Thiago de Lima

Funções
Uma Introdução

Universidade Federal do ABC
Santo André
Maio 2017

Escrito em LaTeX.

Conteúdo

Introdução **8**

I Principais conceitos **12**

1 Funções **14**

1.1 Brevíssima história sobre funções . . . 14
1.2 Alguns exemplos . . . 16
1.2.1 Desempenho dos computadores . . . 16
1.2.2 Forças interatômicas . . . 16
1.2.3 Crescimento populacional . . . 18
1.3 O conceito de função . . . 21
1.3.1 Nomenclatura . . . 28
1.3.2 Ingredientes de uma função: domínio, contradomínio e lei de correspondência . . . 29
1.3.2.1 Domínio . . . 32
1.3.2.2 Contradomínio . . . 34
1.3.2.3 Lei de correspondência . . . 34
1.3.2.4 Misturando todos os ingredientes . . . 35
1.3.3 Natureza da "regra" que define uma função . . . 37
1.4 Revisitando o conceito de função . . . 38
1.4.1 Bourbaki: simplicidade e rigor . . . 41
1.4.2 Pares ordenados . . . 41
1.4.3 Produto cartesiano . . . 42
1.4.4 Relações . . . 44
1.4.5 Funções . . . 47
1.5 Gráficos de funções . . . 49
1.5.1 Diagrama de setas . . . 50
1.5.2 Representação geométrica . . . 51
1.5.2.1 Plano numérico $\mathbb{R}^2$ . . . 51
1.5.2.2 Representação paramétrica . . . 54
1.5.2.3 Coordenadas polares . . . 59

1.5.2.4 Considerações gerais sobre o espaço numérico $\mathbb{R}^3$ 61
1.6 Funções injetivas, sobrejetivas e bijetivas 67
1.6.1 Funções injetivas 67
1.6.2 Funções sobrejetivas 69
1.6.3 Funções bijetivas 70
1.7 Combinando funções: composição e operações 71
1.7.1 Funções compostas 71
1.7.2 Operações com funções 76
1.8 Função inversa 77
1.9 Imagens e imagens inversas 81
1.10 Restrições e extensões de funções 86
1.10.1 Restrições de funções 86
1.10.2 Extensões de funções 87
1.11 Famílias 91
1.11.1 Partições de conjuntos 95
1.11.2 Produtos cartesianos gerais 99
1.12 Funções e conjuntos: propriedades elementares 102
1.13 Funções implícitas: uma noção 104
1.14 Mais alguns conceitos 105
1.14.1 Funções pares, ímpares, simétricas e antissimétricas 105
1.14.2 Periodicidade 111
1.14.3 Monotonias 114
1.14.4 Limitações 116
1.14.5 Taxa de variação 116

II Funções afins e quadráticas **119**

2 Funções afins **121**
2.1 Introdução 121
2.2 Função linear e proporcionalidade 122
2.2.1 Grandeza proporcional a várias outras 128
2.3 Função afim 132
2.3.1 Caracterização da função afim 134
2.4 Gráfico da função afim 136
2.5 Funções afins e progressões aritméticas 139
2.6 Funções poligonais 141

3 Funções quadráticas **147**
3.1 Introdução 147

3.2 Função quadrática 149
3.3 Forma canônica e suas consequências 151
3.4 Gráfico da função quadrática 155
3.5 Algumas aplicações 161

Apêndices **169**

A Alfabeto Grego **171**

B Sistema de Coordenadas Cartesiano **173**
B.1 Na reta 174
B.2 No plano 176
B.3 No espaço 178

C Distâncias **183**
C.1 Distâncias no sistema cartesiano 183
C.2 Generalizando distâncias: a noção de métrica 186

D Eixos não ortogonais no plano **194**

E Princípio da Indução **197**
E.1 Comentários sobre o Princípio da Indução 204

Referências Bibliográficas **210**

Índice Remissivo **210**

Lista de Figuras

1.1 Número de transistores de 1970 a 2014 e curva prevista pela Lei de Moore. 17
1.2 Força entre dois átomos em função da distância interatômica. 18
1.3 Diagrama de $f : X \to Y$. 22
1.4 Função sinal. 24
1.5 Funções piso e teto. 25
1.6 Função módulo. 26
1.7 Interpretação geométrica do módulo. 26
1.8 Contradomínio e lei de correspondência iguais, e domínios diferentes. . . 35
1.9 Taxa do título público *Tesouro IPCA+ 2019* ao longo do tempo[1]. 39
1.10 Ideia do número "dezessete". 40
1.11 Produto cartesiano dos segmentos AB e CD. 43
1.12 Produto cartesiano da circunferência γ e do segmento AB. 44
1.13 Diagrama de uma função $f : X \to Y$. 50
1.14 Coordenadas retangulares. 51
1.15 Gráfico de uma função $f : X \subset \mathbb{R} \to \mathbb{R}$. 52
1.16 Distância entre dois pontos no plano. 53
1.17 Circunferência e disco correspondente. 54
1.18 À esquerda, exemplo do gráfico de uma função $f : X \subset \mathbb{R} \to \mathbb{R}$. À direita, uma curva que não pode ser o gráfico de uma função, pois uma paralela ao eixo OY a intersecta mais de uma vez. 55
1.19 Circunferência de centro na origem e raio r. 55
1.20 Interpretação geométrica da parametrização da elipse. 57
1.21 Parametrização de uma curva um pouco mais sofisticada. 57
1.22 Parábola de Neile (parábola semicúbica). 58
1.23 Sistema de eixos não ortogonais. 59
1.24 Coordenadas polares. 60
1.25 Lemniscata. 60
1.26 63
1.27 Parabolóide de revolução (esquerda) e hiperbólico (direita). 65
1.28 Exemplos de linhas de nível. 66
1.29 Funções injetiva e não injetiva num diagrama de setas. 67

1.30 Funções sobrejetiva e não sobrejetiva num diagrama de setas. 69
1.31 Função bijetiva num diagrama de setas. 70
1.32 Composição de funções. 72
1.33 Composição de funções “numa única etapa”. 73
1.34 Associatividade da composição de funções. 74
1.35 Ideia da restrição de funções. 87
1.36 Exemplo de uma extensão. 89
1.37 Exemplo de uma extensão periódica. 90
1.38 Exemplo de uma extensão periódica par. 90
1.39 Exemplo de uma extensão periódica ímpar. 91
1.40 Exemplo de partição de um conjunto. 96
1.41 Aspectos gerais dos gráficos de uma função par e outra ímpar. 106
1.42 Paridade de algumas funções simples. 106
1.43 Paridade das funções seno e cosseno. 107
1.44 Exemplo de função sem paridade. 109
1.45 Aspectos gerais dos gráficos de uma função simétrica e outra antissimétrica. 109
1.46 Simetria de algumas funções simples. 110
1.47 Uma das mais notáveis funções simétricas: distribuição normal. 111
1.48 Função par obtida de uma função simétrica. 111
1.49 Exemplos de funções que não são periódicas. 112
1.50 Periodicidade de algumas funções trigonométricas. 113
1.51 Função dente de serra. 113
1.52 Função de Euler: exemplo de função periódica. 114
1.53 Gráficos de funções que não são monótonas. 115
1.54 Taxa de variação média de uma função num intervalo. 117
1.55 Secante ao gráfico de uma função e taxa de variação média. 118

2.1 Teorema de Tales e proporcionalidade. 123
2.2 Retângulo decomposto em n retângulos de mesma altura. 127
2.3 Bloco retangular. 130
2.4 Gráfico da função afim. 137
2.5 Equação da reta que é o gráfico de uma função afim. 138
2.6 Interpretação geométrica de uma progressão aritmética. 139
2.7 Gráfico de uma função poligonal. 141
2.8 Esboço do gráfico de imposto devido em função da renda mensal. 142
2.9 Exemplo de COE modelado por função poligonal. 142
2.10 Protótipos de funções poligonais. 143
2.11 Função-rampa. 143
2.12 Gráfico de imposto de renda na faixa de x_0 a x_3 reais. 146

2.13 Funções-rampa “componentes” do gráfico da Figura 2.12. 146

3.1 Colinearidade. 151
3.2 Principais elementos da parábola. 155
3.3 Uma parábola seria o gráfico de uma função quadrática? 156
3.4 Translação vertical do gráfico de f. 159
3.5 Translação horizontal do gráfico de ψ. 159
3.6 Reflexão em torno do eixo horizontal. 159
3.7 Parábolas semelhantes, porém não congruentes. 160
3.8 Trajetória de um projétil na vizinhança da superfície da Terra. 163
3.9 Parabolóide de revolução. 164
3.10 Propriedade refletora da parábola. 165
3.11 Ângulo entre uma reta r e uma curva γ. 165
3.12 Tangente à parábola em (x_0, y_0). 166
3.13 A reta FQ é perpendicular à reta TT'. 167
3.14 Retas cujas equações são $y = ax$ e $y = a'x$. 168
3.15 Os ângulos $A\widehat{P}T$ e $F\widehat{P}T'$ são congruentes. 168

B.1 Eixo. 174
B.2 Sistema de eixos ortogonais no plano. 176
B.3 Sistema de eixos ortogonais no espaço. 179

C.1 Primeiro caso: X e Y à equerda da origem. 183
C.2 Segundo caso: X e Y em lados opostos da origem. 184
C.3 Terceiro caso: X e Y à direita da origem. 184
C.4 Distância entre dois pontos no plano. 185
C.5 Distância entre dois pontos no espaço. 186
C.6 Comparação entre a métrica do taxista e a métrica usual em $\mathbb{R}^2$. 191

D.1 Eixos não ortogonais e coordenadas covariantes e contravariantes. 194

E.1 Polígono convexo. 203
E.2 Decomposição de um polígono por meio de diagonais. 204
E.3 O efeito dominó como analogia para o Princípio da Indução. 205

Introdução

"(...) é inquestionável que quanto antes se familiarize um estudante com o conceito de função, tanto melhor para sua formação matemática."

Howard Eves

O conceito de função está presente em quase todas as áreas da Matemática. Sua abrangência, generalidade e qualidade bastante abstrata lhe conferem um poder unificador e basilar para muitos assuntos matemáticos. Esses atributos todos podem ser as razões pelas quais o leitor leve algum tempo para se acostumar ao conceito e às suas propriedades. Posto isso, a conclusão imediata é a de que, parafraseando Howard Eves, quanto antes se familiarize um estudante com o conceito de função, tanto melhor para sua formação matemática.

O leitor possui em mãos um livro-texto introdutório sobre funções. Para a sua leitura, o mínimo recomendado é a formação equivalente à do primeiro ano do Ensino Médio, além, é claro, de um pouco de paciência e de força de vontade.

O intuito do autor é de que este texto seja utilizado tanto por alunos do Ensino Médio quanto por alunos de cursos de nivelamento do Ensino Superior.

O texto foi escrito de tal modo que se assemelha a uma conversa minimamente formal com esse público-alvo. Por meio dessa conversa, procurou-se transmitir os principais conceitos e resultados sobre funções ao mesmo tempo em que buscou-se questionar e instigar o leitor.

Tentou-se fazê-lo com um mínimo de rigor matemático (que acredita-se ser compatível com a maturidade desses alunos) na expectativa de que isso possa ajudá-los não apenas na transição de um nível escolar para o outro, no caso dos alunos do Ensino Médio, mas também a se prepararem para os cursos subsequentes, nos casos dos alunos de nivelamento.

Na primeira parte do texto, abordam-se as principais noções envolvendo o conceito de função: definições, proposições, notações e representação gráfica.

Na segunda parte, estuda-se, sob um ponto de vista elementar e aplicando-se os conceitos desenvolvidos na primeira parte, as funções afins e quadráticas como modelos para situações reais.

O autor espera que este texto cumpra seu objetivo de auxiliar os estudantes e que valha a pena a sua releitura!

Parte I

Principais conceitos

1 Funções

"A história do termo função proporciona outro exemplo interessante da tendência dos matemáticos de generalizar e ampliar os conceitos."

Howard Eves

O corpo principal da matemática moderna gira em torno de um conceito de destacada importância: o conceito de *função*. Esse conceito pode ser reconhecido aqui e ali ao longo da história da matemática, mas é somente na matemática moderna que sua essência e significado são completamente trazidos à tona. Neste capítulo, tentaremos explicar esse conceito da maneira mais simples e clara possível.

Iniciaremos com alguns exemplos que ilustram sua utilidade e ao mesmo tempo motivam seu uso. A seguir, apresentaremos sua definição, notações, diagramas, gráficos (quando admitirem essa representação), estudaremos algumas de suas propriedades gerais, e veremos como as funções podem ser utilizadas como modelos para aplicação em situações concretas.

Antes de prosseguir, é bastante instrutivo saber, ainda que brevemente, como evoluiu o conceito de função.

1.1 Brevíssima história sobre funções

Nos mapas de locais públicos, a primeira coisa que você procura é aquele ponto em que está escrito "Você está aqui". A única maneira de você se localizar num mapa é inicialmente saber onde você está. Depois você procura o local de destino e o melhor caminho para chegar lá.

De forma semelhante, porém guardadas as devidas limitações da analogia, o objetivo desta seção é esboçar um pequeno mapa histórico sobre a evolução do conceito de função destacando os principais eventos (as principais sínteses, os grandes nomes por trás delas e as épocas) para que você possa se localizar e ter uma ideia para onde vai e por qual caminho. É um tanto curioso, mas esse processo evolutivo coincide com os vários refinamentos do conceito que acompanham os progressos escolares dos estudantes de Matemática. A história tende à rima não ao plágio, mas ainda sim nos indica a direção correta.

Nosso pequeno mapa resume-se a cerca de três séculos de história recente e seis grandes sínteses. Mais do que uma peça de erudição, ele nos servirá para mostrar de que ponto ou estágio o conhecimento sobre esse assunto partiu e em que ponto ou situação ele se encontra atualmente. Como trata-se de um resumo, o estudante que desejar se aprofundar, pode recorrer a [2], que foi nossa principal fonte para esta seção.

O encadeamento dos eventos proporcionará um exemplo interessante da tendência dos matemáticos de generalizar e ampliar os conceitos.

Resumimos a evolução do conceito de função assim:

- 1694 - Leibniz [2]: parece ter introduzido a palavra *função*, na sua forma latina equivalente, para expressar qualquer quantidade associada a uma curva. (Exemplo: as coordenadas de um ponto, a inclinação ou o raio de curvatura de uma curva.)

- 1718 - Johann Bernoulli [2]: chegou a considerar uma função como uma expressão qualquer formada de uma variável e algumas constantes.

- 1748 - Euler [2]: considerou uma função como uma equação ou fórmula qualquer envolvendo variáveis e constantes. (Ideia semelhante que a maioria dos alunos dos cursos elementares de Matemática têm.)

- 1822 - Fourier [2]: foi levado a considerar as séries trigonométricas em suas pesquisas sobre a propagação do calor.

- 1852 - Riemann [2]: chama atenção para o fato de que é indiferente se uma função é definida por uma fórmula ou não.

- Lejeune Dirichlet (1805-1859) [2]: tenta generalizar a definição o suficiente para englobar, entre outras coisas, as séries trigonométricas.

- 1932 - O grupo Bourbaki [2]: propõe a mais rigorosa e abstrata definição.

É muito provável que hoje para você, estudante, o conceito de função tenha o mesmo significado que tinha para Euler em 1748. É verdade que algumas funções possuem como lei de correspondência uma ou mais fórmulas, mas, como veremos, há funções que não podem ser assim definidas. Nas próximas seções, refinaremos essa ideia e, no lugar dela, você disporá de um conceito mais sólido, geral e abstrato.

1.2 Alguns exemplos

1.2.1 Desempenho dos computadores

Em abril de 1965, Gordon E. Moore[1] publicou um artigo [13] no qual observou que o poder de processamento P de um computador, medido pelo número de transistores num *chip*, dobraria a cada ano. Essa observação, que daria origem ao termo *Lei de Moore*[2] por volta da década de 1970, e cuja taxa de crescimento seria modificada para um fator de dois a cada dois anos, pode ser expressa matematicamente da seguinte maneira:

$$P = C \cdot 2^{A/2}, \tag{1.2.1}$$

em que C é o poder de processamento do computador no ano de referência $A = 0$.

Essa fórmula por si só não diz nada sobre as quantidades A e P, mas possui o seguinte significado: se A possui um valor definido, escolhido arbitrariamente num certo conjunto (esse conjunto sendo determinado empiricamente e não matematicamente), então P pode ser determinado. Dizemos que P é uma *função* de A.

Essa regra, que tem funcionado muito bem até hoje, nos diz algo sobre a possibilidade de se produzir computadores milhões de vezes mais potentes do que na época do seu surgimento. Por exemplo, tome como referência o ano de 1970, para o qual $A = 0$, e experimente calcular P para o ano 2014, para o qual $A = 44$. Você deve encontrar um número da ordem de quatro milhões, o que significa que em 2014 as indústrias de semicondutores conseguiam produzir processadores cerca de quatro milhões de vezes mais potentes do que em 1970 (ver Figura 1.1). Uma consequência imediata desse fato foi a redução do tamanho, do custo e do gasto energético dos processadores (apenas para se ter uma ideia, imagine que se não fosse assim, seu *smartphone* precisaria ser do tamanho de uma geladeira, custaria milhões de dólares e utilizaria milhares de vezes mais energia para funcionar razoavelmente!).

1.2.2 Forças interatômicas

Atualmente, sabemos reduzir todos os tipos de forças conhecidas a apenas quatro tipos de interações fundamentais: gravitacional, eletromagnética, forte e fraca. Todas as demais forças que aparecem na natureza podem, em princípio, ser reduzidas a uma dessas quatro interações.

[1] Gordon Earle Moore (1929–), físico e químico cofundador da Intel Corporation em 1968 juntamente com o físico Robert Norton Noyce (1927–1990).

[2] Apesar do termo, trata-se mais de uma observação do que propriamente de uma lei, que curiosamente acabou se tornando um objetivo para as indústrias de semicondutores, fazendo-as dispenderem muitos recursos para poder alcançar as previsões de Moore. Isso torna a Lei de Moore realmente importante, pois sem ela talvez não houvesse um desenvolvimento tão acelerado de hardware com custos cada vez mais acessíveis.

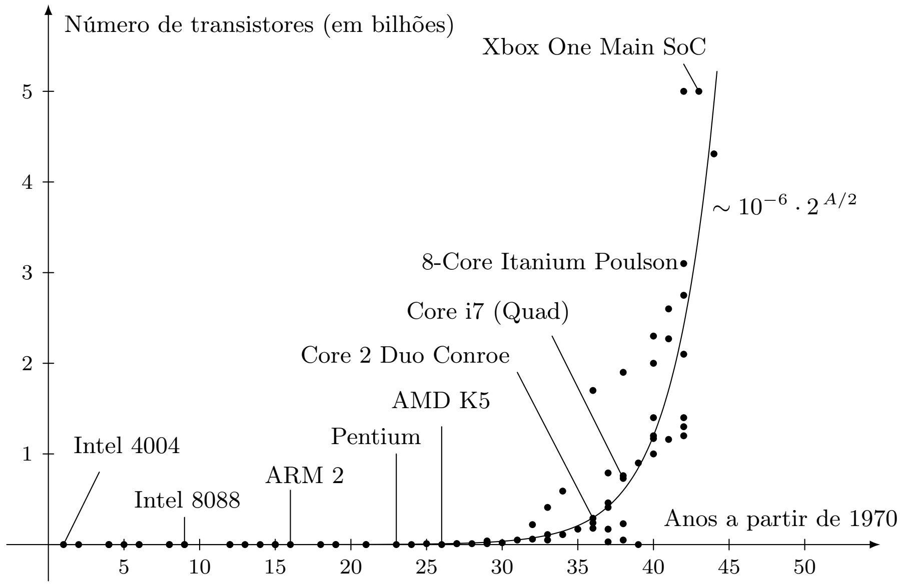


Figura 1.1: Número de transistores de 1970 a 2014 e curva prevista pela Lei de Moore.

Dessas, as interações forte e fraca só desempenham um papel relevante na escala nuclear devido a seu curto alcance. Assim, do ponto de vista macroscópico, prevalecem as interações eletromagnética e gravitacional. A estrutura dos átomos e as forças interatômicas dependem predominantemente da interação eletromagnética combinada com os princípios da mecânica quântica.

Consideremos a interação entre dois átomos que podem formar uma molécula diatômica, admitindo que seus centros possam se deslocar apenas ao longo de uma reta, a fim de tornar o problema unidimensional. Em 1931, Lennard-Jones [3] propôs [6] que essa força fosse da forma:

$$F(R) = \lambda_{rep.}\, R^{-13} - \lambda_{atr.}\, R^{-7},$$

em que $\lambda_{rep.}$ e $\lambda_{atr.}$ são constantes positivas que representam respectivamente a repulsão e a atração que surge na interação e R é a distância interatômica. Dizemos que F é uma *função* de R.

Um gráfico da força em função da distância entre eles tem o aspecto da Figura 1.2.

[3] John Edward Lennard-Jones (1894–1954), matemático britânico, foi professor de física teórica na Universidade de Bristol e depois de química teórica na Universidade de Cambridge. Considerado o criador da química computacional moderna.

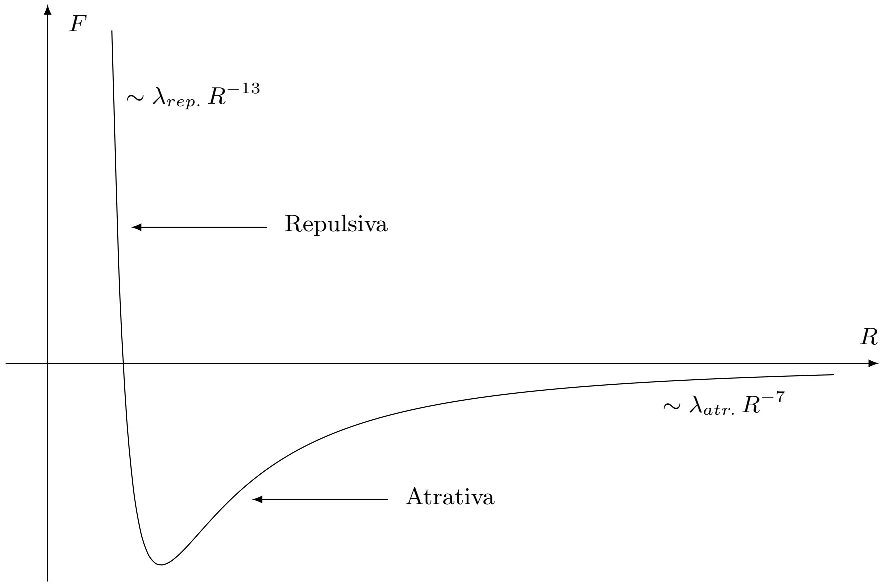


Figura 1.2: Força entre dois átomos em função da distância interatômica.

1.2.3 Crescimento populacional

Considere uma população de tamanho descrito pela variável N_t (número de indivíduos no tempo t, onde t assume valores inteiros: $t = 1, 2, 3, \ldots$). Vamos supor que a população se reproduza a intervalos iguais (semanas, meses, estações, gerações ou qualquer outra unidade). A *taxa de crescimento* R_t referente ao tempo t será definida por

$$R_t = \frac{N_{t+1} - N_t}{N_t}, \tag{1.2.2}$$

ou seja, a variação do número de indivíduos de uma população $(N_{t+1} - N_t)$ em relação ao seu número inicial (N_t).

Por exemplo, se no ano $t = 2013$ a população é de $N_{2013} = 1000$ indivíduos e no ano posterior é de $N_{2014} = 1100$ indivíduos, então sua taxa de crescimento com relação ao ano $t = 2013$ é de

$$R_{2013} = \frac{N_{2014} - N_{2013}}{N_{2013}} = \frac{1100 - 1000}{1000} = \frac{100}{1000} = 0,10 \text{ ou } 10\%.$$

Considere agora o caso mais simples possível, o modelo com taxa de crescimento constante, isto é,

$$R_t = r \ (constante). \tag{1.2.3}$$

Qual a consequência disso para uma descrição da forma como o tamanho da população varia com o tempo? Substituindo (1.2.3) em (1.2.2) e isolando N_{t+1} encontraremos uma regra[4] que nos dará a população N_{t+1} do ano $t + 1$ em função da população do

[4] A seguir veremos precisamente o que isso significa.

ano anterior N_t (verifique!):

$$N_{t+1} = (1 + r)N_t.$$

Para simplificar a notação, podemos introduzir uma nova constante $\xi = 1+r$. Assim, a equação acima fica:

$$N_{t+1} = \xi N_t. \tag{1.2.4}$$

Note que para $\xi > 1$ ($r > 0$), a população aumenta a cada intervalo de tempo; para $\xi < 1$ ($r < 0$), diminui; e para $\xi = 1$ ($r = 0$), não varia.

Em outras palavras, a hipótese de taxa de crescimento constante nos permitiu deduzir que o tamanho da população num instante $t + 1$ é dada pelo produto da população no instante anterior t e um parâmetro constante ξ.

Por exemplo, se N_t representa o número de besouros (em alguma unidade) na geração t (lembre-se que t representa uma unidade de tempo arbitrária, portanto podemos escolher "geração" como tal) e ξ representa o número médio de larvas produzido por besouro, então

$$N_t = N_1\, \xi^{t-1}.$$

De fato, a geração t descende da $(t-1)$-ésima geração. Como os besouros da $(t-1)$-ésima geração produzem ξ larvas cada, $N_t = \xi\, N_{t-1}$. Aplicando o mesmo raciocínio à $(t-1)$-ésima geração, obtemos $N_{t-1} = \xi\, N_{t-2}$. Logo, podemos escrever $N_t = \xi^2\, N_{t-2}$. Continuando esse raciocínio, obtemos $N_t = N_1\, \xi^{t-1}$, provando nossa afirmação sobre o crescimento da população de besouros.

Note que, dados N_1 e ξ, para cada t natural menor do que um certo $T \in \mathbb{N}$ (T sendo determinado empiricamente e não matematicamente), N_t pode ser determinado. Dizemos, então, que N_t é uma *função* de t. Da mesma forma, como para todo t natural num certo intervalo (determinado empiricamente) R_t pode ser determinado em (1.2.3), – na verdade, nesse caso R_t sempre vale r – dizemos também que R_t é uma função de t.

Neste ponto, há vários motivos para que o estudante fique frustrado. Várias queixas podem ser e serão escutadas:

1. Como sabemos que é assim?
2. Posso mencionar vários contraexemplos[5].
3. Certamente ξ não é constante.
4. Como devo ler ξ?

[5]Para provar que uma afirmação é falsa, basta exibir uma exceção. Tal exceção é chamada um contraexemplo.

Comecemos pelo mais fácil. Os matemáticos usam letras para representar grandezas numéricas e outras coisas também (teremos oportunidade de constatar isso mais adiante). Ocorre porém que letras diferentes às vezes acabam antes de terminarmos de descrever as grandezas que desejamos. O que se faz, então, é emprestar letras de outros alfabetos. A letra ξ emprestamos do alfabeto grego. Em português ela é pronunciada "csi". No Apêndice A você pode encontrar uma lista com as letras maiúsculas e minúsculas do alfeto grego e seus respectivos nomes.

As outras queixas são mais sérias. O que se faz ao construir um modelo matemático é tentar ver as consequências das hipóteses feitas e se há algum comportamento interessante. Se não há, jogamos fora o modelo. Caso contrário, algumas mudanças podem conduzir a um modelo melhor. Algumas vezes somos guiados por evidências experimentais, outras por motivos teóricos.

Por exemplo, no caso particular em que supusemos que a taxa de crescimento populacional não variava com o tempo (equação (1.2.3)), a conclusão a que chegamos é de que teremos um crescimento ou decrescimento exponencial (equação (4)).

No entanto, ainda que esse tipo de comportamento seja observado em diversas situações, é óbvio que ele não pode se manter por muito tempo no caso de populações de organismos. Mesmo na ausência de predadores ou catástrofes naturais, uma população não pode crescer indefinidamente num ambiente com recursos finitos.

Talvez o modelo mais simples que inclui essa característica é aquele que considera que a taxa de crescimento depende do tamanho da população da seguinte maneira:

$$R_t = r - sN_t, \tag{1.2.5}$$

onde $s > 0$ é uma constante que descreve como o tamanho da população afeta a taxa de crescimento (quanto maior s, mais rapidamente R_t decresce).

Note que existe um valor de N_t, digamos N_{eq}, para o qual a taxa de crescimento é zero. Nesse caso,

$$N_{eq} = \frac{r}{s}.$$

Assim, podemos reescrever a equação (1.2.5):

$$R_t = r - r\frac{N_t}{N_{eq}} = r(1 - x_t),$$

onde $x_t = N_t/N_{eq}$ é o tamanho da população relativamente à população de equilíbrio.

Agora, dividindo o numerador e o denominador do segundo membro da equação (1.2.2) por N_{eq} e substituindo R_t, obtemos a seguinte dinâmica populacional (verifique!):

$$x_{t+1} = (1 + r)x_t - rx_t^2.$$

Essa equação é conhecida como *mapa logístico* ou *processo de Verhulst*[6] para t discreto[7]. O parâmetro r, característico do tipo de organismo modelado, é a taxa de crescimento quando a população é pequena em relação à população de equilíbrio e é denominado *taxa de crescimento natural*, ou seja, é aquela na ausência de efeitos de competição por recursos naturais devido ao crescimento populacional.

Dado o parâmetro r e a condição incial x_1, podemos determinar x_t para todo t natural menor do que um certo $T \in \mathbb{N}$ (T sendo determinado empiricamente e não matematicamente). Dizemos, então, que x_t é uma *função* de t. Tal função modela o crescimento populacional sob a hipótese que fizemos em (1.2.5), que representa a taxa de crescimento como função de t (note que para cada $t \in \mathbb{N}$ corresponderá um único valor de N_t – e nesse sentido N_t também é função de t – e consequentemente um único valor de R_t).

Agora o estudante pode questiornar-se: o que a evolução do desempenho dos computadores, as forças interatômicas e a dinâmica de populações têm em comum? Como já deve ter suspeitado, do ponto de vista da Matemática, todos esses fenômenos podem ser modelados e estudados utilizando-se um conceito de importância fundamental para as ciências e para a própria Matemática: o conceito de *função*.

1.3 O conceito de função

A Matemática "fornece uma variedade de conceitos abstratos que servem de modelos para situações concretas, permitindo assim analisar, prever e tirar conclusões de forma eficaz em circunstâncias nas quais a abordagem empírica muitas vezes não conduz a nada (ver [10])."

A escolha do modelo depende da comparação das características do problema a ser estudado com as propriedades do próprio modelo. Esse processo requer que se conheçam os teoremas de caracterização do modelo. Nos próximos capítulos, falaremos sobre as funções afins e quadráticas como exemplos de modelos matemáticos para representar situações específicas e apresentaremos seus teoremas de caracterização.

Convém, portanto, (re)apresentar ao leitor uma das noções matemáticas mais básicas que lhe acompanhará pelos próximos capítulos e além, o conceito de *função*. Esse é o conceito subjacente aos exemplos acima.

Nos cursos mais básicos de matemática elementar, geralmente somos apresentados à seguinte definição:

[6]Em homenagem a Pierre François Verhulst (1804-1849). O modelo proposto por Verhulst complementa a teoria do crescimento exponencial ao considerar fatores de inibição. Seu modelo atualmente chama a atenção como exemplo importante da teoria do caos.

[7]Para t contínuo o modelo seria dado por uma *equação diferencial* conhecida como *equação logística*.

Definição 1.1. *Sejam* X *e* Y *conjuntos. Uma* função *de* X *em* Y, $f : X \to Y$, *é uma regra que diz como associar a cada elemento* $x \in X$ *um único elemento* $y = f(x) \in Y$.

Além disso,

1. os conjuntos X e Y chamam-se o *domínio* (ou o *campo de existência* ou o *campo de definição*) e o *contradomínio* (ou o *codomínio*) da função f, respectivamente;

2. para cada elemento $x \in X$, o elemento $f(x) \in Y$ chama-se a *imagem* de x por f, ou o *valor* que f *assume* (ou *toma*) no *argumento* $x \in X$; equivalentemente, diz-se que f *manda* ou *aplica* ou *transforma* ou *leva* x em $f(x)$ e escreve-se $x \mapsto f(x)$;

3. o conjunto de todo $y \in Y$ tal que $y = f(x)$ para algum $x \in X$ chama-se a *imagem* (ou o *conjunto de valores*, ou, ainda, o *campo de valores*) de f, e denota-se: $Im(f) := f(X) = \{y \in Y; \exists x \in X, y = f(x)\}$;

Observação 1.2. Note que, na definição acima, X e Y são conjuntos quaisquer, inclusive vazios. Embora as situações em que o domínio ou o contradomínio ocorram como o conjunto vazio sejam raras e pouco interessantes, elas são permitidas pela teoria e mantê-las evita que precisemos provar que X ou Y não são vazios cada vez que considerarmos funções arbitrárias de X em Y. Após estudar a Seção 1.4.5, você será capaz de provar o seguinte: no caso em que $X = \emptyset$, para qualquer Y (vazio ou não), $f = \emptyset$; no caso em que $Y = \emptyset$, para qualquer $X \neq \emptyset$, não existe f. Naquela seção, retomaremos esta observação. ○

Uma forma de ilustrar a ideia geral de uma função é aquela do diagrama da Figura 1.3. Trata-se de uma imagem evocativa para ajudar a pensar no conceito de função, que a cada elemento x do domínio X associa um só valor $f(x)$ no contradomínio Y. Nessa figura, assim como na notação $f : X \to Y$, a seta representa a "regra" que "associa" $x \in X$ a $f(x) \in Y$.

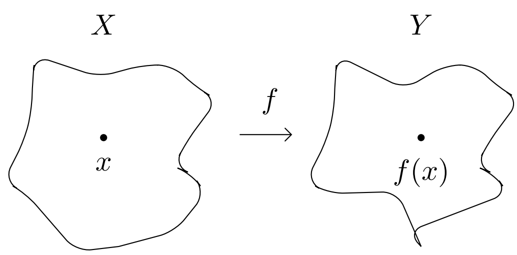


Figura 1.3: Diagrama de $f : X \to Y$.

Se considerarmos o símbolo x variando à vontade no conjunto X e $y = f(x)$, x recebe o nome de *variável independente* e y recebe o nome de *variável dependente*.

É importante ressaltar que $f(x)$ é a imagem do elemento $x \in X$ pela função f, ou o valor da função f em $x \in X$. O leitor encontrará muitos textos nos quais os autores dizem "a função $f(x)$" ou, em alguma situação específica, algo como "a função $x^2 + 1$", quando deveriam dizer respectivamente "a função f" ou "a função $p : \mathbb{R} \to \mathbb{R}$ tal que $p(x) = x^2 + 1$", especificando, portanto, nesse último caso, o domínio, o contradomínio e a regra de associação $x \mapsto y$. Essa linguagem inexata agiliza a comunicação, mas é indispensável sempre ter a noção precisa do que se está fazendo.

Exemplo 1.3. São exemplos particularmente simples e bem conhecidos de funções:

(a) a *função inclusão* $i : X \subset Y \to Y$, definida por $i(x) = x$ para todo $x \in X \subset Y$;

(b) a *função identidade* $id_X : X \to X$, definida por $id_X(x) = x$ para todo $x \in X$ (note que a função identidade nada mais é do que a função inclusão de X sobre X);

(c) as *funções constantes* $f : X \to Y$, definidas por $f(x) = c$, para todo $x \in X$ e algum $c \in Y$;

(d) as *projeções* de um produto cartesiano $A \times B$ nos fatores A e B: $\pi_1 : A \times B \to A$, $\pi_1(a,\, b) = a$, chamada *primeira projeção*, e $\pi_2 : A \times B \to B$, $\pi_2(a,\, b) = b$, chamada *segunda projeção*;

(e) a *função indicadora* (ou *função característica*) de um conjunto $A \subset X$, $\chi_A : X \to \{0,\, 1\}$, definida por:

$$\chi_A(x) = \begin{cases} 1 & \text{se } x \in A \\ 0 & \text{se } x \in X \setminus A \end{cases} .$$

A letra grega χ é lida "qui". A função indicadora também pode ser representada mais compactamente se for utilizada a notação dos colchetes de Iverson[8]: $\chi_A(x) = [x \in A]$. $\diamond$

Exemplo 1.4. Algumas funções bastante utilizadas em programação:

(a) a *função sinal*, $\operatorname{sgn} : \mathbb{R} \to \{-1,\, 0,\, 1\}$, que retorna o sinal de um número real, ou seja:

$$\operatorname{sgn} x = \begin{cases} -1 & \text{se } x < 0 \\ 0 & \text{se } x = 0 \\ 1 & \text{se } x > 0 \end{cases} ,$$

e se relaciona com a função indicadora por meio da equação:

$$\operatorname{sgn} x = \chi_{x<0}(x) - \chi_{x>0}(x)$$

[8]Kenneth Eugene Iverson (1920–2004), matemático canadense, vencedor do prêmio Turing em 1979 por seu esforço pioneiro nas linguagens de programação e na notação matemática. Os colchetes de Iverson são definidos da seguinte maneira: seja P uma proposição; $[P] = 1$ ou 0 conforme P seja verdadeira ou falsa.

e com a função módulo (que veremos no Exemplo 1.6) por meio da equação (veja o gráfico dessa função na Figura 1.4):

$$x = |x| \operatorname{sgn} x \quad \text{ou} \quad |x| = x \operatorname{sgn} x;$$

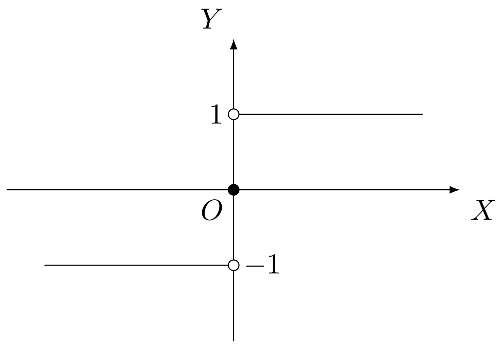


Figura 1.4: Função sinal.

(b) a *função chão* (ou *piso* ou ainda *floor* em inglês), que associa um número real x ao maior número inteiro menor ou igual a x, $\lfloor \cdot \rfloor = \mathbb{R} \to \mathbb{Z}$,

$$\lfloor x \rfloor = \max\{n \in \mathbb{Z};\ n \leqslant x\},$$

por exemplo, $\lfloor 1,3 \rfloor = 1$, $\lfloor 17,999 \rfloor = 17$, $\lfloor -5,5 \rfloor = -6$ (veja o gráfico dessa função na Figura 1.5a);

(c) a *função teto* (ou *ceil* em inglês), que associa um número real x ao menor número inteiro maior ou igual a x, $\lceil \cdot \rceil = \mathbb{R} \to \mathbb{Z}$,

$$\lceil x \rceil = \min\{n \in \mathbb{Z};\ n \geqslant x\},$$

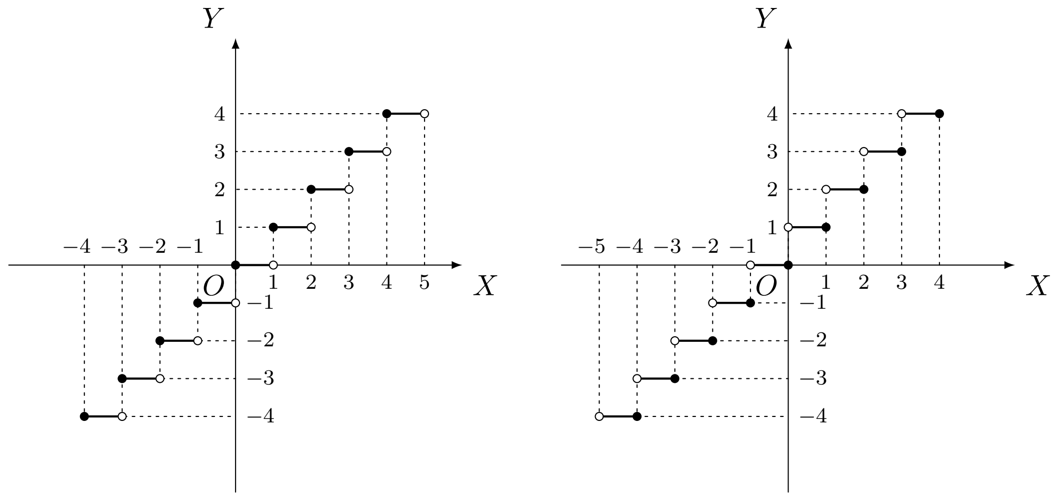
por exemplo, $\lceil 6,67 \cdot 10^{-11} \rceil = 1$, $\lceil 17,2 \rceil = 18$, $\lceil -5,5 \rceil = -5$ (veja o gráfico dessa função na Figura 1.5b). ◇

Exemplo 1.5. A partir das funções dos itens (b) e (c) do Exemplo 1.4, podem ser definidas outras, como:

(a) a *função parte inteira* (ou *valor inteiro*) de um número real, $\operatorname{int} : \mathbb{R} \to \mathbb{Z}$,

$$\operatorname{int} x = \begin{cases} \lfloor x \rfloor & \text{se } x \geqslant 0 \\ \lceil x \rceil & \text{se } x < 0 \end{cases},$$

(alguns autores definem simplesmente $\operatorname{int} x = \lfloor x \rfloor$ para todo x real, mas isso é uma questão de gosto; adotaremos a definição acima); e



(a) Função piso.   (b) Função teto.

Figura 1.5: Funções piso e teto.

(b) a *função parte fracionária* (ou *valor fracionário*) de um número real, $\{\cdot\} : \mathbb{R} \to \mathbb{Z}$,

$$\{x\} = x - \operatorname{int} x.$$

(alguns autores, também por questões de gosto, preferem adotar a definição $\{x\} = x - \lfloor x \rfloor$ para todo x real; adotaremos a definição acima.) ◇

Exemplo 1.6. Uma função especialmente útil é a *função módulo* de um número real: $|\cdot|\colon X \subset \mathbb{R} \to \mathbb{R}$, definida da seguinte maneira:

$$|x| = \begin{cases} x & \text{se } x \geqslant 0 \\ -x & \text{se } x < 0, \end{cases}$$

que também pode ser definida como:

$$|x| = \max\{x, -x\},$$

ou seja, o maior dos números x e $-x$ (claro que se $x = 0$, então $x = -x = |x| = 0$), ou, em termos da função sinal, conforme já dissemos no item (a) do Exemplo 1.4:

$$|x| = x \operatorname{sgn} x.$$

Há, ainda, outra caracterização:

$$|x| = \sqrt{x^2},$$

ou seja, $|x|$ é o número não negativo cujo quadrado é x^2. Em algumas raras situações envolvendo expressões em que o módulo aparece como algum termo, essa última caracterização evita que precisemos considerar caso a caso os subconjuntos do domínio de $|\cdot|$ nos quais o seu argumento muda de sinal.

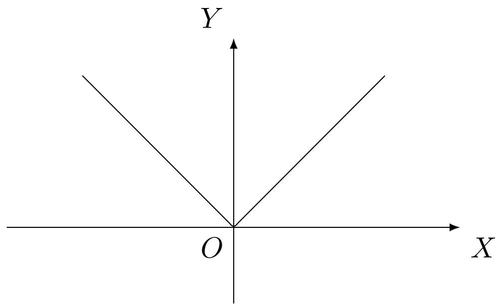


Figura 1.6: Função módulo.

O valor da função módulo em x, i.e., $f(x) = |x|$ é chamado o *módulo* ou o *valor absoluto* do número real x. (Veja o gráfico dessa função na Figura 1.6.)

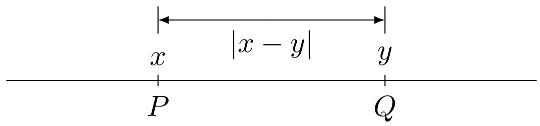


Figura 1.7: Interpretação geométrica do módulo.

Geometricamente, o módulo é interpretado como a distância entre dois pontos na reta da seguinte maneira: se x e y são as coordenadas dos pontos P e Q, respectivamente, então a distância entre esses pontos é $d(P, Q) = |x - y|$ (ver Figura 1.7 e para mais detalhes sobre distâncias na reta, veja a Seção B.1). A vantagem dessa interpretação é permitir enxergar intuitivamente o significado e a resposta de algumas questões envolvendo módulos.

Por exemplo, a igualdade $|x - 17| = 1$ significa que o número x (ou o ponto correspondende no eixo) está a uma distância 1 do número 17, devendo ser $x = 16$ (se estiver à esquerda de 17) ou $x = 18$ (se estiver à direita). A desigualdade $|x - a| < \delta$, com $\delta > 0$, significa que a distância de x até a é menor do que δ e portanto que x deve estar entre $a - \delta$ e $a + \delta$, ou seja, o conjunto $\{x \in \mathbb{R};\ |x - a| < \delta\}$ é igual ao intervalo $(a - \delta,\ a + \delta)$ (falaremos sobre intervalos na Seção 1.3.2). ◇

Exemplo 1.7. Um valioso exemplo de função, ou melhor, de uma classe de funções, consiste nas *sequências* (ou *sucessões*), funções cujo domínio é o conjunto dos números naturais.

Em poucos símbolos e palavras, uma função $f : \mathbb{N} \to Y$ cujo domínio é o conjunto dos números naturais chama-se uma *sequência* ou *sucessão* de elementos de Y.

A notação usada para representar uma tal sequência é $(y_1,\ y_2,\ \ldots,\ y_n,\ \ldots)$, ou $(y_n)_{n\in\mathbb{N}}$ (o porquê dessa notação ficará mais claro quando falarmos sobre famílias na Seção 1.11), ou simplesmente (y_n), em que o valor da função f no número n, $f(n)$, é indicado pelo símbolo y_n, chamado o *n-ésimo* termo da sequência.

Se Y é um conjunto de números reais, dizemos que (y_n) é uma sequência de números reais. A Seção 1.2.3 contém alguns exemplos. Algumas sequências famosas e particularmete interessantes de números reais são:

(a) a *progressão aritmética*, em que cada termo, a partir do segundo, é a soma $y_{n+1} = y_n + r$ do termo anterior com uma constante r, chamada a *razão* da progressão (os anos de passagem do cometa Halley[9] pela Terra formam uma sequência desse tipo);

(b) a *progressão geométrica*, em que cada termo, a partir do segundo, é o produto $y_{n+1} = y_n \cdot r$ do termo anterior por uma constante r, também chamada a *razão* da progressão (os números de transistores num *chip* previstos pela lei de Moore, equação (1.2.1), ou o tamanho de uma população ao longo das gerações num ambiente com recursos abundantes e livre de predadores, pragas ou competidores, equação (1.2.4), formam progressões desse tipo);

(c) a *sequência de Fibonacci*[10], em que cada termo, a partir do terceiro, é a soma $y_{n+2} = y_{n+1} + y_n$ dos dois termos anteriores, com $y_1 = y_2 = 1$. Essa sequência aparece em diversas configurações biológicas (como na concha do Nautilus, na dentição humana, no arranjo do cone da alcachofra, no desenrolar da samambaia, na reprodução de abelhas[11] etc.) e tem aplicações na análise de mercados financeiros, na ciência da computação e na teoria de jogos apenas para mencionar algumas.

As sequências (leia-se também: funções) acima têm uma característica em comum que servirá para ilustrar uma lição importante. Como vimos até aqui, para se definir uma função $f : X \to Y$ exige-se, em geral, que seja dada uma regra bem determinada, que mostre como se deve associar a cada elemento $x \in X$ um único elemento $y = f(x) \in Y$. No entanto, no caso particular em que $X = \mathbb{N}$, não é necessário dizer, de uma só vez, qual é a receita que dá o valor $f(n)$ para todo $n \in \mathbb{N}$. Para um dado $k \in \mathbb{N}$, basta sabermos:

1. os valores $y_1 = f(1)$, $y_2 = f(2)$, $\ldots$, $y_k = f(k)$; e

[9] Em honra a Edmond Halley (1656–1742), astrônomo e matemático britânico que calculou a órbita do cometa que hoje leva o seu nome e defensor de Newton.

[10] Leonardo de Pisa (1170–1250) ou Fibonacci (contração do italiano *filius Bonacci*, que significa filho de Bonacci, de onde provém o apelido) desempenhou um papel importante na revitalização da Matemática antiga e fez contribuições significativas de sua autoria. Seu livro, *Liber Abaci*, foi responsável pela introdução do sistema de numeração indo-arábico na Europa e pelo posterior desenvolvimento da álgebra e da aritmética no ocidente.

[11] Uma recente análise histórica-matemática [12] sugere que foram as linhagens de abelhas que realmente inspiraram os números de Fibonacci, em vez do famigerado modelo de reprodução de coelhos tal qual aparece no livro *Liber Abaci*.

2. a regra que permite calcular, para todo $n \in \mathbb{N}$, $y_{n+k} = f(n+k)$ quando se conhece $y_n = f(n)$, $y_{n+1} = f(n+1)$, ..., $y_{n+k-2} = f(n+k-2)$, $y_{n+k-1} = f(n+k-1)$.

Esses dados permitem que se conheça $y_n = f(n)$ para todo $n \in \mathbb{N}$ e se diz que a função f foi definida por *recorrência*. Por sua vez, a sequência correspondente, (y_n), é chamada de *sequência recorrente de ordem* k (assim, progressões aritméticas e geométricas são exemplos de sequências recorrentes de primeira ordem, enquanto a sequência de Fibonacci é um exemplo de sequência recorrente de segunda ordem). Essa maneira de se definir funções (recursivamente) se apoia num fato fundamental sobre a estrutura dos números naturais, mais especificamente, no chamado axioma da indução, um dos axiomas de Peano[12] sobre os números naturais. Para a demonstração do fato acima e mais detalhes sobre esse assunto, recomendamos muito fortemente ao estudante a leitura de [8], um artigo elegante, de leitura fácil, rápida e agradável, destinado a jovens que se preparam para a Olimpíada Brasileira de Matemática; esse texto também é uma ótima oportunidade para o estudande reforçar os conceitos aqui estudados sobre funções. Uma versão (bem) resumida, mas que não dispensa a leitura de [8], pode ser encontrada no Apêndice E. $\diamond$

1.3.1 Nomenclatura

Sobre a nomenclatura, há grande quantidade de palavras usadas para *função*: *aplicação*, *correspondência*, *forma*, *funcional*, *mapa*, *mapeamento*, *operação*, *operador*, *produto*, *transformação*, e talvez ainda outras.

O que difere seu uso é por vezes a tradição de certas áreas e os tipos de domínio ou imagem. Por exemplo: *função* é comumente utilizada quando se trata de funções numéricas como as de $\mathbb{R}$ em $\mathbb{R}$ ou as de $\mathbb{C}$ em $\mathbb{C}$; *aplicação*, *mapa* e *mapeamento* são frequentemente empregadas para designar funções em áreas como Geometria Diferencial, Topologia ou Sistemas Dinâmicos; *forma* designa por vezes certas funções bilineares de $V \times V$ em $\mathbb{R}$ ou $\mathbb{C}$, em que V é um espaço vetorial; *funcional*[13] geralmente refere-se a funções que levam vetores ou funções numéricas em números, como a função que leva funções reais contínuas f nas suas integrais no intervalo [0, 1]: $f \mapsto \int_0^1 f(x)\, dx$; *operações* e *produtos* frequentemente designam funções de $X \times X$ em X, para um conjunto X não vazio qualquer; *operador* tipicamente designa funções lineares entre espaços vetoriais, como as matrizes, que são funções lineares entre espaços vetoriais de dimensão finita; *transformações* são utilizadas com frequência na teoria de semigrupos

[12] Giuseppe Peano (1858–1932), matemático italiano, considerado o fundador da lógica simbólica; seus interesses estavam centrados nos fundamentos da Matemática e no desenvolvimento de uma linguagem lógica formal.

[13] Essa palavra foi empregada pela primeira vez na Matemática por Jacques Salomon Hadamard (1865–1963).

para designar funções $f : X \to X$ que mapeiam um conjunto nele mesmo, sendo alguns exemplos as transformações lineares e afins, translações, rotações e reflexões.

Há ainda certos termos para designar certas funções com propriedades especiais. Por exemplo, um *homeomorfismo* é uma função bijetora entre dois espaços topológicos que seja contínua e cuja inversa seja também contínua. Um *difeomorfismo* é um homeomorfismo entre duas variedades diferenciáveis que seja infinitamente diferenciável. Há ainda outros "morfismos".

Essa abundância de palavras causa frequentemente confusão e mesmo perplexidade em estudantes recém-iniciados, mas, em essência, todos esses objetos são funções como definimos acima. É conveniente dispormos por vezes de certa quantidade de palavras sinônimas para evitarmos o uso monótono e descolorido da palavra *função*. Com uma suave ironia, lembremos da definição circular de Edward Teller[14]: *"An intellectual is someone who thinks the same things and uses the same words as other intellectuals"*.

1.3.2 Ingredientes de uma função: domínio, contradomínio e lei de correspondência

Intervalos No caso de funções reais de uma variável real, há certos subconjuntos de $\mathbb{R}$ muito comuns utilizados como domínio (ou contradomínio): os *intervalos*. Já os utilizamos acima, mas uma pausa para uma palavrinha sobre eles antes de prosseguirmos será útil. Trata-se da seguinte definição:

Definição 1.8. *Sejam a, $b \in \mathbb{R}$ com $a \leqslant b$. Chamam-se* intervalos *os seguintes subconjuntos de $\mathbb{R}$:*

1. $[a, b] = \{x \in \mathbb{R};\ a \leqslant x \leqslant b\}$;
2. $(a, b) = \{x \in \mathbb{R};\ a < x < b\}$;
3. $[a, b) = \{x \in \mathbb{R};\ a \leqslant x < b\}$;
4. $(a, b] = \{x \in \mathbb{R};\ a < x \leqslant b\}$;
5. $[a, +\infty) = \{x \in \mathbb{R};\ a \leqslant x\}$;
6. $(a, +\infty) = \{x \in \mathbb{R};\ a < x\}$;
7. $(-\infty, b] = \{x \in \mathbb{R};\ x \leqslant b\}$;
8. $(-\infty, b) = \{x \in \mathbb{R};\ x < b\}$;
9. $(-\infty, +\infty) = \mathbb{R}$.

Além disso, chamamos de:

1. *limitados* os quatro primeiros intervalos; desses, chamamos de:
   (a) *fechado* o intervalo $[a, b]$;
   (b) *aberto* o intervalo (a, b);

[14] Edward Teller (1908–2003), físico nuclear norte-americano de origem húngara, conhecido como "o pai da Bomba H".

   (c) *fechado à esquerda* (ou *aberto à direita*) o intervalo $[a, b)$;

   (d) *fechado à direita* (ou *aberto à esquerda*) o intervalo $(a, b]$;

   a e b são ditos os *extremos* desses intervalos;

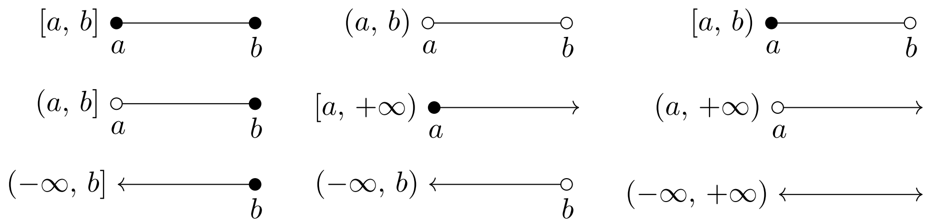
2. *ilimitados* os demais; dizemos que $[a, +\infty)$ é a semirreta direita, fechada, de origem a; analogamente para os outros intervalos ilimitados.

3. *intervalo degenerado* o intervalo fechado $[a, b]$ quando $a = b$, ou seja, quando esse intervalo se reduz ao conjunto unitário $\{a = b\}$; note que os demais intervalos limitados se tornam o conjunto vazio quando $a = b$.

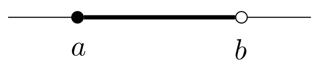
Observação 1.9. Muitos textos brasileiros de Matemática elementar utilizam a notação $]a, b[$ ao invés de (a, b), $[a, b[$ ao invés de $[a, b)$, e assim por diante. ○

Observação 1.10. Os símbolos $\pm\infty$ não são números! São apenas parte da notação dos intervalos ilimitados. ○

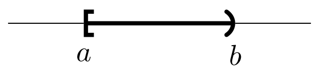
Como existe uma correspondência biunívoca entre a reta e $\mathbb{R}$ (veja Apêndice B), costuma-se representar assim os intervalos não degenerados:



Quando se quer representar um intervalo numa reta, costuma-se destacá-lo de alguma maneira, geralmente com uma cor mais forte, para evitar ambiguidade; por exemplo, o intervalo $[a, b)$ seria indicado assim:



Por vezes, ao invés das bolas abertas e fechadas, são utilizados os colchetes e os parênteses da notação dos intervalos nessa representação gráfica, por exemplo, o intervalo $[a, b)$ poderia ser representado assim:



Um fato particularmente importante sobre os intervalos não degenerados e que utilizaremos para demonstrar o Teorema Fundamental da Proporcionalidade no capítulo sobre funções afins é que eles sempre contêm números racionais e irracionais, ou seja, racionais e irracionais estão por toda parte em $\mathbb{R}$. Vejamos.

Proposição 1.11. *Todo intervalo não degenerado contém números racionais e irracionais.*

Demonstração 1.

Seja I um intervalo não degenerado e $(a,\, b) \subset I$ com $a < b$ um intervalo aberto qualquer em I. Vamos mostrar que existem um racional e um irracional em $(a,\, b)$.

Há três casos a considerar: $a \geqslant 0$; $b \leqslant 0$; e $a < 0 < b$. Sem perda de generalidade, podemos considerar apenas o primeiro deles, pois se $b \leqslant 0$, nossos argumentos se aplicarão a $(-b,\, -a)$ e consequentemente a $(a,\, b)$, e se $a < 0 < b$, poderemos simplesmente nos restringir a $(0,\, b) \subset (a,\, b)$.

$(a,\, b)$ **contém um racional.** Considere as representações decimais de $a = a_0, a_1 a_2 \ldots$ e $b = b_0, b_1 b_2 \ldots$, descartadas as que terminam por uma sequência infinita de noves (veja o porquê no Exemplo 1.21). Como $a < b$, existe um menor inteiro m tal que $a_i = b_i$ para $i = 0,\, 1,\, \ldots m-1$ e $a_m < b_m$ e, após o dígito a_m do número a, existe, no máximo, um número finito de noves, ou seja, existe um menor inteiro $n > m$ tal que $a_n < 9$. Agora considere o número $c = c_0, c_1 c_2 \ldots c_m \ldots c_n$ em que $c_i = a_i$ para $i = 0,\, 1,\, \ldots n-1$ e $c_n = a_n + 1$. Note que c é racional (pois equivale à soma dos racionais $c_i/10^i$, $i = 0,\, \ldots,\, n$) e $a < c < b$ ($a < c$ porque $a_i = c_i$ para $i = 0,\, \ldots,\, n-1$ e $a_n < c_n$; $c < b$ porque $c_i = a_i = b_i$ para $i = 0,\, 1,\, \ldots m-1$ e $c_m = a_m < b_m$). Construímos assim um racional que pertence a $(a,\, b)$.

$(a,\, b)$ **contém um irracional.** Agora, escolha o número irracional de sua preferência e considere apenas os seus dígitos (ou seja, desconsidere a parte inteira). Para fixar ideias, tome o número $\pi = 3,141592\ldots$ e descarte a parte inteira, ficando apenas com os dígitos $d_1 = 1,\; d_2 = 4,\; d_3 = 1,\; d_4 = 5,\; d_5 = 9,\; d_6 = 2,\; \ldots$. Defina, então, $c' = c_0, c_1 \ldots c_n d_1 d_2 \ldots$. Pelos mesmos argumentos do parágrafo anterior, $a < c' < b$; além disso, c' é um irracional, pois consiste na soma do racional c com o irracional $\{\pi\}/10^n$ (se necessário, veja a definição da função parte fracionária de um número real, $\{\cdot\} : \mathbb{R} \to \mathbb{Z}$, no Exemplo 1.5; e como exercício, prove que a soma de um racional com um irracional é um irracional[15]). Construímos assim um irracional que pertence a $(a,\, b)$.

Por nossas considerações iniciais, podemos afirmar que todo intervalo aberto $(a,\, b)$, e portanto qualquer intervalo não degenerado I, contém um racional e um irracional, o que completa a demonstração. ■

[15] Dica: caracterize os racionais como um quociente p/q de inteiros p e $q \neq 0$, e note que a soma de dois racionais $p/q + p'/q' = (pq' + p'q)/qq'$ é um racional; em seguida, demonstre por absurdo.

Demonstração 2.

Seja I um intervalo não degenerado e $(a,\, b) \subset I$ com $a < b$ um intervalo aberto qualquer em I. Vamos mostrar que existem um racional e um irracional em $(a,\, b)$.

$(a,\, b)$ **contém um racional.** Seja $n \in \mathbb{N}$ tal que $1/n < b - a$ e seja $I_m = [m/n,\, (m+1)/n]$, $m \in \mathbb{Z}$. Como $\mathbb{R} = \cup_{m\in\mathbb{Z}} I_m$, existe $m \in \mathbb{Z}$ tal que $a \in I_m$ e $a \notin I_{m+1}$. Se $a = m/n$, como $1/n < b - a$, então $a < (m+1)/n < b$; caso contrário, $m/n < a < (m+1)/n$, e como $1/n < b - a$, segue-se novamente que $a < (m+1)/n < b$. Em todo caso, o racional $(m+1)/n$ pertence a $(a,\, b)$ e portanto a I.

$(a,\, b)$ **contém um irracional.** Seja $n \in \mathbb{N}$ tal que $\sqrt{2}/n < b - a$ e considere agora $I_m = [m\sqrt{2}/n,\, (m+1)\sqrt{2}/n]$, $m \in \mathbb{Z}$. Utilizando os mesmos argumentos do parágrafo anterior (complete os detalhes!), concluímos que o irracional $(m+1)\sqrt{2}/n$ pertence a $(a,\, b)$ e portanto a I.

Isso completa a prova de que todo intervalo não degenerado I contém um racional e um irracional. ∎

Observação 1.12. Na segunda demonstração da Proposição 1.11, fizemos uso implícito da propriedade arquimediana do corpo dos números reais, i.e., a propriedade de que, dado qualquer $x > 0$ em $\mathbb{R}$, existe $n \in \mathbb{N}$ tal que $0 < 1/n < x$. ○

1.3.2.1 Domínio

Sejam $f : X \to Y$ uma função e X' um superconjunto de X (diz-se equivalentemente que X é um subconjunto de X' e denota-se esse fato por $X \subset X'$ ou $X' \supset X$).

Se X estiver contido propriamente em X' (o que significa que todo elemento de X pertence a X', mas há elementos em X' que não pertencem a X – denota-se $X \subsetneq X'$ ou $X' \supsetneq X$), a função f não precisa (e talvez não possa) estar definida para todos os valores de x em X'.

Por exemplo, se $f : (0,\, 1) \to \mathbb{R}$, $x \mapsto \sqrt{x}$, por algum motivo essa função não foi definida para todo $x \in \mathbb{R}^+$ $(\supsetneq (0,\, 1))$ e na verdade não pode ser definida para todo $x \in \mathbb{R}$ $(\supsetneq (0,\, 1))$ – qual seria, por exemplo, o número real igual a $f(-1)$?

Assim, o *domínio* X de f nos diz em quais valores $x \in X'$ a função pode ser aplicada, deixando de lado aqueles elementos $x' \in X'$ para os quais eventualmente $f(x')$ não existe. Isso nos previne o constrangimento de responder questões como a que está acima. Mais ainda, evita que cometamos erros elementares ao manipular cegamente fórmulas (ou, falando mais geralmente, leis de correspondência) que não definem um objeto matemático, como $y = \sqrt{-x}/\sqrt{x}$ ou $y = \log(\log(\operatorname{sen}(x)))$, se estivermos trabalhando com $y \in \mathbb{R}$. Note que ambas são uma fraude e que as funções correspondentes não existem (por quê?[16]).

Muitos textos apresentam perguntas do tipo "Qual o domínio da função $f(x) = \sqrt{x}$?". Estritamente falando, perguntas desse tipo não fazem sentido; em primeiro lugar, porque não basta apenas fornecer a lei de correspondência, é preciso também dizer qual o domínio e o contradomínio, caso contrário f ainda não é uma função; em segundo lugar, porque $f(x)$ é o valor que f assume em x, e não a função f em si; em terceiro lugar, porque mesmo o domínio e o contradomínio podem ser livremente escolhidos, contanto que a lei de correspondência possa ser definida. Neste exemplo, qualquer subconjunto de $\mathbb{R}^+$ poderia ser domínio de f. O correto seria perguntar "Qual o maior subconjunto $X \subset \mathbb{R}$ tal que a fórmula $f(x) = \sqrt{x}$ define uma função $f : X \to \mathbb{R}$?". A perguta incorreta é mais simples de formular e é difícil resistir à tentação de utilizá-la, mas se for feita assim, é preciso conhecer seu significado.

Fazendo a pergunta da maneira correta, podemos considerar o seguinte exemplo:

Exemplo 1.13. Qual o maior subconjunto $X \subset \mathbb{R}$ tal que a fórmula

$$f(x) = \frac{\log(x^2 - 1)}{x^2 - 5x + 6}$$

define uma função $f : X \to \mathbb{R}$? ◇

Solução. Por um lado, como o argumento do logaritmo precisa ser positivo, uma condição é que $x \notin [-1, 1]$; por outro lado, como o denominador não pode ser igual a zero, a outra condição é que não seja $x = 2$ nem $x = 3$. Portanto, $X = \mathbb{R} \backslash ([-1, 1] \cup \{2, 3\})$.◆

Dependendo do domínio, dizemos, por exemplo, que a nossa função $f : X \to Y$ é uma função de uma variável inteira se $X \subset \mathbb{Z}$, ou real se $X \subset \mathbb{R}$, ou complexa se $X \subset \mathbb{C}$, ou, mais exoticamente, quaternária[17] se $X \subset \mathbb{H}$, ou p-ádica[18] (p sendo um número primo) se $X \subset \mathbb{Q}_p$, e assim por diante. Embora essa nomenclatura seja comum, trata-se dos casos específicos em que o domínio é um conjunto numérico. Mas note que, na definição de função, nada foi dito sobre ser assim. Basta que X seja um *conjunto*. Por exemplo, X pode ser um subconjunto de pontos do plano, ou o conjunto de todas as circunferências, ou o conjunto de todas as funções reais, e assim por diante.

[16] Dica para a primeira fórmula: considere $x < 0$, $x = 0$ e $x > 0$ separadamente. Resposta para a segunda: para qualquer $x \in \mathbb{R}$, $\operatorname{sen} x \leqslant 1$, assim $\log(\operatorname{sen}(x)) \leqslant 0$; como o logaritmo de números reais não positivos não pode ser definido (pelo menos no contexto de funções reais), não existe $\log(\log(\operatorname{sen}(x)))$.

[17] O conjunto dos números quatérnios, comumente denotado por $\mathbb{H}$, foi descrito pela primeira vez, em 1843, pelo matemático irlandes William Rowan Hamilton (1805–1865) e utilizado por ele na mecânica.

[18] Os chamados números p-ádicos foram introduzidos por Kurt Wilhelm Sebastian Hensel (1861–1941) entre 1897 e 1899.

1.3.2.2 Contradomínio

Como vimos no item 3 da Definição 1.1, além do domínio, há outro conjunto associado a uma função: a sua *imagem*, que consiste de todos os valores possíveis que a função pode tomar quando aplicada nos elementos do domínio. Por exemplo, a imagem da função "log" consiste de todos os números reais; da função " sen ", de todos os números reais entre -1 e 1; da função "raiz quadrada de", de todos os números reais não negativos.

Ocorre, porém, que mesmo funções simples podem ter imagens muito complicadas. Por exemplo, a função cujo domínio é o conjunto dos inteiros não negativos e cuja lei de correspondência é $f(x) = \sqrt{x!}$ (raiz quadrada positiva) tem como imagem o conjunto de todas as raízes quadradas de fatoriais. É difícil dar uma caracterização melhor do que esta!

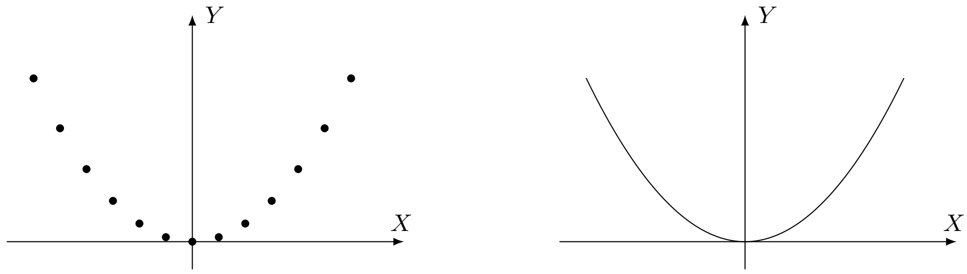
Por esse motivo, não nos interessa precisamente qual a imagem de uma função para que possamos defini-la. Basta apresentar um conjunto em que todos os valores do domínio são levados (e essa é a prática comum e mais útil). Qualquer conjunto Y que contenha os valores $f(x)$, para todo $x \in X$, serve para desempenhar esse papel. Tal conjunto é o *contradomínio* e quando desejamos caracterizar uma função f, devemos dizer algo como: f é uma função do domínio X no contradomínio Y (mais abreviadamente, que f é uma função de X em Y).

Analogamente ao que ocorre com a denominação das variáveis dependendo do domínio, em relação ao contradomínio, dizemos, por exemplo, que a nossa função $f : X \to Y$ é uma função inteira se $Y \subset \mathbb{Z}$, ou real se $Y \subset \mathbb{R}$, ou complexa se $Y \subset \mathbb{C}$, ou, mais exoticamente, quaternária se $Y \subset \mathbb{H}$, ou p-ádica se $Y \subset \mathbb{Q}_p$, e assim por diante. E assim como observado acima, para o domínio, embora essa nomenclatura seja comum, trata-se dos casos específicos em que o contradomínio é um conjunto numérico. Na definição de função, nenhuma restrição foi feita sobre o tipo de conjunto que o contradomínio deve ser. Basta que Y seja um *conjunto*. Por exemplo, Y pode ser o conjunto de todos os conjuntos de números primos, ou o conjunto de todos os pontos do plano, ou o conjunto de todos os conjuntos, etc.

1.3.2.3 Lei de correspondência

Assim como não se faz um bolo com apenas um ingrediente, não se define uma função somente com a lei de correspondência. Não basta, por exemplo, dizer algumas palavras mágicas como "seja a função $f(x) = x^2$". De fato, a lei de correspondência, por si só, não define uma função. Se escolhemos como domínio $\mathbb{Z}$ ou $\mathbb{R}$, ainda que o contradomínio seja o mesmo (por exemplo, $\mathbb{R}$) , obteremos funções distintas: a primeira levará todos os inteiros em seus quadrados, enquanto a segunda levará todos os números reais nos seus quadrados (veja a Figura 1.8).

Para algumas funções é fácil e natural utilizar a terminologia correta, por exemplo, costuma-se escrever $\operatorname{sen} : \mathbb{R} \to \mathbb{R}$ e $\log : \mathbb{R}^+ \to \mathbb{R}$ para as funções e $\operatorname{sen} x$ e $\log x$ para denotar os respectivos números reais que são os valores dessas funções num dado ponto x. Para outras, como as funções polinomias, é mais simples e cômodo dizer algo como "a função $x^2 - 6x + 9$" em vez da forma mais correta e pedante "a função $p : \mathbb{R} \to \mathbb{R}$ tal que $p(x) = x^2 - 6x + 9$ para todo $x \in \mathbb{R}$". Uma outra forma de expressar uma função de forma breve e cômoda (e também incompleta) quando estão subentendidos o domíno e o contradomínio é "a função $x \to y$" em vez da frase "a função que associa a cada x de X o elemento y de Y". Seja como for, quando você se deparar com ou for utilizar frases semelhantes a essas, tenha sempre em mente que trata-se apenas de uma forma mais curta de transmitir a ideia de qual a função em questão, mas que devem estar subentendidos o domínio e o contradomínio a fim de que o objeto em consideração (a função) esteja completo.



(a) $f_1 : \mathbb{Z} \to \mathbb{R}$, $x \mapsto x^2$.

(b) $f_2 : \mathbb{R} \to \mathbb{R}$, $x \mapsto x^2$.

Figura 1.8: Contradomínio e lei de correspondência iguais, e domínios diferentes.

Quanto à lei de correspondência em si, esta deve apenas cumprir um requisito de extrema importância: *associar a cada elemento $x \in X$ um único elemento $y = f(x) \in Y$*. Esse requisito é o cerne da questão. É importante que $f(x)$ seja definido de modo *único*, *sem* que haja *ambiguidade*. Por exemplo, extrair a raiz quadrada não define uma função a menos que especifiquemos se desejamos a raiz quadrada positiva ou negativa. Também é importante que a função esteja definida para *todo x* no domínio, *sem exceção*, a fim de que o fato de conhecermos o domínio possa nos dizer quando é seguro aplicar f.

1.3.2.4 Misturando todos os ingredientes

Como já dissemos, um bolo não se faz com apenas um ingrediente; da mesma forma, uma função não se define com apenas um dos objetos acima. É importante destacar que uma função consiste de três ingredientes: domínio, contradomínio e lei de correspondência $x \mapsto f(x)$. Sendo assim, mesmo quando dizemos simplesmente "a função f" em vez de "a função $f : X \to Y$, definida por $x \mapsto f(x)$", a rigor, estamos utilizando uma terminologia incompleta. Antes mesmo de ser fornecida a lei de correspondência, devem ficar subentendidos o domínio X e o contradomínio Y de f. Sem que qualquer um dos três ingredientes seja especificado, a função não existe.

Todas as considerações feitas até aqui motivam a seguinte definição de igualdade de funções.

Definição 1.14. *Sejam $f_1 : X \to Y$ e $f_2 : X' \to Y'$ funções. Dizemos que f_1 é igual a f_2 se $X = X'$, $Y = Y'$ e $f_1(x) = f_2(x)$ para todo $x \in X$.*

Em outras palavras, duas funções são iguais quando têm o mesmo domínio, o mesmo contradomínio e a mesma regra de correspondência. Consequentemente, $f_1 \neq f_2$ se $X \neq X'$ ou $Y \neq Y'$ ou $f_1(x) \neq f_2(x)$ para algum $x \in X \cap X'$. Note que quando temos $f_1 = f_2$ para funções com mesmo domínio e mesmo contradomínio, a igualdade $f_1(x) = f_2(x)$ deve ser satisfeita ponto a ponto, i.e., para todo $x \in X$, ao passo que para termos $f_1 \neq f_2$, ainda que seus domínios e contradomínios sejam os mesmos, é suficiente que $f_1(x) \neq f_2(x)$ para somente um $x \in X$.

Exemplo 1.15. Que função é $y = x^2$? Como acabamos de discutir, essa pergunta, por si só, não faz sentido sem que o domínio e o contradomínio sejam fornecidos. Além disso, dependendo do domínio e do contradomínio dados teremos diferentes funções (tantas quantas forem as combinações dos diferentes conjuntos escolhidos para os papéis de domínio e contradomínio para os quais a definição de função faça sentido), ou sequer teremos uma função. Considere, por exemplo, as seguintes tentativas de se definir uma função com essa fórmula:

$$\begin{aligned} f_1 : \mathbb{Z} &\to \mathbb{R} \\ x &\mapsto x^2 \end{aligned} \qquad \begin{aligned} f_2 : \mathbb{R} &\to \mathbb{R} \\ x &\mapsto x^2 \end{aligned} \qquad \begin{aligned} f_3 : \mathbb{R} &\to \mathbb{R}^- \\ x &\mapsto x^2 \end{aligned} \qquad \begin{aligned} f_4 : \mathbb{R} &\to \mathbb{R} \\ x &\mapsto \pm x^2 \end{aligned}$$

As duas primeiras, f_1 e f_2, são funções, porém, apesar de possuirem a mesma fórmula e contradomínio, são totalmente diferentes: f_1 está definida apenas para os números inteiros, equanto f_2 está definida para todos os números reais (veja os gráficos delas nas Figuras 1.8a e 1.8b, respectivamente). Já as duas últimas, f_3 e f_4, não são funções, pois violam, cada uma, alguma condição essencial da definição de função: f_3 não está definida em todo elemento do domínio (em particular, não o está para elemento algum) e f_4 faz corresponder a cada elemento do domínio mais de um elemento do contradomínio. ◇

Nos exemplos dados até aqui, o domínio e o contradomínio foram considerados conjuntos numéricos. Mas, note que a Definição 1.1 não faz nenhuma restrição nesse sentido. Basta que X e Y sejam *conjuntos*. Na verdade, é muito comum e natural surgirem regras, no sentido geral da Definição 1.1, em situações nas quais X e Y não são conjuntos numéricos. O estudante econtrará muitos exemplos ao longo de sua jornada. A seguir exibimos alguns.

Exemplo 1.16. Sejam X o conjunto de todas as figuras fechadas do plano e Y o conjunto de números reais $\mathbb{R}$. Se a cada $\varphi \in X$ fizermos corresponder o número real $A(\varphi) =$ área da figura φ, obteremos uma função $A : X \to \mathbb{R}$. $\diamond$

Exemplo 1.17. Seja $X = Y = M_2(\mathbb{R})$ o conjunto de todas as matrizes 2×2 reais. Se a cada $\mathcal{M} \in M_2(\mathbb{R})$ fizermos corresponder a matriz $T(\mathcal{M}) =$ transposta da matriz $\mathcal{M}$, obteremos uma função $T : M_2(\mathbb{R}) \to M_2(\mathbb{R})$. $\diamond$

Exemplo 1.18. Sejam X um subconjunto do plano, Y o plano Π e, para cada $x \in X$, seja $f(x) =$ o ponto 1 cm à direita de x. Obteremos assim uma função $f : X \to \Pi$ (conhecida como *translação de* X). $\diamond$

Exemplo 1.19. Sejam X o conjunto de todas as funções reais de uma variável real, Y o conjunto de todos os subconjuntos de $\mathbb{R}$ (ou seja, o conjunto das partes de $\mathbb{R}$, $\mathcal{P}(\mathbb{R})$) e, para cada função x, seja $f(x) =$ o domínio de x. Essa é uma função $f : X \to \mathcal{P}(\mathbb{R})$. $\diamond$

1.3.3 Natureza da “regra” que define uma função

Neste ponto, por ocasião do que acabamos de discutir no Exemplo 1.15 a respeito de f_3 e f_4, é interessante chamarmos a atenção para a natureza da regra à qual nos referimos na Definição 1.1. Nessa definição, não foi feita restrição alguma a essa regra (que pode ser totalmente arbitrária e tão complicada quanto quisermos, sendo essa generalidade uma vantagem em investigações teóricas). Ela deve estar sujeita a apenas duas condições, uma de *existência* e outra de *unicidade*, ou seja, ela deve:

1. estar definida em *todo* elemento do domínio (*existência*);

2. fazer corresponder a cada elemento do domínio *apenas um* elemento do contradomínio (*unicidade*).

Dito de outra maneira, mais simples e informal:

1’. não deve haver exceção: a regra deve fornecer o valor da função para *todo* elemento do domínio;

2’. não deve haver ambiguidade: a regra deve fazer corresponder um *único* valor no contradomínio para cada elemento do domínio.

Exemplo 1.20. Considere a tentativa de definir uma função $Inv : M_2(\mathbb{R}) \to M_2(\mathbb{R})$, estipulando que, para toda $\mathcal{M} \in M_2(\mathbb{R})$, $Inv(\mathcal{M}) = \mathcal{M}^{-1}$ é a inversa de $\mathcal{M}$. Ora, mas nem toda matriz é invertível, logo a regra dada não define uma função com domínio $M_2(\mathbb{R})$, pois há exceções. Entretanto, se considerarmos X o conjunto das matrizes de $M_2(\mathbb{R})$ equivalentes à matriz identidade I_2 (ou, se preferir, cujo determinante é diferente de zero), então a mesma regra define uma função $Inv : X \to M_2(\mathbb{R})$. ◇

Exemplo 1.21. Sejam $Y = \{a_0, a_1a_2 \ldots a_n \ldots ;\ a_0 \in \mathbb{N}\cup\{0\} \text{ e } a_n \in \{0,\ 1,\ \ldots,\ 9\},\ n \in \mathbb{N}\}$ o conjunto de todas as expressões decimais. Para cada $x \in \mathbb{R}$ vamos associar $f(x) =$ expressão decimal de x. Essa regra não define uma função $f : \mathbb{R} \to Y$ porque é ambígua: há números reais com duas representações decimais; de fato, se $0 \leqslant a_n \leqslant 8$, então as expressões decimais

$$a_0, a_1 \ldots a_n 999 \ldots \quad \text{e} \quad a_0, a_1 \ldots (a_n + 1)000 \ldots$$

definem o mesmo número real. Por exemplo, $2,718999\ldots = 2,719000\ldots$ e $0,999\ldots = 1,000\ldots$ (por quê?[19]). Se descartarmos as expressões decimais que terminam por uma sequência de noves, isto é, se considerarmos $Z = Y \setminus \{$expressões decimais que terminam por uma sequência de noves$\}$, então $f : \mathbb{R} \to Z$ será uma função. ◇

Uma consequência imediata de 2 (ou 2') é que não existem funções "plurívocas". De fato, por essa condição, se $f : X \to Y$ é uma função e $x = x'$ em X, então $f(x) = f(x')$ em Y.

1.4 Revisitando o conceito de função

Agora, devemos revisitar a Definição 1.1, pois ela é problemática. O problema está no uso das palavras "regra" e "associar", que não possuem definição matemática. Por exemplo, seja $X = Y = \mathbb{N}$ e considere

$$f(x) = \begin{cases} 1 & \text{se existe vida como a conhecemos fora da Terra} \\ 0 & \text{se não exisite vida como a conhecemos fora da Terra.} \end{cases}$$

Trata-se de uma função? O conteúdo do membro direito pode ser considerado uma regra? Será preciso esperar até encontrarmos vida extraterrestre ou provar que ela não existe para podermos considerar f como uma função?

Além disso, pensar sobre funções como "regras" é muito restritivo. As funções $f, g : \mathbb{R} \to \mathbb{R}$ definidas por $f(x) = 3x^2 + 2x - 1$ e $g(x) = \operatorname{sen} x$ são sem dúvida alguma dadas por "regras". Mas observe o gráfico da Figura 1.9. Ele certamente é o gráfico de uma função, mas qual a "regra" que define essa função?

[19] Dica: para provar que $a_0, a_1 \ldots a_n 999 \ldots = a_0, a_1 \ldots (a_n + 1)000 \ldots$, reescreva essas expressões na forma $\sum_{k=0}^{\infty} a_k 10^k$ e calcule a soma dos termos de uma progressão geométrica.

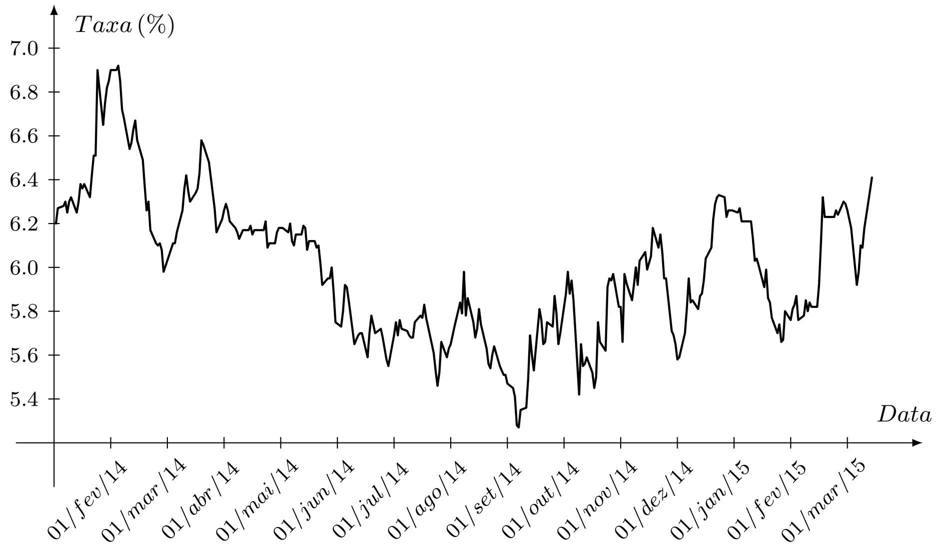


Figura 1.9: Taxa do título público *Tesouro IPCA+ 2019* ao longo do tempo[20].

Ainda mais restritivo seria pensar em funções como fórmulas. As funções f e g do parágrafo anterior certamente são dadas por fórmulas e muitas das quais nos deparamos no ensino fundamental e médio. Porém, há muitos exemplos de funções que não são dadas por fórmulas. A função da Figura 1.9 é um deles. Vejamos mais alguns exemplos.

Exemplo 1.22. Sejam X o conjunto das representações decimais dos números naturais, i.e., $X = \{1, 2, 3, \ldots\}$, Y o conjunto das representações romanas dos números naturais, i.e., $Y = \{I, II, III, IV, \ldots\}$, e vamos convencionar o seguinte para evitar ambiguidade na representação romana de um número: não repetiremos mais de três vezes consecutivas o mesmo algarismo romano, por exemplo, escreveremos IV ao invés de $IIII$. Agora, para cada $x \in X$ vamos associar $f(x) =$ numeral romano que representa o mesmo número que x (por exemplo, $1 \mapsto I$, $2 \mapsto II$, $3 \mapsto III$, $4 \mapsto IV$, $\ldots$). Como todo número tem (pelo menos em tese) uma representação romana e essa representação é única (respeitada a nossa convenção de não repetir mais de três vezes consecutivas o mesmo algarismo romano), f é uma função $f : X \to Y$. Qual a fórmula dessa função? ◇

A propósito desse último exemplo, note que todo numeral romano está associado a um numeral decimal e, além disso, não há dois numerais romanos associados ao mesmo numeral decimal. A função f é o que chamamos uma *bijeção* (estudaremos bijeções

[20]Fonte: Tesouro Direto: Balanço e Estatísticas. Disponível em: <http://www.tesouro.fazenda.gov.br/ tesouro-direto-balanco-e-estatisticas>. Acesso em: 12 mar. 2015.

em detalhes mais adiante). Ainda que tivéssemos escolhido Y como o conjunto das representações binárias ou o conjunto dos numerais cardinais, f continuaria sendo uma bijeção. Isso suscita uma discussão sobre a "representação" do número e sua "essência". Veja as diferentes representações do número "dezessete" na Figura 1.10.

Figura 1.10: Ideia do número "dezessete".

Nenhuma dessas representações em si é o próprio número, mas todas são ideias do número que chamamos "dezessete". Um número, assim como outros conceitos matemáticos, existe independentemente de sua representação. É uma entidade abstrata!

Exemplo 1.23. Sejam X o conjunto de todas as árvores genealógicas e $f : X \to \mathbb{N}$, $f(x) =$ número de pessoas que não têm filhos. Qual a fórmula que descreve tal função? Observe que se trocarmos X por um conjunto de árvores, f é a função que dá o número de folhas de uma árvore. $\diamond$

Exemplo 1.24. Seja $\mathcal{P}_n(\mathbb{R})$ o conjunto de todos os polinômios reais de uma variável real de grau menor do que ou igual a n e mais o polinômio identicamente nulo. Vamos definir $\varphi : \mathcal{P}_n(\mathbb{R}) \to \{0, 1\}$ da seguinte maneira: $\varphi(p) = 1$ se, e somente se, o polinômio p tem solução inteira, e $\varphi(p) = 0$ caso contrário. O leitor tem alguma sugestão de descrição algébrica para este exemplo? $\diamond$

Exemplo 1.25. Todo número natural maior do que 1 que não é primo é divisível por mais de dois números naturais, enquanto os números primos são divisíveis por exatamente dois números naturais: eles próprios e 1. Certamente, podemos considerar a função $T : \mathbb{N} \to \mathbb{N}$ que a cada n faz corresponder o número $T(n)$ de seus divisores. Para os primeiros números naturais, essa função é dada pela seguinte tabela:

n	1	2	3	4	5	6	7	8	9	10	11	12
$T(n)$	1	2	2	3	2	4	2	4	3	4	2	6

Note que não há uma fórmula aqui. É preciso algum algoritmo para calcular $T(n)$. $\diamond$

Por tudo o que foi dito até aqui, vimos que é preciso uma maneira alternativa para pensar sobre funções que evite palavras subjetivas como "regra" e "associar". Essa é a nossa motivação para redefinir esse conceito na seção a seguir.

1.4.1 Bourbaki: simplicidade e rigor

A definição alternativa que estamos prestes a enunciar é devida ao grupo de matemáticos Bourbaki[21], que lutou por mais rigor e simplicidade na Matemática, criando uma nova terminologia e novos conceitos ao longo do tempo.

Antes, no entanto, convém enunciarmos alguns conceitos preliminares, a saber, de *pares ordenados*, *produto cartesiano* e *relações*.

1.4.2 Pares ordenados

O conceito de par ordenado (x, y) formado por dois elementos genéricos $x, y \in X$ é intuitivo. Intuitivamente, trata-se de uma lista de dois elementos na qual um deles ocupa a "primeira" posição da lista (no caso, x) e o outro a "segunda" posição (no caso, y). Formalmente, há várias definições. Uma, em particular, que merece destaque pela simplicidade, rigor e precisão, é a de Kuratowski[22]:

Definição 1.26. *O par ordenado* (x, y) *é o conjunto* $\{\{x\}, \{x, y\}\}$.

Note que é possível recobrar a ordem pretendida no par (x, y) examinando-se os elementos de $\{\{x\}, \{x, y\}\}$: aquele contido em ambos, no caso $\{x\}$, é o que desempenha o papel de primeiro elemento. Mas o fato de o par ordenado ser um conjunto é irrelevante. Há outras maneiras de se abordar essa ideia[23], inclusive ela poderia ter sido introduzida como um conceito primitivo com as propriedades certas. É uma questão de escolha. O que realmente nos importa é provar, a partir dessa definição, a propriedade intrínseca que o par ordenado deve possuir e que faz ele merecer o nome que possui. Trata-se da seguinte proposição.

Proposição 1.27. *Se* (x, y) *e* (u, v) *são pares ordenados e* $(x, y) = (u, v)$, *então* $x = u$ *e* $y = v$.

[21] Nicolas Bourbaki é um pseudônimo de um grupo de matemáticos formado em 1935 com o objetivo de fundamentar toda a Matemática na teoria dos conjuntos. Em 1952 foi criada a *Associação dos colaboradores de Nicolas Bourbaki*, fundada por grandes nomes como Henri Cartan (1904–2008), Claude Chevalley (1909–1984), Jean Coulomb (1904–1999), Jean Delsarte (1903–1968), Jean Dieudonné (1906-1992), Charles Ehresmann (1905–1979), René de Possel (1905–1974), Szolem Mandelbrojt (1899–1983) e André Weil (1906–1998). Essa associação possui ainda hoje um gabinete na *École Normale Supérieure*, em Paris.

[22] Kazimierz Kuratowski (1896–1980), matemático polonês.

[23] Apenas para mencionar duas: a definição de Norbert Wiener (1894–1964), matemático estadunidense, segundo a qual $(x, y) = \{\{x, \varnothing\}, \{y\}\}$ (historicamente, essa definição antecede a de Kuratowski); e outra, baseada no Axioma da Escolha, que consiste em afirmar que cada par ordenado (x, y) é uma função $f : \{1, 2\} \to X \cup Y$ tal que $x = f(1) \in X$ e $y = f(2) \in Y$, ou seja, em que o primeiro membro do par é a imagem de 1 (por ser o primeiro) e o segundo a imagem de 2 (por ser o segundo); esta última definição, assim como a de Kuratowski, tem a vantagem de se generalizar facilmente.

Demonstração. Se $x = y$, então $(x, y) = \{\{x\}\}$; como $(u, v) = (x, y)$, segue que $\{u, v\} \in \{\{x\}\}$, logo $u = v = x$ e, portanto, $x = u$ e $y = v$. Por outro lado, se $x \neq y$, então (x, y) e (u, v) contêm exatamente um conjunto unitário, a saber, $\{x\}$ e $\{u\}$, logo $x = u$; como (x, y) e (u, v) também contêm exatamente um par não ordenado, a saber, $\{x, y\}$ e $\{u, v\}$, segue que esses conjuntos são iguais e, em particular, $y \in \{u, v\}$; e como $y \neq u$ (pois, do contrário, teríamos $x = u$ e $y = u$, logo $x = y$, contradição), segue-se que $y = v$, completando a demonstração. ■

Note que $\{x, y\} = \{y, x\}$ sempre, mas $(x, y) = (y, x)$ se, e somente se, $x = y$.

Apesar de existir a definição formal acima (ou, ainda, outras como mencionamos *en passant*), recomenda-se ao estudante fiar-se inicialmente na intuição por trás do conceito e na propriedade intrínseca (Proposição 1.27). Na verdade, após definirmos pares ordenados e demonstrarmos sua propriedade fundamental, estas definições (que de algum modo parecem artificiais) podem e, em certo sentido, devem ser esquecidas.

1.4.3 Produto cartesiano

Dados os conjuntos X e Y, surge uma questão natural: existe um conjunto que contém todos os pares ordenados (x, y) com $x \in X$ e $y \in Y$? A resposta é positiva e ela é dada muito elegantemente em [4]. Isso é o que motiva nossa próxima definição.

Definição 1.28. *Sejam os conjuntos X e Y. Definimos o* produto cartesiano[a] *de X e Y, denotado por $X \times Y$, como sendo o conjunto de todos os pares ordenados (x, y), sendo $x \in X$ e $y \in Y$. Simbolicamente,*

$$X \times Y := \{(x, y);\, x \in X,\, y \in Y\}.$$

[a] Assim denominado em honra a René Descartes (1596–1650). O adjetivo *cartesiano* provém da latinização de seu nome como *Cartesius*.

Observe que, em geral, $X \times Y \neq Y \times X$. (Por quê?) O exemplo abaixo nos fornece uma resposta.

Notação 1.29. É usual representarmos o produto cartesiano de um conjunto X com ele mesmo, ou seja, $X \times X$, como X^2. ○

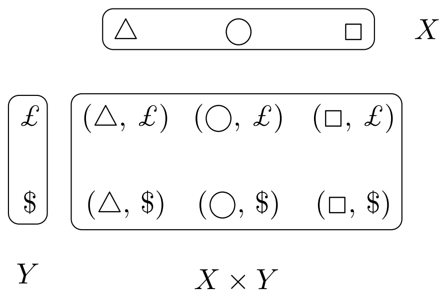
Exemplo 1.30. Seja $X = \{\triangle, \bigcirc, \square\}$ e $Y = \{\pounds, \$\}$. Então,

$$X \times Y = \{(\triangle, \pounds), (\triangle, \$), (\bigcirc, \pounds), (\bigcirc, \$), (\square, \pounds), (\square, \$)\},$$

$$Y \times X = \{(\pounds, \triangle), (\$, \triangle), (\pounds, \bigcirc), (\$, \bigcirc), (\pounds, \square), (\$, \square)\}.$$

$X \times Y \neq Y \times X$ porque $Y \times X$ contém, dentre outros, o elemento $(\pounds, \triangle)$, que não pertence a $X \times Y$. ◇

Quando os conjuntos X e Y são ambos finitos (adiante, definiremos esse conceito, mas, por ora, vamos supor que sabemos do que se trata, fiando-se na nossa intuição), digamos, com m e com n elementos, respectivamente, então o conjunto $X \times Y$ possui exatamente mn elementos (também provaremos isso na ocasião apropriada). Uma forma de enxergar isso é representar o produto cartesiano $X \times Y$ dos conjuntos X e Y finitos como um quadro. No caso do Exemplo 1.30 ficaria assim:



Exemplo 1.31. Sejam AB e CD segmentos de reta. O produto cartesiano $AB \times CD$ pode ser interpretado como um retângulo da seguinte maneira: tome ambos segmentos perpendiculares e considere cada $(x, y) \in A \times B$ sendo representado pelo ponto P, obtido pela interseção das perpendiculares aos segmentos tiradas pelos pontos $x \in AB$ e $y \in CD$. O que estamos dizendo tem o aspecto da Figura 1.11. $\diamond$

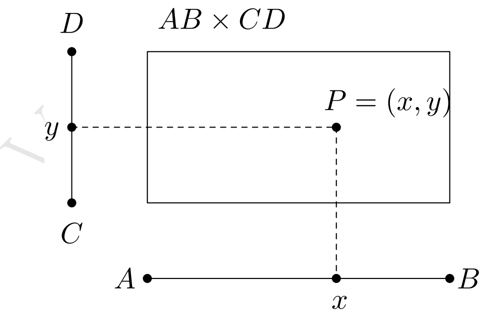


Figura 1.11: Produto cartesiano dos segmentos AB e CD.

Exemplo 1.32. Analogamente ao exemplo anterior, se considerarmos agora uma circunferência γ e um segmento de reta AB, o produto cartesiano pode ser interpretado como um cilindro da seguinte maneira: considere AB perpendicular ao plano que contém γ e cada $(x, y) \in \gamma \times AB$ representado pelo ponto P obtido pela interseção da perpendicular a γ por x com o plano perpendicular a AB por y (Figura 1.12). $\diamond$

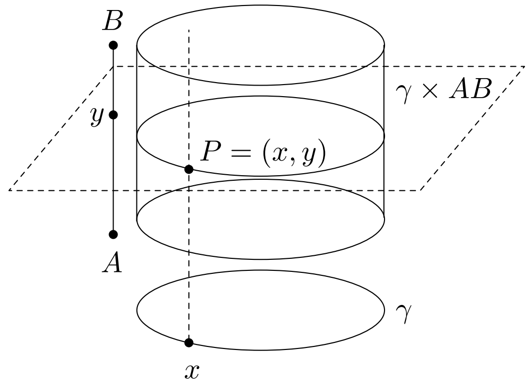


Figura 1.12: Produto cartesiano da circunferência γ e do segmento AB.

1.4.4 Relações

Definição 1.33. *Sejam os conjuntos X e Y e o produto cartesiano $X \times Y$. Uma* relação binária*, ou simplesmente* relação*, entre X e Y é um subconjunto R de $X \times Y$.*

Além disso,

1. se x está relacionado com y segundo R, expressa-se o fato $(x,\, y) \in R$ escrevendo-se
$$x\, R\, y;$$
2. define-se como domínio de R o conjunto
$$Dom(R) := \{x \in X;\ (x,\, y) \in R \text{ para algum } y \in Y\};$$
3. define-se como imagem de R o conjunto
$$Im(R) := \{y \in Y;\ (x,\, y) \in R \text{ para algum } x \in X\}.$$

Note que $Dom(R) \subset X$ e que $Im(R) \subset Y$.

Exemplo 1.34. Sejam X o conjunto de homens vivos e Y o conjunto de mulheres vivas e seja $R \subset X \times Y$ o conjunto $R := \{(x,\, y);\ x \text{ é casado com } y\}$. Assim, R representa uma relação (de casamento) entre homens e mulheres. O domínio dessa relação é o conjunto de todos os homens casados e sua imagem é o conjunto de todas as mulheres casadas. ◇

Outros exemplos bem conhecidos são os seguintes.

Exemplo 1.35. Defina $< := \{(x, y) \in \mathbb{R} \times \mathbb{R};\ y - x \text{ é positivo}\}$. Essa é a relação entre $\mathbb{R}$ e $\mathbb{R}$ que nos permite escrever $x < y$, ou seja, trata-se da relação "menor do que". ◇

Exemplo 1.36. Seja $\mathcal{R}$ o conjunto de todas as retas e $\mathcal{P}$ o conjunto de todos os planos do espaço. Defina $\| := \{(r, \Pi) \in \mathcal{R} \times \mathcal{P};\ r \text{ é paralela a } \Pi\}$. Essa é a relação entre $\mathcal{R}$ e $\mathcal{P}$ que nos permite escrever $r \parallel \Pi$, ou seja, trata-se da relação de "paralelismo entre retas e planos". ◇

Há um tipo de relação muito notável e útil, que ocorre em um sem-número de construções matemáticas importantes como você constatará em vários exemplos neste e em outros textos. Trata-se de uma noção muito antiga na Matemática, encontrada até mesmo (não sob o nome atual) nos *Elementos* de Euclides[24]. Estamos nos referindo à definição que se segue.

Definição 1.37. *Uma relação $E \subset X \times X$ é chamada uma* relação de equivalência *num conjunto X se os seguintes quesitos forem satisfeitos:*

1. *Reflexividade: $(x, x) \in E$, para todo $x \in X$;*

2. *Simetria: $(x, y) \in E$ implica $(y, x) \in E$;*

3. *Transitividade: $(x, y) \in E$ e $(y, z) \in E$ implica $(x, z) \in E$.*

Se o par (x, y) pertence a uma relação de equivalência E, então x e y são ditos *equivalentes segundo E*.

Notação 1.38. A notação mesofixa $x \overset{E}{\sim} y$, ou $x \sim_E y$, é frequentemente empregada para indicar que dois elementos são equivalentes segundo uma relação de equivalência E dada. Quando essa relação está subentendida, escreve-se simplesmente $x \sim y$. ○

Com a última dessas notações, as propriedades definidoras de uma relação de equivalência ficam da seguinte forma:

1. Reflexividade: $x \sim x$, para todo $x \in X$;

2. Simetria: $x \sim y$ implica $y \sim x$;

3. Transitividade: $x \sim y$ e $y \sim z$ implica $x \sim z$.

Exercício 1.39. Considere a seguinte relação no conjunto $\mathbb{R}$ dos números reais:

$$E := \{(x, y) \in \mathbb{R} \times \mathbb{R};\ x - y \in \mathbb{Z}\},$$

[24] Euclides de Alexandria (c. 325 a.C. – c. 265 a.C.), matemático grego, famoso por seu tratado em Geometria, *Os Elementos*, que influenciou o desenvolvimento da matemática ocidental por mais de 2000 anos.

em que $\mathbb{Z}$ é o conjunto dos números inteiros. Prove que E é uma relação de equivalência. $\diamond$

Exercício 1.40. Fixado um inteiro positivo n, os possíveis restos da divisão euclidiana de números inteiros são $0, 1, \ldots, n-1$. Dizemos que os números inteiros x e y são *congruentes módulo* n (e escrevemos $x \equiv y \pmod{n}$ ou simplesmente $x \equiv y$ se n já estiver subentendido) se eles têm o mesmo resto da divisão euclidiana por n, i.e., se $y - x$ é um múltiplo de n. Mostre que a congruência é uma relação de equivalência. $\diamond$

Toda relação de equivalência definida num conjunto não vazio X permite definir uma família de subconjuntos especiais de X (chamados classes de equivalência), que, como veremos adiante (quando falarmos especificamente sobre famílias) particiona o conjunto X, i.e., todo elemento de X pertence a uma e somente uma dessas classes.

Definição 1.41. *Seja E uma relação de equivalência em X. Para todo $x \in X$ podemos definir o conjunto:*

$$E(x) := \{x' \in X;\, (x, x') \in E\},$$

chamado de classe de equivalência de x *(pela relação de equivalência E).*

Notação 1.42. Além da notação $E(x)$ para a classe de equivalência de x, também encontra-se frequentemente na literatura $[x]_E$ ou $[x]$ (quando E está subentendido). $\circ$

Exemplo 1.43. Na Geometria Analítica, aprendemos que dois segmentos orientados são equipolentes se ambos são nulos ou se possuem igual direção, sentido e comprimento. A equipolência é na verdade uma relação de equivalência no conjunto de todos os segmentos orientados do espaço. (Prova?) Por sua vez, o conjunto de todos os segmentos orientados equipolentes a um segmento orientado (A, B) dado é uma classe de equivalência, especialmente denominada *classe de equipolência de* (A, B). Essa classe, geralmente representada por $\overrightarrow{AB}$, é o que se denomina um *vetor*, sendo (A, B) um de seus representantes. Também se representa essa classe de equipolência (vetor) por letras latinas minúsculas, por exemplo $\vec{v}$, para não se fazer referência a algum de seus representantes, dentre os quais (A, B) é um deles. $\diamond$

Proposição 1.44. *Seja E uma relação de equivalência em $X \neq \emptyset$. As seguintes afirmações são equivalentes:*

(a) $x \sim y$;

(b) $[x] = [y]$;

(c) $[x] \cap [y] \neq \emptyset$.

Demonstração. ((a) $\Rightarrow$ (b)) Se $z \in [x]$, então $x \sim z$. Como $x \sim y$ e E é simétrica, então $y \sim x$. Como também E é transitiva, $y \sim z$. Assim, $z \in [y]$. Logo, $[x] \subset [y]$. Trocando os papeis de x e y, chega-se a $[y] \subset [x]$. Portanto, $[x] = [y]$.

((b) $\Rightarrow$ (c)) Suponha que $[x] = [y]$. Como $x \in [x]$, então $x \in [y]$, logo $x \in [x] \cap [y]$ e portanto $[x] \cap [y] \neq \emptyset$.

((c) $\Rightarrow$ (a)) Se $[x] \cap [y] \neq \emptyset$, então existe z tal que $z \in [x]$ e $z \in [y]$, ou seja, $x \sim z$ e $y \sim z$. Por simetria, $z \sim y$. Por transitividade, como $x \sim z$ e $z \sim y$, segue-se que $x \sim y$. ∎

Corolário 1.45. *Seja E uma relação de equivalência em $X \neq \emptyset$. Uma condição necessária e suficiente para que $[x] \neq [y]$ é $[x] \cap [y] = \emptyset$.*

Adiante, quando tratarmos de famílias de funções, esses conceitos serão importantes na demonstração de que partições induzem relações de equivalência e classes de equivalência particionam um conjunto.

1.4.5 Funções

De posse do conceito de relação, estamos aptos a enunciar a definição alternativa de função que prometemos.

Definição 1.46. *Sejam X e Y conjuntos e f uma relação entre X e Y. Então, a relação f é dita ser uma* função *de X em Y se:*

1. *$Dom(f) = X$ e*
2. *se $(x, y) \in f$ e $(x, y') \in f$ só for possível caso $y = y'$.*

Em outras palavras, a primeira condição impõe que *cada* elemento de X esteja associado a algum elemento de Y (ou seja, não deve haver exceção – condição de existência) e a segunda condição, de uma maneira formal, impõe que a cada elemento x de X a função associe um e apenas um elemento y de Y que faz o papel de segundo elemento do par ordenado (x, y) (ou seja, não deve haver ambiguidade – condição de unicidade). Esse segundo elemento associado a x pela função f é mais convenientemente denotado por $f(x)$, ao invés de $(x, y) \in f$ ou da notação mesofixa (ou infixa) $x\,f\,y$. Assim, por essa perpectiva (Bourbaki), uma função f nada mais é do que um conjunto qualquer de pares ordenados sujeitos apenas à condição:

$$(x, y) \in f, (x', y') \in f, x = x' \Rightarrow y = y',$$

ou seja, f é o conjunto $\{(x, y) \in X \times Y;\, y = f(x)\} = \{(x, f(x));\, x \in X\}$. Note que a Definição 1.46 está livre dos conceitos subjetivos de “associar” e “regra”.

Como dissemos na Observação 1.2, tanto o domínio quanto o contradomínio de uma função pode ser o conjunto vazio. Note que a Definição 1.46 não faz nenhuma restrição nesse sentido, o que, como já comentamos naquela observação, é uma vantagem do ponto de vista teórico, pois podemos tratar de funções arbitrárias sem nos preocupar em provar que X e Y não são vazios. Resta apenas provar que:

1. se $X = \emptyset$ (sendo Y vazio ou não), então existe uma função $f : X \to Y$ e $f = \emptyset$; e

2. para $X \neq \emptyset$, se $Y = \emptyset$, então não existe uma função $f : X \to Y$.

Certamente, existe uma (única) relação f entre $X = \emptyset$ e Y, que é o próprio conjunto vazio (único subconjunto de $\emptyset \times Y = \emptyset$) e o domínio dessa relação (vazia), por conseguinte, também é o conjunto vazio, o que significa que $Dom(f) = X$. Para que f seja uma função de X em Y, basta que ela cumpra a condição 2 da Definição 1.46, ou seja, que $(x, y), (x, y') \in f$ implique $y = y'$; mas isso segue por vacuidade, afinal, para que fosse falso, seria necessário existir um $x \in Dom(f)$ para o qual $(x, y), (x, y') \in f$ com $y \neq y'$, impossível, uma vez que $Dom(f) = \emptyset$. Portanto f é uma função de X em Y. Isso completa a prova da afirmação 1 acima.

Analogamente, no caso da afirmação 2, existe apenas uma relação f entre $X \neq \emptyset$ e $Y = \emptyset$: a relação vazia, cujo domínio é, por conseguinte, o conjunto vazio. No entanto, como, por hipótese, $X \neq \emptyset$, segue-se que $Dom(f) \neq X$ (descumprindo assim a condição 1 da Definição 1.46) e, portanto, f não é uma função de X em Y. Como f é a única relação entre esses conjuntos, não existe uma função de X em Y. Isso completa a prova.

Denotanto o conjunto de todas as funções de X em Y por Y^X, acabamos de mostrar que $Y^{\emptyset} = \{\emptyset\}$, sendo Y vazio ou não, e que se X não é vazio, então $\emptyset^X = \emptyset$.

Exemplo 1.47. Sejam $X = \{1, 2, 3, 4, 5\}$ e $Y = \{a, b, c, d, e\}$. Façamos

$$f = \{(1, b), (2, a), (3, c), (4, b), (5, a)\}.$$

Trata-se de uma função, pois tal f cumpre as condições da Definição 1.46. Você pode achar mais conveniente olhar para essa função da seguinte forma: $f(1) = b$, $f(2) = a$, $f(3) = c$, $f(4) = b$, $f(5) = a$. Note que cada elemento do domínio está "associado" a um, e apenas um, elemento do contradomínio. No entanto, a definição de função nos permite associar dois elementos distintos do domínio a um mesmo elemento do contradomínio (veja, por exemplo, que $f(1) = f(4) = b$), bem como nos permite não utilizar todos os elementos do contradomínio: observe que $Im(f) = \{a, b, c\}$ ao passo que o contradomínio é $Y = \{a, b, c, d, e\}$. ◇

Exemplo 1.48. Sejam X e Y conjuntos. Considere uma função $f : X \times Y \to X$ escrevendo $f(x, y) = x$. Esse é um exemplo do que chamamos de uma função de duas variáveis; note que, a rigor, ela é como as demais funções que vimos até o presente, pois, na verdade, $(x,\, y)$ é um elemento do conjunto $X \times Y$, e portanto, quando varia nesse conjunto, nada mais é do que uma só variável; note, portanto, que a rigor deveríamos escrever $f((x,\, y))$ ao invés de $f(x,\, y)$, mas ninguém faz assim. Voltando ao nosso exemplo, f é chamada de *projeção* de $X \times Y$ sobre X. Analogamente, $g : X \times Y \to Y$, $g(x,\, y) = y$ é a projeção de $X \times Y$ sobre Y. $\diamond$

Exemplos de funções no sentido preciso da Definição 1.46 são abundantes, tanto na matemática quanto nas ciências ou na vida comum. Tudo o que é preciso fazer é buscar informação, não necessariamente numérica, mas na forma tabulada. Um exemplo é uma agenda de aniversários; os argumentos da função neste caso são os aniversariantes e os valores são as datas de aniversário deles.

O fato de uma função ser um conjunto possui uma vantagem. Se conhecemos a função, conhecemos o conjunto e, melhor ainda, se conhecemos o conjunto, conhecemos a função. Assim, por exemplo, se nos for apresentado o conjunto de pares ordenados que correspondem aos aniversários das pessoas, mesmo que tenhamos esquecido o que significa aniversário, seríamos capazes de dizer se uma pessoa x faz ou não faz aniversário numa determinada data y: bastaria verificar se o par ordenado $(x,\, y)$ pertence ou não ao conjunto. Essa observação serve para relações em geral.

Outra implicação de uma função ser um conjunto é que ela não FAZ algo, como podem sugerir os verbos listados no item 2 logo após a Definição 1.1, mas ela apenas É. Por esse motivo, tal definição deixa alguns estudiosos insatisfeitos e essa insatisfação se reflete nos diferentes usos dos seguintes vocábulos: *função* fica reservada para o objeto indefinido que é de algum modo ativo e o conjunto de pares ordenados que chamamos de função passa a ser denominado o *gráfico* da função.

1.5 Gráficos de funções

O *gráfico* de uma função $f : X \to Y$ é o subconjunto $\Gamma(f)$ do produto cartesiano $X \times Y$ formado por todos os pares ordenados $(x,\, y)$ tais que x é um ponto qualquer de X e $y = f(x)$. Mais formalmente:

Definição 1.49. *Seja $f : X \to Y$ uma função. Definimos o gráfico de f, denotado por $\Gamma(f)$, como sendo o seguinte subconjunto de $X \times Y$:*

$$\Gamma(f) := \{(x,\, y) \in X \times Y;\, y = f(x)\} = \{(x,\, f(x));\, x \in X\}.$$

Note que, de acordo com essa definição, um subconjunto $\Gamma \subset X \times Y$ é o gráfico de uma função se, e somente se, para todo $x \in X$ existe um, e apenas um, $y \in Y$ tal que $(x,\, y) \in \Gamma$.

Há quem prefira relegar essa formal e estática definição apenas para o gráfico de uma função e pensar em função como um objeto dinâmico, que transforma, associa, dá a ideia de dependência ou o do resultado de um movimento. O ônus dessa escolha foi discutido na Seção 1.4. Não há problema em se fazer assim. Essa é a prática comum dos matemáticos e dos usuários da Matemática. O estudante apenas precisa ter em mente que é possível (e até mesmo desejável quando se trata de fazer Matemática mais avançada) possuir meios de se evitar termos subjetivos como os da Definição 1.1, mas, no dia-a-dia, ele pode fiar-se na usual, humana e intuitiva compreensão desse conceito. Afinal, exceto os lógicos, cuja intenção seja mostrar que todas as noções matemáticas se reduzem, em última análise, à ideia de conjunto, quem mais pensaria numa translação ou numa rotação como um conjunto de pares ordenados?

Observação 1.50 (importante). Ao estudarmos funções, é muito comum utilizarmos desenhos sempre que possível, como veremos a seguir. No entanto, os desenhos devem ser considerados apenas como um instrumento de auxílio à intuição e à linguagem, nunca como dados para demonstrações. ○

1.5.1 Diagrama de setas

Observe que a Definição 1.49 é geral o suficiente para admitir que utilizemos quaisquer conjuntos no lugar de X e Y, não necessariamente conjuntos numéricos. Ela admite, por exemplo, que X e Y sejam conjuntos de figuras do plano, ou conjuntos de matrizes, ou conjuntos de funções, ou o que desejarmos. Neste caso, não é possivel desenhar o gráfico como fizemos, por exemplo, nas Figuras 1.1 ou 1.2.

Por razões evolutivas, nosso cérebro só é capaz de produzir e desenvolver imagens em uma, duas ou três dimensões. Assim, se $\Gamma(f) \not\subset \mathbb{R}^2$ nem $\Gamma(f) \not\subset \mathbb{R}^3$, não é possível desenharmos o gráfico da função, sendo necessário pensarmos em maneiras alternativas de representarmos alguns conceitos.

Uma maneira alternativa e útil de pensar sobre conceitos envolvendo o próprio conceito de função é utilizar um *diagrama de setas*. (Ver Figura 1.13.)

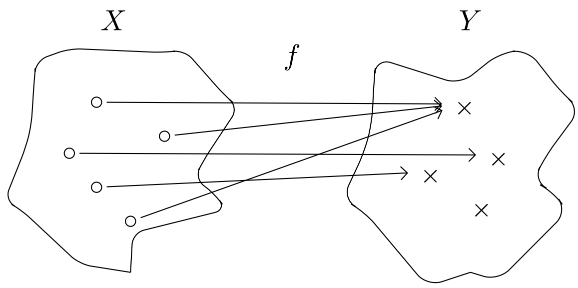


Figura 1.13: Diagrama de uma função $f : X \to Y$.

Embora não seja parte da teoria, essa imagem evocativa ajuda a clarear as ideais acerca de muitos conceitos envolvendo funções, além da própria definição de função: sobrejetividade, injetividade, bijetividade, composição e inversas de funções são alguns deles (que adiante veremos com detalhes).

No diagrama da Figura 1.13, note que todo elemento de X está associado a um, e somente um, elemento de Y, logo f é uma função. As setas nessa figura representam a "regra" que nos diz, para cada $x \in X$, qual o valor de $f(x)$.

1.5.2 Representação geométrica

1.5.2.1 Plano numérico $\mathbb{R}^2$

Quando se introduz no plano Π um par de eixos ortogonais OX e OY, que se intersectam no ponto O, chamado a *origem* do sistema de coordenadas, os pontos P desse plano passam a ser representados por suas *coordenadas* x, y, que são números reais. Tem-se $x > 0$ (resp. $x < 0$) se P está à direita (resp. esquerda) do eixo OY; analogamente, $y > 0$ (resp. $y < 0$) se P está acima (resp. abaixo) do eixo OX.

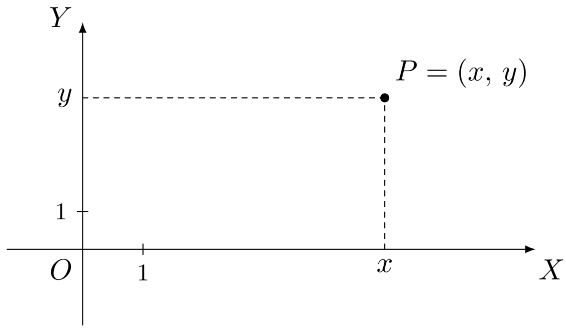


Figura 1.14: Coordenadas retangulares.

As coordenadas x e y são chamadas, respectivamente, de *abcissa* (ou *abscissa*) e *ordenada* de P e escreve-se $P = (x, y)$ ou $P(x, y)$. A primeira é a coordenada do pé da perpendicular baixada de P sobre o eixo OX; a segunda é a coordenada do pé da perpendicular baixada de P sobre o eixo OY. Diz-se, então, que (x, y) é o par de *coordenadas* do ponto P relativamente ao *sistema de eixos ortogonais* OXY. A correspondência entre pontos do plano e pares ordenados de números reais se chama *sistema de coordenadas retangular* ou *sistema de coordenadas cartesiano*[25] e, por conseguinte, as coordenadas x, y são chamadas de *coordenadas retangulares* ou *coordenadas cartesianas.* Cada uma das quatro regiões em que o plano fica dividido pelos eixos OX e OY chama-se um *quadrante* e é caracterizado pelos sinais das coordenadas de seus pontos:

[25]Devido a René Descartes (1596–1650), que, em 1637, desenvolveu as ideias para esse sistema na sua obra *Discours de la méthode.*

primeiro quadrante ($x \geqslant 0$ e $y \geqslant 0$); segundo quadrante ($x \leqslant 0$ e $y \geqslant 0$); terceiro quadrante ($x \leqslant 0$ e $y \leqslant 0$); e quarto quadrante ($x \geqslant 0$ e $y \leqslant 0$).

O que torna legítima essa correspondência entre pontos do plano e pares ordenados de números reais, ou seja, elementos de $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$, é a existência de uma bijeção entre Π e $\mathbb{R}^2$, isto é, existe uma função $f : \Pi \to \mathbb{R}^2$, $P \mapsto (x,\, y)$, que é bijetiva (para a definição de função bijetiva, veja a Seção 1.6.3; para uma demonstração dessa afirmação, veja Apêndice B).

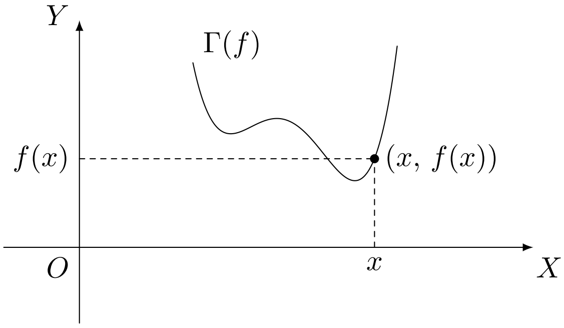


Figura 1.15: Gráfico de uma função $f : X \subset \mathbb{R} \to \mathbb{R}$.

Devido à existência dessa bijeção, costuma-se dizer que $\mathbb{R}^2$ é o modelo aritmético-algébrico do plano enquanto Π é o modelo geométrico de $\mathbb{R}^2$. Assim, dada, por exemplo, uma curva definida por alguma propriedade geométrica, pode-se procurar uma maneira de representar essa curva analiticamente exprimindo-se uma das coordenadas em função da outra, por exemplo, y em função de x: $y = f(x)$ (neste caso, estamos olhando para $\mathbb{R}^2$ como modelo para Π). Reciprocamente, se ao invés de partirmos de uma curva definida geometricamente, considerarmos uma função $f : X \subset \mathbb{R} \to \mathbb{R}$, com $y = f(x)$, podemos representar graficamente a dependência de y em relação a x no sistema de coordenadas OXY (Figura 1.15): para cada abcissa x, basta determinarmos a ordenada correspondente a $y = f(x)$; o gráfico de f, $\Gamma(f)$, é o conjunto de pontos $(x,\, f(x))$ assim obtidos (neste caso, estamos olhando para Π como modelo para $\mathbb{R}^2$).

A ideia é tirar vantagem desse caminho de mão dupla que existe entre a Aritmética e a Álgebra de um lado e a Geometria de outro, a fim de obtermos um melhor entendimento das funções reais que iremos estudar. Desse ponto de vista, olharemos para $\mathbb{R}^2$ como um plano, chamaremos seus elementos $(x,\, y)$ de pontos, e usaremos essa linguagem e os resultados da Geometria para nos ajudar nessa tarefa.

Nesse sentido, a primeira questão que se coloca é: se $P_1 = (x_1,\, y_1)$ e $P_2 = (x_2,\, y_2)$, como a distância entre P_1 e P_2, $d(P_1,\, P_2)$, pode ser expressa em termos das coordenadas desses pontos?

A resposta é dada pelo Teorema de Pitágoras. Seja $P_3 = (x_2,\, y_1)$ um ponto auxiliar (Figura 1.16). Como P_1 e P_3 têm mesma ordenada, o segmento P_1P_3 é paralelo ao eixo

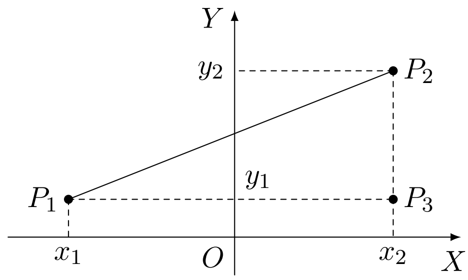


Figura 1.16: Distância entre dois pontos no plano.

OX; analogamente, P_2P_3 é paralelo a OY. Logo, $P_1P_2P_3$ é um triângulo retângulo cujos catetos são P_1P_3 e P_2P_3 e medem respectivamente $|x_1 - x_2|$ e $|y_1 - y_2|$ (veja Seção C.1). Portanto:

$$d(P_1,\, P_2)^2 = (x_1 - x_2)^2 + (y_1 - y_2)^2, \tag{1.5.1}$$

ou seja,

$$d(P_1,\, P_2) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}. \tag{1.5.2}$$

Exemplo 1.51. Sejam $Q = (u,\, v)$ um ponto dado do plano e $r > 0$ um número real positivo. Uma *circunferência* $\mathcal{C}$ de centro Q e raio r (Figura 1.17) é, por definição, o conjunto de pontos $P = (x,\, y)$ tais que $d(P,\, Q) = r$, ou seja, pela equação (1.5.2),

$$\mathcal{C} = \{(x,\, y) \in \mathbb{R}^2;\, (x - u)^2 + (y - v)^2 = r^2\},$$

e

$$(x - u)^2 + (y - v)^2 = r^2 \tag{1.5.3}$$

é chamada a equação da circunferência $\mathcal{C}$.

Um *disco*[26] $\mathcal{D}$ de centro Q e raio r (Figura 1.17), por sua vez, é o conjunto de pontos $P = (x,\, y)$ tais que $d(P,\, Q) \leqslant r$, ou seja,

$$\mathcal{D} = \{(x,\, y) \in \mathbb{R}^2;\, (x - u)^2 + (y - v)^2 \leqslant r^2\}. \qquad \diamond$$

Nos cursos mais elementares, a construção do gráfico de funções $f : X \subset \mathbb{R} \to \mathbb{R}$ no plano munido do sistema de coordenadas OXY é feita por meio da marcação de uma rede suficientemente densa de pontos $P_i = (x_i,\, y_i)$, em que $y_i = f(x_i)$ $(i = 0,\, 1,\, 2,\, \ldots)$, e pela união desses pontos por uma linha cuja aspecto deve considerar a posição dos

[26] Como bem observa o autor de [10], a palavra *círculo* é ambígua, pois ora significa a circunferência, ora o disco que tem essa circunferência como fronteira. O mesmo ocorre com as palavras polígono, elipse, triângulo, quadrado, e assim por diante. Sempre que utilizadas, deve-se explicar o seu sentido, a fim de se evitar mal-entendidos.

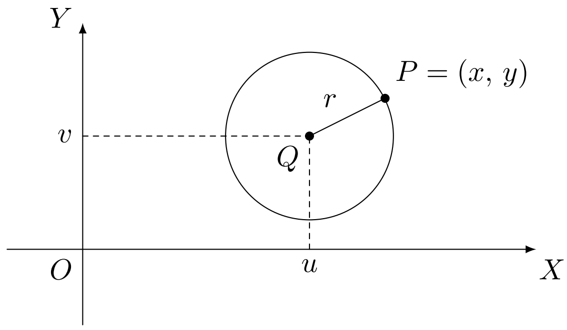


Figura 1.17: Circunferência e disco correspondente.

pontos intermediários. Na verdade, esse é o procedimento utilizado pelos computadores, porém com a capacidade de obter uma rede com milhares de pontos numa fração do tempo que um ser humano costuma levar para obter apenas um. O estudande com senso crítico mais aguçado notará que este método é precário e incompleto, afinal, por maior que seja o valor de i acima, ele ainda será finito, e portanto o será o número de pontos P_i. Que garantias teremos de que, ao ligá-los, a linha que os conecta não deveria se comportar de maneira diferente da qual a traçamos? As técnicas de que dispomos não são suficientes para responder essa pergunta. É necessário recorrer ao Cálculo Diferencial para se aprender técnicas mais sofisticadas e eficientes para esboçar o gráficos de certas funções.

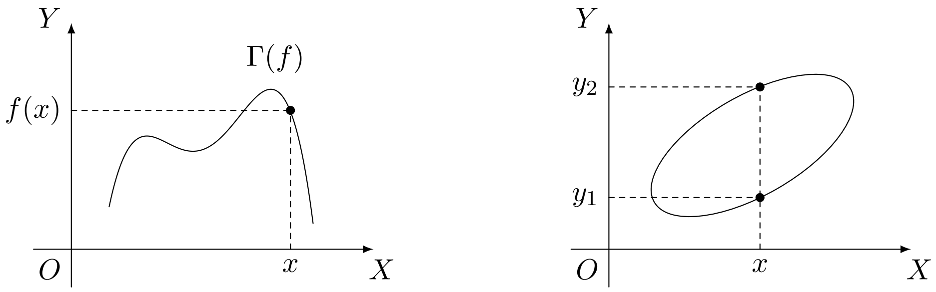
1.5.2.2 Representação paramétrica

Como já discutimos anteriormente (logo após a Definição 1.49), para que $\Gamma \subset X \times Y$ seja o gráfico de uma função $f : X \to Y$, é necessário e suficiente que, para cada $x \in X$, exista um único ponto $(x,\, y) \in \Gamma$. No caso em que $X, Y \subset \mathbb{R}$ e podemos representar a função num plano munido de um sistema de coordenadas OXY, isso significa que toda paralela ao eixo OY traçada por um ponto de X deve cortar Γ num único ponto (veja Figura 1.18).

Dada uma curva definida por alguma propriedade geométrica, pode ocorrer, como no caso da Figura 1.18b, que esta curva não represente o gráfico de uma função, pelo menos não de uma coordenada em relação à outra, digamos y em relação a x. O exemplo mais simples e comum é o da circunferência de raio r centrada na origem do sistema de coordenadas (Figura 1.19).

Sua equação (dada pela equação (1.5.3) quando $Q = (u,\, v) = (0,\, 0) = O$),

$$x^2 + y^2 = r^2,$$



(a) Gráfico de uma função.

(b) Não é o gráfico de uma funcão.

Figura 1.18: À esquerda, exemplo do gráfico de uma função $f : X \subset \mathbb{R} \to \mathbb{R}$. À direita, uma curva que não pode ser o gráfico de uma função, pois uma paralela ao eixo OY a intersecta mais de uma vez.

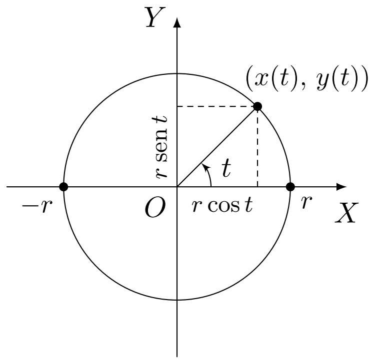


Figura 1.19: Circunferência de centro na origem e raio r.

após uma pequena manipulação nos fornece y:

$$y = \sqrt{r^2 - x^2} \quad \text{e} \quad y = -\sqrt{r^2 - x^2}.$$

Se considerarmos $y_1 : [-r, r] \to \mathbb{R}$, $y_1(x) = \sqrt{r^2 - x^2}$, e $y_2 : [-r, r] \to \mathbb{R}$, $y_2(x) = -\sqrt{r^2 - x^2}$, a circunferência toda será descrita analiticamente por duas funções. Portanto, essa curva não nos fornece uma função de forma única.

Existem, no entanto, outros métodos analíticos de representação para lidar com esse tipo de problema. O mais geral e útil deles (sobretudo no caso de curvas fechadas) é o da *representação paramétrica*. Nessa representação, ao invés de considerar uma coordenada como função da outra, o que se faz é pensar em ambas x e y como função de uma terceira variável independente t chamada *parâmetro*. Isso motiva a próxima definição.

Definição 1.52. *Seja* Γ *uma curva plana. Dizemos que uma aplicação* $\gamma : T \subset \mathbb{R} \to \mathbb{R}^2$, $\gamma(t) = (x(t),\, y(t))$ *é uma* parametrização *de* Γ *se a sua imagem* $Im(\gamma) = \gamma(T)$ *coincide com* Γ, *i.e., se*

$$\Gamma = Im(\gamma) = \{(x(t),\, y(t)) \in \mathbb{R}^2;\, t \in T\}.$$

A imagem $Im(\gamma) = \gamma(T)$ chama-se *traço* de γ. O domínio $T \subset \mathbb{R}$ de γ geralmente é um intervalo ou uma reunião finita de intervalos.

Em muitos casos, o parâmetro t recebe uma interpretação física imediata: a de tempo. Com isso, as coordenadas x e y se tornam funções do tempo e, portanto, o movimento de qualquer ponto no plano pode ser expresso matematicamente: $\gamma(t) = (x(t),\, y(t))$ é a posição do ponto no instante t e $\gamma(T)$ é o seu *caminho* ou *trajetória*.

Exemplo 1.53 (Parametrização da circunferência de centro na origem). Voltanto à circunferência da Figura 1.19, como ficaria sua parametrização? Como sugerido na figura, se t é a medida, em radianos, do ângulo entre o eixo OX e o segmento OP, em que $P = (x(t),\, y(t))$, tomada no sentido anti-horário, então:

$$x(t) = r \cos t \quad \text{e} \quad y(t) = r \operatorname{sen} t,$$

sendo $t \in T = [0,\, 2\pi[$, é uma parametrização da circunferência. Na nossa interpretação física: o ponto de coordenadas x, y descreve completamente a circunferência conforme o tempo t varia de 0 a 2π (em qualquer unidade de tempo); note que não haveria problema algum em t ser maior do que 2π (a rigor, a função seria outra, pois o domínio seria diferente, mas ainda assim a curva seria a mesma e poderíamos interpretar que o ponto P continua sua trajetória circular). Assim, a parametrização nos permitiu descrever, de uma única vez (i.e., com uma única função $\gamma : T \to \mathbb{R}^2$), a curva (nesse caso particular, a circunferência) que antes era descrita por duas funções quando pensávamos numa coordenada como função da outra. ◇

Exemplo 1.54 (Parametrização da elipse de centro na origem). Outro caso particularmente bem conhecido de parametrização é o da elipse centrada na origem do sistema de coordenadas. Da Geometria Analítica, sabemos que sua equação é $x^2/a^2 + y^2/b^2 = 1$; uma possível representação paramétrica é da forma:

$$x(t) = a \cos t \quad \text{e} \quad y(t) = b \operatorname{sen} t,$$

$t \in [0,\, 2\pi[$. (Por quê?) Geometricamente, o parâmetro t é interpretado como a medida (tomada no sentido anti-horário) do ângulo formado pelo eixo OX e o segmento OP', em que P' é ponto logo acima ou abaixo de P na circunferência circunscrita à elipse; nesse caso, t recebe o nome de ângulo excêntrico (Figura 1.20). ◇

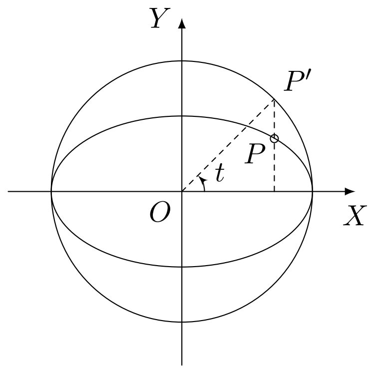


Figura 1.20: Interpretação geométrica da parametrização da elipse.

Uma das grandes vantagens da parametrização é permitir representarmos graficamente e estudarmos por meio da linguagem das funções curvas mais sofisticadas (e razoavelmente complicadas), como, por exemplo, a dada pelas seguintes equações:

$$x(t) = t + 2\,\mathrm{sen}\,2t, \quad y(t) = t + 2\cos 5t,$$

cujo traçado tem o aspecto ilustrado na Figura 1.21. Note que cada uma das coordenadas, $x = x(t)$ e $y = y(t)$, é uma função do parâmetro t, mas a curva $\gamma(t) = (x(t),\, y(t))$ não é o gráfico de uma função.

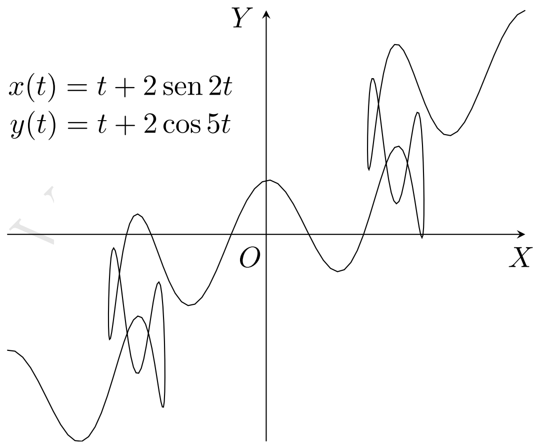


Figura 1.21: Parametrização de uma curva um pouco mais sofisticada.

Para o que segue, considere T, X, $Y \subset \mathbb{R}$.

Em geral, para representar uma curva parametricamente, procura-se por funções do parâmetro t, $\phi : T \to X$ e $\psi : T \to Y$, uma para cada coordenada[27]:

$$x = \phi(t) = x(t), \quad y = \psi(t) = y(t).$$

[27]As letras gregas ϕ e ψ leem-se “fi” e “psi”, respectivamente.

Ambas devem ser determinadas de tal modo que a totalidade de pares $(x(t), y(t))$ obtidos quando t varia em T esteja na curva e nenhum deles esteja fora.

Se uma curva é dada por uma função $f : X \to Y$, $y = f(x)$, podemos obter a sua representação paramétrica escolhendo qualquer função estritamente monótona[28] e sobrejetiva[28] $\phi : T \to X$, $x = \phi(t)$; dessa forma, $\psi : T \to Y$, $y = \psi(t) = f(\phi(t))$, i.e., ψ é a composta[29] de f com ϕ. Note, portanto, que há bastante liberdade na representação paramétrica, devido à liberdade de escolha da função ϕ. Podemos escolher, por exemplo, $\phi = id_X$, a função identidade: basta fazer $x = t$; assim podemos pensar na representação inicial $y = f(x)$ como sendo uma representação paramétrica com parâmetro $t = x$. Uma vantagem dessa liberdade de escolha de ϕ é seu uso para propósitos de simplificação, como ilustra o proximo exemplo.

Exemplo 1.55. Seja $f : \mathbb{R} \to \mathbb{R}$, $f(x) = \sqrt[3]{x^2}$. Podemos escolher $\phi : \mathbb{R} \to \mathbb{R}$, $\phi(t) = t^3$, assim $\psi : \mathbb{R} \to \mathbb{R}$, $\psi(t) = f(\phi(t)) = t^2$ e, portanto, $x = t^3$, $y = t^2$ descrevem toda a curva (parábola de Neile[30] – Figura 1.22) conforme t varia de $-\infty$ a ∞. ◇

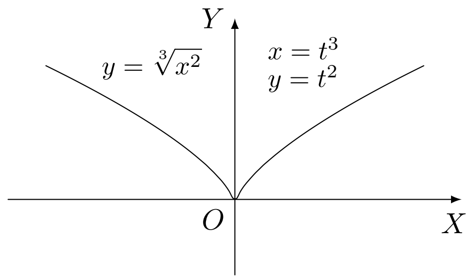


Figura 1.22: Parábola de Neile (parábola semicúbica).

Reciprocamente, dada uma curva representada parametricamente por $\phi : T \to X$, $x = \phi(t)$, e $\psi : T \to Y$, $y = \psi(t)$, se desejamos obter sua representação na forma não paramétrica, por meio da função $f : X \to Y$, $y = f(x)$, basta eliminarmos t das equações $x = \phi(t)$, $y = \psi(t)$. Nos casos particulares da circunferência e da elipse (Exemplos 1.53 e 1.54), isso pode ser feito de uma só vez elevando ao quadrado e utilizando a relação $\operatorname{sen}^2\xi + \cos^2\xi = 1$. Mas, em geral, devemos encontrar uma expressão para t a partir da equaçao $x = \phi(t)$ por meio da função inversa[1.7] $t = \phi^{-1}(x)$ e substituir em $y = \psi(t)$ a fim de obtermos a representação $y = \psi(\phi^{-1}(x)) = f(x)$. Neste processo, devemos nos restringir à porção da curva que é intersectada apenas uma vez pela paralela ao

[28] Veja conceito na Seção 1.14.
[29] Veja conceito na Seção 1.7.
[30] Em honra a William Neile (1637–1670), cujo sobrenome aparece por vezes como Neil, matemático inglês, um dos doze cofundadores da *Royal Society*, que aos dezenove anos obteve a retificação (i.e., calculou o comprimento de um arco qualquer) dessa curva, também chamada parábola semicúbica.

eixo OY, pois, com essa restrição, ϕ é sobrejetiva e injetiva, logo é bijetiva e, portanto, invertível (veja Teorema 1.85).

1.5.2.3 Coordenadas polares

Naturalmente, o uso de um par de eixos ortogonais não é a única maneira de se estabelecer uma correspondência entre pontos do plano e pares ordenados de números reais. Poderíamos utilizar qualquer par de eixos concorrentes (como o da Figura 1.23) e um processo similar ao descrito no início da Seção 1.5.2 para obter uma correspondência $f : \Pi \to \mathbb{R}^2$ (veja no Apêndice D como, curiosamente, um mesmo ponto pode ter duas representações num sistema de eixos não ortogonais dependendo da escolha de f).

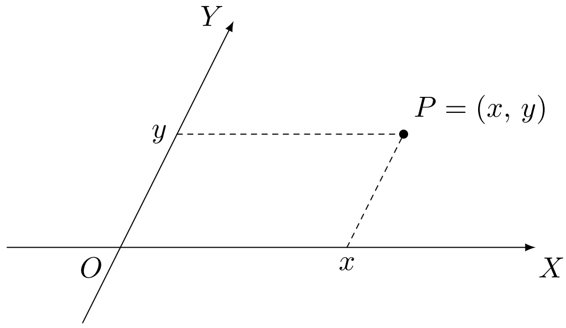


Figura 1.23: Sistema de eixos não ortogonais.

Outra alternativa seria dispensar um dos eixos. No sistema de *coordenadas polares*, utiliza-se apenas OX. A correspondência $f : \mathbb{R}^2 \to \Pi$ se faz da seguinte maneira: dado um par ordenado de números reais (ρ, θ)[31], obtém-se o ponto correspondente P do plano a partir da interseção da circunferência de centro O e raio $|\rho|$ com a semirreta de origem O que faz, com o semieixo positivo (resp. negativo) OX, se $\rho \geqslant 0$ (resp. $\rho < 0$), um ângulo de θ radianos medido a partir do semieixo positivo (resp. negativo) no sentido anti-horário, se $\theta \geqslant 0$; no sentido horário, se $\theta < 0$. Se $\rho = 0$, então $P = O$, qualquer que seja $\theta \in \mathbb{R}$. O ponto O chama-se *polo* ou *origem*.

As *coordenadas polares* ρ, θ estão conectadas com as coordenadas cartesianas x, y de um ponto P por meio das equações:

$$x = \rho\cos\theta, \quad y = \rho\,\mathrm{sen}\,\theta, \quad \rho^2 = x^2 + y^2, \quad \mathrm{tg}\,\theta = y/x, \tag{1.5.4}$$

cuja interpretação geométrica fica clara na Figura 1.24.

A transformação do sistema de coordenadas polares para o sistema de coordenadas cartesianas é um exemplo de uma função $f : \mathbb{R}^2 \to \mathbb{R}^2$, $f(\rho, \theta) = (x, y)$, cuja lei

[31] Essas letras gregas leem-se "rô" e "teta", respectivamente.

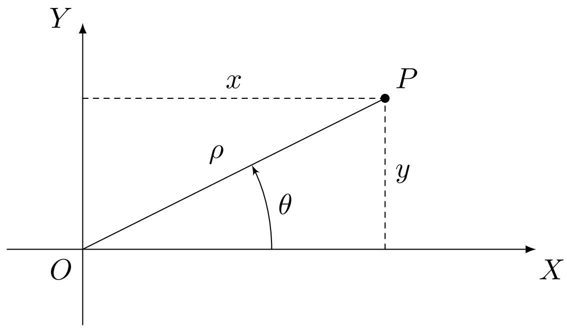


Figura 1.24: Coordenadas polares.

de correspondência é dada pelas duas primeiras das equações (1.5.4). Cada uma das coordenadas x, y, por sua vez, é uma função de $\mathbb{R}^2$ em $\mathbb{R}$:

$$x = x(\rho,\, \theta) = \rho \cos\theta, \quad y = y(\rho,\, \theta) = \rho \operatorname{sen}\theta.$$

A vantagem de se utilizar o sistema de coordenadas polares é obter uma descrição mais simples de certas curvas. Considere, por exemplo, a *lemniscata*[32]. Geometricamente, essa curva é o lugar geométrico dos pontos P para os quais o produto das distância r_1 e r_2 em relação aos pontos $F_1 = (a,\, 0)$ e $F_2 = (-a,\, 0)$, respectivamente, é igual à constante a^2 (Figura 1.25).

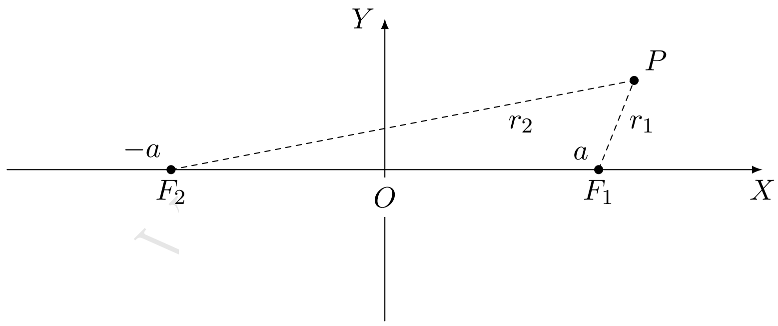


Figura 1.25: Lemniscata.

Como, pela equação (1.5.1),

$$r_1^2 = (x-a)^2 + y^2, \quad r_2^2 = (x+a)^2 + y^2,$$

um simples cálculo nos fornece a equação da lemniscata em coordenadas cartesianas:

$$(x^2+y^2)^2 - 2a^2(x^2-y^2) = 0.$$

[32]Em 1694, Jacob (Jacques) Bernoulli (1654–1705) publicou um artigo em *Acta Eruditorum* (uma revista científica mensal alemã publicada entre 1682 e 1782) no qual descrevera uma curva *com a forma de um "8", ou um nó, ou o arco de uma fita*, que ele chamou pela palavra latina *lemniscus* ("fita pingente").

Introduzindo coordenadas polares (duas primeiras das equações (1.5.4)), obtemos:

$$\rho^4 - 2a^2\rho^2(\cos^2\theta - \operatorname{sen}^2\theta) = 0;$$

agora, dividindo por ρ^2 e utilizando a fórmula trigonométrica $\cos^2\theta - \operatorname{sen}^2\theta = \cos 2\theta$, a equação acima fica:

$$\rho^2 = 2a^2\cos 2\theta,$$

que é mais simples do que em coordenadas cartesianas. Note que são funções $\rho_+,\ \rho_- : [-\pi/4, \pi/4] \to \mathbb{R}$:

$$\rho_+(\theta) = a\sqrt{2\cos 2\theta}, \quad \rho_-(\theta) = -a\sqrt{2\cos 2\theta},$$

que correspondem respectivamente ao lado direito e ao lado esquerdo da lemniscata.

Uma desvantagem do uso desse sistema reside no fato de que a correspondência entre as coordenadas polares e os pontos do plano não é biunívoca (veja esse conceito na Seção 1.6.3), já que todos os pares da forma $(\rho,\ \theta + 2k\pi)$, $k \in \mathbb{Z}$, são associados ao mesmo ponto do plano; em particular, todos os pares $(0,\ \theta)$, $\theta \in \mathbb{R}$, estão associados ao polo.

É importante o estudante ter em mente que os sistemas de coordenadas acima não esgotam todas as possíbilidades. Eles são apenas os mais comuns. Outros podem ser construídos e utilizados conforme a conveniência. Basta encontrar funções $f : X \subset \mathbb{R}^2 \to \Pi$ sobrejetivas (para esse conceito, veja a Seção 1.6.2), ou seja, basta econtrar uma maneira de mapear pares ordenados de números reais em pontos do plano.

1.5.2.4 Considerações gerais sobre o espaço numérico $\mathbb{R}^3$

Embora o foco deste texto esteja sobre as funções (elementares) reais de uma variável real e sua representação no plano numérico $\mathbb{R}^2$, uma palavra sobre funções reais de várias variáveis reais, em particular, de duas variáveis reais e sua representação no espaço numérico $\mathbb{R}^3$, vale a pena visando estudos um pouco mais avançados futuramente.

Em quase todas as relações que aparecem na natureza, de fato, as funções em questão não dependem de uma única variável independente; ao contrário, a variável dependente geralmente é determinada por duas, três, ou mais variáveis independentes. Por exemplo, a lei da gravitação universal, de Newton[33], afirma que dois corpos, um de massa m e o outro de massa m', situados a uma distância d um do outro, se atraem segundo uma força cuja intensidade F é proporcional a essas massas e inversamente proporcional ao quadrado da distância entre eles, i.e.:

$$F = F(m,\ m',\ d) = c\frac{mm'}{d^2},$$

[33] Isaac Newton (1643–1727) foi o maior matemático inglês de sua geração. Ele lançou as bases para o cálculo diferencial e integral. Seu trabalho em óptica e gravitação fazem dele um dos maiores cientistas que o mundo já conheceu.

em que c é uma constante que depende do sistema de unidades utilizado. Ou seja, F é uma função de três variáveis independentes: $F : \mathbb{R}^3_+ \to \mathbb{R}$.

Tanto do ponto de vista da matemática pura quanto aplicada, o estudo das funções de várias variáveis independentes é sem dúvida indispensável. Uma boa notícia é que podemos tirar vantagem do que já aprendemos e vamos aprender sobre funções de apenas uma variável independente, de modo que em muitos casos basta uma simples extensão dos argumentos. Ademais, geralmente basta considerar apenas duas variáveis independentes x e y, uma vez que sua extensão para três ou mais variáveis essencialmente não requer novidades. Assim, podemos nos deter a apenas duas variávies independentes a fim de simplificar a notação e as afirmações.

Considere equações da forma

$$z = x + y, \quad z = xy, \quad z = \log(1 - x^2 - y^2), \quad \text{com}\, x,\, y,\, z \in \mathbb{R}.$$

As duas primeiras associam um único valor $z \in \mathbb{R}$ a pares $(x, y) \in \mathbb{R}^2$; a última associa um único $z \in \mathbb{R}$ a pares $(x, y) \in \mathbb{R}^2$ tais que $x^2 + y^2 < 1$. São, portanto, exemplos de leis de formação de *funções* $f : \mathcal{R} \to \mathbb{R}$, $\mathcal{R} \subset \mathbb{R}^2$, $f(x, y) = z$.

Nesses casos, dizemos que x e y são as *variáveis independentes* e que z é a *variável dependente*. Representamos o par de valores (x, y) por um ponto no plano numérico $\mathbb{R}^2$ e chamamos esse par de *argumento* da função. O conjunto $\mathcal{R}$ é o *domínio* ou *campo de definição* da função.

Como já discutido anteriormente, a relação funcional entre as variáveis pode ser estabelecida por meio de uma equação como as acima, uma descrição verbal, como "z é a área de um retângulo cujos lados medem x e y", um algoritmo ou por observações físicas, como no caso da declinação magnética, que depende da latitude e da longitude (fixado o tempo). O essencial é que exista uma correspondência, como a da Definição 1.1 ou, mais precisamente, da Definição 1.46, em que x nessas definições pode ser interpretado como o ponto (x, y) no contexto desta seção. A extensão para mais variáveis é natural: $f : \mathcal{R} \to Z$ é uma função de três variáveis independentes se para todo terno $(x, y, z) \in \mathcal{R} \subset \mathbb{R}^3$ existe apenas um $u \in Z$ tal que $f(x, y, z) = u$; analogamente para n variáveis independentes $x_1, x_2, \ldots, x_n$.

Além do próprio $\mathbb{R}^2$, alguns dos exemplos mais simples e comuns de domínios de funções de duas variáveis reais são:

1. a *região retangular* (Figura 1.26a), definida por desigualdades da forma:

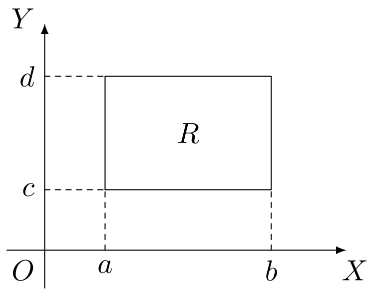
   $$a \leqslant x \leqslant b, \quad c \leqslant y \leqslant d,$$

   na qual cada variável independente está restrita a um intervalo e o argumento (x, y) varia num retângulo;

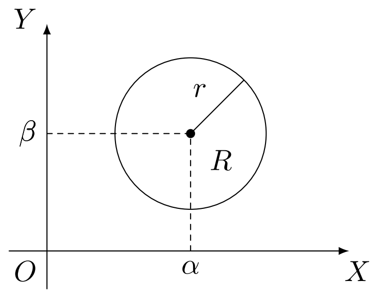
2. e a *região circular* (Figura 1.26b), definida por uma desigualdade do tipo:

$$(x-\alpha)^2 + (y-\beta)^2 \leqslant r^2,$$

na qual o argumento $(x,\, y)$ varia no disco de centro $(\alpha,\, \beta)$ e raio r.



(a) Uma região retangular.



(b) Uma região circular.

Figura 1.26

Analogamente, no caso de três variáveis independentes x, y e z, teríamos regiões retangulares

$$a \leqslant x \leqslant b, \quad c \leqslant y \leqslant d, \quad e \leqslant z \leqslant f$$

e regiões esféricas

$$(x-\alpha)^2 + (y-\beta)^2 + (z-\gamma)^2 \leqslant r^2.$$

Quando lidamos com mais de três variáveis independentes, a intuição geométrica falha, mas é conveniente utilizar a mesma terminologia. Assim, para funções de n variáveis independentes $x_1,\ x_2,\ \ldots,\ x_n$, chamamos de regiões retangular e esférica respectivamente as regiões

$$a_1 \leqslant x_1 \leqslant b_1, \quad a_2 \leqslant x_2 \leqslant b_2, \quad \ldots, \quad a_n \leqslant x_n \leqslant b_n$$

e

$$(x_1-\alpha_1)^2 + (x_2-\alpha_2)^2 + \cdots + (x_n-\alpha_n)^2 \leqslant r^2.$$

Alguns exemplos simples de funções reais de duas variáveis reais $\phi : \mathbb{R}^2 \to \mathbb{R}$ são os *polinômios*. O polinômio geral do primeiro grau tem a forma

$$\phi(x,\, y) = ax + by + c,$$

em que a, b e c são constantes. O polinômio geral do segundo grau tem o aspecto

$$\phi(x,\, y) = ax^2 + bxy + cy^2 + dx + ey + f,$$

em que a, b, c, d, e e f são constantes. Em Geometria Analítica, você verá que essas são, respectivamente, a equação geral de um plano no espaço e a equação de uma classe de superfícies que recebem a denominação de quádricas. Em geral, um polinômio de grau qualquer consiste na soma de termos da forma $a_{mn}x^m y^n$, em que os a_{mn} são constantes.

Outros exemplos são as *funções racionais*, que consistem no quociente de polinômios; a esta classe pertencem as *funções racionais lineares*, $\phi : \mathcal{R} \to \mathbb{R}$, que têm a forma:

$$\phi(x,\, y) = \frac{ax + by + c}{dx + ex + f},$$

em que $a,\, b,\, c,\, d,\, e,\, f \in \mathbb{R}$ e cujo domínio $\mathcal{R}$ consiste no conjunto de pontos de $\mathbb{R}^2$ que não anulam o denominador dessa expressão.

Por meio da extração de raízes podemos passar das funções racionais para as *funções algébricas*[34], por exemplo,

$$\phi(x,\, y) = \sqrt[3]{\frac{ax + by + c}{dx + ex + f}}.$$

E recorrendo às bem conhecidas funções de uma variável, podemos construir funções mais complicadas de várias variáveis, por exemplo:

$$\phi(x,\, y) = \log(\,\mathrm{sen}\,(xy) + 1).$$

Assim como podemos representar as funções reais de uma variável real por meio de pontos no plano, que podem ou não formar curvas, também podemos representar as funções reais de duas variáveis reais por pontos no espaço, que por sua vez podem ou não formar *superfícies* ou curvas. Para a construção tridimensional, consideramos um sistema de coordenadas retangular no espaço, com coordenadas $x,\, y,\, z$ (chamadas, respectivamente, de *abscissa*, *ordenada* e *cota*) e, para cada ponto $(x,\, y)$ do domínio $\mathcal{R}$ da função marcamos o ponto $P = (x,\, y,\, z)$ cuja terceira coordenada é $z = f(x,\, y)$. Para detalhes sobre a construção da correspondência (biunívoca) entre pontos do espaço e ternos ordenados de números reais, veja o Apêndice B.

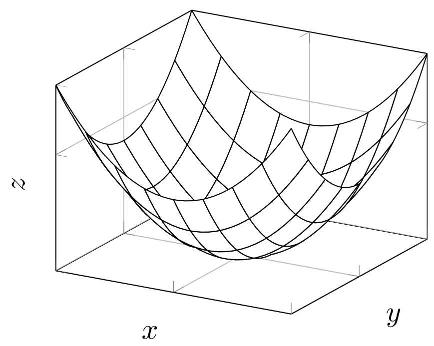
Para certas funções reais de duas variáveis reais, quando $(x,\, y)$ "se move" em $\mathcal{R}$, o ponto P descreve uma superfície no espaço; reciprocamente, na Geometria Analítica, as superfícies são descritas por funções desse tipo. Há assim uma relação de reciprocidade entre certas superfícies no espaço e essas funções.

[34] Apenas por curiosidade, uma função $f : \mathbb{R}^n \to \mathbb{R}$, $y = f(x_1,\, x_2,\, \ldots,\, x_n)$ é dita uma *função algébrica* das variáveis independentes $x_1,\, x_2,\, \ldots,\, x_n$ se y pode ser definida implicitamente por uma equação $F(x_1,\, x_2,\, \ldots,\, x_n,\, y) = 0$, em que F é um polinômio em $x_1,\, x_2,\, \ldots,\, x_n,\, y$, i.e., se y satisfaz uma "equação algébrica". Funções que não safisfazem uma equação algébrica são chamadas *transcendentes.* A palavra "transcendente" não significa nada particularmente profundo ou misterioso; significa apenas que um objeto matemático (um número, uma função etc.) não pode ser obtido por meio de operações algébricas elementares, "*quod algebrae vires transcendit*" ("o que transcende o poder da álgebra"), como colocado por Leibniz.

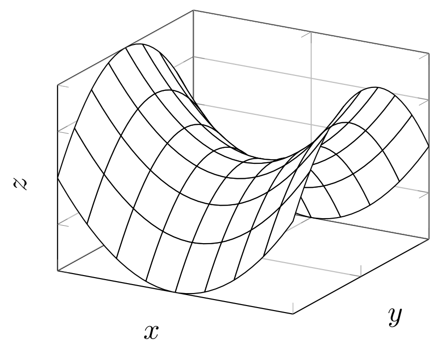
Por exemplo, a função $\phi : \{(x,\, y) \in \mathbb{R}^2;\, x^2 + y^2 \leqslant 1\} \to \mathbb{R}$,

$$\phi(x,\, y) = \sqrt{1 - x^2 - y^2},$$

corresponde a um *hemisfério* sobre o plano OXY de raio unitário e centro na origem; a função linear $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x,\, y) = ax{+}by{+}c$, tem por gráfico um *plano* que intersecta o eixo OZ no ponto de coordenada c; a função $g : \mathbb{R}^2 \to \mathbb{R}$, $g(x,\, y) = x^2 + y^2$, corresponde ao chamado *parabolóide de revolução*, que pode ser obtido por meio da rotação da parábola $z = x^2$ em torno do eixo z (Figura 1.27a); a função $h : \mathbb{R}^2 \to \mathbb{R}$, $h(x,\, y) = x^2 - y^2$, corresponde ao chamado *parabolóide hiperbólico*, também conhecido como *sela de cavalo* devido ao seu aspecto (Figura 1.27b).



(a) $z = g(x,\, y) = x^2 + y^2$.



(b) $z = h(x,\, y) = x^2 - y^2$.

Figura 1.27: Parabolóide de revolução (esquerda) e hiperbólico (direita).

No entanto, esse tipo de representação aplica-se a funções de no máximo duas variáveis, pois mais variáveis implicaria representar objetos em quatro ou mais dimensões e, por razões evolutivas, o cérebro humano não consegue fazê-lo. Além disso, fazer esboços e construções geométricas no plano é mais simples do que no espaço. Desse ponto de vista, é preferível adotar uma representação geométrica alternativa das funções de duas variáveis, que consiste nas *linhas de nível*[35].

Uma linha de nível c de uma função f consiste do conjunto de todos os pontos do plano numérico $\mathbb{R}^2$ tais que $f(x,\, y) = c$, em que c é uma constante. Podemos dizer isso um pouco mais formalmente como na seguinte definição.

Definição 1.56. *Seja uma função $f : \mathcal{D} \subset \mathbb{R}^2 \to \mathbb{R}$. Dizemos que o conjunto*

$$f^{-1}(c) = \{(x,\, y) \in \mathcal{D};\, f(x,\, y) = c\}$$

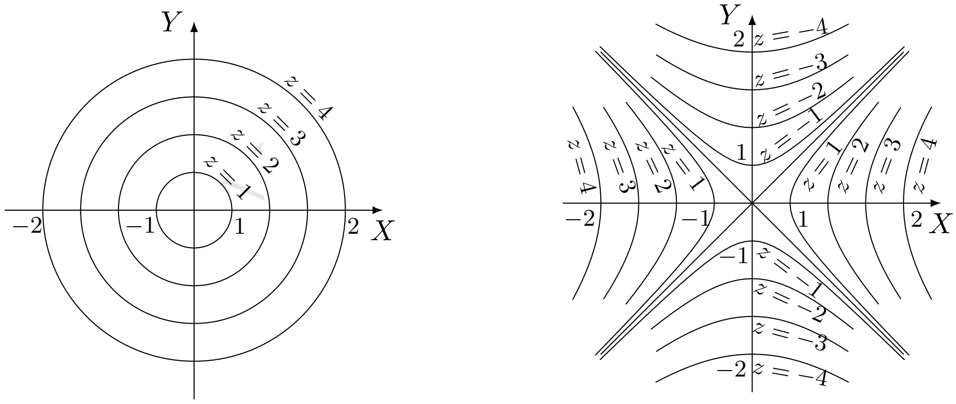
[35]A palavra "nível" provém dos mapas de relevo, nos quais a função em estudo mede a altitude em relação ao nível do mar, que é considerado o nível zero.

é a linha de nível *c da função f, em que* $c \in \mathbb{R}$.

Do ponto de vista geométrico, a linha de nível c de uma função real de duas variáveis reais também pode ser obtida a partir da intersecção do gráfico dessa função com o plano $z = c$, paralelo ao plano-XY, seguida da projeção ortogonal dessa curva no plano-XY.

A representação gráfica das funções reais de duas variáveis reais por esse método se obtém traçando sucessivas linhas de nível no plano numérico $\mathbb{R}^2$ para diferentes constantes c_1, c_2, ..., que representam diferentes alturas do gráfico da função. Geralmente, adota-se uma progressão aritmética para esses valores: $c_i = ik$, onde $i = 1, 2, \ldots$ e k é uma constante. Cada linha de nível representa os pontos do gráfico da função que estão à mesma "altura"; por esse motivo, o espaçamento entre as linhas sucessivas nos conta algo sobre a declividade da superfície: linhas próximas indicam uma superfície íngreme; linhas mais afastadas sugerem maior suavidade.

Alguns exemplos devem tornar as coisas mais claras. A função linear $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = ax + by + c$, é representada no plano numérico $\mathbb{R}^2$ por uma família de retas $ax+by+c = k$, $k \in \mathbb{R}$; a função $g : \mathbb{R}^2 \to \mathbb{R}$, $g(x, y) = x^2+y^2$ (parabolóide de revolução – Figura 1.27a), por uma família de círculos concêntricos, de centro na origem e raio r: $x^2+y^2 = r^2$, $r \in \mathbb{R}$ (Figura 1.28a); a função $h : \mathbb{R}^2 \to \mathbb{R}$, $h(x, y) = x^2-y^2$ (parabolóide de hiperbólico – Figura 1.27b), que possui um ponto de sela na origem, por uma família de hipérboles $x^2 - y^2 = k^2$, $k \in \mathbb{R}$ (Figura 1.28b).



(a) Linhas de nível de $z = x^2 + y^2$.

(b) Linhas de nível de $z = x^2 - y^2$.

Figura 1.28: Exemplos de linhas de nível.

Esse tipo de representação tem ainda a vantagem de poder ser extendido para funções reais de três variáveis reais $f : \mathbb{R}^3 \to \mathbb{R}$, $f(x, y, z) = u$; seguindo a mesma terminologia, ao invés de linhas, passamos a ter *superfícies de nível* c: $f(x, y, z) = c$, em que c é uma constante. Por exemplo, as superfícies de nível da função $f : \mathbb{R}^3 \to \mathbb{R}$, $f(x, y, z) = x^2 + y^2 + z^2$, são superfícies esféricas concêntricas, de centro na origem e raio r: $x^2 + y^2 + z^2 = r^2$.

1.6 Funções injetivas, sobrejetivas e bijetivas

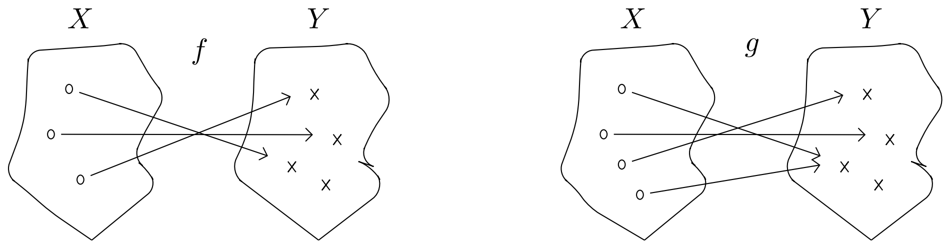
Há algumas palavras ligadas ao conceito de função cuja familiaridade é indispensável para o prosseguimento dos nossos estudos. Vamos apresentá-las nas três próximas definições.

1.6.1 Funções injetivas

Definição 1.57. *Dizemos que uma função $f : X \to Y$ é* injetiva *(ou* biunívoca*, ou uma* correspondência um-a-um*, ou simplesmente* um-a-um*) quando elementos diferentes em X são transformados em elementos diferentes em Y, ou seja,*

$$x_1,\, x_2 \in X,\, x_1 \neq x_2 \Rightarrow f(x_1) \neq f(x_2)$$

Para ilustrar o conceito dessa definição, considere o modelo do diagrama de setas descrito na Seção 1.5.1. Por exemplo, as funções esquematizadas nas Figuras 1.13 e 1.29b não são injetivas, pois em ambos os casos há elementos no contradomínio Y para os quais mais de uma flecha aponta, indicando que há, em cada um dos dois casos, mais de um elemento no domínio X cuja imagem é o mesmo elemento de Y. Já a Figura 1.29a ilustra uma função injetiva, pois para cada elemento do contradomínio está apontada no máximo uma flecha, indicando que elementos diferentes do domínio possuem diferentes imagens. Nessa figura, fica claro o porquê da nomenclatura *um-a-um.*



(a) Diagrama de uma função injetiva.

(b) Diagrama de uma função não injetiva.

Figura 1.29: Funções injetiva e não injetiva num diagrama de setas.

Exemplo 1.58. Exemplos simples de funções injetivas são as funções inclusão e identidade: respectivamente itens (a) e (b) do Exemplo 1.3. ◇

No caso de funções reais de uma variável real, se você dispusesse apenas do gráfico delas no plano numérico, como você poderia dizer se se trata de uma função injetiva ou não? A resposta é simples: a função não será injetiva se existir uma reta paralela ao eixo OX que intersecte o gráfico em mais de um ponto. Como exercício, aponte todas as figuras vistas até aqui que representam os gráficos de funções reais de uma variável real não injetivas[36]. Pense sobre o análogo para funções reais de duas variáveis reais e note que os gráficos da Figura 1.27 representam funções não injetivas. Apenas tome cuidado para não confundir essa maneira intuitiva de visualizar funções injetivas com uma demonstração! Para demonstrar se uma fução é injetiva ou mostrar que ela não é, é preciso raciocinar logicamente. Veja adiante.

Há outras formas equivalentes de se enunciar a Definição 1.57; por exemplo, dizemos que uma função $f : X \to Y$ é *injetiva* quando,

1. para todo $y \in f(X)$, existe um único $x \in X$ tal que $f(x) = y$; ou

2. para dois elementos quaisquer x_1 e x_2 do domínio de f, $f(x_1) = f(x_2)$ implica $x_1 = x_2$.

Este último enunciado é a forma *contrapositiva*[37] do enunciado da Definição 1.57. Dependendo da situação, um desses enunciados pode ser mais conveniente do que o outro.

Quando se deseja mostrar que uma função não é injetiva, e para tanto basta exibir um exemplo, é preferível utilizar a forma do enunciado da Definição 1.57. Para ilustrar o que acaba de ser dito, considere a função do Exemplo 1.16. A fim de mostrar que ela não é injetiva, basta exibirmos duas figuras planas distintas φ_1 e φ_2 tais que $A(\varphi_1) = A(\varphi_2)$; por exemplo, considere φ_1 um quadrado de lado unitário e φ_2 um retângulo de lados meio e dois; dessa forma $A(\varphi_1) = A(\varphi_2) = 1$ e portanto A não é injetiva.

Por outro lado, quando desejamos demonstrar que certa função é injetiva, convém utilizar a forma contrapositiva, pois, do ponto de vista operacional, ela nos permite manipular uma equação, enquanto o enunciado original (da Definição 1.57) não. Para exemplificar isso, considere agora a função do Exemplo 1.17. Se $\mathcal{M}_1$ e $\mathcal{M}_2$ são duas matrizes arbitrárias de $M_2(\mathbb{R})$, precisamos apenas mostrar que $T(\mathcal{M}_1) = T(\mathcal{M}_2)$ implica $\mathcal{M}_1 = \mathcal{M}_2$, o que é muito simples, pois basta recordar a definição de igualdade de matrizes: duas matrizes são iguais quando seus elementos correspondentes são iguais;

[36] Resposta: Figuras 1.2, 1.6, 1.8, 1.9, 1.15, 1.18a, 1.22.

[37] Dada uma implicação $p \Rightarrow q$, chama-se sua *contrapositiva* a implicação $q' \Rightarrow p'$, em que p' e q' são as negações de p e q respectivamente. Ambas, $p \Rightarrow q$ e $q' \Rightarrow p'$, são logicamente equivalentes, ou seja, são duas maneiras diferentes de se dizer a mesma coisa.

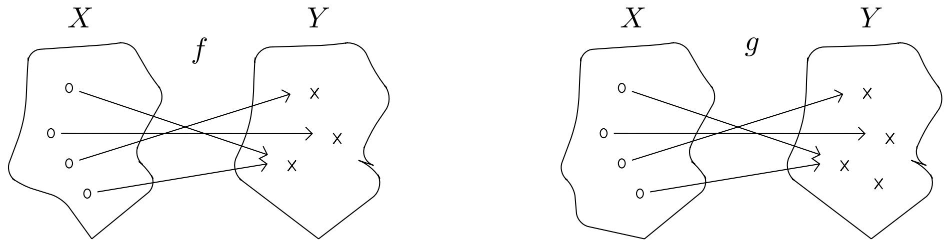
supor que as transpostas são iguais ($T(\mathcal{M}_1) = T(\mathcal{M}_2)$) implica justamente na igualdade dos elementos correspondentes das duas matrizes e, portanto, na iguadade delas.

1.6.2 Funções sobrejetivas

Definição 1.59. *Dizemos que uma função $f : X \to Y$ é* sobrejetiva *(ou* sobre *Y) quando $f(X) = Y$, ou seja, quando, para todo $y \in Y$, existe (pelo menos) um $x \in X$ tal que $y = f(x)$.*

Em outras palavras, f é sobrejetiva quando todo elemento do contradomínio é imagem de (pelo menos) um elemento do domínio.

Vamos utilizar novamente um diagrama de setas para ilustrar esse conceito e comparar uma função sobrejetiva com uma função não sobrejetiva (Figura 1.30).



(a) Diagrama de uma função sobrejetiva. (b) Diagrama de uma função não sobrejetiva.

Figura 1.30: Funções sobrejetiva e não sobrejetiva num diagrama de setas.

Dada uma função $f : X \to Y$, para saber se um certo elemento $b \in Y$ pertence ou não à imagem $f(X)$, escrevemos a "equação" $f(x) = b$ e procuramos encontrar algum $x \in X$ que a satisfaça; desse modo, para mostrar que f é sobrejetiva, deve-se provar que a equação $f(x) = y$ possui solução para todo $y \in Y$.

Exemplo 1.60. Exemplos simples de funções sobrejetivas são as projeções de um produto cartesiano num dos fatores, item (d) do Exemplo 1.3. As funções dos Exemplos 1.17 e 1.19 também são sobrejetivas, enquanto as dos Exemplos 1.16 e 1.18 não. (Verifique!) ◇

No caso particular das funções reais de uma variável real, se dispuséssemos novamente apenas do gráfico delas no plano numérico, como identificar uma função sobrejetiva (ou não sobrejetiva)? Basta que toda reta paralela ao eixo OX por um ponto do contradomínio intersecte (pelo menos) um ponto do gráfico da função. Como exercício, observe todos os gráficos desse tipo de função vistos até aqui e, sempre que possível, diga se se trata de uma função sobrejetiva ou não. Como um desenho é sempre limitado, não é possível saber o que ocorre em pontos do contadomínio que não estão representados na imagem. Essa é apenas uma maneira de o estudante munir-se de alguns exemplos e contraexemplos e alimentar sua intuição (sempre sujeita a falácias) antes de realmente escrever uma demonstração (seu antídoto às falácias). Jamais considere essa ou qualquer outra maneira intuitiva como a palavra final!

1.6.3 Funções bijetivas

Definição 1.61. *Uma função $f : X \to Y$ chama-se* bijetiva *(ou uma* bijeção *ou uma* correspondência biunívoca *entre X e Y) quando é ao mesmo tempo injetiva e sobrejetiva, ou seja, quando, para todo $y \in Y$, existe um único $x \in X$ tal que $f(x) = y$.*

Nenhum dos diagramas das Figuras 1.29 e 1.30 representam funções bijetivas: na Figura 1.29a, a função é injetiva mas não é sobrejetiva; nas Figuras 1.29b e 1.30b, as funções não são injetivas nem sobrejetivas; e na Figura 1.30a a função é sobrejetiva mas não é injetiva. Um exemplo de diagrama de uma função bijetiva seria como o da Figura 1.31.

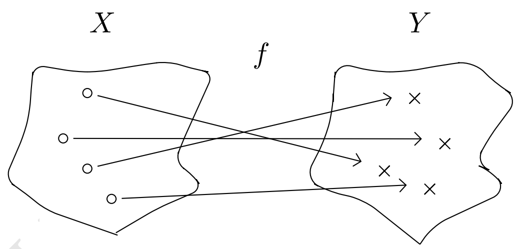


Figura 1.31: Função bijetiva num diagrama de setas.

Exemplo 1.62. Talvez o exemplo mais simples de bijeção seja a função identidade, item (b) do Exemplo 1.3. ◇

Exemplo 1.63. Dados a, b em $\mathbb{R}$, com $a \neq 0$, a função $f : \mathbb{R} \to \mathbb{R}$, definida por $f(x) = ax + b$, é uma bijeção. De fato, se $f(x) = f(y)$, ou seja, $ax + b = ay + b$, somando-se $-b$ a ambos os membros e, em seguida, multiplicando-se ambos por $1/a$, obtemos $x = y$. Assim, f é injetiva. Além disso, dado $y \in \mathbb{R}$ qualquer, o número real $x = (y - b)/a$ é tal que $ax + b = y$, i.e., para todo $y \in \mathbb{R}$, existe $x \in \mathbb{R}$ tal que $f(x) = y$. Logo, f é sobrejetiva. ◇

Naturalmente, sempre é possível se obter uma função bijetiva a partir de uma função injetiva dada. Esse é o conteúdo da próxima proposição, cuja demonstração é deixada como exercício.

Proposição 1.64. *Se $f : X \to Y$ é uma função injetiva, então a função $g : X \to f(X)$, definida por $g(x) = f(x)$ para todo $x \in X$, é bijetiva.*

Observe que f e g são funções diferentes: embora tenham o mesmo domínio e a mesma lei de correspondência, seus contradomínios são diferentes; além disso, g é bijetiva, enquanto f não precisa ser.

Exemplo 1.65. Seja $f : \mathbb{N} \to \mathbb{N}$, $x \mapsto 2x$. Essa função é injetiva mas não é sobrejetiva. (Verifique!) Se chamarmos de P o conjunto $f(\mathbb{N}) = \{$números naturais pares$\}$, então $g : \mathbb{N} \to P$, definida por $x \mapsto 2x$, é bijetiva. Essa correspondência biunívoca de $\mathbb{N}$ em P foi descoberta pelo cientista Galileu Galilei[38]; ela é particularmente curiosa porque P é um subconjunto próprio de $\mathbb{N}$. ◇

Observação 1.66 (Nomenclatura). Muitos autores utilizam as palavras "sobrejetora", "injetora" e "bijetora". Embora sinônimas de "sobrejetiva", "injetiva" e "bijetiva", respectivamente, preferimos estas últimas por associação aos substantivos correspondentes "sobrejetividade", "injetividade" e "bijetividade". Por se tratar meramente de uma questão de gosto, sinta-se à vontade para escolher o seu próprio vocabulário. ○

1.7 Combinando funções: composição e operações

1.7.1 Funções compostas

Suponha que dispomos de duas funções, digamos $f : X \to Y$ e $g : U \to V$, e que desejamos definir uma nova função calculando $g(f(x))$ para todo x que seja relevante, ou seja, aplicando primeiro f e depois g.

Para que isso faça sentido, algumas condições precisam ser satisfeitas. Em primeiro lugar, o valor $f(x)$ precisa estar definido, ou seja,

(a) x deve pertencer ao domínio de f;

em segundo lugar, para podermos calcular $g(f(x))$,

(b) $f(x)$ deve pertencer ao domínio de g.

[38] Galileu Galilei (1564–1642) foi um cientista italiano que formulou a lei básica da queda dos corpos, que ele verificou por medidas cuidadosas. Ele construiu um telescópio com o qual estudou crateras lunares e descobriu quatro luas orbitando Júpiter, abraçado a causa de Copérnico.

Devido a (a) e como o domínio de f é X, o máximo que podemos esperar é que nossa nova função também tenha domínio X. E para que possamos defini-la em todo X, (b) deve valer para todo $x \in X$, ou seja, a imagem de f deve estar contida no domínio de g: $f(X) \subset U$. Se essa condição for satisfeita, nossa função ficará bem definida pela fórmula $g(f(x))$ para todo $x \in X$. Essa heurística motiva a seguinte definição.

Definição 1.67. *Sejam* $f : X \to Y$ *e* $g : U \to V$ *funções, com* $f(X) \subset U$*. A* função composta de g com f *é a função denotada por* $g \circ f : X \to V$*, que a cada elemento* $x \in X$ *associa o elemento* $v = (g \circ f)(x) = g(f(x)) \in V$*, ou seja,*

$$\begin{array}{cccccc} g \circ f: & X & \to & f(X) \subset U & \to & V \\ & x & \mapsto & f(x) & \mapsto & g(f(x)) \end{array}.$$

Note que foi definida uma nova função, $g \circ f$, em termos das funções componentes f e g. Diga-se de passagem, cada uma dessas funções componentes é uma coleção de pares ordenados, $f \subset X \times Y$ e $g \subset U \times V$, sujeitos às condições da Definição 1.46 de funções. Na linguagem dos pares ordenados, $g \circ f$ nada mais é do que o conjunto:

$$g \circ f = \{(x, v) \in X \times V; \exists u \in U \,\text{tal que}\, (x, u) \in f \,\text{e}\, (u, v) \in g\},$$

que é consistente com a definição de função composta dada acima: dado um $x \in X$ qualquer, para algum $u \in U$, $(x, u) \in f$ significa o mesmo que $u = f(x)$ e, para algum $v \in V$, $(u, v) \in g$ significa o mesmo que $v = g(u)$; ou seja, para todo $x \in X$, existe algum $v \in V$ tal que $v = g(f(x))$.

Incidentalmente, ao longo desta seção, você perceberá o valor que possui uma boa notação matemática, i.e., a notação $g \circ f$ será utilizada para denotar a composta de g com f, enquanto a notação $g \cdot f$ ou simplesmente gf será utilizada para denotar a multiplicação de g por f, que definiremos adiante.

Vamos evocar novamente nosso diagrama de setas para ilustrar a ideia de composição de funções (Figura 1.32).

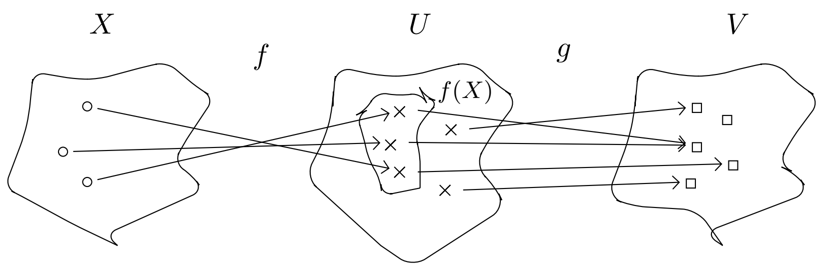


Figura 1.32: Composição de funções.

Em certo sentido, "numa única etapa", a composição acima ficaria mais ou menos como ilustrado na Figura 1.33.

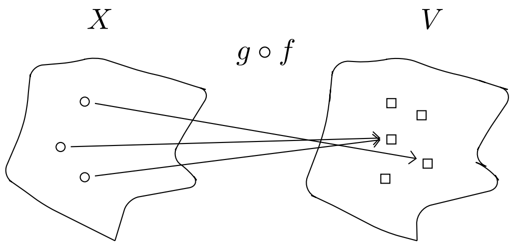


Figura 1.33: Composição de funções “numa única etapa”.

Note que, nessa definição, $g(f(x))$ só faz sentido devido à relação entre a imagem de f e o domínio de g. Em particular, $Y \subset U \Rightarrow f(X) \subset U$, portanto, numa situação em que o contradomínio de f esteja contido no domínio de g (hipótese mais comum apresentada nos livros-textos), a composição $g \circ f$ fará sentido. Observe que, em geral, provar que o contradomínio de f está contido no domínio de g é uma tarefa menos difícil do que provar que a imagem de f é um subconjunto do domínio de g pelo simples fato de que mesmo a imagem de uma simples função como $f : \mathbb{N} \to \mathbb{R}$, $f(x) = \sqrt{x}!$, pode ser muito complicada de se determinar.

Exemplo 1.68. Considere o conjunto X de todas as pessoas (vivas ou que já viveram) e $f : X \to X$, $x \mapsto y$, a função que significa “pai”, i.e., a função que associa a cada pessoa $x \in X$ o seu pai $y \in X$. O que significa a função composta de f com f, $f \circ f : X \to X$? Essa função associa cada pessoa $x \in X$ ao pai do seu pai $z = f(f(x)) \in X$, ou seja, ao seu avô. ◇

O estudante atento deve ter observado que a ordem dos eventos é importante na composição de funções. A fim de que $g \circ f$ esteja definida, a imagem de f deve estar contida no domínio de g, o que pode ocorrer sem que necessariamente o mesmo aconteça quando se trocam os papeis de f e g.

Exemplo 1.69. Seja $f : [0, +\infty) \to \mathbb{R}$ definida por $f(x) = \sqrt[4]{x-1}$ e seja $g : \mathbb{R} \to \mathbb{R}$ dada por $g(x) = \operatorname{sen} x^5$. A função $g \circ f$ faz sentido, pois a imagem de f está contida no domínio de g e $(g \circ f)(x) = \operatorname{sen}(x-1)^{5/4}$, mas não podemos considerar $f \circ g$ porque a imagem de g (a saber, $[-1, 1]$) não está contida no domínio de f. ◇

Ainda que ambas $g \circ f$ e $f \circ g$ estejam definidas ao mesmo tempo, se f mapeia X em Y e g mapeia Y em X, as funções $g \circ f$ e $f \circ g$ não precisam ser iguais; em outras palavras, a composição de funções não é necessariamente comutativa.

Exemplo 1.70. Seja $f : \mathbb{R} \to [0, +\infty)$ dada por $f(x) = x^2 + 2x + 1$ e seja $g : [0, +\infty) \to \mathbb{R}$ dada por $g(x) = \sqrt{x} - 4$. Note que f e g satisfazem o paradigma especificado na definição de composição de funções. Então, $g \circ f : \mathbb{R} \to \mathbb{R}$ é dada por:

$$(g \circ f)(x) = g(f(x)) = g(x^2 + 2x + 1) = \sqrt{x^2 + 2x + 1} - 4 = |x + 1| - 4.$$

Note que $f \circ g : [0, +\infty) \to [0, +\infty)$ também faz sentido e é dada por:

$$(f \circ g)(x) = f(g(x)) = f(\sqrt{x} - 4) = (\sqrt{x} - 4)^2 + 2(\sqrt{x} - 4) + 1 = (\sqrt{x} - 3)^2.$$

Este exemplo ilustra bem isto: $f \circ g$ e $g \circ f$, quando ambas fizerem sentido, serão, em geral, diferentes. $\diamond$

Por outro lado, a composição de funções sempre é associativa. Esse é o conteúdo da próxima proposição.

Proposição 1.71. *Sejam $f : X \to Y$, $g : U \to V$ e $h : W \to Z$ funções, com $f(X) \subset U$ e $g(U) \subset W$. Então as funções compostas $h \circ (g \circ f)$, $(h \circ g) \circ f : X \to Z$ estão definidas e vale a lei associativa: $h \circ (g \circ f) = (h \circ g) \circ f$.*

Demonstração. As hipóteses $f(X) \subset U$ e $g(U) \subset W$ garantem que as funções em questão estão bem definidas. Agora, para todo $x \in X$,

$$(h \circ (g \circ f))(x) = h((g \circ f)(x)) = h(g(f(x))) = (h \circ g)(f(x)) = ((h \circ g) \circ f)(x),$$

ou seja, $h \circ (g \circ f) = (h \circ g) \circ f$ para todo $x \in X$. $\blacksquare$

Para ilustrar essa ideia, considere o diagrama da Figura 1.34.

$$\begin{tikzcd}
X \arrow[rr, "g \circ f"] \arrow[d, "f"'] \arrow[drr, "(h \circ g) \circ f = h \circ (g \circ f)"] & & g(f(X)) \arrow[d, "h"] \\
f(X) \arrow[urr, "g"] \arrow[rr, "h \circ g"'] & & h(g(f(X)))
\end{tikzcd}$$

Figura 1.34: Associatividade da composição de funções.

Exemplo 1.72. Dissemos acima que podemos combinar f, g e h contanto que suas imagens e domínios se "encaixem" apropriadamente. É fácil ver o que isso significa: a imagem de f deve estar contida no domínio de g e a imagem de g, por sua vez, deve estar contida no domínio de h. Considere novamente a fórmula

$$\log(\log(\operatorname{sen}(x))).$$

Ela é obtida a partir da combinação das funções sen, $\log$ e $\log$ novamente. Se fizermos $f = \operatorname{sen}\colon \mathbb{R} \to \mathbb{R}$ e $g = h = \log\colon \mathbb{R}^+ \to \mathbb{R}$, teremos $\log(\log(\operatorname{sen}(x))) = (h \circ g \circ f)(x)$. Isso faz algum sentido? As condições que nos permitem fazer a composição dessa funções são violadas em várias ocasiões: a imagem de f é $[-1, 1]$, que não está contida no domínio de g; a imagem de g, por sua vez, é $\mathbb{R}$, que não está contido no domínio de h. Portanto a fórmula acima é uma fraude. Você deve sempre se assegurar de que o paradigma da definição de composição de funções é válido, sob pena de estar trabalhando com uma função que não existe! Aliás, essa lição vale para qualquer objeto matemático! $\diamond$

Outros fatos sobre a composição de funções estão contidos nas proposições a seguir.

Proposição 1.73. *Sejam $f : X \to Y$ e $g : U \to V$ funções, com $f(X) \subset U$, e $g \circ f : X \to V$ a composta de g com f.*

(c) se f e g são injetivas, então $g \circ f$ é injetiva;

(d) se f e g são sobrejetivas, então $g \circ f$ é sobrejetiva.

A demonstração dessa proposição é imediata.

Corolário 1.74. *Sejam $f : X \to Y$ e $g : U \to V$ funções, com $f(X) \subset U$. Se f e g são bijetivas, então $g \circ f$ é bijetiva.*

Proposição 1.75. *Toda função $f : X \to Y$ pode ser escrita como composta de uma função injetiva com uma função sobrejetiva.*

Demonstração. Basta considerar as funções $g : X \to f(X)$, definida por $g(x) = f(x)$ para todo $x \in X$, e a inclusão $i : f(X) \to Y$. Assim, $f = i \circ g$. $\blacksquare$

Proposição 1.76. *Toda função $f : X \to Y$ pode ser escrita como composta de uma função sobrejetiva com uma função injetiva.*

Demonstração. Basta considerar as funções $\phi : X \to X \times Y$, definida por $\phi(x) = (x, f(x))$ para todo $x \in X$, e $\pi_2 : X \times Y \to Y$, com $\pi_2(x, y) = y$ para todo $(x, y) \in X \times Y$ (segunda projeção). Agora considere $f = \pi_2 \circ \phi$. $\blacksquare$

Proposição 1.77. *Dadas as funções $f : X \to Y$ e $g : U \to V$, com $f(X) \subset U$ e $A \subset X$, tem-se $(g \circ f)(A) = g(f(A))$.*

Demonstração. Isso é fácil de ver. Basta considerar a definição de imagem de um conjunto por uma função:

$$\begin{aligned}(g \circ f)(A) &:= \{v \in V;\, v = (g \circ f)(a) = g(f(a)), \text{ para algum } a \in A\} \\ &= \{v \in V;\, v = g(f(a)), \text{ para algum } f(a) \in f(A)\} := g(f(A)).\end{aligned}$$ $\blacksquare$

Note que, na Definição 1.67, não há restrição quanto ao conjuntos que são os domínios e contradomínios das funções, nem quanto à natureza delas. Particularmente, isso significa que a composição de funções também vale para funções de várias variáveis.

Sejam as funções ϕ, ψ, $\ldots : X \to Y$, $\xi = \phi(x, y, \ldots)$, $\eta = \psi(x, y, \ldots)$, $\ldots$ definidas numa certa região X das variáveis independentes x, y, $\ldots$. Suponha que, conforme o argumento $(x, y, \ldots)$ varia em X, o ponto de coordenadas $(\xi, \eta, \ldots)$ varia numa região $Z \subset U$ na qual a função $f : U \to V$ está definida. Então a função composta de f com ϕ, ψ, $\ldots$, que chamaremos $F : X \to V$,

$$u = f(\phi(x, y, \ldots), \psi(x, y, \ldots), \ldots) = F(x, y, \ldots),$$

está definida em X.

1.7.2 Operações com funções

No caso particular em que $f : U \to V$ é uma função binária (uma função de duas variáveis), se $\phi, \psi : X \to Y$ são funções tais que $\phi(X) \subset U$ e $\psi(X) \subset U$, podemos definir:

$$f(\phi, \psi)(x) := f(\phi(x), \psi(x)),$$

ou, na notação mesofixa (ou infixa):

$$(\phi f \psi)(x) := \phi(x) f \psi(x).$$

Isso é o que chamamos de *operação com as funções componentes ϕ e ψ.*

Exemplo 1.78. Sejam $\phi, \psi : X \to Y$ e $+, -, \cdot, /: U \to V$ funções que satisfazem o paradigma da definição de composição de funções. Utilizando a notação mesofixa, definimos:

(a) $(\phi + \psi)(x) := \phi(x) + \psi(x)$;

(b) $(\phi - \psi)(x) := \phi(x) - \psi(x)$;

(c) $(\phi \cdot \psi)(x) := \phi(x) \cdot \psi(x)$;

(d) $(\phi/\psi)(x) := \phi(x)/\psi(x)$, contanto que $\psi(x) \neq 0$ em X; ◇

Note que em cada um dos itens acima foi definida uma nova função ($\phi + \psi$, $\phi - \psi$, $\phi \cdot \psi$ e ϕ/ψ) em termos das funções componentes ϕ e ψ. Cada uma dessas funções componentes é uma coleção de pares ordenados de $X \times Y$ sujeitos às condições do paradigma da definição de funções. Na linguagem dos pares ordenados, $\phi + \psi$ nada mais é do que o conjunto:

$$\phi + \psi = \{(x, y + y'); (x, y) \in \phi, (x, y') \in \psi\}.$$

Assim, de uma maneira (bastante) informal, podemos dizer que a operação é feita "ponto a ponto": para cada x no domínio de ϕ e ψ, aplica-se f nas imagens correspondentes.

As outras operações se expressam de maneira similar nesses termos e representá-las assim fica como exercício.

Exemplo 1.79. Sejam $+, -, \cdot, / : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ e $\phi,\ \psi : \mathbb{R} \to \mathbb{R}$ tais que $\phi(x) = x - 1$ e $\psi(x) = \operatorname{sen} x^2$. Vamos definir $\phi + \psi$, $\phi - \psi$, $\phi \cdot \psi$ e ϕ/ψ:

(a) $\phi + \psi : \mathbb{R} \to \mathbb{R}$, $(\phi + \psi)(x) := (x + 1) + \operatorname{sen} x^2$;

(b) $\phi - \psi : \mathbb{R} \to \mathbb{R}$, $(\phi - \psi)(x) := (x + 1) - \operatorname{sen} x^2$;

(c) $\phi \cdot \psi : \mathbb{R} \to \mathbb{R}$, $(\phi \cdot \psi)(x) := (x + 1) \cdot \operatorname{sen} x^2$;

(d) $\phi/\psi : \mathbb{R}/\{x \in \mathbb{R};\ x \neq \pm\sqrt{k\pi},\ k = 0, 1, \ldots\} \to \mathbb{R}$, $(\phi/\psi)(x) := (x + 1)/\operatorname{sen} x^2$. $\diamond$

1.8 Função inversa

Considere as funções

$$\begin{aligned} u : \mathbb{R} &\to [0, +\infty) \\ x &\mapsto x^2 \end{aligned} \qquad \text{e} \qquad \begin{aligned} v : [0, +\infty) &\to \mathbb{R} \\ x &\mapsto \sqrt{x}. \end{aligned}$$

Você já deve ter ouvido expressões do tipo "$\sqrt{x}$ é a função inversa de x^2". Mas também já deve ter sido apresentado ao teorema segundo o qual "uma função possui inversa se, e somente se, é bijetiva". Como u não é injetiva, não é bijetiva, logo não pode ser a inversa de v. Há alguma incoerência, então? Precisamos recorrer à definição de função inversa para responder essa questão.

Definição 1.80. *Dizemos que uma função $f : X \to Y$ é* invertível *quando existe uma função $g : Y \to X$ tal que*

(a) $f \circ g = id_Y : y \in Y \mapsto y \in Y$, ou seja, quando $f(g(y)) = y$ para todo $y \in Y$;

(b) $g \circ f = id_X : x \in X \mapsto x \in X$, ou seja, quando $g(f(x)) = x$ para todo $x \in X$.

Também dizemos que:

1. g é uma *inversa à direita* de f ou f é uma *inversa à esquerda* de g quando se cumpre a condição (a) acima;

2. g é uma *inversa à esquerda* de f ou f é uma *inversa à direita* de g quando se cumpre a condição (b) acima;

3. g é a[39] *inversa* de f ou f é a[39] *inversa* de g quando se cumprem ambas as condições (a) e (b) acima.

Agora, vamos retomar o exemplo inicial desta seção. De acordo com a Definição 1.80, para verificar se u e v são inversas uma da outra, devemos determinar as compostas $u \circ v$ e $v \circ u$:

$$\begin{array}{rcccc} u \circ v: & [0, +\infty) & \to & \mathbb{R} & \to & [0, +\infty) \\ & x & \mapsto & \sqrt{x} & \mapsto & (\sqrt{x})^2 = x, \end{array}$$

$$\begin{array}{rcccc} v \circ u: & \mathbb{R} & \to & [0, +\infty) & \to & \mathbb{R} \\ & x & \mapsto & x^2 & \mapsto & \sqrt{x^2} = |x|, \end{array}$$

ou seja, $u \circ v = id_{[0,\, +\infty)}$ mas $v \circ u \neq id_{\mathbb{R}}$, portanto u e v não são inversas uma da outra e, de modo mais geral, concluímos que u e v não são invertíveis.

Provar que uma função não é invertível a partir da Definição 1.80, por via de regra, não é uma tarefa fácil, pois devemos mostrar que não existe função alguma satisfazendo as duas condições daquela definição. Por isso, é importante entender que injetividade e sobrejetividade são condições suficientes (e necessárias) para garantir a existência da função inversa, como provaremos mais adiante.

Antes, note que u é sobrejetiva, mas não é injetiva; e que v é injetiva, mas não é sobrejetiva. Há uma íntima relação entre injetividade (resp. sobrejetividade) e a existência de inversas à esquerda (resp. à direita).

Apenas para ilustrar as ideias, pense por um momento no diagrama de setas de uma função $f : X \to Y$. Inverter o sentido de todas as setas que representam f pode ou não resultar numa nova função. Se f não for sobrejetiva, haverá algum elemento de Y do qual seta alguma partirá, portanto, haverá exceção; por outro lado, se f não for injetiva, haverá algum elemento de Y do qual mais de uma seta partirá, portanto, haverá ambiguidade; em qualquer caso, não obteremos uma função. Assim, intuitivamente, vemos a necessidade que f tem de ser bijetiva para admitir inversa. Agora, vamos tornar essa ideia de inverter o sentido das setas algo matematicamente respeitável.

Para que uma função admita uma inversa à esquerda, é necessário e suficiente que ela seja injetiva. Esse é o conteúdo da proposição adiante.

Proposição 1.81. *Uma função $f : X \to Y$ possui inversa à esquerda se, e somente se, é injetiva.*

[39] O uso do artigo definido "a" aqui se justifica porque a inversa de uma função, quando existe, é única, como provaremos adiante.

Demonstração. ($\Rightarrow$) Se existe $g : Y \to X$ com $g \circ f = id_X$, considerando-se $x_1, x_2 \in X$ tais que $f(x_1) = f(x_2)$, então $f(x_1) = f(x_2) \Rightarrow (g \circ f)(x_1) = (g \circ f)(x_2) \Rightarrow id_X(x_1) = id_X(x_2) \Rightarrow x_1 = x_2$. Logo, f é injetiva.

($\Leftarrow$) Considere $y \in Y$ qualquer. Ou $y \in f(X)$ ou $y \notin f(X)$. Se f é injetiva, no primeiro caso, existe apenas um $x \in X$ tal que $y = f(x)$, de modo que podemos definir $g(y) = x$ se $y \in f(X)$; no segundo caso, podemos fixar qualquer $x_0 \in X$ e definir $g(y) = x_0$ se $y \in Y \setminus f(X)$. Assim, construímos uma função $g : Y \to X$ com a propriedade $g \circ f = id_X$. Logo, f possui uma inversa à esquerda. ■

A inversa à esquerda, quando existe, não é única, como mostra o próximo exemplo.

Exemplo 1.82 (Inversa à esquerda). Considere a função $p : [0, +\infty) \to \mathbb{R}$, definida por $p(x) = x^2$, e a função $q : \mathbb{R} \to [0, +\infty)$, definida por $q(y) = \sqrt{x}$ se $y \geqslant 0$ e $q(y) = 0$ se $y < 0$. Certamente, $q(p(x)) = q(x^2) = \sqrt{x^2} = x$, para todo $x \in [0, +\infty)$. Assim, $q \circ p = id_{[0, +\infty)}$, logo q é uma inversa à esquerda de p. Note-se que qualquer função $r : \mathbb{R} \to [0, +\infty)$, tal que $r(y) = \sqrt{y}$ para $y \geqslant 0$, é uma inversa à esqueda de p, pois a definição de r para $y < 0$ não afeta a igualdade $r \circ p = id_{[0, +\infty)}$. Note que p é injetiva, mas não é sobrejetiva. ◇

Para que uma função admita uma inversa à direita, é necessário e suficiente que ela seja sobrejetiva. Mostramos isso na proposição a seguir.

Proposição 1.83. *Uma função $f : X \to Y$ possui inversa à direita se, e somente se, é sobrejetiva.*

Demonstração. ($\Rightarrow$) Se existe $g : Y \to X$ com $f \circ g = id_Y$, então, para cada $y \in Y$, pondo $x = g(y)$, temos $f(x) = f(g(y)) = (f \circ g)(y) = id_Y(y) = y$. Logo, f é sobrejetiva.

($\Leftarrow$) Se f é sobrejetiva, então, dado $y \in Y$ qualquer, existe $x \in X$ tal que $f(x) = y$. Para cada $y \in Y$, escolha (apenas) um $x \in X$ tal que $f(x) = y$ e faça $g(y) = x$. Isso define uma funçao $g : Y \to X$ tal que $f \circ g = id_Y$. Logo g é uma inversa à direita de f. ■

Assim como a inversa à esquerda, a inversa à direita, quando existir, não será única, como ilusta o exemplo abaixo.

Exemplo 1.84 (Inversa à direita, [7]). Seja $p : \mathbb{N} \to \mathbb{N}$ definida por $p(1) = 1$ e, se $x > 1$, $p(x) =$ *número de fatores primos distintos que entram na composição de* x. Seja $q : \mathbb{N} \to \mathbb{N}$ definida por $q(y) =$ *menor número natural que é o produto de* y *fatores primos distintos.* Então, para todo $n \in \mathbb{N}$, $p(q(y)) = y$, ou seja, $p \circ q = id_{\mathbb{N}}$ e, portanto, q é uma inversa à direita para p. Mas poderíamos ter definido outra função $r : \mathbb{N} \to \mathbb{N}$ com a propriedade $p \circ r = id_{\mathbb{N}}$. Por exemplo, fazendo $r(y) =$ *menor número natural* *divisível por* 17 *que é o produto de y fatores primos distintos.* Note que p é sobrejetiva, mas não é injetiva. ◇

Assim, o próximo teorema segue das duas proposições anteriores.

Teorema 1.85. *Uma função é invertível se, e somente se, é bijetiva.*

Ao contrário das inversas à esquerda e à direita, a inversa de uma função, se existir, será única.

Proposição 1.86. *Se uma função $f : X \to Y$ possui uma inversa, ela é única.*

Demonstração. De fato, vamos supor que $g : Y \to X$ e $h : Y \to X$ são ambas inversas de f. Então, usando a Proposição 1.71 (na terceira igualdade a seguir), $h = h \circ id_Y = h \circ (f \circ g) = (h \circ f) \circ g = id_X \circ g = g$. ■

Como você deve ter notado, ficou provado um pouco mais do que o enunciado na proposição, a saber: se f possui uma inversa à esquerda, h, e uma inversa à direita, g, então $h = g$ e f possui uma inversa.

Exemplo 1.87. Seja a bijeção $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto ax + b$, com $a \neq 0$. Então a função $g : \mathbb{R} \to \mathbb{R}$, $x \mapsto (x - b)/a$, é a inversa de f. ◇

Exemplo 1.88 ([7]). Dada uma função arbitrária $f : X \to Y$, seja $\Gamma(f)$ o gráfico de f. Definimos $F : X \to \Gamma(f)$ pondo $F(x) = (x, f(x))$. Seja $\pi : \Gamma(f) \to X$ definida por $\pi(x, f(x)) = x$. Então $F \circ \pi = id_{\Gamma(f)}$ e $\pi \circ F = id_X$. Portanto, π é a inversa de F. ◇

Notação 1.89. Se $g : Y \to X$ é a inversa de uma bijeção $f : X \to Y$, devido a essa unicidade da inversa e às igualdades $f \circ g = id_Y$ e $g \circ f = id_X$, adota-se a notação do tipo exponencial f^{-1} para a função g, i.e., escrevemos $f^{-1} : Y \to X$ para indicar a inversa da bijeção $f : X \to Y$. ○

Há uma conexão entre compostas e inversas como mostra a proposição a seguir.

Proposição 1.90. *Se $f : X \to Y$ e $g : Y \to Z$ são bijeções, então $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.*

Demonstração. Basta utilizar a propriedade associativa da composição (Proposição 1.71):

$$(f^{-1} \circ g^{-1}) \circ (g \circ f) = f^{-1} \circ [(g^{-1} \circ g) \circ f] = f^{-1} \circ (id_Y \circ f) = f^{-1} \circ f = id_X;$$

$$(g \circ f) \circ (f^{-1} \circ g^{-1}) = g \circ [(f \circ f^{-1}) \circ g^{-1}] = g \circ (id_Y \circ g^{-1}) = g \circ g^{-1} = id_Z.$$

Portanto, $f^{-1} \circ g^{-1}$ é a inversa de $g \circ f$. ■

1.9 Imagens e imagens inversas

Se f é uma função de X em Y e A é um subconjunto de X, podemos estar interessados no conjunto de todos aqueles elementos y de Y para os quais existe algum x em A tal que $y = f(x)$. Isso motiva nossa próxima definição.

Definição 1.91. *Sejam $f : X \to Y$ uma função e $A \subset X$. Chamamos de* imagem de *A pela função f ao conjunto $f(A)$ formado pelos valores $f(x)$ que f assume nos pontos $x \in A$. Assim:*

$$f(A) := \{f(x);\, x \in A\} = \{y \in Y;\, y = f(x),\, x \in A\}.$$

A notação $f(A)$ para a imagem de A por f é ruim mas não totalmente desastrosa. O problema está no fato de o símbolo $f(A)$ ser ambíguo, pois A pode ocorrer tanto como um elemento quanto um subconjunto de X (algo improvável, porém não impossível). Seguindo os usos e costumes matemáticos, utilizaremos a notação ruim, ficando a desambiguação por conta do contexto ou de estipulações verbais quando necessário para evitar confusão.

Evidentemente, $f(A) \subset Y$ e a condição necessária e suficiente para que f seja sobrejetiva é que $f(X) = Y$ (em geral ocorre apenas $f(X) \subset Y$). Como já vimos, $f(X)$ é chamado a *imagem* ou o *conjunto de valores* ou ainda o *campo de valores* de f.

Exemplo 1.92. Seja $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^2$. Então $f(\mathbb{R}) = [0, +\infty)$ (assumindo que todo número real positivo possui uma raiz quadrada). Também, por exemplo, $f(\mathbb{R}^+) = f(\mathbb{R}^-) = \mathbb{R}^+$, $f(\{0\}) = \{0\}$ e $f([0, 1]) = [0, 1]$. $\diamond$

A imagem de um conjunto por uma função possui algumas propriedades de fácil demonstração como veremos a seguir.

Proposição 1.93. *Dada uma função $f : X \to Y$ e sendo $A, B \subset X$, temos:*

(a) $f(A \cup B) = f(A) \cup f(B)$;

(b) $f(A \cap B) \subset f(A) \cap f(B)$;

(c) f *é injetiva* $\Rightarrow f(A \cap B) = f(A) \cap f(B)$;

(d) $f(A) \setminus f(B) \subset f(A \setminus B)$;

(e) f *é injetiva* $\Rightarrow f(A) \setminus f(B) = f(A \setminus B)$;

(f) $A \subset B \Rightarrow f(A) \subset f(B)$;

(g) $f(\emptyset) = \emptyset$;

(h) f *é injetiva* $\Rightarrow f(A^{\complement}) \subset (f(A))^{\complement}$*;*

(i) f *é sobrejetiva* $\Rightarrow (f(A^{\complement}))^{\complement} \subset f(A^{\complement})$*;*

em que $A^{\complement} = X \setminus A$ *e* $(f(A))^{\complement} = Y \setminus f(A)$*.*

Demonstração.

(a) Se $y \in f(A \cup B)$, então existe $x \in A \cup B$ tal que $y = f(x)$. Se $x \in A$, então $y \in f(A)$; se $x \in B$, então $y \in f(B)$. Em qualquer caso, $y \in f(A) \cup f(B)$. Logo, $f(A \cup B) \subset f(A) \cup f(B)$. Reciprocamente, seja $c \in f(A) \cup f(B)$. Então $c \in f(A)$, logo existe $a \in A$ tal que $c = f(a)$, ou $c \in f(B)$, logo existe $b \in B$ tal que $c = f(b)$. Em qualquer hipótese, existe $d \in A \cup B$ tal que $c = f(d) \in f(A \cup B)$, ou seja, $f(A) \cup f(B) \subset f(A \cup B)$. Portanto, $f(A \cup B) = f(A) \cup f(B)$.

(b) Se $y \in f(A \cap B)$, então existe $x \in A \cap B$ tal que $y = f(x)$. Consequentemente, $x \in A$ e portanto $y \in f(A)$, assim como $x \in B$ e $y \in f(B)$. Logo, $f(A \cap B) \subset f(A) \cap f(B)$.

(c) Se $c \in f(A) \cap f(B)$, então $c \in f(A)$, logo existe $a \in A$ tal que $c = f(a)$, e também $c \in f(B)$, logo existe $b \in B$ tal que $c = f(b)$. Como f é injetiva, $f(a) = c = f(b)$ implica $a = b$, logo existe $d(= a = b) \in A \cap B$ tal que $c = f(d)$, ou seja, $f(A) \cap f(B) \subset f(A \cap B)$. Como a recíproca é sempre verdadeira, como mostramos em (b), segue-se que $f(A \cap B) = f(A) \cap f(B)$.

(d) Se $y \in f(A) \setminus f(B)$, então $y \in f(A)$ e $y \notin f(B)$, logo existe $x \in A$ tal que $y = f(x)$ e $x \notin B$, ou seja, $x \in A \setminus B$ tal que $y = f(x)$. Portanto, $f(A) \setminus f(B) \subset f(A \setminus B)$.

(e) Se $y \in f(A \setminus B)$, existe $x \in A \setminus B$, ou seja, $x \in A$ e $x \notin B$ tal que $y = f(x)$. Como f é injetiva, $y \in f(A)$ mas $y \notin f(B)$, logo $y \in f(A) \setminus f(B)$. Assim, $f(A \setminus B) \subset f(A) \setminus f(B)$. Como a recíproca é sempre verdadeira, como mostramos em (d), segue-se que $f(A) \setminus f(B) = f(A \setminus B)$.

(f) Se $y \in f(A)$, existe $x \in A \subset B$, ou seja, $x \in B$ tal que $y = f(x)$, portanto $y \in f(B)$. Logo, $f(A) \subset f(B)$.

(g) Suponha, por absurdo, que $f(\emptyset) \neq \emptyset$. Então, para algum $y \in f(\emptyset)$, existe $x \in \emptyset$ tal que $y = f(x)$, absurdo. Portanto, $f(\emptyset) = \emptyset$.

(h) Temos $f(A^{\complement}) = f(X \setminus A)$. Como f é injetiva, segue de (e) que $f(A^{\complement}) \subset f(X) \setminus f(A)$. Como $f(X) \subset Y$, segue-se que $f(A^{\complement}) \subset Y \setminus f(A) = (f(A))^{\complement}$, ou seja, $f(A^{\complement}) \subset (f(A))^{\complement}$.

(i) Temos $(f(A))^{\complement} = Y \setminus f(A)$. Como f é sobrejetiva, $Y = f(X)$, logo, $(f(A))^{\complement} = f(X) \setminus f(A)$. Segue-se de (d) que $f(X) \setminus f(A) \subset f(X \setminus A) = f(A^{\complement})$. Portanto, $(f(A))^{\complement} \subset f(A^{\complement})$. ■

Corolário 1.94. *Se $f : X \to Y$ é uma função* bijetiva *e $A \subset X$, então $f(A^{\complement}) = (f(A))^{\complement}$.*

Na verdade, dada uma função qualquer f de X em Y, sempre existe associada a ela outra função, frequentemente também chamada f, de $\mathscr{P}(X)$ em $\mathscr{P}(Y)$ (em que $\mathscr{P}(X)$ e $\mathscr{P}(Y)$ são os conjuntos de todos os subconjuntos de X e Y respectivamente), que mapeia todo $A \subset X$ em $f(A)$. No entanto, como você deve ter notado na Proposição 1.93, o comportamento algébrico dessa correspondência deixa algo a desejar. Por exemplo, vimos que se $A, B \subset X$, então $f(A \cup B) = f(A) \cup f(B)$, mas nem sempre a equação correspondente vale para a intersecção, a não ser sob a condição de que f seja injetiva.

Ocorre, porém, que mesmo assim f sempre induz uma correspondência bem comportada entre $\mathscr{P}(X)$ e $\mathscr{P}(Y)$, não diretamente, pela formação de imagens, mas no sentido oposto, pela formação de imagens inversas.

Definição 1.95. *Seja uma função $f : X \to Y$ e $B \subset Y$. Chamamos de* imagem inversa *ou* pré-imagem de *B ao conjunto $f^{-1}(B)$ formado pelos valores $x \in X$ tais que $f(x) \in B$. Assim:*

$$f^{-1}(B) := \{x \in X;\ f(x) \in B\}.$$

Ou seja, $f^{-1}(B)$ consiste exatamente dos elementos de X que f mapeia em B.

Novamente, o uso do símbolo f^{-1} para designar a imagem inversa $f^{-1}(B)$ de um conjunto B é uma escolha muito infeliz, ainda que universalmente aceita, pois pode causar confusão com a noção de função inversa de f, que nem mesmo pode estar definida. Mais uma vez, vamos agir de acordo com os usos e costumes, ficando a distinção de um conceito do outro por conta do contexto ou de estipulações verbais. O estudante deve estar atento.

Exemplo 1.96 ([7]). Os subconjuntos do plano definidos por meio de equações e desigualdades são imagens inversas de conjuntos. Por exemplo, a reta cuja equação é $ax + by = c$ é o conjunto $X = \{(x, y) \in \mathbb{R}^2;\ ax + by = c\}$; considerando a função $f : \mathbb{R}^2 \to \mathbb{R}$, definida por $f(x, y) = ax + by$, a reta X é a imagem inversa de $\{c\}$ por f, ou seja, $X = f^{-1}(\{c\})$. O disco de centro na origem e raio 1 é o conjunto $D = \{(x, y) \in \mathbb{R}^2;\ x^2 + y^2 \leqslant 1\}$; considerando a função $g : \mathbb{R}^2 \to \mathbb{R}$, definida por $g(x, y) = x^2 + y^2$, o disco D é a imagem inversa do intervalo $I = [0, 1]$ pela função g: $D = g^{-1}(I)$. ◇

Vejamos as principais propriedades da imagem inversa na proposição a seguir.

Proposição 1.97. *Dada uma função $f : X \to Y$ e sendo $B, C \subset Y$, temos:*

(a) $f^{-1}(B \cup C) = f^{-1}(B) \cup f^{-1}(C)$*;*

(b) $f^{-1}(B \cap C) = f^{-1}(B) \cap f^{-1}(C)$*;*

(c) $B \subset C \Rightarrow f^{-1}(B) \subset f^{-1}(C)$*;*

(d) $f^{-1}(Y) = X$*;*

(e) $f^{-1}(\emptyset) = \emptyset$*;*

(f) $f^{-1}(B \setminus C) = f^{-1}(B) \setminus f^{-1}(C)$*;*

(g) $f^{-1}(B^{\complement}) = (f^{-1}(B))^{\complement}$*;*

em que $B^{\complement} = Y \setminus B$ *e* $(f^{-1}(B))^{\complement} = X \setminus f^{-1}(B)$*.*

Demonstração.

(a) $x \in f^{-1}(B \cup C) \Leftrightarrow f(x) \in B \cup C \Leftrightarrow f(x) \in B$ ou $f(x) \in C \Leftrightarrow x \in f^{-1}(B)$ ou $x \in f^{-1}(C) \Leftrightarrow x \in f^{-1}(B) \cup f^{-1}(C)$.

(b) $x \in f^{-1}(B \cap C) \Leftrightarrow f(x) \in B \cap C \Leftrightarrow f(x) \in B$ e $f(x) \in C \Leftrightarrow x \in f^{-1}(B)$ e $x \in f^{-1}(C) \Leftrightarrow x \in f^{-1}(B) \cap f^{-1}(C)$.

(c) $x \in f^{-1}(B) \Leftrightarrow f(x) \in B \subset C \Leftrightarrow f(x) \in C \Leftrightarrow x \in f^{-1}(C)$.

(d) Por definição, $f^{-1}(Y) = \{x \in X;\ f(x) \in Y\} = X$.

(e) Suponha, por absurdo, que $f^{-1}(\emptyset) \neq \emptyset$. Então, para algum $x \in f^{-1}(\emptyset)$, existe $y = f(x) \in \emptyset$, absurdo. Logo, $f^{-1}(\emptyset) = \emptyset$.

(f) $x \in f^{-1}(B \setminus C) \Leftrightarrow f(x) \in B \setminus C \Leftrightarrow f(x) \in B$ e $f(x) \notin C \Leftrightarrow x \in f^{-1}(B)$ e $x \notin f^{-1}(C) \Leftrightarrow x \in f^{-1}(B) \setminus f^{-1}(C)$.

(g) $f^{-1}(B^{\complement}) = f^{-1}(Y \setminus B) = f^{-1}(Y) \setminus f^{-1}(B) = X \setminus f^{-1}(B) = (f^{-1}(B))^{\complement}$, em que utilizamos (f) e (d) na segunda e terceira igualdade respectivamente. ■

Além destas, vale mencionar mais outras duas propriedades da imagem inversa que a conectam aos conceitos de funções sobrejetivas e injetivas.

Proposição 1.98. *Seja* $f : X \to Y$ *uma função. Uma condição necessária e suficiente para que* f *seja sobrejetiva é que a imagem inversa por* f *de cada subconjunto não vazio de* Y *seja um subconjunto não vazio de* X*.*

Demonstração. ($\Rightarrow$) Suponha, por absurdo, que existe $B \subset Y$ não vazio tal que $f^{-1}(B) = \emptyset$. Isso significa que existe $y \in B$, e portanto $y \in Y$, que não é imagem de x algum em X, absurdo, pois f é sobrejetiva.

($\Leftarrow$) Considere, em particular, cada subconjunto unitário de Y. Como todos eles possuem imagens inversas não vazias, isso significa que todo elemento $y \in Y$ é a imagem por f de (ao menos) um $x \in X$. Portanto, f é sobrejetiva. ∎

Proposição 1.99. *Seja $f : X \to Y$ uma função. Uma condição necessária e suficiente para que f seja injetiva é que a imagem inversa por f de cada conjunto unitário na sua imagem seja um conjunto unitário no seu domínio.*

Demonstração. ($\Rightarrow$) Suponha, por absurdo, que existe um conjunto unitário $B \subset f(X)$ tal que $f^{-1}(B) = A$ não seja um conjunto unitário (note que A não pode ser vazio, uma vez que B está na imagem de f). Então, existe x_1, $x_2 \in A$, com $x_1 \neq x_2$, tal que $f(x_1) = f(x_2)$, absurdo, pois f é injetiva.

($\Leftarrow$) Suponha, por absurdo, que f não seja injetiva. Então existe um conjunto unitário na sua imagem cuja imagem inversa por f não é unitária, absurdo. ∎

Sob as condições da última proposição, o símbolo f^{-1} assume aquela interpretação segundo a qual ela é a função inversa, cujo domínio é a imagem de f e que associa a cada y na imagem de f o único $x \in X$ tal que $y = f(x)$. Em outras palavras, para funções injetivas, podemos escrever $f^{-1}(y) = x$ (ao invés de $f^{-1}(\{y\}) = \{x\}$) sempre que $f(x) = y$. Como já dissemos acima, essa notação é ligeiramente inconsistente, mas não é suscetível de conduzir a qualquer confusão.

Vale a pena um momento a mais de consideração sobre a conexão entre imagens e imagens inversas.

Proposição 1.100. *Para uma função $f : X \to Y$ qualquer, valem:*

$$f(f^{-1}(B)) \subset B \quad e \quad A \subset f^{-1}(f(A)),$$

para todo $A \subset X$, $B \subset Y$.

Demonstração.
(a) Se $y \in f(f^{-1}(B))$, então $y = f(x)$ para algum $x \in f^{-1}(B)$, ou seja, $y = f(x)$ e $f(x) \in B$, logo $y \in B$. Portanto, $f(f^{-1}(B)) \subset B$.

(b) Se $x \in A$, então $f(x) \in f(A)$, ou seja, $x \in f^{-1}(f(A))$. Portanto, $A \subset f^{-1}(f(A))$. ∎

Proposição 1.101. *Se $f : X \to Y$ é uma função sobrejetiva, então $f(f^{-1}(B)) = B$, para todo $B \subset Y$.*

Demonstração. Se $y \in B$, então $y = f(x)$ para algum $x \in f^{-1}(B)$, pois f é sobrejetiva. Ou seja, $y \in f(f^{-1}(B))$. Assim, $B \subset f(f^{-1}(B))$. Como a recíproca é sempre verdadeira, conforme mostramos na Proposição 1.100, segue-se que $f(f^{-1}(B)) = B$. ∎

Proposição 1.102. *Se $f : X \to Y$ é uma função injetiva, então $A = f^{-1}(f(A))$, para todo $A \subset X$.*

Demonstração. Se $x \in f^{-1}(f(A))$, então $f(x) \in f(A)$, logo $f(x) = f(a)$ para algum $a \in A$. Como f é injetiva, segue-se que $x = a$ e portanto $x \in A$. Assim, $f^{-1}(f(A)) \subset A$. Como a recíproca é sempre verdadeira, conforme mostramos na Proposição 1.100, segue-se que $A = f^{-1}(f(A))$. ■

Corolário 1.103. *Se $f : X \to Y$ é uma função bijetiva, então $f(f^{-1}(B)) = B$ e $A = f^{-1}(f(A))$, para todo $A \subset X$, $B \subset Y$.*

Vamos provar uma última proposição relacionando imagens inversas e compostas.

Proposição 1.104. *Sejam $f : X \to Y$ e $g : U \to V$ funções, com $Y \subset U$. Se $Z \subset V$, então $(g \circ f)^{-1}(Z) = f^{-1}(g^{-1}(Z))$.*

Demonstração. $x \in (g \circ f)^{-1}(Z) \Leftrightarrow (g \circ f)(x) \in Z \Leftrightarrow g(f(x)) \in Z \Leftrightarrow f(x) \in g^{-1}(Z) \Leftrightarrow x \in f^{-1}(g^{-1}(Z))$. ■

Tendo em vista o comentário logo após a Proposição 1.99, note que a Proposição 1.90 é um caso particular da Proposição 1.104.

1.10 Restrições e extensões de funções

1.10.1 Restrições de funções

Considere as funções inclusão de $X \subset Y$ em Y e identidade em Y, respectivamente $i : X \subset Y \to Y$, definida por $i(x) = x$, e $id_Y : Y \to Y$, dada por $id_Y(x) = x$. Você deve ter notado que, nesse caso, a função inclusão é o mesmo que a função identidade quando restringimos o domínio desta última a X.

Essa conexão entre ambas ilustra um procedimento geral de se obter funções menores a partir de funções maiores. Dada uma função $f : Y \to Z$ e $X \subset Y$, esse procedimento consiste em se definir de maneira natural uma função $g : X \to Z$ fazendo-se $g(x) = f(x)$ para cada $x \in X$.

Definição 1.105. *Seja uma função $f : Y \to Z$, $X \subset Y$. Uma função $g : X \to Z$ tal que $g(x) = f(x)$ para cada $x \in X$ chama-se a* restrição *de f a X.*

Notação 1.106. A função g é comumente denotada por $f|X$. Alguns autores utilizam também $f \upharpoonright X$, ou $f|_X$ ou ainda $f \upharpoonright_X$. Adotaremos a primeira notação. ○

Assim, a restrição de uma função $f : Y \to Z$ a um conjunto $X \subset Y$ é expressa como $(f|X)(x) = f(x)$ para cada $x \in X$. Note também que $(f|X) = f(X)$, ou seja, a imagem da restrição de f a X é igual à imagem de X por f.

A ideia da restrição de uma função f a um subconjunto X do seu domínio pode ser ilustrada como na Figura 1.35, em que as setas de traço sólido representam $f|X$.

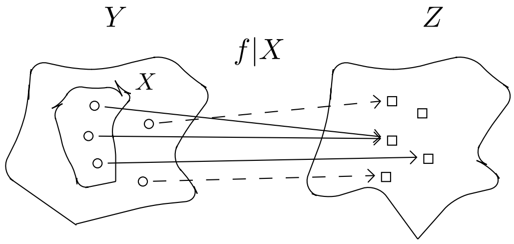


Figura 1.35: Ideia da restrição de funções.

Na linguagem dos pares ordenados, $f|X := \{(x,\, f(x)) \in Y \times Z;\, x \in X \subset Y\} \subset X \times Z$.

Existe uma conexão entre a função inclusão $i : X \subset Y \to Y$ e a restrição de $f : Y \to Z$ a $X \subset Y$, expresso na seguinte proposição, cuja demonstração é imediata.

Proposição 1.107. *Sejam $f : Y \to Z$ uma função qualquer e $i : X \subset Y \to Y$ a função inclusão. Então, $f|X = f \circ i : X \to Z$.*

Exemplo 1.108. Seja a função não injetiva $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto x^2$. A restrição de f a $[0, +\infty)$ é a função injetiva $f|[0, +\infty) : [0, +\infty) \to \mathbb{R}$, $x \mapsto x^2$. $\diamond$

Exemplo 1.109. Seja $f : X \to Y$ uma função arbitrária e $\Gamma(f)$ o seu gráfico. A função $\pi : \Gamma(f) \to X$, definida por $\pi(x,\, f(x)) = x$, é a restrição $\pi = \pi_1|\Gamma(f)$ a $\Gamma(f)$ da primeira projeção $\pi_1 : A \times B \to A$. $\diamond$

Exemplo 1.110. A função fatorial é a restrição da função gama[40] ao conjunto dos números naturais. $\diamond$

1.10.2 Extensões de funções

Considere novamente as funções inclusão e identidade como no início da subseção anterior. Vimos que ambas coincidem na parte comum de seus domínios, i.e., no

[40] A função gama é importante em muitas aplicações, nas ciências e engenharias. Ela é definida para todos os números complexos, exceto para os inteiros não positivos. Para números complexos z cuja parte real é positiva ela é dada pela integral imprópria $\Gamma(z) = \int_0^{+\infty} e^{-x} x^{z-1}\, dx$ e então é extendida analiticamente para todos os complexos, exceto os inteiros não positivos, em que a função tem polos simples. Você terá a oportunidade de estudar a função gama e esses conceitos subjacentes em cursos mais avançados.

conjunto $X \subset Y$. Naquela ocasião, estávamos interessados no processo de obter uma função "menor" a partir de uma função "maior"[41], o que nos motivou olhar para a definição de restrição de uma função. Agora, estamos interessados em abordar esse processo a partir de outra perspectiva, que nos permita ir no sentido oposto e obter funções "maiores" a partir de funções "menores"[41].

Esse é justamente o conteúdo da próxima definição.

Definição 1.111. *Seja $f : X \to Z$ uma função com $X \subset Y$. Uma função $g : Y \to Z$ é dita ser uma* extensão *de f se f e g coincidirem na parte comum de seus domínios, que vem a ser o conjunto X, i.e., se $f(x) = g(x)$ para todo $x \in X$. Nesse caso, dizemos também que f é* estendida *por g.*

Você deve ter notado que estender uma função $f : X \to Z$ a um conjunto Y obtendo uma função $g : Y \to Z$ equivale a dizer que f é a restrição de g a X, ou seja, que $f = g|X$.

Evidentemente existem, em geral, várias extensões da mesma função, o que justifica o uso do artigo indefinido "uma" quando nos referimos a extensões, em oposição ao uso do artigo definido "a" quando restringimos uma função a um conjunto dado.

Em termos de conjuntos, lembrando que uma função $f : X \to Z$ é um subconjunto de $X \times Z$, que uma função $g : Y \to Z$ é um subconjunto de $Y \times Z$ e notando que $X \subset Y$ implica $X \times Z \subset Y \times Z$, poderíamos definir extensão alternativamente da seguinte maneira: dizemos que uma função g é uma extensão de uma função f quando $f \subset g$, ambas entendidas como subconjuntos de $Y \times Z$. Pode ser mostrado que essa definição é equivalente à Definição 1.111.

O conceito de extensão é frequentemente usado em matemática, como na teoria das equações diferenciais parciais, na teoria das funções de variáveis complexas, e na teoria dos operadores lineares em espaços de Hilbert[42] para mencionar apenas algumas. O emprego desse conceito é feito para que as extensões satisfaçam a certas condições adicionais como continuidade, analiticidade etc. A função que se deseja estender é chamada a "condição de contorno".

Exemplo 1.112. A função identidade $id_Y : Y \to Y$ é um exemplo de uma extensão da função inclusão $i : X \subset Y \to Y$ ao conjunto Y, como mencionado no início. $\diamond$

Exemplo 1.113. A extensão da função injetiva $f : [0, +\infty) \to \mathbb{R}$, $x \mapsto x^2$, ao conjunto $\mathbb{R}$ é a função não injetiva $g : \mathbb{R} \to \mathbb{R}$, $x \mapsto x^2$. $\diamond$

[41] Apenas uma observação sobre esses termos informais: se f é uma função e X um subconjunto do seu domínio, estamos dizendo informalmente que $f|X$ é a função "menor" obtida a partir da função "maior" f; na linguagem dos conjuntos, o uso desses termos é motivado pelo fato de que $f|X \subset f$.

[42] David Hilbert (1862–1943), matemático alemão, considerado um dos mais notáveis de todos os tempos. Foi membro estrangeiro da *Royal Society*. Seu trabalho em Geometria é considerado o de maior influência depois de Euclides. Fez contribuições em muitas áreas da Matemática e da Física.

Esse exemplo ilustra o fato de que estender uma função injetiva, em geral, não preserva essa propriedade. Mas se a função for sobrejetiva sim, pois nesse processo o contradomínio não muda.

Exemplo 1.114. A função gama é uma extensão da função fatorial, definida sobre o conjunto dos númeors naturais, ao conjunto dos números complexos exceto os inteiros não positivos. $\diamond$

Exemplo 1.115. Uma função muito encontrada em Cálculo é a função $f : \mathbb{R}^* \to \mathbb{R}$, $f(x) = (\operatorname{sen} x)/x$. Embora ela não esteja definida em $x = 0$, podemos estendê-la de forma natural a todo $\mathbb{R}$ tomando $f(0) = 1$. Com isso, essa função passa a ser contínua em toda a reta (veja esboço na Figura 1.36).

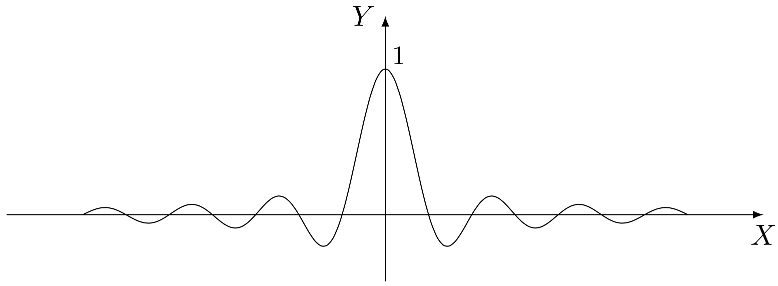


Figura 1.36: Exemplo de uma extensão.

Como dissemos acima, as extensões são empregadas de tal modo que as funções obtidas satisfaçam a certas condições adicionais. Um exemplo disso são as extensões periódicas, em particular as extensões pares e ímpares, de grande utilidade quando se trabalham com séries de Fourier[43]. Veja os próximos exemplos.

Exemplo 1.116 (Extensões periódicas). Se qualquer função arbitrária f é definida num intervalo arbitrário, digamos $[a, b]$, sempre podemos estendê-la a $\mathbb{R}$ como uma função periódica de período $p = b - a$: basta que se defina f fora desse intervalo por meio da equação $f(x + np) = f(x)$ para todo $x \in [a, b]$ e todo $n \in \mathbb{Z}$. Apenas um cuidado deve ser tomado a fim de que a nossa extensão seja de fato uma função, uma vez que, por exemplo, para todo $n \in \mathbb{Z}$, definimos $f(a + np)$ como $f(a)$ e também $f(a + np) = f(a + (n + 1)p) = f((a + p) + np) = f(b + np) = f(b)$, o que nos dá $f(a + np) = f(a)$ e $f(a + np) = f(b)$ para todo $n \in \mathbb{Z}$, e, como em geral $f(a) \neq f(b)$, haveria uma ambiguidade em todos os pontos da forma $a + np$. Esse cuidado adicional consiste em primeiro restringirmos f a $(a, b]$ ou a $[a, b)$ e em seguida procedermos à extensão periódica. $\diamond$

[43]Jean Baptiste Joseph Fourier (1768–1830), matemático francês, estudou a teoria matemática da condução de calor. Ele estabeleceu a equação diferencial parcial que rege a difusão de calor e a resolveu usando séries de funções trigonométricas.

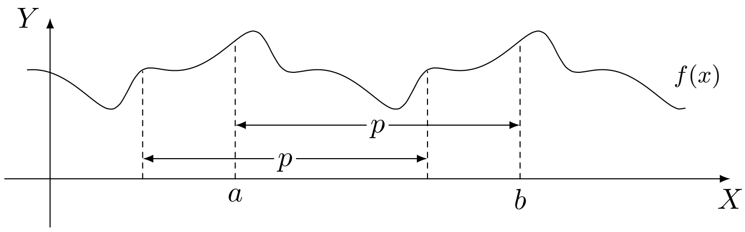


Figura 1.37: Exemplo de uma extensão periódica.

Exemplo 1.117 (Extensões periódicas pares e ímpares). Muitas extensões periódicas são possíveis, a maioria sendo inútil. Porém, há dois tipos particularmente úteis como mencionamos acima: as extensões periódicas pares (Figura 1.38) e ímpares (Figura 1.39). Dada uma função qualquer f definida sobre um intervalo I de extremos 0 e $l > 0$, definimos sua *extensão periódica par* a $\mathbb{R}^{44}$, f_p, por meio das condições:

(a) $f_p(x) = f(x)$, para todo $x \in I$;

(b) $f_p(-x) = f(x)$, para todo $x \in I$;

(c) $f_p(x + 2nl) = f_p(x)$, para todo $x \in -I \cup I$ e $n \in \mathbb{Z}$.

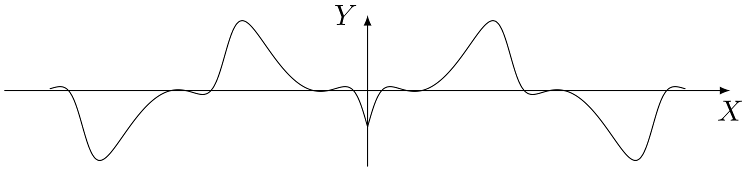


Figura 1.38: Exemplo de uma extensão periódica par.

E, contanto que l não pertença a I, definimos a *extensão periódica ímpar* de f ao conjunto $\mathbb{R}^{44}$, f_i, por meio das condições:

(a) $f_i(x) = f(x)$, para todo $x \in I$;

(b) $f_i(-x) = -f(x)$, para todo $x \in I$;

(c) $f_i(x + 2nl) = f_i(x)$, para todo $x \in -I \cup I$ e $n \in \mathbb{Z}$.

$\diamond$

[44] Exceto talvez aos pontos $(2n+1)l$, $n \in \mathbb{Z}$, conforme l pertença ou não a I; idem para os pontos $2nl$, $n \in \mathbb{Z}$, conforme 0 pertença ou não a I

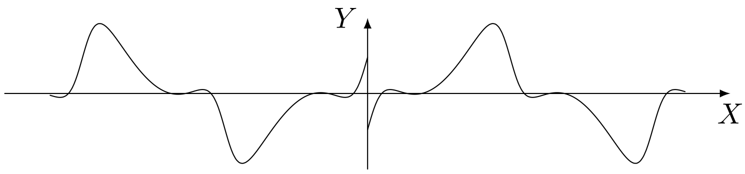


Figura 1.39: Exemplo de uma extensão periódica ímpar.

1.11 Famílias

Em certas ocasiões, a imagem da função é tida como sendo mais importante do que a própria função. Nesses casos, tanto a terminologia quanto a notação sofrem alteração.

Definição 1.118. *Seja $x : I \to X$ uma função. Os elementos de I são chamados de* índices*, I é chamado de o* conjunto de índices*, o conjunto imagem de x é chamado um* conjunto indexado*, a função é chamada uma* família *e o valor da função num índice i é chamado um* termo *da família.*

Chamamos a atenção para o fato de que "família" é a função (ou seja, domínio, contradomínio e lei de formação). Ocorre porém, como de costume, que a maioria dos textos (e este não é diferente nesse sentido) se refere a "uma família $\{x_i\}$ em X" ou a "uma família $\{x_i\}$ de elementos quaisquer de X" e, quando necessário, o conjunto de índices I é indicado por alguma expressão como: $i \in I$; essas e outras frases semelhantes são uma maneira de comunicar a notação e indicar a ênfase a que nos referimos no início desta seção. O verdadeiro significado delas, que você deve subentender, é: "existe uma função x de algum conjunto de índices I em X".

Da mesma forma, a frase "uma família $\{A_i\}$ de subconjuntos de X", por exemplo, deve ser entendida como se referindo a uma função A de algum conjunto de índices I (que não foi necessário mencionar) em $\mathscr{P}(X)$.

Notação 1.119. Utilizamos a notação x_i (ao invés da usual $x(i)$) para os termos da família e, seguindo a prática geralmente aceita, $\{x_i\}_{i \in I}$ para a famíla em si ou simplesmente $\{x_i\}$ quando não houver dúvida sobre o conjunto de índices ou não for necessário deixá-lo explícito. ○

A notação (x_i) em vez de $\{x_i\}$ também é utilizada por alguns autores para se referir a uma família, possivelmente porque, quando o conjunto de índices é um subconjunto dos naturais[45], a família coincide com uma n-upla ordenada ou com uma sequência como veremos mais adiante na Seção 1.11.2.

[45] Mais geralmente, bastaria que o conjunto de índices I fosse enumerável, ou seja, que existisse uma bijeção entre I e um subconjunto (próprio ou não) de $\mathbb{N}$.

Frente a esse conceito (Definição 1.118), note que é equivalente falar em coleções arbitrárias ou em famílias de conjuntos. Isso porque, dada uma coleção não vazia $\mathcal{C}$ de conjuntos, sempre é possível indexá-la: basta considerar a própria coleção $\mathcal{C}$ como o conjunto de índices e adotar a função identidade em $\mathcal{C}$ como família. O que acabamos de provar pode ser enunciado na seguinte proposição.

Proposição 1.120. *A indexação de uma coleção não vazia de conjuntos sempre existe.*

Se $\{A_i\}$ é uma família de subconjuntos de X, diz-se simplesmente que a reunião da imagem da família é a reunião da família $\{A_i\}$ ou a reunião dos conjuntos A_i e utiliza-se a seguinte notação:

$$\bigcup_{i \in I} A_i \quad \text{ou} \quad \bigcup_i A_i,$$

conforme precisemos ou não explicitar o conjunto de índices I. Claro que:

$$\bigcup_i A_i := \{x \in X;\, x \in A_i \,\text{para algum}\, i \in I\}. \tag{1.11.1}$$

Em particular, se $I = \{1, 2\}$, a imagem da família $\{A_i\}$ é o par não ordenado $\{A_1, A_2\}$ e $\bigcup_i A_i = A_1 \cup A_2$.

Uma reunião de uma coleção de conjuntos vazia faz sentido e é vazia, no entanto, a intersecção só faz sentido se a coleção não for vazia. Exceto por este fato trivial, a intersecção é análoga à reunião (terminologia e notação). Se, por exemplo, $\{A_i\}$ é uma família de conjuntos não vazios, a intersecção da imagem da família é chamada de a intersecção da família $\{A_i\}$ ou a intersecção dos conjuntos A_i e utiliza-se a seguinte notação:

$$\bigcap_{i \in I} A_i \quad \text{ou} \quad \bigcap_i A_i,$$

conforme seja importante ou não explicitar o conjunto de índices I. Claro que se $I \neq \emptyset$, nossa definição fica assim:

$$\bigcap_i A_i := \{x \in X;\, x \in A_i \,\text{para todo}\, i \in I\}. \tag{1.11.2}$$

Analogamente às reuniões, se $I = \{1, 2\}$, $\bigcap_i A_i = A_1 \cap A_2$.

As definições acima, equações (1.11.1) e (1.11.2), implicam as propriedades abaixo.

Proposição 1.121. *Sejam $\{I_j\}$ uma família de conjuntos com domínio J e $\{A_k\}$ uma família de conjuntos com domínio $K = \bigcup_j I_j$. Então:*

$$\bigcup_{k \in K} A_k = \bigcup_{j \in J} \Big(\bigcup_{i \in I_j} A_i \Big).$$

Demonstração. Suponha que x pertence ao lado esquerdo. Então existe $k \in K$ tal que $x \in A_k$. Como $K = \bigcup_j I_j$, $k \in K$ significa que $k \in I_j$ para algum $j \in J$. Assim, $x \in A_k$ para algum $k \in I_j$, ou seja, necessariamente $x \in \bigcup_{i \in I_j} A_i$, para algum $j \in J$, e, portanto, $x \in \bigcup_{j \in J}\big(\bigcup_{i \in I_j} A_i\big)$. Suponha agora que x pertence ao lado direito. Então existe algum $j \in J$ tal que $x \in \bigcup_{i \in I_j} A_i$. Isso significa que $x \in A_i$ para algum $i \in I_j$. Como $K = \bigcup_j I_j$, segue-se que $i \in K$, logo $x \in A_k$ para algum $k \in K$ e, portanto, $x \in \bigcup_{k \in K} A_k$. ∎

Com maior pretensão e menor clareza: a ordem em que os índices j e, consequentemente, os conjuntos I_j são escolhidos não tem importância. Essa é a versão generalizada da propriedade associativa da reunião de conjuntos, independentemente dos conjuntos de índices serem finitos ou não, enumeráveis ou não. A versão generalizada da propriedade comutativa pode ser enunciada como se segue.

Proposição 1.122. *Sejam $\{A_i\}$ uma família com domínio $I = \{i_j\}$ indexado por J. Então:*

$$\bigcup_{i \in I} A_i = \bigcup_{j \in J} A_{i_j}.$$

Demonstração. Se x pertence ao lado esquerdo, então $x \in A_i$ para algum $i \in I$. Como i é a imagem de algum j, $x \in A_{i_j}$ para algum $j \in J$ e, portanto, $x \in \bigcup_{j \in J} A_{i_j}$. Reciprocamente, se x pertence ao lado direito, então $x \in A_{i_j}$ para algum $j \in J$. Mas como $i_j \in I$, $x \in A_i$ para algum $i \in I$ e, portanto, $x \in \bigcup_{i \in I} A_{i_j}$. ∎

Como a indexação de I é arbitrária, essa última proposição nos diz que não importa a ordem em que a reunião é feita, o resultado é sempre o mesmo.

Para as intersecções valem propriedades análogas. Considerando-se famílias não vazias (entenda-se por família não vazia aquelas cujo domínio I não é vazio) e utilizando-se as mesmas notações das Proposições 1.121 e 1.122, obtém-se respectivamente:

$$\bigcap_{k \in K} A_k = \bigcap_{j \in J}\left(\bigcap_{i \in I_j} A_i\right) \quad \text{e} \quad \bigcap_{i \in I} A_i = \bigcap_{j \in J} A_{i_j}.$$

A premissa de que os conjuntos da família não são vazios, como já comentamos, serve para assegurar que as intersecções existam.

As demonstrações desses fatos são fáceis e podem ser obtidas utilizando-se diretamente a definição de intersecção ou por meio da aplicação das leis de De Morgan[46] às equações análogas para reuniões (Proposições 1.121 e 1.122).

[46]Augustus De Morgan (1806–1871) foi o primeiro professor de Matemática da *University College London* e deu importantes contribuições à Matemática inglesa. Famoso por suas leis (teoremas) que podemos dizer, com certa liberdade, permitem converter os operadores lógicos "e" e "ou" entre si ou, em termos da linguagem de conjuntos, se A e B são subconjuntos de X, $(A \cup B)' = A' \cap B'$ e $(A \cap B)' = A' \cup B'$, em que A' e B' são os complementares de A e B em relação a X.

Os fatos mais interessantes sobre reuniões e intersecções de famílias são os enunciadas a seguir, envolvendo ambos os conceitos.

Proposição 1.123. *Sejam $B \subset X \neq \emptyset$ e $\{A_i\}$ uma família de subconjuntos de X indexada por I. Então:*

(1.a) $B \setminus \bigcup_i A_i = \bigcap_i (B \setminus A_i)$;

(1.b) $B \setminus \bigcap_i A_i = \bigcup_i (B \setminus A_i)$;

(2.a) $(\bigcup_i A_i) \setminus B = \bigcup_i (A_i \setminus B)$;

(2.b) $(\bigcap_i A_i) \setminus B = \bigcap_i (A_i \setminus B)$;

(3.a) $B \cap \bigcup_i A_i = \bigcup_i (B \cap A_i)$;

(3.b) $B \cup \bigcap_i A_i = \bigcap_i (B \cup A_i)$;

(4.a) $B \cap \bigcap_i A_i = \bigcap_i (B \cap A_i)$;

(4.b) $B \cup \bigcup_i A_i = \bigcup_i (B \cup A_i)$;

(5.a) $(\bigcup_i A_i)' = \bigcap_i A_i'$;

(5.b) $(\bigcap_i A_i)' = \bigcup_i A_i'$.

Demonstração.

(1.a) $x \in B \setminus \bigcup_i A_i \Leftrightarrow x \in B$ e $x \notin \bigcup_i A_i \Leftrightarrow x \in B$ e $x \notin A_i$ para todo $i \in I \Leftrightarrow x \in B \setminus A_i$ para todo $i \in I \Leftrightarrow x \in \bigcap_i (B \setminus A_i)$.

(1.b) $x \in B \setminus \bigcap_i A_i \Leftrightarrow x \in B$ e $x \notin \bigcap_i A_i \Leftrightarrow x \in B$ e $x \notin A_i$ para algum $i \in I \Leftrightarrow x \in B \setminus A_i$ para algum $i \in I \Leftrightarrow x \in \bigcup_i (B \setminus A_i)$.

(2.a) $x \in (\bigcup_i A_i) \setminus B \Leftrightarrow x \in \bigcup_i A_i$ e $x \notin B \Leftrightarrow x \in A_i$ para algum $i \in I$ e $x \notin B \Leftrightarrow x \in A_i \setminus B$ para algum $i \in I \Leftrightarrow x \in \bigcup_i (A_i \setminus B)$.

(2.b) $x \in (\bigcap_i A_i) \setminus B \Leftrightarrow x \in \bigcap_i A_i$ e $x \notin B \Leftrightarrow x \in A_i$ para todo $i \in I$ e $x \notin B \Leftrightarrow x \in A_i \setminus B$ para todo $i \in I \Leftrightarrow x \in \bigcap_i (A_i \setminus B)$.

(3.a) $x \in B \cap \bigcup_i A_i \Leftrightarrow x \in B$ e $x \in \bigcup_i A_i \Leftrightarrow x \in B$ e $x \in A_i$ para algum $i \in I \Leftrightarrow x \in B \cap A_i$ para algum $i \in I \Leftrightarrow x \in \bigcup_i (B \cap A_i)$.

(3.b) $x \in B \cup \bigcap_i A_i \Leftrightarrow x \in B$ ou $x \in \bigcap_i A_i \Leftrightarrow x \in B$ ou $x \in A_i$ para todo $i \in I \Leftrightarrow x \in B \cup A_i$ para todo $i \in I \Leftrightarrow x \in \bigcap_i (B \cup A_i)$.

(4.a) $x \in B \cap \bigcap_i A_i \Leftrightarrow x \in B$ e $x \in \bigcap_i A_i \Leftrightarrow x \in B$ e $x \in A_i$ para todo $i \in I \Leftrightarrow x \in B \cap A_i$ para todo $i \in I \Leftrightarrow x \in \bigcap_i (B \cap A_i)$.

(4.b) $x \in B \cup \bigcup_i A_i \Leftrightarrow x \in B$ ou $x \in \bigcup_i A_i \Leftrightarrow x \in B$ ou $x \in A_i$ para algum $i \in I \Leftrightarrow x \in B \cup A_i$ para algum $i \in I \Leftrightarrow x \in \bigcup_i (B \cup A_i)$.

(5.a) Segue imediatamente de (1.a): basta fazer $B = X$.

(5.b) Segue imediatamente de (1.b): basta fazer $B = X$. ∎

Adotando o símbolo $\bigcup_{i,\,j}$ como uma abreviação para $\bigcup_{(i,\,j)\in I\times J}$, podemos enunciar o seguinte:

Proposição 1.124. *Se $\{A_i\}$ e $\{B_j\}$ são famílias não vazias de conjuntos, então:*

$$\left(\bigcup_i A_i\right) \cap \left(\bigcup_j B_j\right) = \bigcup_{i,\,j} (A_i \cap B_j) \quad e \quad \left(\bigcap_i A_i\right) \cup \left(\bigcap_j B_j\right) = \bigcap_{i,\,j} (A_i \cup B_j).$$

Demonstração.
$x \in (\bigcup_i A_i) \cap (\bigcup_j B_j) \Leftrightarrow x \in (\bigcup_i A_i)$ e $x \in (\bigcup_j B_j) \Leftrightarrow x \in A_i$ para algum $i \in I$ e $x \in B_j$ para algum $j \in J \Leftrightarrow x \in A_i \cap B_j$ para algum $(i,\,j) \in I \times J \Leftrightarrow \bigcup_{i,\,j} (A_i \cap B_j)$.

$x \in (\bigcap_i A_i) \cup (\bigcap_j B_j) \Leftrightarrow x \in (\bigcap_i A_i)$ ou $x \in (\bigcap_j B_j) \Leftrightarrow x \in A_i$ para todo $i \in I$ ou $x \in B_j$ para todo $j \in J \Leftrightarrow x \in A_i \cup B_j$ para todo $(i,\,j) \in I \times J \Leftrightarrow \bigcap_{i,\,j} (A_i \cup B_j)$. ■

1.11.1 Partições de conjuntos

Uma noção que você encontrará repetidas vezes ao longo de sua jornada é a de *partição de um conjunto.* Em poucas palavras, uma partição de um conjunto X é uma coleção disjunta de subconjuntos não vazios de X cuja reunião é ele próprio. Como vimos na Proposição 1.120, uma indexação de uma coleção não vazia de subconjuntos de X sempre existe. Assim, podemos expressar a noção de partição em termos do conceito de família.

Definição 1.125. *Sejam X um conjunto e $\mathcal{P} = \{P_i\}$ uma família de subconjuntos de X indexadas por um conjunto de índices I. Dizemos que $\mathcal{P}$ é uma* partição *de X quando:*

(a) $P_i \neq \emptyset$, *para todo* $i \in I$;

(b) $P_i \cap P_{i'} = \emptyset$ *sempre que* $i \neq i'$ *e*

(c) $\bigcup_{i\in I} P_i = X$.

Evidentemente, dizer que $\mathcal{P}$ é uma partição de X equivale a dizer que cada $x \in X$ pertence a um e somente um P_i.

Se $\mathcal{P}$ é uma partição de X, diz-se um tanto pictoricamente que $\mathcal{P}$ *particiona* X e cada elemento P_i é uma *componente* ou um *bloco* ou, ainda, uma *célula* da partição $\mathcal{P}$ de X. Quando $\mathcal{P}$ é uma partição com exatamente dois elementos, a chamamos uma *bipartição.*

Nunca chame um bloco de uma partição de "partição". Isso seria como confundir um membro de uma casal com o próprio casal. Por exemplo, se João e Maria formam um casal, nos referimos ao par (não ordenado) como "o casal João e Maria" e a um elemento desse par como "o cônjuge masculino" ou "o cônjuge João" e não "o casal masculino" ou "o casal João". Da mesma forma que isso não faz sentido, também não faz sentido dizer que $\{\triangle, \bigcirc\}$ é uma das partições de $X = \{\triangle, \bigcirc, \square\}$ caso $\{\{\triangle, \bigcirc\}, \{\square\}\}$ seja uma partição de X.

Exemplo 1.126. Os fatos mais simples sobre partições são os seguintes:

(a) todo conjunto unitário $\{x\}$ tem exatamente uma partição, a saber, $\{\{x\}\}$;

(b) o conjunto vazio possui uma única partição: o conjunto vazio;

(c) para todo conjunto não vazio X, $\mathcal{P} = \{X\}$ é uma partição de X, chamada partição trivial;

(d) se $A \neq \emptyset$ é um subconjunto próprio de X, A juntamente com seu complemento $A' = X \setminus A$, i.e., o conjunto $\mathcal{P} = \{A, A'\}$, forma uma partição de X. $\diamond$

Exemplo 1.127. Sejam $X = \{\triangle, \bigcirc, \square, \diamond, \pounds, \partial\}$ e $\mathcal{P}_1 = \{\{\triangle, \bigcirc, \square\}, \{\diamond\}, \{\pounds, \partial\}\}$. O conjunto $\mathcal{P}_1$ é uma partição de X, pois cumpre todas as condições da Definição 1.125 (Figura 1.40). Já os conjuntos $\mathcal{P}_2 = \{\emptyset, \{\triangle, \bigcirc\}, \{\square, \diamond, \pounds, \partial\}\}$, $\mathcal{P}_3 =$

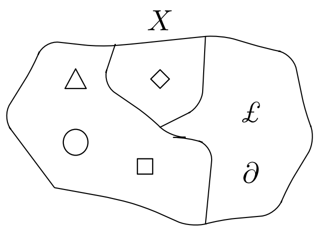


Figura 1.40: Exemplo de partição de um conjunto.

$\{\{\triangle, \bigcirc, \square, \diamond\}, \{\diamond, \pounds, \partial\}\}$ e $\mathcal{P}_4 = \{\{\triangle, \bigcirc\}, \{\diamond, \pounds, \partial\}\}$ não são partições de X porque não cumprem, respectivamente, as condições (a), (b) e (c) da Definição 1.125. $\diamond$

Definição 1.128. *Dadas as partições $\mathcal{P}$ e $\mathcal{Q}$ de um conjunto X, dizemos que $\mathcal{P}$ é um* refinamento *de $\mathcal{Q}$ se todo elemento de $\mathcal{P}$ é subconjunto de algum elemento de $\mathcal{Q}$.*

Exemplo 1.129. A partição $\{\{1, 3\}, \{2, 4\}, \{5\}, \{6\}\}$ de $\{1, 2, 3, 4, 5, 6\}$ é um refinamento da partição $\{\{1, 3, 2, 4\}, \{5, 6\}\}$. $\diamond$

Exemplo 1.130. Seja $I = [a, b]$ um intervalo não degenerado. Considere uma família $\mathcal{P} = \{[x_j, x_{j+1}]\}$ de subconjuntos não degenerados de I indexada por $J =$

$\{1, 2, \ldots, m-1\}$, com $x_1 = a$ e $x_m = b$. Certamente $\mathcal{P}$ é uma partição de I como você pode constatar. Considere agora outra família $\mathcal{Q} = \{[u_k, u_{k+1}]\}$ de subconjuntos não degenerados de I indexada por $K = \{1, 2, \ldots, n-1\}$, com $u_1 = a$ e $u_n = b$, que também é uma partição de I. Suponha que $\{x_j\}$ é subconjunto próprio de $\{u_k\}$. Então, para todo $k \in K$ existe algum $j \in J$ tal que $[u_k, u_{k+1}] \subset [x_j, x_{j+1}]$. (Prova?) Portanto, $\mathcal{Q}$ é um refinamento de $\mathcal{P}$. ◇

Esse último exemplo tem aplicação no cálculo da soma de Riemann[47] na definição da integral definida de uma função contínua num intervalo $[a, b]$. Você estudará isso num curso de Cálculo e de Análise. Do ponto de vista numérico, os refinamentos de um intervalo podem ser vistos como um processo para a obtenção de aproximações mais acuradas da integral definida quando se utiliza uma das técnicas mais diretas para tanto, que se enquadra numa família de técnicas de integração numérica conhecida como fórmulas de Newton-Cotes[48].

Proposição 1.131. *Se $\{P_i\}$ e $\{Q_j\}$ são partições de um conjunto X, então $\mathcal{C} = \{P_i \cap Q_j\} \setminus \emptyset$ também é uma partição de X.*

Demonstração. Por definição, $\mathcal{C}$ não contém o conjunto vazio. Usando a associatividade e a comutatividade da intersecção, obtemos $(P_i \cap Q_j) \cap (P_{i'} \cap Q_{j'}) = (P_i \cap P_{i'}) \cap (Q_j \cap Q_{j'})$; logo, se $(i, j) \neq (i', j')$, então ou $(P_i \cap P_{i'}) = \emptyset$ ou $(Q_j \cap Q_{j'}) = \emptyset$ (ou ambos), de modo que $(P_i \cap P_{i'}) \cap (Q_j \cap Q_{j'}) = \emptyset$ sempre que $(i, j) \neq (i', j')$. Por fim, utilizando a Proposição 1.124, obtemos $\bigcup_{i,j}(P_i \cap Q_j) = (\bigcup_i P_i) \cap (\bigcup_j Q_j)$; como $\{P_i\}$ e $\{Q_j\}$ são partições de X, $\bigcup_i P_i = \bigcup_j Q_j = X$, logo $(\bigcup_i P_i) \cap (\bigcup_j Q_j) = X$ e, portanto, $\bigcup_{i,j}(P_i \cap Q_j) = X$. ■

Definição 1.132. *A nova partição, obtida na Proposição 1.131, é chamada* partição cruzada *das duas partições originais.*

Partições cruzadas são comumente utilizadas em problemas de classificação.

Exemplo 1.133. A partir do conjunto X de todas as formas de vida, podemos formar a partição $\{P, A\}$ em que P é o conjunto de todas as plantas e A o conjunto de todos os animais. Também podemos formar a partição $\{E, F\}$ em que E é o conjunto de todas

[47]Georg Friedrich Bernhard Riemann (1826–1866). Suas ideias em matéria de geometria do espaço tiveram um efeito profundo sobre o desenvolvimento da física teórica moderna. Ele esclareceu a noção de integral ao definir o que hoje chamamos de integral de Riemann.

[48]Em honra a Isaac Newton (1642–1727) e a Roger Cotes (1682–1716). Cotes foi um matemático inglês, membro da *Royal Society*; trabalhou com Newton, revisando a segunda edição do *Principia* e fez avanços na teoria dos logarítmos, no cálculo integral e em métodos numéricos, especialmente em interpolação.

as formas de vida extintas e F é o conjunto de todas as formas de vida ainda existentes. A partição cruzada $\{P \cap E,\ P \cap F,\ A \cap E,\ A \cap F\}$ nos fornece uma classificação completa de acordo com as duas classificações independentes iniciais. $\diamond$

Um fato importante que você deve saber sobre partições é sua relação com as classes de equivalência. Toda partição induz uma relação de equivalência e, reciprocamente, toda relação de equivalência induz uma partição. Esses são os conteúdos das duas próximas proposições.

Proposição 1.134. *Seja $\mathcal{P} = \{P_i\}_{i \in I}$ uma partição de um conjunto X. Defina uma relação R em X como:*

$$R := \{(x,\, y) \in X \times X;\ x,\, y \in P_i \, para\ algum \, i \in I\}.$$

Então, R é uma relação de equivalência em X (determinada por $\mathcal{P}$).

Demonstração. Se $x \in X$, então x está no mesmo bloco que ele mesmo, logo $x\,R\,x$ e, portanto, R é reflexiva. Se $x\,R\,y$, então x e y pertencem ao mesmo bloco, logo $y\,R\,x$ e, portanto, R é simétrica. Se $x\,R\,y$ e $y\,R\,z$, então x, y e z estão no mesmo bloco , logo $x\,R\,z$ e, portanto, R é transitiva. $\blacksquare$

Exemplo 1.135. Sejam $X = \{\diamond,\, \pounds,\, \partial\}$ e a partição $\mathcal{P} = \{\{\diamond,\, \pounds\},\, \{\partial\}\}$. Encontrar a relação de equivalência determinada por $\mathcal{P}$. $\diamond$

Solução. Os blocos de $\mathcal{P}$ são $\{\diamond,\, \pounds\}$ e $\{\partial\}$. Cada elemento de um bloco deve estar relacionado com todos os outros elementos do mesmo bloco e com nenhum outro elemento além destes. Portanto, basta definir R assim:

$$R := \{(\diamond,\, \diamond),\, (\diamond,\, \pounds),\, (\pounds,\, \diamond),\, (\pounds,\, \pounds),\, (\partial,\, \partial)\}. \qquad \blacklozenge$$

Proposição 1.136. *Sejam E uma relação de equivalência em $X \neq \emptyset$ e $\mathcal{P} = \{[x]\}$ a coleção de todas as classes de equivalência dos elementos x de X pela relação de equivalência E. Então P é uma partição de X.*

Demonstração. Primeiramente, como $X \neq \emptyset$, $[x] \neq \emptyset$ para todo $x \in X$. Em segundo lugar, de acordo com o Corolário 1.45, $[x] \cap [y] = \emptyset$ sempre que $[x] \neq [y]$. Por fim, como, por um lado, todo $x \in X$ pertence à sua classe de equivalência $[x]$, então $X \subset \bigcup_{x \in X}[x]$; por outro lado, se $y \in \bigcup_{x \in X}[x]$, então $y \in [x]$ para algum $x \in X$, ou seja, $x \sim y$ para algum $x \in X$ e, portanto $y \in X$, de modo que $\bigcup_{x \in X}[x] \subset X$; assim $\bigcup_{x \in X}[x] = X$. Portanto, como as condições da Definição 1.125 estão satisfeitas, segue-se que $\mathcal{P} = \{[x]\}$ é uma partição de X. $\blacksquare$

Exemplo 1.137. Sejam $X = \{\diamond, \pounds, \partial\}$ e a relação de equivalência em X:

$$R := \{(\diamond, \diamond), (\diamond, \pounds), (\pounds, \diamond), (\pounds, \pounds), (\partial, \partial)\}.$$

Encontrar todas as classes de equivalência de R e constatar que a coleção de todas elas forma uma partição de X. ◇

Solução. Temos $[\diamond] = \{\diamond, \pounds\}$; $[\pounds] = \{\diamond, \pounds\}$; e $[\partial] = \{\partial\}$. A coleção de todas essas classes é o conjunto $\mathcal{P} = \{\{\diamond, \pounds\}, \{\partial\}\}$, que é, de fato, uma partição de X. ◆

1.11.2 Produtos cartesianos gerais

A notação de famílias também pode ser utilizada (e geralmente o é) com vistas a uma generalização do conceito de produto cartesiano. Nossa motivação inicial é a existência de uma bijeção entre o produto cartesiano de dois conjuntos e um certo conjunto de famílias. Em virtude disso, definimos o produto cartesiano para uma família arbitrária de conjuntos.

Vamos à existência da bijeção, cuja demostração é imediata.

Proposição 1.138. *Sejam os conjuntos X, Y e $\{a, b\}$, com $a \neq b$. Seja Z o conjunto de todas as famílias z indexadas por $\{a, b\}$ tal que $z_a \in X$ e $z_b \in Y$. Então existe uma bijeção entre $X \times Y$ e Z.*

Demonstração. Basta definir uma função $f : Z \to X \times Y$ por $f(z) = (z_a, z_b)$. ■

A ideia intuitiva de par ordenado que discutimos na Seção 1.4.2 continua válida, pois a presente ideia consiste em dizer que cada par ordenado (x, y) com $x \in X$ e $y \in Y$ é uma função em que o primeiro membro do par é a imagem de a (por ser o primeiro) e o segundo é a imagem de b (por ser o segundo). Se preferir, troque a por 1 e b por 2 e a ideia intuitiva de "primeiro" e "segundo" por detrás do conceito ficará mais clara.

Como existe a bijeção acima, note que a diferença entre Z e $X \times Y$ é apenas uma matéria de notação. Seguindo essa linha de raciocínio, a generalização é imediata.

Definição 1.139. *Seja $\{X_i\}_{i \in I}$ uma família de conjuntos. Chamamos de* produto cartesiano *dessa família o conjunto de todas as famílias $\{x_i\}$ com $x_i \in X_i$ para cada $i \in I$.*

Notação 1.140. Há vários símbolos para o produto cartesiano, sendo os mais comuns:

$$\bigtimes_{i \in I} X_i \text{ ou } \bigtimes_{i} X_i \quad \text{e} \quad \prod_{i \in I} X_i \text{ ou } \prod_{i} X_i.$$

○

Neste texto, utilizaremos $\bigtimes_{i\in I} X_i$ ou $\bigtimes_i X_i$. Assim, em símbolos, a Definição 1.139 fica:

$$\bigtimes_{i\in I} X_i := \left\{ x : I \to \bigcup_{i\in I} X_i;\ x_i \in X_i \text{ para todo } i \in I \right\}.$$

Se todos os conjuntos X_i forem iguais a X, então denota-se $\bigtimes_i X_i$ por X^I, que é o conjunto de todas as funções de I em X. Se I é o conjunto unitário $\{a\}$, $\bigtimes_i X_i = X_a$; quando I é o par não ordenado $\{a,\, b\}$ com $a \neq b$, $\bigtimes_i X_i = X_a \times X_b$, ou seja, obtemos o produto cartesiano de dois conjuntos como de praxe; analogamente, definimos *triplas ordenadas*, *quádruplas ordenadas*, ..., *n-uplas ordenadas*, ... a partir de triplas, quádruplas, ..., n-uplas, ... não ordenadas.

E como cada n-upla é uma família, isso pode justificar a notação (x_i) ao invés de $\{x_i\}$ adotada por alguns autores. Isso porque a função $x : I \to \bigcup_{i\in I} X_i$ tem, por assim dizer, o papel de ordenar os elementos de $\bigcup_{i\in I} X_i$ tomando-se sucessivamente um elemento de cada X_i por vez, o que é totalmente coerente com a ideia intuitiva por detrás do conceito de n-upla ordenada.

Exemplo 1.141. $\mathbb{R}\times\mathbb{R}$ e $\mathbb{R}^{\{a,\, b\}}$ $(a \neq b)$ ou $\mathbb{R}^{\{1,\, 2\}}$ são notações distintas para o mesmo objeto, também denotado por $\mathbb{R}^2$ como se sabe, que representa todos os pares ordenados de números reais ou todas as funções de $\{1,\, 2\}$ em $\mathbb{R}$. $\diamond$

Quando $I = \{1,\, 2,\, \ldots,\, n\}$, a notação para o produto cartesiano e para um de seus elementos x, ou seja, uma n-upla ordenada, fica assim:

$$\bigtimes_{i\in I} X_i = X_1 \times X_2 \times \cdots \times X_n \quad \text{e} \quad x = \bigtimes_{i\in I} (x_i) = (x_1,\, x_2,\, \ldots,\, x_n),$$

e o elemento x_i é chamado a *i-ésima coordenada da n-upla x*.

Essa discussão motiva uma definição muito útil, já apresentada no Exemplo 1.7, porém com outros termos:

Definição 1.142. *Uma família com índices no conjunto* $\mathbb{N} = \{1,\, 2,\, \ldots\}$ *dos números naturais*[a] *chama-se uma* sequência *ou* sucessão.

[a] De um modo mais geral, o conjunto de índices não precisa ser necessariamente $\mathbb{N}$, mas qualquer conjunto contável totalmente ordenado; ocorre, porém, que esse tipo de conjunto equivale a $\mathbb{N}$ no sentido de que existe uma bijeção entre ele e $\mathbb{N}$, por isso não é necessário mencionar outro conjunto que não o dos naturais.

Assim, uma sequência $x = (x_i)_{i\in\mathbb{N}} = (x_1,\, x_2,\, \ldots,\, x_n,\, \ldots)$ de elementos de um conjunto X nada mais é do que uma função $x : \mathbb{N} \to X$, cujo valor $x(n)$ é indicado por x_n e é chamado de *n-ésimo termo* da sequência.

Como você deve ter notado, o conjunto de índices I na definição de produtos cartesianos gerais é totalmente arbitrário, nem sequer necessariamente finito ou enumerável. No entanto, ainda que ele não seja vazio, tampouco o sejam os conjuntos X_i indexados por ele, surge uma questão um tanto sutil: $\bigtimes_i X_i$ pode ser vazio? Como garantir que sim ou não?

Toda a Matemática se assenta sobre uma série de postulados sobre a noção de conjuntos, os *axiomas*, afirmações aceitas como verdadeiras a partir das quais outras podem ser deduzidas. Há diversos de tais axiomas. Por simplicidade, evitamos listá-los e discuti-los neste texto. No entanto, para poder respoder a questão acima, somos obrigados a abrir uma exceção, pois a resposta está contida no chamado *Axioma da Escolha*, originalmente formulado por Zermelo[49] em 1904 como parte de sua demonstração do chamado *Princípio do Bom Ordenamento*. Trata-se do enunciado a seguir.

> *Seja $\{X_i\}$ uma família de conjuntos não vazios indexada por um conjunto arbitrário não vazio de índices. Então, é possivel construir um conjunto X tomando ("escolhendo") um elemento x_i de cada conjunto X_i.*

Em termos mais técnicos, esse axioma nos diz que:

> *existem funções $x : I \to \bigcup_i X_i$ tais que $x_i \in X_i$ para todo $i \in I$.*

Em outras palavras, como o produto cartesiano é o conjunto de todas as funções $x : I \to \bigcup_i X_i$, o Axioma da Escolha afirma o seguinte:

> *O produto cartesiano de uma família não vazia de conjuntos não vazios não é vazia.*

Essa é a resposta e a garantia para as duas questões acima.

Um último fato interessante sobre famílias está contido na próxima proposição.

Proposição 1.143. *Seja $\{X_i\}_{i \in I}$ uma família de conjuntos, X seu produto cartesiano e $J \subset I$. Então existe uma correspondência entre X e o produto cartesiano parcial $\bigtimes_{i \in J} X_i$.*

Demonstração. Cada $x \in X$ é uma família $\{x_i\}$, que em última análise é uma função $x : I \to X$; o elemento correspondente $y \in \bigtimes_{i \in J} X_i$ é obtido simplesmente restringindo-se x a J, ou seja, basta definirmos $y := x|J$. ∎

[49] Ernst Friedrich Ferdinand Zermelo (1871–1953), matemático e filósofo alemão, cujo trabalho teve influência direta nos fundamentos da Matemática. Famoso por seu papel no desenvolvimento dos axiomas de Zermelo-Fraenkel e na prova do Teorema do Bom Ordenamento.

A função $x \to y$ definida na demonstração acima é chamada a *projeção* de X em $\bigtimes_{i \in J} X_i$ e a denotaremos por π_J. Essa função, como você pode facilmente constatar, é sobrejetiva. (Certifique-se!) Se, em particular, J for um conjunto unitário $J = \{j\}$, escreveremos π_j em vez de $\pi_{\{j\}}$. Há outros usos da palavra "projeção":

1. se $x \in X$, então o valor de π_j em x, i.e., x_j, também é chamado de a projeção de x em X_j ou, alternativamente, a *j-coordenadada* de x (se, em particular, $I \subset \mathbb{N}$, também costuma-se dizer *j-ésima coordenada* de x);

2. se $f : A \to \bigtimes_i X_i$ é uma função de um conjunto qualquer A no produto cartesiano dos conjuntos X_i, podemos definir, para cada $i \in I$, uma função $f_i : A \to X_i$ por $f_i := \pi_i \circ f$, assim, $f(a) = \bigtimes_i (f_i(a))_i$ para cada $a \in A$; reciprocamente, se for dada uma função $f_i : A \to X_i$ para cada $i \in I$, então existe uma única função $f : A \to \bigtimes_i X_i$ tal que $f_i = \pi_i \circ f$ para cada $i \in I$: basta definir $f(a) := \bigtimes_i (f_i(a))_i$ para cada $a \in A$; a função f_i é chamada de a projeção de f em X_i ou a *j-coordenada* de f (ou *j-ésima coordenada* de f se $I \subset \mathbb{N}$).

Uma função cujo domínio é um produto cartesiano é chamada de uma função de *várias variáveis*; em particular, uma função cujo domínio é o produto cartesiano $X \times Y$ de dois conjuntos é chamada de uma função de duas variáveis.

Exercício 1.144 ([4]). Prove que $(\bigcup_i X_i) \times (\bigcup_j Y_j) = \bigcup_{i,j} (X_i \times Y_j)$ e que $(\bigcap_i X_i) \times (\bigcap_j Y_j) = \bigcap_{i,j} (X_i \times Y_j)$ (contanto que, neste último caso, os conjuntos de índices não sejam vazios). Adicionalmente, prove que $\bigcap_i X_i \subset X_j \subset \bigcup_i X_i$ para todo j, contanto que o conjunto de índices da família $\{X_i\}$ não seja vazio. Além disso, prove que a intersecção e a reunião podem, de fato, ser caracterizadas como as soluções extremas dessas inclusões, i.e., se $X_j \subset Y$ para todo j, então $\bigcup_i X_i \subset Y$ e essa reunião é o único conjunto que satisfaz essa condição minimal; analogamente para a intersecção. $\diamond$

1.12 Funções e conjuntos: propriedades elementares

Agora que você foi apresentado ao conceito de família na seção precedente, vamos lhe apresentar mais alguns fatos importantes envolvendo conjuntos e funções nas duas seguintes proposições.

Proposição 1.145. *Sejam $f : X \to Y$ uma função e I um conjunto de índices.*

(a) Se $A_i \subset X$ para todo $i \in I$, então:

$$f\left(\bigcup_i A_i\right) = \bigcup_i f(A_i), \tag{1.12.1}$$

mas

$$f\left(\bigcap_i A_i\right) \subset \bigcap_i f(A_i). \tag{1.12.2}$$

(b) Se $B_i \subset Y$ para todo $i \in I$, então:

$$f^{-1}\left(\bigcup_i B_i\right) = \bigcup_i f^{-1}(B_i) \tag{1.12.3}$$

e

$$f^{-1}\left(\bigcap_i B_i\right) = \bigcap_i f^{-1}(B_i). \tag{1.12.4}$$

Demonstração.

(1.12.1) $y \in f(\bigcup_i A_i) \Leftrightarrow$ existe $x \in \bigcup_i A_i$ tal que $y = f(x) \Leftrightarrow$ para algum $i \in I$, existe $x \in A_i$ tal que $y = f(x) \Leftrightarrow$ para algum $i \in I$, $y \in f(A_i) \Leftrightarrow y \in \bigcup_i f(A_i)$.

(1.12.2) $y \in f(\bigcap_i A_i) \Rightarrow$ existe $x \in \bigcap_i A_i$ tal que $y = f(x) \Rightarrow$ para todo $i \in I$, existe $x \in A_i$ tal que $y = f(x) \Rightarrow$ para todo $i \in I$, $y \in f(A_i) \Rightarrow y \in \bigcap_i f(A_i)$.

(1.12.3) $x \in f^{-1}(\bigcup_i B_i) \Leftrightarrow$ existe $y \in \bigcup_i B_i$ tal que $y = f(x) \Leftrightarrow$ para algum $i \in I$, existe $y \in B_i$ tal que $y = f(x) \Leftrightarrow$ para algum $i \in I$, $x \in f^{-1}(B_i) \Leftrightarrow x \in \bigcup_i f^{-1}(B_i)$.

(1.12.4) $x \in f^{-1}(\bigcap_i B_i) \Leftrightarrow$ existe $y \in \bigcap_i B_i$ tal que $y = f(x) \Leftrightarrow$ para todo $i \in I$, existe $y \in B_i$ tal que $y = f(x) \Leftrightarrow$ para todo $i \in I$, $x \in f^{-1}(B_i) \Leftrightarrow x \in \bigcap_i f^{-1}(B_i)$. ∎

Não vale a igualdade em (1.12.2) porque, se y pertence ao lado direito, então y pertence a $f(A_i)$ para todo $i \in I$, portanto existe um x_i em cada A_i tal que $y = f(x_i)$. Ocorre, porém, que pode não haver x algum em $\bigcap_i A_i$ tal que $y = f(x)$. Um contraexemplo deixará tudo mais claro. Considere $f : [-1, 1] \to \mathbb{R}$, definida por $f(x) = x^2$, e sejam $A_1 = [-1, 0]$ e $A_2 = [0, 1]$. Então $f(A_1) = f(A_2) = [0, 1]$, logo $f(A_1) \cap f(A_2) = [0, 1]$, mas $f(A_1 \cap A_2) = f(\{0\}) = \{0\}$.

Porém, se adicionarmos a hipótese de que a f é injetiva, então é possível obtermos a igualdade, como nos mostra a próxima proposição.

Proposição 1.146. *Seja $f : X \to Y$ injetiva. Então, se $A_i \subset X$ para todo $i \in I$, vale:*

$$f\left(\bigcap_i A_i\right) = \bigcap_i f(A_i).$$

Demonstração. Se y pertence ao lado direito, então y pertence a $f(A_i)$ para todo $i \in I$, portanto existe um x_i em cada A_i tal que $y = f(x_i)$. Como f é injetiva, todos os x_i são iguais a, digamos, x. Logo, $x \in \bigcap_i A_i$. Como $y = f(x)$, segue-se que y pertence ao lado esquerdo. A outra inclusão já foi provada na proposição anterior. ∎

1.13 Funções implícitas: uma noção

O intuito desta seção é apresentar apenas uma noção sobre as funções implícitas. O aprofundamento deste assunto está fora do nosso escopo e requer ferrramentas mais avançadas, mas haverá ocasião para você estudá-las adiante em outros cursos, como um curso mais avançado de Cálculo ou de Análise. Nesses cursos, você será apresentado ao teorema da função inversa, do qual o teorema da função implícita é uma consequência, e verá as aplicações deste último, como, por exemplo, assegurar a existência das soluções de certas equações diferenciais. Por ora, gostaríamos apenas de lhe apresentar o conceito de uma função dada implicitamente.

Até aqui, vimos ou falamos sobre funções definidas *explicitamente*, i.e., funções para as quais a lei de correspondência, consistindo de um conjunto finito de instruções (um algoritmo), não necessariamente, mas geralmente, uma fórmula, permitia obter diretamente a variável independente y dada a variável dependente x ou vice-versa se a função fosse invertível, e costumávamos denotar assim:

$$y = f(x) \quad \text{ou} \quad x = \phi(y).$$

Por outro lado, podemos olhar para a função inversa de $y = f(x)$ como a solução da equação $y - f(x) = 0$ para x. Ou, mais geralmente, podemos falar em equações do tipo $F(x, y) = 0$ e tentar resolvê-la para x ou y. Isso é frequentemente feito na Geometria Analítica. Por exemplo, ao invés de se representar curvas por meio das equações $y = f(x)$ ou $x = \phi(y)$, o que se faz é representá-las por meio de qualquer equação em x e y da forma $F(x, y) = 0$. Alguns exemplos clássicos são os das equações:

(a) do círculo de centro na origem e raio 1: $x^2 + y^2 - 1 = 0$;

(b) da elipse de centro na origem e semieixos a e b: $x^2/a^2 + y^2/b^2 - 1 = 0$;

(c) da lemniscata: $(x^2 + y^2)^2 - 2a^2(x^2 - y^2) = 0$.

Em todos esses casos, a fim de obtermos y como função de x, ou x como função de y, precisamos resolver a equação para y ou para x. Dizemos, nesses casos, que a função $y = f(x)$ ou $x = \phi(y)$ é definida *implicitamente* por meio da equação $F(x, y) = 0$ e que a solução dessa equação nos fornece a função *explicitamente*. Claro que essa ideia pode ser estendida para mais variáveis e poderíamos falar de funções $G(x, y, z)$, $H(x, y, z, u, v)$ e assim por diante, mas é suficiente restringir essa breve discussão a apenas x e y.

Em alguns casos, como os de acima, as equações podem ser resolvidas e podemos exibir a solução explicitamente em termos de funções elementares. Em outros casos, a solução só pode ser obtida lançando-se mão de ferramentas mais sofisticadas que nos permitam obter a solução em termos de uma série infinita ou por meio de outros processos envolvendo limites, i.e., processos que nos permitam obter soluções tão próximas quando desejarmos de $y = f(x)$ ou $x = \phi(y)$. Para muitos propósitos, entretanto, não convém resolver a equação $F(x, y) = 0$, seja a sua solução exata ou aproximada, mas apenas basear a discussão na definição implícita.

Por fim, é importante ressaltar que nem toda função $F(x, y)$ nos fornece uma função $y = f(x)$ ou $x = \phi(y)$ por meio da resolução da equação $F(x, y) = 0$. De fato, é fácil encontrar exemplos de funções $F(x, y)$ que, quando igualadas a zero, não nos fornece funções de uma variável como solução. Por exemplo, $x^2 + y^2 = 0$ só é possível quando ambos x e y são iguais a zero e $x^2 + y^2 + 1 = 0$ não possui solução real. Por esse motivo, antes de tudo, se faz necessário estudar se a equação $F(x, y) = 0$ pode ser resolvida e quais propriedades sua solução possui, mas isso não pode ser feito aqui porque, como dissemos no início, demanda ferramentas mais sofisticadas. Esse estudo será deixado para cursos posteriores.

1.14 Mais alguns conceitos

Nesta seção, vamos falar predominantemente das funções reais de uma variável real, i.e., aquelas cujo domínio e contradomínio são subconjuntos de $\mathbb{R}$.

Algumas dessas funções possuem propriedades especiais ao mesmo tempo simples e úteis. Os primeiros conceitos que estudaremos a seguir são particularmente úteis, por exemplo, no esboço do gráfico de certas funções e no cálculo dos coeficientes das séries de Fourier, utilizadas na resolução de equações diferenciais parciais, como a equação da onda, do calor e problemas envolvendo a Equação de Laplace[50].

1.14.1 Funções pares, ímpares, simétricas e antissimétricas

Antes de prosseguir, vamos convencionar o seguinte: diremos que um conjunto $X \subset \mathbb{R}$ é simétrico em relação à origem (independentemente de 0 pertencer ou não a X) se ele possuir a propriedade: $x \in X \Rightarrow -x \in X$.

Definição 1.147. *Sejam $X, Y \subset \mathbb{R}$, X simétrico em relação à origem e uma função $f : X \to Y$. Dizemos que:*

(a) *f é* par *se $f(-x) = f(x)$ para todo $x \in X$.*

(b) *f é* ímpar *se $f(-x) = -f(x)$ para todo $x \in X$.*

A escolha de um domínio X e um contradomínio Y que sejam subconjuntos de $\mathbb{R}$ na Definição 1.147 não é feita por acaso. Essa escolha se deve ao fato de que precisamos

[50]Pierre-Simon Laplace (1749–1827), matemático, físico e astrônomo francês. Dentre seus grandes feitos, colocou a Teoria da Probabilidade numa base sólida, comprovou a estabilidade do sistema solar e, na Análise, introduziu a função potencial e os coeficientes que hoje levam seu nome.

de parte da estrutura algébrica de $\mathbb{R}$, que consiste na capacidade de tomarmos opostos $(-x)$ e ordenarmos elementos. Por esse motivo, X e Y não poderiam ser conjuntos arbitrários.

Os aspectos gerais dos gráficos de funções pares e ímpares são ilustrados na Figura 1.41.

(a) Exemplo de função par. (b) Exemplo de função ímpar.

Figura 1.41: Aspectos gerais dos gráficos de uma função par e outra ímpar.

Exemplo 1.148. Sejam $X = \{-3, -2, -1, 1, 2, 3\}$, $Y = \{1, 2, 3, 4, 5\}$ e $f : X \to Y$ definida por:

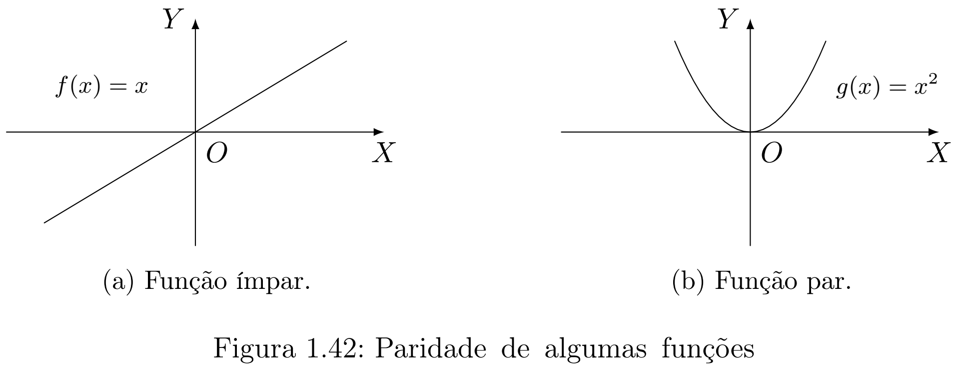
$$f = \{(-3, 2), (-2, 1), (-1, 5), (1, 5), (2, 1), (3, 2)\}.$$

Como você pode facilmente constatar, f é par, uma vez que $f(-x) = f(x)$ para todo $x \in X$. Por outro lado, a função $g : X \to Y$ definida por:

$$g = \{(-3, 2), (-2, 1), (-1, 3), (1, 5), (2, 1), (3, 2)\}.$$

não é par nem ímpar, uma vez que $f(-1)$ não é igual a $f(1)$ nem a $-f(1)$. ◇

Exemplo 1.149. Talvez os exemplos mais simples de uma função ímpar e outra par sejam $f, g : \mathbb{R} \to \mathbb{R}$ definidas por $f(x) = x$ e $g(x) = x^2$, respectivamente, Figuras 1.42a e 1.42b. ◇

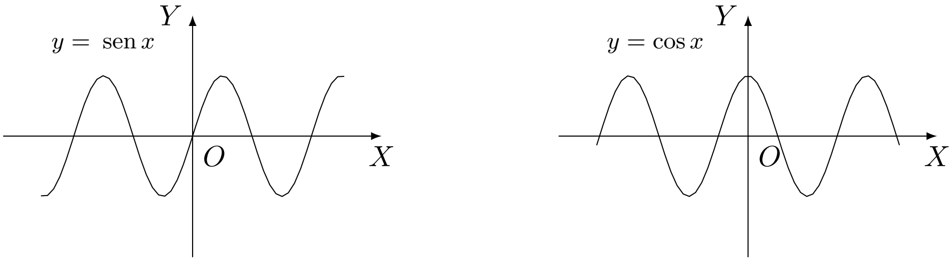


(a) Função ímpar. (b) Função par.

Figura 1.42: Paridade de algumas funções simples.

Esse simples exemplo contém uma dica para você associar os nomes às propriedades. Suponha $X, Y \subset \mathbb{R}$, X simétrico em relação à origem e $f : X \to Y$ uma função definida por $f(x) = x^n$ $(n \in \mathbb{N})$. Então, f será ímpar quando n for ímpar, e paro o caso contrário, f será par.

Exemplo 1.150. Outros exemplos clássicos são os das funções $\operatorname{sen}, \cos : \mathbb{R} \to \mathbb{R}$, respectivamente, Figuras 1.43a e 1.43b. $\diamond$



(a) Seno: função ímpar.

(b) Cosseno: função par.

Figura 1.43: Paridade das funções seno e cosseno.

Exercício 1.151. Verifique quais das funções a seguir são pares e quais são ímpares:

(a) $\sinh : \mathbb{R} \to \mathbb{R}$, $\sinh x = (e^x - e^{-x})/2$;

(b) $\cosh : \mathbb{R} \to \mathbb{R}$, $\cosh x = (e^x + e^{-x})/2$;

(c) $f : \mathbb{R} \to \mathbb{R}$, $f(x) = \sqrt{1 + x + x^2} - \sqrt{1 - x + x^2}$;

(d) $g : (-1,\, 1) \to \mathbb{R}$, $g(x) = \log\left[(1+x)/(1-x)\right]$;

(e) $h : \mathbb{R} \to \mathbb{R}$, $h(x) = \log\left(x + \sqrt{1 + x^2}\right)$. $\diamond$

Exercício 1.152. Seja $f : \mathbb{R} \to \mathbb{R}$ definida por $f(x) = 2x^4 - 3x^3 - 5x^2 + 6x - 10$. Calcule:

$$\varphi(x) = \frac{1}{2}[f(x) + f(-x)] \quad \text{e} \quad \psi(x) = \frac{1}{2}[f(x) - f(-x)]$$

e estude as paridades dessas novas funções. $\diamond$

Um fato notável sobre funções pares e ímpares de fácil demonstração é o seguinte.

Proposição 1.153. *Sejam X, $Y \subset \mathbb{R}$, sendo X simétrico em relação à origem. Qualquer função $f : X \to Y$ pode ser escrita como a soma de uma função par e uma função ímpar.*

Demonstração.
$$f(x) = f(x) + \frac{1}{2}f(-x) - \frac{1}{2}f(-x) = \frac{1}{2}[f(x) + f(-x)] + \frac{1}{2}[f(x) - f(-x)].$$
Definindo f_p, $f_i : X \to Y$ da seguinte maneira:
$$f_p(x) = \frac{1}{2}[f(x) + f(-x)] \quad \text{e} \quad f_i(x) = \frac{1}{2}[f(x) - f(-x)],$$
temos:
$$f_p(-x) = \frac{1}{2}[f(-x) + f(x)] = f_p(x) \quad \text{e} \quad f_i(-x) = \frac{1}{2}[f(-x) - f(x)] = -f_i(x),$$
ou seja, f_p é par e f_i é ímpar. Concluímos, então, que $f = f_p + f_i$. ∎

Além dessa, podemos mencionar mais as seguites propriedades, cuja demonstrção deixamos como exercício.

(a) Uma função ímpar definida na origem vale zero na origem.

(b) A única função par e ímpar ao mesmo tempo é a função nula.

(c) A soma de funções de mesma paridade mantém essa paridade.

(d) O produto de duas funções de mesma paridade é par e de paridades opostas é ímpar.

(e) Se uma função f for par, a composta $g \circ f$ também o será; se f for ímpar, a paridade da composta $g \circ f$ será a mesma de g, caso g possua alguma.

Exercício 1.154. Demonstre as afirmações de acima. Tente generalizar cada uma das três últimas para o caso de n funções ($n > 2$). ◇

Note que os gráficos das funções (reais de uma variável real) pares e ímpares possuem grande apelo visual: as primeiras são simétricas em relação ao eixo das ordenadas OY (reflexão especular); as últimas, em relação à origem (um ponto).

Mas é claro que nem todas as funções de uma variável real podem ser classificadas como pares ou ímpares. Pense, por exemplo, na função $\ln : \mathbb{R}^+ \to \mathbb{R}$ (Figura 1.44).

Entretanto, repare que, como a paridade é um conceito sobre simetria, mesmo dentre as funções sem paridade, podemos encontrar aquelas que possuem simetria em relação a uma reta vertical (como no caso das funções pares) ou em relação a um ponto (como no caso das funções ímpares). E frequentemente é isto o que ocorre: encontramos não as funções pares ou ímpares, mas o que chamamos de funções *simétricas* e *antissimétricas* em relação a um ponto c.

Vamos dizer que um conjunto $X \subset \mathbb{R}$ é simétrico em relação ao ponto c se ele possuir a seguinte propriedade: $c+h \in X \Rightarrow c-h \in X$ (c pode ou não pertencer a X). Assim, de modo análogo à Definição 1.155, temos a seguinte definição:

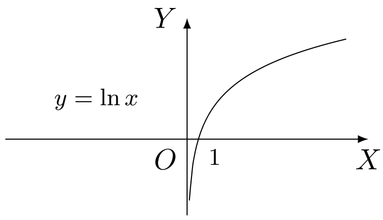


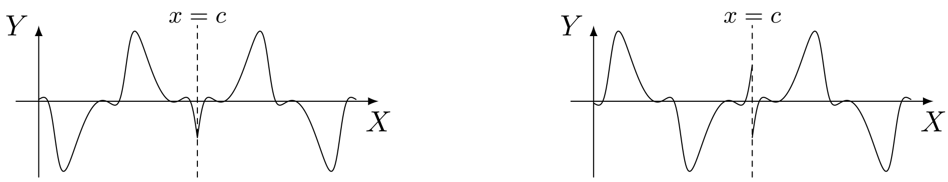
Figura 1.44: Exemplo de função sem paridade.

Definição 1.155. *Sejam $X, Y \subset \mathbb{R}$, X simétrico em relação ao ponto c e uma função $f : X \to Y$. Dizemos que:*

(a) f *é* simétrica em relação a c *se* $f(c+h) = f(c-h)$ *para todo* $c+h \in X$*;*

(b) f *é* antissimétrica em relação a c *se* $f(c+h) = -f(c-h)$ *para todo* $c+h \in X$.

Os aspectos gerais dos gráficos de funções simétricas e antissimétricas são ilustrados na Figura 1.45.



(a) Exemplo de função simétrica. (b) Exemplo de função antissimétrica.

Figura 1.45: Aspectos gerais dos gráficos de uma função simétrica e outra antissimétrica.

Podemos pensar sobre as funções simétricas ou antissimétricas como funções pares ou ímpares transladadas horizontalmente. Compare o Exemplo 1.149 com o próximo exemplo.

Exemplo 1.156. Sejam $\phi, \psi : \mathbb{R} \to \mathbb{R}$ definidas por $\phi(x) = x - 1$ e $\psi(x) = (x-1)^2$. Essas funções são, respectivamente, antissimétrica e simétrica em relação a $x = 1$. (Certifique-se!) Observe seus gráficos na Figura 1.46. Note que essas funções correspondem respectivamente às translações horizontais de f e g do Exemplo 1.149.$\diamond$

Uma função simétrica notável, uma das mais importantes da Estatística Inferencial, com aplicações no estudo de diversos fenômenos das mais variadas áreas, é a *distribuição normal* (também conhecida como distribuição de Gauss ou Gaussiana[51]) com média μ

[51] Em honra a Johann Carl Friedrich Gauss (1777-1855), matemático alemão conhecido como *princeps*

e variância σ^2, $f : \mathbb{R} \to \mathbb{R}$, definida por:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right).$$

Seu gráfico tem o aspecto em forma de sino da Figura 1.47.

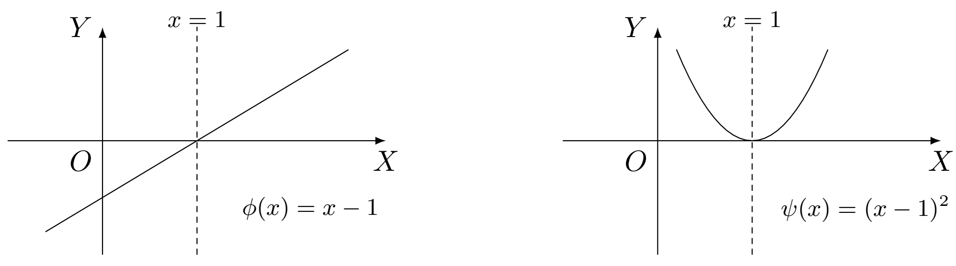
Exercício 1.157. Verifique que a distribuição normal é simétrica, i.e., mostre que $f(c+h) = f(c-h)$ para todo $h \in \mathbb{R}$. $\diamond$

Uma observação final: note que, em última análise, uma função simétrica (resp. antissimétrica) sempre pode ser reduzida a uma função par (resp. ímpar). Basta, para isso, efetuarmos uma mudança de variável $u = x - c$, que nada mais é do que uma bijeção $\phi : X \to U$. Essa transformação pode ser interpretada geometricamente como uma translação horizontal do gráfico da função (ou do sistema de eixos, dependendo do ponto de vista). Repare que, com essa mudança de variável, a propriedade $f(c+h) = f(c-h)$ de função simétrica (resp. $f(c+h) = -f(c-h)$ de função antissimétrica) fica $f(h) = f(-h)$ (resp. $f(h) = -f(-h)$). (Verifique!) Por exemplo, no caso da distribuição normal, com a mudança $u = x - \mu$, obteríamos:

$$f(u) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{u^2}{2\sigma^2}\right),$$

e seu gráfico ficaria como na Figura 1.48.

mathematicorum (em latim: "o príncipe da matemática"ou "o mais notável dos matemáticos"). Gauss trabalhou em muitos campos, tanto da Matemática quando da Física, incluindo Teoria dos Números, Análise, Geometria Diferencial, Geodésia, Magnetismo, Astronomia e Óptica. Seu trabalho teve uma imensa influência em muitas áreas. No que concerne à introdução da distribuição normal, considera-se que sua equação foi publicada num artigo em 1733 por Abraham De Moivre (1667–1754), matemático francês, pioneiro no desenvolvimento da Geometria Analítica e da Teoria da Probabilidade, eleito membro da *Royal Society* em 1697 (12 anos após migrar para a Inglaterra); foi amigo de Isaac Newton e Edmond Halley. Além de Gauss e De Moivre, vale mencionar que outros matemáticos notáveis se debruçaram sobre esse assunto, como Laplace, Robert Adrain (1775–1843) e Lambert Adolphe Jacques Quetelet (1796–1874).



(a) Função antissimétrica. (b) Função simétrica.

Figura 1.46: Simetria de algumas funções simples.

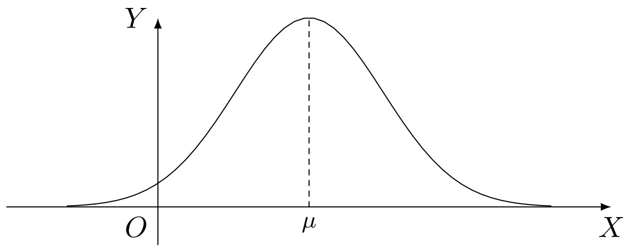


Figura 1.47: Uma das mais notáveis funções simétricas: distribuição normal.

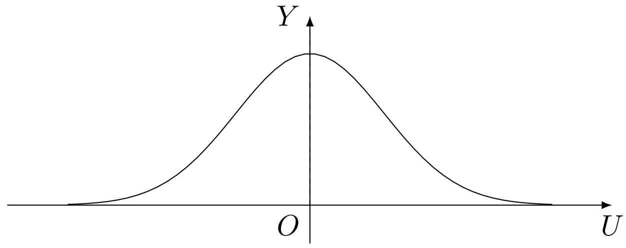


Figura 1.48: Função par obtida de uma função simétrica.

1.14.2 Periodicidade

Em diversas ocasiões, encontramos funções periódicas do espaço ou do tempo, i.e., funções cujos valores se repetem após percorrida certa distância ou após decorrido certo intervalo de tempo. Por exemplo, os pontos das frentes de onda produzidas por fontes sonoras ou luminosas, a configuração de certos ladrilhos regulares revestindo um piso ou a configuração dos reticulados cristalinos podem ser modelados por funções periódicas do espaço. Processos periódicos decorrentes do movimento de um rotor (como a corrente alternada gerada por um dínamo), as estações do ano, ciclos menstruais ou a quantidade de presas e predadores numa população à qual possa se aplicar o modelo de Lotka-Volterra[52] são exemplos de fenômenos periódicos do tempo, bem como todos os fenômenos vibratórios ou oscilatórios.

Definição 1.158. *Dizemos que uma função $f : X \subset U \to Y$ é* periódica *se existir algum $T \in U$ não nulo tal que, para todo $x \in X$, $f(x + T) = f(x)$.*

Repare que o domínio da função não pode ser arbitrário. Sua estrutura algébrica precisa ser tal que $x \in X \Rightarrow x + T \in X$ (T não necessariamente um número e $+$ não necessariamente a soma usual). Não entraremos nos pormenores desse assunto, apenas

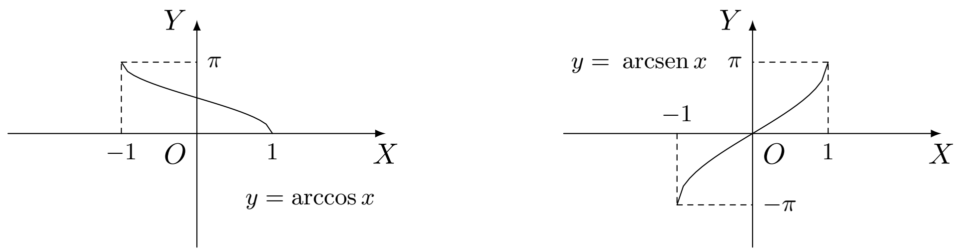
[52]Em honra a Alfred James Lotka (1880–1949) e a Vito Volterra (1860–1940).

queremos ressaltar que esse conjunto não pode ser qualquer. Para quase todos os propósitos (principalmente para as aplicações), será suficiente algum conjunto numérico (como $\mathbb{N}$, $\mathbb{R}$, $\mathbb{C}$, etc.) ou um espaço vetorial (vetores do espaço, espaço de matrizes, espaço de funções, etc.).

Exemplo 1.159. O exemplo mais simples de função periódica talvez seja o da função constante $f : \mathbb{R} \to \mathbb{R}$, $f(x) = c$. Note que $f(x + T) = f(x)$ para qualquer $T > 0$. $\diamond$

Exemplo 1.160. As funções trigonométricas cosseno e seno, $\cos : \mathbb{R} \to \mathbb{R}$ e $\operatorname{sen} : \mathbb{R} \to \mathbb{R}$, são talvez os exemplos mais famosos de funções periódicas. Veja seus gráficos na Figura 1.43. Note que para qualquer x real, $\cos(x + 2\pi) = \cos x$ e $\operatorname{sen}(x + 2\pi) = \operatorname{sen} x$. $\diamond$

Exemplo 1.161. Por outro lado, as inversas das funções $\cos : [0,\, \pi] \to [-1,\, 1]$ e $\operatorname{sen} : [\pi/2,\, \pi/2] \to [-1,\, 1]$, respectivamente, $\operatorname{arccos} : [-1,\, 1] \to [0,\, \pi]$ e $\operatorname{arcsen} : [-1,\, 1] \to [\pi/2,\, \pi/2]$, não são periódicas (Figura 1.49). $\diamond$



(a) Função arco-cosseno.

(b) Função arco-seno.

Figura 1.49: Exemplos de funções que não são periódicas.

Exemplo 1.162. As funções $\operatorname{tg} x = \operatorname{sen} x / \cos x$ (tangente), $\operatorname{cotg} x = \cos x / \operatorname{sen} x$ (cotangente), $\sec x = 1/\cos x$ (secante) e $\operatorname{cossec} x = 1/\operatorname{sen} x$ (cossecante), que derivam das funções seno e cosseno, são periódicas. Observe seus gráficos na Figura 1.50.

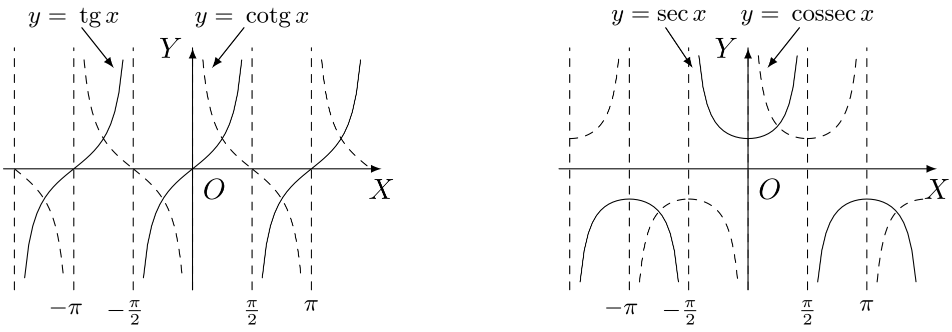
Cumpre observar que os domínios de tais funções consistem dos números reais para os quais os denominadores não se anulam. Note que:

(a) $\operatorname{tg}(x + \pi) = \operatorname{tg} x$;

(b) $\operatorname{cotg}(x + \pi) = \operatorname{cotg} x$;

(c) $\sec(x + 2\pi) = \sec x$;

(d) $\operatorname{cossec}(x + 2\pi) = \operatorname{cossec} x$. $\diamond$



(a) Tangente e cotangente.

(b) Secante e cossecante.

Figura 1.50: Periodicidade de algumas funções trigonométricas.

Exemplo 1.163. A função *dente de serra* $f : \mathbb{R} \to \mathbb{R}$, definida por $f(x) = x - \lfloor x \rfloor$, é outro exemplo de função periódica (veja a Figura 1.51). Um exercício interessante é provar que ela satisfaz: $f(x+1) = f(x)$ para todo $x \in \mathbb{R}$. (Escreva $x = k + \xi$, em que $k \in \mathbb{Z}$ e $\xi \in [0,\, 1)$ são respectivamente a parte inteira e a parte fracionária de x.) $\diamond$

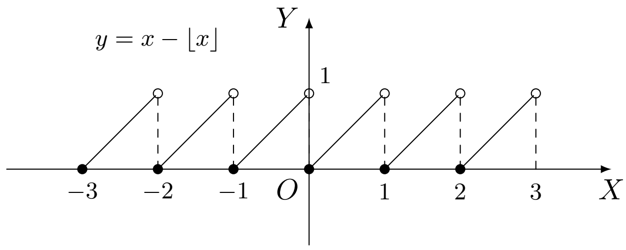


Figura 1.51: Função dente de serra.

Exemplo 1.164. Por último, uma palavra sobre uma função periódica relevante na definição das funções trigonométricas: a função de Euler $E : \mathbb{R} \to C$, que faz corresponder a cada número real t o ponto $E(t) = (x,\, y)$ da circunferência unitária[53] C da seguinte maneira:

1. $E(0) = (1,\, 0)$;

2. se $t > 0$, $E(t)$ será a extremidade de um caminho de comprimento t que parte de $(1,\, 0)$ e percorre C no sentido anti-horário;

3. se $t < 0$, $E(t)$ será a extremidade de um caminho de comprimento $|t|$ que parte de $(1,\, 0)$ e percorre C no sentido horário.

[53] Circunferência de centro na origem e raio unitário.

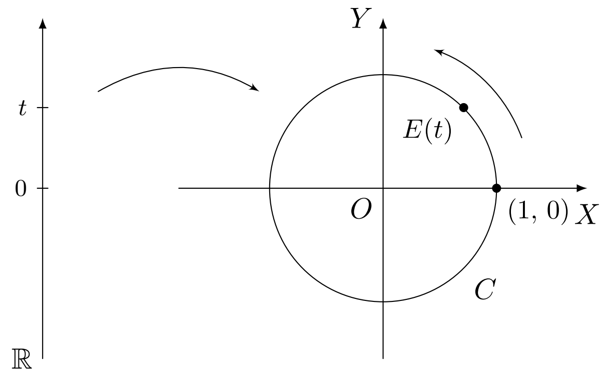


Figura 1.52: Função de Euler: exemplo de função periódica.

Essa função pode ser imaginada como o processo de enrolar a reta sobre a circunferência (Figura 1.52). Cada vez que t percorre um intervalo de comprimento 2π, $E(t)$ percorre na circunferência uma volta completa e, portanto, $E(t + 2\pi) = E(t)$. ◇

1.14.3 Monotonias

Uma das possíveis acepções que o leitor encontrará para a palavra "monotonia" ao abrir um dicionário é: "falta de variação". E naturalmente questionará: falta de variação do quê, se o próprio conceito de função nos induz à ideia de variação? Esclarecemos: monotonia é o conceito utilizado para dizer que uma função preserva ou inverte a relação de ordem (total[54]) do seu domínio. A definição a seguir deixará o conceito mais claro.

Definição 1.165. *Sejam X e Y conjuntos dotados de uma ordem total denotada por $\preceq$ e $f : X \to Y$ uma função de X em Y. Sejam também x_1 e x_2 elementos (arbitrários) de X. Dizemos que f é:*

(a) crescente *quando* $x_1 \prec x_2 \Rightarrow f(x_1) \prec f(x_2)$;

(b) decrescente *quando* $x_1 \prec x_2 \Rightarrow f(x_1) \succ f(x_2)$;

[54]Dizemos que um conjunto X é *totalmente ordenado* ou *linearmente ordenado* por uma relação $\preceq$ se gozar das seguintes propriedades: 1. *reflexividade*: $\forall a \in X$, $a \preceq a$; 2. *transitividade*: $a \preceq b$ e $b \preceq c \Rightarrow a \preceq c$; 3. *antissimetria*: $a \preceq b$ e $b \preceq a \Rightarrow a = b$; e 4. $\forall a, b \in X$, $a \preceq b$ ou $b \preceq a$. O leitor não deve pensar que as ordens usuais $\leqslant$ em $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ ou $\mathbb{R}$ são as únicas existentes! Por exemplo, podemos estabelecer uma relação de ordem em $\mathbb{R}$ da seguinte maneira: sejam $x, y \in \mathbb{R}$; se $x, y \in \mathbb{Q}$, diremos que $x \preceq y$ se $x \leqslant y$; se $x, y \in \mathbb{R} \setminus \mathbb{Q}$, diremos que $x \preceq y$ se $x \leqslant y$; por fim, se $x \in \mathbb{Q}$ e $y \in \mathbb{R} \setminus \mathbb{Q}$, diremos *sempre* que $x \preceq y$. (Verifique que essa é uma relação de ordem total em $\mathbb{R}$.) Para mais detalhes, veja [1].

(c) não decrescente *quando* $x_1 \prec x_2 \Rightarrow f(x_1) \preceq f(x_2)$;

(d) não crescente *quando* $x_1 \prec x_2 \Rightarrow f(x_1) \succeq f(x_2)$.

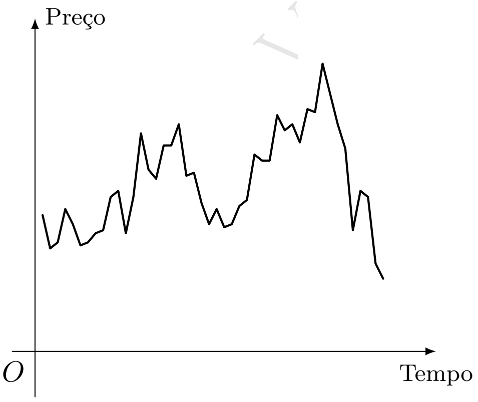
Em todos os casos da Definição 1.165, dizemos que f é *monótona*. Nos dois primeiros, que são casos particulares dos dois últimos, dizemos ainda que f é *estritamente monótona*; note que em ambos f é injetiva.

Certamente, a única função simultaneamente monótona não decrescente e monótona não crescente é a função constante. De fato, se uma função $f : X \to Y$ é monótona não decrescente, então, para $x_1, x_2 \in X$, $x_1 \prec x_2 \Rightarrow f(x_1) \preceq f(x_2)$; se f é também monótona não crescente, $x_1 \prec x_2 \Rightarrow f(x_1) \succeq f(x_2)$. Isso significa que $x_1 \prec x_2 \Rightarrow f(x_1) \preceq f(x_2) \preceq f(x_1)$. Assim, forçosamente, $x_1 \prec x_2 \Rightarrow f(x_1) = f(x_2)$, donde segue o resultado.

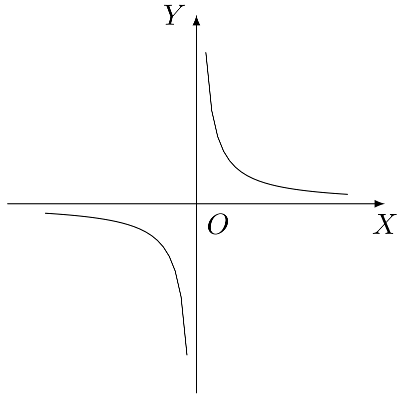
Exemplo 1.166. A função identidade em qualquer conjunto totalmente ordenado é estritamente monótona (crescente). ◇

Recomendamos evitar o uso isolado dos termos *não decrescente* ou *não crescente* para se referir às funções que são dos dois últimos tipos, pois negar que uma função seja decrescente (resp. crescente) não significa que ela seja monótona. (Pense numa função periódica qualquer para ilustrar esse ponto.)

Se você pensou numa função periódica qualquer, deve ter notado que a definição acima não esgota todas as possibilidades, i.e., há funções não monótonas. Pense, por exemplo, no comportamento trimestral do preço de uma *commodity* qualquer, algo como ilustrado na Figura 1.53a, ou na função $f : \mathbb{R} \setminus \{0\} \to \mathbb{R}$, $x \mapsto 1/x$, cujo gráfico é a hipérbole equilátera (Figura 1.53b).



(a) Comportamento do preço de uma *commodity*: um exemplo.



(b) Hipérbole equilátera.

Figura 1.53: Gráficos de funções que não são monótonas.

1.14.4 Limitações

Definição 1.167. *Seja Y um conjunto dotado de uma ordem total denotada por $\preceq$ e uma função $f : X \to Y$. Diremos que f é* limitada *se existirem K, $M \in X$ tais que, para todo $x \in X$, $K \preceq f(x) \preceq M$. Quando $K \preceq f(x)$ $\forall x \in X$ também dizemos que f é* limitada inferiormente*; analogamente, dizemos que f é* limitada superiormente *quando $f(x) \preceq M$ $\forall x \in X$.*

No caso das funções reais de uma variável real, os exemplos mais imediatos são as funções sen e $\cos$.

1.14.5 Taxa de variação

Você deve estar acostumado a escutar ou a ler certas expressões, tais como:

- *quilômetros por litro*: obtidos dividindo-se o número de quilômetros rodados pela quantidade de litros de combustível utilizada;
- *custo por quilowatt-hora*: calculado dividindo-se o custo da eletricidade pela quantidade de quilowatt-hora utilizada;
- *velocidade média*: distância percorrida dividida pelo tempo gasto no percurso.

O que todas essas expressões possuem em comum é o fato de descreverem quanto, em média, uma grandeza varia em relação a outra.

Mais geralmente, temos a seguinte definição:

Definição 1.168. *Seja $f : X \subset \mathbb{R} \to \mathbb{R}$ uma função real de uma variável real. Se x_1, $x_2 \in X$, sendo $x_1 \neq x_2$, dizemos que o número*

$$v = \frac{f(x_2) - f(x_1)}{x_2 - x_1}$$

é a taxa de variação média *da função f entre x_1 e x_2.*

Note que se chamarmos de h a diferença $x_2 - x_1$ da definição acima, sendo $x_1 \neq x_2$, ou seja, $h \neq 0$, podemos dizer equivalentemente que

$$v = \frac{f(x+h) - f(x)}{h}.$$

Alguns autores chamam a taxa de variação média também de *taxa de crescimento média*. Embora nossa preferência seja por aquele nome, não há problema em se utilizar este. Quando a função descrescer, a taxa de crescimento será negativa.

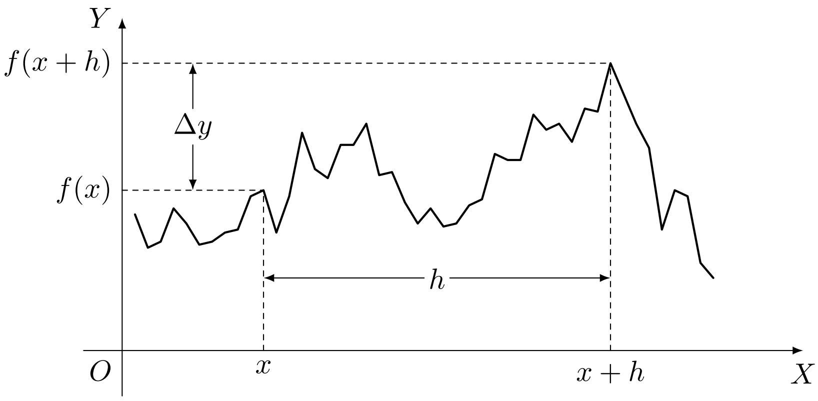


Figura 1.54: Taxa de variação média de uma função num intervalo.

Claramente, v depende dos pontos x_1 e x_2 ou, o que dá no mesmo, do ponto x e de h.

A Figura 1.54 ilustra a ideia da taxa de variação média.

Exemplo 1.169. Suponha que um objeto despenque de um edifício e que a distância em metros que ele percorre na queda após t segundos seja dada pela função $d(t) = 5t^2$.

(a) Qual seria a velocidade média desse objeto entre $t = 1\,s$ e $t = 5\,s$?

(b) E entre $t = x$ e $t = x + h$? ◇

Solução.

(a) Nesse caso,

$$v = \frac{d(5) - d(1)}{5 - 1} = \frac{5 \cdot 5^2 - 5 \cdot 1^2}{5 - 1} = \frac{125 - 5}{4} = 30\,m/s.$$

(b) Agora,

$$v = \frac{d(x+h) - d(x)}{h} = \frac{5(x+h)^2 - 5x^2}{h} = \frac{5(x^2 + 2xh + h^2 - x^2)}{h} = \frac{5h(2x+h)}{h},$$

ou seja,

$$v = 10x + 5h.$$

◆

Nesta ideia de taxa de variação média está a semente para algo um pouco mais sofisticado, que vamos mencionar *en passant*, apenas por curiosidade. Para certas classes de funções, ditas deriváveis, é possível falar sobre a taxa de varição num ponto (ou instantânea, em alusão ao fato de que em muitos problemas a variável independente representa o tempo).

Observe na figura 1.55 o gráfico de uma função f e a reta secante que passa pelos pontos $P = (x_0,\, f(x_0))$ e $Q = (x_0+h,\, f(x_0+h))$. Na figura, foi ilustrada uma situação em que $h > 0$, mas o mesmo vale para $h < 0$.

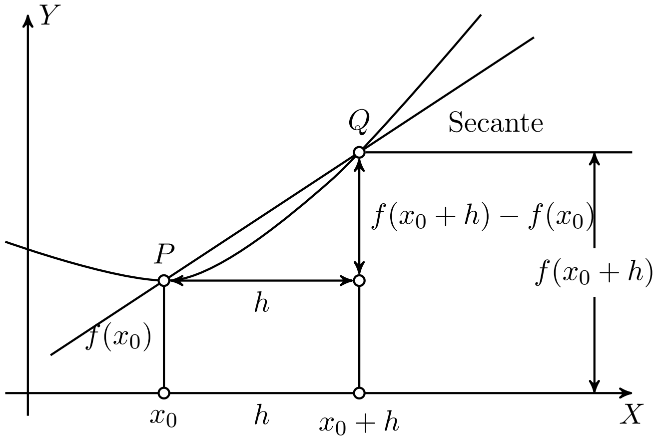


Figura 1.55: Secante ao gráfico de uma função e taxa de variação média.

A taxa de variação média de f entre os pontos P e Q é dada pelo quociente:

$$\frac{f(x_0 + h) - f(x_0)}{h}.$$

Tomando h cada vez mais próximo de zero, obtemos retas secantes que cortam o gráfico de f em dois pontos P e Q cada vez mais próximos. À medida em que $x_0 + h$ se aproxima de x_0, $f(x_0 + h)$ e $f(x_0)$ ficam mais próximos e a secante PQ se aproxima cada vez mais da tangente ao gráfico de f em x_0. No limite desse processo, obtém-se a taxa de variação no ponto x_0:

$$f'(x_0) = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h},$$

que é chamada a *derivada* da função f no ponto x_0.

Note que esta é apenas uma introdução informal deste conceito para instigá-lo. Você deve estudá-lo com diligência num curso de Cálculo.

Parte II

Funções afins e quadráticas

2 Funções afins

"Um certo rei envia 30 homens a seu pomar para plantar árvores. Se eles podem plantar 1000 árvores em 9 dias, em quantos dias 36 homens plantariam 4400 árvores?"

Fibonacci – Problema proposto no Liber Abaci *(1202) para ilustrar a regra de três.*

2.1 Introdução

No capítulo anterior, vimos os principais conceitos envolvendo a definição de função. Continuamos nosso estudo sob o ponto de vista elementar, ou seja, sem o uso do Cálculo Diferencial, com foco nas funções reais de uma variável real que mais aparecem na prática e no desenvolvimento do estudo da Matemática, i.e., funções do tipo $f : X \subset \mathbb{R} \to \mathbb{R}$ que têm como domínio um subconjunto X dos números reais e cujos valores $f(x)$, para todo $x \in X$, são números reais.

A mais simples delas é a função afim, que é o modelo matemático para os problemas cuja taxa de variação de uma grandeza (que será substituida por sua medida, que é um número real) em relação a uma outra é constante. Por exemplo, o preço de uma corrida de táxi em função da distância percorrida, a distância percorrida em função do tempo num movimento uniforme ou como se transformam as medidas de temperatura (ou seja, como uma medida se escreve em função de outra) quando se muda de escala termométrica.

O estudo da função afim será antecedido pelo da função linear, que é o modelo matemático para os problemas de proporcionalidade direta. A função linear nada mais é do que um caso particular da função afim, mas é conveniente começarmos pela função linear para que possamos utilizar o seu teorema de caracterização na demonstração do teorema de caracterização da função afim. Não que esta seja a única forma de se abordar este problema. Ela é apenas a forma que este autor acredita ser conveniente para uma abordagem de um ponto de vista elementar. O leitor fica convidado a experimentar outras abordagens, sobretudo próprias, de um ponto de vista elementar ou não. Você verá mais adiante, num curso de Cálculo, como este problema é facilmente atacado com as ferramentas lá desenvolvidas, utilizando apenas a hipótese de taxa de variação constante.

Iniciaremos com foco nos aspectos algébricos e em seguida estudaremos os aspectos geométricos da função afim. Sua interpretação geométrica é muito profícua, pois permite visualizar o que está sendo feito e aguça a intuição matemática.

2.2 Função linear e proporcionalidade

Definição 2.1. *Uma função* $f : \mathbb{R} \to \mathbb{R}$ *chama-se* linear *quando existe uma constante* $a \in \mathbb{R}$ *tal que* $f(x) = ax$ *para todo* x *no domínio de* f*.*

Como já dissemos, a função linear é o modelo matemático para os problemas de proporcionalidade direta. Dizemos "direta" porque há também o conceito de proporcionalidade inversa. Ambos são casos particulares do conceito geral de proporcionalidade. Na linguagem de funções, temos a seguinte definição.

Definição 2.2. *Uma* proporcionalidade direta *é uma função* $f : \mathbb{R} \to \mathbb{R}$ *tal que* $f(cx) = cf(x)$ *para todo* $c, x \in \mathbb{R}$*. Uma* proporcionalidade inversa *é uma função* $g : \mathbb{R}^* \to \mathbb{R}^*$ *tal que* $g(cx) = g(x)/c$ *para todo* $c, x \in \mathbb{R}^*$ $(\mathbb{R}^* = \mathbb{R} \setminus \{0\})$*. Uma* proporcionalidade *é uma função de qualquer um dos tipos anteriores.*

É claro que se $f(cx) = cf(x)$ para todo $c, x \in \mathbb{R}$, então, escrevendo-se $a = f(1)$, $f(x) = f(x \cdot 1) = xf(1) = xa$, ou seja, $f(x) = ax$ para todo $x \in \mathbb{R}$. Portanto f é linear, provando nossa assertiva acima de que ela é o modelo para os problemas de proporcionalide direta.

Analogamente, se $g(cx) = g(x)/c$ para todo $c, x \in \mathbb{R}^*$, então, escrevendo-se $a = g(1)$, $g(x) = g(x \cdot 1) = g(1)/x = a/x$, ou seja, $g(x) = a/x$ para todo $x \in \mathbb{R}^*$.

Note que a proporcionalidade inversa só tem sentido quando se trata de grandezas não nulas, pois, por definição, $g(cx) = g(x)/c$ para todo $c, x \in \mathbb{R}^*$, logo, se invertermos os papéis de x e c, deve valer $g(cx) = g(c)/x$. Por isso a escolha de $\mathbb{R}^*$ como domínio e contradomínio de g.

O conceito de proporcionalidade é talvez um dos mais difundidos na cultura de todos os povos e um dos mais antigos também. Apesar disso, é interessante observar como muitas pessoas "cultas", mas "incultas" matematicamente, usam a palavra "proporcional" de modo equivocado para dizer que uma grandeza cresce ou decresce em relação a outra. Essa é uma condição necessária, mas não suficiente: conforme seja $a > 0$

ou $a < 0$, f (resp. g) é crescente ou decrescente (resp. decrescente ou crescente). A função $\exp : \mathbb{R} \to \mathbb{R}$, $x \mapsto \exp(x)$, por sua vez, é crescente, mas a relação entre x e sua imagem não é de proporcionalidade. Duas grandezas só podem ser ditas proporcionais se cumprem os termos da Definição 2.2.

Doravante, vamos fixar nossa atenção apenas na proporcionalidade direta, que passaremos a chamar apenas de “proporcionalidade” para simplificar a comunicação.

Em suma, a definição de proporcionalidade equivale a dizer que a grandeza y é proporcional à grandeza x quando existe um número a, chamado *constante de proporcionalidade*, tal que $y = ax$ para todo valor de x.

Na prática, há situações em que a fórmula $y = ax$ é dada explicitamente (ou quase). Por exemplo, se um grama de ouro custa a reais, então x gramas custam $y = ax$ reais.

No entanto, há muitas outras situações em que a constante de proporcionalidade não é dada ou sequer tem relevância alguma para o problema. Por exemplo, nas aplicações do Teorema de Tales.

Considere no plano um ângulo $A\widehat{O}B$ e uma reta r que não é paralela ao lado OA nem ao lado OB (Figura 2.1). Dado qualquer segmento de reta de comprimento x contido em OA, as paralelas a r traçadas por suas extremidades determinam sobre o lado OB um segmento de comprimento y. O teorema de Tales assegura que y é proporcional a x. (A seguir, após estudar o teorema de caracterização da função linear, também chamado de teorema fundamental da proporcionalidade, seremos capazes de justificar por quê.) Mas qual a importância da constante de proporcionalidade $a = y/x$ neste problema? Por acaso, $a = \operatorname{sen}\alpha / \operatorname{sen}\beta$, mas isso não tem muita importância.

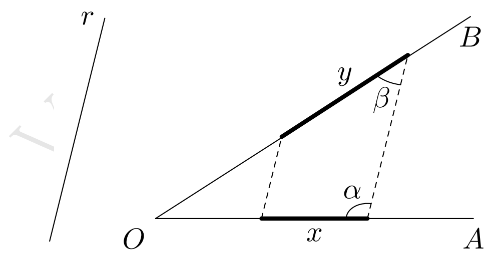


Figura 2.1: Teorema de Tales e proporcionalidade.

O que esse problema ilusta é o fato de que nos problemas de proporcionalidade o que interessa é saber apenas que se $y = f(x)$ e $y' = f(x')$ então $y/x = y'/x'$ é constante. Isso nos permite determinar uma das quatro grandezas se forem conhecidas as outras três. Nisso consiste a famigerada “regra de três”.

Antes de utilizá-la, porém, precisamos garantir que a correspondência $x \mapsto y$ é uma proporcionalidade. Para tanto, de acordo com a Definição 2.2, precisamos que a condição $f(cx) = cf(x)$ seja cumprida por todos os valores reais de c e x, em particular para todo c real. Isso é fácil de verificar quando c é inteiro, mas o mesmo não pode ser dito para os demais casos, por exemplo, quando c for irracional. A boa notícia é que $f(nx) = nf(x)$ para todo $x \in \mathbb{R}$ e todo $n \in \mathbb{Z}$, contanto que f seja estritamente monótona, é condição suficiente para garantir que f é uma função linear. Esse é o conteúdo do próximo teorema.

Teorema 2.3 (Teorema Fundamental da Proporcionalidade). *Seja $f : \mathbb{R} \to \mathbb{R}$ uma função estritamente monótona. As seguintes afirmações são equivalentes:*

(1) $f(nx) = nf(x)$, $\forall n \in \mathbb{Z}$, $\forall x \in \mathbb{R}$;

(2) $f(x) = ax$, $\forall x \in \mathbb{R}$, *em que* $a = f(1)$;

(3) $f(x + y) = f(x) + f(y)$, $\forall x, y \in \mathbb{R}$.

Demonstração. Basta demonstrarmos as implicações: (1) ⇒ (2), (2) ⇒ (3) e (3) ⇒ (1).

(1) ⇒ (2) Seja $r = p/q$ ($p, q \in \mathbb{Z}$, $q \neq 0$) um número racional. Como $p = qr$,

$$qf(rx) = f(qrx) = f(px) = pf(x),$$

logo

$$f(rx) = \frac{p}{q}f(x) = rf(x).$$

Seja $a = f(1)$. Como $f(0) = f(0 \cdot 0) = 0 \cdot f(0) = 0$, a monotonicidade de f implica que $a = f(1) > f(0) = 0$ caso f seja crescente ou $a = f(1) < f(0) = 0$ caso f seja decrescente. Sem perda de generalidade, suponha f crescente. (Para f decrescente, a demonstração seria análoga.) Assim, $a > 0$. Além disso,

$$f(r) = f(r \cdot 1) = r \cdot f(1) = r \cdot a = ar, \forall r \in \mathbb{Q}.$$

Suponha agora, por absurdo, que $f(x) \neq ax$, para algum x real (necessariamente irracional). Apenas para fixar ideias, vamos considerar $f(x) < ax$. (O caso $f(x) > ax$ é análogo.) Segue-se que $f(x)/a < x$. De acordo com a Proposição 1.11, existe um racional r tal que $f(x)/a < r < x$, ou seja, $f(x) < ar = f(r) < ax$, absurdo, pois como f é crescente e $r < x$, deveríamos ter $f(r) < f(x)$. Portanto, devemos ter $f(x) = ax$, $\forall x \in \mathbb{R}$. Isto completa a prova de que (1) ⇒ (2).

(2) ⇒ (3) Trivial. Supondo (2) verdadeira e tomando x e y reais quaisquer:

$$f(x + y) = a(x + y) = ax + ay = f(x) + f(y).$$

(3) ⇒ (1) Primeiro, note que, por (3), $f(0) = f(0 + 0) = f(0) + f(0)$, logo:

$$f(0) = 0.$$

Ainda, por (3), $0 = f(0) = f(x + (-x)) = f(x) + f(-x)$, ou seja,

$$f(-x) = -f(x).$$

(Isto é, f é uma função ímpar.) O restante da demonstração se faz por indução (veja o Apêndice E) sobre os naturais e, por essas duas últimas igualdades, estende-se o resultado para os demais inteiros. Vejamos. Em primeiro lugar, é óbvio que $f(1 \cdot x) = 1 \cdot f(x)$. Suponha agora que, para $n \in \mathbb{N}$, $f(nx) = nf(x)$; segue-se que $f((n+1)x) = f(nx + x) = f(nx) + f(x) = nf(x) + f(x) = (n+1)f(x)$, em que utilizamos (3) na segunda igualdade e a hipótese de indução na igualdade seguinte. Como $f(1 \cdot x) = 1 \cdot f(x)$ e $f(nx) = nf(x) \Rightarrow f((n+1)x) = (n+1)f(x)$, $\forall n \in \mathbb{N}$, segue-se, pelo Princípio da Indução, que $f(nx) = nf(x)$, $\forall n \in \mathbb{N}$. Agora, suponha n um inteiro negativo. Como $f(-x) = -f(x)$, segue-se que $f(nx) = -f(-nx) = -(-n)f(x) = nf(x)$, pois, na penúltima igualdade, utilizamos o fato de que $-n > 0$. Como x é um real arbitrário, isso completa a demonstração de que $f(nx) = nf(x)$, $\forall n \in \mathbb{Z}$, $\forall x \in \mathbb{R}$. ■

Observação 2.4. No enunciado do Teorema 2.3, utilizamos a hipótese de que f fosse estritamente monótona (poderíamos ter dito monótona injetiva, o que dá no mesmo), apenas para provar que se $f(r) = ar$ para todo racional r, então $f(x) = ax$ para todo x real. Outra hipótese possível, na verdade equivalente neste caso, seria a de que f fosse contínua, pois, como você estudará num curso de Cálculo ou Análise, todo número real x é limite de uma sequência de números racionais r_n, logo a continuidade nos dá:

$$f(x) = \lim_{n\to\infty} f(r_n) = \lim_{n\to\infty} ar_n = ax.$$

Deve-se esclarecer que uma dessas hipóteses – monotonicidade, continuidade ou algo equivalente – deve estar presente no enunciado do Teorema Fundamental da Proporcionalidade, pois existem funções $f : \mathbb{R} \to \mathbb{R}$ tais que $f(nx) = nf(x)$ para todo $n \in \mathbb{Z}$ e todo $x \in \mathbb{R}$, mas f não é da forma $f(x) = ax$, como mostramos no exemplo a seguir. ○

Exemplo 2.5. Defina uma função $f : \mathbb{R} \to \mathbb{R}$ da seguinte maneira:

$$f(x) = \begin{cases} x, & \text{se } x \text{ é racional} \\ -x, & \text{se } x \text{ é irracional} \end{cases}.$$

Note que $f(nx) = nf(x)$ para todo $n \in \mathbb{Z}$ e todo $x \in \mathbb{R}$, mas f não é linear. ◇

Em grande parte das aplicações do Teorema 2.3, as grandezas envolvidas são tais que suas medidas só podem ser expressas por números reais positivos. Nesses casos, temos uma função crescente $f : \mathbb{R}^+ \to \mathbb{R}^+$, em que $\mathbb{R}^+ = \{x \in \mathbb{R};\ x > 0\}$, e as afirmações do Teorema 2.3 leem-se assim:

(1') $f(nx) = nf(x)$, $\forall n \in \mathbb{N}$, $\forall x \in \mathbb{R}^+$;

(2') $f(x) = ax$, $\forall x \in \mathbb{R}^+$, em que $a = f(1)$;

(3') $f(x + y) = f(x) + f(y)$, $\forall x, \, y \in \mathbb{R}^+$.

Para demonstrar que o Teorema Fundamental da Proporcionalidade continua válido neste novo contexto, basta fazermos a seguinte extensão ímpar de f:

$$F(x) = \begin{cases} f(x) & \text{se } x > 0 \\ 0 & \text{se } x = 0 \\ -f(-x) & \text{se } x < 0 \end{cases} .$$

Note que $F : \mathbb{R} \to \mathbb{R}$ está bem definida e é de fato uma função (não há excessão nem ambiguidade na sua definição). Além disso, F preserva a monotonicidade de f. Considere $x_1 < x_2$ e um dos seguintes casos: $0 = x_1 < x_2$, $x_1 < x_2 = 0$, $x_1 < x_2 < 0$, $x_1 < 0 < x_2$ ou $0 < x_1 < x_2$. É fácil ver $-f(x) < 0 < f(x)$ para todo x do domínio de f, pois f assume valores em $\mathbb{R}^+$; sendo assim, no caso em que $0 = x_1 < x_2$, temos $F(x_1 = 0) = 0 < f(x_2) = F(x_2)$. Analogamente, $x_1 < x_2 = 0 \Rightarrow F(x_1) < F(x_2 = 0)$. No caso em que $x_1 < x_2 < 0$, essa desigualdade equivale a $0 < -x_2 < -x_1$; como f é crescente, $f(-x_2) < f(-x_1)$, ou seja, $F(x_1) = -f(-x_1) < -f(-x_2) = F(x_2)$. Os dois últimos casos são análogos.

Com F assim definida, cada uma das afirmações (1'), (2') e (3') para f equivale a uma das afirmações (1), (2) e (3) para F. Você pode mostrar isso uma a uma. Como as afirmações de cada um desses trios se equivalem entre si, um jeito simples de verificar essa afirmação é mostrando a equivalência de uma das afirmações de um terno com uma das afirmações do outro, por exemplo, que (2) $\Leftrightarrow$ $(2')$.

De fato, (2) implica $F(x) = ax$, $\forall x \in \mathbb{R}^+$, com $a = F(1)$; como $F(x) = f(x)$ se $x > 0$, segue-se que $f(x) = ax$, $\forall x \in \mathbb{R}^+$, em que $a = F(1) = f(1)$. Reciprocamente, se vale (2'), então

$$F(x) = \begin{cases} ax & \text{se } x > 0 \\ a \cdot 0 & \text{se } x = 0 \\ -a(-x) & \text{se } x < 0 \end{cases} ,$$

ou seja, $F(x) = ax$, $\forall x \in \mathbb{R}$, em que $a = f(1) = F(1)$.

A importância do Teorema 2.3 reside no fato de que, para saber se uma função $f : \mathbb{R} \to \mathbb{R}$ é linear ou $g : \mathbb{R}^+ \to \mathbb{R}^+$ é uma restrição de uma função linear, basta apenas verificar se a função dada é estritamente monótona e se (2) no caso de f ou (2') no caso de g é satisfeita. Poderíamos reformular isso no seguinte teorema, equivalente ao Teorema 2.3:

Teorema 2.6 (Teorema de Caracterização da Função Linear). *Seja $f : \mathbb{R} \to \mathbb{R}$ uma função estritamente monótona. Então $f(nx) = nf(x)$, $\forall x \in \mathbb{R}$, $\forall n \in \mathbb{Z}$, se, e somente se, f é uma função linear.*

Alternativamente, nos casos em que as grandezas envolvidas são medidas por números reais positivos, podemos utilizar a seguinte variante do teorema anterior:

Teorema 2.7. *Seja $f : \mathbb{R}^+ \to \mathbb{R}^+$ uma função crescente. Então $f(nx) = nf(x)$, $\forall x \in \mathbb{R}^+$, $\forall n \in \mathbb{N}$, se, e somente se, $f(x) = ax$, em que $a = f(1)$.*

O exemplo a seguir ilustra uma interessante aplicação do Teorema Fundamental da Proporcionalidade na dedução da fórmula de área do retângulo.

Exemplo 2.8. Seja $f(x)$ a área do retângulo de altura a e base x. Certamente $f : \mathbb{R}^+ \to \mathbb{R}^+$ é uma função, pois, fixada a altura a, a cada número positivo x corresponde uma única área $f(x)$. Além disso, f é crescente. Com efeito, se $x_1 < x_2$, então a área $f(x_2)$ do retângulo de base x_2 é igual à área $f(x_1)$ do retângulo de base x_1 acrescida da área de um retângulo de base $x_2 - x_1$, i.e., $f(x_2) = f(x_1) + f(x_2 - x_1)$, logo $x_1 < x_2 \Rightarrow f(x_1) < f(x_2)$.

Chame agora $u = x_1$ e $v = x_2 - x_1$. Certamente $u, v \in \mathbb{R}^+$. Temos então: $f(u+v) = f(u) + f(v)$, $\forall u, v \in \mathbb{R}^+$. Logo, pelo item (3) do Teorema 2.3, $f(x) = Ax$, em que $A = f(1)$ é a área de um retângulo de altura a e base 1.

Outra forma de se chegar à mesma conclusão é notar que um retângulo de base nx pode ser decomposto em n retângulos de mesma altura a, cada um com base x, logo $f(nx) = nf(x)$, isso para qualquer $x \in \mathbb{R}^+$ e todo $n \in \mathbb{N}$. Assim, pelo Teorema 2.7, $f(x) = Ax$, em que $A = f(1)$. (Figura 2.2.)

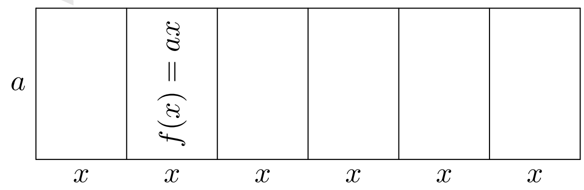


Figura 2.2: Retângulo decomposto em n retângulos de mesma altura.

O mesmo raciocínio aplicado aos retângulos de altura a e base 1 nos faz concluir que $A = Ua$, em que U é a área do retângulo de base e altura iguais a 1; ora, mas este é o quadrado de lado 1, o qual é, por definição, a unidade de área. Portanto, $U = 1$ e $A = a$.

Provamos, assim, que a área de um retângulo de altura a e base x é igual a ax. $\diamond$

Exemplo 2.9. Se comprarmos um lote de ações de uma determinada empresa listada na bolsa de valores pagando x reais hoje, depois de um ano, supondo que a empresa permaneça listada, teremos $f_p(x)$ reais, sendo p o preço por ação no momento da compra. Evidentemente, f_p é uma função crescente de x, pois quanto mais ações ao preço unitário p se adquire hoje, maior o valor que se tem ao final de um ano, independentemente da trajetória de preços e da cotação final. Além disso, $f_p(nx) = nf_p(x)$ para todo $n \in \mathbb{N}$ e todo x. De fato, para um dado preço unitário p, essa igualdade significa que é indiferente comprar de uma vez n lotes pagando-se nx ou adquirir n lotes separadamente, cada um pelos mesmos x reais. O Teorema 2.7 nos permite concluir que $f_p(x)$ é proporcional ao valor x pago pelas ações e nos diz mais: se cada 1 real aplicado se transforma(ria) em a reais depois de findo um ano, então x se transforma(riam) em $f_p(x) = ax$ reais.

Supondo que investimos 8 mil reais e, após um mês, temos $11,5$ mil reais, quanto teríamos se tivéssemos investido 10 mil reais? Como já vimos que trata-se de uma proporcionalidade entre o valor inicial e o valor final, é fácil concluir que $11,5/8 = y/10$ e, portanto, a resposta é $y = 14,375$ mil reais. Note que conhecer a constante a é completamente desnecessário neste caso. $\diamond$

Exercício 2.10. Mostre que se $f : \mathbb{R}^* \to \mathbb{R}^*$ é uma função estritamente monótona, as seguintes afirmações são equivalentes:

(1) $f(nx) = f(x)/n$, $\forall n \in \mathbb{Z}^*$, $\forall x \in \mathbb{R}^*$;

(2) $f(x) = a/x$, $\forall x \in \mathbb{R}^*$, em que $a = f(1)$;

(3) $f(x + y) = f(x)f(y)/(f(x) + f(y))$, $\forall x, y \in \mathbb{R}^*$. $\diamond$

2.2.1 Grandeza proporcional a várias outras

Em muitas situações, deparamo-nos com uma grandeza y de tal modo relacionada com outras, digamos $x_1, x_2, \ldots, x_n$, que, a cada escolha de valores para estas últimas, corresponde um valor bem determinado para a primeira, ou seja, temos uma função $f : \mathbb{R}^n \to \mathbb{R}$, tal que $y = f(x_1, x_2, \ldots, x_n)$.

Nestas ocasiões, também podemos falar em proporcionalidade direta ou inversa.

Definição 2.11.

1. *Dizemos que $f : \mathbb{R}^n \to \mathbb{R}$ é uma* proporcionalidade (direta) *em relação a x_i quando:*

   (a) para quaisquer valores fixados de $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n$, f é uma função

estritamente monótona de x_i; e

(b) *para todo $n \in \mathbb{Z}$ e todo $x_1, \ldots, x_n \in \mathbb{R}$,*

$$f(x_1, \ldots, nx_i, \ldots, x_n) = nf(x_1, \ldots, x_i, \ldots, x_n).$$

2. *Analogamente, dizemos que $g : (\mathbb{R}^*)^n \to \mathbb{R}^*$ é uma* proporcionalidade inversa *em relação a x_i quando:*

(a') *para quaisquer valores fixados de $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n$, g é uma função estritamente monótona de x_i; e*

(b') *para todo $n \in \mathbb{Z}^*$ e todo $x_1, \ldots, x_n \in \mathbb{R}^*$,*

$$g(x_1, \ldots, nx_i, \ldots, x_n) = \frac{1}{n} g(x_1, \ldots, x_i, \ldots, x_n).$$

Segue-se do Teorema Fundamental da Proporcionalidade que as propriedades (b) e (b') acima valem respectivamente para c e $d \neq 0$ reais quaisquer no lugar de n.

Corolário 2.12. *Seja a função $f : \mathbb{R}^m \times (\mathbb{R}^*)^{n-m} \to \mathbb{R}$, com*

$$y = f(x_1, \ldots, x_m, x_{m+1}, \ldots, x_n).$$

Se y é (diretamente) proporcional a $x_1, \ldots, x_m$ e inversamente proporcional a $x_{m+1}, \ldots, x_n$, então:

$$f(x_1, \ldots, x_m, x_{m+1}, \ldots, x_n) = a \cdot \frac{x_1 \cdots x_m}{x_{m+1} \cdots x_n},$$

em que $a = f(1, 1, \ldots, 1)$.

Demonstração.

$$f(x_1, \ldots, x_m, x_{m+1}, \ldots, x_n) = f(x_1 \cdot 1, \ldots, x_m \cdot 1, x_{m+1} \cdot 1, \ldots, x_n \cdot 1).$$

Aplicando-se sucessivas vezes a Definição 2.11 ao segundo membro, obtém-se:

$$f(x_1, \ldots, x_m, x_{m+1}, \ldots, x_n) = \frac{x_1 \cdots x_m}{x_{m+1} \cdots x_n} \cdot f(1, 1, \ldots, 1).$$

Por fim, definindo-se $a := f(1, 1, \ldots, 1)$, segue-se a tese:

$$f(x_1, \ldots, x_m, x_{m+1}, \ldots, x_n) = a \cdot \frac{x_1 \cdots x_m}{x_{m+1} \cdots x_n}.$$

■

Para fixar ideias, considere, por exemplo, que $y = f(x_1, x_2, x_3, x_4, x_5)$ é (diretamente) proporcional a x_1 e x_2 e inversamente proporcional a x_3, x_4 e x_5. Então, tomando-se $a = f(1, 1, 1, 1, 1)$, tem-se:

$$y = f(x_1, x_2, x_3, x_4, x_5) = a \cdot \frac{x_1 \cdot x_2}{x_3 \cdot x_4 \cdot x_5}.$$

Exemplo 2.13 (Adaptado de [9]). A lei da gravitação universal, de Newton, afirma que dois corpos, de massas m e m' respectivamente, cujos centros de massa estão situados a uma distância d um do outro, se atraem segundo uma força cuja intensidade F é proporcional a essas massas e inversamente proporcional a d^2. Resulta do acima exposto que:

$$F = G \cdot \frac{mm'}{d^2},$$

em que a constante (gravitacional) G depende do sistema de unidades utilizado. $\diamond$

Exemplo 2.14 (Adaptado de [9]). A noção de grandeza proporcional a várias outras permite deduzir a fórmula do volume de um bloco retangular. O volume de um sólido geométrico Ω, que se escreve $\mathrm{vol}(\Omega)$, é um número real com as seguintes propriedades:

1. se o sólido Ω está contido propriamente no sólido Ω', então:

$$\mathrm{vol}(\Omega) < \mathrm{vol}(\Omega');$$

2. se o sólido Σ é a reunião de dois sólidos adjacentes Ω e Ω', então:

$$\mathrm{vol}(\Sigma) = \mathrm{vol}(\Omega) + \mathrm{vol}(\Omega').$$

Seja Ω um bloco retangular (Figura 2.3) cujas arestas medem x, y e z. Considere $\mathcal{R}$ o conjunto de todos os blocos retangulares e seja a função $\mathrm{vol} : \mathcal{R} \to \mathbb{R}^+$ que associa cada bloco retangular Ω ao número positivo $w = \mathrm{vol}(\Omega)$. Das propriedades 1 e 2 acima

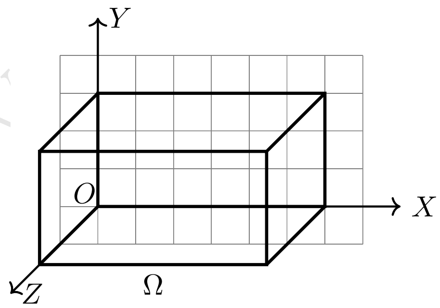


Figura 2.3: Bloco retangular.

e da Definição 2.11 resulta que w é proporcional a x, y e z. Portanto,

$$\mathrm{vol}(\Omega) = a \cdot xyz,$$

em que a é o volume do bloco retangular cujas arestas medem 1. Como o volume de tal bloco é, por definição, igual à unidade, segue-se que $a = 1$ e portanto:

$$\mathrm{vol}(\Omega) = xyz. \qquad \diamond$$

Por último, vamos resolver o problema da epígrafe deste capítulo, retomando-o no exemplo a seguir.

Exemplo 2.15. Um certo rei envia 30 homens a seu pomar para plantar árvores. Se eles podem plantar 1000 árvores em 9 dias, em quantos dias 36 homens plantariam 4400 árvores? ◇

Solução. O número de dias D para plantar árvores é uma função $D = D(a,\, h)$ da quantidade a de árvores a serem plantadas e do número h de homens que trabalharão. Em condições normais, é razoável admitir as seguintes hipóteses:

1. D é uma função crescente de a (quanto mais árvores a serem plantadas, mais tempo levará);
2. D é uma função decrescente de h (quanto mais trabalhadores, menos dias são necessários);
3. fixado h, $D(na,\, h) = nD(a,\, h)$ (se um lote de a árvores leva $D(a,\, h)$ dias para ser plantado, n lotes devem levar $nD(a,\, h)$ dias); e
4. fixado a, $D(a,\, nh) = D(a,\, h)/n$ (se um grupo de h homens precisa de $D(a,\, h)$ dias para cumprir a tarefa, n grupos podem dividir igualmente entre si a tarefa e devem levar $D(a,\, h)/n$ dias para cumprí-la).

Satisfeitas essas hipóteses e tendo em vista a Definição 2.11, é plausível admitir que D seja proporcional a a e inversamente proporcional a h, isto é:

$$D(a,\, h) = k \cdot \frac{a}{h}.$$

Sabemos que

$$9 = D(1000,\, 30) = k \cdot \frac{1000}{30} = \frac{100k}{3}$$

e queremos saber

$$x = D(4400,\, 36) = k \cdot \frac{4400}{36} = \frac{1100k}{9};$$

dividindo x por 9, obtemos:

$$\frac{x}{9} = \frac{1100k/9}{100k/3} = \frac{1100k \cdot 3}{100k \cdot 9} = \frac{11}{3},$$

ou seja, $x = 9 \cdot 11/3 = 33$ dias é o tempo que 36 homens plantariam 4400 árvores. ◆

Observação 2.16. Note que não foi necessário calcular a constante k para resolver o problema acima. Poderíamos tê-lo feito, mas é desnecessário, como o é em geral nos problemas de proporcionalidade. (Relembre a discussão sobre o teorema de Tales acima. A ideia é análoga.) ○

Observação 2.17. Neste ponto, o estudante poderia se queixar levantando as seguintes questões:

1. Como podemos saber se os homens manterão o mesmo ritmo com o passar dos dias (ou mesmo durante um dia)?

2. Como garantir que algum fator externo, como uma chuva, não irá atrasar a tarefa?

3. Alguns homens não poderiam ficar doentes e faltar?

Poderíamos tentar argumentar respectivamente que, na média, os homens mantêm o mesmo ritmo (a velocidade dos menos motivados seria compensada pela agilidade dos mais motivados), que eles poderiam trabalhar somente em dias favoráveis e que haveria mão de obra reserva. Porém, o mais justo é reconhecer (e esta é a lição importante para as quais estas justas queixas chamam a atenção) que a proporcionalidade pode ocorrer dentro de um certo intervalo mas não valer fora dele. Quanto mais nos afastamos do domínio e validade do modelo, mais perdemos em precisão.

Um exemplo típico é o da Lei de Hooke, segundo a qual o comprimento de uma mola é proporcional à intensidade da força que se aplica sobre ela. Se a força for muito pequena, não conseguirá mover a mola e, caso seja muito grande, cada acréscimo de sua intensidade provocará acréscimos cada vez menores no comprimento da mola, terminando por rompê-la. ◦

2.3 Função afim

Definição 2.18. *Uma função $f : \mathbb{R} \to \mathbb{R}$ chama-se* afim *quando existem constantes $a, b \in \mathbb{R}$ tais que $f(x) = ax + b$ para todo x no domínio de f.*

Os números (reais abitrários) a e b são chamados de *coeficientes*.

Frequentemente, o leitor encontrará a nomenclatura "função do primeiro grau" para a função afim em textos didáticos ou em exames. Essa terminologia não é boa e, a rigor, está equivocada, afinal, o que é o "grau" de uma função? Qual seria, por exemplo, o grau da função seno ou da função exponencial? Funções não têm grau. Grau é um atributo dos polinômios. Por existir uma correspondência biunívoca entre polinômios e funções polinomiais, como estudaremos adiante, podemos tomar a liberdade de atribuir grau às funções dessa classe, mas somente a elas. O mais provável é que o uso (equivocado) da expressão "função do primeiro grau" para se referir às funções afins se deva ao fato de alguns autores terem em mente as funções polinomias de grau 1. Ocorre, porém, que isso obrigaria o coeficiente a a ser diferente de zero na Definição 2.18, o que excluiria daquela definição as funções constantes. Por essas razões, recomendamos evitar a expressão "função do primeiro grau" para se referir às funções afins.

Exemplo 2.19. Os exemplos mais simples possíveis de funções afins são os seguintes casos particulares:

(a) as *funções constantes* $f : \mathbb{R} \to \mathbb{R}$, $f(x) = b$;

(b) as *funções lineares* $L : \mathbb{R} \to \mathbb{R}$, $L(x) = ax$;

(c) as *translações* $T : \mathbb{R} \to \mathbb{R}$, $T(x) = x + b$; e

(d) a *função identidade* $id_{\mathbb{R}} : \mathbb{R} \to \mathbb{R}$, $id_{\mathbb{R}}(x) = x$. ◇

Usualmente, encontramos não as funções afins na letra estrita da definição, em que o domínio é o conjunto dos números reais, mas sim uma restrição da função afim a um subconjunto de $\mathbb{R}$. Veja o próximo exemplo.

Exemplo 2.20. Sendo b o preço da bandeirada e a o preço do quilômetro rodado, o custo total de uma corrida de táxi que percorre x quilômetros é uma função $g : \mathbb{R}^+ \to \mathbb{R}$, dada por $g(x) = ax + b$. Note: sendo $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto ax + b$, $g = f|\mathbb{R}^+$, ou seja, g é a restrição de f ao conjunto $\mathbb{R}^+$. ◇

Por serem comuns situações como a do exemplo acima, em que encontramos uma função afim restrita a um subconjunto de $\mathbb{R}$, e por tornar a comunicação mais rápida, fica difícil resistir à tentação de usar a linguagem inexata e chamar as restrições também de funções afins. Seguindo o bom-senso, faremos uso dessa linguagem (inexata) daqui em diante, evitando a forma mais longa e pedante (e correta) do tipo "a função $g : X \to \mathbb{R}$ é a restrição da função afim $f : \mathbb{R} \to \mathbb{R}$ ao conjunto $X \subset \mathbb{R}$" ou "a função afim $f : \mathbb{R} \to \mathbb{R}$ é a extensão da função $g : X \to \mathbb{R}$, $X \subset \mathbb{R}$". Mas é importante, como sempre, a cada momento, ter em mente a noção precisa do que se está dizendo.

Como vimos no capítulo anterior, para conhecermos uma função f é preciso que sejam dados seu domínio, seu contradomínio e sua lei de correspondência (explícita ou implicitamente). Dados os dois primeiros ingredientes, quando a lei de correspondência é fornecida de modo que possamos obter a variável independente conhecida a(s) variável(is) dependente(s), seja por uma fórmula ou um algoritmo, dizemos que a função foi dada na forma explícita. No caso das funções afins, isso significa fornecer os coeficientes a e b.

No entanto, é possível, mediante critérios que estudaremos logo a seguir, saber se uma função $f : \mathbb{R} \to \mathbb{R}$ é afim sem que esses coeficientes (e portanto a lei de correspondência) sejam fornecidos explicitamente. Isso é de grande importância, pois, diferentemente do problema da corrida do táxi, o mais comum é nos depararmos com situações em que não são feitas menções explícitas aos coeficientes a nem b e, por isso, não sabemos *a priori* se a função afim é o modelo adequado a se utilizar naquelas situações. Precisamos, portanto, dispor de critérios que nos permitam comparar as características do problema em estudo com as propriedades típicas da função que temos em mente. É esse processo que nos permite aplicar satisfatoriamente os conceitos e métodos aprendidos.

Uma vez constatada que a função a ser empregada em determinada situação é a afim e, portanto, que a lei de correspondência em questão é do tipo $f(x) = ax + b$, obtém-se b como o valor que a função f assume quando $x = 0$, i.e., $b = f(0)$. Por essa razão, o coeficiente b é por vezes chamado de *valor inicial* da função f.

O valor do coeficiente a, por sua vez, pode ser obtido desde que se conheçam os valores $f(x_1)$ e $f(x_2)$ que a função f assume em dois pontos distintos (porém arbitrários) x_1 e x_2. De fato, uma vez que:

$$f(x_1) = ax_1 + b \quad \text{e} \quad f(x_2) = ax_2 + b,$$

segue-se que:

$$f(x_2) - f(x_1) = a(x_2 - x_1),$$

e portanto:

$$a = \frac{f(x_2) - f(x_1)}{x_2 - x_1}.$$

que é a taxa de variação média de f no intervalo de extremos x_1 e x_2.

Naturalmente, uma função afim é crescente quando $a > 0$, decrescente quando $a < 0$ e constante quando $a = 0$. Veja: no caso em que $a = (f(x_2) - f(x_1))/(x_2 - x_1) > 0$, se $x_1 < x_2$, então $f(x_2) - f(x_1) > 0$ e, portanto, $f(x_1) < f(x_2)$, com x_1 e x_2 arbitrários. Analogamente no caso em que $a < 0$ e trivialmente $f(x_1) = f(x_2)$ para x_1 e x_2 arbitrários quando $a = 0$, justificando nossa assertiva.

2.3.1 Caracterização da função afim

Como saber se, numa determinada situação, o modelo matemático a ser adotado é uma função afim? Precisamos do seguite teorema.

Teorema 2.21 (Teorema de Caracterização da Função Afim). *Seja $f : \mathbb{R} \to \mathbb{R}$ uma função estritamente monótona. Se o acréscimo $f(x+h) - f(x) = \varphi(h)$ depender apenas de h*[a]*, mas não de x, então f é uma função afim.*

[a] A hipótese de que $f(x+h) - f(x)$ não depende de x às vezes se exprime dizendo que "a acréscimos iguais de x correspondem acréscimos iguais de $f(x)$" ou que "os acréscimos sofridos por $f(x)$ são proporcionais aos acréscimos dados a x".

Demonstração. Sem perda de generalidade, suponha que f seja crescente. Então $\varphi : \mathbb{R} \to \mathbb{R}$ também o é, com $\varphi(0) = 0$. Além disso, para quaisquer h, $k \in \mathbb{R}$,

$$\begin{aligned}\varphi(h+k) &= f(x+h+k) - f(x)\\ &= f((x+k)+h) - f(x+k) + f(x+k) - f(x)\\ &= \varphi(h) + \varphi(k).\end{aligned}$$

Logo, pelo Teorema Fundamental da Proporcionalidade, pondo-se $a = \varphi(1)$, tem-se $\varphi(h) = ah$, $\forall h \in \mathbb{R}$. Assim, $f(x+h) - f(x) = ah$; pondo-se $b = f(0)$, segue-se que $f(h) = ah + b$, ou seja, $f(x) = ax + b$ para todo $x \in \mathbb{R}$. ■

A recíproca desse teorema é óbvia: se $f(x) = ax + b$, então $f(x+h) - f(x) = a(x+h) + b - (ax+b) = ah$ não depende de x.

O teorema acima nos permite comparar as características de um determinado problema com as propriedades típicas da função afim e por meio dessa comparação nos permite decidir se ela é o modelo apropriado para o nosso problema.

Exemplo 2.22 ([9]). As escalas termométricas mais familiares baseiam-se na altura de uma coluna de mercúrio, a qual aumenta ou diminui conforme a temperatura sobe ou desce. Na escala Celsius, o valor 0 corresponde à temperatura de fusão do gelo e o valor 100 corresponde à temperatura de ebulição da água (ambas à pressão do nível do mar). Na escala Fahrenheit, esses valores são 32 e 212 respectivamente. Os demais valores na escala Celsius são marcados dividindo-se o intervalo entre aquelas duas temperaturas em 100 partes de igual comprimento e, na escala Fahrenheit, em 180 partes também de comprimentos iguais. Usando-se esses comprimentos em cada caso, as escalas são estendidas para assinalarem valores de temperaturas superiores à da ebulição da água e inferiores à da fusão do gelo. Pergunta-se: (a) em que temperatura as duas escalas assinalam o mesmo valor e (b) qual a temperatura Celsius igual à metade do valor correspondente em Fahrenheit? ◇

Solução. Em última análise, os graus Celsius (C) e Fahreinheit (F) são diferentes unidades de comprimento com as quais se mede a temperatura de uma coluna de mercúrio. Assim, a mudança de escala, de C para F, é uma função $f : \mathbb{R} \to \mathbb{R}$ que associa à medida x, segundo C, a medida $f(x)$, segundo F, da mesma coluna de mercúrio. Evidentemente, f é crescente. Além disso, a diferença $f(x+h) - f(x)$ é a medida, segundo F, do segmento de reta de extremos $f(x)$ e $f(x+h)$, o qual, segundo C, tem extremos x e $x+h$, logo seu C-comprimento é igual a h. Ora, a medida deste segmento depende apenas de h mas não de x e o mesmo se dá com a diferença $f(x+h) - f(x)$. Pelo Teorema de Caracterização da Função Afim, concluímos que f é uma função afim: $f(x) = ax + b$. Sabemos que $f(0) = 32$ e $f(100) = 212$. Então, $b = 32$ e $100a + 32 = 212$, donde $a = 1,8$. Portanto, $f(x) = 1,8x + 32$ é a fórmula que permite passar da temperatura x na escala Celsius para a temperatura $f(x)$ na escala Fahrenheit. Assim:

(a) para que tenhamos $f(x) = x$, deve-se ter $1,8x + 32 = x$, donde $x = -40$, ou seja, -40 graus Celsius é o mesmo que -40 graus Fahrenheit; e

(b) para que tenhamos $f(x) = 2x$, deve-se ter $1,8x + 32 = 2x$, donde $x = 160$, ou seja, 160 graus Celsius equivalem a 320 graus Fahrenheit. ◆

Exemplo 2.23 ([10]). Suponha que um ponto se movimenta sobre um eixo. Sua posição, em cada instante t, é determinada pela coordenada (abscissa) $x(t)$. Dizemos que esse movimento é *uniforme* quando o ponto se desloca sempre no mesmo sentido, i.e., $x : \mathbb{R} \to \mathbb{R}$ é uma função estritamente monótona, e, além disso, em intervalos de tempo iguais percorre distâncias iguais. Isso significa que $x(t+h) - x(t)$, distância percorrida no intervalo de tempo de duração h, a partir da posição $x(t)$, depende apenas de h, mas não de t. Então, x é uma função afim: $x(t) = vt + x_0$, em que $v = x(t+1) - x(t)$, distância percorrida na unidade de tempo, chama-se a *velocidade* e $x_0 = x(0)$ é a posição inicial. $\diamond$

Observação 2.24. Na definição usual de movimento uniforme, não se impõe a condição de que o ponto móvel se desloque sempre no mesmo sentido, pois se supõe que a função $x(t)$ seja contínua. Como já chamamos a atenção na Observação 2.4 sobre o Teorema Fundamental da Proporcionalidade, a condição de monotonicidade pode ser substituída pela condição de continuidade, sem alterar a conclusão. Ao menos uma delas ou alguma forma equivalente é imprescindível, pois existem funções incrivelmente complicadas para as quais vale $f(x+y) = f(x) + f(y)$ para todo $x, y \in \mathbb{R}$ mas f não é da forma $f(x) = ax$. $\circ$

2.4 Gráfico da função afim

Proposição 2.25. *O gráfico de uma função afim $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto ax + b$, é uma reta não vertical.*

Demonstração. Para ver que o gráfico de f é uma reta, basta mostrar que três pontos quaisquer do gráfico de f, $P_i = (x_i, y_i)$, em que $y_i = ax_i + b$, $i = 1, 2, 3$, são colineares. Para tanto, é necessário e suficiente que a maior das distâncias

$$d(P_i, P_j)_{i \neq j} = \sqrt{(x_j - x_i)^2 + a^2(x_j - x_i)^2} = (x_j - x_i)\sqrt{1 + a^2},$$

$i, j \in \{1, 2, 3\}$, seja igual à soma das outras duas.

Sem perda de generalidade, suponha que os pontos foram escolhidos de modo que suas abscissas satisfazem: $x_1 < x_2 < x_3$. É fácil então ver que $d(P_1, P_3) = d(P_1, P_2) + d(P_1, P_2)$.

Para ver que o gráfico de f não pode ser uma reta vertical, basta mostrar que $x_1 = x_2 \Leftrightarrow f(x_1) = f(x_2)$ quando $a \neq 0$, o que é imediato, pois, supondo $a \neq 0$, $x_1 = x_2 \Leftrightarrow ax_1 = ax_2 \Leftrightarrow ax_1 + b = ax_2 + b \Leftrightarrow f(x_1) = f(x_2)$. E quando $a = 0$, basta notar que se tem uma função constante, e que portanto o gráfico de f é uma reta horizontal. $\blacksquare$

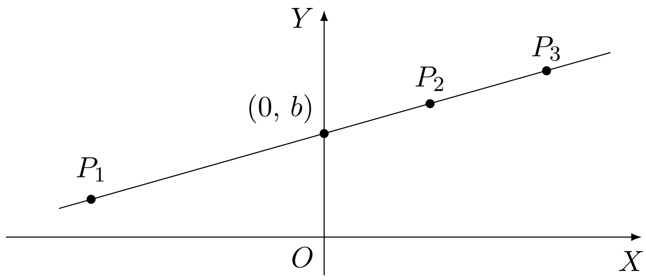


Figura 2.4: Gráfico da função afim.

A assertiva de que o gráfico de f não é uma reta vertical é de certo modo intuitiva. Se o gráfico de f fosse uma reta vertical, violaria a própria definição de função.

Corolário 2.26. *Uma função afim f fica completamente determinada se conhecermos os valores $f(x_1)$ e $f(x_2)$ que f assume em dois números distintos e arbitrários x_1 e x_2.*

Demonstração. Como uma reta fica completamente determinada por dois de seus pontos e o gráfico de uma função afim $f : \mathbb{R} \to \mathbb{R}$ é uma reta, segue-se a tese. ■

Em outras palavras, o corolário acima nos diz que existe uma, e somente uma, função afim $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto ax + b$, tal que $f(x_i) = y_i$, $i = 1, 2$. Podemos também provar isso construtivamente, resolvendo o sistema $ax_i + b = y_i$, $i = 1, 2$, para a e b, cuja (única e imediata) solução é:

$$a = \frac{y_2 - y_1}{x_2 - x_1}, \quad b = \frac{x_2 y_1 - x_1 y_2}{x_2 - x_1}.$$

Essa é uma particularidade da função afim, ou seja, enquanto precisamos de uma regra que permita, ao menos teoricamente, determinar o valor $f(x)$ para todo $x \in X$ a fim de se conhecer uma função $f : X \to Y$, no caso da função afim bastam dois pontos para que ela fique inteiramente determinada.

A Proposição 2.25 possui recíproca.

Proposição 2.27. *Toda reta não vertical é o gráfico de uma função afim.*

Demonstração. Seja uma reta não vertical r e sejam $P_i = (x_i, y_i)$, $i = 1, 2$, dois pontos distintos de r. Como r não é vertical, necessariamente $x_1 \neq x_2$. Então, de acordo com o Corolário 2.26, existe uma, e somente uma, função afim $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto ax + b$, tal que $f(x_i) = y_i$, $i = 1, 2$. Isso significa que o gráfico de f passa pelos pontos P_i, $i = 1, 2$. Como o gráfico de f é uma reta (Proposição 2.25), forçosamente ela coincide com r e, portanto, r é o gráfico de f. Sendo r uma reta não vertical arbitrária, segue-se a tese. ■

Juntas, as duas últimas proposições nos mostram que existe uma correspondência biunívoca entre o conjunto de todas as funções afins e o conjunto de todas as retas não verticais do plano. Esses conjuntos se equivalem. A cada função afim corresponde uma única reta não vertical do plano e a cada reta não vertical do plano corresponde uma única função afim.

Geometricamente, podemos interpretar o coeficiente b como a ordenada do ponto onde a reta, que é o gráfico de f, intersecta o eixo OY (Figura 2.4). O coeficiente a, por sua vez, é dito a *inclinação* ou *coeficiente angular* dessa reta em relação ao eixo OX. Quanto maior o valor absoluto de a, mais a reta se afasta do eixo horizontal. Além disso, se $a > 0$, a reta é ascendente; se $a < 0$, ela é descendente; e se $a = 0$, ela é horizontal. Ademais, diz-se que $y = ax + b$ é a *equação* da reta r que é o gráfico de f.

É importante ressaltar que os nomes *coeficiente angular* e *inclinação* são apropriados para o parâmetro a da equação de uma reta, $y = ax + b$, mas não da lei de formação de uma função afim, $f(x) = ax + b$. Na maioria dos problemas modelados pela função afim, sequer há um ângulo envolvido. E mesmo considerando o gráfico de f, o ângulo que ele faz com o eixo OX depende das unidades escolhidas para medir x e $f(x)$. Quando falamos da lei de formação, o nome adequado para a é *taxa de variação* ou *taxa de crescimento*.

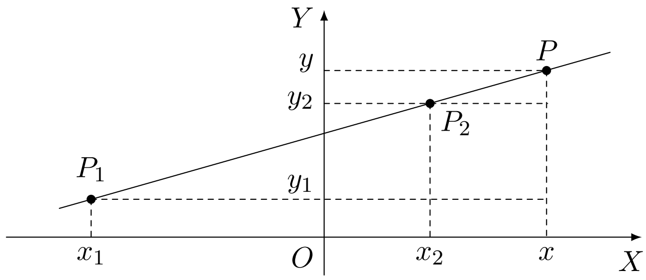


Figura 2.5: Equação da reta que é o gráfico de uma função afim.

No caso de uma função afim f, o coeficiente

$$a = \frac{y_2 - y_1}{x_2 - x_1}$$

tem claramente o significado de taxa de crescimento e a interpretação geométrica acima permite concluir que a equação da reta, gráfico de f, que passa pelos pontos (x_i, y_i), $i = 1, 2$, não situados na mesma vertical, ou seja, com $x_1 \neq x_2$, é:

$$y = y_i + \frac{y_2 - y_1}{x_2 - x_1}(x - x_i),\ i = 1,\ 2,$$

que tem o seguinte significado: partindo-se do ponto (x_i, y_i) e incrementando-se x de $x - x_i$, y sofre um incremento, a partir de y_i, igual ao incremento de x vezes a taxa de variação a, $i = 1, 2$. As equações para i igual a 1 e 2 são iguais. Sendo assim, podemos dizer, de modo geral, que a equação da reta que passa por um ponto (x_0, y_0) e tem inclinação a é:

$$y = y_0 + a(x - x_0).$$

2.5 Funções afins e progressões aritméticas

Há uma conexão interessante entre funções afins e progressões aritméticas.

Progressões aritméticas são sequências nas quais o incremento (ou decremento) entre termos sucessivos é sempre o mesmo. Noutras palavras, uma *progressão aritmética* (P.A.) é uma sequência $(x_1, x_2, \ldots, x_i, \ldots)$, finita ou infinita, na qual a diferença entre termos sucessivos é uma constante k chamada de *razão*:

$$k = x_2 - x_1 = x_3 - x_2 = \cdots = x_{i+1} - x_i = \cdots.$$

Além da notação usual $(x_1, x_2, \ldots, x_i, \ldots)$, também representamos uma P.A. abreviadamente por $(x_n)_{n \in \mathbb{N}}$ ou simplesmente (x_n). Note: uma P.A. nada mais é do que uma função que faz corresponder a cada $n \in \mathbb{N}$ o número real x_n, i.e., $1 \mapsto x_1$, $2 \mapsto x_2$, $\ldots$, $n \mapsto x_n$, $\ldots$

Quando $k \neq 0$, dizemos que a P.A. é *não estacionária* ou simplesmente que ela não é constante.

Geometricamente, podemos interpretar uma P.A. não estacionária como uma sequência de pontos igualmente espaçados na reta:

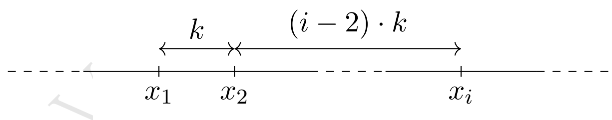


Figura 2.6: Interpretação geométrica de uma progressão aritmética.

Afirmamos o seguinte:

Proposição 2.28. *Seja $f : \mathbb{R} \to \mathbb{R}$ uma função monótona injetiva. Uma condição necessária e suficiente para que f seja uma função afim é que f transforme uma P.A. em outra.*

Demonstração. Que essa é uma condição necessária é um fato trivial: se (x_i) é uma P.A. de razão k e f é da forma $f(x) = ax + b$, então $(y_i = f(x_i))$ também é uma P.A., pois

$$y_{i+1} - y_i = (ax_{i+1} + b) - (ax_i + b) = a(x_{i+i} - x_i) = ak.$$

Para provar que também é uma condição suficiente, vamos considerar uma função auxiliar $g : \mathbb{R} \to \mathbb{R}$, definida como $g(x) := f(x) - f(0)$ para todo $x \in \mathbb{R}$. Como f transforma uma P.A. em outra P.A., a função g também o faz, além de ser ímpar, pois $g(0) = f(0) - f(0) = 0$ e, para todo $x \in \mathbb{R}$, como $-x$, 0, x forma uma P.A., o mesmo se dá com suas imagens por g, i.e., $g(-x)$, 0, $g(x)$, logo $g(-x) = -g(x)$. Vamos mostrar que g é linear.

Sejam $n \in \mathbb{N}$ e $x \in \mathbb{R}$. Como os números 0, x, $2x$, ..., nx formam uma P.A., suas imagens por g também o fazem: 0, $g(x)$, $g(2x)$, ..., $g(nx)$. A diferença entre os dois primeiros termos nos dá a razão desta P.A., a saber, $g(x)$. Assim, $g(nx) = ng(x)$. Se n for um inteiro negativo, $-n \in \mathbb{N}$ e portanto $g(nx) = -g(-nx) = -(-ng(x)) = ng(x)$. (Foi utilizada a paridade de g na primeira igualdade.) A conclusão é que $g(nx) = ng(x)$ para todo $n \in \mathbb{Z}$ e todo $x \in \mathbb{R}$, logo, pelo Teorema Fundamental da Proporcionalidade (Teorema 2.3), segue-se que g é linear, i.e., existe $a \in \mathbb{R}$ tal que $g(x) = ax$.

Agora, pondo-se $f(0) = b$, tem-se $f(x) = g(x) + f(0) = ax + b$ para todo $x \in \mathbb{R}$, o que conclui a demonstração. ∎

Para finalizar esta seção, vamos provar que dada uma P.A. não estacionária $(a_i)_{i\in\mathbb{N}}$ e outra P.A. qualquer $(b_i)_{i\in\mathbb{N}}$ sempre existe uma e única função afim f tal que os termos desta são as imagens por f dos termos daquela.

Proposição 2.29. *Dadas a P.A. não estacionária $(a_i)_{i\in\mathbb{N}}$ e a P.A. $(b_i)_{i\in\mathbb{N}}$ arbitrária, existe uma, e somente uma, função afim $f : \mathbb{R} \to \mathbb{R}$ tal que $f(a_i) = b_i$ para todo $i \in \mathbb{N}$.*

Demonstração. Essa afirmação é consequência imediata do Corolário 2.26. Vejamos. Seja $k_a \neq 0$ a razão da primeira P.A, k_b a razão da segunda e $f : \mathbb{R} \to \mathbb{R}$ a função afim de lei de correspondência $x \mapsto ax + b$ para algum a, $b \in \mathbb{R}$, tal que $f(a_1) = b_1$ e $f(a_2) = b_2$. Temos assim o seguinte sistema linear nas incógnitas a e b:

$$\begin{cases} a_1 a + b = b_1 \\ a_2 a + b = b_2 \end{cases}.$$

Como $a_2 - a_1 = k_a \neq 0$, esse sistema possui solução e essa solução é única, a saber,

$$a = \frac{b_2 - b_1}{a_2 - a_1} = \frac{k_b}{k_a}, \quad b = \frac{a_2 b_1 - a_1 b_1 + a_1 b_1 - a_1 b_2}{a_2 - a_1} = b_1 - \frac{k_b}{k_a} a_1,$$

e portanto

$$f(x) = \frac{k_b}{k_a} x + \left(b_1 - \frac{k_b}{k_a} a_1 \right),$$

o que conclui a demonstração. ∎

2.6 Funções poligonais

Há várias situações nas quais nos deparamos não com uma função afim pura e simplesmente como temos visto até agora, mas sim com funções tais que é possível particionar os seus domínios em intervalos de forma que, restritas a cada um desses intervalos, elas coincidem com uma função afim. Trata-se das *funções poligonais.* Mais precisamente:

Definição 2.30. *Dizemos que* $f : \mathbb{R} \to \mathbb{R}$ *é uma* função poligonal *quando existem* $x_0 < x_1 < \cdots < x_n$*, chamados de* pontos singulares *ou* singularidades de f*, tais que, em cada um dos intervalos* $I_0 = (-\infty,\, x_0]$, $I_1 = [x_0,\, x_1]$, $\ldots$, $I_n = [x_{n-1},\, x_n]$, $I_{n+1} = [x_n,\, +\infty)$*, a função* f *coincide com uma função afim* $f_i : I_i \to \mathbb{R}$ *e, além disso, para evitar descontinuidades,* $f_{i+1}(x_i) = f_i(x_i)$ *para todo* $i = 0,\, 1,\, \ldots,\, n$.

Evidentemente, o gráfico de uma função poligonal é uma linha poligonal (Figura 2.7).

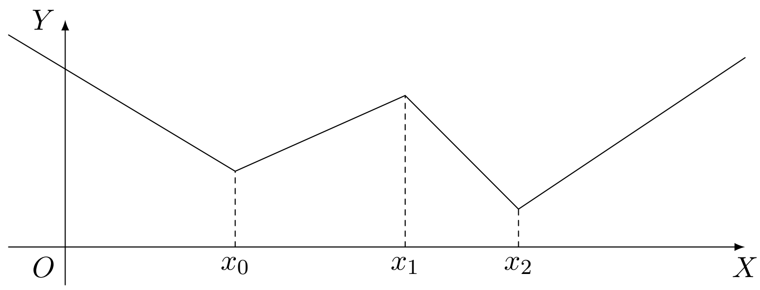


Figura 2.7: Gráfico de uma função poligonal.

Exemplo 2.31. O imposto de renda y pago por uma pessoa que, durante o ano-calendário de 2015 a partir do mês de abril, teve uma renda líquida mensal x é dada por uma expressão da forma $y = ax - b$, em que a alíquota a e a parcela a deduzir b dependem da renda x e são dadas pela tabela a seguir.

Base de cálculo (R$)	Alíquota (%)	Parcela a deduzir (R$)
Até $1.903,98$	–	–
De $1.903,99$ a $2.826,65$	$7,5$	$142,80$
De $2.826,66$ a $3.751,05$	15	$354,80$
De $3.751,06$ a $4.664,68$	$22,5$	$636,13$
Acima de $4.664,68$	$27,5$	$869,36$

Esse é um exemplo de função poligonal e o seu gráfico tem o aspecto da Figura 2.8, em que $x_0 = 1.903,99$, $x_1 = 2.826,66$, $x_2 = 3.751,06$ e $x_3 = 4.664,69$ são os pontos em que ocorrem as mudanças das faixas de renda de acordo com a tabela acima e $y_0 = 0,00$, $y_1 = 69,20$, $y_2 = 207,86$ e $y_3 = 413,43$ são os valores correspondentes dos impostos.

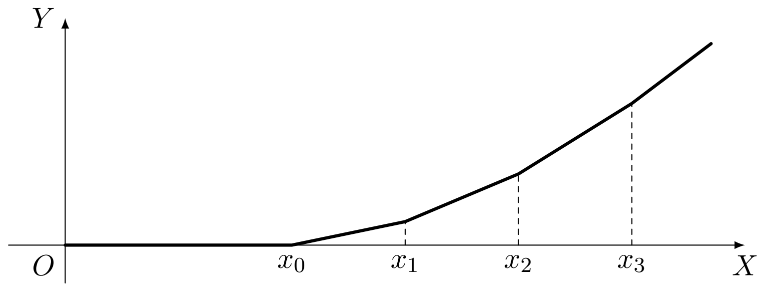


Figura 2.8: Esboço do gráfico de imposto devido em função da renda mensal.

Restrito a cada um dos intervalos $[0,\, x_0]$, $[x_0,\, x_1]$, $[x_1,\, x_2]$, $[x_2,\, x_3]$ e $[x_3,\, +\infty]$, o gráfico dessa função se reduz ao gráfico de uma função afim. ◇

Exemplo 2.32. O Certificado de Operações Estruturadas (COE) é um tipo de investimento que combina elementos de renda fixa e variável, com retorno atrelado a ativos e índices, como câmbio, inflaxão, ações, ativos internacionais, etc. É a versão brasileira das Notas Estruturadas, populares na Europa e nos Estados Unidos. Numa das possíveis estruturas, o investidor ganha proporcionalmente com a alta ou com a baixa do ativo de referência, sendo que os ganhos são limitados e, caso a cotação do ativo termine entre dois valores denominados *strikes*, o investidor recebe o percentual do capital protegido. A ideia dessa estrutura encontra-se esquematizada na Figura 2.9. (Na figura, l_i := *limitador* i e s_i := *strike* i, para $i = 1,\, 2$.)

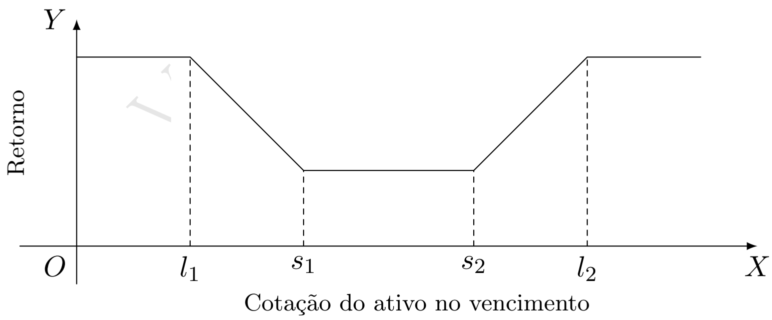
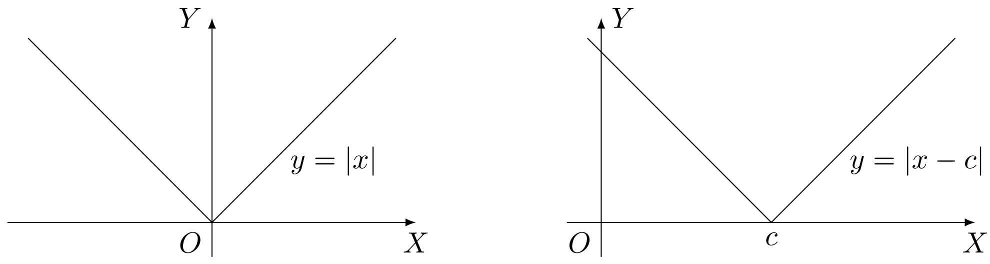


Figura 2.9: Exemplo de COE modelado por função poligonal.

Note que a função subjacente a essa operação de investimento é uma função poligonal. ◇

Alguns protótipos de função poligonal são a função módulo, $f : \mathbb{R} \to \mathbb{R}$, $f(x) = |x|$, ou alguma de suas variantes: $g : \mathbb{R} \to \mathbb{R}$, $g(x) = |x - c|$ para algum $c \in \mathbb{R}$; $h : \mathbb{R} \to \mathbb{R}$, $h(x) = |\alpha x + \beta|$ ou $h(x) = |x - \alpha| + |x - \beta|$, para algum $\alpha,\, \beta \in \mathbb{R}$.

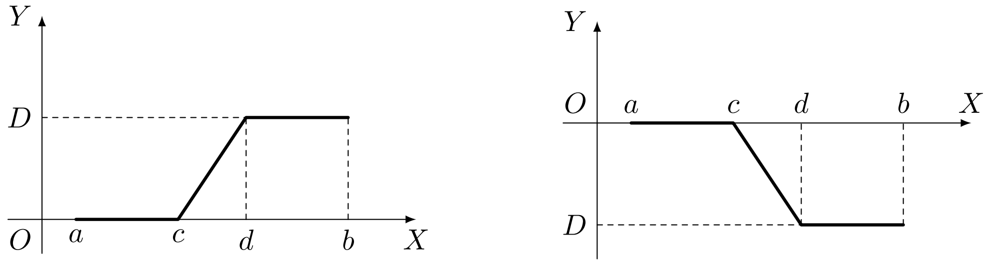


(a) Função módulo. (b) Translado da função módulo.

Figura 2.10: Protótipos de funções poligonais.

Isso nos leva a conjecturar que toda função poligonal pode ser definida combinando-se valores absolutos de funções afins. Essa suposição é verdadeira, como provaremos a seguir. Antes, porém, precisamos estudar (brevemente) um tipo especial de função poligonal, chamada *função-rampa*.

Considere os gráficos ilustrados nas Figuras 2.11a e 2.11b.



(a) (b)

Figura 2.11: Função-rampa.

A função φ representada por um qualquer desses gráficos possui dois patamares, um no intervalo $[a,\, c]$ e outro no intervalo $[d,\, b]$, nos quais assume respectivamente os valores 0 e D, conectados por uma rampa de inclinação

$$\alpha = \frac{\varphi(d) - \varphi(c)}{d - c} = \frac{D - 0}{d - c} = \frac{D}{d - c}.$$

Claro que $\varphi(x) = 0$ se $x \in [a,\, c]$, $\varphi(x) = D$ se $x \in [d,\, b]$ e se $x \in [c,\, d]$ vale

$$\alpha = \frac{\varphi(x) - \varphi(c)}{x - c} = \frac{\varphi(x) - 0}{x - c} = \frac{\varphi(x)}{x - c},$$

logo $\varphi(x) = \alpha(x - c)$ se $x \in [c,\, d]$. Podemos condensar essas informações na seguinte:

Definição 2.33. *Considere um intervalo não degenerado* $[a, b]$*, um número real* $D \neq 0$ *e* $c, d \in [a, b]$ *tais que* $c < d$*. Chamamos de* função-rampa *uma função poligonal* $\varphi : [a, b] \to \mathbb{R}$ *assim definida:*

$$\varphi(x) := \begin{cases} 0, & \text{se } a \leqslant x \leqslant c \\ \alpha(x-c), & \text{se } c \leqslant x \leqslant d \\ D, & \text{se } d \leqslant x \leqslant b \end{cases},$$

em que $\alpha = D/(d-c)$ *é a inclinação da rampa.*

Observação 2.34. Para sermos mais abrangentes, consideraremos as funções constantes como um tipo especial de função-rampa, em que a inclinação α da rampa vale zero, e diremos que se trata de uma função-rampa degenerada. ○

Um fato interessante sobre a função-rampa está contido no seguinte:

Proposição 2.35. *Nos termos da Definição 2.33, podemos expressar uma função-rampa na forma:*

$$\varphi(x) = \frac{\alpha}{2}[(d-c) + |x-c| - |x-d|].$$

Demonstração. Para ver que isso é verdade, basta calcular os valores de φ em cada um dos intervalos $[a, c]$, $[c, d]$ e $[d, b]$:

- se $x \in [a, c]$, então $\varphi(x) = \frac{\alpha}{2}[(d-c) - (x-c) + (x-d)] = 0$;
- se $x \in [c, d]$, então $\varphi(x) = \frac{\alpha}{2}[(d-c) + (x-c) + (x-d)] = \alpha(x-c)$;
- se $x \in [d, b]$, então $\varphi(x) = \frac{\alpha}{2}[(d-c) + (x-c) - (x-d)] = \alpha(d-c) = D$. ■

Proposição 2.36. *Uma função poligonal definida num intervalo* $[a, b]$ *pode ser expressa como uma soma de um número finito de funções-rampa.*

Demonstração. Seja uma função poligonal $f : [a, b] \to \mathbb{R}$ e sejam $a = x_0 < x_1 < \cdots < x_n = b$ as singularidades de f. Por definição, f é afim quando restrita a cada um dos intervalos $X_i = [x_{i-1}, x_i]$, i.e.:

$$(f|X_i)(x) = \alpha_i(x - x_{i-1}) + f(x_{i-1}),$$

em que

$$\alpha_i = \frac{f(x_i) - f(x_{i-1})}{x_i - x_{i-1}}, \tag{2.6.1}$$

para $i = 1, 2, \ldots, n$.

Considere as funções-rampa $\varphi_i : [a,\, b] \to \mathbb{R}$, $\varphi_0(x) := f(a)$ e, para $i = 1,\, 2,\, \ldots,\, n$,

$$\varphi_i(x) := \begin{cases} 0, & \text{se } a \leqslant x \leqslant x_{i-1} \\ \alpha_i(x - x_{i-1}), & \text{se } x_{i-1} \leqslant x \leqslant x_i \\ f(x_i) - f(x_{i-1}), & \text{se } x_i \leqslant x \leqslant b \end{cases},$$

em que α_i é como em (2.6.1). Considere também a função $\sigma : [a,\, b] \to \mathbb{R}$, assim definida:

$$\sigma := \varphi_0 + \varphi_1 + \cdots + \varphi_n.$$

No intervalo $[x_{i-1},\, x_i]$,

$$\begin{aligned} \sigma(x) &= \varphi_0(x) + \varphi_1(x) + \cdots + \varphi_{i-1}(x) + \varphi_i(x) + \varphi_{i+1}(x) + \cdots + \varphi_n(x) \\ &= f(a) + [f(x_1) - f(a)] + [f(x_2) - f(x_1)] + \cdots + \\ &+ [f(x_{i-1}) - f(x_{i-2})] + \alpha_i(x - x_{i-1}) + 0 + \cdots + 0 \\ &= f(x_{i-1}) + \alpha_i(x - x_{i-1}) \\ &\overset{(\textbf{??})}{=} (f|X_i)(x), \end{aligned}$$

para $i = 1,\, 2,\, \ldots,\, n$. Isso significa que $\sigma = f|X_i$ em cada intervalo $[x_{i-1},\, x_i]$, $i = 1,\, 2,\, \ldots,\, n$, logo $\sigma = f$ em $[a,\, b]$.

Provamos assim que:

$$f = \sum_{i=0}^{n} \varphi_i,$$

ou seja, que uma função poligonal definida em um intervalo $[a,\, b]$ pode ser expressa como uma soma de um número finito de funções-rampa. ∎

Como consequência imediata das duas últimas proposições, temos o seguinte:

Corolário 2.37. *Toda função poligonal $f : [a,\, b] \to \mathbb{R}$ pode ser escrita na forma:*

$$f(x) = c + c_0|x - x_0| + c_1|x - x_1| + \cdots + c_n|x - x_n|,$$

para todo $x \in [a,\, b]$, em que $a = x_0,\, x_1,\, x_2,\, \ldots,\, x_n = b$ são as singularidades de f.

Demonstração. Das Proposições 2.35 e 2.36, temos:

$$\begin{aligned} f(x) &= \sum_{i=0}^{n} \varphi_i(x) = \varphi_0(x) + \sum_{i=1}^{n} \frac{\alpha_i}{2}[(x_i - x_{i-1}) + |x - x_{i-1}| - |x - x_i|] \\ &= f(x_0) + \sum_{i=1}^{n} \frac{f(x_i) - f(x_{i-1})}{2} + \sum_{i=1}^{n} \frac{\alpha_i}{2}|x - x_{i-1}| + \sum_{i=1}^{n} \frac{\alpha_i}{2}|x - x_i| \\ &= \frac{f(x_n) - f(x_0)}{2} + \frac{\alpha_1}{2}|x - x_0| + \sum_{i=1}^{n-1} \frac{\alpha_{i+1}}{2}|x - x_i| - \sum_{i=1}^{n-1} \frac{\alpha_i}{2}|x - x_i| - \frac{\alpha_n}{2}|x - x_n| \\ &= \frac{f(x_n) - f(x_0)}{2} + \frac{\alpha_1}{2}|x - x_0| + \sum_{i=1}^{n-1} \frac{\alpha_{i+1} - \alpha_i}{2}|x - x_i| + \left(-\frac{\alpha_n}{2}\right)|x - x_n|. \end{aligned}$$

Fazendo $c = [f(x_n) - f(x_0)]/2$, $c_0 = \alpha_1/2$, $c_i = (\alpha_{i+1} - \alpha_i)/2$ ($i = 1, 2, \ldots, n-1$) e $c_n = -\alpha_n/2$, resulta a tese. ■

Exemplo 2.38. Considere a função poligonal f do Exemplo 2.31, desta vez definida no intervalo $[x_0, x_3]$. (Figura 2.12.) Neste caso, $\varphi_0 \equiv 0$ e as funções-rampa φ_1, φ_2,

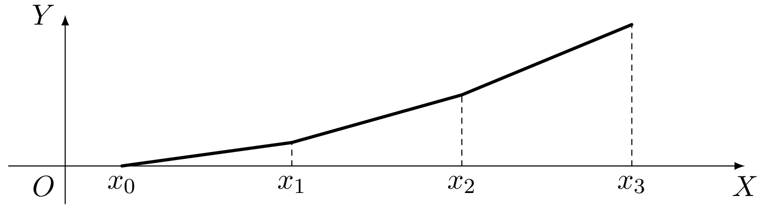


Figura 2.12: Gráfico de imposto de renda na faixa de x_0 a x_3 reais.

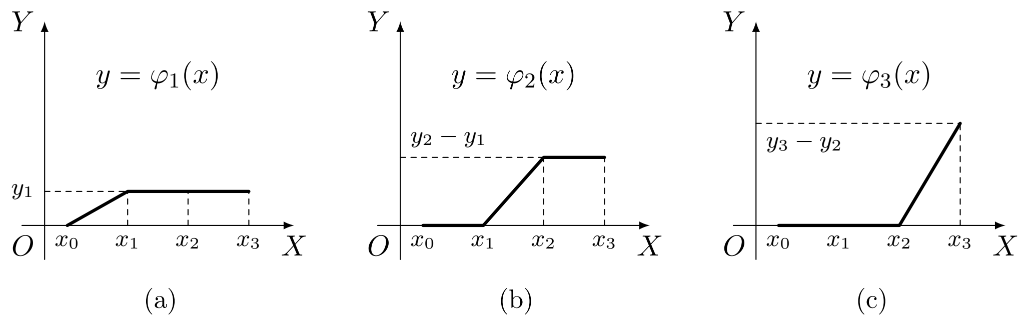


Figura 2.13: Funções-rampa "componentes" do gráfico da Figura 2.12.

φ_3, cujos gráficos estão representados na Figura 2.13, são dadas por:

$$\varphi_1(x) = 0{,}0375 \cdot (922{,}67 + |x - 1.903{,}99| - |x - 2.826{,}66|),$$
$$\varphi_2(x) = 0{,}0750 \cdot (924{,}40 + |x - 2.826{,}66| - |x - 3.751{,}06|),$$
$$\varphi_3(x) = 0{,}1125 \cdot (913{,}63 + |x - 3.751{,}06| - |x - 4.664{,}69|).$$

Decerto, pela Proposição 2.36, $f = \varphi_1 + \varphi_2 + \varphi_3$, ou seja:

$$\begin{aligned} f(x) &= \varphi_1(x) + \varphi_2(x) + \varphi_3(x) \\ &= 0{,}0375 \cdot (5.512{,}36 + |x - 1.903{,}99| + |x - 2.826{,}66| + \\ &+ |x - 3.751{,}06| - 3|x - 4.664{,}69|). \end{aligned}$$

(Compare os coeficientes aqui obtidos com os previstos pelo Corolário 2.37.) ◇

3 Funções quadráticas

"Um restaurante a quilo vende 100 kg/dia de comida a R$ 12,00/kg. Uma pesquisa de opinião revelou que, a cada real de aumento no preço, o restaurante perderia 10 clientes, com o consumo médio de 500 g cada um. Qual deve ser o preço do quilo de comida para que o restaurante tenha a maior receita possível?"

Epígrafe de abertura do Capítulo 2 de [9] para ilustrar uma das aplicações das funções quadráticas.

3.1 Introdução

O estudo das funções quadráticas tem possivelmente como uma de suas maiores motivações a resolução de equações do segundo grau, i.e., equações do tipo $ax^2+bx+c = 0$, em que a, b e c são números reais dados, com $a \neq 0$, e deseja-se saber quais valores de x, caso exista algum, tornam essa equação verdadeira. Os "x" que a tornam verdadeira chamam-se as *raízes* dessa equação.

Problemas que recaem nesse tipo de equação estão entre os mais antigos da Matemática. Por exemplo, como mencionado em [10], em textos cuneiformes, escritos pelos babilônios há quase quatro mil anos, pode-se encontrar, por exemplo, a questão de achar dois números conhecendo sua soma s e seu produto p, que, em termos geométricos, equivale a procurar as medidas dos lados de um retângulo conhecendo-se o semiperímetro s e a área p.

Evidentemente, se um dos número é x, o outro é $s-x$ e seu produto é $p = x(s-x) = sx - x^2$, logo o problema se reduz a encontrar a(s) solução(ões) da equação

$$x^2 - sx + p = 0.$$

Certamente, se α é uma raiz dessa equação, i.e., $\alpha^2 - s\alpha + p = 0$, então $\beta = s - \alpha$

também é raiz, pois

$$\begin{aligned}\beta^2 - s\beta + p &= (s-\alpha)^2 - s(s-\alpha) + p\\ &= s^2 - 2s\alpha + \alpha^2 - s^2 + s\alpha + p\\ &= \alpha^2 - s\alpha + p = 0.\end{aligned}$$

Não consta registro de como os autores dos textos cuneiformes teriam chegado às raízes da equação acima, mas eles as conheciam e especula-se que tenha sido algo mais ou menos como se segue (abstraia-se a notação e preserve-se o raciocínio, pois a representação de números por letras só viria a ser introduzida por François Viète[1] no século XVI).

Sejam α e β os números procurados. Sem perda de generalidade, suponha $\alpha \leqslant \beta$. Ambos são equidistantes de sua média aritmética $s/2 = (\alpha+\beta)/2$. Seja $d = \beta - s/2 = s/2 - \alpha$ essa distância. Se encontrarmos d, obteremos os números procurados: $\alpha = s/2 - d$ e $\beta = s/2 + d$. Mas d é fácil de obter, pois

$$p = \alpha\beta = \left(\frac{s}{2} - d\right)\left(\frac{s}{2} + d\right) = \left(\frac{s}{2}\right)^2 - d^2,$$

donde segue-se que

$$d = \sqrt{\left(\frac{s}{2}\right)^2 - p}.$$

Portanto,

$$\alpha = \frac{s}{2} - d = \frac{s}{2} - \sqrt{\left(\frac{s}{2}\right)^2 - p} \quad \text{e} \quad \beta = \frac{s}{2} + d = \frac{s}{2} + \sqrt{\left(\frac{s}{2}\right)^2 - p}.$$

Ao que parece, s e p eram sempre números positivos, de maneira que soluções negativas não teriam sido motivo de preocupação para os babilônios. Mas certamente devem ter ocorrido situações em que $(s/2)^2 < p$, como no problema de encontrar dois números cuja soma e produto são ambos iguais a 2. No entanto, isso não levou à invenção dos números complexos[2]. Quando isso ocorria, os babilônios diziam que a solução não existia, o que é correto no âmbito dos números reais.

A seguir, apresentaremos a definição de função quadrática e, como quem tenta iluminar diferentes aspectos do mesmo objeto, vamos explorar algumas representações algébricas (forma canônica e fatorada), sua representação geométrica (seu gráfico, mostrando que realmente é uma parábola) e apresentar algumas aplicações clássicas (obtenção de máximos e mínimos, uso dos seus zeros para resolver problemas, trajetória de projéteis e propriedade do refletor parabólico).

[1] François Viète (1540–1603), administrador público, advogado e matemático. Considerado responsável por introduzir a primeira notação algébrica sistemática e contribuir para a teoria das equações. Em serviço a Henrique IV na guerra com a Espanha, decodificou um código complexo, utilizado pelo rei Filipe II, de mais de 500 caracteres.

[2] A introdução dos números complexos só aconteceu no século XVI com o advento da fórmula para raízes de equações de terceiro grau, que fornecia as raízes reais por meio de uma expressão que continha raízes quadradas de números negativos.

3.2 Função quadrática

Definição 3.1. *Chamamos uma função* $f : \mathbb{R} \to \mathbb{R}$ *de* quadrática *quando existem* a, b, $c \in \mathbb{R}$, *com* $a \neq 0$, *tais que* $f(x) = ax^2 + bx + c$ *para todo* $x \in \mathbb{R}$.

Diferentemente das funções afins, em que o coeficiente a pôde ser considerado nulo e as funções constantes eram tidas como casos particulares de funções afins, agora não admitiremos $a = 0$, pois as naturezas das funções quadráticas e afins são completamente diferentes e estas não podem ser consideradas casos particulares daquelas, como ficará mais claro durante o desenvolvimento deste capítulo, principalmente quando estudarmos o gráfico da função quadrática.

Além disso, os coeficientes a, b, c da função quadrática f ficam completamente determinados pelos valores que essa função assume. Noutras palavras,

$$ax^2 + bx + c = a'x^2 + b'x + c',\ \forall x \in \mathbb{R} \Leftrightarrow a = a',\ b = b',\ c = c'.$$

De fato, se o antecedente da proposição acima é verdadeiro, tomar $x = 0$ implica $c = c'$. Assim, $ax^2 + bx = a'x^2 + b'x$ para todo $x \in \mathbb{R}$, em particular para todo $x \neq 0$. Neste caso, cancelando x, obtemos $ax + b = a'x + b'$. Tomando $x = 1$ e depois $x = -1$, obtemos $a + b = a' + b'$ e $-a + b = -a' + b'$, donde $a = a'$ e $b = b'$. A recíproca é trivial.

Mas não é necessária a hipótese "para todo $x \in \mathbb{R}$". Note que, no argumento precedente, foram necessários apenas três pontos arbitrários ($x = -1$, 0 e 1) para mostrar que os coeficientes da função quadrática são únicos. Efetivamente, há uma proposição mais forte:

Proposição 3.2. *Se duas funções quadráticas assumem os mesmos valores em três pontos distintos, então essas funções são iguais.*

Demonstração. Sejam $f_i : \mathbb{R} \to \mathbb{R}$, $f_i(x) = a_i x^2 + b_i x + c_i$ ($i = 1, 2$) funções quadráticas e x_j ($j = 1, 2, 3$) os três pontos distintos nos quais elas assumem os mesmos valores, i.e., para $j = 1, 2, 3$:

$$f_1(x_j) = f_2(x_j).$$

Sejam $a = a_2 - a_1$, $b = b_2 - b_1$ e $c = c_2 - c_1$. Queremos mostrar que $a = b = c = 0$. Decerto, para $j = 1, 2, 3$:

$$f_2(x_j) - f_1(x_j) = 0,$$

ou seja,

$$ax_j^2 + bx_j + c = 0.$$

Subtraindo a equação que corresponde a $j = 1$ das outras duas, obtemos:

$$\begin{cases} a(x_2^2 - x_1^2) + b(x_2 - x_1) = 0 \\ a(x_3^2 - x_1^2) + b(x_3 - x_1) = 0 \end{cases}.$$

Como $x_k - x_j \neq 0$ quando $k \neq j$, podemos escrever:

$$\begin{cases} a(x_2 + x_1) + b = 0 \\ a(x_3 + x_1) + b = 0 \end{cases}.$$

Subtraindo a primeira da segunda equação, segue-se que $a(x_3 - x_2) = 0$ e, como $x_3 - x_2 \neq 0$, resulta $a = 0$. Substituindo a nas equações anteriores, obtemos sucessivamente $b = 0$ e $c = 0$, donde concluímos que os coeficientes de f_1 e f_2 são os mesmos e portanto $f_1 = f_2$. ∎

Noutras palavras, isso significa que uma função quadrática fica inteiramente determinada pelos valores que ela assume em três pontos distintos do seu domínio. Ainda mais do que isso: dados três pontos não colineares do plano numérico $\mathbb{R}^2$, a função quadrática fica inteiramente determinada. Esse é o conteúdo da próxima proposição.

Proposição 3.3. *Sejam $P_i = (x_i, y_i)$, $i = 1, 2, 3$, pontos não colineares em $\mathbb{R}^2$. Então existe uma, e somente uma, função quadrática $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto ax^2 + bx + c$, tal que $f(x_i) = y_i$, $i = 1, 2, 3$.*

Demonstração. Aplica-se, *mutatis mutandis*, o mesmo raciocínio da demonstração anterior. Basta considerar o sistema de equações: $ax_i^2 + bx_i + c = y_i$, $i = 1, 2, 3$, em que as incógnitas são a, b, c. Já vimos, na demonstração precedente, que o sistema homogêneo correspondente só admite a solução trivial, logo a solução do sistema não homogêneo existe e é única, o que implica a existência e unicidade de uma f cuja lei de correspondência é como acima. Resta provar que f é quadrática, ou seja, que $a \neq 0$. Seguindo o paradigma utilizado na demonstração anterior para resolução do sistema homogêneo (deixamos os detalhes como exercício), obtemos:

$$a = \frac{1}{x_3 - x_2}\left(\frac{y_3 - y_1}{x_3 - x_1} - \frac{y_2 - y_1}{x_2 - x_1}\right).$$

Evidentemente,

$$a = 0 \Leftrightarrow \frac{y_3 - y_1}{x_3 - x_1} = \frac{y_2 - y_1}{x_2 - x_1}.$$

Note que essa condição equivale à colinearidade dos P_i, $i = 1, 2, 3$, pois nos diz que as inclinações das retas P_1P_3 e P_1P_2 são iguais e, como P_1 é ponto comum a ambas, $P_1P_3 = P_1P_2$. (A Figura 3.1 ilustra o que acabou de ser dito.)

Como, por hipótese, os P_i ($i = 1, 2, 3$) não são colineares, $a \neq 0$ e portanto f é quadrática. ∎

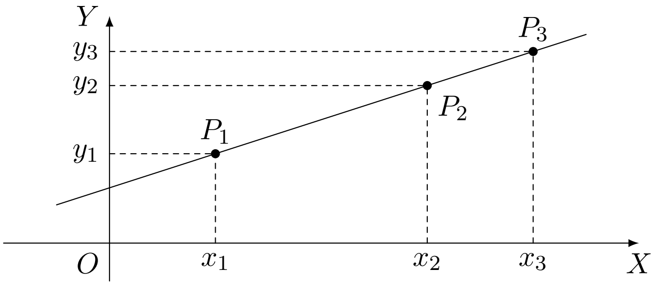


Figura 3.1: Colinearidade.

3.3 Forma canônica e suas consequências

Numa de suas famosas aulas, Feynman[3] teria dito que qualquer grande descoberta de uma nova lei só seria útil de conseguíssemos extrair mais do que introduzimos (ver [3]). Com esse espírito e guardadas as devidas proporções, vamos explorar as consequências da lei de correspondência da função quadrática.

Considere o trinômio ax^2+bx+c, com $a \neq 0$. Completando o quadrado, escrevemos:

$$\begin{aligned} ax^2+bx+c &= a\left(x^2+\frac{b}{a}x+\frac{c}{a}\right) = a\left(x^2+2\frac{b}{2a}x+\frac{b^2}{4a^2}-\frac{b^2}{4a^2}+\frac{c}{a}\right) \\ &= a\left[\left(x+\frac{b}{2a}\right)^2+\frac{4ac-b^2}{4a^2}\right]. \end{aligned}$$

Podemos, ainda, tomar $m=-b/2a$ e $k=(4ac-b^2)/4a$. Com isso, obtemos:

$$ax^2+bx+c = a(x-m)^2+k.$$

Essa é a chamada *forma canônica do trinômio do segundo grau* ax^2+bx+c. Ela não só nos permite olhar para a lei de correspondência da função quadrática de uma forma alternativa, mas também (e isto é o que nos interessa) extrair informações importantes. A primeira informação está na seguinte:

Proposição 3.4 (Ponto extremo). *Toda função quadrática possui um, e somente um, ponto extremo*[4].

[3]Richard Phillips Feynman (1918–1988), físico norte-americano, laureado com o Nobel de Física de 1965, juntamente com Julian Schwinger (1918–1994) e Shin'ichiro Tomonaga (1906–1979), pelo trabalho fundamental na eletrodinâmica quântica, com profundas consequências para a física das partículas elementares. Participou do Projeto Manhattan, de pesquisa e desenvolvimento das primeiras bombas atômicas durante a Segunda Guerra Mundial. Seu dom para explicações acessíveis fez dele um divulgador da Física de distinção aos leigos.

[4]Ponto de máximo ou de mínimo.

Demonstração. Na notação utilizada até aqui, seja $f(x) = a(x-m)^2 + k$ o valor que a função quadrática f assume em x. Se $a > 0$, como $(x-m)^2 \geqslant 0$ para todo $x \in \mathbb{R}$, então $f(x) = a(x-m)^2 + k \geqslant k$ para todo $x \in \mathbb{R}$, só valendo a igualdade para $x = m$; logo f assume um valor mínimo e o faz quando $x = m$. Analogamente, se $a < 0$, $f(x) = a(x-m)^2 + k \leqslant k$ para todo $x \in \mathbb{R}$, só valendo a igualdade para $x = m$; desta vez, f assume um valor máximo e o faz quando $x = m$. ■

Obviamente, o valor da função quadrática no ponto extremo é

$$f\left(m = -\frac{b}{2a}\right) = k = \frac{4ac - b^2}{4a}.$$

Como vimos na demonstração acima, se $a > 0$ esse é um ponto de mínimo; se $a < 0$ trata-se de um ponto de máximo.

Se por um lado a função quadrática possui ponto de mínimo (resp. máximo) quando $a > 0$ (resp. $a < 0$), por outro ela é ilimitada superiormente (resp. inferiormente).

Proposição 3.5 (Não limitações). *A função quadrática é ilimitada superiormente (resp. inferiormente) se $a > 0$ (resp. $a < 0$).*

Demonstração. Inicialmente, considere $a > 0$. Suponha, por absurdo, que a função seja limitada superiormente, i.e., que existe $L \in \mathbb{R}$ tal que $f(x) \leqslant L$ para todo $x \in \mathbb{R}$. Então, para todo $x \in \mathbb{R}$:

$$a(x-m)^2 + k \leqslant L \Leftrightarrow (x-m)^2 \leqslant \frac{L-k}{a} \Leftrightarrow m - \sqrt{\frac{L-k}{a}} \leqslant x \leqslant m + \sqrt{\frac{L-k}{a}}.$$

(Observe que $L - k \geqslant 0$, pois k é o valor mínimo de f.) A última desigualdade nos diz que x está contido num intervalo limitado, o que é uma contradição, pois x pode assumir qualquer valor real. Logo, a função quadrática é ilimitada superiormente se $a > 0$. O caso $a < 0$ é análogo e é deixado como exercício. ■

Para o que segue, considere o *discriminante*[5]:

$$\Delta = b^2 - 4ac.$$

Com essa notação, $k = -\Delta/4a$.

Proposição 3.6 (Zeros). *A função quadrática:*

(a) não possui zeros se $\Delta < 0$;

(b) possui um zero se $\Delta = 0$;

(c) possui dois zeros se $\Delta > 0$.

[5]Assim chamado porque serve para discriminar ou distinguir os zeros da função quadrática.

Demonstração.

$$f(x) = a(x-m)^2 + k = 0 \Leftrightarrow (x-m)^2 = -\frac{k}{a} = \frac{\Delta}{4a^2}.$$

Como $4a^2 > 0$ e $(x-m)^2 \geqslant 0$ para todo $x \in \mathbb{R}$, a última equação só terá solução se $\Delta \geqslant 0$. Assim:

(a) se $\Delta < 0$, não existe $x \in \mathbb{R}$ tal que $f(x) = 0$, ou seja, f não possui zeros;

(b) se $\Delta = 0$, segue-se que $(x-m)^2 = 0$ se, e somente se, $x = m = -b/2a$, ou seja, f possui apenas um zero: $x = -b/2a$;

(c) se $\Delta > 0$, a última igualdade acima equivale a

$$\left(x + \frac{b}{2a}\right)^2 = \frac{\Delta}{4a^2} \Leftrightarrow x + \frac{b}{2a} = \pm\frac{\sqrt{\Delta}}{2a} \Leftrightarrow x = \frac{-b \pm \sqrt{\Delta}}{2a},$$

ou seja, f tem dois zeros distintos:

$$\alpha = \frac{-b - \sqrt{\Delta}}{2a} \quad \text{e} \quad \beta = \frac{-b + \sqrt{\Delta}}{2a}.$$

■

Observação 3.7. Uma equação do segundo grau possui (eventualmente) *raízes*; uma função quadrática possui (eventualmente) *zeros*. Os x que são zeros de uma função quadrática coincidem com as raízes da equação do segundo grau correspondente, mas usamos esses nomes distintos. ○

Observação 3.8. Por economia de linguagem, mesmo quando uma equação do segundo grau possuir apenas uma raiz ou uma função quadrática possuir apenas um zero, por vezes diremos "duas raízes" ou "dois zeros", que devem ser entendidos de modo mais amplo como "uma ou duas raízes" ou "um ou dois zeros". Quando duas raízes são iguais também dizemos que a equação do segundo grau possui uma "raiz dupla". ○

Corolário 3.9. *Seja uma função quadrática $f : \mathbb{R} \to \mathbb{R}$, $f(x) = ax^2 + bx + c$, com $\Delta \geqslant 0$ e zeros α e β. Então $\alpha \leqslant \beta$ e:*

$$\alpha + \beta = -\frac{b}{a} \quad e \quad \alpha\beta = \frac{c}{a}.$$

Corolário 3.10 (Forma fatorada). *Seja uma função quadrática $f : \mathbb{R} \to \mathbb{R}$, $f(x) = ax^2 + bx + c$, com $\Delta \geqslant 0$ e zeros α e β. Então:*

$$f(x) = a(x-\alpha)(x-\beta), \forall x \in \mathbb{R}.$$

Demonstração. De fato, como $a \neq 0$, temos, para todo $x \in \mathbb{R}$,

$$f(x) = ax^2 + bx + c = a\left(x^2 + \frac{b}{a}x + \frac{c}{a}\right) = a[x^2 - (\alpha + \beta)x + \alpha\beta],$$

em que na última igualdade utilizamos o corolário anterior. Como o termo entre colchetes é o desenvolvimento de $(x - \alpha)(x - \beta)$, segue-se:

$$f(x) = a(x - \alpha)(x - \beta),\ \forall x \in \mathbb{R},$$

que é a chamada *forma fatorada* de f. ∎

Corolário 3.11 (Sinal da função quadrática). *Se x está entre dois zeros de uma função quadrática f, $f(x)$ tem o sinal oposto ao do coeficiente a. Caso contrário, ou $f(x) = 0$ ou $f(x)$ tem o mesmo sinal de a.*

Demonstração. Vimos no Corolário 3.10 que a lei de correspondência de f se escreve na forma $f(x) = a(x - \alpha)(x - \beta)$, $\forall x \in \mathbb{R}$. Como o produto $(x - \alpha)(x - \beta)$ é negativo se, e somente se, $\alpha < x < \beta$, segue-se a tese. ∎

Observação 3.12. Note que o corolário acima inclui também os casos em que f possui apenas um zero (caso no qual $x \neq \alpha$ implica que $f(x)$ tem o mesmo sinal de a) ou em que f não possui zeros (e portanto $f(x)$ tem o mesmo sinal de a para todo $x \in \mathbb{R}$). ○

Como você já deve ter notado, a função quadrática não é injetiva. Gostaríamos de saber quando $f(x) = f(y)$ para $x \neq y$. A resposta está na próxima proposição.

Proposição 3.13 (Simetria). *Uma função quadrática assume o mesmo valor em pontos distintos x e y se, e somente se, esses pontos são equidistantes de $-b/2a$.*

Demonstração. Uma condição necessária e suficiente para que $f(x) = f(y)$ é $a(x - m)^2 + k = a(y - m)^2 + k$, ou seja, $(x - m)^2 = (y - m)^2$. Como, por hipótese, $x \neq y$, devemos ter $x - m = -y + m$, i.e., a equação anterior equivale a

$$\frac{x + y}{2} = m = -\frac{b}{2a}.$$

Noutras palavras, $-b/2a$ é o ponto médio entre x e y e portanto equidista de ambos. Assim, a condição necessária e suficiente para que $f(x) = f(y)$, com $x \neq y$, é x e y serem equidistantes de $-b/2a$. ∎

Outro desdobramento da forma canônica nos conduzirá à forma do gráfico de uma função quadrática.

3.4 Gráfico da função quadrática

O principal objetivo desta seção é mostrar que o gráfico de uma função quadrática é uma parábola. Cumprido esse objetivo, veremos como o gráfico facilita a visualização de alguns fatos relevantes vistos acima sob um ponto de vista algébrico e faremos a interpretação geométrica dos coeficientes da lei de correspondência. Falaremos sobre a congruência e semelhança dos gráficos das funções quadráticas e apresentaremos o conceito de reta tangente a uma parábola, que depois utilizaremos para estudar o refletor parabólico. Vamos iniciar recordando a seguinte:

Definição 3.14 (Parábola). *Num plano* Π*, sejam um ponto* F *e uma reta* r *que não o contém. Dizemos que a* parábola de foco F e diretriz r *é o conjunto dos pontos de* Π *que distam igualmente de* F *e de* r*.*

Se preferir expressar isso na linguagem dos conjuntos, você poderia dizer que a parábola $\mathcal{P}(F, r)$ de foco F e diretriz r é:

$$\mathcal{P}(F, r) = \{P \in \Pi;\ d(P, F) = d(P, r)\},$$

em que d denota a distância entre dois pontos ou entre um ponto e uma reta de Π.

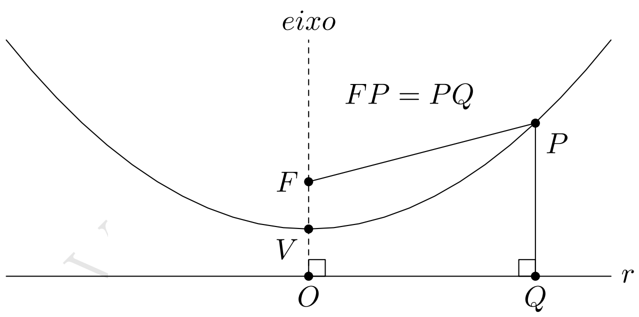


Figura 3.2: Principais elementos da parábola.

Há dois elementos da parábola que merecem especial destaque: a reta perpendicular a r baixada a partir de F, que chamamos o *eixo* da parábola, e o ponto desta mais próximo de r, que chamamos o *vértice* dessa parábola. Note que o vértice é o ponto médio do segmento que une F à intersecção do eixo com a diretriz. (Ver Figura 3.2.)

Considere f uma função quadrática totalmente arbitrária. Já vimos que sua lei de correspondência pode ser expressa na forma canônica: $f(x) = a(x - m)^2 + k$, sendo a, m e k números reais quaisquer. Considere como um esboço do seu gráfico a Figura 3.3, sem qualquer compromisso com a forma que escolhemos para o traçado (inclusive com o sinal de a, que não nos faz perder generalidade). Foque apenas nos elementos geométricos que deveriam estar presentes se tal fosse o traçado de uma parábola, i.e., o foco F e a diretriz r, e no ponto extremo $(m,\, k)$ que já estudamos acima, e se questione: seria esse gráfico uma parábola?

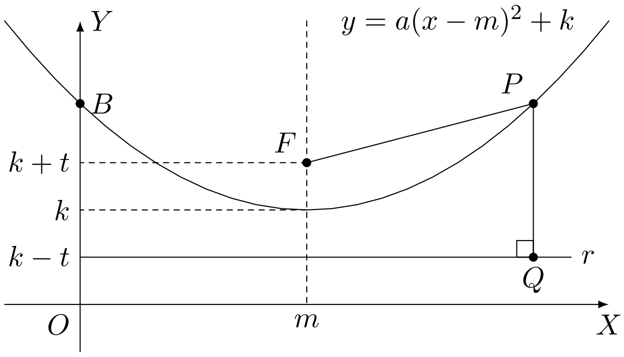


Figura 3.3: Uma parábola seria o gráfico de uma função quadrática?

Se for verdade que esse gráfico é uma parábola, algumas condições devem ser satisfeitas (considere a Figura 3.3):

- o ponto B de intersecção do gráfico de f com o eixo OY deve ter coordenadas:

$$B = (0,\, am^2 + k);$$

- o ponto F, supostamente o foco da parábola, deve ter coordenadas:

$$F = (m,\, k+t),$$

  para algum $t > 0$;

- por simetria, a equação da diretriz r deve ser:

$$y = k - t;$$

- um ponto $P = (x,\, a(x-m)^2 + k)$ qualquer do gráfico de f deve satisfazer:

$$d(P,\, F) = d(P,\, r),$$

  ou seja, a distância de P até F deve ser a mesma de P a r. Como se tratam de números positivos, essa equação equivale a:

$$[d(P,\, F)]^2 = [d(P,\, r)]^2. \tag{3.4.1}$$

Dito isso, nossa primeira tarefa é descobrir o valor de t. A equação (3.4.1) deve valer quando, em particular, $P = B$, caso em que temos:

$$m^2 + (am^2 - t)^2 = (am^2 + t)^2.$$

Desenvolvendo ambos membros e simplificando (faça como exercício!), obtemos:

$$t = 1/4a.$$

De posse do valor de t, para mostrar que o gráfico de f é uma parábola, basta agora verificar que a equação (3.4.1) é satisfeita para um ponto P qualquer do gráfico de f. Façamos isso. Primeiramente, note que:

$$[d(P,\, r)]^2 = \left[a(x-m)^2 + \frac{1}{4a}\right]^2.$$

Em seguida, vamos calcular:

$$\begin{aligned}
[d(P,\, F)]^2 &= (x-m)^2 + \left[a(x-m)^2 - \frac{1}{4a}\right]^2 \\
&= (x-m)^2 + [a(x-m)^2]^2 - 2a(x-m)^2\frac{1}{4a} + \left(\frac{1}{4a}\right)^2 \\
&= [a(x-m)^2]^2 + 2a(x-m)^2\frac{1}{4a} + \left(\frac{1}{4a}\right)^2 = \left[a(x-m)^2 + \frac{1}{4a}\right]^2.
\end{aligned}$$

Assim, $[d(P,\, F)]^2 = [d(P,\, r)]^2$ para um ponto P qualquer do gráfico de f. Isso significa que, de fato, o gráfico de f é uma parábola. Como também os parâmetros a, m e k que caracterizam f são números reais arbitrários, a nossa conclusão vale para qualquer função quadrática. Demonstramos, assim, a seguinte

Proposição 3.15. *O gráfico de uma função quadrática é uma parábola.*

O gráfico da função quadrática f nos ajuda a entender vários aspectos do comportamento dessa função de um ponto de vista geométrico. Por exemplo:

- se o gráfico de f intersecta o eixo OX em dois pontos, as abscissas α e β desses pontos são as raízes da equação $f(x) = 0$;
- a abscissa do ponto médio do intervalo $[\alpha,\, \beta]$ é a abscissa do vértice e do eixo da parábola;
- se $\alpha < x < \beta$, então $f(x)$ tem sinal contrário ao do coeficiente a; se $x < \alpha$ ou $x > \beta$, então $f(x)$ tem o mesmo sinal de a (já havíamos visto isso no Corolário 3.11);
- se o gráfico de f tangencia o eixo OX, a equação $f(x) = 0$ tem uma raiz dupla (única);
- se o gráfico de f está inteiramente acima ou abaixo do eixo OX, a equação $f(x) = 0$ não possui raízes reais.

Vamos examinar a seguir em que condições os gráficos das funções quadráticas são parábolas congruentes ou semelhantes.

Congruência

Intuitivamente, diremos que duas figuras são congruentes se elas forem geometricamente equivalentes, no sentido de que as distâncias entre os seus pontos (e consequentemente as amplitudes dos seus ângulos) são preservados. Há certas transformações que nos permitem obter uma das figuras a partir da outra satisfazendo essa condição. Essas transformações chamam-se isometrias e consistem em: translações, reflexões e rotações. Para nosso estudo será suficiente consideraremos apenas as:

- translações horizontais (paralelas ao eixo OX): $(x,\, y) \mapsto (x + m,\, y)$;
- translações verticais (paralelas ao eixo OY): $(x,\, y) \mapsto (x,\, y + k)$;
- reflexões em relação ao eixo OX: $(x,\, y) \mapsto (x,\, -y)$.

Diremos que duas figuras do plano são congruentes se, após aplicarmos uma ou algumas dessas transformações a uma das figuras, obtivermos a outra figura e vice-versa.

Nossa figuras de interesse aqui são os gráficos de funções. Queremos falar especificamente do gráfico da função quadrática. Mas, antes, vamos observar, de um modo breve, porém um pouco mais geral, como o gráfico de uma função $f : \mathbb{R} \to \mathbb{R}$ se comporta sob essas transformações.

Observação 3.16 (Translação horizontal). Aplicando a translação horizontal $(x,\, y) \mapsto (x + m,\, y)$ ao gráfico de $f : \mathbb{R} \to \mathbb{R}$, obtém-se o gráfico de $\varphi : \mathbb{R} \to \mathbb{R}$ tal que $\varphi(x) = f(x - m)$, $\forall x \in \mathbb{R}$. De fato, essa translação transforma um ponto qualquer $(x,\, f(x))$ do gráfico de f no ponto $(x + m,\, f(x))$; escrevendo $u = x + m$, esse ponto corresponde ao ponto $(u,\, f(u - m)) = (u,\, \varphi(u))$ do gráfico de φ. ○

Observação 3.17 (Translação vertical). Aplicando a translação vertical $(x,\, y) \mapsto (x,\, y + k)$ ao gráfico de $f : \mathbb{R} \to \mathbb{R}$, obtém-se o gráfico de $\psi : \mathbb{R} \to \mathbb{R}$ tal que $\psi(x) = f(x) + k$, $\forall x \in \mathbb{R}$. Com efeito, essa translação transforma um ponto qualquer $(x,\, f(x))$ do gráfico de f no ponto $(x,\, f(x) + k) = (x,\, \psi(x))$ do gráfico de ψ. ○

Observação 3.18 (Reflexão em relação ao eixo OX). Por fim, a reflexão em torno do eixo horizontal $(x,\, y) \mapsto (x,\, -y)$ leva o gráfico de $f : \mathbb{R} \to \mathbb{R}$ no gráfico de $\omega : \mathbb{R} \to \mathbb{R}$ tal que $\omega(x) = -f(x)$, $\forall x \in \mathbb{R}$. ○

Agora considere a função quadrática $f : \mathbb{R} \to \mathbb{R}$, com $f(x) = ax^2 + bx + c$. Claro que podemos escrever $f(x) = a(x - m)^2 + k$.

Aplicando-se a translação vertical $(x,\, y) \mapsto (x,\, y - k)$ ao gráfico de $f : \mathbb{R} \to \mathbb{R}$, $f(x) = a(x - m)^2 + k$, obtém-se o gráfico de $\psi : \mathbb{R} \to \mathbb{R}$, $\psi(x) = f(x) - k = a(x - m)^2$. (Figura 3.4.)

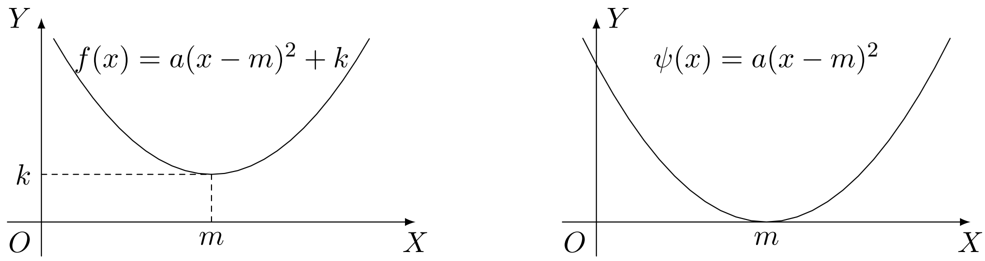


Figura 3.4: Translação vertical do gráfico de f.

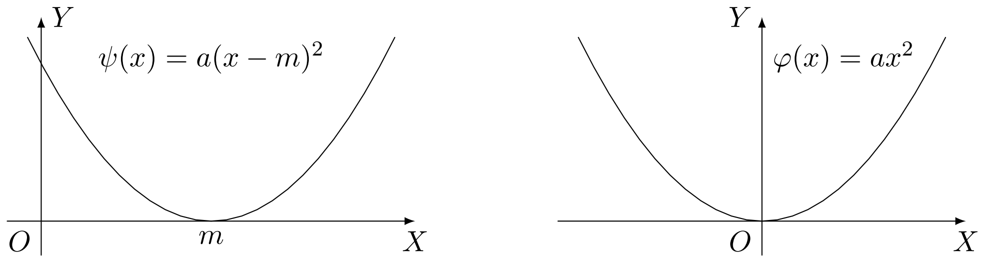


Figura 3.5: Translação horizontal do gráfico de ψ.

Na sequência, aplicando-se a tranlação horizontal $(x,\, y) \mapsto (x-m,\, y)$ ao gráfico de ψ, obtém-se o gráfico da função $\varphi : \mathbb{R} \to \mathbb{R}$, $\varphi(x) = \psi(x+m) = ax^2$. (Figura 3.5.)

Assim, a parábola que é o gráfico da função f transforma-se na parábola que é o gráfico de φ. As duas são, portanto, congruentes. E mais ainda: aplicando-se ao gráfico de φ uma reflexão em torno do eixo horizontal, obtém-se o gráfico da função $\omega : \mathbb{R} \to \mathbb{R}$, com $\omega(x) = -ax^2$. (Figura 3.6.)

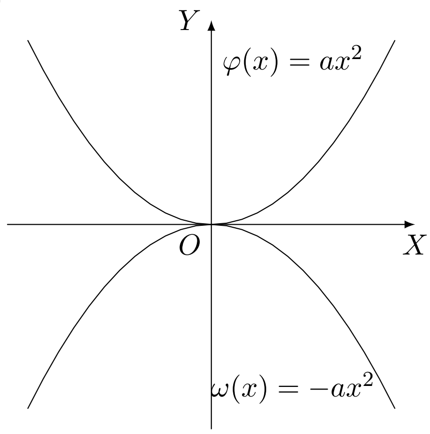


Figura 3.6: Reflexão em torno do eixo horizontal.

Podemos resumir essa discussão na seguinte:

Proposição 3.19. *Os gráficos das funções* $f_i : \mathbb{R} \to \mathbb{R}$, $f_i(x) = a_i x^2 + b_i x + c_i$, $i = 1, 2$, *são congruentes se, e somente se,* $a_1 = \pm a_2$.

Ou seja, os coeficientes b_i e c_i, $i = 1, 2$, não importam no que diz respeito à forma. Eles apenas determinam a posição da parábola em relação ao sistema de eixos, como veremos abaixo ao fazermos a interpretação dos coeficientes.

Corolário 3.20. *Qualquer parábola pode ter equação da forma* $y = ax^2$, *bastando para isso escolher convenientemente o sistema de eixos.*

Semelhança

A questão natural que se coloca após se enunciar a Proposição 3.19 é se as parábolas que são os gráficos de duas funções quadráticas podem ser congruentes mesmo quando $a_1 \neq a_2$. A resposta é negativa. Para ver isso, considere, por exemplo, as funções f_1 e f_2 dadas por $f_1(x) = a_1 x^2$ e $f_2(x) = a_2 x^2$, com $a_1 > 0$ e $a_2 > 0$. Se $a_1 > a_2$, então $a_1 x^2 > a_2 x^2$ para todo $x \neq 0$; da mesma forma, se $a_1 < a_2$, então $a_1 x^2 < a_2 x^2$ para todo $x \neq 0$. (Figura 3.7.) Quanto maior a, mais fechada a parábola e, vice-versa, quanto menor esse parâmetro, mais aberta é a curva, como se pode ver na figura 3.7. Aqui, "maior" e "menor" devem ser entendidos no sentido de valor absoluto.

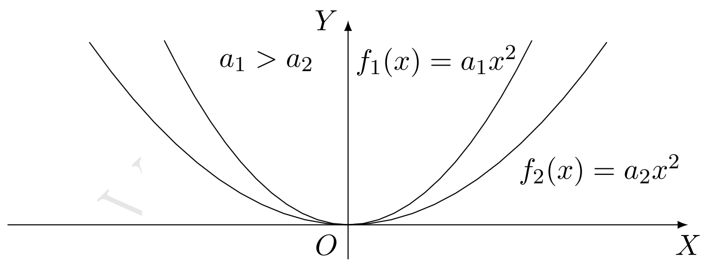


Figura 3.7: Parábolas semelhantes, porém não congruentes.

Aqui entra o conceito de *homotetia* (semelhança) de razão k e centro na origem, que consiste na transformação $(x, y) \mapsto (x', y') = (kx, ky)$, que transforma a parábola de equação $y = ax^2$ na parábola de equação (substituindo x por x'/k e y por y'/k):

$$\frac{y'}{k} = a\left(\frac{x'}{k}\right)^2,$$

ou seja,

$$y' = \frac{a}{k} x'^2.$$

Isso significa que se duas parábolas de equações $y = a_1x^2$ e $y' = a_2x'^2$ são semelhantes, a razão de semelhança entre as duas é k tal que $a_2 = a_1/k$, i.e., $k = a_1/a_2$. Quanto $k = 1$, temos uma isometria (congruência).

Tendo em vista o Corolário 3.20, segue-se que quaisquer duas parábolas são semelhantes. Podemos sintetizar isso na seguinte

Proposição 3.21. *As parábolas que são os gráficos das funções $f_i : \mathbb{R} \to \mathbb{R}$, $f_i(x) = a_ix^2 + b_ix + c_i$, $i = 1, 2$, são semelhantes entre si.*

Interpretação dos coeficientes

Este é também o momento conveniente para interpretarmos o significado dos coeficientes a, b e c da lei de correspondência de f.

O significado de do coeficiente c é óbvio: $c = f(0)$ é a ordenada do ponto em que o gráfico de f intersecta o eixo OY.

3.5 Algumas aplicações

Vamos começar resolvendo o problema da epígrafe deste capítulo.

Exemplo 3.22 (Máximo lucro). Um restaurante a quilo vende 100 kg/dia de comida a R$ 12,00/kg. Uma pesquisa de opinião revelou que, a cada real de aumento no preço, o restaurante perderia 10 clientes, com o consumo médio de 500 g cada um. Qual deve ser o preço do quilo de comida para que o restaurante tenha a maior receita possível? ◇

Solução. Supondo um consumo médio diário de meio quilograma por cliente, esse restaurante atende em média 200 clientes por dia. Aumentando x reais no preço do quilo, ele atenderá $200 - 10x$ clientes, logo venderá $(200 - 10x)/2$ quilos por dia a $12 + x$ reais o quilo, arrecadando $R = (200 - 10x)(12 + x)/2 = -5x^2 + 40x + 1200$ reais por dia. A receita máxima, conforme vimos na demonstração da Proposição 3.4, é alcançada quando $x = -b/2a = -40/(-10) = 4$, ou seja, quando o restaurante cobrar R$ 16,00/kg. ◆

Um outro problema clássico de otimização é o de encontrar a área máxima de um retângulo quando o seu perímetro (ou o semiperímetro) é fixado.

Exemplo 3.23 (Máxima área [10]). Com 80 metros de cerca, um fazendeiro deseja circundar uma área retangular junto a um rio para confinar alguns animais. Quais devem ser as medidas do retângulo para que a área cercada seja a maior possível? ◇

Solução. Digamos que os lados perpendiculares ao rio medem x e o lado paralelo mede y. Devemos ter $2x + y = 80$, ou seja, $y = 80 - 2x$. A área cercada mede $xy = x(80 - 2x) = -2x^2 + 80x$ e deve ser máxima quando $x = -b/2a = -80/(-4) = 20$; com isso, $y = 80 - 2 \cdot 20 = 40$. Logo, os lados perpendiculares ao rio devem ter 20 metros e o lado paralelo 40 metros. ◆

Também podemos ter o interesse de minimizar uma certa quantidade. Se essa quantidade for modelada por uma função quadrática, isso é possível com as ferramentas de que dispomos, como ilustra o próximo exemplo.

Exemplo 3.24 (Mínimo erro). Sejam x_1, x_2, …, x_n os valores encontrados para medições igualmente precisas de uma grandeza x. É razoável que o valor adotado para essa grandeza seja escolhido de modo que o erro incorrido pelas medições seja o menor possível. Em geral, esse erro é medido pelo chamado desvio quadrático total, que consiste na soma dos quadrados dos erros de cada medição. Ele é assim definido:

$$\delta(x) := \sum_{i=1}^{n} (x - x_i)^2.$$

Demonstre que δ é minimizado quando x é a média aritmética das medições. ◇

Solução. Desenvolvendo o membro direito, obtemos:

$$\delta(x) = \sum_{i=1}^{n} (x^2 - 2x_i x + x_i^2) = nx^2 + \left(-2\sum_{i=1}^{n} x_i\right) x + \sum_{i=1}^{n} x_i^2,$$

que é da forma $\delta(x) = ax^2 + bx + c$, em que os coeficientes a e b são respectivamente n e $-2\sum x_i$. Como $a > 0$, δ possui um valor mínimo quando $x = -b/2a$, ou seja, quando

$$x = -\frac{1}{2n}\left(-2\sum_{i=1}^{n} x_i\right) = \frac{1}{n}\sum_{i=1}^{n} x_i,$$

que nada mais é do que a média aritmética das medições. ◆

O próximo exemplo ilustra o uso dos zeros da função quadrática.

Exemplo 3.25 ([10]). Nas águas paradas de um lago, Marcelo rema seu barco a 12 km/h. Num certo rio, com o mesmo barco e as mesmas remadas, ele percorreu 12 km a favor da corrente e 8 km contra a corrente, num tempo total de 2 horas. Qual era a velocidade do rio, quanto tempo ele levou para ir e quanto tempo para voltar? ◇

Solução. Como (tempo de percurso) = (distância percorrida)/(velocidade), chamando de x a velocidade da corrente, os tempos gastos por Marcelo foram $12/(12+x)$ horas para ir e $8/(12-x)$ horas para voltar. Podemos então escrever o tempo total de 2 horas assim:

$$\frac{12}{12+x} + \frac{8}{12-x} = 2.$$

Multiplicando ambos os membros por $(12+x)(12-x)$ e simplificando, obtemos a equação $x^2 - 2x - 24 = 0$. Suas raízes são:

$$\alpha = \frac{2-\sqrt{100}}{2} = -4 \quad \text{e} \quad \beta = \frac{2+\sqrt{100}}{2} = 6.$$

Sendo 6 a única raiz positiva, a velocidade da corrente é 6 km/h e os tempos gastos por Marcelo são 12/18 h = 2/3 h = 40 min na ida e 8/6 h = 1h20min na volta. ◆

No exemplo seguinte, você verá por que a trajetória de projéteis na vizinhança da superfície da Terra é parabólica.

Exemplo 3.26 (Trajetória de um projétil). Imagine um projétil lançando por uma força instantânea e, após, sujeito apenas à força da gravidade, sendo desprezada a resistência do ar. Vamos descrever cada ponto $P = (x, y)$ da trajetória por meio das coordenadas $x = x(t)$ e $y = y(t)$. (Notações e convenções na Figura 3.8)

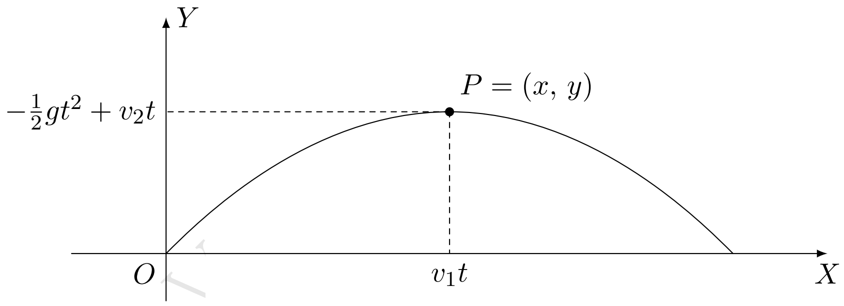


Figura 3.8: Trajetória de um projétil na vizinhança da superfície da Terra.

Como nenhuma força atua na direção horizontal, temos um movimento uniforme descrito por:

$$x(t) = v_1t + x_0,$$

em que v_1 é a velocidade inicial na direção do eixo OX e $x_0 = 0$ é a posição vertical inicial.

Já na direção vertical, como atua somente a força da gravidade, que pode ser muito bem considerada constante na vizinhança da superfície da Terra, o movimento é modelado por uma função quadrática:

$$y(t) = -\frac{1}{2}gt^2 + v_2t + y_0,$$

em que $-g$ é a aceleração da gravidade (com o sinal negativo por termos adotado o sentido do eixo OY oposto ao da gravidade), v_2 é a velocidade inicial na direção do eixo OY e $y_0 = 0$ é a posição vertical inicial.

Se $v_1 = 0$, então, para todo t, $x = 0$, logo $P = (0, y)$, com $y = -gt^2/2 + v_2 t$ e, neste caso, a trajetória do projétil é vertical.

Por outro lado, se $v_1 \neq 0$, de $x = v_1 t$, obtemos $t = x/v_1$. Substituindo t por esse valor na expressão de y, obtemos

$$y = \left(-\frac{g}{2v_1^2}\right)x^2 + \left(\frac{v_2}{v_1}\right)x,$$

que nos mostra que a trajetória do projétil é uma parábola. ◇

Por fim, um (belo) exemplo que faz uso de uma notável propriedade da parábola.

Exemplo 3.27 (Refletor parabólico [10]). Este exemplo diz respeito à propriedade refletora da parábola. Se a girarmos em torno de seu eixo, ela vai gerar uma superfície chamada *parabolóide de revolução* ou *superfície parabólica* (Figura 3.9).

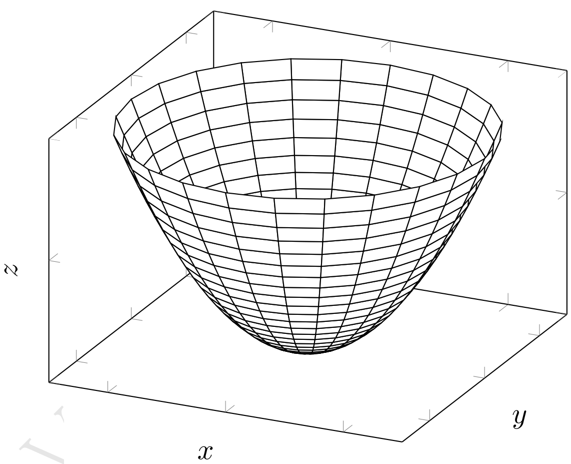


Figura 3.9: Parabolóide de revolução.

Essa superfície possui diversas aplicações interessantes, tanto no sentido de encaminhar na direção paralela ao eixo raios de luz que emanam do foco (como nos holofotes, nos faróis de automóveis ou em simples lanternas de mão) quanto no sentido de fazer raios ou sinais paralelos ao eixo que incidem sobre sua superfície interna convergirem para o foco a fim de concentrá-los e reforçá-los (como nas antenas utilizadas na rádio-astronomia ou nos aparelhos de televisão). (Figura 3.10.)

Vamos analisar o fundamento matemático dessas aplicações.

Para o nosso estudo, vamos considerar apenas os raios que chegam ao (ou partem do) refletor parabólico na direção paralela ao seu eixo.

Partindo do princípio de que o raio incidente sobre uma superfície refletora e o correspondente raio refletido estão contidos no mesmo plano e dada a simetria do

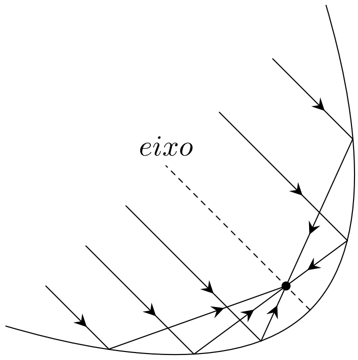


Figura 3.10: Propriedade refletora da parábola.

refletor parabólico em relação a seu eixo, podemos substituir a superfície parabólica pela parábola que é a intersecção dessa superfície com o plano que contém o raio incidente, o raio refletido e o eixo de rotação (que coincide com o eixo da parábola).

Por definição, o ângulo entre uma reta e uma curva que se intersectam num ponto P é o ângulo entre essa reta e a reta tangente à curva em P. (Figura 3.11.)

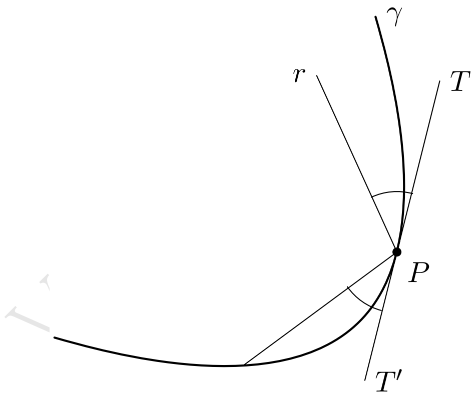


Figura 3.11: Ângulo entre uma reta r e uma curva γ.

A tangente a uma parábola γ no ponto P é a reta TT' que tem em comum com γ apenas o ponto P e tal que os demais pontos de γ estão no mesmo semiplano determinado por t. (Figura 3.11.)

Se γ é o gráfico da função cuja lei de correspondência é $f(x) = ax^2 + bx + c$, $P = (x_0, y_0)$, em que $y_0 = ax_0^2 + bx_0 + c$, então a inclinação de TT' é igual a $2ax_0 + b$.

Para ver que isso é verdade, basta mostrarmos que todos os pontos de γ que têm abscissa diferente de x_0 não pertencem a TT' e estão no mesmo semiplano determinado por essa reta.

Sem perda de generalidade, suponha que $a > 0$. Vamos mostrar então que, para todo $x \neq x_0$, o ponto (x, y) de γ, com $y = ax^2 + bx + c$, está acima do ponto $(x, y_0 + (2ax_0 + b)(x - x_0))$, de mesma abscissa x sobre TT'. Em outras palavras, supondo $a > 0$, queremos mostrar que:

$$x \neq x_0 \Rightarrow ax^2 + bx + c > ax_0^2 + bx_0 + c + (2ax_0 + b)(x - x_0).$$

Mas,

$$x \neq x_0 \Rightarrow ax^2 + bx + c - [ax_0^2 + bx_0 + c + (2ax_0 + b)(x - x_0)] = a(x - x_0)^2 > 0.$$

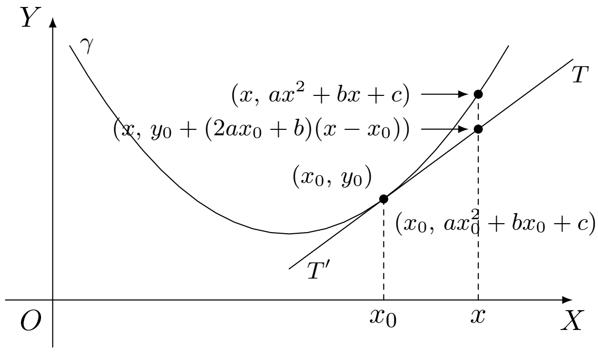


Figura 3.12: Tangente à parábola em (x_0, y_0).

Isso prova que TT', de inclinação $2ax_0 + b$ e que passa por $P = (x_0, y_0)$, com $y_0 = f(x_0)$, tem P como único ponto em comum com γ, que é o gráfico de f, e que todos os pontos de γ estão acima de TT'. Logo TT' é a tangente a γ em P. (Figura 3.12.)

Quando $a > 0$ (respec. $a < 0$), a parábola se situa acima (respec. abaixo) de qualquer de suas tangentes.

Claro que se uma reta r é paralela ao eixo de uma parábola, r tem apenas um ponto em comum com essa parábola, mas não é tangente, pois há pontos da parábola em ambos semiplanos determinados por r.

Agora, sabendo que a parábola γ, gráfico da função cuja lei de correspondência é $f(x) = ax^2 + bx + c$, tem no ponto $P = (x, y)$ uma tangente cuja inclinação é $2ax + b$, vamos calcular a inclinação da reta FQ que une o foco F ao ponto Q, pé da perpendicular baixada de P sobre a diretriz d de γ. (Figura 3.13.)

Vamos supor que P não é o vértice da parábola, i.e., que sua abscissa é diferente de $-b/2a$, e portanto $2ax + b \neq 0$. Se P fosse o vértice, a reta FQ seria vertical e a tangente TT' no ponto P teria inclinação zero, logo seria horizontal e FQ e TT' seriam perpendiculares.

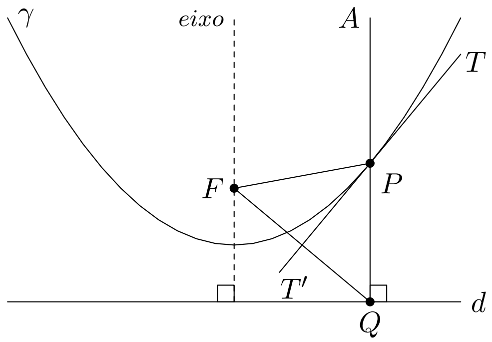


Figura 3.13: A reta FQ é perpendicular à reta TT'.

Já vimos que $F = (m,\, k + 1/4a)$ e $Q = (x,\, k - 1/4a)$, em que $m = -b/2a$ e k é a ordenada do vértice da parábola. Logo, a inclinação de FQ é:

$$\frac{k - \frac{1}{4a} - \left(k + \frac{1}{4a}\right)}{x - m} = \frac{-1}{2a(x-m)} = \frac{-1}{2a\left(x + \frac{b}{2a}\right)} = -\frac{1}{2ax+b}.$$

Isso significa que a reta FQ é perpendicular à reta TT', tangente a γ em P.

Para provar esse ponto, vamos mostrar que as retas cujas equações são $y = ax + b$ e $y = a'x + b'$, com $a \neq 0$ e $a' \neq 0$, são perpendiculares se, e somente se, $a' = -1/a$.

Com efeito, como as retas cujas equações são $y = ax$ e $y = a'x$ são paralelas às retas dadas, aquelas serão perpendiculares se, e somente se, estas o forem.

Se estas retas forem perpendiculares, tomando $x = 1$, o ponto $(1,\, a)$ pertence a uma delas e o ponto $(1,\, a')$ pertence à outra (ver Figura 3.14). Então o triângulo cujos vértices são os pontos $(0,\, 0)$, $(1,\, a)$ e $(1,\, a')$ é retângulo, logo a altura baixada do vértice do ângulo reto é a média geométrica dos segmentos que ela determina sobre a hipotenusa. Mas, por um lado, o comprimento da altura é 1. Por outro lado, um dos números a e a' é negativo, enquanto o outro é positivo. Sem perda de generalidade, suponhamos que a' seja negativo e que a seja positivo. Logo os segmentos medem a e $-a'$. Assim, $1 = -aa'$ e, portanto, $a' = -1/a$.

Reciprocamente, se $a' = -1/a$, considere a reta de equação $y = bx$, perpendicular à reta de equação $y = ax$ a partir da origem. Pelo que acabamos de ver, $b = -1/a$, logo $b = a'$. Assim, $y = a'x$ coincide com $y = bx$ e, portanto, a reta de equação $y = a'x$ é perpendicular à reta de equação $y = ax$.

Neste ponto, podemos finalmente enunciar a propriedade geométrica na qual se baseiam as aplicações da superfície parabólica, a saber: *a tangente à parábola num ponto P faz ângulos iguais com a paralela ao eixo e com a reta que une o foco F a esse ponto.*

De fato, se Q é o pé da perpendicular baixada de P sobre a diretriz, pela definição da parábola, os segmentos FP e PQ têm o mesmo comprimento, logo o triângulo FPQ é isósceles. Ademais, como já vimos, FQ é perpendicular à tangente, i.e., a tangente é

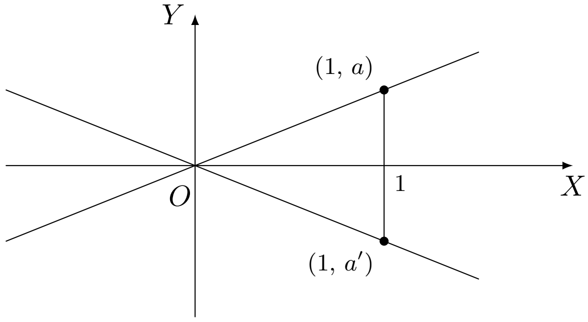


Figura 3.14: Retas cujas equações são $y = ax$ e $y = a'x$.

altura desse triângulo isósceles, logo é também bissetriz. Portanto, os ângulos $F\widehat{P}T'$ e $T'\widehat{P}Q$ são congruentes. Assim, $A\widehat{P}T = F\widehat{P}T' = \alpha$. (Ver Figura 3.15.)

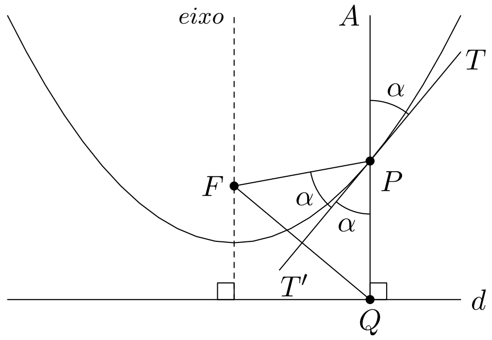


Figura 3.15: Os ângulos $A\widehat{P}T$ e $F\widehat{P}T'$ são congruentes.

Se a superfície parabólica estiver voltada, por exemplo, para a posição estacionária de um satélite, a grande distância faz com que os sinais por ele emitidos sigam trajetórias praticamente paralelas ao eixo da superfície e, assim que esses sinais são refletidos na superfície, convirjam para o foco. $\diamond$

Apêndices

A Alfabeto Grego

Não há razão obrigatória para o uso, em qualquer situação, de letras maiúsculas ou minúsculas, de fontes pomposas ou não, latinas, gregas ou provenientes de qualquer outro alfabeto. O uso de diferentes tipos de letra pode ser motivado por várias razões, dentre as quais diferenciar o tipo ou hierarquia de determinado objeto matemático[1], o esgotamento de letras de um determinado tipo, simples costume, convenção ou ainda o gosto de quem escreve. Seja qual for o motivo, caso você não esteja habituado, aqui está o alfabeto grego (minúsculas e maiúsculas) para lhe munir destes símbolos, tanto para a escrita quanto para a leitura de textos matemáticos.

$A\alpha$	Alfa	$H\eta$	Eta	$N\nu$	Ni	$T\tau$	Tau
$B\beta$	Beta	$\Theta\theta\vartheta$	Teta	$\Xi\xi$	Csi	$\Upsilon\upsilon$	Úpsilon
$\Gamma\gamma$	Gama	$I\iota$	Iota	Oo	Ômicron	$\Phi\phi\varphi$	Fi
$\Delta\delta$	Delta	K, $\kappa\varkappa$	Capa	$\Pi\pi\varpi$	Pi	$X\chi$	Qui
$E\epsilon\varepsilon$	Épsilon	$\Lambda\lambda$	Lambda	$P\rho\varrho$	Rô	$\Psi\psi$	Psi
$Z\zeta$	Zeta	$M\mu$	Mi	$\Sigma\sigma\varsigma$	Sigma	$\Omega\omega$	Ômega

[1]Por exemplo, na Geometria Euclidiana, é praxe indicar pontos por letras latinas maiúsculas, retas por letras latinas minúsculas e planos por letras gregas minúsculas; quando se trata de conjuntos, costuma-se diferenciar elementos dos conjuntos que os contém utilizando-se para estes letras do final do alfabeto, ou maiúsculas ou de fontes pomposas, e, para aqueles, letras do início do alfabeto, ou minúsculas ou de fontes mais simples.

B Sistema de Coordenadas Cartesiano

O que torna possível fazer a reta (resp. o plano, o espaço) corresponder a $\mathbb{R}$ (resp. $\mathbb{R}^2$, $\mathbb{R}^3$) e assim estabelecer uma via de mão dupla entre a Aritmética e a Álgebra de um lado e a Geometria de outro como comentamos na Seção 1.5.2 é a existência de uma bijeção entre esses conjuntos. A prova disso depende de umas poucas noções e proposições primitivas da Geometria Euclidiana acerca dos seus elementos básicos (pontos, retas e planos) e de alguns axiomas sobre a medida de segmentos. Vamos adotar aqueles de [11]. Por conveniência, vamos resumi-los abaixo e, em seguida, utilizá-los para demonstrar a existência dessa bijeção.

Axioma B.1. *A cada par de pontos distintos está associado um único número real positivo; a um par de pontos iguais está associado o número zero.* □

Definição B.2. *O número a que se refere o Axioma B.1 é chamado* medida *ou* comprimento *do segmento formado pelos pontos ou, ainda, a* distância *entre esses pontos.*

Notação B.3. A distância entre os pontos X e Y será denotada por $d(X, Y)$ ou $d(Y, X)$. ○

Observação B.4. Dito em outros termos, o Axioma B.1 afirma que se $\mathcal{C}$ é um conjunto de pontos, existe uma função $d : \mathcal{C} \times \mathcal{C} \to \mathbb{R}$, $d(X, Y) =$ *distância entre* X *e* Y. Observe que não há exceção nem ambiguidade, portanto d é de fato uma função. ○

Axioma B.5. *Se um ponto* Y *está entre dois pontos* X *e* Z*, então* $d(X, Z) = d(X, Y) + d(Y, Z)$. □

Axioma B.6. *A todo número real positivo fica associado um par de pontos cuja distância é igual a esse número; ao número zero fica associado um par de pontos iguais.* □

Observação B.7. Diferentemente do Axioma B.1, o Axioma B.6 não determina uma função. Considere, por exemplo, um quadrado $P_1P_2P_3P_4$ cujo lado mede l; note que l corresponde aos pares $(P_i,\, P_{i+1})$, $i \in \{1,\, 2,\, 3,\, 4\}$, com $P_5 = P_1$. ○

Axioma B.8 (Transporte de segmentos). *Fixado um par de pontos quaisquer P e Q, para todo par de pontos distintos X e Z, existe um único ponto Y que pertence à semirreta $\overrightarrow{XZ}$ tal que $d(P,\, Q) = d(X,\, Y)$.* □

De posse desses poucos axiomas, estamos aptos a demonstrar as bijeções entre a reta e $\mathbb{R}$, o plano e $\mathbb{R}^2$, o espaço e $\mathbb{R}^3$. Tenha sempre em mente o que foi dito na Observação 1.50 sobre desenhos!

B.1 Na reta

Teorema B.9. *Existe uma correspondência biunívoca entre pontos de uma reta e números reais.*

Demonstração.

(a) *Correspondência entre os pontos da reta e os números reais.*

Seja uma reta r e um ponto $O \in r$. Esse ponto divide r em três conjuntos: $\{O\}$ e outros dois que estão em semirretas opostas, aos quais chamaremos de *parte positiva* (r_+) e *parte negativa* (r_-) de r. Tais escolhas fazem da reta o que chamamos um *eixo* (Figura B.1).

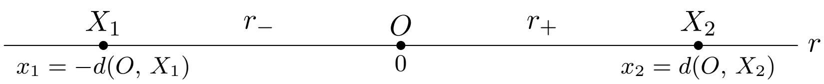


Figura B.1: Eixo.

Vamos colocar os pontos X de r em correspondência com os números reais por meio de uma função $\varphi : r \to \mathbb{R}$, definida da seguinte maneira:

$$\varphi(X) = \begin{cases} d(O,\, X) & \text{se } X \in r_+ \\ 0 & \text{se } X = O \\ -d(O,\, X) & \text{se } X \in r_- \end{cases}.$$

Como não há exceção (pois todo ponto de r está dessa forma associado a algum número real dado pelo Axioma B.1) nem ambiguidade (pois, novamente pelo Axioma B.1, cada número assim associado aos pontos de r é único), φ é de fato uma função.

(b) *Injetividade.*

De fato, sejam X_1 e X_2 pontos distintos quaisquer de r. Se um desses pontos for igual a O, claramente $\varphi(X_1) \neq \varphi(X_2)$. Suponhamos, então, que ambos são distintos de O. Nesse caso, há seis posições relativas entre X_1, X_2 e O. Considere, por exemplo, aquela em que $X_1 \in r_-$ e $X_2 \in r_+$ (Figura B.1). Então, pelo Axioma B.5, $d(X_1, X_2) = d(X_1, O) + d(O, X_2)$, ou seja, pela definição de φ, $d(X_1, X_2) = -\varphi(X_1) + \varphi(X_2)$; como, por hipótese, $X_1 \neq X_2$, segue-se que $d(X_1, X_2) \neq 0$; logo, $\varphi(X_1) \neq \varphi(X_2)$ (caso contrário, seria $d(X_1, X_2) = 0$, contradição). Os outros cinco casos são totalmente análogos. Portanto, como $X_1 \neq X_2 \Rightarrow \varphi(X_1) \neq \varphi(X_2)$, para quaisquer $X_1, X_2 \in r$, segue-se que φ é injetiva.

(c) *Sobrejetividade.*

Dado um número real positivo qualquer x, o Axioma B.6 nos garante que existe um par de pontos, digamos, P e Q, tal que $d(P, Q) = x$; pelo Axioma B.8, existe um único ponto X sobre r_+ tal que $d(O, X) = d(P, Q)$, ou seja, tal que $x = \varphi(X)$; se x for um número real negativo qualquer, pelos mesmos argumentos anteriores existe um único ponto X sobre r_- tal que $d(O, X) = d(P, Q) = -x$, ou seja, $x = -d(O, X) = \varphi(X)$; por fim, se $x = 0$, basta tomar $X = O$: $\varphi(O) = 0$. Assim, para todo $x \in \mathbb{R}$ existe um $X \in r$ tal que $x = \varphi(X)$, logo φ é sobrejetiva.

(d) *Bijetividade.*

Portanto, pela Definição 1.61, a correspondência φ entre a reta e o conjunto dos números reais é bijetiva, completando assim a demonstração. ■

Definição B.10. *Uma reta na qual se fixou um ponto O e se escolheu uma parte positiva e outra negativa chama-se um* eixo.

Observação B.11. Costuma-se indicar a parte positiva da reta por meio de uma seta. ○

Definição B.12. *A bijeção do Teorema B.9 é chamada* sistema de coordenadas *para a reta.*

Definição B.13. *Cada número x assim associado a um ponto P da reta é chamado* coordenada *desse ponto em relação à reta e o ponto associado ao número zero é chamado* origem *do sistema de coordenadas.*

Definição B.14. *Se X e Y são pontos de uma reta e x e y são suas respectivas coordenadas, dizemos que Y está* à direita *de X (ou que X está* à esquerda *de Y) quando $x < y$.*

B.2 No plano

Teorema B.15. *Existe uma correspondência biunívoca entre pontos de um plano e pares ordenados de números reais.*

Demonstração.

(a) *Correspondência entre os pontos do plano e os pares ordenados de números reais.*

Seja um plano Π e um par de eixos OX e OY, tomados em Π e que se intersectam perpendicularmente no ponto O.

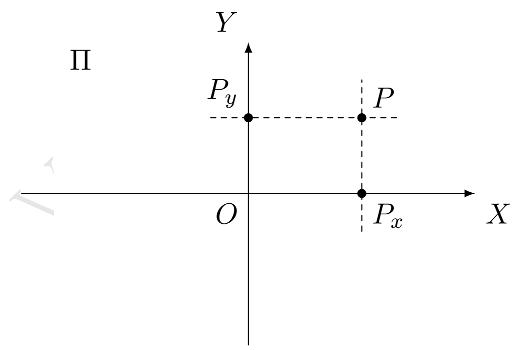


Figura B.2: Sistema de eixos ortogonais no plano.

Vamos colocar os pontos de Π em correspondência com os pares ordenados de números reais da seguinte maneira. Seja P um ponto qualquer de Π. Passando por esse ponto, considere uma reta perpendicular a OX e outra perpendicular a OY, a primeira intersectando OX no ponto P_x e a segunda intersectando OY em P_y (figura B.2). Segue do Teorema B.9 que a cada um dos pontos P_x e P_y está associado um único número real, digamos x e y, respectivamente. Diremos, então, que o ponto P está associado ao par ordenado de números reais (x, y). Por um lado, a existência das perpendiculares (teorema da Geometria Euclidiana) assegura que todo ponto do plano está associado dessa forma a algum par ordenado de números reais (não há exceção); por outro, a unicidade das perpendiculares (teorema da Geometria Euclidiana) implica que cada ponto do plano está associado dessa maneira a apenas um par ordenado de números reais (não há ambiguidade). Portanto, o modo como associamos pontos P do plano a pares ordenados de números reais (x, y) é uma função $\psi : \Pi \to \mathbb{R}^2$, $\psi(P) = (x, y)$.

(b) *Injetividade.*

Sejam dois pontos quaisquer $P, P' \in \Pi$ tais que $\psi(P) = \psi(P')$. Se $\psi(P) = (x, y)$ e $\psi(P') = (x', y')$, então $(x, y) = (x', y')$. Logo, pela Proposição 1.27, $x = x'$ e $y = y'$. Assim, as perpendiculares a OX pelos pontos P e P', digamos r e r' respectivamente, são iguais; analogamente, as perpendiculares a OY por esses pontos, digamos s e s' respectivamente, também são iguais. Então, como $r = r'$, $s = s'$, $r \cap s = \{P\}$ e $r' \cap s' = \{P'\}$, segue-se que $P = P'$. Portanto, como para todo $P, P' \in \Pi$, $\psi(P) = \psi(P') \Rightarrow P = P'$, segue-se que ψ é injetiva.

(c) *Sobrejetividade.*

Seja (x, y) um elemento abitrário de $\mathbb{R}^2$. Considere o ponto P_x de coordenada x sobre OX e o ponto P_y de coordenada y sobre OY. Um dos teoremas da Geometria Euclidiana garante que a perpendicular a OX por P_x existe e é única assim como a perpendicular a OY por P_y. Como essas retas não são paralelas, elas se encontram num único ponto, digamos P. Além disso, por essa construção, P é tal que $\psi(P) = (x, y)$. Como $(x, y) \in \mathbb{R}^2$ é arbitrário, podemos afirmar que todo par de números reais é a imagem de algum ponto do plano por ψ. Portanto, ψ é sobrejetiva.

(d) *Bijetividade.*

Por fim, a bijetividade de ψ segue da Definição 1.61. ■

Notação B.16. Se ψ é a função do Teorema B.15 e $\psi(P) = (x, y)$, então escrevemos $P = (x, y)$ ou $P(x, y)$. ○

Definição B.17. *A bijeção do Teorema B.15 chama-se* sistema de coordenadas retangular *(para o plano) ou* sistema de coordenadas cartesiano *(para o plano) ou, ainda,* plano cartesiano.

Definição B.18. *Os eixos OX e OY chamam-se respectivamente* eixo das abcissas *(ou abscissas) e* eixo das ordenadas*; ambos também recebem o nome de* eixos coordenados*.*

Definição B.19. *Os números x, $y \in \mathbb{R}$ do par ordenado (x, y) são as* coordenadas *(cartesianas ou retangulares) do ponto P; dizemos, ainda, que x é a* abcissa *(ou* abscissa*) ou* primeira coordenada *e y a* ordenada *ou* segunda coordenada *de P.*

Definição B.20. *Cada uma das quatro regiões em que o plano fica dividido pelos eixos OX e OY chama-se um* quadrante *e é caracterizado pelos sinais das coordenadas de seus pontos: primeiro quadrante ($x \geqslant 0$ e $y \geqslant 0$); segundo quadrante ($x \leqslant 0$ e $y \geqslant 0$); terceiro quadrante ($x \leqslant 0$ e $y \leqslant 0$); e quarto quadrante ($x \geqslant 0$ e $y \leqslant 0$).*

B.3 No espaço

Em analogia ao teorema anterior, um sistema de eixos ortogonais no espaço $\mathcal{E}$ da Geometria Euclidiana estabelece uma correspondência biunívoca entre os pontos $P \in \mathcal{E}$ e os ternos ordenados $(x, y, z) \in \mathbb{R}^3$. Vejamos.

Teorema B.21. *Existe uma correspondência biunívoca entre pontos do espaço e ternos ordenados de números reais.*

Demonstração.

(a) *Correspondência entre os pontos do espaço e os ternos ordenados de números reais.*

Considere o espaço $\mathcal{E}$ da Geometria Euclidiana. Sejam três eixos mutuamente perpendiculares, OX, OY e OZ, com a mesma origem O.
Seja P um ponto qualquer de $\mathcal{E}$. Passando por esse ponto, considere uma reta perpendicular a OZ e outra perpendicular ao plano Π_{XY} (determinado por OX e OY), a primeira intersectando OZ no ponto P_z e a segunda intersectando Π_{XY} em P_{xy} (Figura B.3). Segue do Teorema B.9 que ao ponto P_z está associado um único número real, digamos z, e, do Teorema B.15, segue que ao ponto P_{xy} está associado um único par ordenado de números reais, digamos (x, y). Diremos, então, que ao ponto P está associado o terno ordenado de números reais (x, y, z), sendo os números reais x, y e z obtidos dessa forma que acabamos de descrever. A existência e unicidade das perpendiculares (teorema da Geometria Euclidiana) implica na existência e unicidade dos

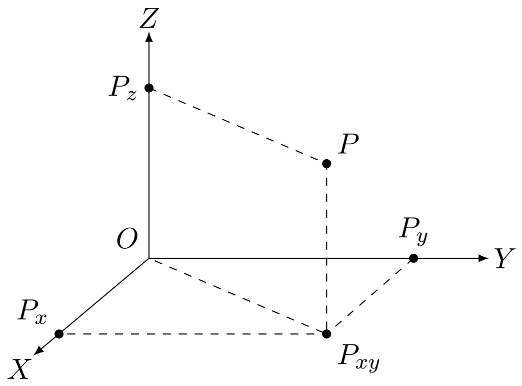


Figura B.3: Sistema de eixos ortogonais no espaço.

números reais x, y e z assim obtidos, e, portanto, na existência e unicidade do terno ordenado $(x, y, z) \in \mathbb{R}^3$ desse modo associados a P. Portanto, a maneira como associamos pontos P do espaço a ternos ordenados (x, y, z) de números reais é uma função $\zeta : \mathcal{E} \to \mathbb{R}^3$, $\zeta(P) = (x, y, z)$.

(b) *Injetividade.*

Para demonstrar a injetividade de ζ, observamos que $(x, y, z) = (u, v, w)$ se, e somente se, $x = u$, $y = v$ e $z = w$. Podemos provar esse fato definindo ternos ordenados como fizemos com pares ordenados (Definição 1.26, de Kuratowski, que é facilmente generalizável): $(x, y, z) := \{\{x\}, \{x, y\}, \{x, y, z\}\}$. (Tente como exercício!) Outra alternativa mais fácil seria definir ternos ordenados recursivamente a partir da definição de par ordenado: podemos definir (x, y, z) como $((x, y), z)$, ou seja, nosso terno ordenado consistiria num outro par ordenado. Dessa última definição, a propriedade fundamental decorre imediatamente da Proposição 1.27.

Para facilitar a demonstração da injetividade de ζ, também convém observar que os números x e y, do terno ordenado (x, y, z) correspondente ao ponto P, podem ser obtidos diretamente por meio da intersecção das perpendiculares por P a OX e a OY, respectivamente. De fato, usando as notações da figura B.3, considere a reta PP_{xy}; por construção, ela é perpendicular a Π_{XY}, logo é ortogonal a OX; por outro lado, também por construção, a reta $P_{xy}P_x$ é perpendicular a OX por P_x; segue-se que o plano $PP_{xy}P_x$ é perpendicular a OX por P_x e, consequentemente, a reta PP_x é a única (como já mencionamos e fizemos uso anteriormente, esse é um dos teoremas da Geometria Euclidiana) perpendicular a OX por P_x. Analogamente, PP_y é a única perpendicular a OY por P_y.

Feitas essas duas observações, vamos à demonstração da injetividade de ζ.

Suponha $P, P' \in \mathcal{E}$ tais que $\zeta(P) = \zeta(P')$. Se $\zeta(P) = (x,\, y,\, z)$ e $\zeta(P') = (x',\, y',\, z')$, então $(x,\, y,\, z) = (x',\, y',\, z')$. Logo, como observado acima, $x = x'$, $y = y'$ e $z = z'$. Assim, se r, s e t são as perpendiculares, por P, a OX, OY e OZ, respectivamente, e r', s' e t' as perpendiculares, por P', a OX, OY e OZ, respectivamente, o fato de $x = x'$, $y = y'$ e $z = z'$ implica $r = r'$, $s = s'$ e $t = t'$. Como também $r \cap s \cap t = \{P\}$ e $r' \cap s' \cap t' = \{P'\}$, segue-se que $P = P'$. Logo, como $\zeta(P) = \zeta(P') \Rightarrow P = P'$, quaiquer que sejam $P, P' \in \mathcal{E}$, segue-se que ζ é injetiva.

(c) *Sobrejetividade.*

Seja $(x,\, y,\, z)$ um elemento abitrário de $\mathbb{R}^3$. Considere os pontos P_x, P_y, P_z de coordenada x, y, z sobre os eixos OX, OY, OZ, respectivamente. Sabemos que a perpendicular a OX por P_x existe e é única assim como a perpendicular a OY por P_y. Como essas retas não são paralelas, elas se encontram num único ponto, digamos P_{xy} contido no plano Π_{XY}. Considere a perpendicular r a Π_{XY} por P_{xy}; ela é paralela a OZ, porque ambas (r e OZ) são perpendiculares a Π_{XY}, portanto determinam um plano; agora, considere a perpendicular s a OZ por P_z contida no plano determinado por r e OZ; ora, r e s são coplanares mas não paralelas, logo se encontram no mesmo ponto, digamos P. Mas essa construção de P é tal que $\zeta(P) = (x,\, y,\, z)$, sendo $(x,\, y,\, z)$ um elemento abitrário de $\mathbb{R}^3$. Portanto, ζ é sobrejetiva.

(d) *Bijetividade.*

Da Definição 1.61 decorre que ζ é bijetiva, completando assim a demonstração. ■

Notação B.22. Se ζ é a função do Teorema B.21 e $\zeta(P) = (x,\, y,\, z)$, então escrevemos $P = (x,\, y,\, z)$ ou $P(x,\, y,\, z)$. ○

Definição B.23. *A bijeção do Teorema B.21 chama-se* sistema de coordenadas retangular *(para o espaço da Geometria Euclidiana) ou* sistema de coordenadas cartesiano *(para essa espaço) ou, ainda,* espaço cartesiano.

Definição B.24. *Os eixos OX, OY e OZ chamam-se respectivamente* eixo das abcissas *(ou abscissas),* eixo das ordenadas *e* eixo das cotas*, ou genericamente,* eixos coordenados.

Definição B.25. *Os números* $x, y, z \in \mathbb{R}$ *do terno ordenado* (x, y, z) *são as* coordenadas *(cartesianas ou retangulares) do ponto* P*; dizemos, ainda, que* x *é a* abcissa *(ou* abscissa*) ou* primeira coordenada*,* y *a* ordenada *ou* segunda coordenada *e* z *a* cota *ou* terceira coordenada *de* P*.*

Definição B.26. *Cada um dos planos determinados pelos pares de eixos coordenados chama-se um* plano coordenado *ou* plano cartesiano*.*

Definição B.27. *Cada uma das oito regiões em que o espaço fica dividido pelos planos cartesianos chama-se um* octante*, caracterizado pelos sinais das coordenadas de seus pontos:*

$$\begin{aligned}
&\{(x, y, z) \in \mathbb{R}^3;\, x \geqslant 0,\, y \geqslant 0,\, z \geqslant 0\}, && \{(x, y, z) \in \mathbb{R}^3;\, x \leqslant 0,\, y \geqslant 0,\, z \geqslant 0\},\\
&\{(x, y, z) \in \mathbb{R}^3;\, x \leqslant 0,\, y \leqslant 0,\, z \geqslant 0\}, && \{(x, y, z) \in \mathbb{R}^3;\, x \geqslant 0,\, y \leqslant 0,\, z \geqslant 0\},\\
&\{(x, y, z) \in \mathbb{R}^3;\, x \geqslant 0,\, y \geqslant 0,\, z \leqslant 0\}, && \{(x, y, z) \in \mathbb{R}^3;\, x \leqslant 0,\, y \geqslant 0,\, z \leqslant 0\},\\
&\{(x, y, z) \in \mathbb{R}^3;\, x \leqslant 0,\, y \leqslant 0,\, z \leqslant 0\}, && \{(x, y, z) \in \mathbb{R}^3;\, x \geqslant 0,\, y \leqslant 0,\, z \leqslant 0\}.
\end{aligned}$$

O primeiro octante é aquele cujas coordenadas não são negativas; os demais não têm uma designação padrão. Em alguns livros mais antigos, os octantes são chamados de triedros.

C Distâncias

C.1 Distâncias no sistema cartesiano

Reta

Sejam X, Y e Z pontos quaisquer de um eixo. Lembrando que $d(X, Y) \geqslant 0$, $d(X, Y) = d(Y, X)$ (Axioma B.1) e $d(X, Z) = d(X, Y) + d(Y, Z)$ (Axioma B.5), podemos caracterizar a distância entre dois pontos de um eixo em termos de suas coordenadas.

Teorema C.1. *Se x e y são respectivamente as coordenadas dos pontos X e Y de um eixo, então*

$$d(X, Y) = |x - y| = |y - x|.$$

Demonstração. Seja O a origem do sistema de coordenadas. Se $X = Y$ ou $X = O$ ou $Y = O$, é fácil verificar o resultado. Suponha, então, que X, Y e O são pontos distintos. Sem perda de generalidade, suponha que X está à esquerda de Y. Há três casos a considerar:

(a) *X e Y estão à equerda da origem, ou seja, $x < y < 0$. (Figura C.1.)*

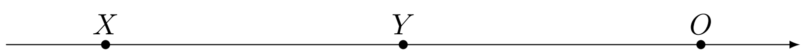


Figura C.1: Primeiro caso: X e Y à equerda da origem.

Neste caso, Y está entre X e O. Como $d(O, X) = -x$ e $d(O, Y) = -y$, segue-se que $d(X, O) = d(X, Y) + d(Y, O)$, ou seja, $d(X, Y) = d(O, X) - d(O, Y) = -x + y = |y - x|$.

(b) *X e Y estão em lados opostos em relação à origem, ou seja, $x < 0 < y$. (Figura C.2.)*

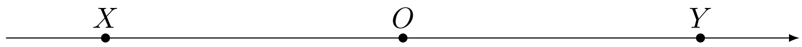


Figura C.2: Segundo caso: X e Y em lados opostos da origem.

Neste caso, O está entre X e Y, sendo agora $d(O, X) = -x$ e $d(O, Y) = y$. Então $d(X, Y) = d(X, O) + d(O, Y)$, ou seja, $d(X, Y) = d(O, X) + d(O, Y) = -x + y = |y - x|$.

(c) *X e Y estão à direita da origem, ou seja, $0 < x < y$. (Figura C.3.)*

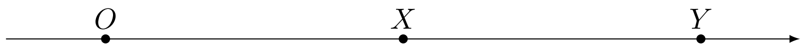


Figura C.3: Terceiro caso: X e Y à direita da origem.

Neste caso, X está entre O e Y, com $d(O, X) = x$ e $d(O, Y) = y$. Então $d(O, Y) = d(O, X) + d(X, Y)$, ou seja, $d(X, Y) = d(O, Y) - d(O, X) = y - x = |y - x|$.

Se X estiver à direita de Y, a demonstração se faz de modo análogo. ∎

Observação C.2. Note que a distância entre dois pontos de um eixo E é uma função: $d : E^2 \to \mathbb{R}$, $d(X, Y) = |x - y|$, em que x e y são respectivamente as coordenadas dos pontos X e Y de E. ○

Plano

Seja um plano Π munido de um sistema de eixos ortogonal OXY. Com o resultado do Teorema C.1 e utilizando o Teorema de Pitágoras vamos provar o seguinte teorema.

Teorema C.3. *Se (x_1, y_1) e (x_2, y_2) são respectivamente as coordenadas dos pontos P_1 e P_2 de Π, então*

$$d(P_1, P_2) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}.$$

Demonstração. Seja $P_3 = (x_2, y_1)$ um ponto auxiliar (Figura C.4). Como P_1 e P_3 têm mesma ordenada, o segmento P_1P_3 é paralelo ao eixo OX; analogamente, P_2P_3 é paralelo a OY. Logo, $P_1P_2P_3$ é um triângulo retângulo cujos catetos são P_1P_3 e P_2P_3 e medem respectivamente $|x_1 - x_2|$ e $|y_1 - y_2|$ (Teorema C.1). Portanto, pelo Teorema de Pitágoras:

$$d(P_1, P_2)^2 = (x_1 - x_2)^2 + (y_1 - y_2)^2,$$

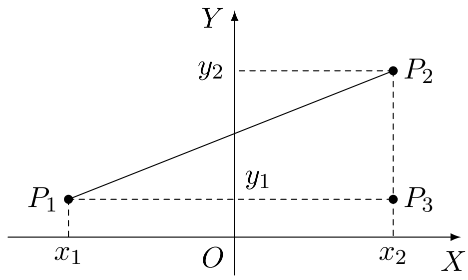


Figura C.4: Distância entre dois pontos no plano.

ou seja, como deve ser $d(P_1, P_2) \geqslant 0$, segue-se que:

$$d(P_1, P_2) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \qquad \blacksquare$$

Observação C.4. Cabe aqui uma observação análoga à Observação C.2, ou seja, é uma função $d : \Pi^2 \to \mathbb{R}$, $d(P_1, P_2) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$, em que (x_1, y_1) e (x_2, y_2) são respectivamente as coordenadas dos pontos P_1 e P_2 de Π. ○

Espaço

Seja o espaço $\mathcal{E}$ da Geometria Euclidiana munido de um sistema de eixos ortogonal $OXYZ$. Com os resultados dos dois últimos Teoremas (C.1 e C.3) e utilizando o Teorema de Pitágoras vamos provar o seguinte teorema.

Teorema C.5. *Se (x_1, y_1, z_1) e (x_2, y_2, z_2) são respectivamente as coordenadas dos pontos P_1 e P_2 de $\mathcal{E}$, então*

$$d(P_1, P_2) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2}.$$

Demonstração. Se P_1 e P_2 estiverem sobre uma reta paralela a um dos eixos, aplica-se o resultado do Teorema C.1 e não há o que demonstrar. Então, vamos supor que P_1 e P_2 não estão nessa condição. Para calcular a distância desses pontos, vamos considerar os pontos auxiliares (Figura C.5):

$$P_3 = (x_1, y_1, z_3), \quad P_4 = (x_2, y_2, 0), \quad P_5 = (x_1, y_1, 0), \quad P_4 = (x_1, y_2, 0).$$

Como P_4 e P_6 estão sobre uma reta paralela a um dos eixos, assim como P_5 e P_6,

$$d(P_4, P_6) = |x_1 - x_2| \quad \text{e} \quad d(P_5, P_6) = |y_1 - y_2|.$$

Aplicando o Teorema de Pitágoras ao triângulo $P_4P_5P_6$, obtemos:

$$d(P_4, P_5)^2 = d(P_4, P_6)^2 + d(P_5, P_6)^2 = |x_1 - x_2|^2 + |y_1 - y_2|^2 = (x_1 - x_2)^2 + (y_1 - y_2)^2.$$

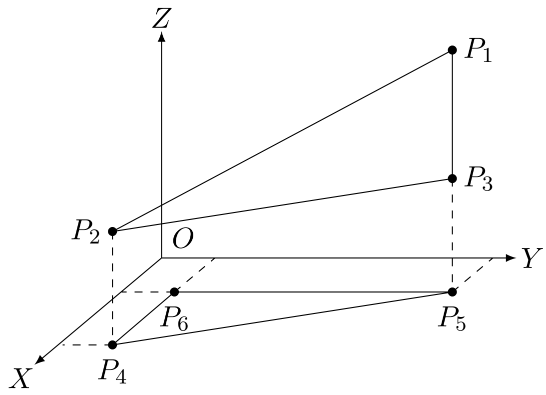


Figura C.5: Distância entre dois pontos no espaço.

Como os segmetos P_2P_3 e P_4P_5 são lados opostos de um retângulo, $d(P_2,\, P_3) = d(P_4,\, P_5)$, logo

$$d(P_2,\, P_3)^2 = (x_1 - x_2)^2 + (y_1 - y_2)^2,$$

e também, como P_1 e P_3 estão sobre uma mesma reta paralela a OZ,

$$d(P_1,\, P_3) = |z_1 - z_2|.$$

Aplicando novamente o Teorema de Pitágoras, dessa vez ao triângulo $P_1P_2P_3$, obtemos:

$$d(P_1,\, P_2)^2 = d(P_2,\, P_3)^2 + d(P_1,\, P_3)^2 = (x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2,$$

ou seja, como $d(P_1,\, P_2) \geqslant 0$,

$$d(P_1,\, P_2) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2}. \qquad \blacksquare$$

Observação C.6. Novamente temos uma função,

$$d : \mathcal{E}^2 \to \mathbb{R},\; d(P_1,\, P_2) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2},$$

em que $(x_1,\, y_1,\, z_1)$ e $(x_2,\, y_2,\, z_2)$ são respectivamente as coordenadas dos pontos P_1 e P_2 de $\mathcal{E}$. ○

C.2 Generalizando distâncias: a noção de métrica

Até o momento, vimos as noções de distância entre pontos da reta real $\mathbb{R}$, do plano bidimensional $\mathbb{R}^2$ e do espaço tridimensional $\mathbb{R}^3$, em relação às quais estamos razoavelmente familiarizados. Como o estudante terá oportunidade de reconhecer, boa parte do material tratado em cursos de cálculo de funções de uma ou várias variáveis, reais ou complexas, como os conceitos de derivação e integração, assentam-se sobre as definições de convergência e limite, que, por sua vez, assentam-se sobre a noção intuitiva de distância entre pontos. Por exemplo, dizemos que uma sequência x_n de pontos da reta real converge a um ponto x se a distância $|x_n - x|$ entre x_n e x torna-se cada vez menor à medida que n cresce.

Por um lado, os familiares espaços (conjuntos) de dimensão finita $\mathbb{R}$, $\mathbb{R}^2$ e $\mathbb{R}^3$ não são (e estão longe de ser) os únicos objetos aos quais a noção intuitiva de distância é útil para o seu estudo; então seria interessante dispor de uma maneira de lidar com esse conceito que também permitisse aplicá-lo a outros tipos de conjuntos. Por outro lado, razões evolutivas impedem que o cérebro humano produza e desenvolva imagens que não em uma, duas ou três dimensões; portanto, para o estudo de espaços com mais dimensões (finitas ou não) faz-se necessário dispor de instrumentos que permitam desenvolver raciocínios os mais próximos possíveis daqueles empregados em espaços de dimensão 1, 2 ou 3.

O reconhecimento da importância de abstrair e generalizar a noção de distância, de modo a aplicá-la a outros tipos de conjunto que não os familiares, foi o que conduziu às noções de *métrica* e *espaços métricos*, que serão definidos a seguir, e permitiu aplicar muitas das noções geométricas e instrumentos analíticos, originalmente desenvolvidos em espaços mais familiares, a conjuntos menos acessíveis à intuição, como os espaços vetoriais de dimensão infinita, por exemplo, os espaços de funções ou de sequências.

Os conceitos de *métrica* e *espaços métricos* foram introduzidas por Fréchet[1] em sua tese de doutorado *Sur quelques points du calcul fonctionnel*, apresentada em 1906 sob a supervisão de Hadamard[2] à *École Normale Supérieure* em Paris. A expressão *espaço métrico*, no entanto, não foi sua invenção, tendo sido cunhada por Hausdorff[3] em 1914.

A questão relevante que se coloca afinal é: quais as propriedades básicas que a noção intuitiva de distância possui a fim de que possa ser empregada em diversas instâncias? O desenvolvimento da Matemática mostrou que a resposta consiste num conjunto de quatro propriedades. Essas propriedades definem a noção de *métrica*, que abstrai e generaliza a noção intuitiva de distância. Vejamos.

[1]René Maurice Fréchet (1878–1973), matemático francês, famoso por suas contribuições à Topologia Geral e introdução dos espaços abstratos.

[2]Jacques Salomon Hadamard (1865–1963), matemático francês, famoso pela prova do seu teorema sobre números primos, em que afirma que os números primos menores do que n tendem ao infinito mais rápido do que $n/\ln n$.

[3]Felix Hausdorff (1868–1942). Desenvolveu a teoria de espaços métricos e topológicos. Trabalhou na Teoria dos Conjuntos e introduziu o conceito de conjunto parcialmente ordenado.

Definição C.7. *Seja X um conjunto*[a]*. Uma função $d : X^2 \to \mathbb{R}$ é dita ser uma* métrica *em X se possuir as seguintes propriedades:*

1. *positividade: $d(x,\, y) \geqslant 0$, $\forall x,\, y \in X$;*

2. *condição de distância nula: $d(x,\, y) = 0 \Leftrightarrow x = y$, $\forall x,\, y \in X$;*

3. *simetria: $d(x,\, y) = d(y,\, x)$, $\forall x,\, y \in X$;*

4. *desigualdade triangular: $d(x,\, z) \leqslant d(x,\, y) + d(y,\, z)$, $\forall x,\, y,\, z \in X$.*

[a] Inclusive $X = \varnothing$. Isso é simplesmente uma questão de gosto. Se excluirmos essa possibilidade, precisaremos provar que X não é vazio cada vez que definirmos um espaço métrico $(X,\, d)$. Por outro lado, se a considerarmos, precisaremos pensar sobre o conjunto vazio cada vez que iniciarmos alguma demonstração e será preciso dizer mais ou menos assim no início: "Se X é vazio, então o resultado segue por vacuidade, logo podemos supor que X não é vazio." Obviamente pode ocorrer de alguma proposição ser falsa para $X = \varnothing$ e, nesse caso, devemos mencionar essa exceção no corpo da proposição. Para o leitor menos experiente, apenas uma observação sobre o significado de "algum resultado seguir por vacuidade": para provar que algo sobre o conjunto vazio é verdadeiro, basta provar que não pode ser falso, por exemplo, se P é uma proposição referente aos elementos do conjunto vazio, para provar que P é falsa, é necessário existir algum elemento de $\varnothing$ que não goze de P, absurdo, pois $\varnothing$ não contém elemento algum.

Definição C.8. *Seja X um conjunto e d uma métrica em X. Dizemos que o par $(X,\, d)$ é um* espaço métrico*.*

Em outras palavras, um espaço métrico é um conjunto munido de uma métrica.

Cada propriedade da Definição C.7 possui uma interpretação intuitiva: a primeira diz que a distância de um lugar a outro nunca é negativa; a segunda, que a distância de um lugar a ele mesmo é nula; a terceira, que ir de x a y não é mais fácil nem mais difícil do quer ir de y a x; a quarta propriedade, de grande importância, afirma que ir de x para y e depois para z não pode resultar num caminho mais curto do que a rota direta de x para z e seu nome se deve ao seu significado geométrico nos espaços $\mathbb{R}^2$ e $\mathbb{R}^3$ com a métrica usual, i.e., a soma dos comprimentos de dois lados de um triângulo não é menor do que o comprimento do terceiro lado. As quatro propriedades da definição são aquelas identificadas como essenciais à noção intuitiva de distância e qualquer função d que as possua, ou seja, qualquer métrica, pode potencialmente ser empregada como equivalente a essa noção.

A propriedade de positividade é, na verdade, reduntante, como mostra a seguinte proposição.

Proposição C.9. *Se $d : X^2 \to \mathbb{R}$ é uma função tal que*

1. $d(x,\, y) = 0 \Leftrightarrow x = y$, $\forall x,\, y \in X$*;*

2. $d(x,\, y) = d(y,\, x)$, $\forall x,\, y \in X$*;*

3. $d(x,\, z) \leqslant d(x,\, y) + d(y,\, z)$, $\forall x,\, y,\, z \in X$*;*

então d é uma métrica em X.

Demonstração. Fazendo $z = x$ na condição 3 e usando 1 e 2, obtemos:

$$0 = d(x,\, x) \leqslant d(x,\, y) + d(y,\, x) = 2d(x,\, y),$$

ou seja, $d(x,\, y) \geqslant 0$; como x e y são abritrários, $d(x,\, y) \geqslant 0$, $\forall x,\, y \in X$. Essa propriedade, em conjunto com as propriedades 1 a 3, fazem de d, conforme a Definição C.7, uma métrica em X. ∎

Na verdade, a condição de simetria também poderia ser dispensada dependendo da forma como exibimos a desigualdade triangular. É o que mostra a seguinte proposição.

Proposição C.10. *Se $d : X^2 \to \mathbb{R}$ é uma função tal que*

1. $d(x,\, y) = 0 \Leftrightarrow x = y$, $\forall x,\, y \in X$*;*

2. $d(x,\, z) \leqslant d(x,\, y) + d(z,\, y)$, $\forall x,\, y,\, z \in X$*;*

então d é uma métrica em X.

Demonstração. Fazendo $y = x$ em 2 e utilizando 1, obtemos $d(x,\, z) \leqslant d(z,\, x)$, com x e z quaisquer em X; como x e z são arbitrários, podemos trocar seus papeis, obtendo $d(z,\, x) \leqslant d(x,\, z)$, ou seja,

$$d(x,\, z) = d(z,\, x),\ \forall x,\, z \in X,$$

estabelecendo assim a condição de simetria.

Agora, utilizando-se a condição de simetria e a desigualdade triangular, vamos estabelecer a condição de positividade. Para isso, vamos provar o seguinte fato mais forte:

$$d(x,\, y) \geqslant |d(x,\, z) - d(z,\, y)|,\ \forall x,\, y,\, z \in X, \tag{C.2.1}$$

que, em particular, garante que $d(x,\, y) \geqslant 0$, $\forall x,\, y \in X$. Para tanto, basta notar que, pela desigualdade triangular e pela condição de simetria, temos $d(x,\, z) \leqslant d(x,\, y) + d(y,\, z)$, $\forall x,\, y,\, z \in X$, ou seja,

$$d(x,\, y) \geqslant d(x,\, z) - d(y,\, z),\ \forall x,\, y,\, z \in X;$$

agora, trocando-se x por y (e vice-versa) e utilizando-se novamente a condição de simetria, resulta

$$d(x,\, y) = d(y,\, x) \geqslant d(y,\, z) - d(x,\, z),\ \forall x,\, y,\, z \in X.$$

Em conjunto, essas duas últimas inequações nos garantem a validade de (C.2.1), como queríamos mostrar.

As propriedades de simetria e positividade juntamente com os itens 1 e 2 fazem de d uma métrica em X. ■

Dessa forma, a definição de métrica pode ser reduzida à Proposição C.10. Embora reduntantes, a positividade e a simetria são incluídas na Definição C.7 apenas para enfatizar sua importância como propriedades das métricas.

Observação C.11. Note a diferença entre a condição 4 da Definição C.7 e a condição 2 da Proposição C.10: na definição aparece $d(y,\, z)$, na proposição, $d(z,\, y)$. Essa diferença que nos permitiu demonstrar a condição de simetria na proposição. ○

Exemplo C.12. O exemplo mais básico de uma métrica é o da função $d : \mathbb{R}^2 \to \mathbb{R}$, $d(x,\, y) = |y - x|$. Outro exemplo básico é oferecido pela função $d : \mathbb{C}^2 \to \mathbb{R}$, $d(z,\, w) = |z - w|$. Essas são as chamadas *métricas usuais em* $\mathbb{R}$ *e* $\mathbb{C}$, respectivamente. Fica como exercício a simples tarefa de verificar que essas funções satisfazem as condições da definição de métrica. ◇

Exemplo C.13. As funções dos Teoremas C.3 e C.5 são as métricas usuais em $\mathbb{R}^2$ e $\mathbb{R}^3$, respectivamente. De maneira geral, a função $d_2 : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, definida por:

$$d_2(x,\, y) := \sqrt{(y_1 - x_1)^2 + \cdots + (y_n - x_n)^2} = \sqrt{\sum_{i=1}^{n} (y_i - x_i)^2},$$

em que $x = (x_1,\, \ldots,\, x_n)$ e $y = (y_1,\, \ldots,\, y_n)$, é a métrica usual em $\mathbb{R}^n$ (também conhecida como métrica Euclidiana). ◇

Exemplo C.14. Imaginando o plano $\mathbb{R}^2$ como a planta de uma cidade cujas ruas são retas paralelas aos eixos coordenados, o menor caminho ligando $x = (x_1,\, x_2)$ a $y = (y_1,\, y_2)$ através das ruas tem comprimento $|x_1 - y_1| + |x_2 - y_2|$. Fica como exercício ao leitor verificar que $d : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ definida por

$$d(x,\, y) := |x_1 - y_1| + |x_2 - y_2|,$$

é de fato uma métrica. Essa métrica é conhecida como *métrica do táxi* ou *métrica do taxista* ou, ainda, *métrica de Manhattan* (em alusão ao formato quadriculado de grande

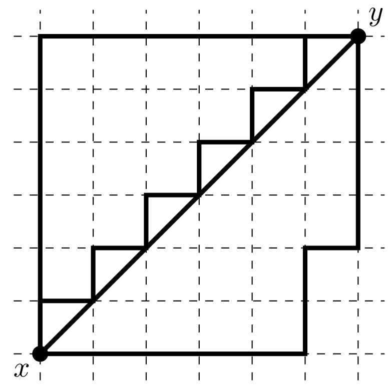


Figura C.6: Comparação entre a métrica do taxista e a métrica usual em $\mathbb{R}^2$.

parte das ruas na ilha de Manhattan) e é geralmente atribuída a Minkowski[4]. Veja a figura C.6 para uma comparação intuitiva entre a métrica do taxista e a métrica usual ambas em $\mathbb{R}^2$. Ela pode ser generalizada para pontos do $\mathbb{R}^n$ da seguinte maneira: sejam $x = (x_1, \ldots, x_n)$ e $y = (y_1, \ldots, y_n)$; definimos $d_1 : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ por:

$$d_1(x, y) := |x_1 - y_1| + \cdots + |x_n - y_n| = \sum_{i=1}^{n} |x_i - y_i|.$$

Exercício: demonstrar que d_1 é uma métrica em $\mathbb{R}^n$ e comparar com a métrica usual d_2 provando que $d_2(x, y) \leqslant d_1(x, y)$, $\forall x, y \in \mathbb{R}^n$ (dica: compare $[d_2(x, y)]^2$ com $[d_1(x, y)]^2$). ◇

Exercícios

Exercício C.15. Compare o Axioma B.1 com as propriedades 1 a 3 da definição de métrica (Definição C.7) e o Axioma B.5 com a propriedade 4 dessa definição. ◇

Exercício C.16 (Adaptado de [5]). Se desejamos ir da cidade de Santo André – SP à cidade de São Paulo, podemos estar interessados em um ou mais dos seguintes números:

(a) a distância, em quilômetros, de Santo André a São Paulo em linha reta;

[4] Hermann Minkowski (1864–1909), matemático alemão, desenvolveu uma nova visão do espaço e do tempo e lançou as bases matemáticas da teoria da relatividade. Trabalhou por alguns anos com Hilbert e outros notáveis matemáticos da época. Einstein foi sei aluno em vários cursos no Instituto Politécnico de Zurique.

(b) a distância, em quilômetros, do percurso mais curto de Santo André a São Paulo de carro;

(c) o tempo, em minutos, do percurso mais curto de Santo André a São Paulo de trem e metrô;

(d) o custo, em reais, do trajeto mais barato de Santo André a São Paulo de trem e metrô.

Cada um desses números pode ser de interesse para quem deseja se deslocar de uma cidade para a outra e nenhum deles é facilmente obtido a partir do outro.

Seja X o conjunto de cidades do estado de São Paulo. Considere d correspondendo aos itens (a) a (d). Discuta se os itens 1 a 4 da Definição C.7 se aplicam.

(Uma questão aberta como esta será mais útil se abordada com um espírito de boa vontade.) $\diamond$

Exemplo C.17 (Adaptado de [5]). Seja $X = \{\triangle, \bigcirc, \square\}$ com $\triangle$, $\bigcirc$ e $\square$ distintos. Escreva funções $d_i : X^2 \to \mathbb{R}$ satisfazendo a condição 1 da Definição C.7 tal que:

1. d_1 satisfaça as condições 2 e 3, mas não a 4;

2. d_2 satisfaça as condições 3, 4 e $d_2(x, y) = 0 \Rightarrow x = y$, mas $x = y \not\Rightarrow d_2(x, y) = 0$;

3. d_3 satisfaça as condições 3, 4 e $x = y \Rightarrow d_3(x, y) = 0$, mas $d_3(x, y) = 0 \not\Rightarrow x = y$;

4. d_4 satisfaça as condições 2 e 4, mas não a 3.

Você deve provar suas afirmações.

Observação: d_3 é dita ser uma *pseudo-métrica* em X. $\diamond$

Exercício C.18. Seja X um conjunto e considere uma função $f : X^2 \to \mathbb{R}$ com as seguintes propriedades:

1. $f(x, y) \geqslant 0, \forall x, y \in X$;

2. $f(x, y) = 0 \Leftrightarrow x = y, \forall x, y \in X$;

3. $f(x, z) \leqslant f(x, y) + f(y, z), \forall x, y, z \in X$.

Mostre que se $d(x, y) = f(x, y) + f(y, x)$, então (X, d) é um espaço métrico. $\diamond$

Exercício C.19. Seja X um conjunto e considere a seguinte função $d_t : X^2 \to \mathbb{R}$:

$$d_t(x, y) := \begin{cases} 0 & \text{se } x = y \\ 1 & \text{se } x \neq y \end{cases} .$$

Mostre que d_t é uma métrica em X. Essa métrica é denominada *métrica trivial.*

Esse é um exemplo simples e interessante que não tem uma interpretação clara de distância como o das métricas anteriores. Apesar de esquisita, essa métrica serve para treinar e ilustrar a generalidade dos conceitos. $\diamond$

Exercício C.20. Verifique que a função $d_\infty : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$,

$$d_\infty(x, y) := \max\{|y_1 - x_1|, \ldots, |y_n - x_n|\} = \max_{1 \leqslant i \leqslant n} |y_i - x_i|,$$

em que $x = (x_1, \ldots, x_n)$ e $y = (y_1, \ldots, y_n)$, é uma métrica em $\mathbb{R}^n$. Compare-a com a métrica do taxista e a métrica usual em $\mathbb{R}^n$, mostrando que

$$d_\infty(x, y) \leqslant d_2(x, y) \leqslant d_1(x, y) \leqslant n d_\infty(x, y),\ \forall x, y \in \mathbb{R}^n.$$

A segunda desigualdade você já deve ter verificado no Exemplo C.14. $\diamond$

O intuito desta pequena seção é fornecer ao estudante de início de graduação uma noção mais ampla do conceito de distância, para além daquela usual nos espaços de dimensão 1, 2 ou 3.

Adiantamos que é essa noção que permite desenvolver conceitos como os de convergência e limite de sequências em espaços além dos usuais e, eventualmente, prosseguir desenvolvendo em tais espaços outros ingredientes do Cálculo e da Análise.

Que esta pequena introdução sirva para estimular seus estudos e aguçar sua curiosidade sobre o assunto!

D Eixos não ortogonais no plano

Como dito na Seção 1.5.2, o uso de um par de eixos ortogonais não é a única maneira de se estabelecer uma correspondência entre pontos do plano e pares ordenados de números reais. Vimos as coordenadas polares e comentamos que poderíamos utilizar um par qualquer de eixos concorrentes e o processo descrito no início daquela seção (ou, se preferir, o processo detalhado da demonstração do Teorema B.15). Embora um tal sistema não seja utilizado na maior parte dos casos, pode haver situações em que ele seja vantajoso.

Quando se faz tal opção (por eixos não ortogonais), uma surpresa aparece: um ponto pode ter dois tipos de coordenadas.

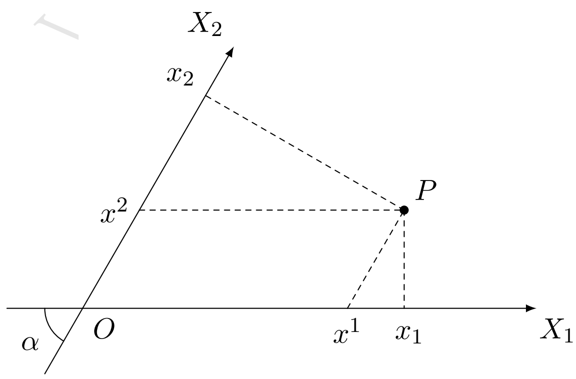


Figura D.1: Eixos não ortogonais e coordenadas covariantes e contravariantes.

Isso ocorre porque, como mostra a Figura D.1, podemos identificar um ponto tanto por suas projeções ortogonais quanto por suas projeções paralelas aos eixos (em eixos ortogonais, essas duas projeções coincidem). Assim, o ponto P, referido a eixos oblíquos, pode ser caracterizado tanto pelo par ordenado de números reais (x^1, x^2), obtidos pelas projeções de P paralelamente aos eixos, quanto pelo par ordenado de números reais (x_1, x_2), obtidos pelas projeções ortogonais. Esses pares ordenados recebem, respectivamente, o nome de *coordenadas contravariantes* e *coordenadas covariantes*[1]. A relação entre elas pode ser obtida facilmente a partir da Figura D.1:

$$x_1 = x^1 + x^2 \cos\alpha, \tag{D.0.1}$$

$$x_2 = x^2 + x^1 \cos\alpha, \tag{D.0.2}$$

assim como a relação inversa:

$$x^1 = \frac{1}{\operatorname{sen}^2\alpha}(x_1 - x_2 \cos\alpha), \tag{D.0.3}$$

$$x^2 = \frac{1}{\operatorname{sen}^2\alpha}(x_2 - x_1 \cos\alpha), \tag{D.0.4}$$

onde α é o ângulo entre os semieixos OX_1 e OX_2.

Observação D.1. Na notação adotada aqui, os índices sobrescritos não são expoentes. Por exemplo, x^2 deve ser lido "x-dois", não "x ao quadrado", e a notação (x^1, x^2) será equivalente à tradicional (x, y). ○

Exercício D.2. Obtenha as equações (D.0.1) e (D.0.2). Analise todos os casos, i.e., considere o ponto P em cada quadrante. ◇

Exercício D.3. Obtenha as equações (D.0.3) e (D.0.4). ◇

Lidar com esse sistema de coordenadas não retangulares afeta todas as propriedades ligadas ao conceito de distância. Pela lei dos cossenos, podemos obter a distância entre o ponto P e a origem O em termos das coordenadas contravariantes:

$$d(O,\, P)^2 = (x^1)^2 + (x^2)^2 + 2x^1x^2\cos\alpha = \sum_{i,\,j} g_{ij}x^ix^j, \tag{D.0.5}$$

onde $i,\, j \in \{1,\, 2\}$ e g_{ij} assume os valores tabulados na seguinte matriz:

$$(g_{ij}) = \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix} = \begin{pmatrix} 1 & \cos\alpha \\ \cos\alpha & 1 \end{pmatrix}.$$

Também podemos obter a distância de O a P em termo das coordenadas covariantes. Basta utilizar as equações (D.0.3) e (D.0.4) na equação (D.0.5):

$$d(O,\, P)^2 = \frac{1}{\operatorname{sen}^2\alpha}[(x_1)^2 + (x_2)^2 - 2x_1x_2\cos\alpha] = \sum_{i,\,j} g^{ij}x_ix_j, \tag{D.0.6}$$

[1]Esses termos foram introduzidos por James Joseph Sylvester (1814 – 1897) em 1853.

em que i, $j \in \{1,\, 2\}$ e g^{ij} assume os valores tabulados na seguinte matriz:

$$(g^{ij}) = \begin{pmatrix} g^{11} & g^{12} \\ g^{21} & g^{22} \end{pmatrix} = \frac{1}{\operatorname{sen}^4\alpha} \begin{pmatrix} 1 & -\cos\alpha \\ -\cos\alpha & 1 \end{pmatrix}.$$

Exercício D.4. Complete os detalhes para obter a equação (D.0.6). $\diamond$

E Princípio da Indução

A seguir, um breve resumo do Princípio da Indução para entender as demonstrações feitas no texto envolvendo os números naturais. Para um estudo mais detalhado, consulte por exemplo o brilhante artigo [8] ou o Capítulo 2 do livro-texto [10], ambos do Prof. Elon Lages Lima.

Esse princípio é um eficiente instrumento para a demonstração de fatos que dizem respeito aos números naturais. Entendê-lo é praticamente o mesmo que compreender os números naturais e a sua estrutura.

Ele se assenta sobre quatro axiomas, conhecidos como *axiomas de Peano*[1], dos quais resultam, como consequências lógicas, todas as demais afirmações verdadeiras que se podem fazer sobre os números naturais.

Esses axiomas são os seguintes:

Axioma E.1. *Exite uma função* $s : \mathbb{N} \to \mathbb{N}$ *que associa cada* $n \in \mathbb{N}$ *a um elemento* $s(n) \in \mathbb{N}$ *chamado o sucessor de* n. □

Axioma E.2. *A função* $s : \mathbb{N} \to \mathbb{N}$ *é injetiva.* □

Axioma E.3. *Existe um único elemento no conjunto* $\mathbb{N}$*, representado pelo símbolo* 1 *e chamado "um", tal que* $1 \neq s(n)$ *para todo* $n \in \mathbb{N}$. □

Axioma E.4. *Se* $X \subset \mathbb{N}$ *é tal que* $1 \in X$ *e* $s(X) \subset X$ *(i.e.:* $n \in X \Rightarrow s(n) \in X$*), então* $X = \mathbb{N}$. □

Na definição da operação de adição de números naturais, conclui-se que $s(n) = n+1$. Veja, por exemplo, [8].

Dispomos de um sistema de numeração que nos permite representar, mediante o uso apropriado dos símbolos 0, 1, 2, 3, 4, 5, 6, 7, 8 e 9, todos os números naturais: $2 = s(1)$, $3 = s(2)$, $4 = s(3)$, $5 = s(4)$, etc.

A igualdade $2 = s(1)$ significa apenas que estamos usando o símbolo 2 para representar o sucessor de 1, a igualdade $3 = s(2)$ significa que estamos utilizando o símbolo

[1]Em horna a Giuseppe Peano (1858–1932).

3 para representar o sucessor de 2 (ou seja, o sucessor do sucessor de 1) e assim por diante. A sequência dos números naturais pode ser indicada assim:

$$1 \xrightarrow{s} 2 \xrightarrow{s} 3 \xrightarrow{s} 4 \xrightarrow{s} 5 \xrightarrow{s} \cdots$$

As flechas ligam cada número ao seu sucessor e nenhuma flecha aponta para 1, pois esse número não é sucessor de nenhum outro. O diagrama acima diz muito sobre a estrutura do conjunto $\mathbb{N}$ dos números naturais.

Como este apêndice consiste apenas num breve resumo, não vamos desenvolver as consequências dos axiomas de Peano (por isso incentivamos fortemente as leituras indicadas no primeiro parágrafo), mantendo o nosso foco apenas no último deles, que possui uma natureza mais elaborada do que a dos demais.

O último axioma acima é conhecido como *axioma da indução.* Informalmente, seu significado é o seguinte: a partir de 1 e aplicando-se repetidamente a operação de se tomar o sucessor, podemos obter qualquer número natural.

Seu significado e utilidade ficam mais claros após a exibição de um exemplo em que a função s é modificada e essa estrutura dos números naturais é perdida. Vejamos.

Exemplo E.5. Seja $s : \mathbb{N} \to \mathbb{N}$, tal que $s(n) = n + 2$. Então, se começarmos com 1 e aplicarmos repetidas vezes a operação de tomar o "sucessor" nesta nova acepção, obteremos $s(1) = 3$, $s(3) = 5$, etc., de maneira que nunca chegaremos a qualquer número par. O diagrama

$$1 \xrightarrow{s} 3 \xrightarrow{s} 5 \xrightarrow{s} \cdots 2 \xrightarrow{s} 4 \xrightarrow{s} 6 \xrightarrow{s} \cdots$$

exibe uma função injetiva $s : \mathbb{N} \to \mathbb{N}$ para a qual não é verdade que todo número natural n pode ser obtido a partir de 1 mediante repetidas aplicações da operação de passar de k para $s(k)$. $\diamond$

Um comentário sobre a nomenclatura: um subconjunto $X \subset \mathbb{N}$ chama-se *indutivo* quando $s(X) \subset X$, ou seja, quando $n \in X \Rightarrow s(n) \in X$ ou, ainda, quando o sucessor de qualquer elemento de X também pertence a X.

O axioma da indução afirma que o único conjunto indutivo de $\mathbb{N}$ que contém 1 é o próprio $\mathbb{N}$.

O ponto alto dessa discussão toda é o fato de que o axioma da indução pode ser visto como um método de demonstração, chamado *Método de Indução Matemática*, ou *Princípio da Indução Finita*, ou ainda *Princípio da Indução*, desempenhando papel fundamental na teoria dos números naturais e, de modo mais geral, em toda a Matemática. Explicamos.

Considere P uma propriedade que se refere a números naturais. Um certo número natural pode ou não gozar de P. Exemplo: se P for a propriedade de possuir um sucessor, 1 não goza de P, mas todos os demais números naturais sim.

Dito isso, uma das formas de se enunciar o Princípio da Indução é a seguinte:

Proposição E.6 (Princípio da Indução). *Seja P uma propriedade referente a números naturais. Se 1 goza de P e se, além disso, o fato de o número natural n gozar de P implica que seu sucessor $s(n)$ também goza, então todos os números naturais gozam da propriedade P.*

Para averiguar sua veracidade (uma vez admitidos os axiomas de Peano), basta observar que, dada a propriedade P cumprindo as condições estipuladas no enunciado acima, o conjunto X dos números naturais que gozam de P contém 1 e é indutivo. Logo $X = \mathbb{N}$, i.e., todo número natural goza de P.

Nas demonstrações por indução, a hipótese de que a propriedade P é válida para o número natural n (da qual deve decorrer que P vale também para $s(n)$) chama-se *hipótese de indução.*

As propriedades básicas dos números naturais são demonstradas por indução. Vejamos um exemplo bem simples.

Exemplo E.7 ([8]). Seja P a afirmação de que todo número natural é diferente do seu sucessor. Dado o número natural n, escrevamos $P(n)$ para significar, abreviadamente, a afirmação $n \neq s(n)$. Então $P(1)$ é verdadeira, pois $1 \neq s(1)$, já que 1 não é sucessor de número algum; em particular, 1 não é sucessor de si próprio. Além disso, supondo $P(n)$ verdadeira, i.e., se admitirmos que $n \neq s(n)$, então $s(n) \neq s(s(n))$, pois a função $s : \mathbb{N} \to \mathbb{N}$ é injetiva. Mas a afirmação $s(n) \neq s(s(n))$ significa que $P(s(n))$ é verdadeira. Assim, a verdade de $P(n)$ acarreta a verdade de $P(s(n))$. Pelo Princípio da Indução, todos os números naturais gozam de P, ou seja, são diferentes de seus sucessores. $\diamond$

A seguir um exemplo de como não utilizar a indução.

Exemplo E.8. Quer-se provar que todas as pessoas têm a mesma altura. Em outras palavras, quer-se provar que se X é um conjunto de n ($n \geqslant 1$) pessoas, então todos os elementos de X têm a mesma altura. Evidentemente, se $n = 1$ a afirmação é verdadeira, pois se X é um conjunto unitário, todos os seus elementos têm a mesma altura. Suponhamos agora que a afirmação seja verdadeira para todos os conjuntos de n elementos. Consideremos um conjunto com $n + 1$ pessoas, $\{a_1, \ldots, a_{n+1}\}$. Ora, $\{a_1, \ldots, a_n\}$ é um conjunto de n pessoas, logo $a_1, \ldots, a_n$ têm a mesma altura. Da mesma forma, $\{a_2, \ldots, a_{n+1}\}$ também é um conjunto de n pessoas, logo todos os seus elementos têm a mesma altura, em particular a_n e a_{n+1}. Como $a_1, \ldots, a_n$ têm a mesma altura e a_n e a_{n+1} têm a mesma altura, segue-se que todos os elementos de $\{a_1, \ldots, a_{n+1}\}$ têm a mesma altura, como queria-se provar. $\diamond$

O erro desse exemplo consiste na passagem $P(n) \Rightarrow P(n + 1)$, que é falsa quando $n = 1$. Não é verdade que $P(1) \Rightarrow P(2)$, ou, para dizer mais exatamente, $P(2)$ certamente é falsa. O raciocínio empregado supõe implicitamente que X tem pelo menos 3 elementos. (Releia o exemplo substituindo n por 1 e isso ficará patente.)

Há uma forma equivalente à Proposição E.6, que é por vezes mais conveniente do que ela própria quando desejamos provar algumas proposições. Trata-se do Princípio da Boa Ordenação. Antes de enunciá-lo e demonstrá-lo, vamos precisar de duas definições e dois lemas.

Definição E.9. *Dados dois números naturais m e n, diremos que m é menor do que n e escreveremos $m < n$ para significar que existe $p \in \mathbb{N}$ tal que $n = m + p$. Neste caso, diremos também que n é maior do que m e escreveremos $n > m$ para exprimir que $m < n$.*

Definição E.10. *Dado $X \subset \mathbb{N}$, diz-se que o número natural n é o* menor *(ou* primeiro*) elemento de X quando $n \in X$ e, além disso, $n \leqslant x$ para todo $x \in X$.*

Lema E.11. *O número 1 é o menor dos naturais, i.e., $n \neq 1 \Rightarrow n > 1$.*

Demonstração. Pelo Axioma E.3, $n \neq 1$ implica que n é sucessor de algum número natural m, ou seja, $n = m + 1 = 1 + m$, logo $n > 1$. ■

Lema E.12. *Não existem números naturais entre um número natural e o seu sucessor.*

Demonstração. Se existisse m natural tal que $n < m < n + 1$, existiriam naturais p e q tais que $m = n + p$ e $n + 1 = m + q$. Logo, $n + 1 = n + p + q$, donde $1 = p + q$, ou seja, teríamos $p < 1$. Mas $p < 1$ significa $p \neq 1$, que, de acordo com o Lema E.11, implica $p > 1$, contradição. Portanto, não podem existir números naturais entre um número natural e o seu sucessor. ■

Na proposição a seguir, indicaremos por I_n o conjunto dos números naturais m tais que $1 \leqslant m \leqslant n$. Exemplos: $I_1 = \{1\}$, $I_2 = \{1, 2\}$, $I_3 = \{1, 2, 3\}$, etc.

Proposição E.13 (Princípio da Boa Ordenação [8]). *Todo subconjunto não vazio $X \subset \mathbb{N}$ possui um menor elemento.*

Demonstração. Sem perda de generalidade, podemos admitir que $1 \notin X$, do contrário 1 seria evidentemente o menor elemento de X (Lema E.11). O menor elemento de X, cuja existência queremos provar, deverá ser da forma $n + 1$. Devemos encontrar um número natural n tal que $n + 1 \in X$ e, além disso, todos os elementos de X sejam maiores do que n, logo maiores do que $1, 2, \ldots, n$. Noutras palavras, procuramos um número natural n tal que $I_n \subset \mathbb{N} \setminus X$ e $n + 1 \in X$. Com esse objetivo, consideramos o conjunto

$$A = \{n \in \mathbb{N};\ I_n \subset \mathbb{N} \setminus X\},$$

ou seja, todos os elementos de X são maiores do que n. Como estamos supondo que $1 \notin X$, sabemos que $1 \in A$. Por outro lado, como X não é vazio, nem todos os números naturais pertencem a A, ou seja, $A \neq \mathbb{N}$. Pelo Axioma E.4, vemos que A não é indutivo, i.e., deve existir um $n \in A$ tal que $n+1 \notin A$. Isso significa que todos os elementos de X são maiores do que n mas nem todos são maiores do que $n+1$. Como não há números naturais entre n e $n+1$ (Lema E.12), concluímos que $n+1$ pertence a X e é o seu menor elemento. ■

Na verdade, o Princípio da Boa Ordenação e o Princípio da Indução são equivalentes. Vimos que este implica aquele na demonstração acima. Resta mostrar a recíproca. Seja X um conjunto indutivo que contém 1. Vamos mostrar que $X = \mathbb{N}$. Para isso, suponha, por absurdo, que $X \neq \mathbb{N}$. Seja Y o complementar de X em relação a $\mathbb{N}$, ou seja, $Y = \mathbb{N} \setminus X$. Como estamos supondo que $X \neq \mathbb{N}$, segue-se que $Y \neq \emptyset$, logo, pelo Princípio da Boa Ordenação, Y possui um menor elemento, digamos y. Por outro lado, como $1 \notin Y$, pois $1 \in X$, segue-se que $y \neq 1$, logo y é sucessor de algum número natural x. Uma vez que y é o menor elemento de Y, $x \notin Y$, portanto devemos ter $x \in X$. Ora, mas X é indutivo, logo o sucessor de x, i.e., y, deve pertencer a X, contradição. Portanto, $X = \mathbb{N}$.

Exemplo E.14 (Teorema Fundamental da Aritmética [10]). Um número natural p chama-se *primo* quando não pode ser expresso como o produto mn de dois números naturais, a menos que um deles seja igual a 1 (e o outro, naturalmente, igual a p); isso equivale a dizer que os fatores m e n não podem ser ambos menores do que p. Um resultado fundamental em Aritmética afirma que todo número natural é primo ou é um produto de fatores primos. Provaremos essa assertiva por boa ordenação. Seja X o conjunto dos números naturais que são primos ou produtos de fatores primos. Observe que se m e n pertencem a X, então o produto mn também pertence a X. Seja Y o complementar de X. Assim, Y é conjunto dos números naturais que não são primos nem são produtos de fatores primos. Queremos provar que Y é vazio. Isso será feito por redução ao absurdo, como sempre se dá com demonstrações por boa ordenação. Com efeito, se Y não fosse vazio, haveria um menor elemento $y \in Y$. Então todos os números menores do que y pertenceriam a X. Como y não é primo, ter-se-ia $y = mn$, com $m < y$ e $n < y$, logo $m \in X$ e $n \in X$. Assim, $mn \in X$. Mas $mn = y$, o que daria $y \in X$, contradição. Logo $Y = \emptyset$, concluindo a demonstração. ◇

Exemplo E.15 ([8]). Toda função monótona não crescente $f : \mathbb{N} \to \mathbb{N}$ é constante a partir de um certo ponto. (I.e., existe $n_0 \in \mathbb{N}$ tal que $f(n) = f(n_0)$ para todo $n \geqslant n_0$.) Para demonstrar isso, considere $f(n_0)$ o menor elemento de $X = \{f(1), \ldots, f(n), \ldots\}$. Então $n > n_0 \Rightarrow f(n) \geqslant f(n_0)$, pois f é não crescente. Mas isso acarreta que $f(n) = f(n_0)$ para todo $n \geqslant n_0$, pois $f(n_0)$ é o menor elemento de X. ◇

Apenas como curiosidade, vale a pena mencionar um corolário da proposição provada no exemplo acima, a saber, que toda sequência estritamente decrescente $n_1 > n_2 > \cdots$ de números naturais é finita. Com efeito, do contrário, definindo-se $f : \mathbb{N} \to \mathbb{N}$ pela fórmula $f(k) = n_k$, obteríamos uma função monótona não crescente, que, de acordo com o teorema anterior, seria constante a partir de um certo ponto. Isso significa que nossa sequência seria constante a partir de um certo ponto, absurdo, pois, por hipótese, ela é estritamente decrescente. Assim, a suposição de que alguma sequência estritamente decrescente de números naturais é infinita leva a uma contradição. Portanto toda sequência estritamente decrescente de números naturais é finita.

Há ainda outras variantes do Princípio da Indução, duas das quais enunciamos a seguir por serem as mais comuns. Suas demonstrações serão feitas utilizando-se o Princípio da Boa Ordenação.

Proposição E.16 (Princípio da Indução Generalizado). *Se P uma propriedade referente a números naturais, cumprindo as seguintes condições:*

1. *o número natural k goza da propriedade P;*

2. *se um número natural n goza da propriedade P, seu sucessor $n+1$ também goza de P.*

Então, todos os números naturais maiores do que ou iguais a k gozam da propriedade P.

Demonstração. Seja X o conjunto dos números naturais que gozam de P. Assim, X é um conjunto indutivo que contém k. Suponha, por absurdo, que existam números naturais maiores do que k não pertencentes a X. Seja m o menor desses números. Como $m > k$, podemos escrever $m = n + 1$, em que, pela definição de m, tem-se necessariamente $n \in X$. Mas como X é indutivo, necessariamente $m = n + 1 \in X$, contradição. ∎

Exemplo E.17. Como aplicação do Princípio da Indução Generalizado, vamos mostrar que um polígono convexo de n lados ($n \geqslant 3$) possui $n(n-3)/2$ diagonais. Se $n = 3$, o polígono é um triângulo e possui $3 \cdot (3-3)/2 = 0$ diagonais, ou seja, não possui diagonais. Vamos supor então que, para algum $n > 3$, seja verdade que o número de diagonais de um polígono convexo de n lados seja $n(n-3)/2$ e consideremos um polígono de $n+1$ lados, com vétices V_1, V_2, ..., V_n, V_{n+1}. (Figura E.1.) Se unirmos V_1 a V_n, teremos um polígono convexo de n lados que, por hipótese, possui $n(n-3)/2$ diagonais. Assim, o número de diagonais do polígono de $n+1$ lados será:

$$\frac{n(n-3)}{2} + 1 + (n + 1 - 3),$$

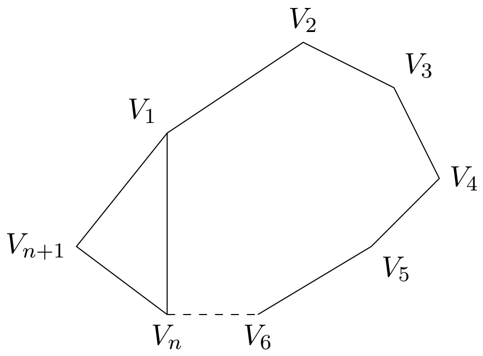


Figura E.1: Polígono convexo.

em que a primeira parcela é o número de diagonais do polígono de n lados (hipótese de indução), a segunda parcela refere-se ao lado V_1V_n do polígono de n lados (que é uma diagonal do polígono de $n+1$ lados), e a última parcela refere-se ao fato de que o vértice V_{n+1} se une a todos os vértices para formar diagonais excetuando-se V_1, V_n e ele próprio. Mas a soma acima é igual a $(n+1)(n+1-3)/2$, o que completa nossa demonstração. $\diamond$

Proposição E.18 (Segundo Princípio da Indução). *Seja $X \subset \mathbb{N}$ um conjunto com a seguinte propriedade: dado $n \in \mathbb{N}$, se todos os números naturais menores do que n pertencem a X, então $n \in X$. Então $X = \mathbb{N}$.*

Demonstração. Suponha, por absurdo, que $X \neq \mathbb{N}$, i.e., que $\mathbb{N} \setminus X \neq \emptyset$ e seja n o menor elemento de $\mathbb{N} \setminus X$, ou seja, n é o menor número natural que não pertence a X. Isto quer dizer que todos os números naturais menores do que n pertencem a X. Mas, pela propriedade do teorema, $n \in X$, contradição. Logo $\mathbb{N} \setminus X = \emptyset$ e portanto $X = \mathbb{N}$. $\blacksquare$

Exemplo E.19. [8] Qualquer que seja a maneira de decompor um polígono P (convexo ou não convexo), de n lados, em triângulos justapostos por meio de diagonais internas que não se intersectam, o número de diagonais utilizadas é sempre $n-3$. Com efeito, dado n, suponhamos que a proposição acima seja verdadeira para todo polígono com menos de n lados. Seja então dada uma decomposição do polígono P, de n lados, em triângulos justapostos, mediante diagonais internas. Fixemos uma dessas diagonais. (Veja a Figura E.2) Ela decompõe P como reunião de dois polígonos justapostos P_1, de n_1 lados, e P_2, de n_2 lados, em que $n_1 < n$ e $n_2 < n$, logo a proposição vale para os polígonos P_1 e P_2. Evidentemente, $n_1 + n_2 = n + 2$. As d diagonais que efetuam a decomposição de P se agrupam assim: $n_1 - 3$ delas decompõem P_1, $n_2 - 3$ decompõem P_2 e uma foi usada para separar P_1 de P_2. Portanto $d = n_1 - 3 + n_2 - 3 + 1 = n_1 + n_2 - 5$. Como $n_1 + n_2 = n + 2$, resulta que $d = n - 3$. Isso completa a demonstração. $\diamond$

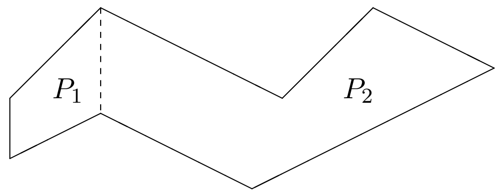


Figura E.2: Decomposição de um polígono por meio de diagonais.

E.1 Comentários sobre o Princípio da Indução

Como já dissemos, o Princípio da Indução é um eficiente instrumento para a demonstração de fatos que dizem respeito aos números naturais. Ele é um método que permite demonstrar a verdade de um número infinito de proposições basicamente a partir de dois passos. Na forma mais simples e comum do Princípio da Indução, esses passos consistem:

1. no *caso base*: mostrar que a proposição vale para $n = 1$ e

2. no *passo indutivo*: mostrar que se a proposição vale para $n = k$, vale também para $n = k + 1$.

A Figura E.3 ilustra intuitivamente essa ideia ao fazer um paralelo entre os números naturais e uma hipotética sequência infinita de dominós enfileirados. O caso base consitiria no fato de que o primeiro dominó poderia ser derrubado; o passo indutivo, no fato de que após derrubar um dominó qualquer, o próximo da fileira seria derrubado. Dessa forma, todo dominó (assim como todo número natural) seria derrubado (alcançado) se o primeiro o fosse.

Analogias à parte, em geral, ao entrar em contato com o Princípio da Indução pela primeira vez, fica a (falsa) impressão para o estudante de que há algo fundamentalmente errado ao se admitir a hipótese de indução. A seguinte queixa é bastante comum: ao se admitir que a proposição vale para um n arbitrário no passo indutivo, não estaríamos incorrendo numa demonstração circular, uma vez que provar a veracidade da proposição para todo n natural é o nosso objetivo?

A resposta é um sonoro NÃO!

Na verdade, o passo indutivo consiste em se demonstrar a proposição "$P(k) \Rightarrow P(k+1)$", e não $P(k)$ isoladamente.

Para tanto, recorremos às leis básicas da lógica. Quando uma proposição do tipo "p implica q" é falsa? Somente quando o antecedente, p, é verdadeiro e o consequente, q, é falso. Veja:

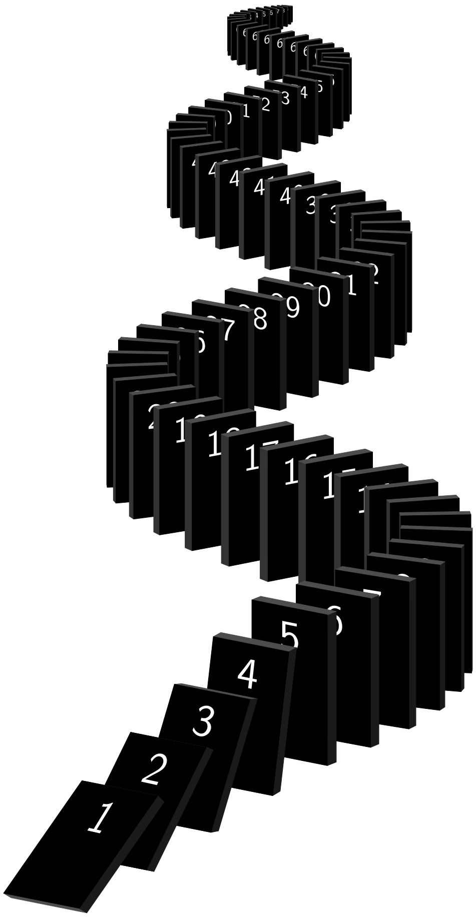


Figura E.3: O efeito dominó como analogia para o Princípio da Indução.

p	q	p implica q
Verdadeira	Verdadeira	Verdadeira
Verdadeira	Falsa	Falsa
Falsa	Verdadeira	Verdadeira
Falsa	Falsa	Verdadeira

Por isso, a única coisa a se fazer para mostrar que "p implica q" é verdadeira é provar que se p é verdadeira, q também o é.

Dessa forma, não está sendo feito demonstração circular, como o novato poderia supor.

Por fim, mas não menos importante, nunca faça generalizações apressadas sobre números naturais ou, de modo mais geral, sobre qualquer assunto! O exemplo a seguir deixa isso claro.

Exemplo E.20. [8] Considere o polinômio $p(n) = n^2 - n + 41$ e afirmação "o valor de $p(n)$ é sempre um número primo para $n = 1, 2, \ldots$". Embora isso seja verdadeiro para

$n = 1, 2, \ldots, 40$, temos $p(41) = 41^2 - 41 + 41$, que não é primo, logo a afirmação não é verdadeira. O mesmo ocorre com $q(n) = n^2 - 79n + 1601$, que fornece primos para $n = 1, 2, \ldots, 79$, mas $q(80) = 80^2 - 79 \cdot 80 + 1601$ não é primo, pois é divisível por 41.$\diamond$

Moral da história: só aceite uma afirmação (sobre números naturais ou mais geralmente sobre qualquer assunto) se ela tiver sido demonstrada de fato!

Bibliografia

1 BARATA, J. C. A. Noções Conjuntivistas Básicas. Em: CURSO de Física-Matemática: versão de 27 de junho de 2014. São Paulo: USP, 2014. capítulo 1, páginas 31–71. Ver página 114.

2 EVES, H. **Introdução à História da Matemática**. 5. edição. Campinas: Unicamp, 2011. página 848. Tradução de: Hygino H. Domingues. Ver página 15.

3 FEYNMAN, R. P. The Theory of Gravitation. Em: FEYNMAN, R. P.; LEIGHTON, R. B.; SANDS, M. (Org.). **Feynman Lectures on Physics, Vol. 1**: Mainly Mechanics, Radiation and Heat. New York: Basic Books, 2013. capítulo 7. Ver página 151.

4 HALMOS, P. R. **Naive Set Theory**. New York: Van Nostrand Reinhold Company, 1960. página 104. Ver páginas 42, 102.

5 KÖRNER, T. W. **Metrical and Topological Spaces**. Cambridge: University of Cambridge, 2015. página 109. Ver páginas 191, 192.

6 LENNARD-JONES, J. E. Cohesion. **Proceedings of the Physical Society**, IOP Publishing, volume 43, número 5, páginas 461–482, set. de 1931. Ver página 17.

7 LIMA, E. L. **Curso de Análise**: Volume 1. 3. edição. Rio de Janeiro: IMPA, 1982. página 344. Ver páginas 79, 80, 83.

8 ________. O Princípio da Indução. **Eureka!**, SBM, número 3, páginas 26–43, 1998. Ver páginas 28, 197, 199–201, 203, 205.

9 ________. **Temas e Problemas**. Rio de Janeiro: SBM, 2001. página 193. Ver páginas 130, 135, 147.

10 LIMA, E. L. et al. **A Matemática do Ensino Médio**: Volume 1. 10. edição. Rio de Janeiro: SBM, 2012. página 264. Ver páginas 21, 53, 136, 147, 161, 162, 164, 197, 201.

11 MANFIO, F. **Fundamentos da Geometria**. São Paulo: Universidade de São Paulo, [entre 2009 e 2014]. página 164. Ver página 173.

12 MARKETOS, P.; SCOTT, T. C. On the origin of the Fibonacci sequence. MacTutor History of Mathematics, mar. de 2014. Ver página 27.

13 MOORE, G. E. Cramming more components onto integrated circuits, reprinted from Electronics, v. 38, n. 8, April 19, 1965, p. 114–117. **IEEE Solid-State Circuits Society Newsletter**, IEEE, volume 11, número 5, páginas 33–35, set. de 2006. Ver página 16.

Índice

aplicação, 28
argumento, 22
axioma da escolha, 101

bijetividade, 70
boa ordem, 101, 200

campo
  de definição, 22
  de existência, 22
  de valores, 22
codomínio, 22
composição, 72
conjunto de valores, 22
contradomínio, 22, 34
coordenadas
  polares, 59
  retangulares, 51, 173–181
correspondência, 28

diagrama de setas, 50
distância, 183–186
domínio, 22, 32

espaço métrico, 188
extensão
  de funções, 88
  periódica, 89

família, 91
forma, 28
função,
  afim, 132
  característica, 23
  chão, 24
  conceito, 22, 47
  constante, 23
  de Euler, 113
  dente de serra, 113
  identidade, 23
  inclusão, 23
  indicadora, 23
  linear, 122
  módulo, 25
  parte fracionária, 25
  parte inteira, 24
  poligonal, 141
  projeção, 23
  quadrática, 149
  rampa, 144
  sinal, 23
  teto, 24
funcional, 28

gráfico de uma função, 49
  afim, 136
  quadrática, 155

imagem,
  conjunto, 22, 81
  inversa, 83

por uma função, 22
implícita, função, 104
injetividade, 67
intervalo, 29
aberto, 29
degenerado, 30
fechado, 29
ilimitado, 30
limitado, 29
semiaberto, 30
inversa
à direita, 79
à esquerda, 79
função, 77

lei de correspondência, 34, 37
limitação, 116

métrica, 188
mapa, 28
mapeamento, 28
monotonia, 114

operação, 28, 76
operador, 28

parábola, 155
parametrização, 56
pares ordenados, 41
paridade, 105
partição, 95
cruzada, 97
periodicidade, 111
princípio da indução, 199, 202, 203
produto, 28
produto cartesiano, 42, 99
proporcionalidade, 122
direta, 128
inversa, 129
teorema fundamental, 124

refinamento, 96
relação, 44
restrição de funções, 86

sequência, 26, 100
simetria, 109
sobrejetividade, 69

taxa de variação, 116
transformação, 28

variável
dependente, 22
independente, 22