UNIVERSIDADE FEDERAL DE MATO GROSSO DO SUL
INSTITUTO DE MATEMÁTICA
PROGRAMA DE PÓS GRADUAÇÃO
MESTRADO PROFISSIONAL EM MATEMÁTICA
EM REDE NACIONAL

MAIARA NASCIMENTO BORGES MARQUES

FUNÇÕES REAIS E FUNÇÕES CONVEXAS

CAMPO GRANDE - MS
2018

UNIVERSIDADE FEDERAL DE MATO GROSSO DO SUL
INSTITUTO DE MATEMÁTICA
PROGRAMA DE PÓS GRADUAÇÃO
MESTRADO PROFISSIONAL EM MATEMÁTICA
EM REDE NACIONAL

MAIARA NASCIMENTO BORGES MARQUES

FUNÇÕES REAIS E FUNÇÕES CONVEXAS

Orientadora: Profª. Drª. Rubia Mara de Oliveira Santos

Dissertação apresentada ao Programa de Pós-Graduação Mestrado Profissional em Matemática em Rede Nacional do Instituto de Matemática da Universidade Federal de Mato Grosso do Sul-INMA/UFMS como parte dos requisitos para obtenção do título de Mestre.

Campo Grande - MS
2018

FUNÇÕES REAIS E FUNÇÕES CONVEXAS

MAIARA NASCIMENTO BORGES MARQUES

Dissertação apresentada ao Programa de Pós-Graduação Mestrado Profissional em Matemática em Rede Nacional do Instituto de Matemática da Universidade Federal de Mato Grosso do Sul-INMA/UFMS como parte dos requisitos para obtenção do título de Mestre.

Aprovado pela Banca Examinadora:

Prof[a]. Dr[a]. Rubia Mara de Oliveira Santos -UFMS
Prof. Dr. André Nagamine - UESB
Prof. Dr. Claudemir Aniz - UFMS

Campo Grande – MS, Outubro de 2018

Agradecimentos

Agradeço primeiramente à Deus pois, sem o amor e a misericórdia Dele a me sustentar todos os dias, nada seria possível.

Agradeço ao meu marido, Mateus, que nunca mediu esforços para me ajudar a concluir esse sonho. Obrigada por ser paciente e compreensivo, entendendo minhas ausências nesses anos e sendo o maior responsável pelos cuidados de nossa casa e de nossa família. Se eu não tivesse seu apoio desde o início, essa conquista não seria possível. À minha filha, Alice, obrigada pelo seu amor incondicional e por me dar tanto orgulho de ser sua mãe, você é o meu maior incentivo para tudo na vida.

Aos meus pais, Maria Aparecida e Alcides, muito obrigada. O amor de vocês para comigo e a preocupação em sempre me proporcionar uma educação de qualidade, me motivando a estudar e lutar pelos meus sonhos, foram fundamentais para atingir esse e tantos outros objetivos. À minha irmã, Maieza, obrigada por me apoiar, ouvir e aconselhar nessa caminhada.

Agradeço aos meus amigos de mestrado, em especial ao Thiago, André, Eder, Clayton e Estevão, pelas horas de convivência e de muito estudo. Poder trilhar esse caminho com vocês tornou a jornada mais leve.

À todos os meus amigos que acompanharam minha dedicação ao mestrado, sempre me apoiando e fornecendo palavras de carinho e motivação nos momentos em que eu mais precisei, meus mais sinceros agradecimentos.

Agradeço à minha orientadora, profª Drª Rubia Mara de Oliveira Santos, por acreditar em mim desde o início, me motivar, me orientar e não medir esforços para que esse trabalho fosse feito da melhor forma possível.

À todos os professores do mestrado, obrigada por todo o conhecimento e dedicação compartilhados nesses anos.

Resumo

Essa dissertação aborda os principais conceitos de Conjuntos e Funções Reais, Conjuntos Convexos e Funções Convexas. O desenvolvimento é feito inicialmente com uma abordagem histórica do conceito de função. Em seguida, é apresentada uma revisão bibliográfica contendo definições, propriedades, teoremas e aplicações reais. Espera-se que esse material possa servir como um material didático de enriquecimento do conhecimento de professores de Matemática da educação básica.

Palavras-chave: Conjuntos, Funções Reais, Funções Convexas.

Abstract

This dissertation approaches the main concepts of Real Sets and Functions, Convex Sets and Convex Functions. The development is done initially with a historical approach of the function concept. Then, a bibliographic review is presented along with definitions, properties, theorems and real applications. It is hoped that this text can serve as a didactic material for the enrichment of the knowledge of Mathematic teachers of basic education.

Keywords: Sets, Real Functions, Convex Functions.

Sumário

Lista de Figuras

Lista de Tabelas

Capítulo 1

Introdução

O desenvolvimento do conceito de função aconteceu de forma lenta. Desde a primeira utilização do termo função, os conceitos que o definiam foram sendo modificados com o passar dos anos. O surgimento de funções como objetos de estudo matemático e sua conceituação individualizada teve início no final do século XVII [4].

No século XIV, Nicole Oresme (1323-1382) fez uma representação gráfica da variação da velocidade em relação ao tempo no movimento de um corpo com aceleração constante. Essa representação sugeria o que hoje denomina-se de representação gráfica de funções [2].

O conceito de função tem sua origem confundida com o início do Cálculo Infinitesimal [4]. Inicialmente, o Cálculo desenvolvido por Newton (1642-1727) e Leibniz (1646-1716) não tratava-se de um cálculo de funções. O objetivo principal desse estudo no século XVII eram as curvas, devido ao fato de os problemas que deram origem ao cálculo serem de natureza geométrica e cinemática [1]. Em 1692, Leibniz utilizou o termo "função"para designar um objeto geométrico relacionado a uma curva [1]. À ele também são devidas as introduções dos termos "constante","variável"e "parâmetro"[4].

À medida que o estudo de curvas foi se desenvolvendo e tornando-se mais algébrico, tornou-se necessário um termo que representasse, através de uma expressão analítica, quantidades dependentes de uma variável. Por volta de 1718, Jean Bernoulli (1667-1748) considerou como função de uma variável uma expressão qualquer formada por essa variável e por constantes [4]. Para expressar uma função de x, Bernoulli utilizou várias notações, dentre as quais a que mais se aproxima da utilizada atualmente foi ϕx [2].

A primeira obra em que o conceito de função desempenha papel central é de autoria de Euler (1707-1783). Nesse trabalho, Euler define função de uma quantidade variável como uma expressão analítica qualquer formada por essa variável e por números ou quantidades constantes [2]. A notação utilizada atualmente $f(x)$ para uma função de x é devida à Euler.

Nos séculos XVIII e XIX, a noção de função era relacionada com a de expressão analítica. Essa noção, ligada às noções de continuidade e de desenvolvimento em série passou por mudanças que alteraram seu significado [4]. Fourier (1768-1830), em seu trabalho sobre a propagação de calor em determinados materiais, afirmou que toda função definida em um intervalo finito poderia ser representada nesse intervalo por uma série trigonométrica, hoje denominada de série de Fourier [3]. Posteriormente, Dirichlet (1805-1859) foi o primeiro a empregar uma definição de função sem o uso de uma expressão analítica ou fórmula fechada, considerando como função uma correspondência entre duas variáveis x e y de modo que a cada valor de x existe uma regra tal que fica determinado um único y.

Após o desenvolvimento da teoria dos conjuntos, o conceito de função foi ampliado no século XX, de modo a incluir todas as relações arbitrárias entre conjuntos, sejam eles numéricos ou não [4]. No século XX, teve início a matemática moderna onde, entre os estudiosos desse século, destacou-se um grupo de matemáticos que utilizava o pseudônimo Nicolas Bourbaki [3]. Em suas obras, Bourbaki reformula diversos conceitos e fornece a definição de função como um conjunto de pares ordenados.

1.1 O Ensino das Funções na Educação Básica

De acordo com a Lei de Diretrizes e Bases da Educação Nacional (LDB 9.394/96) [5] a educação escolar brasileira é composta de dois níveis escolares, a educação básica e a educação superior, sendo a educação básica constituída pela educação infantil, ensino fundamental e ensino médio.

Com a necessidade de criar referências nacionais comuns ao processo educacional no Brasil, foram elaborados os Parâmetros Curriculares Nacionais (PCN) [6] que organizam os conteúdos de Matemática a serem trabalhados no ensino fundamental em quatro blocos: números e operações, espaço e forma, grandezas e medidas e tratamento da informação. O conteúdo de **funções** está presente nos PCN no bloco números e operações, onde o documento menciona o ensino da noção de funções nas séries finais do ensino fundamental, ressaltando, que a abordagem formal deve ser estudada no ensino médio.

Os Parâmetros Curriculares Nacionais para o Ensino Médio (PCNEM) [7] e as Orientações

Educacionais Complementares aos Parâmetros Curriculares Nacionais (PCN+) [8] determinam as diretrizes para o Ensino Médio e as habilidades básicas esperadas no aprendizado dos alunos, além de orientarem sobre metodologias no ensino da Matemática. Ambos os documentos abordam a necessidade da interdisciplinaridade e contextualização dos temas em Matemática.

Os PCNEM reconhecem que o ensino de funções não deve ser feito de forma isolada, uma vez que esse tema está relacionado à vários outros. O documento ressalta a importância do conceito de função pois este pode ser aplicado em eventos do cotidiano e em outras áreas, como a Física, Geografia ou Economia [7]. Nos PCN+, o conteúdo de funções está presente no eixo Álgebra: números e funções, e neste documento a orientação é que o enfoque no estudo das funções deve ser no conceito de função e em suas propriedades em relação às operações, na interpretação de gráficos e nas aplicações dessas funções [8].

Foram criadas em 2006 as Orientações Curriculares para o Ensino Médio [9], que abordam a escolha de conteúdos, a forma de trabalhar os conteúdos, o projeto pedagógico e a organização curricular. Nesse documento a organização dos conteúdos básicos é feita em quatro blocos: Números e operações; Funções; Geometria; Análise de dados e probabilidade. Embora os conteúdos sejam assim organizados, não devem ser explorados de forma isolada, mas, ao contrário, deve-se buscar conectá-los. O estudo de Funções pode ser trabalhado inicialmente com a análise das relações entre duas grandezas em diferentes situações, esboçando também qualitativamente os gráficos dessas relações e prosseguindo o estudo com os diferentes modelos que devem ser estudados na escola: linear, quadrático, exponencial e periódico relacionando-os em diferentes áreas do conhecimento [9].

Em dezembro de 2017 foi aprovada e homologada a Base Nacional Comum Curricular (BNCC) [10] referente à Educação Infantil e Ensino Fundamental. A BNCC é um documento de caráter normativo que define as aprendizagens essenciais aos alunos da Educação Básica e é uma referência nacional obrigatória para a elaboração dos currículos a serem trabalhados em todas as escolas do país. Nesse documento, as habilidades a serem desenvolvidas ao longo do Ensino Fundamental e os objetos de conhecimento estão organizados em cinco unidades temáticas: Números; Álgebra; Geometria; Grandezas e medidas; Probabilidade e estatística. O ensino de funções encontra-se na unidade temática Álgebra do 9º ano do Ensino Fundamental.

A tecnologia encontra-se cada dia mais presente no cotidiano. O uso de recursos tecnológicos em sala de aula proporciona ao aluno uma aproximação dos conteúdos abordados com a prática. Contudo, no ensino da Matemática, ainda se predomina a forma tradicional, com o pouco relacionamento entre a teoria e a prática.

As Orientações Curriculares para o Ensino Médio citam a relevância da relação entre tecnologia e Matemática: "É importante contemplar uma formação escolar nesses dois sentidos, ou seja, a Matemática como ferramenta para entender a tecnologia, e a tecnologia como ferramenta para entender a Matemática". No uso da tecnologia para a Matemática há programas de computadores (*softwares*) que contém recursos motivadores do "pensar matematicamente"onde os alunos fazem tentativas e constroem raciocínios na resolução de problemas [9].

O uso de recursos tecnológicos contribui no processo de ensino e aprendizagem, pois além de pôr em perspectiva a ineficiência da simples manipulação simbólica e do cálculo mecânico se comparados com a resolução de cálculos utilizando instrumentos tecnológicos, permite diferentes métodos de abordagem de problemas ao destacar para os alunos benefícios da linguagem gráfica e de novas formas de representação, despertando neles entusiasmo pela prática de exercícios de investigação e exploração [6].

Os *softwares* facilitam de maneira simultânea o estudo algébrico e gráfico através de seus recursos, o que contribui para o entendimento do aluno do conceito de função [9].

Um programa de computador que pode ser utilizado no processo de ensino e aprendizagem de funções é o *GeoGebra*. O *GeoGebra* é um *software* matemático dinâmico, de fácil utilização e que reúne recursos associados à geometria, álgebra, gráficos, entre outros [11]. Esse programa é gratuito para uso não comercial; pode ser instalado em computadores e também em *smartphones*, e após instalado, o acesso à internet não é necessário para sua utilização. O *software GeoGebra* pode ser obtido para instalação em seu *site* oficial: https://www.geogebra.org/. O processo de aprendizagem dos alunos pode ser enriquecido com esse *software* pois, através dele é possível construir com precisão gráficos de funções definidas por quaisquer expressões analíticas, além de facilitar a compreensão de conceitos abordados em sala de aula.

1.2 Objetivos

Este trabalho tem como objetivo apresentar conceitos acerca de Conjuntos e Funções Reais, Conjuntos Convexos e Funções Convexas para ser um material de apoio no aprofundamento do conhecimento do professor de Matemática nesses temas. Dessa forma, o trabalho está organizado como segue:

O capítulo 2 apresenta os principais conceitos de conjuntos e funções em $\mathbb{R}$, incluindo definições, proposições e teoremas. Também estão presentes nesse capítulo aplicações de funções afim, quadrática, exponencial e logarítmica.

O capítulo 3 aborda conceitos de conjuntos e funções convexas em $\mathbb{R}^n$. Nesse capítulo apresenta-se através de problemas, funções definidas em $\mathbb{R}$ relacionadas aos conceitos de convexidade.

No capítulo 4 serão feitas as conclusões dessa dissertação.

Capítulo 2

Conjuntos, Funções e Aplicações

Este capítulo apresenta definições, propriedades e teoremas acerca de Conjuntos e Funções reais, incluindo aplicações em situações presentes na realidade. Os conceitos aqui abordados foram obtidos em [12], [13], [14], [15], [17] e [18].

2.1 Conjuntos

Um **conjunto** é uma coleção de objetos que são chamados de elementos. A relação entre um conjunto e um objeto é a relação de pertinência.

Sendo A um conjunto e x um objeto qualquer, se x é um elemento do conjunto A diz-se que x pertence a A e escreve-se $x \in A$, caso contrário, diz-se que x não pertence a A e escreve-se $x \notin A$.

Para determinar um conjunto A é necessário uma regra que permite julgar se um objeto qualquer x pertence ou não ao conjunto A. Na maioria dos casos, os elementos de um conjunto são determinados através de uma propriedade comum e exclusiva entre eles. Quando um conjunto não possui elementos que satisfazem a regra determinada, esse é chamado de conjunto vazio e é representado por $\emptyset$.

Definição 1 *Dados os conjuntos A e B, se todos os elementos do conjunto A são pertencentes também ao conjunto B, diz-se que A é um subconjunto de B, que A está contido em B ou que A é parte de B e denota-se por $A \subset B$.*

A relação $A \subset B$ dada na definição 1 é chamada de relação de inclusão. Quando o conjunto A não é um subconjunto de B, denota-se por $A \not\subset B$.

Dados conjuntos quaisquer A,B e C, na relação de inclusão são válidas as seguintes propriedades:

a) reflexividade: $A \subset A$;

b) antissimetria: se $A \subset B$ e $B \subset A$ então $A = B$;

c) transitividade: se $A \subset B$ e $B \subset C$ então $A \subset C$.

Definição 2 *Fixado um conjunto U, chamado de conjunto universo, tal que todos os elementos a serem considerados pertencerão a U e todos os conjuntos estarão contidos em U, então, dado um conjunto A, o conjunto formado pelos elementos de U que não pertencem a A é chamado de* ***complementar*** *de A e é representado por A^c. Ou seja, $A^c = \{x \ : \ x \notin A\}$*

Proposição 1 *Sendo A e B conjuntos dados, são válidas as seguintes regras operatórias:*

a) Para todo $A \subset U$, tem-se $(A^c)^c = A$;

b) Se $A \subset B$ então $B^c \subset A^c$.

Demonstração: a) Seja $A \subset U$. Se $a \in (A^c)^c$, então $a \notin (A^c)$, logo $a \in A$ e, daí, $(A^c)^c \subset A$. Por outro lado, se $a \in A$ tem-se que $a \notin A^c$, assim segue que $a \in (A^c)^c$ e, consequentemente, $A \subset (A^c)^c$. Portanto, $(A^c)^c = A$ para todo $A \subset U$.
b) Suponha que $A \subset B$. Sendo $b \in B^c$ tem-se que $b \notin B$ e, por hipótese, $b \notin A$, logo $b \in A^c$. Então, $B^c \subset A^c$.

Definição 3 *Dados os conjuntos A e B, o conjunto formado pelos elementos de A juntamente com os elementos de B é denominado* ***reunião*** *de A e B e é representado por $A \cup B$. Ou seja, $A \cup B = \{x \ : \ x \in A \text{ ou } x \in B\}$.*

Definição 4 *Sejam A e B dois conjuntos dados. Então, o conjunto formado pelos elementos comuns de A e B é chamado de* ***interseção*** *de A e B e é denotado por $A \cap B$. Desta forma, $A \cap B = \{x \ : \ x \in A \text{ e } x \in B\}$.*

Definição 5 *Dados os conjuntos A e B. O conjunto formado pelos elementos de A que não pertencem a B é denominado* ***diferença*** *entre A e B e é indicado por $A - B$. Assim, $A - B = \{x \ : \ x \in A \text{ e } x \notin B\}$.*

Um conjunto X, pode ser classificado em finito ou infinito, conforme a definição 6.

Definição 6 *Um conjunto X é* ***finito*** *quando possui um número finito de elementos, ou seja, quando é possível determinar o número de elementos desse conjunto e é* ***infinito*** *quando não é finito.*

O número de elementos de um conjunto X é chamado o número cardinal de X. O conjunto vazio $\emptyset$ é definido como um conjunto finito e que possui zero elementos, logo seu número cardinal é zero.

Proposição 2 *Indicando por $n(X)$ o número cardinal de um conjunto finito X, são válidas as seguintes propriedades:*

a) O número cardinal de um conjunto finito é único, independente da contagem adotada.

b) Todo subconjunto Y de um conjunto finito X é finito e $n(Y) \leq n(X)$. A igualdade $n(Y) = n(X)$ só ocorre quando $Y = X$.

c) Se X e Y são finitos então $X \cup Y$ é finito e tem-se $n(X \cup Y) = n(X) + n(Y) - n(X \cap Y)$.

2.2 Funções

Definição 7 *Dados os conjuntos não vazios X, Y, a regra que associa a cada elemento $x \in X$ um único elemento $y = f(x) \in Y$ é denominada* ***função*** *f de X em Y e denota-se por $f : X \to Y$.*

O conjunto X é chamado **domínio** e Y é o **contradomínio** de f. Para cada $x \in X$, o elemento $f(x) \in Y$ chama-se a imagem de x por f.

Dado um conjunto não vazio X, a função denotada por $Id_X : X \to X$ e definida por $Id_X(x) = x$ para todo $x \in X$ é denominada função identidade de X.

Dados os conjuntos não vazios X, Y e fixado um elemento $c \in Y$, a função $f : X \to Y$ tal que $f(x) = c$ para todo $x \in X$, é chamada função constante.

Definição 8 *Dada a função $f : X \to Y$, a* ***imagem*** *de f é o conjunto formado pelos elementos $f(x) \in Y$, onde $x \in X$, ou seja, $Im(f) = \{f(x) \in Y \ : \ x \in X\}$.*

A imagem de f é um conjunto que está contido no contradomínio da função e pode ser diferente deste, ou seja, $Im(f) \subset Y$ e é possível que $Im(f) \neq Y$.

Definição 9 *Dados os conjuntos não vazios X e Y, o conjunto formado por todos os pares ordenados (x, y), onde $x \in X$ e $y \in Y$ é denominado* ***produto cartesiano*** *de X por Y. Ou seja, $X \times Y = \{(x, y) \ : \ x \in X \text{ e } y \in Y\}$.*

Definição 10 *Dados os conjuntos não vazios X e Y, todo subconjunto $R \subset X \times Y$, é denominado **relação** de X em Y.*

Uma função é um caso particular de uma relação entre dois conjuntos como vê-se a seguir.

Definição 11 *Dados os conjuntos não vazios X e Y, uma relação f de X em Y é uma função se para todo $x \in X$ existir um único $y \in Y$ tal que $(x, y) \in f$.*

Através da definição 12 tem-se as condições para que ocorra a igualdade entre duas funções.

Definição 12 *Duas funções $f : X \to Y$ e $g : X' \to Y'$ são **iguais** se, e somente se, $X = X', Y = Y'$ e $f(x) = g(x)$ para todo $x \in X$.*

A partir de funções já conhecidas pode-se construir novas funções como vê-se na definição 13.

Definição 13 *Dadas as funções $f : X \to \mathbb{R}$ e $g : Y \to \mathbb{R}$, com $X, Y \subset \mathbb{R}$, tais que $X \cap Y$ seja diferente do vazio e um número real c, são definidas as funções:*

a) $f + g : X \cap Y \to \mathbb{R}$ dada por $(f + g)(x) = f(x) + g(x)$;

b) $f \cdot g : X \cap Y \to \mathbb{R}$ dada por $(f \cdot g)(x) = f(x) \cdot g(x)$;

c) $\frac{f}{g}$ dada por $\left(\frac{f}{g}\right)(x) = \frac{f(x)}{g(x)}$ e domínio igual a $\{x \in X \cap Z \ : \ g(x) \neq 0\}$.

d) $c \cdot f : X \to \mathbb{R}$ dada por $(c \cdot g)(x) = c \cdot g(x)$.

Sendo $f : X \to Y$ e $g : Y \to Z$ duas funções, é possível formar uma nova função com elementos $x \in X$, determinando suas imagens $f(x)$ e, em seguida, calculando $g(f(x))$. Os resultados compõem uma função obtida pela substituição de f em g chamada de função *composta* de f e g.

Definição 14 *Dadas as funções $f : X \to Y$ e $g : Y \to Z$, onde a imagem de f está contida no domínio de g, a função $g \circ f : X \to Z$ definida por $(g \circ f)(x) = g(f(x))$, para cada $x \in X$ é denominada **função composta** de f e g, nessa ordem.*

Proposição 3 *Dadas as funções $f : X \to Y$, $Id_x : X \to X$ e $Id_y : Y \to Y$, tem-se que $f \circ Id_x = f$ e $Id_y \circ f = f$.*

Demonstração: A função $f \circ Id_x$ onde $f \circ Id_x : X \to Y$ é tal que $(f \circ Id_x)(x) = f(Id_x(x)) = f(x)$ para todo $x \in X$. Por outro lado, a função $Id_y \circ f$ com $Id_y \circ f : Y \to X$ é tal que, para todo $y \in Y$, tem-se $(Id_y \circ f)(y) = Id_y(f(y)) = f(y)$. ∎

Proposição 4 *Dadas as funções $f : X \to Y$, $g : Y \to Z$ e $h : Z \to W$, tem-se:*

a) $h \circ (g \circ f) = (h \circ g) \circ f$;

b) $f^{(n)} = \underbrace{f \circ f \circ f \cdots \circ f}_{n}$, *sendo* $f : X \to X$ *e* $n \in \mathbb{N}$.

Demonstração: a) Primeiramente, note que $h \circ (g \circ f)$ e $(h \circ g) \circ f$ são ambas funções de X em W. Assim, pela definição 12 a igualdade ocorrerá se elas associarem a cada $x \in X$ um mesmo elemento em W. De fato, para todo $x \in X$ tem-se que

$$[h \circ (g \circ f)](x) = (h \circ g)(f(x)) = h[g(f(x))] = h[(g \circ f)(x)] = [(h \circ g) \circ f](x). \qquad \blacksquare$$

Definição 15 *Uma função $f : X \to Y$ é:*

a) ***injetiva****, ou* ***injetora****, quando elementos distintos do domínio X são transformados por f em elementos distintos em Y. Ou seja, f é injetiva quando $x \neq x'$ em $X \Rightarrow f(x) \neq f(x')$. Equivalentemente, a condição pode ser expressa por $f(x) = f(x') \Rightarrow x = x'$.*

b) ***sobrejetiva****, ou* ***sobrejetora****, se para todo elemento $y \in Y$ existir pelo menos um elemento $x \in X$ tal que $y = f(x)$, ou seja, quando $Y = Im(f)$.*

c) ***bijetiva****, ou* ***bijetora****, quando for ao mesmo tempo injetiva e sobrejetiva.*

Uma função $f : X \to Y$ que é ao mesmo tempo injetiva e sobrejetiva pode ser chamada de **bijeção**, ou **correspondência biunívoca** entre X e Y.

Estabelecidos os conceitos de função e correspondência biunívoca é possível escrever de modo mais formal as definições de conjunto finito e conjunto infinito dadas na definição 6.

Definição 16 *Um conjunto X é* ***finito*** *quando existe uma correspondência biunívoca $f : I_n \to X$, onde $I_n = \{1, 2, \ldots, n\}$ para $n \in \mathbb{N}$. O número natural n é o número cardinal de X ou o número de elementos de X.*

O conjunto I_n é finito e tem n como seu número cardinal. Desta forma, todo conjunto finito possui um número natural n como seu número cardinal.

Definição 17 *Um conjunto X é* ***infinito*** *quando não é vazio e não existe uma correspondência biunívoca $f : I_n \to X$, com $I_n = \{1, 2, \ldots, n\}$ qualquer seja $n \in \mathbb{N}$.*

Proposição 5 *Dadas as funções $f : X \to Y$ e $g : Y \to Z$. Então:*

a) $g \circ f$ *injetora* $\Rightarrow f$ *injetora.*

b) $g \circ f$ *sobrejetora* $\Rightarrow g$ *sobrejetora.*

c) g, f injetoras $\Rightarrow g \circ f$ injetora.

d) g, f sobrejetoras $\Rightarrow g \circ f$ sobrejetora.

e) g, f bijetoras $\Rightarrow g \circ f$ bijetora.

Demonstração: a) Supondo $g \circ f$ injetora, para $x_1, x_2 \in X$, tem-se que

$$f(x_1) = f(x_2) \Rightarrow g(f(x_1)) = g(f(x_2) \Rightarrow (g \circ f)(x_1) = (g \circ f)(x_2) \Rightarrow x_1 = x_2,$$

portanto, f é injetora.

b) Supondo $g \circ f$ sobrejetora, dado $z \in Z$ existe pelo menos um $x \in X$ tal que $z = (g \circ f)(x)$. Assim, $z = g(f(x))$ e, portanto, g é sobrejetora.

c) Supondo f, g injetoras e tomando $x_1, x_2 \in X$ tais que $(g \circ f)(x_1) = (g \circ f)(x_2)$ tem-se

$$(g \circ f)(x_1) = (g \circ f)(x_2) \Rightarrow g(f(x_1)) = g(f(x_2))$$

como g é injetora, segue que $g(f(x_1)) = g(f(x_2)) \Rightarrow f(x_1) = f(x_2)$ e, da injetividade de f, conclui-se que $f(x_1) = f(x_2) \Rightarrow x_1 = x_2$. Logo, $(g \circ f)(x_1) = (g \circ f)(x_2) \Rightarrow x_1 = x_2$, o que mostra que $g \circ f$ é injetora.

d) Supondo f e g sobrejetoras, para $z \in Z$, o fato de g ser sobrejetora garante que existe $y \in Y$ tal que $z = g(y)$. Em contrapartida, a sobrejetividade de f atesta a existência de $x \in X$ tal que $f(x) = y$. Assim, para todo $z \in Z$ existe $x \in X$ de modo que $(g \circ f)(x) = g(f(x)) = g(y) = z$. Portanto, $g \circ f$ é sobrejetora.

e) Supondo f e g bijetoras, segue que f e g são injetoras e sobrejetoras. Do item c) tem-se que se g e f são injetoras implica que $g \circ f$ é injetora, além disso, pelo item d), se g e f são sobrejetoras então $g \circ f$ é sobrejetora. Portanto, g e f bijetoras $\Rightarrow g \circ f$ bijetora. ■

Definição 18 *Seja $f : X \to Y$ uma função bijetora. Existe a função $g : Y \to X$ tal que, para $x \in X$, $y \in Y$, tem-se $g(y) = x \Leftrightarrow y = f(x)$ denominada* ***função inversa*** *de f.*

A inversa de uma função bijetora $f : X \to Y$ pode ser denotada por $f^{-1} : Y \to X$.

Dadas as funções $f : X \to Y$ e $g : Y \to X$. Diz-se que g é a inversa de f quando $g \circ f = Id_x$ e $f \circ g = Id_x$.

A proposição 6 relaciona as operações de composição e inversão de funções.

Proposição 6 *Se $f : X \to Y$ e $g : Y \to Z$ são bijeções, então $g \circ f : X \to Z$ é uma bijeção e*

$$(g \circ f)^{-1} = f^{-1} \circ g^{-1}$$

Demonstração: Como f, g são bijeções, segue do item (e) da Proposição 5, que $g \circ f$ é uma bijeção. No entanto, como $(g\circ f)^{-1}$ e $f^{-1}\circ g^{-1}$ são ambas funções de Z em X, pela unicidade da inversa, para comprovar que $(g\circ f)^{-1} = f^{-1}\circ g^{-1}$ é suficiente mostrar que $(f^{-1}\circ g^{-1})\circ(g\circ f) = Id_x$ e $(g \circ f) \circ (f^{-1} \circ g^{-1}) = Id_z$. De fato, note que

$$f^{-1} \circ f = Id_X,\ f \circ f^{-1} = Id_Y,\ g^{-1} \circ g = Id_Y \text{ e } g \circ g^{-1} = Id_Z.$$

Assim, tem-se

$$(f^{-1} \circ g^{-1}) \circ (g \circ f) = [(f^{-1} \circ g^{-1}) \circ g] \circ f = [f^{-1} \circ (g^{-1} \circ g)] \circ f = [f^{-1} \circ Id_Y] \circ f = f^{-1} \circ f = Id_X$$

e

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

∎

Definição 19 *Seja $X \subset \mathbb{R}$. Uma função $f : X \to \mathbb{R}$ é:*

a) ***crescente*** *se $x_1 < x_2 \Rightarrow f(x_1) < f(x_2)$, para quaisquer $x_1, x_2 \in X$;*

b) ***decrescente*** *se $x_1 < x_2 \Rightarrow f(x_1) > f(x_2)$, quaisquer sejam $x_1, x_2 \in X$;*

c) ***monótona não-decrescente*** *se $x_1 < x_2 \Rightarrow f(x_1) \leq f(x_2)$, para quaisquer $x_1, x_2 \in X$;*

d) ***monótona não-crescente*** *se $x_1 < x_2 \Rightarrow f(x_1) \geq f(x_2)$, quaisquer sejam $x_1, x_2 \in X$;*

Em qualquer um dos casos, a função f é dita monótona. Nos casos (a) e (b), f é uma função injetiva.

É importante reassaltar que existem funções que não são definidas por nenhuma das quatro condições dadas na definição 19.

Definição 20 *Seja $f : I \to \mathbb{R}$ uma função dada onde $I \subset \mathbb{R}$ é um intervalo. O* ***valor mínimo*** *da função f no intervalo I é dado por $y_0 \in \mathbb{R}$ quando ocorrerem simultaneamente:*

a) $Im(f) \subset [y_0, +\infty)$.

b) $y_0 \in Im(f)$.

Os ***pontos de mínimo*** *da função f no intervalo I são os números reais $x_0 \in I$ tais que $f(x_0) = y_0$.*

De forma análoga é possível definir os conceitos de **valor máximo** e **ponto de máximo** de uma função $f : I \to \mathbb{R}$, onde I é um intervalo tal que $I \subset \mathbb{R}$. A definição 21 conceitua a paridade de uma função.

Definição 21 *Sejam $X \subset \mathbb{R}$ tal que $x \in X \Rightarrow -x \in X$ e $f : X \to Y$ uma função. Diz-se que:*

*a) f é **par** se, para todo $x \in X$, $f(x) = f(-x)$.*

*b) f é **ímpar** se, para todo $x \in X$, $f(x) = -f(-x)$.*

Existem funções que não podem ser definidas por nenhuma das condições dadas na definição 21

Definição 22 *O gráfico de uma função $f : X \to Y$ é definido como o subconjunto do produto cartesiano $X \times Y$ dado por*

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

Um subconjunto $G \subset X \times Y$ é o gráfico de uma função $f : X \to Y$ se para cada $x \in X$ existir um, e apenas um, $y \in Y$ tal que $(x, y) \in G$.

2.2.1 Função afim

Definição 23 *Uma função $f : \mathbb{R} \to \mathbb{R}$ tal que $f(x) = ax + b$ para todo $x \in \mathbb{R}$ e $a, b \in \mathbb{R}$ é denominada **afim**.*

A função $f : X \to Y$ tal que $f(x) = ax$ chamada de **função linear** é um exemplo particular de uma função afim, onde $b = 0$. Essa função é utilizada em problemas de proporcionalidade. O teorema a seguir, intitulado Teorema Fundamental da Proporcionalidade, permite determinar sempre se uma dada função é ou não uma função linear.

Teorema 1 *Seja $f : \mathbb{R} \to \mathbb{R}$ uma função crescente. São equivalentes as seguintes afirmações:*

(i) Para todo $n \in \mathbb{Z}$ e todo $x \in \mathbb{R}$ tem-se $f(nx) = nf(x)$.

(ii) Pondo $a = f(1)$, tem-se $f(x) = ax$, $\forall x \in \mathbb{R}$. (Logo $f(cx) = cf(x)$, $\forall c, x \in \mathbb{R}$.)

(iii) $f(x + y) = f(x) + f(y)$ quaisquer sejam $x, y \in \mathbb{R}$.

Demonstração: Com o objetivo de demonstrar que (i) $\Rightarrow$ (ii), será provado que, para todo número racional $r = \frac{m}{n}$, a hipótese (i) ocasiona que $f(rx) = rf(x)$, qualquer que seja $x \in \mathbb{R}$. De fato, como $nr = m$, tem-se:

$$n \cdot f(rx) = f(nrx) = f(mx) = m \cdot f(x),$$

logo $f(rx) = \dfrac{m}{n} f(x) = r \cdot f(x)$.

Seja $a = f(1)$. Como $f(0) = f(0 \cdot 0) = 0 \cdot f(0) = 0$, a monotonicidade de f garante que $a = f(1) > f(0) = 0$. Assim, a é positivo. Além disso, tem-se $f(r) = f(r \cdot 1) = r \cdot f(1) = r \cdot a = ar$ para todo $r \in \mathbb{Q}$. Resta mostrar que se tem $f(x) = ax$ para todo $x \in \mathbb{R}$. Suponha, por absurdo, que exista algum real x (necessariamente irracional) tal que $f(x) \neq ax$. A fim de fixar ideias, admita $f(x) < ax$. Tem-se $\dfrac{f(x)}{a} < x$.

Tomando um racional r no intervalo $\left[\dfrac{f(x)}{a}, x\right]$, ou seja, $\dfrac{f(x)}{a} < r < x$, segue-se que $f(x) < ar < ax$, assim, $f(x) < f(r) < ax$. Contudo isso é um absurdo, pois f é crescente logo, como $r < x$, deveria ocorrer $f(r) < f(x)$. Portanto, conclui-se que (i) $\Rightarrow$ (ii).

Para provar que (ii)$\Rightarrow$ (iii), suponha $f(x) = ax$, para todo $x \in \mathbb{R}$ e $a = f(1)$. Desta forma, para quaisquer $x, y \in \mathbb{R}$ tem-se $f(x+y) = a(x+y) = ax + ay = f(x) + f(y)$. Logo, (ii)$\Rightarrow$ (iii).

Por fim, para mostrar que (iii)$\Rightarrow$(i), suponha $f(x + y) = f(x) + f(y)$ quaisquer sejam $x, y \in \mathbb{R}$. Com isso, sendo $n \in \mathbb{N}$, tem-se

$$
\begin{aligned}
f(nx) &= f(x + (n-1)x) \\
&= f(x) + f((n-1)x) \\
&= f(x) + f(x + (n-2)x) \\
&= f(x) + f(x) + f((n-2)x) \\
&\vdots \\
&= \underbrace{f(x) + \cdots + f(x)}_{n} \\
&= nf(x).
\end{aligned}
$$

Resta mostrar que $f(-nx) = -nf(x)$. Desta forma, $0 = f(0) = f[x+(-x)] = f(x)+f(-x)$, o que implica $f(x) = -f(-x) = -f(x)$. Logo, $f(-nx) = -f(nx) = -nf(x)$. ∎

O teorema 2 apresenta uma caracterização das funções afins.

Teorema 2 *Seja $f : \mathbb{R} \to \mathbb{R}$ uma função monótona injetiva. Se o acréscimo $f(x+h) - f(x) = \varphi(h)$ depender somente de h, e não de x, então f é uma função afim.*

Demonstração: Supondo f crescente, então $\varphi : \mathbb{R} \to \mathbb{R}$ também é crescente, com $\varphi(0) = 0$.

Além disso, para quaisquer $h, k \in \mathbb{R}$ tem-se

$$\begin{aligned}\varphi(h+k) &= f(x+h+k) - f(x)\\ &= f((x+k)+h) - f(x+k) + f(x+k) - f(x)\\ &= \varphi(h) + \varphi(k).\end{aligned}$$

Logo, pelo Teorema 1, pondo-se $a = \varphi(1)$, tem-se $\varphi(h) = a \cdot h$ para todo $h \in \mathbb{R}$. O que significa que $f(x+h) - f(x) = ah$. Chamando $f(0)$ de b, resulta $f(h) = ah + b$, isto é, $f(x) = ax + b$ para todo $x \in \mathbb{R}$. ∎

É possível determinar se uma função $f : \mathbb{R} \to \mathbb{R}$ é afim, sem o conhecimento de seus coeficientes a e b. Nessa situação, o valor de b é obtido como o valor da função quando $x = 0$, ou seja, $b = f(0)$. Em relação ao coeficiente a, o mesmo pode ser determinado conhecendo dois valores $f(x_1)$ e $f(x_2)$ assumidos pela função em dois pontos distintos quaisquer x_1 e x_2. De fato, sendo conhecidos $f(x_1) = ax_1 + b$ e $f(x_2) = ax_2 + b$ tem-se que,

$$f(x_2) - f(x_1) = a(x_2 - x_1)$$

logo

$$a = \frac{f(x_2) - f(x_1)}{x_2 - x_1}$$

Dados $x, x+h \in \mathbb{R}$, com $h \neq 0$, o número $a = \dfrac{[f(x+h) - f(x)]}{h}$ é chamado de **taxa de crescimento** ou taxa de variação de f em $[x, x+h]$. O coeficiente b pode ser chamado de **valor inicial** da função f.

Proposição 7 *O gráfico de uma função afim $f : \mathbb{R} \to \mathbb{R}$, $f(x) = ax + b$ é uma reta.*

Demonstração: Para verificar isso basta mostrar que três pontos quaisquer deste gráfico $P_1 = (x_1, ax_1 + b)$, $P_2 = (x_2, ax_2 + b)$ e $P_3 = (x_3, ax_3 + b)$ são colineares. Uma condição necessária e suficiente para que isso ocorra é que o maior dos três números $d(P_1, P_2)$, $d(P_2, P_3)$ e $d(P_1, P_3)$ seja igual à soma dos outros dois. Desta forma, supondo $x_1 < x_2 < x_3$ e calculando a distância entre os pontos dois a dois, tem-se:

$$d(P_1, P_2) = \sqrt{(x_2 - x_1)^2 + a^2(x_2 - x_1)^2} = \sqrt{(x_2 - x_1)^2(1 + a^2)} = (x_2 - x_1)\sqrt{1 + a^2},$$

$$d(P_1, P_3) = \sqrt{(x_3 - x_1)^2 + a^2(x_3 - x_1)^2} = (x_3 - x_1)\sqrt{1 + a^2}$$

e

$$d(P_2, P_3) = \sqrt{(x_3 - x_2)^2 + a^2(x_3 - x_2)^2} = (x_3 - x_2)\sqrt{1 + a^2}.$$

Com isso, segue que:

$$\begin{aligned} d(P_1, P_2) + d(P_2, P_3) &= (x_2 - x_1)\sqrt{1 + a^2} + (x_3 - x_2)\sqrt{1 + a^2} \\ &= (x_3 - x_1)\sqrt{1 + a^2} = d(P_1, P_3). \end{aligned}$$

Portanto, P_1, P_2 e P_3 são colineares. ∎

Proposição 8 *Seja $f : \mathbb{R} \to \mathbb{R}$ uma função afim tal que $f(x) = ax + b$, com $a, b \in \mathbb{R}$ e $a \neq 0$. O conjunto imagem de f é $\mathbb{R}$.*

Demonstração: A imagem de f é obtida ao encontrar os $y \in \mathbb{R}$ tais que $f(x) = y$. Note que, qualquer que seja $y \in \mathbb{R}$ existe $x = \dfrac{y - b}{a} \in \mathbb{R}$ tal que

$$f(x) = f\left(\frac{y - b}{a}\right) = a \cdot \frac{y - b}{a} + b = y.$$

Logo, $Im(f) = \mathbb{R}$ ∎

Proposição 9 *Seja $f : \mathbb{R} \to \mathbb{R}$ uma função afim tal que $f(x) = ax + b$, com $a, b \in \mathbb{R}$ e $a \neq 0$. A função f é:*

(a) crescente, se $a > 0$.

(b) decrescente, se $a < 0$.

Demonstração: (a) Supondo $a > 0$ e considerando $x_1, x_2 \in \mathbb{R}$ tais que $x_1 < x_2$, tem-se

$$f(x_1) - f(x_2) = ax_1 + b - (ax_2 + b) = a(x_1 - x_2) < 0,$$

pois $x_1 < x_2$ implica $x_1 - x_2 < 0$. Logo, f é crescente quando $a > 0$.

(b) Analogamente, suponha $a < 0$. Se $x_1, x_2 \in \mathbb{R}$ são tais que $x_1 < x_2$, então

$$f(x_1) - f(x_2) = ax_1 + b - (ax_2 + b) = a(x_1 - x_2) > 0,$$

uma vez que $x_1 < x_2$ é equivalente a $x_1 - x_2 < 0$. Portanto, f é decrescente quando $a < 0$. ∎

2.2.2 Funções quadráticas

Definição 24 *Uma função $f : \mathbb{R} \to \mathbb{R}$ tal que $f(x) = ax^2 + bx + c$ para todo $x \in \mathbb{R}$, com $a, b, c \in \mathbb{R}$ e $a \neq 0$ é denominada* ***quadrática****.*

Em relação ao trinômio $ax^2 + bx + c$, é possível escrevê-lo de uma forma conveniente. Para isso, note que:

$$ax^2 + bx + c = a\left[x^2 + \frac{b}{a}x + \frac{c}{a}\right].$$

Como $x^2 + \dfrac{b}{a}x + \dfrac{c}{a}$ é parte do quadrado $\left[x + \dfrac{b}{a}\right]^2$, então, adicionando e subtraindo $\dfrac{b^2}{4a^2}$ à expressão entre colchetes, tem-se

$$ax^2 + bx + c = a\left[x^2 + 2 \cdot \frac{b}{2a} \cdot x + \frac{b^2}{4a^2} - \frac{b^2}{4a^2} + \frac{c}{a}\right]$$

Logo:

$$ax^2 + bx + c = a\left[\left(x + \frac{b}{2a}\right)^2 + \frac{4ac - b^2}{4a^2}\right] \tag{2.1}$$

A equação 2.1 é chamada de **forma canônica** do trinômio do segundo grau. Utilizando essa forma, as raízes da equação $ax^2 + bx + c = 0$ podem ser facilmente obtidas como vê-se na definição 25.

Definição 25 *As raízes da equação $ax^2 + bx + c = 0$, com $a \neq 0$ são os valores $x \in \mathbb{R}$ tais que $ax^2 + bx + c = 0$.*

A partir da forma canônica do trinômio do segundo grau, tem-se que

$$\begin{aligned} ax^2 + bx + c = 0 &\Leftrightarrow \left(x + \frac{b}{2a}\right)^2 + \frac{4ac - b^2}{4a^2} = 0 \\ &\Leftrightarrow \left(x + \frac{b}{2a}\right)^2 = \frac{b^2 - 4ac}{4a^2} \qquad (2.2) \\ &\Leftrightarrow x + \frac{b}{2a} = \pm\frac{\sqrt{b^2 - 4ac}}{2a} \qquad (2.3) \\ &\Leftrightarrow x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \end{aligned}$$

A passagem de 2.2 para 2.3 só faz sentido quando o **discriminante** $\Delta = b^2 - 4ac$ é maior ou igual a zero. Quando $\Delta < 0$, a equação dada não possui solução real pois $\left(x + \dfrac{b}{2a}\right)^2$ não pode ser negativo.

A Definição 26 é essencial para o Teorema 3 dado a seguir.

Definição 26 *Uma progressão aritmética de segunda ordem é uma sequência* $(y_1, y_2, y_3, y_4, \ldots)$ *na qual as diferenças entre cada termo e o termo anterior dadas por* $d_1 = y_2 - y_1$, $d_2 = y_3 - y_2$, $d_3 = y_4 - y_3$ *e assim por diante, formam uma progressão aritmética usual.*

Exemplo 1 *A sequência dos números triangulares dada por* $(T_n) = (1, 3, 6, 10, 15, ...)$ *é uma progressão aritmética de segunda ordem.*

Determinando a diferença entre cada termo e o termo anterior em (T_n) obtém-se $d_1 = 3-1 = 2$, $d_2 = 6 - 3 = 3$, $d_3 = 10 - 6 = 4$, e assim por diante, que formam a progressão aritmética $(2, 3, 4, 5, ...)$ de razão igual a 1. Portanto, de acordo com a definição 26, (T_n) é uma progressão aritmética de segunda ordem.

O Teorema 3 é uma caracterização das funções quadráticas.

Teorema 3 *A fim de que a função contínua* $f : \mathbb{R} \to \mathbb{R}$ *seja quadrática é necessário e suficiente que toda progressão aritmética não-constante* $x_1, x_2, \ldots, x_n, \ldots$, *seja transformada por* f *numa progressão aritmética de segunda ordem não-degenerada*[1] $y_1 = f(x_1), y_2 = f(x_2), \ldots, y_n = f(x_n), \ldots$

Demonstração: Para provar a necessidade, note inicialmente que uma progressão aritmética é uma função afim restrita aos números naturais. Com efeito, sendo $(x_1, x_2, \ldots, x_n, \ldots)$ uma progressão aritmética de razão r então seu n-ésimo termo $x_n = x_1 + (n-1)r$ pode ser escrito como $x_n = an + b$, onde $a = r$ e $b = x_1 - r$. Assim, ao restringir a função afim $h(x) = ax + b$ aos números naturais, obtém-se os termos $x_1 = h(1), x_2 = h(2), \ldots, x_n = h(n), \ldots$ da P.A.

Analogamente, se $(y_1, y_2, \ldots, y_n, \ldots)$ é uma P.A. de segunda ordem, existem $a, b, c \in \mathbb{R}$ tais que $y_n = an^2 + bn + c$ para todo $n \in \mathbb{N}$. Então, considerando a função $f(x) = ax^2 + bx + c$, ao restringir f aos números naturais, tem-se os termos $y_1 = f(1), y_2 = f(2), \ldots, y_n = f(n), \ldots$ da P.A. de segunda ordem. De fato, as diferenças sucessivas $d_1 = y_2 - y_1$, $d_2 = y_3 - y_2$,...,$d_n = y_{n+1} - y_n$,... formam uma progressão aritmética de primeira ordem com primeiro termo d_1 e razão r. Desta forma, seu n-ésimo termo é $d_n = d_1 + (n-1)r$, ou seja, $y_{n+1} - y_n = y_2 - y_1 + (n-1)r$, onde $n \geq 1$. Escrevendo $y_{n+1} = (y_{n+1} - y_n) + (y_n - y_{n-1}) + \cdots + (y_3 - y_2) + (y_2 - y_1) + y_1$, segue que $y_{n+1} = [d_1 + (n-1)r] + [d_1 + (n-2)r] + \cdots + (d_1 + r) + d_1 + y_1$. Logo

$$y_{n+1} = nd_1 + \frac{n(n-1)}{2} r + y_1 \tag{2.4}$$

[1] Uma progressão aritmética de segunda ordem é não-degenerada quando a diferença entre cada termo e o termo anterior formam uma P.A. com razão não nula.

para todo $n \in \mathbb{N}$. Como a igualdade 2.4 é válida para $n = 0$, então é possível escrever

$$\begin{aligned}
y_n &= (n-1)d_1 + \frac{(n-1)(n-2)}{2}r + y_1 \\
&= \frac{r}{2}n^2 + \left(d_1 - \frac{3r}{2}\right)n + r - d_1 + y_1 \\
&= an^2 + bn + c,
\end{aligned}$$

para todo $n \in \mathbb{N}$, onde $a = \frac{r}{2}$, $b = d_1 - \frac{3r}{2}$ e $c = r - d_1 + y_1$.

Para provar a suficiência, seja $f : \mathbb{R} \to \mathbb{R}$ uma função contínua que possui a propriedade de transformar toda P.A. não constante em uma P.A. de segunda ordem não-degenerada. Substituindo $f(x)$ por $g(x) = f(x) - f(0)$, nota-se que g tem as mesmas propriedades de f além de possuir a propriedade adicional de que $g(0) = 0$. Assim, existem reais $a \neq 0$ e b tais que $g(n) = an^2 + bn$ para todo $n \in \mathbb{N}$.

Seja $p \in \mathbb{N}$ um número fixado de forma arbitrária e considere a progressão aritmética $\left(\frac{1}{p}, \frac{2}{p}, \ldots, \frac{n}{p}, \ldots\right)$. Analogamente, existem $a' \neq 0$ e b' tais que $g\left(\frac{n}{p}\right) = a'n^2 + b'n$ para todo $n \in \mathbb{N}$. Logo, para todo $n \in \mathbb{N}$ tem-se

$$an^2 + bn = g(n) = g\left(\frac{np}{p}\right) = a'(np)^2 + b'(np) = (a'p^2)n^2 + (b'p)n.$$

Portanto, para todo $x = n \in \mathbb{N}$ as funções quadráticas $ax^2 + bx$ e $(a'p^2)x^2 + (b'p)x$ são coincidentes. Consequentemente, $a = a'p^2$ e $b = b'p$, ou seja, $a' = \frac{a}{p^2}$ e $b' = \frac{b}{p}$. Então, para quaisquer $n, p \in \mathbb{N}$ é válido $g\left(\frac{n}{p}\right) = a'n^2 + b'n = \frac{a}{p^2}n^2 + \frac{b}{p}n = a\left(\frac{n}{p}\right)^2 + b\left(\frac{n}{p}\right)$. Desta forma, as funções contínuas $g(x)$ e $ax^2 + bx$ são tais que $g(r) = ar^2 + br$ para todo número racional positivo $r = \frac{n}{p}$. Segue-se que $g(x) = ax^2 + bx$ para todo número real positivo x. De forma análoga, considerando a P.A. $(-1, -2, -3, \ldots)$, seria possível concluir que $g(x) = ax^2 + bx$ para todo $x \leq 0$. Portanto, pondo $f(0) = c$, tem-se $f(x) = g(x) + c$, ou seja, $f(x) = ax^2 + bx + c$ para todo $x \in \mathbb{R}$.

Definição 27 *Dados um ponto F e uma reta d, tais que $F \notin d$. O conjunto dos pontos do plano que são equidistantes de F e d é chamado de parábola, ou seja,*

$$\{P = (x, y) \ : \ d(P, F) = d(P, d)\}.$$

onde $d(P, d)$ denota a distância de P à reta d. A reta d é denominada diretriz e o ponto F chama-se foco.

A reta que passa por F e é perpendicular à diretriz d é chamada de eixo da parábola. O ponto médio do segmento que possui como extremos o foco e a interseção do eixo com a diretriz chama-se o vértice da parábola. Esse é o ponto da parábola mais próximo da diretriz.

Definição 28 *O gráfico da função quadrática $f : \mathbb{R} \to \mathbb{R}$ definida por $f(x) = ax^2 + bx + c$, onde $a \neq 0$, é o conjunto dado por $G_f = \{(x, y) \in \mathbb{R} \times \mathbb{R} \ : \ y = ax^2 + bx + c\}$*

O teorema 4 mostra que o gráfico da função quadrática, dado na definição 28, é uma parábola.

Teorema 4 *Para $a, b, c \in \mathbb{R}$, com $a \neq 0$, o gráfico da função quadrática $f : \mathbb{R} \to \mathbb{R}$ dada por $f(x) = ax^2+bx+c$ é a parábola de eixo $\left\{x = -\frac{b}{2a}\right\}$ e vértice $V = \left(-\frac{b}{2a}, -\frac{\Delta}{4a}\right)$, com concavidade voltada para cima se $a > 0$, e concavidade voltada para baixo se $a < 0$.*

Demonstração: Pela definição 27, deve-se determinar $x_0, y_0, k \in \mathbb{R}$, $y_0 \neq k$, tais que para $F = (x_0, y_0)$ e d a reta $y = k$, tem-se $P \in G_f \Leftrightarrow d(P, F) = d(P, d)$. Sendo $P = (x, y)$, pela definição 28, segue que $P \in G_f \Leftrightarrow y = ax^2 + bx + c$. Determinando a distância entre os pontos P e F e a distância entre o ponto P e a diretriz d, tem-se que

$$d(P, F) = d(P, d) \Leftrightarrow \sqrt{(x - x_0)^2 + (y - y_0)^2} = |y - k|.$$

Logo,

$$y = ax^2 + bx + c \Leftrightarrow (x - x_0)^2 + (y - y_0)^2 = (y - k)^2.$$

Note que,

$$\begin{aligned}
(x - x_0)^2 + (y - y_0)^2 = (y - k)^2 &\Leftrightarrow x^2 - 2xx_0 + x_0^2 - 2yy_0 + y_0^2 = -2yk + k^2 \\
&\Leftrightarrow -2yy_0 + 2yk = -x^2 + 2x_0x - x_0^2 - y_0^2 + k^2 \\
&\Leftrightarrow -2(y_0 - k)y = -x^2 + 2x_0x - x_0^2 - y_0^2 + k^2 \\
&\Leftrightarrow y = \frac{1}{2(y_0 - k)}x^2 - \frac{x_0}{y_0 - k}x + \frac{x_0^2 + y_0^2 - k^2}{2(y_0 - k)}.
\end{aligned}$$

Então,

$$y = ax^2 + bx + c \Leftrightarrow y = \frac{1}{2(y_0 - k)}x^2 - \frac{x_0}{y_0 - k}x + \frac{x_0^2 + y_0^2 - k^2}{2(y_0 - k)}.$$

Os valores de x_0, y_0 e k ficam determinados, ao ser resolvido o sistema

$$\begin{cases} \dfrac{1}{2(y_0-k)} = a \\ -\dfrac{x_0}{y_0-k} = b \\ \dfrac{x_0^2}{2(y_0-k)} + \dfrac{1}{2}(y_0+k) = c \end{cases}$$

então, dividindo membro a membro as duas primeiras equações, obtém-se $x_0 = -\frac{b}{2a}$. Substituindo a primeira igualdade e o valor encontrado de x_0 na terceira equação, tem-se que

$$\frac{x_0^2}{2(y_0-k)} + \frac{1}{2}(y_0+k) = c \Rightarrow y_0 + k = 2\left(c - x_0^2 \cdot \frac{1}{2(y_0-k)}\right) = 2\left(c - \frac{b^2}{4a^2} \cdot a\right) = -\frac{\Delta}{2a};$$

da primeira equação, segue que $\dfrac{1}{2(y_0-k)} = a \Rightarrow y_0 - k = \dfrac{1}{2a}$. Então, resolvendo o sistema

$$\begin{cases} y_0 - k = \dfrac{1}{2a} \\ y_0 + k = -\dfrac{\Delta}{2a} \end{cases}$$

encontra-se $y_0 = \dfrac{1-\Delta}{4a}$ e $k = -\dfrac{1+\Delta}{4a}$.

Como o vértice V da parábola é a interseção da reta $x = -\dfrac{b}{2a}$ com o gráfico, então sua ordenada é dada por

$$y_v = a\left(-\frac{b}{2a}\right)^2 + b\left(-\frac{b}{2a}\right) + c = -\frac{\Delta}{4a}.$$

■

Proposição 10 *Sendo $f : \mathbb{R} \to \mathbb{R}$ tal que $f(x) = ax^2 + bx + c$ a função quadrática, tem-se:*

(a) Se $a > 0$, então a imagem da função f é dada por Im(f) $= \left[-\dfrac{\Delta}{4a}, +\infty\right)$.

(b) Se $a < 0$, então a imagem de f é dada por Im(f) $= \left(-\infty, -\dfrac{\Delta}{4a}\right]$.

Além disso, em qualquer um dos casos (a) e (b), tem-se $f(x) = -\dfrac{\Delta}{4a} \Leftrightarrow x = -\dfrac{b}{2a}$.

Demonstração: Para obter a imagem de f, é necessário encontrar os $y \in \mathbb{R}$ para os quais a equação $ax^2 + bx + c = y$, isto é, a equação de segundo grau $ax^2 + bx + (c - y) = 0$, tenha solução. Tal equação tem solução se, e somente se, seu discriminate é não negativo, isto é, $b^2 - 4a(c - y) \geq 0$. Mas, como $b^2 - 4ac = \Delta$, os y procurados são exatamente as soluções da inequação (em y) $\Delta + 4ay \geq 0$. Agora, considerando separadamente os casos $a > 0$ e $a < 0$, segue que, se $a > 0$, então $4ay + \Delta \geq 0 \Leftrightarrow y \geq -\frac{\Delta}{4a}$ e daí, tem-se que

$$Im(f) = \left\{y \in \mathbb{R}; y \geq -\frac{\Delta}{4a}\right\} = \left[-\frac{\Delta}{4a}, +\infty\right);$$

se $a < 0$, então $4ay + \Delta \geq 0 \Leftrightarrow y \leq -\frac{\Delta}{4a}$, de forma que

$$Im(f) = \left\{y \in \mathbb{R}; y \leq -\frac{\Delta}{4a}\right\} = \left(-\infty, -\frac{\Delta}{4a}\right].$$

Para o que falta, note que, para $y \in Im(f)$, as soluções da equação $ax^2 + bx + c = y (\Leftrightarrow ax^2 + bx + (c - y) = 0)$ são

$$x = \frac{-b \pm \sqrt{b^2 - 4a(c - y)}}{2a} = \frac{-b \pm \sqrt{\Delta + 4ay}}{2a}. \tag{2.5}$$

Portanto, a equação $f(x) = y$ admite uma solução única se, e só se, $\Delta + 4ay = 0$, ou, equivalentemente, se, e só se, $y = -\frac{\Delta}{4a}$; sendo esse o caso, tem-se, a partir de 2.5, que $x = -\frac{b}{2a}$.

■

Corolário 1 *A função quadrática $f(x) = ax^2 + bx + c$ tem sinal constante se, e somente se, $\Delta \leq 0$. Neste caso, tem-se $af(x) \geq 0$, para todo $x \in \mathbb{R}$. De outro modo:*

(a) Se $\Delta \leq 0$ e $a > 0$, então $f(x) \geq 0$ para todo $x \in \mathbb{R}$.

(b) Se $\Delta \leq 0$ e $a < 0$, então $f(x) \leq 0$ para todo $x \in \mathbb{R}$.

Demonstração: Será analisado o caso $a > 0$, sendo a análise do outro caso totalmente análoga. Sendo $a > 0$ e $\Delta \leq 0$, segue da Proposição 10 que

$$f(x) \geq -\frac{\Delta}{4a} \geq 0, \ \forall x \in \mathbb{R}.$$

Reciprocamente, suponha que $a > 0$ e que f tem sinal constante. Pelo item (a) da proposição 10, a imagem de f contém números positivos, de forma que a constância de sinal de f garante que deve-se ter $f(x) \geq 0$, para todo $x \in \mathbb{R}$. Em particular, deve-se ter

$$-\frac{\Delta}{4a} = f\left(\frac{-b}{2a}\right) \geq 0.$$

Logo, $\Delta \leq 0$. ■

Proposição 11 *Sejam $a, b, c \in \mathbb{R}$, com $a \neq 0$, e $f(x) = ax^2 + bx + c$.*

(a) Se $a > 0$, então f é decrescente em $\left(-\infty, -\frac{b}{2a}\right]$ e crescente em $\left[-\frac{b}{2a}, +\infty\right)$.

(b) Se $a < 0$, então f é crescente em $\left(-\infty, -\frac{b}{2a}\right]$ e decrescente em $\left[-\frac{b}{2a}, +\infty\right)$.

Demonstração: Será feita a demonstração do item (a), sendo a demonstração do item (b) análoga. Considerando $x_2 > x_1 \geq -\frac{b}{2a}$, tem-se

$$\begin{aligned} f(x_2) - f(x_1) &= (ax_2^2 + bx_2 + c) - (ax_1^2 + bx_1 + c) \\ &= a(x_2^2 - x_1^2) + b(x_2 - x_1) \\ &= a(x_2 - x_1)\left(x_2 + x_1 + \frac{b}{a}\right) > 0, \end{aligned}$$

uma vez que $x_2 > x_1 \geq -\frac{b}{2a}$ implica $x_2 - x_1 > 0$ e $x_2 + x_1 + \frac{b}{a} > 0$. Logo, f é crescente em $\left[-\frac{b}{2a}, +\infty\right)$ quando $a > 0$. Analogamente, mostra-se que f é decrescente em $\left(-\infty, -\frac{b}{2a}\right]$. ∎

Proposição 12 *Em relação à função quadrática $f(x) = ax^2 + bx + c$, se $a < 0$ (resp. $a > 0$), então $-\frac{b}{2a}$ é o único ponto de mínimo (resp. máximo) de f. Além disso, o valor mínimo (resp. valor máximo) de f é $-\frac{\Delta}{4a}$.*

2.2.3 Funções exponenciais

Definição 29 *Dado um número real $a > 0$, com $a \neq 1$, a função $f : \mathbb{R} \to \mathbb{R}$ tal que $f(x) = a^x$ é chamada de **função exponencial de base a**.*

Definição 30 *Uma função $g : \mathbb{R} \to \mathbb{R}$ é de **tipo exponencial** quando para todo $x \in \mathbb{R}$ tem-se $g(x) = ba^x$, onde a e b são constantes positivas.*

Teorema 5 *Se $a > 1$, a sequência $(a_n)_{n \geq 1}$ dada por $a_n = a^n$, é ilimitada superiormente. Se $0 < a < 1$ então a sequência $(a_n)_{n \geq 1}$ dada por $a_n = a^n$, é descrescente e limitada inferiormente.*

Demonstração: Se $a > 1$ então, multiplicando os membros desta desigualdade por a^n, obtém-se $a^{n+1} > a^n$. Logo,

$$a > 1 \Rightarrow 1 < a < a^2 < \cdots < a^n < a^{n+1} < \cdots ,$$

o que mostra que a sequência é crescente. Para mostrar que (a_n) é ilimitada deve-se provar que não existe $c \in \mathbb{R}$ superior a todas as potências a^n, ou seja, qualquer que seja $c \in \mathbb{R}$ existe $n \in \mathbb{N}$ tal que $a^n > c$. Com efeito, escrevendo $a = 1 + d$, $d > 0$, pela desigualdade de Bernoulli, tem-se que $a^n > 1 + nd$. Assim, tomando $n > \frac{c-1}{d}$, segue que

$$n > \frac{c-1}{d} \Rightarrow 1 + nd > c.$$

Portanto, $a^n > c$ e a sequência (a_n) é ilimitada superiormente. De forma análoga, se $0 < a < 1$, multiplicando a desigualdade por a^n, tem-se $a^n > a^{n+1}$. Com isso,

$$0 < a < 1 \Rightarrow 1 > a > a^2 > \cdots > a^n > a^{n+1} > \cdots ,$$

logo (a_n) é descrescente. Como $0 < a < 1$, então $a^n > 0$ para todo $n \in \mathbb{N}$, mostrando que a sequência é limitada inferiormente por 0. ■

Lema 1 *Dado $a \in \mathbb{R}^+$, com $a \neq 1$, em todo intervalo de $\mathbb{R}^+$ existe uma potência a^r, com $r \in \mathbb{Q}$.*

Demonstração: Dados $\alpha, \beta \in \mathbb{R}$ tais que $0 < \alpha < \beta$, deve-se encontrar $r \in \mathbb{Q}$ tal que $\alpha \leq a^r \leq \beta$. Suponha $a, \alpha > 1$. Como pelo teorema 5 as potências de expoente natural de números maiores que 1 crescem acima de qualquer cota prefixada, é possível obter números M e n, com $M, n \in \mathbb{N}$, tais que

$$\alpha < \beta < a^M \text{ e } 1 < a < \left(1 + \frac{\beta - \alpha}{a^M}\right)^n$$

com isso, resulta que $1 < a^{\frac{1}{n}} < 1 + \dfrac{\beta - \alpha}{a^M}$ e $0 < a^M(a^{\frac{1}{n}} - 1) < \beta - \alpha$. Por isso $\dfrac{m}{n} \leq M \Rightarrow 0 < a^{\frac{m}{n}}(a^{\frac{1}{n}} - 1) < \beta - \alpha \Leftrightarrow 0 < a^{\frac{m+1}{n}} - a^{\frac{m}{n}} < \beta - \alpha$. Desta forma, as potências $a^0, a^{\frac{1}{n}}, a^{\frac{2}{n}}, \cdots, a^M$ são extremos de intervalos consecutivos, todos de comprimento menor do que o comprimento $\beta - \alpha$ do intervalo $[\alpha, \beta]$. Como $[\alpha, \beta] \subset [1, a^M]$, pelo menos um desses extremos, por exemplo $a^{\frac{m}{n}}$, está contido no intervalo $[\alpha, \beta]$. Os demais casos são análogos. ■

A função exponencial de base a é definida de modo a ter as seguintes propriedades, quaisquer sejam $x, y \in \mathbb{R}$:

(1) $a^x \cdot a^y = a^{x+y}$;

(2) $a^1 = a$;

(3) $x < y \Rightarrow a^x < a^y$ quando $a > 1$ e $x < y \Rightarrow a^y < a^x$ quando $0 < a < 1$.

Se uma função $f : \mathbb{R} \to \mathbb{R}$ tem a propriedade (1), ou seja, $f(x + y) = f(x) \cdot f(y)$, então f não assume o valor 0, só assumindo esse valor quando f é identicamente nula.

A partir da propriedade (3), é possível definir o valor $f(x) = a^x$ quando x é irracional. Supondo $a > 1$, tem-se que a^x é tal que $r < x < s,$ com $r, s \in \mathbb{Q} \Rightarrow a^r < a^x < a^s$. Se $0 < a < 1$, então a^x é tal que $r < x < s,$ com $r, s \in \mathbb{Q} \Rightarrow a^s < a^x < a^r$ No caso de $a > 1$, a^x é o número real cujas aproximações por falta são a^r, com $r < x, r \in \mathbb{Q}$, e cujas aproximações por excesso são a^s, com $x < s, s \in \mathbb{Q}$. E, sendo $0 < a < 1$, a^x é o número real cujas aproximações por falta são $a^s, s \in \mathbb{Q}$, e cujas aproximações por excesso são $a^r, r \in \mathbb{Q}$. Não é possível existir dois números reais diferentes, A, B com a propriedade (3), pois, caso existissem, supondo $A < B$ e $a > 1$, seria possível ter

$$r < x < s, r, s \in \mathbb{Q} \Rightarrow a^r < A < B < a^s$$

com isso o intervalo $[A, B]$ não conteria nenhuma potência de a com expoente racional, contrariando o Lema 1. Ademais, tem-se

(4) A função $f : \mathbb{R} \to \mathbb{R}^+$, definida por $f(x) = a^x$, é ilimitada superiormente;

(5) A função exponencial é contínua;

(6) A função exponencial $f : \mathbb{R} \to \mathbb{R}^+, f(x) = a^x, a \neq 1$, é sobrejetiva.

O Teorema 6 apresenta a caracterização da função exponencial.

Teorema 6 *Seja $f : \mathbb{R} \to \mathbb{R}^+$ uma função monótona injetiva. São equivalentes as seguintes afirmações:*

(1) $f(nx) = f(x)^n$ para todo $n \in \mathbb{Z}$ e todo $x \in \mathbb{R}$;

(2) $f(x) = a^x$ para todo $x \in \mathbb{R}$, onde $a = f(1)$;

(3) $f(x + y) = f(x) \cdot f(y)$ para quaisquer $x, y \in \mathbb{R}$.

Demonstração: Serão demonstradas as implicações (1) $\Rightarrow$ (2) $\Rightarrow$ (3) $\Rightarrow$ (1). Para mostrar que (1) $\Rightarrow$ (2) note inicialmente que a hipótese (1) ocasiona que, para todo número racional $r = \dfrac{m}{n}$, com $m \in \mathbb{Z}$ e $n \in \mathbb{N}$, tem-se $f(rx) = f(x)^r$. De fato, como $nr = m$, é possível escrever

$$f(rx)^n = f(nrx) = f(mx) = f(x)^m,$$

logo $f(rx) = f(x)^{\frac{m}{n}} = f(x)^r$. Assim, se fizer $f(1) = a$, tem-se $f(r) = f(r \cdot 1) = f(1)^r = a^r$ para todo $r \in \mathbb{Q}$. Para completar a demonstração de que (1) $\Rightarrow$ (2) suponha que f é crescente, logo $1 = f(0) < f(1) = a$. Admitindo, por absurdo, que existe $x \in \mathbb{R}$ tal que $f(x) \neq a^x$, por exemplo, que seja $f(x) < a^x$. (O caso $f(x) > a^x$ seria feito de maneira análoga.) Então, pelo Lema 1, existe $r \in \mathbb{Q}$ tal que $f(x) < a^r < a^x$, ou seja, $f(x) < f(r) < a^x$. Como f é crescente, tendo $f(x) < f(r)$ conclui-se que $x < r$. Por outro lado, tem-se $a^r < a^x$, logo $r < x$, o que é uma contradição. Logo, segue que (1) $\Rightarrow$ (2). Será feita agora, a prova de que (2) $\Rightarrow$ (3). Sendo $f(x) = a^x$ para todo $x \in \mathbb{R}$ e $a = f(1)$. Se $y \in \mathbb{R}$ então

$$f(x + y) = a^{x+y} = a^x \cdot a^y = f(x) \cdot f(y).$$

Por fim, para mostrar que (3) $\Rightarrow$ (1), seja $f(x+y) = f(x)\cdot f(y)$, quaisquer sejam $x, y \in \mathbb{R}$. Sendo $n \in \mathbb{N}$, tem-se

$$f(nx) = f(\underbrace{x + x + x + \cdots + x}_{\text{n vezes}}) = \underbrace{f(x)\cdot f(x)\cdot f(x)\cdots f(x)}_{\text{n vezes}} = f(x)^n$$

Resta mostrar que $f(-nx) = f(x)^{-n}$. Desta forma, determinando $f(-x)$

$$f(-x)\cdot f(x) = f(-x+x) = f(0) = 1$$

logo, $f(-x) = \left(\dfrac{1}{f(x)}\right)$. Portanto,

$$\begin{aligned} f(-nx) &= f(\underbrace{(-x) + (-x) + (-x) + \cdots + (-x)}_{\text{n vezes}}) \\ &= \underbrace{f(-x)\cdot f(-x)\cdot f(-x)\cdots f(-x)}_{\text{n vezes}} \\ &= \frac{1}{f(x)}\cdot\frac{1}{f(x)}\cdot\frac{1}{f(x)}\cdots\frac{1}{f(x)} \\ &= \frac{1}{f(x)^n} = f(x)^{-n} \end{aligned}$$

■

O Teorema 7 fornece uma caracterização para as funções do tipo exponencial dadas na definição 30.

Teorema 7 *Seja $g : \mathbb{R} \to \mathbb{R}^+$ uma função monótona injetiva tal que, para quaisquer $x, h \in \mathbb{R}$, o acréscimo relativo $\dfrac{g(x+h) - g(x)}{g(x)}$ dependa apenas de h, mas não de x. Então, se $b = g(0)$ e $a = g(1)/g(0)$, tem-se $g(x) = ba^x$ para todo $x \in \mathbb{R}$.*

Demonstração: Suponha que $\varphi(h) = \dfrac{g(x+h)}{g(x)}$ independe de x. Substituindo, se necessário, $g(x)$ por $f(x) = \dfrac{g(x)}{b}$, onde $b = g(0)$, obtém-se $f : \mathbb{R} \to \mathbb{R}^+$ monótona injetiva, com $\dfrac{f(x+h)}{f(x)}$ independente de x e com f(0)=1. Assim, pondo $x = 0$ em $\varphi(h) = \dfrac{f(x+h)}{f(x)}$, obtém-se $\varphi(h) = f(h)$ para todo $h \in \mathbb{R}$. Com isso, percebe-se que a função monótona injetiva f cumpre $f(x+h) = f(x)\cdot f(h)$, ou seja $f(x+y) = f(x)\cdot f(y)$ para quaisquer $x, y \in \mathbb{R}$. Segue-se então do teorema 6 que $f(x) = a^x$, portanto $g(x) = bf(x) = ba^x$ ■

2.2.4 Funções logarítmicas

Definição 31 *Sendo a e x números reais positivos, com $a \neq 1$, chama-se* ***logaritmo*** *de x na base a o expoente ao qual se deve elevar a base a de modo que a potência obtida seja igual a x. Ou seja, $\log_a x = y \Leftrightarrow a^y = x$*

Definição 32 *A função $\log_a : \mathbb{R}^+ \to \mathbb{R}$ que associa a cada número real positivo x o número real $log_a x$ é chamada* ***função logarítmica de base a****.*

Proposição 13 *Para quaisquer $x, y \in \mathbb{R}^+$ e $c \in \mathbb{R}$, tem-se:*

a) $\log_a(xy) = \log_a x + \log_a y$.

b) $\log_a \left(\dfrac{x}{y}\right) = \log_a x - \log_a y$.

c) $\log_a x^c = c \cdot \log_a x$.

Demonstração: a) Considere $m = \log_a x$ e $n = \log_a y$. Então, $a^m = x$ e $a^n = y$, portanto, da relação $a^m \cdot a^n = a^{m+n}$, segue que $xy = a^m \cdot a^n = a^{m+n}$, ou seja,

$$\log_a(xy) = \log_a(a^{m+n}) = m + n = \log_a x + \log_a y.$$

b) De forma análoga, sejam $m = \log_a x$ e $n = \log_a y$, logo $a^m = x$ e $a^n = y$. Então, da propriedade $\dfrac{a^m}{a^n} = a^{m-n}$, tem-se

$$\frac{x}{y} = \frac{a^m}{a^n} = a^{m-n}$$

portanto,

$$\log_a \left(\frac{x}{y}\right) = \log_a(a^{m-n}) = m - n = \log_a x - \log_a y.$$

c) Fazendo $m = \log_a x$, segue que $a^m = x$ e, de $(a^m)^c = a^{m \cdot c}$, tem-se $x^c = (a^m)^c = a^{m \cdot c}$ logo,

$$log_a x^c = \log_a(a^{m \cdot c}) = m \cdot c = c \cdot \log_a x.$$

Proposição 14 *A função $\log_a : \mathbb{R}^+ \to \mathbb{R}$ é*

a) crescente quando $a > 1$;

b) decrescente quando $0 < a < 1$.

Como $a^0 = 1$, então $\log_a 1 = 0$. Para $a > 1$, quando $0 < x < 1$ tem-se $\log_a x < 0$ e se $x > 1$ tem-se $\log_a x > 0$. Por outro lado, se $0 < a < 1$ $x > 1$ então $\log_a x > 0$ quando $0 < x < 1$ e $\log_a x < 0$ quando $x > 1$.

Teorema 8 *Seja $f : \mathbb{R}^+ \to \mathbb{R}$ uma função monótona injetiva tal que $f(xy) = f(x) + f(y)$, quaisquer que sejam $x, y \in \mathbb{R}^+$. Então, existe $a > 0$ tal que $f(x) = \log_a x$ para todo $x \in \mathbb{R}^+$.*

Demonstração: Suponha f crescente e que exista $a \in \mathbb{R}$ tal que $f(a) = 1$. Tem-se $f(1) = f(1 \cdot 1) = f(1) + f(1)$, logo $f(1) = 0$. Como f é crescente e $f(a) = 1 > 0 = f(1)$, tem-se $a > 1$. Para todo $m \in \mathbb{N}$ é válido

$$
\begin{aligned}
f(a^m) &= f(\underbrace{a \cdot a \cdot a \cdots a}_{\text{m vezes}}) \\
&= \underbrace{f(a) + f(a) + f(a) + \cdots + f(a)}_{\text{m vezes}} \\
&= 1 + 1 + 1 + \cdots + 1 = m, \\
0 &= f(1) = f(a^m \cdot a^{-m}) \\
&= f(a^m) + f(a^{-m}) = m + f(a^{-m}),
\end{aligned}
$$

logo $f(a^{-m}) = -m$. Se $r = \frac{m}{n}$, com $m \in \mathbb{Z}$ e $n \in \mathbb{N}$, então $rn = m$, portanto

$$m = f(a^m) = f(a^{rn}) = f((a^r)^n) = n \cdot f(a^r)$$

daí $f(a^r) = \frac{m}{n} = r$. Se $x \in \mathbb{R}$ é irracional então, para $r, s \in \mathbb{Q}$ tem-se:

$$r < x < s \Rightarrow a^r < a^x < a^s \Rightarrow f(a^r) < f(a^x) < f(a^s) \Rightarrow r < f(a^x) < s.$$

Assim, $f(a^x) = x$ para todo $x \in \mathbb{R}$. Portanto $f(y) = log_a y$ para todo $y > 0$.

Considere o caso geral, em que $g : \mathbb{R}^+ \to \mathbb{R}$ é uma função crescente tal que $g(xy) = g(x) + g(y)$. Desta forma $g(1) = 0$ e, como $1 < 2$, deve-se ter $g(2) = b > 0$. A nova função $f : \mathbb{R}^+ \to \mathbb{R}$ definida por $f(x) = \dfrac{g(x)}{b}$, é crescente, transforma soma em produtos e cumpre $f(2) = 1$. Portanto, pela primeira parte da demonstração, tem-se $f(x) = \log_2 x$ para todo $x > 0$. O que significa que, para todo $x > 0$, é válido

$$x = 2^{f(x)} = 2^{\frac{g(x)}{b}} = (2^{\frac{1}{b}})^{g(x)} = a^{g(x)},$$

onde $a = 2^{\frac{1}{b}}$. Tomando $\log_a$ de ambos os membros da igualdade $a^{g(x)} = x$ segue-se que $g(x) = \log_a x$. O caso de f decrescente é feito de forma análoga. ∎

Definição 33 *A função $f : \mathbb{R}^+ \to \mathbb{R}$ dada por $f(x) = \log_e x$ para todo $x \in \mathbb{R}^+$ é denominada* ***função logarítmica de base*** *e, onde e é o número irracional equivalente ao limite de $\left(1 + \frac{1}{n}\right)^n$ quando n tende ao infinito.*

O logaritmo de base e, $\log_e x$, pode ser representado por $\ln x$, denominado logaritmo natural de x.

2.3 Aplicações

De acordo com os PCNEM [7], dentre os objetivos do Ensino está o desenvolvimento de conhecimentos práticos que auxiliem nas necessidades da vida contemporânea. Conforme o documento, o ensino da Matemática no Ensino Médio deve propiciar aos educandos o desenvolvimento da capacidade de utilizar e aplicar a Matemática em situações reais, com destaque para as outras áreas do conhecimento.

Por isso, essa seção tem como objetivo, apresentar aplicações de funções afim, quadrática, exponencial e logarítmica em situações presentes na realidade.

2.3.1 Imposto de Renda

O imposto de renda é um tributo cobrado anualmente sobre os rendimentos de uma pessoa física ou jurídica. Para calcular o valor do imposto a ser pago por uma pessoa física, desconta-se do total de vencimentos a contribuição previdenciária e as deduções legais, por exemplo dependentes declarados. Ao valor obtido após as deduções, chamado de base de cálculo para o imposto de renda, aplica-se a alíquota, que é a taxa percentual incidente sobre a base de cálculo. A alíquota que deve ser aplicada à cada faixa salarial é determinada de acordo com a tabela de incidência mensal do IRPF determinada pela Receita Federal, de modo que o imposto a ser pago varie continuamente com a renda.

Desde o ano de 2015, a tabela de incidência mensal para cálculo do imposto sobre a renda das pessoas físicas é:

Tabela 2.1: Incidência mensal para cálculo do imposto sobre a renda das pessoas físicas

Base de cálculo (R$)	Alíquota (%)
Até 1903,98	-
De 1903,99 até 2826,65	7,5
De 2826,66 até 3751,05	15
De 3751,06 até 4664,68	22,5
Acima de 4664,68	27,5

Fonte: Receita Federal do Brasil

Como o imposto mensal devido deve variar continuamente com a renda, então, o seu cálculo deve ser feito separando a base de cálculo nas faixas determinadas na Tabela 2.1 e aplicando a alíquota referente à cada faixa. Assim, considerando como x a base de cálculo para o imposto de renda e a a alíquota a ser aplicada, o imposto mensal (I) a ser pago pode ser calculado da seguinte forma:

1. Para um salário com base de cálculo x de R$ 1903,99 a R$ 2826,65, desconta-se a parcela isenta de R$ 1903,98 e, ao resultado, aplica-se a alíquota de 7,5 %. Logo:

$$I = 0,075 \cdot (x - 1903,99) = 0,075x - 142,80$$

2. Considerando $2826,66 \leq x \leq 3751,05$, deve-se aplicar a alíquota de 7,5 % à R$ 922,66 referente à parte da base de cálculo que se encontra na faixa de R$ 1903,99 a R$ 2826,65, subtrair de x a parcela isenta de R$ 1903,98 e a parcela de R$ 922,66 e, à diferença, empregar a alíquota de 15% adicionando os resultados encontrados. Portanto:

$$I = 0,075 \cdot 922,66 + 0,15 \cdot (x - 1903,98 - 922,66) = 0,15x - 354,80$$

3. Analogamente, se $3751,06 \leq x \leq 4664,68$, tem-se que:

$$I = 0,075 \cdot 922,66 + 0,15 \cdot 924,39 + 0,225 \cdot (x - 1903,98 - 922,66 - 924,39) = 0,225x - 636,13$$

4. E, por fim, se $x > 4664,68$ então:

$$\begin{aligned} I &= 0,075 \cdot 922,66 + 0,15 \cdot 924,39 + 0,225 \cdot 913,62 + \\ &\quad 0,275 \cdot (x - 1903,98 - 922,66 - 924,39 - 913,62) \\ I &= 0,275x - 869,36 \end{aligned}$$

Portanto, o imposto de renda mensal a ser pago para uma base de cálculo qualquer pode ser definido através da função $I : \mathbb{R}^+ \to \mathbb{R}$ dada por

$$I = \begin{cases} 0, & \text{se} \quad x < 1903,99, \\ 0,075x - 142,8, & \text{se} \quad 1903,99 \leq x \leq 2826,65, \\ 0,15x - 354,8, & \text{se} \quad 2826,66 \leq x \leq 3751,05, \\ 0,225x - 636,13, & \text{se} \quad 3751,06 \leq x \leq 4664,68, \\ 0,275x - 869,36, & \text{se} \quad x > 4664,68. \end{cases}$$

Os valores fixos que são subtraídos em cada intervalo são chamados de parcela a deduzir do imposto. Esses valores já ficam pré-determinados na tabela de incidência mensal para cálculo do imposto sobre a renda das pessoas físicas fornecida pela Receita Federal do Brasil.

Em cada intervalo de I o gráfico é uma reta, onde cada uma possui uma inclinação diferente. Os extremos x de cada intervalo satisfazem a igualdade da faixa anterior e da faixa a que pertencem, por isso, essa retas ficam ligadas umas às outras. Na Figura 2.1 é possível observar uma representação gráfica da função I.

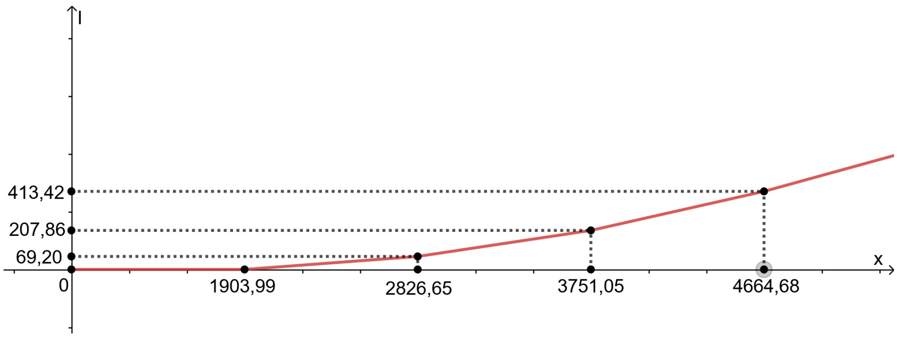

Figura 2.1: Gráfico da função I

Pode-se concluir que o imposto de renda mensal devido por uma pessoa física pode ser calculado através de uma função definida em partes, onde em cada intervalo o valor de I é determinado através de uma **função afim**.

2.3.2 Movimento Uniformemente Variado

Quando um corpo se move com velocidade variando ao longo do tempo, diz-se que esse corpo está sofrendo uma aceleração. E, no caso da aceleração no movimento ser constante e não nula, o movimento é chamado uniformemente variado.

A aceleração escalar média, dada por a_m, de um corpo em um intervalo qualquer de tempo é a taxa de variação da velocidade nesse intervalo, ou seja, considerando como v_0 a velocidade inicial, v a velocidade no instante t e t_0 o instante inicial, tem-se:

$$a_m = \frac{v - v_0}{t - t_0} \tag{2.6}$$

Como no movimento uniformemente variado a aceleração é constante em todo o movimento então, considerando $t_0 = 0$, tem-se da equação 2.6 que a aceleração a é dada por:

$$a = \frac{v - v_0}{t}$$

O que implica:

$$v = v_0 + at \tag{2.7}$$

Nesse tipo de movimento, a velocidade média entre os intantes $t_0 = 0$ e t é a média aritmética das velocidades nesses instantes e, em qualquer movimento, a velocidade média é dada pela razão entre o espaço percorrido e o tempo gasto. Logo, sendo s_0 a posição inicial e s a posição do corpo no instante t do movimento, tem-se que:

$$\frac{s - s_0}{t - t_0} = \frac{v_0 + v}{2} \Rightarrow s - s_0 = \tfrac{1}{2}t(v_0 + v) \Rightarrow s - s_0 = \tfrac{1}{2}v_0 t + \tfrac{1}{2}vt$$

Assim,

$$s = s_0 + \frac{1}{2}v_0 t + \frac{1}{2}vt \tag{2.8}$$

Substituindo a equação 2.7 em 2.8, segue que:

$$s = s_0 + \frac{1}{2}v_0 t + \frac{1}{2}(v_0 + at)t$$

Portanto,

$$s = \frac{1}{2}at^2 + v_0 t + s_0. \tag{2.9}$$

O que mostra que a posição s em função do instante t de um corpo se movendo em movimento uniformemente variado é dada por uma **função quadrática**.

A trajetória de um projétil (uma pedra, uma bola, uma bala, etc.) em um lançamento oblíquo é dada em um plano, onde a velocidade inicial é dada pelo vetor v_0, a aceleração é constante e igual à aceleração da gravidade dirigida para baixo. Nesse lançamento, o projétil não possui aceleração horizontal e considere a resistência do ar desprezada.

O movimento descrito pelo projétil pode ser decomposto em dois movimentos: horizontal, com aceleração nula, e vertical, com aceleração constante para baixo.

No movimento horizontal, como a aceleração é nula, a componente horizontal da velocidade é igual à velocidade inicial v_{0x} em toda a trajetória, assim, o movimento é uniforme. Sendo x_0 a posição inicial, a posição horizontal do projétil no instante t é dada por $x = x_0 + v_{0x}t$.

No movimento vertical, a aceleração é constante, igual à da gravidade e dirigida para baixo, então pode-se afirmar que $a = -g$ e, portanto, o movimento é uniformemente variado. Sendo y_0, a posição inicial e v_{0y}, a velocidade inicial do movimento, tem-se da Equação 2.9 que a posição vertical do projétil no instante t é dada por $y = -\dfrac{1}{2}gt^2 + v_{0y}t + y_0$

Desta forma, conclui-se que a trajetória do projétil é uma parábola. Considerando um lançamento em $x_0 = 0$ e $y_0 = 0$ com velocidade inicial dada pelo vetor v_0, sua trajetória pode ser observada na Figura 2.2.

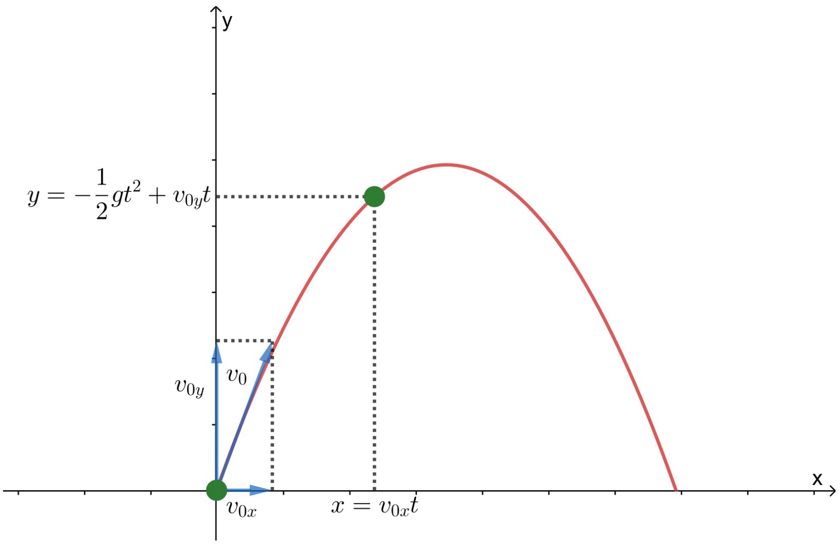

Figura 2.2: Trajetória do projétil

2.3.3 Desintegração radioativa

A desintegração radioativa é um fenômeno natural na qual os átomos de uma substância radioativa emitem partículas e se desintegram até transformarem-se em uma substância não radioativa. Isso acontece de tal forma que, em um dado instante, a quantidade de matéria desintegrada é proporcional à massa da substância original presente no corpo naquele instante [23].

A constante de desintegração α é determinada experimentalmente e é relativa à cada substância. Essa constante é determinada, na prática, pela meia-vida da substância radioativa, que é o tempo necessário para que metade da massa de um corpo formado por aquela substância se desintegre [23]. Seja M_0 a massa de um corpo formado por uma substância radioativa cuja taxa de desintegração é igual a α. Como a desintegração se processa continuamente, é possível obter uma relação entre a massa restante da substância e o tempo decorrido. Sendo assim, fixado um inteiro $n > 0$ e supondo que a desintegração se dê em cada intervalo $\frac{1}{n}$ de segundo, após a primeira fração $\frac{1}{n}$ de segundo a massa do corpo se reduziria a

$$M_0 - \left(\frac{\alpha}{n}\right) M_0 = M_0 \left(1 - \frac{\alpha}{n}\right).$$

Ao se passar 1 segundo, teriam ocorridos n desintegrações instantâneas e, efetuadas as n reduções, a massa restante do corpo seria $M_0 \left(1 - \frac{\alpha}{n}\right)^n$. Dividindo o intervalo $[0, 1]$ em um número n cada vez maior de partes iguais, chega-se à conclusão de que, ao final de 1 segundo, a massa do corpo estará reduzida a

$$\lim_{n\to\infty} M_0 \left(1 - \frac{\alpha}{n}\right)^n$$

Como, por definição, $e = \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n$, segue que

$$\lim_{n\to\infty} M_0\left(1-\frac{\alpha}{n}\right)^n = M_0\cdot e^{-\alpha}.$$

Para obter a massa desse corpo ao final de t segundos, divide-se o intervalo $[0,t]$ em n partes iguais. Logo, a perda de massa em cada intervalo parcial será $M_0\cdot\left(\frac{\alpha t}{n}\right)$. Usando o argumento dado anteriormente, a massa restante do corpo após t segundo é expressa por

$$M(t) = M_0\cdot e^{-\alpha t} \tag{2.10}$$

Portanto, conclui-se que a massa restante de uma substância radioativa que se desintegra em função do tempo é dada por uma **função exponencial**.

Para uma melhor compreensão, pode-se observar na Figura 2.3 o gráfico da função $M:\mathbb{R}^+\to\mathbb{R}^+$ dada pela Equação 2.10.

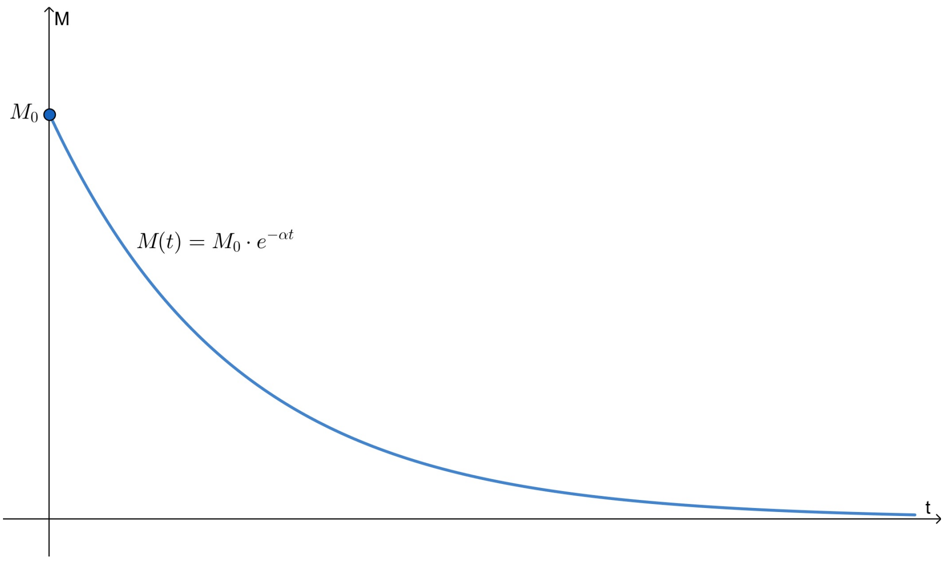

Figura 2.3: Gráfico da desintegração de uma substância radioativa.

No caso de ser conhecida a meia vida t_0 de uma substância, é possível calcular a taxa de desintegração α dessa substância. Desta forma,

$$\frac{M_0}{2} = M_0\cdot e^{-\alpha t_0} \Rightarrow \frac{1}{2} = e^{-\alpha t_0}.$$

Aplicando logaritmos, segue que

$$\frac{1}{2} = e^{-\alpha t_0} \Rightarrow \ln\left(\frac{1}{2}\right) = -\alpha t_0 \Rightarrow -\ln 2 = -\alpha t_0 \Rightarrow \alpha = \frac{\ln 2}{t_0}.$$

Reciprocamente, a meia vida t_0 é dada em função da taxa α por $t_0 = \dfrac{\ln 2}{\alpha}$.

2.3.4 Nível Sonoro

O som é formado por um conjunto de ondas mecânicas que se propagam com velocidade dependente das condições do meio de propagação. O ouvido humano é sensível a ondas sonoras com frequências entre 20 Hz e 20000 Hz, aproximadamente. Esse intervalo pode variar de acordo com a pessoa e a idade de cada um [25].

Como o som é uma propagação ondulatória, então ele é um processo de transporte de energia. A energia transportada por uma onda sonora pode ser representada pela intensidade que é definida como a quantidade de energia sonora que atravessa a unidade de área de uma superfície disposta perpendicularmente à direção de propagação, na unidade de tempo.

Para uma pessoa com audição normal, a sensação sonora aumenta à medida que a intensidade de um som aumenta. Além de depender da intensidade sonora, a sensação sonora depende ainda do ouvinte e da frequência do som. O limiar de sensação auditiva ou limiar de audibilidade é a intensidade mínima necessária por um som para que este seja ouvido. O som passa a ser mais percebido quando a intensidade sonora vai sendo aumentada a partir desse limiar, até que, a partir de certo valor da intensidade, à sensação sonora é acrescentada uma sensação de desconforto ou de dor. Esse valor recebe o nome de limiar de dor ou limiar de sensação dolorosa.

O aparelho auditivo humano, em uma audição normal, percebe sons com intensidades variando de 10^{-12} W/m^2 a 1 W/m^2, onde W representa a unidade de potência *watt* [25]. Como o ouvido humano é sensível a um grande intervalo de intensidade, adota-se geralmente uma escala logarítima para as intensidades. Assim, o nível da intensidade sonora β de uma onda sonora é definido como

$$\beta = (10\ db) \log \frac{I}{I_0} \tag{2.11}$$

onde I_0 é a intensidade de referência cujo valor é igual a 10^{-12} próximo ao limiar da audição humanada de 1000 Hz e dB corresponde a decibel que é a fração equivalente a $\frac{1}{10}$ do bel, unidade criada en homenagem a Alexandre Graham Bell. A unidade usual para indicar o nível de intensidade sonora é o decibel [26].

À partir da Equação 2.11 são obtidos os níveis da intensidade sonora de algumas fontes indicados na Tabela 2.2.

Tabela 2.2: Nível de intensidade sonora de diversas fontes (valores típicos)

Fonte ou descrição do som	I (W/m^2)	β (db)
Limiar de dor	1	120
Tráfego pesado	10^{-5}	70
Conversa comum	$3,2 \cdot 10^{-6}$	65
Rádio com volume baixo	10^{-8}	40
Limiar da audição a 1000 Hz	10^{-12}	0

Fonte: [26]

Conclui-se então, que o nível de intensidade sonora de um som é uma aplicação de **logaritmos**.

Capítulo 3

Conjuntos e Funções Convexas

Esse capítulo apresenta conceitos e resultados sobre análise convexa. A revisão aqui apresentada foi baseada em [27], [28], [29] e [33].

3.1 Conjuntos Convexos

Definição 34 *Um conjunto $C \subseteq \mathbb{R}^n$ é dito **convexo** se, para quaisquer $x_1, x_2 \in C$ e $\alpha \in [0, 1]$ tem-se $\alpha x_1 + (1 - \alpha)x_2 \in C$.*

Geometricamente, um conjunto é convexo quando o segmento de reta de extremos $x_1, x_2 \in C$ está inteiramente contido em C. Por definição, o conjunto vazio é um conjunto convexo. Na Figura 3.1, (a) representa um conjunto convexo e (b) um conjunto não convexo.

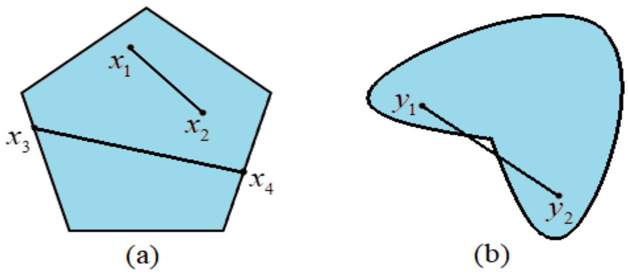

Figura 3.1: (a) conjunto convexo e (b) conjunto não convexo.

Definição 35 *Um conjunto $M \subseteq \mathbb{R}^n$ é um **conjunto afim** se para quaisquer x_1, $x_2 \in M$ e $\alpha \in \mathbb{R}$ tem-se $\alpha x_1 + (1 - \alpha)x_2 \in M$.*

Um conjunto $M \subseteq \mathbb{R}^n$ é afim se a reta que passa por quaisquer dois pontos de M está em M. O conjunto vazio é um exemplo de conjunto afim.

Definição 36 *Dados $x_1, x_2, \ldots, x_k$ vetores do $\mathbb{R}^n$. Um vetor da forma $x = \alpha_1 x_1 + \ldots + \alpha_k x_k$ é denominado* ***combinação linear*** *dos vetores $x_1, x_2, \ldots, x_k$, onde $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$. Caso tenha-se:*

(a) $\sum_{i=1}^{k} \alpha_i = 1$, então o vetor x é chamado ***combinação afim*** *dos vetores x_i;*

(b) $\alpha_i \geq 0$ para todo $1 \leq i \leq k$, então o vetor x é dito ***combinação positiva****, ou combinação cônica, dos vetores x_i.*

Um conjunto afim contém todas as combinações afins de seus pontos.

Definição 37 *Dados $x_1, x_2, \ldots, x_m$ vetores de um conjunto qualquer $C \subseteq \mathbb{R}^n$. As* ***combinações convexas*** *dos vetores $x_1, x_2, \ldots, x_m \in C$ são os vetores da forma $x = \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_m x_m$ tais que $\alpha_i \in \mathbb{R}$, com $\alpha_i \geq 0$, e $\sum_{i=1}^{m} \alpha_i = 1$.*

De acordo com a definição 37, um vetor é uma combinação convexa de vetores de um conjunto $C \subseteq \mathbb{R}^n$ quando for simultaneamente uma combinação afim e positiva.

Teorema 9 *Um conjunto $C \subset \mathbb{R}^n$ é convexo se, e somente se, C contém todas as combinações convexas de seus elementos.*

Lema 2 *O conjunto solução de um sistema de equações lineares $S = \{x \in \mathbb{R}^n \ : \ Ax = b\}$ sendo, $A \in \mathbb{R}^{m \times n}$ e $b \in \mathbb{R}^m$ é um conjunto convexo.*

Demonstração: Suponha $x_1, x_2 \in S$, assim $Ax_1 = b$ e $Ax_2 = b$. Então, para $\alpha \in [0,1]$ tem-se

$$A[\alpha x_1 + (1-\alpha)x_2] = \alpha A x_1 + (1-\alpha)Ax_2 = \alpha b + (1-\alpha)b = b$$

o que mostra que $\alpha x_1 + (1-\alpha)x_2 \in C$ e, portanto, S é convexo. ■

Definição 38 *Um ponto $x \in C$ é um* ***ponto extremo*** *de C se, e somente se, não pode ser escrito como uma combinação convexa $(1-\alpha)x_1 + \alpha x_2$, onde $x_1, x_2 \in C$ e $\alpha \in (0,1)$, exceto quando toma-se $x_1 = x_2 = x$.*

Definição 39 *Um conjunto $C \subseteq \mathbb{R}^n$ é um* ***cone*** *(com vértice na origem) se $x \in C$ implica que $\alpha x \in C$, para todo $\alpha \geq 0$, onde $\alpha \in \mathbb{R}$.*

Definição 40 *Um conjunto C é um* ***cone convexo*** *se para todo $x_1, x_2 \in C$ e $\alpha_1, \alpha_2 \in \mathbb{R}$ tais que $\alpha_1, \alpha_2 \geq 0$ tem-se $\alpha_1 x_1 + \alpha_2 x_2 \in C$.*

Definição 41 *Seja $S \subset \mathbb{R}^n$ um conjunto qualquer. A* ***casca convexa*** *do conjunto S, denotada por $conv(S)$, é o menor conjunto convexo que contém S.*

A casca convexa de S pode ser definida como a interseção de todos os conjuntos convexos que contêm S. O conceito de casca convexa dado na definição 41 pode ser compreendido geometricamente através da representação no $\mathbb{R}^2$ dada na figura 3.2.

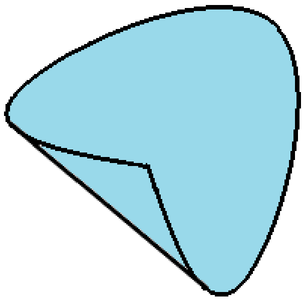

Figura 3.2: Representação da casca convexa da figura 3.1(b)

Teorema 10 *A casca convexa de S é formada por todas as combinações convexas dos vetores de S, qualquer seja $S \subset \mathbb{R}^n$.*

Demonstração: Pela definição 41 os elementos de S pertencem a $conv(S)$ portanto, pelo Teorema 9, todas as combinações convexas de S pertencem a $conv(S)$. Por outro lado, dados os vetores $x = \alpha_1 x_1 + \cdots + \alpha_m x_m$ e $y = \theta_1 y_1 + \cdots + \theta_k y_k$, com $x_1, \ldots, x_m \in S$ e $y_1, \ldots, y_k \in S$, o vetor

$$(1-\alpha)x + \alpha y = (1-\alpha)\alpha_1 x_1 + \cdots + (1-\alpha)\alpha_m x_m + \alpha\theta_1 y_1 + \cdots + \theta_k y_k$$

onde $\alpha \in [0, 1]$ também é uma combinação dos vetores de S. Desta forma, o conjunto de todas as combinações convexas de S é um conjunto convexo e contém o conjunto S, logo deve coincidir com o menor conjunto convexo, $conv(S)$. ∎

Proposição 15 *Sendo $C, D \subset \mathbb{R}^n$ conjuntos convexos. São convexos também os seguintes conjuntos:*

a) $C + D := \{z \ : \ z = c + d, \ c \in C, \ d \in D\}$;

b) $C \cap D$;

c) $\lambda C := \{z \ : \ z = \lambda c, \ c \in C\}$, para todo $\lambda \in \mathbb{R}$.

Demonstração: a) Dados $c, d \in C + D$, existem vetores $x_1, y_1 \in C$ e $x_2, y_2 \in D$ tais que $c = x_1 + x_2$ e $d = y_1 + y_2$. Para $\alpha \in [0, 1]$, tem-se

$$
\begin{aligned}
z &= (1-\alpha)c + \alpha d \\
&= (1-\alpha)(x_1 + x_2) + \alpha(y_1 + y_2) \\
&= [(1-\alpha)x_1 + \alpha y_1] + [(1-\alpha)x_2 + y_2].
\end{aligned}
$$

Como C e D são convexos, $(1-\alpha)x_1+\alpha y_1 \in C$ e $(1-\alpha)x_2+y_2 \in D$, logo $(1-\alpha)c+\alpha d \in C+D$.

b) Dados $c, d \in C \cap D$, pela definição de interseção de conjuntos tem-se que $c, d \in C$ e $c, d \in D$. Assim, como C e D são convexos, para $\alpha \in [0, 1]$, segue que $z = (1-\alpha)c + \alpha d \in C$ e $z = (1-\alpha)c + \alpha d \in D$. Portanto, $z = (1-\alpha)c + \alpha d \in C \cap D$, o que mostra que $C \cap D$ é convexo.

c) Dados $x, y \in \lambda C$ tais que $x = \lambda c_1$ e $y = \lambda c_2$, onde $c_1, c_2 \in C$, para $\alpha \in [0, 1]$ pela convexidade de C, tem-se

$$z = (1-\alpha)x + \alpha y = (1-\alpha)(\lambda c_1) + \alpha(\lambda c_2) = \lambda[(1-\alpha)c_1 + \alpha c_2] = \lambda c,$$

com $c \in C$. Logo, λC é convexo. ∎

Teorema 11 *A interseção de uma coleção de conjuntos convexos é um conjunto convexo.*

Definição 42 *Sejam $c \in \mathbb{R}^n$, $c \neq 0$ e $\beta \in \mathbb{R}$. Chama-se um **hiperplano** do $\mathbb{R}^n$ o conjunto*

$$H := \{x \ : \ c^T x = \beta\}.$$

*Cada hiperplano do $\mathbb{R}^n$ caracteriza dois **semiespaços fechados***

$$H_{\geq} := \{x \ : \ c^T x \geq \beta\}, \ H_{\leq} := \{x \ : \ c^T x \leq \beta\}$$

*e dois **semiespaços abertos***

$$H_{>} := \{x \ : \ c^T x > \beta\}, \ H_{<} := \{x \ : \ c^T x < \beta\}$$

Lema 3 *Todo hiperplano H do $\mathbb{R}^n$ é um conjunto convexo.*

Definição 43 *Um conjunto $C \subset \mathbb{R}^n$ é denominado **poliedro** quando é definido por um número finito de semiespaços fechados em $\mathbb{R}^n$. Um poliedro limitado e não vazio é chamado de **politopo**.*

Definição 44 *Para qualquer* $a \in \mathbb{R}^n$ *o conjunto dado por* $B(a, \epsilon) = \{x \in \mathbb{R}^n \ : \| x - a \| < \epsilon\}$ *é uma* ***bola*** *de raio* $\epsilon > 0$ *e centro* a.

Definição 45 *Um* ***ponto interior*** *de um conjunto* $X \subset \mathbb{R}^n$ *é um elemento* $x \in X$ *se para algum* $\epsilon > 0$ *tem-se* $B(a, \epsilon) \subset X$.
O conjunto formado pelos pontos interiores de X *é chamado de interior de* X *e é denotado por* $int(X)$.

Quando $\text{int}(X) = X$ o conjunto X é definido como um conjunto aberto. Por outro lado, quando o complementar do conjunto $X \subset \mathbb{R}^n$, dado por $\mathbb{R}^n - X$, for um conjunto aberto, então o conjunto X é dito fechado.

Definição 46 *A* ***fronteira*** *de um conjunto* $X \subset \mathbb{R}^n$*, dado por* $bd(X)$*, é o conjunto formado pelos pontos de* X *que não são interiores a* X *e pelos pontos de* $\mathbb{R}^n - X$ *que não são interiores a* $\mathbb{R}^n - X$.

Definição 47 *Um hiperplano* H *é dito um* ***hiperplano separador*** *se dados conjuntos convexos não vazios* $C_1, C_2 \subset \mathbb{R}^n$ *existe* H *tal que* $C_1 \subset H_{\geq}$ *e* $C_2 \subset H_{\leq}$.

Definição 48 *Sejam* $C \subseteq \mathbb{R}^n$ *um conjunto convexo e* $x_0 \in bd(C)$*. Se* $c \neq 0$ *satisfaz* $c^T x \leq c^T x_0$ *para todo* $x \in C$*, então o hiperplano* $\{x \ : \ c^T x = c^T x_0\}$ *é um* ***hiperplano suporte*** *de* C *em* x_0.

Geometricamente, o hiperplano suporte $\{x \ : \ c^T x = c^T x_0\}$ de um conjunto convexo C é tangente a C no ponto x_0 e um dos semiespaços, $\{x \ : \ c^T x \leq c^T x_0\}$ ou $\{x \ : \ c^T x \geq c^T x_0\}$, contém C. Essa interpretação pode ser melhor compreendida através da Figura 3.3.

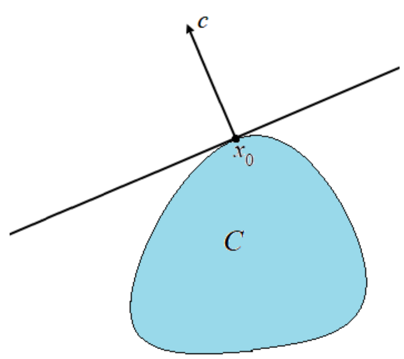

Figura 3.3: Hiperplano suporte de C.

Dizer que $\{x \ : \ c^T x = c^T x_0\}$ é um hiperplano suporte do conjunto C em x_0 é equivalente a dizer que esse é um hiperplano que separa o ponto x_0 e o conjunto C.

Quando para $C_1, C_2 \subseteq \mathbb{R}$ tem-se $c^T x > \beta$ para todo $x \in C_1$ e $c^T x < \beta \in C_2$ diz-se que a separação entre os conjuntos C_1 e C_2 é estrita.

3.2 Funções convexas

Definição 49 *Seja $C \subseteq \mathbb{R}^n$ um conjunto convexo não vazio. Uma função $f : C \to \mathbb{R}$ é **convexa**, se para todo $x_1, x_2 \in C$ e $\alpha \in [0,1]$, tem-se*

$$f(\alpha x_1 + (1-\alpha)x_2) \leq \alpha f(x_1) + (1-\alpha) f(x_2) \tag{3.1}$$

Se a desigualdade 3.1 for estrita sempre que $x_1 \neq x_2$ e $\alpha \in (0,1)$, a função f é denominada estritamente convexa. Diz-se que f é côncava se $-f$ é convexa, e estritamente côncava se $-f$ é estritamente convexa.

Geometricamente, a desigualdade 3.1 significa, que o segmento de reta de extremos $(x_1, f(x_1))$ e $(x_2, f(x_2))$ está acima do gráfico da função f. Veja Figura 3.4.

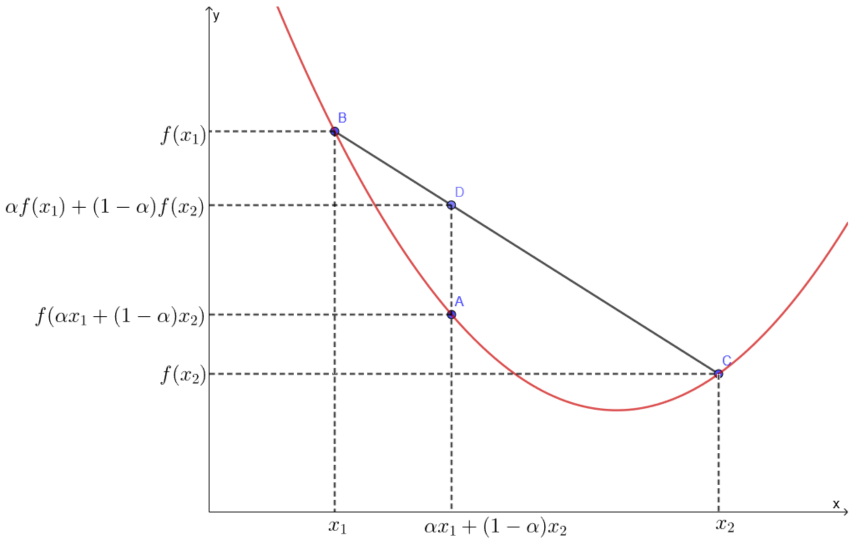

Figura 3.4: Gráfico de uma função convexa representada no espaço $\mathbb{R}^2$

A função f é dita **fortemente convexa** se para todo $x_1, x_2 \in C$ e $\alpha \in [0,1]$ existe um $c > 0$ tal que

$$f(\alpha x_1 + (1-\alpha)x_2) \leq \alpha f(x_1) + (1-\alpha) f(x_2) - \tfrac{1}{2} c\alpha(1-\alpha) \parallel x_1 - x_2 \parallel^2$$

Definição 50 *O **epígrafo** de uma função $f : C \to \mathbb{R}$, denotado por epif, é definido como o conjunto* $\operatorname{epi} f := \{(x, \lambda) \ : \ f(x) \leq \lambda, \ x \in \mathcal{C}, \ \lambda \in \mathbb{R}\}$.

Teorema 12 *Uma função $f : C \to \mathbb{R}$ é convexa se, e somente se, $\operatorname{epi} f$ é um conjunto convexo em $\mathbb{R}^n \times \mathbb{R}$.*

Demonstração: Suponha que f é uma função convexa. Sendo $(x_1, y_1), (x_2, y_2) \in \operatorname{epi} f$ e $\alpha \in [0, 1]$ tem-se que $f(x_1) \leq y_1$ e $f(x_2) \leq y_2$. Então, pela convexidade de f, segue que

$$f(\alpha x_1 + (1 - \alpha)x_2) \leq \alpha f(x_1) + (1 - \alpha)f(x_2) \leq \alpha y_1 + (1 - \alpha)y_2.$$

Com isso,

$$\begin{gathered} \alpha(x_1, y_1) + (1 - \alpha)(x_2, y_2) = (\alpha x_1, \alpha y_1) + ((1 - \alpha)x_2, (1 - \alpha)y_2) = \\ (\alpha x_1 + (1 - \alpha)x_2, \alpha y_1 + (1 - \alpha)y_2) \in \operatorname{epi} f. \end{gathered}$$

Por outro lado, suponha que $\operatorname{epi} f$ é um conjunto convexo. Dados $x_1, x_2 \in C$ e $\alpha \in [0, 1]$, $(x_1, f(x_1)) \in \operatorname{epi} f$ e $(x_2, f(x_2)) \in \operatorname{epi} f$. Pela convexidade de $\operatorname{epi} f$,

$$(\alpha x_1 + (1 - \alpha)x_2, \alpha f(x_1) + (1 - \alpha)f(x_2) = \alpha(x_1, f(x_1)) + (1 - \alpha)(x_2, f(x_2)) \in \operatorname{epi} f.$$

Desta forma, pela definição de epígrafo, conclui-se que

$$f(\alpha x_1 + (1 - \alpha)x_2) \leq \alpha f(x_1) + (1 - \alpha)f(x_2).$$

Portanto, f é convexa. ∎

Teorema 13 *Uma condição necessária e suficiente para que a função $f : C \to \mathbb{R}$ seja convexa é que para quaisquer $a, b \in C$ a função $g : [0, 1] \to \mathbb{R}$ definida por $g(t) = f(a + tv)$, $v = b - a$, seja convexa.*

Demonstração: Se f é convexa então, para $s, t, \alpha \in [0, 1]$ tem-se

$$\begin{aligned} g((1 - \alpha)s + \alpha t) &= f(a + [(1 - \alpha)s + \alpha t]v) \\ &= f[(1 - \alpha)(a + sv) + \alpha(a + tv)] \\ &\leq (1 - \alpha)f(a + sv) + \alpha f(a + tv) \\ &= (1 - \alpha)g(s) + \alpha g(t) \end{aligned}$$

portanto g é convexa. Reciprocamente se todas as funções g, definidas anteriormente, são convexas, então dados $x_1, x_2 \in C$ e $\alpha \in [0,1]$, pondo $g(t) = f(x_1 + t(x_2 - x_1))$ segue-se que:

$$\begin{aligned} f((1-\alpha)x_1 + \alpha x_2) &= f(x_1 + \alpha(x_2 - x_1)) = g((1-\alpha)\cdot 0 + \alpha \cdot 1) \\ &\leq (1-\alpha)g(0) + \alpha g(1) = (1-\alpha)f(x_1) + \alpha f(x_2), \end{aligned}$$

logo, f é convexa. ∎

Algumas operações preservam a convexidade das funções.

Proposição 16 *Sejam f, g funções convexas definidas sobre conjuntos convexos $C \subseteq \mathbb{R}^n$ e $D \subseteq \mathbb{R}^n$, respectivamente e $\theta \in \mathbb{R}$. Então*

a) θf é convexa sobre C, qualquer seja $\theta \geq 0$;

b)$f + g$ é convexa sobre o conjunto $C \cap D$.

Demonstração: a) Sejam $\theta \geq 0$ e f uma função convexa. Então, para todos $x_1, x_2 \in C$ e $\alpha \in [0,1]$ tem-se

$$(\theta f)(\alpha x_1 + (1-\alpha)x_2) = \theta f(\alpha x_1 + (1-\alpha)x_2) \leq \theta(\alpha f(x_1) + (1-\alpha)f(x_2)) = \alpha(\theta f(x_1)) + (1-\alpha)(\theta f(x_2)).$$

Portanto, θf é uma função convexa.

b) Sejam f, g funções convexas. Logo, para todos $x_1, x_2 \in C$ e $\alpha \in [0,1]$,

$$(f+g)(\alpha x_1 + (1-\alpha)x_2) = f(\alpha x_1 + (1-\alpha)x_2) + g(\alpha x_1 + (1-\alpha)x_2)$$

pela convexidade de f e g segue que

$$\begin{aligned} f(\alpha x_1 + (1-\alpha)x_2) + g(\alpha x_1 + (1-\alpha)x_2) &\leq \alpha f(x_1) + (1-\alpha)f(x_2) + \alpha g(x_1) + (1-\alpha)g(x_2) \\ &= \alpha(f(x_1) + g(x_1)) + (1-\alpha)(f(x_2) + g(x_2)) \\ &= \alpha(f+g)(x_1) + (1-\alpha)(f+g)(x_2) \end{aligned}$$

com isso, $(f+g)(\alpha x_1 + (1-\alpha)x_2) \leq (f+g)(x_1) + (1-\alpha)(f+g)(x_2)$ e, portanto, $f+g$ é uma função convexa. ∎

Teorema 14 *Seja $f : C \to \mathbb{R}$ uma função convexa com $C \subseteq \mathbb{R}^n$. Então para todos $x_1, \ldots, x_n \in C$ e $\alpha_1, \ldots, \alpha_n \geq 0$ onde $\alpha_1 + \ldots + \alpha_n = 1$, tem-se*

$$f(\alpha_1 x_1 + \cdots + \alpha_n x_n) \leq \alpha_1 f(x_1) + \cdots + \alpha_n f(x_n) \tag{3.2}$$

Demonstração: Suponha que f é uma função convexa. A prova será feita por indução sobre $n > 1$. Para $n = 2$, a desigualdade 3.2 é válida, uma vez que para $\alpha_1 + \alpha_2 = 1$, segue-se que $\alpha_1 = \alpha$ e $\alpha_2 = 1 - \alpha$. Suponha, por hipótese de indução, que para algum $n > 1$ e todos $x_1, \ldots, x_n \in C$ e $\alpha_1, \ldots, \alpha_n \geq 0$ onde $\alpha_1 + \ldots + \alpha_n = 1$, tem-se

$$f(\alpha_1 x_1 + \cdots + \alpha_n x_n) \leq \alpha_1 f(x_1) + \cdots + \alpha_n f(x_n).$$

Considere $x_1, \ldots, x_n, x_{n+1} \in C$ e $\alpha_1, \ldots, \alpha_n, \alpha_{n+1} \geq 0$ tais que $\alpha_1 + \cdots + \alpha_n + \alpha_{n+1} = 1$. Pelo menos um dos $\alpha_1, \ldots, \alpha_n, \alpha_{n+1}$ deve ser menor que 1, do contrário a desigualdade seria trivial. Suponha, sem perda de generalidade, $\alpha_{n+1} < 1$. Definindo

$$y = \frac{\alpha_1 x_1 + \cdots + \alpha_n x_n}{1 - \alpha_{n+1}} = s_1 x_1 + \cdots + s_n x_n,$$

com $s_j = \dfrac{t_j}{1 - t_{n+1}}$, para $1 \leq j \leq n$. Então, $s_1, \ldots, s_n > 0$ e

$$s_1 + \cdots + s_n = \frac{\alpha_1}{1 - \alpha_{n+1}} + \cdots + \frac{\alpha_n}{1 - \alpha_{n+1}} = \frac{1}{1 - \alpha_{n+1}}(\alpha_1 + \cdots + \alpha_n).$$

De $\alpha_1 + \cdots + \alpha_n + \alpha_{n+1} = 1$ segue que $\alpha_1 + \cdots + \alpha_n = 1 - \alpha_{n+1}$. Logo,

$$s_1 + \cdots + s_n = \frac{1}{1 - \alpha_{n+1}}(\alpha_1 + \cdots + \alpha_n) = \frac{1}{1 - \alpha_{n+1}}(1 - \alpha_{n+1}) = 1.$$

Daí, por hipótese de indução, $y \in C$ e

$$\alpha_1 x_1 + \cdots + \alpha_n x_n + \alpha_{n+1} x_{n+1} = (1 - \alpha_{n+1}) y + \alpha_{n+1} x_{n+1} \in C.$$

Pela convexidade de f, obtém-se

$$f(\alpha_1 x_1 + \cdots + \alpha_{n+1} x_{n+1}) = f((1 - \alpha_{n+1}) y + \alpha_{n+1} x_{n+1}) \leq (1 - \alpha_{n+1}) f(y) + \alpha_{n+1} f(x_{n+1}). \tag{3.3}$$

Por outro lado, também por hipótese de indução, tem-se que

$$f(y) = f(s_1 x_1 + \cdots + s_n x_n) \leq s_1 f(x_1) + \cdots + s_n f(x_n) = \frac{1}{1 - \alpha_{n+1}}(\alpha_1 f(x_1) + \cdots + \alpha_n f(x_n)). \tag{3.4}$$

Assim, juntando as desigualdades 3.3 e 3.4, conclui-se que

$$f(\alpha_1 x_1 + \cdots + \alpha_{n+1} x_{n+1}) \leq (1 - \alpha_{n+1}) f(y) + \alpha_{n+1} f(x_{n+1}) \leq$$
$$(\alpha_1 f(x_1) + \cdots \alpha_n f(x_n)) + \alpha_{n+1} f(x_{n+1}).$$

Desse modo, a desigualdade 3.2 é válida para $n + 1$. Portanto, pelo Princípio de Indução Matemática, a desigualdade 3.2 é válida para todo natural $n > 1$.

Corolário 2 *Seja $f : C \to \mathbb{R}$ uma função convexa com $C \subseteq \mathbb{R}^n$. Então, para $x_1, \ldots, x_k \in C$, tem-se*

$$f\left(\frac{x_1 + \cdots + x_n}{n}\right) \leq \frac{f(x_1) + \cdots + f(x_n)}{n}.$$

Demonstração: O resultado é obtido fazendo $\alpha_i = \dfrac{1}{n}$, com $i = 1, \ldots, n$, no teorema 14.

Definição 51 *Seja $f : C \to \mathbb{R}$ uma função. O conjunto definido por $C_\lambda = \{x \in C \ : \ f(x) \leq \lambda\}$ é denominado **conjunto de nível** da função f.*

Proposição 17 *Seja $f : C \to \mathbb{R}$ uma função convexa. Então, o conjunto de nível da função f, $C_\lambda = \{x \in C \ : \ f(x) \leq \lambda\}$, é convexo, qualquer que seja $\lambda \in \mathbb{R}$.*

Definição 52 *Uma função $f : C \to \mathbb{R}$, onde $C \subseteq \mathbb{R}^n$, é dita **quase-convexa** se C for um conjunto convexo e se $C_\lambda = \{x \in C \ : \ f(x) \leq \lambda\}$ for convexo, para qualquer $\lambda \in \mathbb{R}$.*

Uma função f definida sobre um subconjunto convexo $C \subseteq \mathbb{R}^n$ é quase-côncava se $(-f)$ é quase-convexa, ou seja, se $\{x \in C \ : \ f(x) \geq \lambda\}$. Uma função f que é simultaneamente quase-convexa e quase-côncava, ou seja, definida sobre um conjunto convexo e $\{x \ \in C : \ f(x) = \lambda\}$ é convexo para qualquer $\lambda \in \mathbb{R}$, é chamada de quase-linear.

Proposição 18 *Uma função $f : C \to \mathbb{R}$, onde $C \subseteq \mathbb{R}^n$, é quase-convexa se, e somente se, C é um conjunto convexo e para todo $x_1, x_2 \in C$ e para todo $\alpha \in [0, 1]$ tem-se*

$$f(\alpha x_1 + (1 - \alpha) x_2) \leq max \ \{f(x_1), f(x_2)\}$$

3.3 Funções convexas em $\mathbb{R}$

Essa seção, tem como objetivo provar, através de problemas, a convexidade ou a concavidade de algumas funções reais definidas no capítulo 2.

$\boxed{\textbf{Problema 1}}$

A função **afim** $f : \mathbb{R} \to \mathbb{R}$, $f(x) = ax + b$, $a, b \in \mathbb{R}$ e $a \neq 0$, é uma função convexa.

Demonstração: Sejam x_1, x_2 números reais quaisquer. Para todo $\alpha \in \mathbb{R}$, onde $\alpha \in [0, 1]$, tem-se:

$$\begin{aligned} f(\alpha x_1 + (1-\alpha)x_2) &= a(\alpha x_1 + (1-\alpha)x_2) + b \\ &= a\alpha x_1 + ax_2 - a\alpha x_2 + b \\ &= a\alpha x_1 + \alpha b + ax_2 - a\alpha x_2 + b - \alpha b \\ &= \alpha(ax_1 + b) + ax_2(1-\alpha) + b(1-\alpha) \\ &= \alpha(ax_1 + b) + (1-\alpha)(ax_2 + b) \\ &= \alpha f(x_1) + (1-\alpha)f(x_2). \end{aligned}$$

Portanto, f é uma função convexa. ∎

Como para as funções afins vale sempre a igualdade em 3.1 então, essas funções são simultamente convexas e côncavas. Reciprocamente, toda função convexa e côncava é afim.

$\boxed{\textbf{Problema 2}}$

A função **quadrática** $f : \mathbb{R} \to \mathbb{R}$, $f(x) = ax^2 + bx + c$, onde $b, c \in \mathbb{R}$ e $a \in \mathbb{R}^+$, é estritamente convexa.

Demonstração: Dados $x_1, x_2 \in \mathbb{R}$ e $\alpha \in (0, 1)$, tem-se que

$$\begin{aligned} (\alpha x_1 + (1-\alpha)x_2)^2 < \alpha x_1^2 + (1-\alpha)x_2^2 &\Leftrightarrow (\alpha x_1 + (1-\alpha)x_2)^2 - [\alpha x_1^2 + (1-\alpha)x_2^2] < 0 \\ &\Leftrightarrow \alpha^2 x_1^2 + 2\alpha(1-\alpha)x_1x_2 + (1-\alpha)^2x_2^2 - \alpha x_1^2 - (1-\alpha)x_2^2 < 0 \\ &\Leftrightarrow -\alpha(1-\alpha)x_1^2 + 2\alpha(1-\alpha)x_1x_2 - \alpha(1-\alpha)x_2^2 < 0 \\ &\Leftrightarrow \alpha(1-\alpha)(x_1^2 - 2x_1x_2 + x_2^2) > 0 \\ &\Leftrightarrow \alpha(1-\alpha)(x_1 - x_2)^2 > 0. \end{aligned}$$

Desta forma, segue que

$$\begin{aligned} f(\alpha x_1 + (1-\alpha)x_2) &= a(\alpha x_1 + (1-\alpha)x_2)^2 + b(\alpha x_1 + (1-\alpha)x_2) + c \\ &< a(\alpha x_1^2 + (1-\alpha)x_2^2) + b(\alpha x_1 + (1-\alpha)x_2) + c \\ &= \alpha a x_1^2 + \alpha b x_1 + \alpha c + (1-\alpha)a x_2^2 + (1-\alpha)b x_2 + (1-\alpha)c \\ &= \alpha f(x_1) + (1-\alpha)f(x_2). \end{aligned}$$

Portanto, f é uma função estritamente convexa. ∎

Curiosidade

O estudo de funções convexas teve início através da desigualdade 3.5 conhecida como **Desigualdade do Ponto Médio**. Se $f : C \to \mathbb{R}$ é uma função convexa em C, tal que $f(x) = ax^2 + bx + c$, $a, b, c \in \mathbb{R}$ e $a \neq 0$, então é válida a desigualdade

$$f\left(\frac{x_1 + x_2}{2}\right) < \frac{f(x_1) + f(x_2)}{2} \tag{3.5}$$

De fato, note inicialmente que, se $x_1 \neq x_2$, então

$$\frac{x_1^2 + x_2^2}{2} - \left(\frac{x_1 + x_2}{2}\right)^2 = \frac{x_1^2 - 2x_1x_2 + x_2^2}{4} = \left(\frac{x_1 - x_2}{2}\right)^2 > 0$$

logo $\left(\dfrac{x_1 + x_2}{2}\right)^2 < \dfrac{x_1^2 + x_2^2}{2}$.

Portanto, se $x_1 \neq x_2$ e $a > 0$, tem-se

$$\begin{aligned} f\left(\frac{x_1 + x_2}{2}\right) &= a\left(\frac{x_1 + x_2}{2}\right)^2 + b\left(\frac{x_1 + x_2}{2}\right) + c \\ &< a\frac{x_1^2 + x_2^2}{2} + b\left(\frac{x_1 + x_2}{2}\right) + c \\ &= \frac{(ax_1^2 + bx_1 + c) + (ax_2^2 + bx_2 + c)}{2} \\ &= \frac{f(x_1) + f(x_2)}{2} \end{aligned}$$

Problema 3

A função **exponencial de base** e, $f:\mathbb{R}\to\mathbb{R}$ tal que $f(x)=e^x$ é convexa.

Demonstração: Dados x_1, x_2 números reais quaisquer tais que $x_2 < x_1$. Para todo $\alpha \in \mathbb{R}$, com $\alpha \in [0,1]$, tem-se

$$\begin{aligned}
f(\alpha x_1 + (1-\alpha)x_2) \leq \alpha f(x_1) + (1-\alpha) f(x_2) \;&\Leftrightarrow\; e^{\alpha x_1 + (1-\alpha)x_2} \leq \alpha e^{x_1} + (1-\alpha)e^{x_2}\\
&\Leftrightarrow\; e^{\alpha x_1}\cdot e^{(1-\alpha)x_2} \leq \alpha e^{x_1} + (1-\alpha)e^{x_2}\\
&\Leftrightarrow\; (e^{x_1})^{\alpha}\cdot (e^{x_2})^{(1-\alpha)} \leq \alpha e^{x_1} + (1-\alpha)e^{x_2}\\
&\Leftrightarrow\; (e^{x_1})^{\alpha}\cdot \frac{e^{x_2}}{(e^{x_2})^{\alpha}} \leq \alpha e^{x_1} + (1-\alpha)e^{x_2}\\
&\Leftrightarrow\; e^{x_2}\cdot \left(\frac{e^{x_1}}{e^{x_2}}\right)^{\alpha} \leq \alpha e^{x_1} + (1-\alpha)e^{x_2}\\
&\Leftrightarrow\; \left(\frac{e^{x_1}}{e^{x_2}}\right)^{\alpha} \leq \alpha\left(\frac{e^{x_1}}{e^{x_2}}\right) + 1 - \alpha\\
&\Leftrightarrow\; \alpha - \alpha\left(\frac{e^{x_1}}{e^{x_2}}\right) \leq 1 - \left(\frac{e^{x_1}}{e^{x_2}}\right)^{\alpha}\\
&\Leftrightarrow\; \alpha\left(1 - \left(\frac{e^{x_1}}{e^{x_2}}\right)\right) \leq 1 - \left(\frac{e^{x_1}}{e^{x_2}}\right)^{\alpha}\\
&\Leftrightarrow\; \alpha \leq \frac{1 - \left(\frac{e^{x_1}}{e^{x_2}}\right)^{\alpha}}{1 - \left(\frac{e^{x_1}}{e^{x_2}}\right)}.
\end{aligned}$$

Porém,

$$\frac{1 - \left(\frac{e^{x_1}}{e^{x_2}}\right)^{\alpha}}{1 - \left(\frac{e^{x_1}}{e^{x_2}}\right)} = \frac{\left(\frac{e^{x_1}}{e^{x_2}}\right)^{\alpha} - 1}{\left(\frac{e^{x_1}}{e^{x_2}}\right) - 1} \leq \frac{\left(\frac{e^{x_1}}{e^{x_2}}\right) - 1}{\left(\frac{e^{x_1}}{e^{x_2}}\right) - 1} = 1$$

A última desigualdade é verdadeira pois $\frac{e^{x_1}}{e^{x_2}} > 1$ e $\alpha < 1$, o que implica que $\left(\frac{e^{x_1}}{e^{x_2}}\right)^{\alpha} < \frac{e^{x_1}}{e^{x_2}}$. Logo, $\alpha \leq 1$. Portanto,

$$f(\alpha x_1 + (1-\alpha)x_2) \leq \alpha f(x_1) + (1-\alpha) f(x_2) \Leftrightarrow \alpha \leq 1, \text{ o que é sempre verdadeiro.}$$

O que mostra que f é convexa. ∎

A Proposição enunciada a seguir é fundamental para a demonstração do Problema 4.

Proposição *Se $f: I \to \mathbb{R}$, sendo $I \subset \mathbb{R}$ um intervalo, é uma função contínua tal que para todo $x_1, x_2 \in I$ tem-se*

$$f\left(\frac{x_1 + x_2}{2}\right) \leq \frac{f(x_1) + f(x_2)}{2}$$

então, f é uma função convexa.

A concavidade da função f é garantida quando $-f$ é uma função convexa, ou seja, quando para $I \subset \mathbb{R}$ e $f : I \to \mathbb{R}$ contínua tem-se

$$f\left(\frac{x_1 + x_2}{2}\right) \geq \frac{f(x_1) + f(x_2)}{2}.$$

para todo $x_1, x_2 \in I$.

$\boxed{\textbf{Problema 4}}$

A função **logarítmica** $\log : \mathbb{R}^+ \to \mathbb{R}$ dada por $f(x) = \log x$ é uma função côncava.

Demonstração: Dados $x_1, x_2 \in \mathbb{R}^+$, tem-se:

$$\begin{aligned}
(\sqrt{x_1} - \sqrt{x_2})^2 \geq 0 \quad &\Leftrightarrow \quad x_1 - 2\sqrt{x_1}\sqrt{x_2} + x_2 \geq 0 \\
&\Leftrightarrow \quad x_1 + x_2 \geq 2 \cdot x_1^{\frac{1}{2}} \cdot x_2^{\frac{1}{2}} \\
&\Leftrightarrow \quad \frac{x_1 + x_2}{2} \geq x_1^{\frac{1}{2}} \cdot x_2^{\frac{1}{2}} \\
&\Leftrightarrow \quad \log\left(\frac{x_1 + x_2}{2}\right) \geq \log\left(x_1^{\frac{1}{2}} \cdot x_2^{\frac{1}{2}}\right) \\
&\Leftrightarrow \quad \log\left(\frac{x_1 + x_2}{2}\right) \geq \log\left(x_1^{\frac{1}{2}}\right) + \log\left(x_2^{\frac{1}{2}}\right) \\
&\Leftrightarrow \quad \log\left(\frac{x_1 + x_2}{2}\right) \geq \frac{1}{2}\log x_1 + \frac{1}{2}\log x_2 \\
&\Leftrightarrow \quad \log\left(\frac{x_1 + x_2}{2}\right) \geq \frac{\log x_1 + \log x_2}{2} \\
&\Leftrightarrow \quad f\left(\frac{x_1 + x_2}{2}\right) \geq \frac{f(x_1) + f(x_2)}{2}
\end{aligned}$$

Portanto, a função logarítmica é côncava. ■

Conclusão

Nessa dissertação foram apresentados de forma objetiva uma abordagem histórica do conceito de função, a análise dos principais documentos orientadores da educação brasileira, definições, proposições, propriedades e teoremas sobre Conjuntos e Funções Reais e Conjuntos Convexos e Funções Convexas, além de aplicações de funções reais em situações reais.

A fim de possibilitar um aprofundamento no conhecimento dos professores de Matemática, no capítulo 2 apresentou-se a fundamentação teórica de Conjuntos e Funções Reais, com ênfase nas funções afim, quadrática, exponencial e logarítmica, onde a abordagem dos conceitos foi dada de maneira mais formal, com propriedades e teoremas que comumente são desconhecidos dos profissionais da área. Nesse capítulo, a ligação existente entre a Matemática e outras áreas pôde ser evidenciada através das aplicações no imposto de renda, no movimento uniformemente variado, na desintegração radioativa e no nível de intensidade sonoro, mostrando a importância da interdisciplinariedade.

No capítulo 3, foram mostrados os principais conceitos de convexidade, com destaque para os Conjuntos Convexos e as Funções Convexas em $\mathbb{R}^n$, incluindo definições, proposições e teoremas. O intuito desse capítulo foi de proporcionar um material didático aos profissionais do ensino da Matemática para enriquecer seus conhecimentos, em especial nas funções convexas, uma vez que essa funções podem ser utilizadas em demonstrações de desigualdades e são amplamente utilizadas na otimização. Os problemas apresentados no final desse capítulo, relacionam funções reais com o conceito de convexidade.

Como trabalhos futuros pode-se explorar mais o uso da tecnologia no ensino de funções, através da utilização de *softwares* como o *GeoGebra*, para uma melhor compreensão dos alunos sobre os conceitos trabalhados em sala de aula e para a construção e o estudo do comportamento de gráficos de diferentes tipos de funções. O estudo de aplicações de funções reais em outras áreas deve ser considerado, a fim de proporcionar ao aluno o entendimento da Matemática em situações reais. Sugere-se também um aprofundamento no estudo das Funções Convexas.

Referências Bibliográficas

[1] *KLEINER, Israel.* Evolution of the Function Concept: A Brief Survey. The College Mathematics Journal, v.20, n.4, 1989, p. 282-300. 1989. Disponível em : <http://www.maa.org/pubs/Calc_articles/ma001.pdf> Acesso em: 12 jan. 2018.

[2] *BOYER, Carl B.* Tradução de Elza F. Gomide. História da Matemática. São Paulo: Edgard Blücher, Ed. da Universidade de São Paulo, 1974.

[3] *HOWARD, Eves.* tradução de Hygino H. Domingues. 5 ed. Campinas, São Paulo: Editora da Unicamp, 2011.

[4] *PONTE, João P.* O conceito de função no currículo de Matemática. Educação e Matemática, n. 15, p. 3-9, 1990. Disponível em <http://hdl.handle.net/10451/4473> Acesso em: 30 ago. 2018.

[5] *BRASIL.* Lei de Diretrizes e Bases da Educação Nacional, Lei n. 9394, de 20 de dezembro de 1996 (LDB). Disponível em <http://portal.mec.gov.br/seesp/arquivos/pdf/lei9394_ldbn1.pdf> Acesso em: 12 jan. 2018.

[6] *BRASIL.Secretaria de Educação Fundamental.* Parâmetros curriculares nacionais: Matemática / Secretaria de Educação Fundamental.- Brasilía: MEC/SEF, 1998. Disponível em <http://portal.mec.gov.br/seb/arquivos/pdf/matematica.pdf> Acesso em: 12 jan. 2018.

[7] *BRASIL.Secretaria de Educação Média e Tecnológica.* Parâmetros curriculares nacionais: Ciências da Natureza e Matemática e suas tecnologias.- Brasília: MEC, 2000. Disponível em: <http://portal.mec.gov.br/seb/arquivos/pdf/ciencian.pdf> Acesso em: 12 jan. 2018.

[8] *BRASIL.* Orientações Educacionais Complementares aos Parâmetros Curriculares Nacionais (PCN+). Ciências da Natureza e Matemática e suas tecnologias. Brasília: MEC,

2006. Disponível em: <http://portal.mec.gov.br/seb/arquivos/pdf/CienciasNatureza.pdf> Acesso em: 12 jan. 2018.

[9] *BRASIL. Secretaria de Educação Básica. Ministério da Educação.* Orientações Curriculares para o Ensino Médio, vol. 2: Ciências da Natureza, Matemática e suas tecnologias. Brasília: MEC, 2006. Disponível em: <http://portal.mec.gov.br/seb/arquivos/pdf/book_volume_02_internet.pdf> Acesso em: 16 jan. 2018.

[10] *BRASIL.* Base Nacional Comum Curricular (BNCC). Educação é a Base. Brasília, MEC/CONSED/UNDIME, 2017. Disponível em: <http://basenacionalcomum.mec.gov.br/wp-content/uploads/2018/02/bncc-20dez-site.pdf> Acesso em: 4 mar. 2018.

[11] *GeoGebra.* Disponível em: <https://www.geogebra.org/about> Acesso em: 08 set. 2018.

[12] *LIMA, Elon L.* Números e Funções Reais, Coleção PROFMAT, 1 ed., Rio de Janeiro: SBM, 2014.

[13] *LIMA, Elon L.; CARVALHO, Paulo C. P.; WAGNER, Eduardo; MORGADO, Augusto C.* A Matemática do Ensino Médio, vol. 1, Coleção do Professor de Matemática, 10 ed. Rio de Janeiro: SBM, 2012.

[14] *MUNIZ NETO, Antonio C.* Fundamentos de Cálculo, Coleção PROFMAT. 1 ed. Rio de Janeiro: SBM, 2015.

[15] *STEWART, James.* Tradução de Antonio Carlos Moretti, Antonio Carlos Gilli Martins. Cálculo, vol. 1. 6 ed. americana. São Paulo: Cengage Learning, 2011.

[16] *GUIDORIZZI, Hamilton L.* Um Curso de Cálculo, vol. 1, 5 ed. Rio de Janeiro: LTC, 2011.

[17] *LIMA, Elon L.* Um Curso de Análise, vol. 1. 1 ed., Projeto Euclides. Rio de Janeiro: IMPA, 2014.

[18] *IEZZI, Gelson; MURAKAMI; Carlos.* Fundamentos de Matemática Elementar- Conjuntos, Funções, vol. 1. 3 ed. São Paulo: Atual, 1977.

[19] *COLLI, Eduardo* Imposto progressivo. Revista do Professor de Matemática, Rio de Janeiro, n. 54, p. 10-15, 2004.

[20] *RECEITA FEDERAL DO BRASIL.* IRPF (Imposto sobre a renda das pessoas físicas). Disponível em: <http://idg.receita.fazenda.gov.br/acesso-rapido/tributos/irpf-imposto-de-renda-pessoa-fisica#calculo_mensal_IRPF> Acesso em: 10 set. 2018

[21] *DOCA, Ricardo H.; BISCUOLA, Gualter J.; BÔAS, Newton V.* Física, vol. 1 : mecânica. 3 ed. São Paulo: Saraiva, 2016.

[22] *HALLIDAY, David; RESNICK, Robert; WALKER, Jearl.* Fundamentos de Física: mecânica. vol. 1, 8 ed. Rio de Janeiro: LTC, 2009.

[23] *LIMA, Elon L.* Logaritmos. 2 ed. Coleção do Professor de Matemática. Rio de Janeiro: SBM, 1996.

[24] *RUSSELL, John B.* Química Geral, vol. 2, 2 ed. São Paulo: Pearson, 2000.

[25] *DOCA, Ricardo H.; BISCUOLA, Gualter J.; BÔAS, Newton V.* Física, vol. 2 : termologia, ondulatória, óptica. 3 ed. São Paulo: Saraiva, 2016.

[26] *YOUNG, Hugh D.; FREEDMAN, Roger A.* tradução de Claúdia Santana Martins. Física II: Termodinâmica e Ondas. 12 ed. São Paulo: Addison Wesley, 2008.

[27] *BOYD, Stephen; VANDENBERGHE, Lieven.* Convex Otimization. New York, USA: Cambridge University Press, 2004.

[28] *ROCKAFELLAR, Ralph T.* Convex Analysis. New Jersey, USA: Princeton University, 1970.

[29] *HIRIART-URRUTY, Jean-Batiste; LEMARÉCHAL, Claude* Fundamentals of Convex Analysis. 2 ed. Berlin: Springer-Verlag, 2004.

[30] *OLIVEIRA, Rubia M.* Algoritmos de Busca Global para Problemas de Otimização Geométricos e Multiplicativos. 2005. 160 f. Tese (Doutorado em Engenharia Elétrica) – Universidade Estadual de Campinas, Campinas, 2005.

[31] *AMORIM, Ronan G. de.* Introdução à Análise Convexa: Conjuntos e Funções Convexas. 2013. 79 f. Dissertação (Mestrado Profissional em Matemática em Rede Nacional) – Instituto de Matemática e Estatística, Universidade Federal de Goiás, Goiânia, 2013.

[32] *SILVA, Alvaro A. da.* Funções convexas e desigualdades: uma abordagem no ensino médio. 2015. 68 f. Dissertação (Mestrado Profissional em Matemática em Rede Nacional) – Universidade Estadual Paulista "Júlio de Mesquita Filho", São José do Rio Preto, 2015.

[33] *LIMA, Elon L.* Análise Real, vol. 2. Rio de Janeiro: IMPA, 2004.