6.30), valem para todo $m\in\mathbb{N}$,

$$J_{2m} \;=\; \frac{\pi}{2}\prod_{k=1}^{m}\frac{2k-1}{2k} \;=\; \frac{\pi}{2}\frac{(2m-1)!!}{(2m)!!}, \tag{6.31}$$

$$J_{2m+1} \;=\; \prod_{k=1}^{m}\frac{2k}{2k+1} \;=\; \frac{(2m)!!}{(2m+1)!!}. \tag{6.32}$$

Verifique! Dividindo-se a linha de cima pela de baixo obtém-se facilmente que

$$\frac{\pi}{2} \;=\; \left[\prod_{k=1}^{m}\frac{2k}{2k-1}\,\frac{2k}{2k+1}\right]\frac{J_{2m}}{J_{2m+1}},$$

identidade essa que também pode ser expressa na forma

$$\frac{\pi}{2} \;=\; \frac{\big[(2m)!!\big]^2}{(2m-1)!!(2m+1)!!}\,\frac{J_{2m}}{J_{2m+1}} \;=\; \left[\frac{(2m)!!}{(2m-1)!!}\right]^2 \frac{1}{2m+1}\,\frac{J_{2m}}{J_{2m+1}} \,. \tag{6.33}$$

Como essas identidades valem para todo $m \in \mathbb{N}$, é natural perguntarmos o que sucede se tomarmos o limite $m \to \infty$ do lado direito das mesmas. Para responder a isso precisamos entender o que acontece no limite $m \to \infty$ com a razão $J_{2m}/J_{2m+1}$.

Como $0 < \operatorname{sen}x < 1$ para $x \in (0,\ \pi/2)$, é evidente que nesse intervalo vale também $0 < \big(\operatorname{sen}x\big)^{2m+1} < \big(\operatorname{sen}x\big)^{2m} < \big(\operatorname{sen}x\big)^{2m-1}$. Logo,

$$1 \;<\; \frac{\int_0^{\pi/2}\big(\operatorname{sen}x\big)^{2m}dx}{\int_0^{\pi/2}\big(\operatorname{sen}x\big)^{2m+1}dx} \;<\; \frac{\int_0^{\pi/2}\big(\operatorname{sen}x\big)^{2m-1}dx}{\int_0^{\pi/2}\big(\operatorname{sen}x\big)^{2m+1}dx}$$

o que, face a (6.30), significa que

$$1 \;<\; \frac{J_{2m}}{J_{2m+1}} \;<\; 1+\frac{1}{2m}\,.$$

Verifique! Logo, $\lim_{m\to\infty}\frac{J_{2m}}{J_{2m+1}} = 1$ e, portanto, por (6.33),

$$\lim_{m\to\infty}\frac{\frac{\pi}{2}}{\prod_{k=1}^m \frac{2k}{2k-1}\frac{2k}{2k+1}} \;=\; 1$$

o que implica que o limite $\lim_{m\to\infty}\prod_{k=1}^m \frac{2k}{2k-1}\frac{2k}{2k+1}$ existe e que vale

$$\frac{\pi}{2} \;=\; \lim_{m\to\infty}\prod_{k=1}^m \frac{2k}{2k-1}\,\frac{2k}{2k+1}\,. \tag{6.34}$$

Essa é a *fórmula de Wallis*[25], também denominada *fórmula do produto de Wallis* ou *produto de Wallis*, obtida em 1665, e que pode ainda ser expressa como

$$\frac{\pi}{2} \;=\; \lim_{m\to\infty}\prod_{k=1}^m \frac{4k^2}{4k^2-1} \tag{6.35}$$

ou como

$$\frac{\pi}{2} \;=\; \lim_{m\to\infty}\left[\frac{(2m)!!}{(2m-1)!!}\right]^2\frac{1}{2m+1}\,. \tag{6.36}$$

Note-se que a convergência da produtória infinita $\prod_{k=1}^\infty \frac{2k}{2k-1}\frac{2k}{2k+1}$ pode ser provada diretamente do critério listado na Proposição 6.1, página 325, pois

$$\sum_{k=1}^\infty\left|\frac{2k}{2k-1}\,\frac{2k}{2k+1}-1\right| \;=\; \sum_{k=1}^\infty\frac{1}{4k^2-1} \;<\; \infty\,.$$

Usando os fatos bem conhecidos que $(2m)!! = 2^m m!$ e $(2m)! = (2m)!!(2m-1)!!$ (Exercício E. 7.14, página 380), podemos também reescrever (6.35) na forma

$$\frac{\pi}{2} \;=\; \lim_{m\to\infty}\frac{2^{4m}(m!)^4}{\big((2m)!\big)^2}\,\frac{1}{2m+1}\,. \tag{6.37}$$

Verifique! Essa expressão é útil na obtenção da chamada aproximação de Stirling, como tratado na Seção 7.6.1, página 370.

[25] John Wallis (1616–1703).



## 6.3 A Fórmula de Inversão de Möbius. As Fórmulas de Viète

Seja $X$ um conjunto não vazio e seja $\mathbb{P}(X)$ a coleção de todos os subconjuntos de $X$, incluindo o vazio. Em muitas situações deparamos com funções definidas em $\mathbb{P}(X)$ assumindo valores em $\mathbb{R}$ ou em $\mathbb{C}$. Nesta seção demonstraremos a elegante *fórmula de inversão de Möbius*[26], a qual possui diversas aplicações em Mecânica Estatística (expansões de polímeros e de "clusters", expansões de Mayer[27]), em Teorias de Campos, na Teoria de Probabilidades e na Análise Combinatória.

Se $A$ é um conjunto finito, denotaremos por $|A|$ a cardinalidade de $A$, ou seja, o número de elementos de $A$. Naturalmente, $|\emptyset| = 0$. Note que se $A\cap B = \emptyset$, então $|A\cup B| = |A|+|B|$.

**Lema 6.1** *Seja $M$ um conjunto finito (podendo eventualmente ser vazio) e seja $f$ uma função real ou complexa definida em $M$. Então, vale*

$$\prod_{m\in M}\big(1+f(m)\big) \;=\; \sum_{N\subset M}\prod_{n\in N} f(n)\,, \tag{6.38}$$

*com a convenção $\prod_{x\in\emptyset} f(x) \equiv 1$. Uma consequência importante de (6.38) é a relação*

$$(1+a)^{|M|} \;=\; \sum_{N\subset M} a^{|N|}\,, \tag{6.39}$$

*válida para todo $a$ real ou complexo, a qual tem por implicações a expressão*

$$|\mathbb{P}(M)| \;=\; 2^{|M|}\,, \tag{6.40}$$

*assim como a importante relação*

$$\sum_{N\subset M}(-1)^{|N|} \;=\; \begin{cases} 0\,, & \text{se } M\neq\emptyset\,,\\[2mm] 1\,, & \text{se } M=\emptyset\,.\end{cases} \tag{6.41}$$

□

Prova do Lema 6.1. A relação (6.38) pode ser provada por indução no número de elementos de $M$. Ela é naturalmente válida para $|M| = 0$ e para $|M| = 1$ e é válida também para $|M| = 2$, pois se $M = \{m_1,\ m_2\}$, $(1+f(m_1))(1+f(m_2)) = 1+f(m_1)+f(m_2)+f(m_1)f(m_2)$, que tem a forma do lado direito de (6.38), como facilmente se vê, pois os subconjuntos de $M$ são $\emptyset$, $\{m_1\}$, $\{m_2\}$ e $\{m_1,\ m_2\}$. Seja, então, $M = \{m_1,\ \ldots,\ m_l\}$, com $|M| = l$, sendo $l > 2$, e suponha que (6.38) já tenha sido estabelecida para conjuntos com $l-1$ elementos. Teremos

$$\prod_{m\in M}\big(1+f(m)\big) \;=\; \left[\prod_{j=1}^{l-1}\big(1+f(m_j)\big)\right]\big(1+f(m_l)\big) \;=\; \left[\sum_{N_0\subset\{m_1,\ \ldots,\ m_{l-1}\}}\prod_{n\in N_0} f(n)\right]\big(1+f(m_l)\big)$$

$$=\; \sum_{N_0\subset\{m_1,\ \ldots,\ m_{l-1}\}}\prod_{n\in N_0} f(n) + \sum_{N_0\subset\{m_1,\ \ldots,\ m_{l-1}\}}\left[\prod_{n\in N_0} f(n)\right] f(m_l) \;=\; \sum_{N\subset\{m_1,\ \ldots,\ m_l\}}\prod_{n\in N} f(n)\,,$$

como facilmente se vê. A relação (6.39) é evidente caso $M=\emptyset$ e, caso $M\neq\emptyset$, ela segue de (6.38) para a função constante definida por $f(m) = a$ para todo $m\in M$. A relação (6.40) segue de (6.39) tomando-se $a = 1$. A relação (6.41) é evidente no caso $M = \emptyset$. No caso $M\neq\emptyset$, ela segue de (6.39) tomando-se $a = -1$. ■

[26] August Ferdinand Möbius (1790–1868).

[27] Joseph Eduard Mayer (1904–1983) e Maria Goeppert-Mayer (1906–1972). Vide J. E. Mayer and M. G. Mayer, "Statistical Mechanics", Wiley, New York, (1940). Vide também [188].

**Proposição 6.2 (Fórmula de Inversão de Möbius)** *Seja $X$ um conjunto não vazio e finito e seja $F : \mathbb{P}(X) \to \mathbb{C}$ uma função complexa (ou real) definida em $\mathbb{P}(X)$. Seja $G : \mathbb{P}(X) \to \mathbb{C}$ definida por*

$$G(B) := \sum_{A:\ A\subset B} (-1)^{|A|} F(A) \tag{6.42}$$

*para todo $B \subset X$. Então, vale*

$$F(C) = \sum_{B:\ B\subset C} (-1)^{|B|} G(B) , \tag{6.43}$$

*para todo $C \subset X$.* □

A função $G$ definida em (6.42) é por vezes dita ser a *transformada de Möbius* da função $F$. A expressão (6.43) é denominada *fórmula de inversão de Möbius*. Por (6.43) vemos que a transformada de Möbius tem a si mesma como inversa. A fórmula de inversão de Möbius permite obter facilmente identidades combinatórias cuja demonstração pode não ser tão simples por outros meios.

Prova da Proposição 6.2. Pelas definições,

$$\begin{aligned}
\sum_{B:\ B\subset C} (-1)^{|B|} G(B) \quad &= \quad \sum_{B:\ B\subset C}\ \sum_{A:\ A\subset B} (-1)^{|B|}(-1)^{|A|} F(A) \\
&= \quad \sum_{A:\ A\subset C}\ \sum_{B:\ C\supset B\supset A} (-1)^{|B|}(-1)^{|A|} F(A) \\
&= \quad \sum_{A:\ A\subset C} (-1)^{|A|} F(A) \left( \sum_{B:\ C\supset B\supset A} (-1)^{|B|} \right) \\
&= \quad \sum_{A:\ A\subset C} (-1)^{|A|} F(A) \left( \sum_{B':\ B'\subset C\setminus A} (-1)^{|A\cup B'|} \right) \\
&\overset{B'\cap A=\emptyset}{=} \quad \sum_{A:\ A\subset C} F(A) \left( \sum_{B':\ B'\subset C\setminus A} (-1)^{|B'|} \right) \\
&\overset{(6.41)}{=} \quad F(C) .
\end{aligned}$$

Para a última igualdade, note que $A \subset C$ e $C \setminus A = \emptyset$ implicam $A = C$. Na quarta igualdade usamos que se $B$ é tal que $C \supset B \supset A$, então $B$ pode ser escrito como a união disjunta $B = A \cup B'$ com $B' \subset C \setminus A$ e, portanto, $|B| = |A| + |B'|$. ■

**E. 6.10** *Exercício*. Usando (6.43), prove que se $C$ é um conjunto finito, então

$$\sum_{B:\ B\subset C} (-1)^{|B|} \big|\mathbb{P}(B)\big| = (-1)^{|C|} , \tag{6.44}$$

com a convenção que $\mathbb{P}(\emptyset) = \{\emptyset\}$ e, portanto, $\big|\mathbb{P}(\emptyset)\big| = 1$. Por (6.40), a expressão (6.44) afirma que

$$\sum_{B:\ B\subset C} (-2)^{|B|} = (-1)^{|C|} , \tag{6.45}$$

A identidade (6.45) pode também ser provada diretamente de (6.39) tomando-se $a = -2$. ✣



### 6.3.1 As Fórmulas de Viète

Vamos agora apresentar mais uma aplicação útil de (6.38): as chamadas *Fórmulas de Viète*[28], também denominadas *Fórmulas de Girard*[29], que são relações que expressam os coeficientes de um polinômio em função de suas raízes.

Seja $p(x) = x^m + a_{m-1}x^{m-1} + \cdots + a_1 x + a_0$ um polinômio mônico[30] de grau $m \geq 1$. Os coeficientes $a_k$ podem ser reais ou complexos, assim como a variável $x$.

Se suas raízes (não necessariamente distintas) são $x_1, \ldots, x_m$, é bem sabido que podemos escrever $p(x) = \prod_{k=1}^{m}(x - x_k)$. Para $x \neq 0$, vale $p(x) = x^m \prod_{k=1}^{m}\big(1 + f(k)\big)$ onde, para $k \in \{1, \ldots, m\} \equiv M$, temos $f(k) := -x_k/x$.

Fazendo uso de (6.38), escrevemos,

$$p(x) = x^m \sum_{N\subset M} \prod_{l\in N} \left(-\frac{x_l}{x}\right) = \sum_{N\subset M} (-1)^{|N|} x^{m-|N|} \prod_{l\in N} x_l = \sum_{n=0}^{m} (-1)^n x^{m-n} \sum_{\substack{N\subset M \\ |N|=n}} \prod_{l\in N} x_l .$$

Assim, como $p(x) = \sum_{n=0}^{m} a_{m-n} x^{m-n}$ com $a_m = 1$ vemos, comparando os coeficientes, que

$$a_{m-n} = (-1)^n \sum_{\substack{N\subset M \\ |N|=n}} \prod_{l\in N} x_l , \qquad n = 1, \ldots, m . \tag{6.46}$$

Essas expressões são denominadas *Fórmulas de Viète*, ou *Fórmulas de Girard*. Elas afirmam que

$$\begin{aligned}
a_{m-1} &= -\big(x_1 + \cdots + x_m\big) , \\
a_{m-2} &= \Big(x_1x_2 + x_1x_3 + \cdots + x_1x_m + x_2x_3 + \cdots + x_2x_m + \cdots + x_{m-1}x_m\Big) = \sum_{\substack{i,\, j=1 \\ i<j}}^{m} x_i x_j , \\
a_{m-3} &= -\sum_{\substack{i,\, j,\, k=1 \\ i<j<k}}^{m} x_i x_j x_k , \\
&\ \vdots \\
a_0 &= (-1)^m x_1 x_2 \cdots x_m .
\end{aligned} \tag{6.47}$$

De forma geral, temos

$$a_{m-k} = (-1)^k \sum_{1\leq i_1<i_2<\cdots<i_k\leq m} x_{i_1} x_{i_2} \cdots x_{i_k} , \tag{6.48}$$

para $k = 1, \ldots, m$. Na expressão para $a_{m-2}$, por exemplo, o lado direito envolve somas de todos os produtos possíveis de duas raízes distintas. As demais expressões são similares.

Um fato de grande relevância, e que se percebe mais claramente com as fórmulas de Viète (6.46) e (6.47), é que os coeficientes são polinômios simétricos (por permutações) das raízes. Essa simples observação é a raiz histórica da *Teoria de Grupos* e da área da Álgebra denominada *Teoria de Galois*[31] (vide *e.g.*, [276]), cujo primeiro sucesso foi o de determinar condições necessárias e suficientes para que os zeros de um polinômio possam ser determinadas por radicais, um célebre problema da Matemática que permaneceu aberto por séculos até sua dramática solução por Galois em 1832.

[28] François Viète (1540–1603).
[29] Albert Girard (1595–1632).
[30] Um polinômio é dito *mônico* se o coeficiente de seu monômio de maior grau for 1.
[31] Évariste Galois (1811-1832). Galois morreu tragicamente em um duelo. Suas últimas palavras a seu irmão Alfred teriam sido "Ne pleure pas, Alfred! J'ai besoin de tout mon courage pour mourir à vingt ans!". Vide [36].

#### 6.3.1.1 Uma Aplicação. Localizando Zeros de Certos Polinômios

As fórmulas de Viète (6.47) possuem muitos usos e vamos no que segue apresentar um deles, que fornece informação sobre a localização de raízes de polinômios cujas raízes sejam reais.

No que segue, consideraremos um polinômio mônico $p(x) = x^m + a_{m-1}x^{m-1} + \cdots + a_1 x + a_0$ cujos zeros $x_1, \ldots, x_m$ sejam todos reais. Nossa primeira observação é que, como os zeros de $p$ são supostos reais, então as relações de Viète em (6.47) nos ensinam que os coeficientes $a_0, \ldots, a_{m-1}$ de $p$ são também números reais.

Faremos uso do seguinte lema:

**Lema 6.2** *Seja $p(x) = x^m + a_{m-1}x^{m-1} + \cdots + a_1 x + a_0$ um polinômio mônico cujos zeros $x_1, \ldots, x_m$ sejam todos reais. Então,*

$$\frac{m-1}{m}\big(a_{m-1}\big)^2 \;\geq\; 2a_{m-2} \; . \tag{6.49}$$

Prova. Pelas duas primeiras relações em (6.47), temos

$$a_{m-1} \;=\; -\big(x_1 + \cdots + x_m\big) \; , \tag{6.50}$$

$$a_{m-2} \;=\; \sum_{\substack{i,\, j=1 \\ i<j}}^{m} x_i x_j \; . \tag{6.51}$$

Assim,

$$\big(a_{m-1}\big)^2 \overset{(6.50)}{=} \big(x_1 + \cdots + x_m\big)^2 \;=\; 2\sum_{\substack{i,\, j=1 \\ i<j}}^{m} x_i x_j + \sum_{k=1}^{m} x_k^2 \overset{(6.51)}{=} 2a_{m-2} + \sum_{k=1}^{m} x_k^2 \; .$$

Pela desigualdade de Cauchy-Schwarz,

$$\big|x_1 + \cdots + x_m\big| \;=\; \big|1\cdot x_1 + \cdots + 1\cdot x_m\big| \;\leq\; \sqrt{\underbrace{1^2 + \cdots + 1^2}_{m \text{ vezes}}}\;\sqrt{x_1^2 + \cdots + x_m^2} \; ,$$

e disso segue

$$x_1^2 + \cdots + x_m^2 \;\geq\; \frac{1}{m}\big|x_1 + \cdots + x_m\big|^2 \overset{(6.50)}{=} \frac{1}{m}\big(a_{m-1}\big)^2 \; . \tag{6.52}$$

Logo,

$$\big(a_{m-1}\big)^2 \;\geq\; 2a_{m-2} + \frac{1}{m}\big(x_1 + \cdots + x_m\big)^2 \overset{(6.50)}{=} 2a_{m-2} + \frac{1}{m}\big(a_{m-1}\big)^2 \; ,$$

ou seja, $\frac{m-1}{m}\big(a_{m-1}\big)^2 \geq 2a_{m-2}$, como desejávamos provar. ■

Nosso resultado principal sobre localização de zeros de um polinômio com zeros reais é a proposição a seguir. Na demonstração, seguimos proximamente a referência [7], que atribui esse resultado a Laguerre[32] [33].

A Proposição 6.3 é relevante na determinação numérica de raízes de polinômios (se todas as raízes forem reais), por indicar uma região região específica onde eles podem ser procurados, por exemplo, pelo método de Newton. Usaremos a Proposição 6.3 para obter estimativas sobre o espectro de matrizes autoadjuntas na Seção 10.5.3, página 524.

**Proposição 6.3** *Seja $p(x) = x^m + a_{m-1}x^{m-1} + \cdots + a_1 x + a_0$ um polinômio mônico cujos zeros $x_1, \ldots, x_m$ sejam todos reais. Então, todos os zeros de $p$ estão localizados no intervalo $\big[\alpha_m^-,\, \alpha_m^+\big]$, onde*

$$\alpha_m^{\pm} \;:=\; \frac{-a_{m-1} \pm \sqrt{\big((m-1)a_{m-1}\big)^2 - 2m(m-1)a_{m-2}}}{m} \; , \tag{6.53}$$

[32] Edmond Nicolas Laguerre (1834–1886).

[33] A referência original seria E. Laguerre, "Mémoire pour obtenir par approximation les racines d'une équation algébrique qui a toutes les racines réelles". Nouv. Ann. Math., 2e série, **19**, 161–172 and 193–202 (1880).



*ou seja, todos os zeros pertencem ao intervalo fechado de largura $2\frac{\sqrt{\big((m-1)a_{m-1}\big)^2 - 2m(m-1)a_{m-2}}}{m}$ centrado em $\frac{-a_{m-1}}{m}$. Observe-se que (6.49) garante que $\alpha_m^{\pm}$ são ambos reais.* □

Prova. Pelas duas primeiras relações de Viète em (6.47), temos

$$a_{m-1} \;=\; -\big(x_1 + \cdots + x_m\big) \; , \tag{6.54}$$

$$a_{m-2} \;=\; x_1\Big(x_2 + \cdots + x_m\Big) + \sum_{\substack{i,\, j=2 \\ i<j}}^{m} x_i x_j \; . \tag{6.55}$$

Assim,

$$\begin{aligned}
\big(a_{m-1}\big)^2 &= \big(x_1 + \cdots + x_m\big)^2 \;=\; x_1^2 + 2x_1\big(x_2 + \cdots + x_m\big) + \big(x_2 + \cdots + x_m\big)^2 \\
&\overset{(6.54)}{=} x_1^2 + 2\left(a_{m-2} - \sum_{\substack{i,\, j=2 \\ i<j}}^{m} x_i x_j\right) + \big(x_2 + \cdots + x_m\big)^2 \\
&= x_1^2 + 2a_{m-2} + \sum_{i=2}^{m} x_i^2 \; .
\end{aligned}$$

Portanto,

$$\sum_{i=2}^{m} x_i^2 \;=\; \big(a_{m-1}\big)^2 - x_1^2 - 2a_{m-2} \; . \tag{6.56}$$

Por outro lado, temos, também por (6.54),

$$\big(a_{m-1} + x_1\big)^2 \;=\; \big(x_2 + \cdots + x_m\big)^2 \; , \tag{6.57}$$

mas pela desigualdade de Cauchy-Schwartz

$$\big|x_2 + \cdots + x_m\big| \;=\; \big|1\cdot x_2 + \cdots + 1\cdot x_m\big| \;\leq\; \sqrt{\underbrace{1^2 + \cdots + 1^2}_{m-1 \text{ vezes}}}\;\sqrt{x_2^2 + \cdots + x_m^2} \; .$$

Assim,

$$\big(x_2 + \cdots + x_m\big)^2 \;\leq\; (m-1)\big(x_2^2 + \cdots + x_m^2\big) \; .$$

Retornando com isso a (6.57), obtemos

$$\big(a_{m-1} + x_1\big)^2 \;\leq\; (m-1)\big(x_2^2 + \cdots + x_m^2\big) \overset{(6.56)}{=} (m-1)\Big(\big(a_{m-1}\big)^2 - x_1^2 - 2a_{m-2}\Big) \; ,$$

ou seja,

$$m x_1^2 + 2a_{m-1}x_1 - (m-2)\big(a_{m-1}\big)^2 + 2(m-1)a_{m-2} \;\leq\; 0 \; . \tag{6.58}$$

Façamos agora a observação que o mínimo do polinômio $q(x) = mx^2 + 2a_{m-1}x - (m-2)\big(a_{m-1}\big)^2 + 2(m-1)a_{m-2}$ é negativo. De fato, esse mínimo ocorre no ponto $\xi$ onde $q'(\xi) = 0$, ou seja, em $\xi = -\frac{a_{m-1}}{m}$. Agora,

$$q(\xi) \;=\; (m-1)\Big(2a_{m-2} - \frac{m-1}{m}\big(a_{m-1}\big)^2\Big) \overset{(6.49)}{\leq} 0 \; ,$$

como afirmamos.

Retornando com essa afirmação a (6.58), vemos que $x_1$ deve estar localizada no intervalo fechado entre os dois zeros de $q(x)$, onde esse polinômio é não positivo. Esses zeros são, como trivialmente se verifica,

$$\frac{-a_{m-1} \pm \sqrt{(m-1)^2(a_{m-1})^2 - 2m(m-1)a_{m-2}}}{m} ,$$

que coincidem com $\alpha_m^{\pm}$ apresentados em (6.53). Observe-se que (6.49) garante que esses zeros sejam reais.

Concluímos que $x_1$ encontra-se no intervalo fechado $\left[\alpha_m^-, \ \alpha_m^+\right]$. Como $x_1$ foi escolhido como um zero genérico de $p$ (nenhum ordenamento foi suposto entre os zeros listados como $x_1, \ldots, x_m$), concluímos que os demais zeros devem submeter-se à mesma restrição. □

### 6.3.1.2 As Desigualdades de Samuelson e Algumas Generalizações

A Proposição 6.3, página 331, possui uma leitura muito interessante. Sejam $x_1, \ldots, x_m$ pontos não necessariamente distintos da reta real. Seu valor médio, ou média aritmética, é $\mu_m(x) := \frac{x_1+\cdots+x_m}{m}$ e o desvio padrão desse valor médio é

$$\sigma_m(x) := \sqrt{\frac{1}{m}\sum_{k=1}^{m}\big(x_k - \mu_m(x)\big)^2} = \sqrt{\frac{1}{m}\sum_{k=1}^{m}(x_k)^2 - \big(\mu_m(x)\big)^2} .$$

Suponhamos que não conheçamos os valores de $x_1, \ldots, x_m$, mas saibamos apenas os valores de $m$, de $\mu_m(x)$ e de $\sigma_m(x)$. Podemos apenas com essas informações determinar uma região onde o conjunto se encontra? Talvez um tanto surpreendentemente a resposta é sim e afirmamos que

$$\big\{x_1, \ldots, x_m\big\} \subset \Big[\mu_m(x) - \sqrt{m-1}\,\sigma_m(x), \ \ \mu_m(x) + \sqrt{m-1}\,\sigma_m(x)\Big] . \tag{6.59}$$

Essa afirmação é denominada *desigualdades de Samuelson*[34][35]. Ela afirma que qualquer conjunto finito de $m$ pontos está contido no intervalo fechado demarcado por seu valor médio mais ou menos $\sqrt{m-1}$ vezes seu desvio padrão.

Para prová-la, consideremos o polinômio $p(x) = (x-x_1)\cdots(x-x_m) = x^m + a_{m-1}x^{m-1} + \cdots + a_1x + a_0$. Os coeficientes $a_k$ são dados pelas fórmulas de Viète (6.47) e a Proposição 6.3, página 331, informa que $\{x_1, \ldots, x_m\} \subset \left[\alpha_m^-, \ \alpha_m^+\right]$ com $\alpha_m^{\pm}$ dados em (6.53). Agora, por (6.53) e pelas duas primeiras fórmulas em (6.47),

$$\alpha_m^{\pm} = \frac{x_1+\cdots+x_m}{m} \pm \sqrt{m-1}\sqrt{(m-1)\left(\frac{x_1+\cdots+x_m}{m}\right)^2 - \frac{2}{m}\sum_{1\leq i<j\leq m} x_ix_j} . \tag{6.60}$$

Porém,

$$(m-1)\left(\frac{x_1+\cdots+x_m}{m}\right)^2 = \frac{(m-1)}{m^2}\left(\sum_{k=1}^{n}(x_k)^2 + 2\sum_{1\leq i<j\leq m} x_ix_j\right)$$

e, portanto, o termo dentro da raiz quadrada fica

$$\frac{(m-1)}{m^2}\sum_{k=1}^{n}(x_k)^2 - \frac{2}{m^2}\sum_{1\leq i<j\leq m} x_ix_j = \frac{1}{m}\sum_{k=1}^{n}(x_k)^2 - \frac{1}{m^2}\left(\sum_{k=1}^{n}(x_k)^2 + \sum_{1\leq i<j\leq m} x_ix_j\right)$$

$$= \frac{1}{m}\sum_{k=1}^{n}(x_k)^2 - \frac{\big(x_1+\cdots+x_m\big)^2}{m^2} = \frac{1}{m}\sum_{k=1}^{n}(x_k)^2 - \big(\mu_m(x)\big)^2 = \sigma_m(x)^2 .$$

[34]Paul Anthony Samuelson (1915–2009). Samuelson foi um destacado economista e ganhador do "Prêmio de Ciências Econômicas em Memória de Alfred Nobel" de 1970, incorretamente conhecido como Prêmio Nobel de Economia.

[35]O artigo original é P. A. Samuelson, "How deviant can you be?", J. Amer. Statist. Assoc. **63**, 1522–1525 (1968).



Assim, retornando a (6.60), temos $\alpha_m^{\pm} = \mu_m(x) \pm \sqrt{m-1}\,\sigma_m(x)$, como desejávamos estabelecer.

Comentemos que o resultado final não avança além do de Laguerre, mas é interpretado em termos das quantidades estatísticas $\mu_m(x)$ e $\sigma_m(x)$. Essa foi a contribuição de Samuelson ao resultado[36] e que escapara a Laguerre.

As desigualdades de Samuelson e suas generalizações (vide adiante) podem ser aplicadas ao estudo da localização do espectro de matrizes autoadjuntas, assunto apresentado na Seção 10.5.3, página 524.

- **As desigualdades de Samuelson. Uma segunda demonstração**

As desigualdades de Samuelson podem ser demonstradas com recursos ainda mais simples: elas são uma decorrência da desigualdade de Cauchy-Schwarz, aplicada de forma conveniente. A prova que apresentamos aqui provém de: Barry C. Arnold, "Schwarz, Regression, and Extreme Deviance". The American Statistician, **28**, n. 1, pp. 22–23 (1974).

Empregando a mesma notação de antes, sejam os vetores $\mathbf{x} \in \mathbb{R}^{m-1}$ e $\mathbf{u} \in \mathbb{R}^{m-1}$ definidos por

$$\mathbf{x} := \begin{pmatrix} x_2 - \mu_m(x) \\ \vdots \\ x_m - \mu_m(x) \end{pmatrix} \qquad \text{e} \qquad \mathbf{u} := \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} . \tag{6.61}$$

Observe-se que a componente $x_1 - \mu_m(x)$ foi excluída de $\mathbf{x}$, resultando neste vetor possuir $m-1$ componentes, como também $\mathbf{u}$.

Pela desigualdade de Cauchy-Schwarz, temos $\langle \mathbf{x}, \mathbf{u}\rangle^2 \leq \langle \mathbf{x}, \mathbf{x}\rangle\langle \mathbf{u}, \mathbf{u}\rangle$, ou seja

$$\left(\sum_{j=2}^{m}\big(x_j - \mu_m(x)\big)\right)^2 \leq \left(\sum_{j=2}^{m}\big(x_j - \mu_m(x)\big)^2\right)\left(\sum_{l=2}^{m}1^2\right) = (m-1)\sum_{j=2}^{m}\big(x_j - \mu_m(x)\big)^2 .$$

Somando e subtraindo $x_1 - \mu_m(x)$ e $\big(x_1 - \mu_m(x)\big)^2$ às somas em $j$ do lado direito e esquerdo, respectivamente, obtemos

$$\left(\sum_{j=1}^{m}\big(x_j - \mu_m(x)\big) - \big(x_1 - \mu_m(x)\big)\right)^2 \leq (m-1)\sum_{j=1}^{m}\big(x_j - \mu_m(x)\big)^2 - (m-1)\big(x_1 - \mu_m(x)\big)^2 . \tag{6.62}$$

Pela definição de $\mu_m(x)$, tem-se $\sum_{j=1}^{m}\big(x_j - \mu_m(x)\big) = 0$. Assim, (6.62) fica

$$m\big(x_1 - \mu_m(x)\big)^2 \leq (m-1)\sum_{j=1}^{m}\big(x_j - \mu_m(x)\big)^2 ,$$

ou seja $\big(x_1 - \mu_m(x)\big)^2 \leq (m-1)\sigma_m(x)^2$. Com isso, vemos que

$$\mu_m(x) - \sqrt{m-1}\,\sigma_m(x) \leq x_1 \leq \mu_m(x) + \sqrt{m-1}\,\sigma_m(x) .$$

A escolha de $x_1$ foi arbitrária e, portanto, a desigualdade acima permanece válida se $x_1$ for substituído por qualquer $x_j$, $j = 1, \ldots, m$. Dessa forma, obtemos novamente a desigualdade de Samuelson (6.59).

- **Estendendo as desigualdades de Samuelson**

A demonstração que apresentamos logo acima inspira uma extensão das desigualdades de Samuelson, a saber, utilizando-se a desigualdade de Hölder na forma (24.43), página 1282, em lugar da de Cauchy-Schwarz. Com um

[36]Para extensões desse resultado, vide: Henry Wolkowicz and George P. H. Styan, "Extensions of Samuelson's Inequality", The American Statistician, **33**, no. 3, 143-144, (1979). DOI 10.1080/00031305.1979.10482683.

pouco mais de generalidade consideraremos $m \in \mathbb{N}$ números complexos não necessariamente distintos $x_1, \ldots, x_m \in \mathbb{C}$ cuja média aritmética é $\mu_m(x) := \frac{1}{m}\big(x_1 + \cdots + x_m\big)$. Definimos $\mathbf{x} \in \mathbb{C}^{m-1}$ e $\mathbf{u} \in \mathbb{C}^{m-1}$ por

$$\mathbf{x} := \begin{pmatrix} x_2 - \mu_m(x) \\ \vdots \\ x_m - \mu_m(x) \end{pmatrix} \qquad \text{e} \qquad \mathbf{u} := \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} . \tag{6.63}$$

Observe-se que a componente $x_1 - \mu_m(x)$ foi excluída de $\mathbf{x}$, resultando neste vetor possuir $m-1$ componentes, como também $\mathbf{u}$. Seja também $p \in \mathbb{R}$ com $p > 1$ e seja $q \in \mathbb{R}$ definido por $\frac{1}{p} + \frac{1}{q} = 1$. Temos, pela desigualdade de Hölder (24.43), página 1282,

$$\begin{aligned} \Big|\big\langle \mathbf{u}, \, \mathbf{x}\big\rangle\Big| \;=\; \left|\sum_{j=2}^m \big(x_j - \mu_m(x)\big)\cdot 1\right| \;\leq\; \sum_{j=2}^m \big|x_j - \mu_m(x)\big|\cdot 1 \;&\overset{(24.43)}{\leq}\; \left[\sum_{j=2}^m \big|x_j - \mu_m(x)\big|^p\right]^{1/p} \left[\sum_{j=2}^m 1^q\right]^{1/q} \\ &=\; (m-1)^{1/q}\left[\sum_{j=2}^m \big|x_j - \mu_m(x)\big|^p\right]^{1/p} \end{aligned} \tag{6.64}$$

Assim,

$$\left|\sum_{j=2}^m \big(x_j - \mu_m(x)\big)\right| \;\leq\; (m-1)^{1/q}\left[\sum_{j=2}^m \big|x_j - \mu_m(x)\big|^p\right]^{1/p} .$$

Somando e subtraindo $x_1 - \mu_m(x)$ e $\big|x_1 - \mu_m(x)\big|^p$ às somas em $j$ do lado direito e esquerdo, respectivamente, e usando o fato que $\sum_{j=1}^m \big(x_j - \mu_m(x)\big) = 0$, temos

$$\big|x_1 - \mu_m(x)\big| \;\leq\; (m-1)^{1/q}\left[\left(\sum_{j=1}^m \big|x_j - \mu_m(x)\big|^p\right) - \big|x_1 - \mu_m(x)\big|^p\right]^{1/p} .$$

Com isso,

$$\big|x_1 - \mu_m(x)\big|^p \;\leq\; (m-1)^{p/q}\left[\left(\sum_{j=1}^m \big|x_j - \mu_m(x)\big|^p\right) - \big|x_1 - \mu_m(x)\big|^p\right] ,$$

o que implica

$$\big|x_1 - \mu_m(x)\big| \;\leq\; \left[\frac{(m-1)^{p/q}}{1+(m-1)^{p/q}}\sum_{j=1}^m \big|x_j - \mu_m(x)\big|^p\right]^{1/p} . \tag{6.65}$$

É fácil verificar que $\frac{(m-1)^{p/q}}{1+(m-1)^{p/q}} = \frac{1}{1+(m-1)^{1-p}}$, pois $p/q = p-1$.

Definamos,

$$\nu_p \;\equiv\; \nu_p\big(x_1, \ldots, x_m\big) \;:=\; \frac{1}{m}\sum_{j=1}^m \big|x_j - \mu_m(x)\big|^p$$

e

$$f_p(m) \;:=\; \left[\frac{m}{1+(m-1)^{1-p}}\right]^{1/p} .$$

Observe-se que $f_p(m)$ é, para cada $p$, crescente em $m$ para $m \geq 1$, aproximando-se de $m^{1/p}$ para $m$ "grande". Além disso, $f_p(m)$ é, para cada $m$, descescente em $p$, convergindo a 1 para $p \to \infty$.



Com isso, podemos escrever (6.65) como

$$\big|x_1 - \mu_m(x)\big| \;\leq\; f_p(m)\Big(\nu_p\big(x_1, \ldots, x_m\big)\Big)^{1/p} . \tag{6.66}$$

As expressões acima foram obtidas para todo $p \in \mathbb{R}$ com $p > 1$.

Como a escolha de $x_1$ foi arbitrária, concluímos que (6.65) é válida para todo $x_1, \ldots, x_m$ e, portanto, temos

$$\big|x_k - \mu_m(x)\big| \;\leq\; f_p(m)\Big(\nu_p\big(x_1, \ldots, x_m\big)\Big)^{1/p} \tag{6.67}$$

para todo $k = 1, \ldots, m$ e todo $p > 1$. Essa é uma versão generalizada das desigualdades de Samuelson, pois implica

$$\big\{x_1, \ldots, x_m\big\} \;\subset\; \overline{D}\Big(\mu_m(x), \; f_p(m)\big(\nu_p\big)^{1/p}\Big) ,$$

o disco fechado em $\mathbb{C}$, de raio $f_p(m)\big(\nu_p\big)^{1/p}$, centrado em $\mu_m(x)$. Para $x_1, \ldots, x_m$ reais, isso significa

$$\big\{x_1, \ldots, x_m\big\} \;\subset\; \Big[\mu_m(x) - f_p(m)\big(\nu_p\big)^{1/p}, \;\; \mu_m(x) + f_p(m)\big(\nu_p\big)^{1/p}\Big] . \tag{6.68}$$

No caso em que $x_1, \ldots, x_m$ são reais e $p$ é um inteiro par, a grandeza $\nu_p \equiv \nu_p\big(x_1, \ldots, x_m\big)$ possui uma interpretação estatística: ela é dita ser o $p$-ésimo *momento central* de $x_1, \ldots, x_m$. O qualificativo "central" se deve ao fato de ser o $p$-ésimo momento associado ao conjunto $x_1 - \mu_m(x), \ldots, x_m - \mu_m(x)$, ou seja, ao $p$-ésimo momento centrado no valor médio $\mu_m(x)$. No caso em que $p$ é um inteiro ímpar essa interpretação como momento central é prejudicada pelo aparecimento dos valores absolutos na soma que define $\nu_p$. Pelo mesmo fato, para o caso geral de $p$ real não inteiro, essa interpretação também não está disponível.

Como facilmente se verifica, para $p = 2$ temos $f_2(m) = \sqrt{m-1}$ e $\sqrt{\mu_2} = \sigma_m(x)$, o desvio padrao de $x_1, \ldots, x_m$. Assim, para $p = 2$, (6.67) e (6.68) recuperam as desigualdades de Samuelson originais.

É também interessante analisar o comportamento de $f_p(m)$ e de $\nu_p\big(x_1, \ldots, x_m\big)^{1/p}$ quando $p$ e $m$ são "grandes".

1. Para $m$ fixo e $p \to \infty$ temos $f_p(m) \to 1$. Além disso, nesse mesmo limite, a quantidade $\nu_p\big(x_1, \ldots, x_m\big)^{1/p}$ converge a $\max\big\{|x_1 - \mu_m(x)|, \ldots, |x_m - \mu_m(x)|\big\}$. Logo, nesse limite, (6.67) torna-se

$$\big|x_k - \mu_m(x)\big| \;\leq\; \max\Big\{\big|x_1 - \mu_m(x)\big|, \ldots, \big|x_m - \mu_m(x)\big|\Big\} , \qquad k = 1, \ldots, m, \tag{6.69}$$

   que é um resultado ótimo e que nos garante (no caso em que $x_1, \ldots, x_m$ são reais) que

$$\big\{x_1, \ldots, x_m\big\} \;\subset\; \Big[\mu_m(x) - \big|x_{k_*} - \mu_m(x)\big|, \;\; \mu_m(x) + \big|x_{k_*} - \mu_m(x)\big|\Big] ,$$

   onde $x_{k_*}$ é um elemento de $\{x_1, \ldots, x_m\}$ para o qual $\big|x_k - \mu_m(x)\big|$ atinge o máximo. Assim. Se $x_{k_*} \geq \mu_m(x)$, temos

$$\big\{x_1, \ldots, x_m\big\} \;\subset\; \begin{cases} \Big[2\mu_m(x) - x_{k_*}, \;\; x_{k_*}\Big] , & \text{caso } x_{k_*} \geq \mu_m(x) , \\ \Big[x_{k_*}, \;\; 2\mu_m(x) - x_{k_*}\Big] , & \text{caso } x_{k_*} \leq \mu_m(x) . \end{cases} \tag{6.70}$$

2. Para $p$ fixo e $m \gg 1$, $f_p(m)$ aproxima-se de $m^{1/p}$ de forma crescente. Assim, por (6.67) podemos escrever

$$\big|x_k - \mu_m(x)\big| \;\leq\; \left[\sum_{j=1}^m \big|x_j - \mu_m(x)\big|^p\right]^{1/p} \tag{6.71}$$

   para todo $k = 1, \ldots, m$, uma desigualdade trivial por ser a função $\mathbb{R}_+ \ni x \mapsto x^{1/p}$ uma função crescente.

   Se agora assumirmos que a sequência $\{x_m\}_{m\in\mathbb{N}}$ pertença a $\ell_p(\mathbb{N})$, $p > 1$, teremos pelo Exercício E. 24.60, página 1298, especificamente, por (24.69), que $\lim_{m\to\infty} \mu_m(x) = 0$. Com essas informações, podemos escrever de (6.71), $\big|x_k\big| \leq \Big[\sum_{j=1}^\infty \big|x_j\big|^p\Big]^{1/p} = \|x\|_p$, a norma $p$ da sequência $\{x_m\}_{n\in\mathbb{N}}$. Portanto,

$$\big\{x_j, \; j \in \mathbb{N}\big\} \;\subset\; \overline{D}\big(0, \; \|x\|_p\big) ,$$

o disco fechado em $\mathbb{C}$, de raio $\|x\|_p$, centrado na origem, ou, no caso real,

$$\{x_j,\ j \in \mathbb{N}\} \subset \big[-\|x\|_p,\ \|x\|_p\big] .$$

Pelo Teorema 24.5, página 1289, sabemos que as normas $\|\cdot\|_p$ são decrescentes em $p$: $\|x\|_{p'} \leq \|x\|_p$ se $p' > p$ (lembrar também que $\ell_p(\mathbb{N}) \subset \ell_{p'}(\mathbb{N})$ se $p' > p$, pelo Exercício E. 24.48, página 1280). Assim, a relação $|x_k| \leq \|x\|_p$ torna-se mais restritiva à medida que $p$ cresce. Também sabemos, pela Proposição 24.5, página 1289, que $\lim\limits_{p\to\infty} \|x\|_p = \|x\|_\infty$ e concluímos que para todo $k$, tem-se

$$|x_k| \leq \|x\|_\infty , \qquad \text{ou seja,} \qquad \{x_j,\ j \in \mathbb{N}\} \subset \overline{D}\big(0,\ \|x\|_\infty\big) , \tag{6.72}$$

novamente uma relação ótima.

As relações (6.69)–(6.70) e (6.72) não são surpreendentes, mas a possibilidade de obtê-las a partir de (6.67)–(6.68) ilustra a utilidade daquelas relações.

#### 6.3.1.3 As Identidades de Girard-Newton

Aproveitando nossa discussão sobre as fórmulas de Viète (6.46), (6.47) e (6.48), página 330, vamos apresentar um conjunto de identidades úteis envolvendo certos polinômios simétricos em várias variáveis (reais ou complexas), denominadas *identidades de Girard-Newton*. Essas identidades foram descobertas por Girard[37] em 1629 e redescobertas por Newton[38], sendo publicadas em 1707[39].

Girard e Newton (que aparentemente não conhecia o trabalho anterior de Girard) não consideraram o caso geral, nem apresentaram provas das relações que obtiveram. Newton limitou-se ao caso $m = 3$ e $k = 3$ (em nossa notação, vide abaixo), o que corresponde a equações cúbicas em três variáveis. O caso geral, do qual trataremos abaixo, é também denominado *Teorema de Newton* na literatura (por exemplo, em [112]). Desconhecemos o autor da primeira demonstração do caso geral.

As identidades de Girard-Newton têm aplicações em Álgebra, como na Teoria de Galois (vide, *e.g.*, [112] ou [276]), no estudo do Grupo de Permutações, em Estatística e Álgebra Linear, através de generalizações das desigualdades de Samuelson (vide comentário e referências adiante) e encontraram uso na obtenção de certas soluções exatas na Teoria da Relatividade Geral, as chamadas *soluções de fluido*. Vide para tal [393], um texto especializado em soluções exatas da Teoria da Relatividade Geral, ou o capítulo IX de [76]. Soluções de fluido são empregadas em modelos estelares, em Astrofísica, assim como em modelos cosmológicos.

• **Definições prévias, enunciado principal e sua demonstração**

Comecemos com a definição de dois conjuntos de funções em $m$ variáveis que nos serão úteis.

Seja $m \in \mathbb{N}$ e para $(x_1,\ \ldots,\ x_m) \in \mathbb{C}^m$, defina-se

$$\pi_k\big(x_1,\ \ldots,\ x_m\big) := \sum_{j=1}^{m} (x_j)^k = (x_1)^k + \cdots + (x_m)^k , \tag{6.73}$$

sendo $k \in \mathbb{N}$. Defina-se também, para $k \in \mathbb{N}_0$,

$$\zeta_k\big(x_1,\ \ldots,\ x_m\big) := \begin{cases} 1 , & \text{se} \quad k = 0 , \\ \displaystyle\sum_{1\leq i_1 < \cdots < i_k \leq m} x_{i_1}\cdots x_{i_k} , & \text{se} \quad 0 < k \leq m , \\ 0 , & \text{se} \quad k > m . \end{cases} \tag{6.74}$$

[37] Albert Girard (1595–1632).
[38] Isaac Newton (1642–1726).
[39] Esses resultados foram apresentados em um livro-texto contendo notas de aula de Newton sobre Álgebra e Aritmética, denominado *Arithmetica Universalis*, publicado em 1707 sem o consentimento de Newton, por um certo William Whiston (1667–1752). Historiadores acreditam que Newton já obtivera seus resultados em 1669 ou mesmo em 1666, seu *"Annus mirabilis"*.



Note-se, em particular, que

$$\zeta_1\big(x_1,\ \ldots,\ x_m\big) = x_1 + \cdots + x_m \qquad \text{e} \qquad \zeta_m\big(x_1,\ \ldots,\ x_m\big) = x_1\cdots x_m .$$

Observe-se também que, pela fórmula de Viète em (6.48), tem-se que para $0 < k \leq m$ que

$$\zeta_k\big(x_1,\ \ldots,\ x_n\big) = (-1)^k a_{m-k} , \tag{6.75}$$

onde $a_j$ são os coeficientes do polinômio $(x - x_1)\cdots(x - x_m)$ na variável $x$.

Com isso, podemos agora apresentar e demonstrar as identidades anunciadas:

**Proposição 6.4 (Identidades de Girard-Newton)** *Para todo $m \in \mathbb{N}$, tem-se com as definições acima,*

$$\sum_{i=1}^{k} (-1)^{i-1}\zeta_{k-i}\big(x_1,\ \ldots,\ x_m\big)\pi_i\big(x_1,\ \ldots,\ x_m\big) = k\zeta_k\big(x_1,\ \ldots,\ x_m\big) , \quad \textit{para } 1 \leq k \leq m , \tag{6.76}$$

$$\sum_{i=k-m}^{k} (-1)^{i-1}\zeta_{k-i}\big(x_1,\ \ldots,\ x_m\big)\pi_i\big(x_1,\ \ldots,\ x_m\big) = 0 , \quad \textit{para } k > m . \tag{6.77}$$

*Acima, $\big(x_1,\ \ldots,\ x_m\big) \in \mathbb{C}^m$.* □

Observe-se que a forma das relações independe do número $m$ de variáveis. Após a demonstração exibiremos algumas consequências de (6.76)-(6.77).

Demonstração da Proposição 6.4. Fixemos $m \in \mathbb{N}$. A demonstração é feita em três etapas: a prova de (6.76) no caso $k = m$; a prova de (6.77) para $k > m$ e, finalmente, a prova de (6.76) no caso $1 \leq k < m$.

Etapa 1. Começamos provando (6.76) no caso $k = m$. Considere-se o polinômio mônico de grau $m$ na variável $x$ dado por $(x - x_1)\cdots(x - x_m)$. Sabemos pelas fórmulas de Viète e por (6.75) que

$$(x - x_1)\cdots(x - x_m) = a_0 + \sum_{i=1}^{m-1} a_i x^i + x^m = (-1)^m\zeta_m\big(x_1,\ \ldots,\ x_m\big) + \sum_{i=1}^{m-1}(-1)^{m-i}\zeta_{m-i}\big(x_1,\ \ldots,\ x_m\big)x^i + x^m .$$

Se agora escolhermos $x = x_j$ para algum $j = 1,\ \ldots,\ m$, ficamos com

$$0 = (-1)^m\zeta_m\big(x_1,\ \ldots,\ x_m\big) + \sum_{i=0}^{m-1}(-1)^{m-i}\zeta_{m-i}\big(x_1,\ \ldots,\ x_m\big)(x_j)^i + (x_j)^m .$$

Somando-se essa igualdade para todo $j = 1,\ \ldots,\ m$, obtemos

$$0 = (-1)^m m\zeta_m\big(x_1,\ \ldots,\ x_m\big) + \sum_{i=0}^{m-1}(-1)^{m-i}\zeta_{m-i}\big(x_1,\ \ldots,\ x_m\big)\pi_i\big(x_1,\ \ldots,\ x_m\big) + \pi_m\big(x_1,\ \ldots,\ x_m\big) .$$

Lembrando da convenção que $\zeta_0\big(x_1,\ \ldots,\ x_m\big) = 1$, passando o termo $(-1)^m m\zeta_m\big(x_1,\ \ldots,\ x_m\big)$ para o lado esquerdo e cancelando os fatores $(-1)^m$, extraímos disso que

$$m\zeta_m\big(x_1,\ \ldots,\ x_m\big) = \sum_{i=0}^{m}(-1)^{i-1}\zeta_{m-i}\big(x_1,\ \ldots,\ x_m\big)\pi_i\big(x_1,\ \ldots,\ x_m\big) ,$$

expressão essa que é idêntica a (6.76) no caso $k = m$.

Etapa 2. Provemos agora as relações (6.77), quando $k > m$. Adicionemos às $m$ variáveis $x_1,\ \ldots,\ x_m$ novas variáveis $x_{m+1},\ \ldots,\ x_k$. Devido à validade de (6.76) no caso $k = m$, tem-se

$$k\zeta_k\big(x_1,\ \ldots,\ x_m,\ x_{m+1},\ \ldots,\ x_k\big) = \sum_{i=0}^{k}(-1)^{i-1}\zeta_{k-i}\big(x_1,\ \ldots,\ x_m,\ x_{m+1},\ \ldots,\ x_k\big)\pi_i\big(x_1,\ \ldots,\ x_m,\ x_{m+1},\ \ldots,\ x_k\big) . \tag{6.78}$$

Façamos agora $x_{m+1} = \cdots = x_k = 0$. Teremos, para as funções $\pi_i$,

$$\pi_i\big(x_1, \ldots, x_m, x_{m+1}, \ldots, x_k\big) = \pi_i\big(x_1, \ldots, x_m\big) .$$

Agora, para as funções $\zeta_j$,

$$\zeta_j\big(x_1, \ldots, x_m, x_{m+1}, \ldots, x_k\big) = \sum_{1\leq i_1<\ldots<i_j\leq k} x_{i_1}\cdots x_{i_j}$$

e, portanto, se $j > m$ haverá sempre ao menos um fator do conjunto $\{x_{m+1}, \ldots, x_k\}$ no lado direito, fazendo com que o todo anule-se quando $x_{m+1} = \cdots = x_k = 0$.

Por outro lado, se $1 \leq j \leq m$, temos

$$\zeta_j\big(x_1, \ldots, x_m, 0, \ldots, 0\big) = \sum_{1\leq i_1<\ldots<i_j\leq m} x_{i_1}\cdots x_{i_j} = \zeta_j\big(x_1, \ldots, x_m\big) .$$

Em resumo,

$$\zeta_j\big(x_1, \ldots, x_m, 0, \ldots, 0\big) = \begin{cases} \zeta_j\big(x_1, \ldots, x_m\big) , & \text{se } 1 \leq j \leq m , \\ 0 , & \text{se } j > m . \end{cases} \tag{6.79}$$

Em particular $\zeta_k\big(x_1, \ldots, x_m, 0, \ldots, 0\big) = 0$, posto que $k > m$.

Assim, a relação (6.78) para quando $x_{m+1} = \cdots = x_k = 0$ assume a forma

$$0 = \sum_{i=k-m}^{k} (-1)^{i-1}\zeta_{k-i}\big(x_1, \ldots, x_m\big)\pi_i\big(x_1, \ldots, x_m\big) . \tag{6.80}$$

Essa é precisamente a relação (6.77), válida quando $k > m$.

Etapa 3. Resta-nos provar (6.76) para $1 \leq k < m$. Desejamos provar que o polinômio nas variáveis $x_1, \ldots, x_m$ definido por

$$Q\big(x_1, \ldots, x_m\big) := \sum_{i=1}^{k}(-1)^{i-1}\zeta_{k-i}\big(x_1, \ldots, x_m\big)\pi_i\big(x_1, \ldots, x_m\big) - k\zeta_k\big(x_1, \ldots, x_m\big) \tag{6.81}$$

é identicamente nulo para $1 \leq k < m$. Observemos para tal que se $x_{k+1} = \cdots = x_m = 0$ teremos

$$Q\big(x_1, \ldots, x_k, 0, \ldots, 0\big) := \sum_{i=1}^{k}(-1)^{i-1}\zeta_{k-i}\big(x_1, \ldots, x_k, 0, \ldots, 0\big)\pi_i\big(x_1, \ldots, x_k, 0, \ldots, 0\big)$$
$$- k\zeta_k\big(x_1, \ldots, x_k, 0, \ldots, 0\big) .$$

Agora, $\pi_i\big(x_1, \ldots, x_k, 0, \ldots, 0\big) = \pi_i\big(x_1, \ldots, x_k\big)$ e $\zeta_j\big(x_1, \ldots, x_k, 0, \ldots, 0\big) = \zeta_j\big(x_1, \ldots, x_k\big)$ para todo $1 \leq j \leq k$, pelo mesmo raciocínio que levou a (6.79) (basta usar (6.79) substituindo-se a letra $m$ por $k$). Assim,

$$Q\big(x_1, \ldots, x_k, 0, \ldots, 0\big) := \sum_{i=1}^{k}(-1)^{i-1}\zeta_{k-i}\big(x_1, \ldots, x_k\big)\pi_i\big(x_1, \ldots, x_k\big) - k\zeta_k\big(x_1, \ldots, x_k\big) = 0$$

pela relação (6.76) para $k = m$, que já foi demonstrada na etapa 1.

Como $Q$ é um polinômio simétrico nas suas $m$ variáveis, segue disso que $Q\big(x_1, \ldots, x_m\big) = 0$ sempre que $m - k$ de suas variáveis forem nulas.

Agora, o polinômio $Q\big(x_1, \ldots, x_m\big)$ é uma soma de monômios, cada qual envolvendo no máximo $k$ variáveis distintas[40]. Assim, $Q\big(x_1, \ldots, x_k, 0, \ldots, 0\big)$ contém todos os monômios nas variáveis $x_1, \ldots, x_k$. O fato de

[40]Pela definição de $Q$ em (6.81) e pelo fato de as funções $\zeta_k$, assim como os produtos $\zeta_{k-1}\pi_i$, $i \leq i \leq k$, terem essa propriedade.



$Q\big(x_1, \ldots, x_k, 0, \ldots, 0\big)$ ser nulo significa que os coeficientes desses monômios nas variáveis $x_1, \ldots, x_k$ são identicamente nulos. Como $Q$ é um polinômio simétrico nas suas $m$ variáveis, concluímos que os coeficientes de todos os seus monômios são nulos, o que implica que $Q$ é identicamente nulo. ■

*Comentário sobre a demonstração.* Nossa demonstração adapta ideias próprias a demonstrações elementares previamente conhecidas (elementares por não evocarem recursos especiais ou resultados prévios que não sejam as fórmulas de Viète).

Há diversas outras demonstrações distintas das identidades de Girard-Newton. Uma das mais curtas (uma página!) sendo: Doron Zeilberger, "A Combinatorial Proof of Newton's Identities", Discrete Mathematics, **49**, p. 319 (1984). Para outra demonstração distinta, fazendo uso de Álgebra Linear (particularmente, do Teorema de Hamilton-Cayley, Teorema 10.4, página 489), vide Dan Kalman, "A Matrix Proof of Newton Identities", Mathematics Magazine, **73**, 4, pp. 313–315 (2000). Para outra demonstração simples fazendo uso de transformadas de Laplace, vide Mircea I. Cîrnu, "Newton's Identities and the Laplace Transform", The American Mathematical Monthy, **117**, No. 1 (January 2010), pp. 67–71, ou https://doi.org/10.4169/000298910x474998 ou https://www.jstor.org/stable/10.4169/000298910x474998.

Outras demonstrações fazem uso de funções geratrizes ou de propriedades dos números de Bell, estudados na Seção 6.1.4, página 321. ♣

### • Alguns exemplos

Para termos uma melhor percepção do que afirmam as identidades de Girard-Newton vale escrevê-las explicitamente em alguns casos. Para $m = 2$ e $k = 1, 2, 3$, por exemplo, teremos,

$$(m = 2,\ k = 1): \qquad \zeta_0(x_1, x_2)\pi_1(x_1, x_2) = \zeta_1(x_1, x_2) ,$$
$$\text{ou seja,} \quad 1 \cdot (x_1 + x_2) = x_1 + x_2 .$$
$$(m = 2,\ k = 2): \qquad \zeta_1(x_1, x_2)\pi_1(x_1, x_2) - \zeta_0(x_1, x_2)\pi_2(x_1, x_2) = 2\zeta_2(x_1, x_2) ,$$
$$\text{ou seja,} \quad (x_1 + x_2)(x_1 + x_2) - \big((x_1)^2 + (x_2)^2\big) = 2x_1x_2 .$$
$$(m = 2,\ k = 3): \qquad \zeta_2(x_1, x_2)\pi_1(x_1, x_2) - \zeta_1(x_1, x_2)\pi_2(x_1, x_2) + \zeta_0(x_1, x_2)\pi_3(x_1, x_2) = 0 ,$$
$$\text{ou seja,} \quad x_1x_2(x_1 + x_2) - (x_1 + x_2)\big((x_1)^2 + (x_2)^2\big) + \big((x_1)^3 + (x_2)^3\big) = 0 .$$

É trivial verificar diretamente a validade das identidades acima.

### • Outras consequências

No caso geral, as identidades de Girard-Newton admitem várias leituras. Por exemplo, as relações (6.76) podem ser usadas para fornecer recursivamente $\zeta_k$ em termos de $\zeta_j$, $1 \leq j < k$ e de $\pi_j$, $1 \leq j \leq k$. Elas também permitem escrever recursivamente $\pi_k$ em termos de $\pi_j$'s e $\zeta_j$'s anteriores. De fato, como $\zeta_0 = 1$, vê-se facilmente de (6.76) e (6.77) que

$$\pi_k\big(x_1, \ldots, x_m\big) = \begin{cases} \displaystyle\sum_{i=1}^{k-1}(-1)^{i-k+1}\zeta_{k-i}\big(x_1, \ldots, x_m\big)\pi_i\big(x_1, \ldots, x_m\big) + k(-1)^{k-1}\zeta_k\big(x_1, \ldots, x_m\big) , & \text{para } 1 \leq k \leq m , \\ \displaystyle\sum_{i=k-m}^{k-1}(-1)^{i-k-1}\zeta_{k-i}\big(x_1, \ldots, x_m\big)\pi_i\big(x_1, \ldots, x_m\big) , & \text{para } k > m . \end{cases} \tag{6.82}$$

Verifique! As identidades de Girard-Newton são frequentemente apresentadas na forma (6.82).

Como $\zeta_1\big(x_1, \ldots, x_m\big) = \pi_1\big(x_1, \ldots, x_m\big)$, é possível escrever $\pi_k$ e $\zeta_k$ recursivamente em termos de somas e produtos envolvendo apenas $\zeta_j$'s ou $\pi_j$'s, respectivamente, com $1 \leq j \leq k$. Por exemplo, tem-se (omitindo-se a dependência com as variáveis $x_1, \ldots, x_m$),

$$\zeta_1 = \pi_1 , \qquad \pi_1 = \zeta_1 ,$$

$$\zeta_2 = \frac{1}{2}\big((\pi_1)^2 - \pi_2\big) , \qquad \pi_2 = (\zeta_1)^2 - 2\zeta_2 ,$$

$$\zeta_3 = \frac{1}{6}(\pi_1)^3 - \frac{1}{2}\pi_1\pi_2 + \frac{1}{3}\pi_3 , \qquad \pi_3 = (\zeta_1)^3 - 2\zeta_1\zeta_2 + 3\zeta_3 .$$

Um teorema geral, por vezes denominado *Teorema fundamental dos polinômios simétricos*, informa que todo polinômio simétrico nas variáveis $x_1, \ldots, x_m$ (ou seja, invariante por permutações arbitrárias dessas variáveis) pode ser expresso como um polinômio nas funções $\zeta_a\big(x_1, \ldots, x_m\big)$, $a = 1, \ldots, m$. Para uma demonstração, vide [276].

Vemos pelas considerações acima que todo polinômio simétrico nas variáveis $x_1, \ldots, x_m$ também pode ser escrito como um polinômio nas funções $\pi_a\big(x_1, \ldots, x_m\big)$, $a = 1, \ldots, m$. Assim, fixando $m \in \mathbb{N}$, mesmo os polinômios simétricos $\pi_k$, com $k > m$, podem ser escritos como polinômios em $\pi_1, \ldots, \pi_m$. Por exemplo, é fácil ver que, caso $m = 2$, temos $\pi_3 = \frac{1}{2}\big[3\pi_1\pi_2 - (\pi_1)^3\big]$.

**E. 6.11** *Exercício*. Verifique! Ainda no caso $m = 2$, obtenha $\pi_4$ e $\pi_5$ em termos de $\pi_1$ e $\pi_2$. ✠

Essas considerações mostram que se $A$ é uma matriz complexa $m \times m$, então $\mathrm{Tr}(A^k)$ com $k > m$ pode ser obtido como um polinômio em $\mathrm{Tr}(A)$, $\mathrm{Tr}(A^2)$, $\ldots$, $\mathrm{Tr}(A^m)$.

**E. 6.12** *Exercício*. Justifique essa afirmação. ✠

- **Obtendo as somas $\pi_k$, $k \in \mathbb{N}$, sem o conhecimento das raízes**

Seja dado um polinômio mônico $p(x) = x^m + \sum_{j=0}^{m-1} a_j x^j$, com raízes $x_1, \ldots, x_m$ sendo desconhecidas, mas sendo os coeficientes $a_j$ dados. Sabemos por (6.75) (uma decorrência das fórmulas de Viète) que $\zeta_k\big(x_1, \ldots, x_n\big) = (-1)^k a_{m-k}$, $k = 0, \ldots, m$ (aqui convencionamos que $a_m = 1$, pois o polinômio é mônico, por hipótese). Assim, as relações (6.82) ficam

$$\pi_k\big(x_1, \ldots, x_m\big) = \begin{cases} -\sum_{i=1}^{k-1} a_{m-k+i}\, \pi_i\big(x_1, \ldots, x_m\big) - k a_{m-k} , & \text{para } 1 \le k \le m , \\[2ex] -\sum_{i=k-m}^{k-1} a_{m-k+i}\, \pi_i\big(x_1, \ldots, x_m\big) , & \text{para } k > m . \end{cases} \tag{6.83}$$

Como $\pi_1\big(x_1, \ldots, x_m\big) = \zeta_1\big(x_1, \ldots, x_m\big) = -a_{m-1}$, as relações (6.83) permitem obter recorrentemente todos os valores de $\pi_k$ para todo $k \in \mathbb{N}$, mesmo <u>sem</u> o conhecimento das raízes $x_1, \ldots, x_m$, o que é um tanto surpreendente!

**E. 6.13** *Exercício*. Seja o polinômio $p(x) = x^2 + bx + c$, com $a_1 = b$ e $a_0 = c$, sendo $m = 2$. Temos $\pi_1 = -b$. Mostre usando (6.83) que $\pi_2 = b^2 - 2c$. Obtenha também $\pi_3 = -b^3 + 3bc$. Obtenha, também por esse método, $\pi_4$ e $\pi_5$. Como as raízes do polinômio são conhecidas (e são $\big(-b \pm \sqrt{b^2 - 4c}\big)/2$), é fácil verificar a validade dos resultados. O estudante há de perceber que a situação é completamente outra caso as raízes não sejam conhecidas, o que geralmente ocorre se o grau do polinômio for $m \ge 5$. Mesmo nesse caso, porém, as somas $\pi_k$ podem ser computadas, pelo método apresentado, para qualquer $m \in \mathbb{N}$ e qualquer $k \in \mathbb{N}$. ✠

**E. 6.14** *Exercício*. Usando as ideias descritas acima, obtenha fórmulas fechadas para a expressão $\sum_{a=1}^{m} a^k$, $k \in \mathbb{N}$. Considere, para tal, o polinômio $(x-1)(x-2)\cdots(x-m)$. Compare com o Exercício E. 6.15, página 342. ✠



## 6.4 Exercícios Adicionais

**E. 6.15** *Exercício*. Prove (usando, por exemplo, indução) que para todo $n \in \mathbb{N}$ vale

$$\sum_{a=1}^{n} a = \frac{n(n+1)}{2} , \qquad \sum_{a=1}^{n} a^2 = \frac{n(n+1)(2n+1)}{6} \qquad \text{e} \qquad \sum_{a=1}^{n} a^3 = \left(\sum_{a=1}^{n} a\right)^2 = \frac{n^2(n+1)^2}{4} . \tag{6.84}$$

A expressão geral para tais somas finitas de potências de números naturais é

$$\sum_{a=1}^{m} a^k = \frac{(m+1)^{k+1}}{k+1} + \sum_{j=1}^{k} \binom{k}{j} \frac{B_j}{k-j+1}(m+1)^{k-j+1} , \tag{6.85}$$

para todos $m, k \in \mathbb{N}$, onde $B_j$ são os números de Bernoulli, definidos e estudados na Seção 6.1.3 319. Prove isso! Para mais identidades do gênero, vide, *e.g.* [155]. ✠

**E. 6.16** *Exercício*. Prove (usando, por exemplo, indução) que

$$\sum_{a=0}^{n} (-1)^a a = (-1)^n \lceil n/2 \rceil , \qquad \sum_{a=0}^{n} (-1)^a (2a+1) = (-1)^n (n+1) , \qquad \sum_{a=0}^{n} (-1)^a (2a) = (-1)^n 2 \lceil n/2 \rceil , \tag{6.86}$$

para todo $n \in \mathbb{N}_0$. As funções $\lfloor \cdot \rfloor$ e $\lceil \cdot \rceil$ estão definidas na página 35. ✠

**E. 6.17** *Exercício-dirigido*. O propósito deste exercício dirigido é demonstrar a importante *fórmula da cotangente de Euler*, também conhecida como *expansão em frações parciais da função cotangente*:

$$\pi \operatorname{cotg}(\pi z) = \frac{1}{z} - \sum_{n=1}^{\infty} \frac{2z}{n^2 - z^2} , \tag{6.87}$$

válida para $z \in \mathbb{C} \setminus \mathbb{Z}$. Estabeleceremos primeiro a relação

$$\pi \operatorname{cotg}(\pi x) = \frac{1}{x} - \sum_{n=1}^{\infty} \frac{2x}{n^2 - x^2} , \tag{6.88}$$

para $x \in \mathbb{R} \setminus \mathbb{Z}$. A expressão (6.88) foi obtida pela primeira vez por Euler em 1749. Seguiremos uma demonstração elementar e elegante devida a Herglotz[41] tal como apresentada em [7], texto esse que, por sua vez, segue Elstrodt[42]. Essa demonstração é elegante por fazer uso de poucos ingredientes. Basicamente usa-se apenas o fato de que ambos os lados de (6.88) são funções contínuas (em $\mathbb{R} \setminus \mathbb{Z}$), são periódicas de período 1, têm as mesmas divergências nos inteiros e, *last but not least*, satisfazem uma mesma relação algébrica, a relação (6.89), abaixo. Para uma outra demonstração de (6.87) usando o Teorema de Mittag-Leffler, vide [234] ou outro bom livro de funções de variável complexa. A relação (6.88) pode também ser provada usando séries de Fourier. Vide Exercício E. 36.23, página 1856. Vide também [128]. Para uma demonstração usando a chamada *representação produto da função seno*, vide Exercício E. 7.5, página 359.

Passemos à prova de Herglotz para (6.87). Defina-se, para $x \in \mathbb{R} \setminus \mathbb{Z}$,

$$f(x) := \pi \operatorname{cotg}(\pi x) \qquad \text{e} \qquad g(x) := \lim_{N \to \infty} g_N(x) ,$$

onde

$$g_N(x) := \sum_{n=-N}^{N} \frac{1}{x+n} = \frac{1}{x} - \sum_{n=1}^{N} \frac{2x}{n^2 - x^2} .$$

Desejamos provar que $f(x) = g(x)$ para todo $x \in \mathbb{R} \setminus \mathbb{Z}$. Isso é feito nos passos indicados no que segue.

**a.** Prove que $g_N(x)$ converge uniformemente para $N \to \infty$ e em qualquer intervalo fechado contido em $\mathbb{R} \setminus \mathbb{Z}$. Sugestões: para $n \ge 2$ e $2n - 1 > x^2$ tem-se $n^2 - x^2 > (n-1)^2 > 0$ e

$$0 < \frac{1}{n^2 - x^2} < \frac{1}{(n-1)^2} .$$

Use o Teste $M$ de Weierstraß[43] (Proposição 36.1, página 1798) e use o teste da comparação por uma integral para mostrar que $\sum_{n=1}^{\infty} \frac{1}{n^2}$ é finita.
Isso estabeleceu que $g$ existe em $\mathbb{R} \setminus \mathbb{Z}$.

[41] Gustav Ferdinand Maria Herglotz (1881–1953).
[42] J. Elstrodt, "Partialbruchzerlegung des Kotangens, Herglotz-Trick und die Weierstraßsche stetige, nirgendsdifferenzierbare Funktion". Math. Semesterberichte **45**, 207–220 (1998).
[43] Karl Theodor Wilhelm Weierstraß (1815–1897).

**b.** Convença-se que $f$ e $g$ são contínuas em $\mathbb{R}\setminus\mathbb{Z}$. Para $g$ isso segue da convergência uniforme provada em **1**.

**c.** Mostre que $f$ e $g$ são periódicas de período $1$. Para $f$ isso é evidente. Para $g$ isso segue de

$$g_N(x+1) \;=\; g_N(x) + \frac{1}{x+N} + \frac{1}{x+1+N} \;,$$

para $x\in\mathbb{R}\setminus\mathbb{Z}$. Prove isso e tome $N\to\infty$ para obter $g(x+1)=g(x)$ para todo $x\in\mathbb{R}\setminus\mathbb{Z}$.

**d.** Mostre que $f$ e $g$ são funções ímpares: $f(-x)=-f(x)$ e $g(-x)=-g(x)$ para todo $x\in\mathbb{R}\setminus\mathbb{Z}$. Novamente isso é evidente para $f$ e para $g$ isso segue do fato que $g_N(-x)=-g_N(x)$ para todo $N$.

**e.** Até aqui só lidamos com propriedades elementares de $f$ e $g$ mas agora vem uma passagem crucial. Mostre que $f$ e $g$ satisfazem

$$f\left(\frac{x}{2}\right)+f\left(\frac{x+1}{2}\right) \;=\; 2f(x) \qquad \text{e} \qquad g\left(\frac{x}{2}\right)+g\left(\frac{x+1}{2}\right) \;=\; 2g(x)\;, \tag{6.89}$$

para todo $x\in\mathbb{R}\setminus\mathbb{Z}$. Note que se trata da mesma relação algébrica para $f$ e $g$. Para $f$ isso segue das bem conhecidas fórmulas de adição das funções seno e cosseno. Mostre isso. Para $g$ isso segue da identidade

$$g_N\left(\frac{x}{2}\right)+g_N\left(\frac{x+1}{2}\right) \;=\; 2g_{2N}(x)+\frac{2}{x+2N+1}\;.$$

Prove-a usando a relação trivial

$$\frac{1}{\frac{x}{2}+n}+\frac{1}{\frac{x+1}{2}+n} \;=\; \frac{2}{x+2n}+\frac{2}{x+1+2n}$$

e tome o limite $N\to\infty$.

**f.** Defina a função $h(x):=f(x)-g(x)$ (que desejamos provar ser identicamente nula). Note em primeiro lugar que $h$ é uma função ímpar, contínua e periódica de período $1$ em $\mathbb{R}\setminus\mathbb{Z}$, pois $f$ e $g$ o são.

**g.** Mostre, usando, por exemplo, a regra de l'Hospital[44], que

$$\lim_{x\to 0}\left(\pi\operatorname{cotg}(\pi x)-\frac{1}{x}\right) \;=\; \lim_{x\to 0}\left(\frac{\pi x\cos(\pi x)-\operatorname{sen}(\pi x)}{x\operatorname{sen}(\pi x)}\right) \;=\; 0\;.$$

**h.** Mostre que fato provado em **7** implica $\lim\limits_{x\to 0}h(x)=0$. Como $h$ é periódica de período $1$, isso significa que $\lim\limits_{x\to n}h(x)=0$ para todo $n\in\mathbb{Z}$. Definindo $h(n)=0$ para todo $n\in\mathbb{Z}$, essa propriedade, por sua vez, implica que a função $h$ torna-se contínua e periódica de período $1$ em todo $\mathbb{R}$, não apenas em $\mathbb{R}\setminus\mathbb{Z}$.

**i.** Como $h$ é contínua e periódica em todo $\mathbb{R}$, $h$ possui um máximo, que denotaremos por $H$. Seja $x_0$ um ponto de $\mathbb{R}$ tal que $h(x_0)=H$ (que um tal ponto existe segue da continuidade e periodicidade de $h$). Agora, tem-se por (6.89) que

$$h\left(\frac{x_0}{2}\right)+h\left(\frac{x_0+1}{2}\right) \;=\; f\left(\frac{x_0}{2}\right)+f\left(\frac{x_0+1}{2}\right)-g\left(\frac{x_0}{2}\right)-g\left(\frac{x_0+1}{2}\right)$$

$$\stackrel{(6.89)}{=}\; 2f(x_0)-2g(x_0) \;=\; 2h(x_0) \;=\; 2H\;.$$

Isso está dizendo que a soma de $h\left(\frac{x_0}{2}\right)$ e $h\left(\frac{x_0+1}{2}\right)$ é duas vezes o máximo valor alcançado por $h$ em toda $\mathbb{R}$. Ora, isso só é possível se ambos os termos forem iguais a $H$, pois se um fosse menor que $H$ o outro teria que ser maior que $H$, o que não é possível. Assim, concluímos que $h\left(\frac{x_0}{2}\right)=H$ (e que $h\left(\frac{x_0+1}{2}\right)=H$, mas não usaremos esse segundo fato).
Vimos então que $h(x_0)=H$ implica $h\left(\frac{x_0}{2}\right)=H$. Prosseguindo indutivamente, segue que $h\left(\frac{x_0}{2^m}\right)=H$ para todo inteiro $m$ com $m\geq 0$. Como $h$ é contínua, podemos tomar o limite $m\to\infty$ e obter

$$H \;=\; \lim_{m\to\infty}h\left(\frac{x_0}{2^m}\right) \;\stackrel{\text{continuidade}}{=}\; h\left(\lim_{m\to\infty}\frac{x_0}{2^m}\right) \;=\; h(0) \;=\; 0\;,$$

concluindo que $H=0$.

**j.** Vimos que o máximo de $h$ em $\mathbb{R}$ é nulo. Isso significa que $h(x)\leq 0$ para todo $x\in\mathbb{R}$. Porém, como $h$ é uma função ímpar (observado no item **f**), isso implica que $h(x)=0$ para todo $x\in\mathbb{R}$. Isso provou que $f(x)=g(x)$ para todo $x\in\mathbb{R}\setminus\mathbb{Z}$, ou seja, provou (6.88) em $\mathbb{R}\setminus\mathbb{Z}$, como queríamos.

Que a relação (6.87) vale para todo $z\in\mathbb{C}\setminus\mathbb{Z}$ segue agora do fato que ambos os lados de (6.88) têm extensões analíticas em todo $\mathbb{C}\setminus\mathbb{Z}$ (prove isso!) e são iguais em $\mathbb{R}\setminus\mathbb{Z}$, por (6.88) (justifique!). ✠

[44]Guillaume François Antoine, Marquês de l'Hôpital (ou l'Hospital) (1661–1704).