.

.

# Integrais e aplicações

***Rafael de Freitas Manço***

SERVIÇO DE PÓS-GRADUAÇÃO DO ICMC-USP

Data de Depósito:

Assinatura: ______________________

**Rafael de Freitas Manço**

# Integrais e aplicações

Dissertação apresentada ao Instituto de Ciências Matemáticas e de Computação – ICMC-USP, como parte dos requisitos para obtenção do título de Mestre – Programa de Mestrado Profissional em Matemática. *VERSÃO REVISADA*

Área de Concentração: Matemática

Orientadora: Profa. Dra. Katia Andreia Gonçalves de Azevedo

**USP – São Carlos**
**Outubro de 2016**

Ficha catalográfica elaborada pela Biblioteca Prof. Achille Bassi
e Seção Técnica de Informática, ICMC/USP,
com os dados fornecidos pelo(a) autor(a)

Mi

Manço, Rafael de Freitas
Integrais e aplicações / Rafael de Freitas Manço; orientadora Katia Andreia Gonçalves de Azevedo. - São Carlos - SP, 2016.
111 p.

Dissertação (Mestrado - Programa de Pós-graduação em Mestrado Profissional em Matemática em Rede Nacional) - Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, 2016.

1. Integral de Riemann. 2. Integral de Riemann-Stieltjes. 3. Variáveis Aleatórias. I. Azevedo, Katia Andreia Gonçalves de, orient. II. Título.

**Rafael de Freitas Manço**

# Integral and aplications

Master dissertation submitted to the Instituto de Ciências Matemáticas e de Computação – ICMC-USP, in partial fulfillment of the requirements for the degree of the Master – Program in Mathematics Professional Master. *FINAL VERSION*

Concentration Area: Mathematics

Advisor: Profa. Dra. Katia Andreia Gonçalves de Azevedo

**USP – São Carlos**
**October 2016**

*Aos meus pais que sempre me incentivaram a estudar e a lutar pelos meus sonhos.*

# AGRADECIMENTOS

Agradeço aos meus colegas de curso pela amizade, pelo companheirismo, pela troca de experiências e pela força que me deram durante todo o curso.

Aos meus professores de PROFMAT-USP, pelos ensinamentos.

À Profª. Dra. Katia Andreia Gonçalves de Azevedo pela orientação nessa dissertação, pela paciência, pelos momentos de aprendizado, por toda a dedicação e disponibilidade, sinto-me lisonjeado em ter sido seu aluno.

Aos membros da banca, Sandra Maria Semensato de Godoy, Selma Helena de Jesus Nicola e Rafael Andres Rosales Mitrowsky pelas colaborações.

À CAPES pelo apoio financeiro.

A todos os colegas de trabalho, amigos e familiares que de alguma forma contribuíram durante todo o processo.

*"O teu futuro é duvidoso,*

*eu vejo grana eu vejo dor,*

*no paraíso perigoso*

*que a palma da tua mão mostrou"*

*(Cazuza)*

# RESUMO

NOME PARA REFERÊNCIA. **Integrais e aplicações**. 2016. 111 f. Dissertação (Mestrado – Programa de Mestrado Profissional em Matemática) – Instituto de Ciências Matemáticas e de Computação (ICMC/USP), São Carlos – SP.

O intuito deste trabalho é fazer uma análise sobre o processo de integração de funções. Existem muitas generalizações do conceito de integração abordado inicialmente por meio da integral de Riemann, como por exemplo, a integral de Riemann-Stieltjes, Lebesgue, Henstock-Kurzweil entre outras. Abordaremos especialmente a integral de Riemann-Stieltjes, e mostraremos a limitação da integral de Riemann no estudo de convergência de funções, indicando a necessidade de se generalizar o processo de integração. Faremos uma aplicação da integral de Riemann-Stieltjes no estudo de variáveis aleatórias e apresentamos uma proposta de abordagem, para a sala de aula, sobre o deslocamento e distância percorrida por um objeto em movimento retilíneo uniforme associado a área.

**Palavras-chave:** Integral de Riemann, Integral de Riemann-Stieltjes, Variáveis Aleatórias.

# ABSTRACT

NOME PARA REFERÊNCIA. **Integrais e aplicações**. 2016. 111 f. Dissertação (Mestrado – Programa de Mestrado Profissional em Matemática) – Instituto de Ciências Matemáticas e de Computação (ICMC/USP), São Carlos – SP.

The aim of this work is analizing the process of integration of functions. There are many generalizations of the integration concept originally addressed by Riemann integral such as the Riemann-Stieltjes integral, Lebesgue integral, Henstock-Kurzweil integral, among others. We will be specially concerned with the integral of Riemann-Stieltjes and we will show the limitations of Riemann integral about convergence of functions, leading to the need to generalize the integration process. We will apply Riemann-Stieltjes integral for the study of random variables and present an approach to the classroom, on the displacement and distance traveled by an object in uniform rectilinear motion associated to concept of area.

**Key-words:** Riemann integral, Riemann-Stieltjes integral, Random Variables.

# LISTA DE ILUSTRAÇÕES


Figura 1 – Construindo o conjunto de Cantor. . . . . . . . . . . . . . . . . . . . . . . . 41

Figura 2 – Função de Distribuição para Número de Caras. . . . . . . . . . . . . . . . . 65

Figura 3 – Função de Distribuição para o Exemplo 3.5. . . . . . . . . . . . . . . . . . . 67

Figura 4 – Função de Distribuição Mista. . . . . . . . . . . . . . . . . . . . . . . . . . 68

Figura 5 – Função de Distribuição Mista. . . . . . . . . . . . . . . . . . . . . . . . . . 71

Figura 6 – Gráfico da função de distribuição. . . . . . . . . . . . . . . . . . . . . . . . 83

Figura 7 – Convergência simples ou pontual. . . . . . . . . . . . . . . . . . . . . . . . 87

Figura 8 – A convergência é uniforme se $f_n : [-3,3] \to \mathbb{R}$, com $f_n(x) = x/n$. . . . . . . 88

Figura 9 – Somas Superiores. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 92

Figura 10 – Somas Inferiores. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 93

Figura 11 – Soma inferior e superior com 10 retângulos. . . . . . . . . . . . . . . . . . 94

Figura 12 – Soma inferior e superior com 15 retângulos. . . . . . . . . . . . . . . . . . 95

Figura 13 – Soma inferior e superior com 50 retângulos. . . . . . . . . . . . . . . . . . 95

Figura 14 – A área do retângulo é numericamente igual ao módulo do deslocamento $\Delta S$. 97

Figura 15 – Movimento uniforme. . . . . . . . . . . . . . . . . . . . . . . . . . . . . 98

Figura 16 – Movimento uniforme. . . . . . . . . . . . . . . . . . . . . . . . . . . . . 99

Figura 17 – Calculando o deslocamento de um móvel em MUV. . . . . . . . . . . . . 99

Figura 18 – Calculando o deslocamento de um móvel em MUV. . . . . . . . . . . . . 100

Figura 19 – Neste movimento há mudança de sentido. . . . . . . . . . . . . . . . . . . 101

Figura 20 – Movimento uniformemente variado. . . . . . . . . . . . . . . . . . . . . . 103

Figura 21 – Movimento uniformemente variado. . . . . . . . . . . . . . . . . . . . . . 104

Figura 22 – A área é numericamente igual ao deslocamento. . . . . . . . . . . . . . . . 105

Figura 23 – Espaço percorrido e deslocamento. . . . . . . . . . . . . . . . . . . . . . . 105

Figura 24 – O deslocamento é dado por $A_1 - A_2$ e o espaço percorrido é dado por $A_1 + A_2$. 106

Figura 25 – Gráfico da função $v(t) = -t^2 + t$ com valor 20 para o controle deslizante. . . 107

Figura 26 – Gráfico da função $v(t) = -t^2 + t$ com valor 100 para o controle deslizante. . 107

# SUMÁRIO


**1 Pré-Requisitos . . . . . . . . . . 21**

1.1 Os Números Reais . . . . . . . . . . 21

*1.1.1 Sequências* . . . . . . . . . . 22

*1.1.2 Noções topológicas* . . . . . . . . . . 25

*1.1.3 Conjuntos Compactos* . . . . . . . . . . 28

1.2 Funções Reais . . . . . . . . . . 30

*1.2.1 Limites de Funções* . . . . . . . . . . 30

*1.2.2 Funções Contínuas* . . . . . . . . . . 34

*1.2.3 Continuidade Uniforme* . . . . . . . . . . 37

*1.2.4 Derivadas* . . . . . . . . . . 38

*1.2.5 O conjunto e a função de Cantor* . . . . . . . . . . 40

**2 A integral de Riemann . . . . . . . . . . 45**

2.1 Integral de Riemann . . . . . . . . . . 46

2.2 Condições de Integrabilidade . . . . . . . . . . 49

2.3 Propriedades da Integral . . . . . . . . . . 52

2.4 O Teorema Fundamental do Cálculo . . . . . . . . . . 57

2.5 A integral como limite de somas de Riemann . . . . . . . . . . 58

**3 A integral de Riemann-Stieltjes . . . . . . . . . . 61**

3.1 Probabilidade e Variáveis Aleatórias . . . . . . . . . . 62

3.2 Valor Esperado de uma Variável Aleatória . . . . . . . . . . 69

3.3 Somas de Riemann-Stieltjes . . . . . . . . . . 73

3.4 Existência da integral de Riemann-Stieltjes . . . . . . . . . . 73

3.5 Propriedades da integral de Riemann-Stieltjes . . . . . . . . . . 80

3.6 Valor esperado e Integral de Riemann-Stieltjes . . . . . . . . . . 82

3.7 Sequências de Funções . . . . . . . . . . 86

4 Aplicações da integral de Riemann . . . . . . . . . . . . . . . . . . . . . . . . 91
4.1 Fazendo estimativas com somas finitas . . . . . . . . . . . . . . . . . . . 92
4.2 Deslocamento e espaço percorrido . . . . . . . . . . . . . . . . . . . . . 96
Referências . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 111

# INTRODUÇÃO

Este trabalho tem como proposta fazer um estudo sobre dois tipos de integral, a integral de Riemann e a integral de Riemann-Stieltjes, mostrando uma aplicação de cada uma dessas integrais. Apresentamos a integral de Riemann e faremos uma aplicação desta integral no estudo do espaço percorrido e do deslocamento de um móvel em movimento retilíneo uniforme. De maneira semelhante, apresentamos a integral de Riemann-Stieltjes, uma integral mais geral que a de Riemann e que possui aplicações interessantes em teoria de probabilidades.

O primeiro capítulo é pré-requisito para os demais. Nele serão abordados alguns conceitos básicos de análise na reta.

No segundo capítulo, apresentamos a integral de Riemann, as condições para que uma função seja Riemann integrável e suas propriedades.

No terceiro capítulo tratamos de probabilidade e variáveis aleatórias, com ênfase para o cálculo do valor esperado de uma variável aleatória. Também apresentamos a integral de Riemann-Stieltjes, suas propriedades, condições de existência e como esta integral unifica as expressões para o cálculo do valor esperado.

O quarto capítulo é uma proposta para o estudo do deslocamento e do espaço percorrido por um móvel em movimento retilíneo. Faremos um exemplo que mostra como os alunos podem calcular aproximações para a área de uma região com contorno curvo, a utilização do software GeoGebra vai permiti-los uma visualização das aproximações por somas superiores e inferiores. O cálculo da área de uma região será utilizado para se determinar o deslocamento e o espaço percorrido por um móvel em movimento retilíneo.

CAPÍTULO

1

# PRÉ-REQUISITOS

Neste capítulo, serão apresentados alguns conceitos básicos de análise matemática que são pré-requisitos para os capítulos seguintes. Nossas principais referências para o estudo de análise na reta são [5] e [1]. Na seção 1.1, o ponto de partida será o conjunto dos números reais, com destaque para a propriedade do supremo. Serão apresentados alguns resultados sobre sequências de números reais e algumas noções topológicas. As noções de limite, continuidade e derivada serão abordadas na seção 1.2. Finalizamos o capítulo com o conjunto e a função de Cantor.

## 1.1 Os Números Reais

Designaremos o conjunto dos números reais por $\mathbb{R}$.

Dentre as propriedades do módulo de um número real, destacamos a desigualdade triangular: dados quaisquer reais $x$ e $y$, temos que $|x+y| \leq |x|+|y|$.

Um conjunto $X \subset \mathbb{R}$ diz-se limitado superiormente se existe algum $b \in \mathbb{R}$ tal que $x \leq b$ para todo $x \in X$. Neste caso, diz-se que $b$ é uma cota superior de $X$. Analogamente, diz-se que o conjunto $X \subset \mathbb{R}$ é limitado inferiormente quando existe $a \in \mathbb{R}$ tal que $a \leq x$ para todo $x \in X$. O número $a$ chama-se então cota inferior de $X$. Um conjunto $X \subset \mathbb{R}$ é limitado quando é limitado superiormente e inferiormente.

Seja $X \subset \mathbb{R}$ limitado superiormente e não vazio. O número $b \in \mathbb{R}$ chama-se supremo do conjunto $X$ quando cumpre as seguintes condições:

**I:** Para todo $x \in X$, tem-se $x \leq b$;

**II:** Se $c \in \mathbb{R}$ é tal que $x \leq c$ para todo $x \in X$, então $b \leq c$.

Assim, $b$ é a menor das cotas superiores e escrevemos $b = sup\, X$.

Analogamente, se $X \subset \mathbb{R}$ é um conjunto não vazio limitado inferiormente, o número real $a$ chama-se ínfimo do conjunto $X$, e escreve-se $a = inf\, X$ se $a$ é a maior das cotas inferiores. Isto equivale às duas afirmações:

**I:** Para todo $x \in X$, tem-se $a \leq x$;

**II:** Se $c \leq x$ para todo $x \in X$, então $c \leq a$.

Admitiremos a seguinte importante propriedade dos números reais.

**Propriedade do supremo:** Todo conjunto de números reais, não vazio e limitado superiormente, admite supremo. Analogamente, todo conjunto de números reais, não vazio e limitado inferiormente, admite ínfimo.

**Propriedade dos intervalos encaixantes:** Dada uma sequência de intervalos tais que $I_1 \supset I_2 \supset ... \supset I_n \supset ...$ onde $I_n = [a_n, b_n]$, com $a_n \leq b_n$, existe pelo menos um número real $c$ tal que $c \in I_n$ para todo $n \in \mathbb{N}$. De fato, as inclusões $I_n \supset I_{n+1}$ significam que

$$a_1 \leq a_2 \leq ... \leq a_n \leq ... \leq b_n \leq ... \leq b_2 \leq b_1.$$

O conjunto $A = \{a_1, a_2, ..., a_n, ...\}$ é, portanto, limitado superiormente. Seja $c = sup\, A$. É claro que $a_n \leq c$, para todo $n \in \mathbb{N}$. Além disso, como cada $b_n$ é cota superior de $A$, temos $c \leq b_n$ para todo $n \in \mathbb{N}$. Portanto $c \in I_n$ qualquer que seja $n \in \mathbb{N}$.

### 1.1.1 Sequências

**Definição 1.1.** Uma sequência de números reais é uma função $f : \mathbb{N} \to \mathbb{R}$, que associa a cada número natural $n$ um número real $x_n$, chamado o n-ésimo termo da sequência.

Para representar uma sequência, escrevemos $(x_1, x_2, x_3, ..., x_n, ...)$ ou $(x_n)_{n \in \mathbb{N}}$ ou $(x_n)$.

Uma sequência $(x_n)$ se diz limitada superiormente quando existe $K \in \mathbb{R}$ tal que $x_n \leq K$ e limitada inferiormente se existir $k \in \mathbb{R}$ tal que $k \leq x_n$, para todo $n \in \mathbb{N}$.

É limitada se for limitada superiormente e inferiormente, isto é, se existir $M$ tal que $|x_n| \leq M, \forall n \in \mathbb{N}$.

**Definição 1.2.** Dada uma sequência $x = (x_n)_{n\in\mathbb{N}}$, uma subsequência de $x$ é a restrição da função $x$ a um subconjunto infinito $\mathbb{N}' = \{n_1 < n_2 < ... < n_k < ...\}$ de $\mathbb{N}$. Escreve-se $x' = (x_n)_{n\in\mathbb{N}'}$ ou $(x_{n_1}, x_{n_2}, ..., x_{n_k}, ...)$ ou $(x_{n_k})_{k\in\mathbb{N}}$.

**Definição 1.3.** Diz-se que uma sequência $(x_n)$ converge para o número $L$, ou tem limite $L$ se, dado qualquer $\varepsilon > 0$, é sempre possível encontrar um número $N$ tal que

$$n > N \Rightarrow |x_n - L| < \varepsilon.$$

Escreve-se,

$$\lim_{n\to\infty} x_n = L, \ \lim x_n = L \text{ ou } x_n \to L.$$

Uma sequência que não é convergente é dita divergente.

**Definição 1.4.** Uma sequência $(x_n)$ chama-se monótona não-decrescente quando se tem $x_n \leq x_{n+1}$ para todo $n \in \mathbb{N}$ e monotóna não-crescente se $x_{n+1} \leq x_n$ para todo $n \in \mathbb{N}$. No caso em que $x_{n+1} > x_n$ para todo $n$, a sequência é dita crescente e, no caso em que $x_{n+1} < x_n$ para todo $n$, a sequência é dita decrescente.

**Teorema 1.1.** Toda sequência convergente é limitada.

*Demonstração.* Dado qualquer $\varepsilon > 0$, existe um índice $N$ tal que

$$n > N \Rightarrow L - \varepsilon < x_n < L + \varepsilon.$$

Isto nos diz que, a partir do índice $n = N + 1$, a sequência é limitada superiormente por $L + \varepsilon$ e inferiormente por $L - \varepsilon$. Seja $K = max\{|x_1|, |x_2|, ..., |x_N|, |L - \varepsilon|, |L + \varepsilon|\}$. Então $|x_n| \leq K$, para todo $n \in \mathbb{N}$, o que prova que a sequência é limitada. □

**Teorema 1.2.** Se uma sequência $(x_n)$ converge para um limite $L$, e se $A < L < B$, então, a partir de um índice $N$, $A < x_n < B$.

*Demonstração.* Dado qualquer $\varepsilon > 0$, existe $N$ tal que, a partir desse índice, $L - \varepsilon < x_n < L + \varepsilon$. Considere $\varepsilon < min\{L - A, B - L\}$, desta forma temos $L - \varepsilon > L - (L - A) = A$ e $L + \varepsilon < L + (B - L) = B$. Em consequência, $n > N \Rightarrow A < x_n < B$, o que completa a demonstração. □

**Teorema 1.3.** Toda sequência monótona limitada é convergente.

*Demonstração.* Seja $(x_n)$ monótona não-decrescente e limitada. Considere $X = \{x_1, x_2, ..., x_n, ...\}$ e $a = sup\, X$. Podemos afirmar que $a = \lim x_n$ . De fato, dado $\varepsilon > 0$, o número $a - \varepsilon$ não é cota superior de $X$ ( pois $a - \varepsilon < a$ e $a = sup\, X$). Logo existe $n_0 \in \mathbb{N}$ tal que $a - \varepsilon < x_{n_0} \leq a$. Como a sequência é não-decrescente, se $n > n_0 \Rightarrow a - \varepsilon < x_{n_0} \leq x_n \leq a < a + \varepsilon$ e daí $\lim x_n = a$.

Analogamente, se $(x_n)$ é não-crescente e limitada então $b = \lim x_n$, onde $b$ é o ínfimo do conjunto dos valores $x_n$. De fato, para todo $\varepsilon > 0$, o número $b + \varepsilon$ não é cota inferior de $X$ (pois $b + \varepsilon > b$ e $b = inf\, X$). Logo existe $n_0 \in \mathbb{N}$ tal que $b \leq x_{n_0} < b + \varepsilon$. Como a sequência é não-crescente, se $n > n_0$, então, $b - \varepsilon < b \leq x_n \leq x_{n_0} < b + \varepsilon$.

Os casos em que a sequência é crescente ou decrescente são provados de modo análogo.

□

**Teorema 1.4.** Se $\lim(x_n) = a$ então toda subsequência de $(x_n)$ converge para o limite $a$.

*Demonstração.* Seja $(x_{n_1}, x_{n_2}, ..., x_{n_k}, ...)$ uma subsequência de $(x_n)$. Dado qualquer intervalo aberto $I$ de centro $a$, existe $n_0 \in \mathbb{N}$ tal que todos os termos $x_n$, com $n > n_0$, pertencem a $I$. Em particular, todos os termos $x_{n_k}$, com $n_k > n_0$ também pertencem a $I$. Logo $\lim x_{n_k} = a$. □

**Teorema 1.5. Bolzano-Weierstrass:** Toda sequência limitada de números reais possui uma subsequência convergente.

*Demonstração.* Basta mostrar que toda sequência $(x_n)$ possui uma subsequência monótona. Para isso vamos definir o conceito de termo destacado. Um termo $x_n$ da sequência dada é destacado quando $x_n \geq x_p$ para todo $p > n$. Seja $D \subset \mathbb{N}$ o conjunto dos índices $n$ tais que $x_n$ é um termo destacado. Se $D$ for um conjunto infinito, $D = \{n_1 < n_2 < ... < n_k < ...\}$, então a subsequência $(x_n)_{n \in D}$ será monótona não-crescente. Se, entretanto, $D$ for finito, considere $n_1 \in \mathbb{N}$ maior do que todos os $n \in D$. Então $x_{n_1}$ não é destacado, logo existe $n_2 > n_1$ com $x_{n_1} < x_{n_2}$. Por sua vez, $x_{n_2}$ não é destacado, logo existe $n_3 > n_2$ com $x_{n_1} < x_{n_2} < x_{n_3}$. Prosseguindo, obtemos uma subsequência crescente $x_{n_1} < x_{n_2} < ... < x_{n_k} < ...$, mostrando assim que toda sequência possui uma subsequência monótona que pelo Teorema 1.3, é convergente, já que a sequência é limitada. □

## 1.1.2 Noções topológicas

**Definição 1.5.** Um ponto $x$ é interior a um conjunto $A \subset \mathbb{R}$, se existe um intervalo $(a,b)$ tal que $x \in (a,b) \subset A$.

**Definição 1.6.** Um conjunto $A$ é aberto se todos os seus pontos são interiores.

Por exemplo, o intervalo $(1,2)$ é um conjunto aberto pois todos os seus pontos são interiores. Todo intervalo aberto $(a,b)$ de números reais é um conjunto aberto. Por outro lado, o intervalo $[1,2)$ não é um conjunto aberto, pois 1 não é interior ao conjunto, uma vez que nenhum intervalo aberto com centro em 1 está contido em $[1,2)$.

**Definição 1.7.** Chama-se vizinhança de um ponto $x$, qualquer intervalo aberto que contenha $x$ . Dado $\varepsilon > 0$, o intervalo $V_\varepsilon(x) = (x-\varepsilon,\ x+\varepsilon)$ será chamado, vizinhança simétrica de $x$ de raio $\varepsilon$.

Assim, um ponto $x \in A \subset \mathbb{R}$ é interior a $A$ se existe uma vizinhança $V_\varepsilon(x) \subset A$.

**Teorema 1.6.** Se $A_1, A_2, ..., A_n$ são subconjuntos abertos de $\mathbb{R}$, então $A_1 \cup A_2 \cup ... \cup A_n = \bigcup_{\lambda=1}^{n} A_\lambda$, $\lambda \in \mathbb{N}$, também é um conjunto aberto.

*Demonstração.* Seja $x$ um elemento do conjunto $\bigcup_{\lambda=1}^{n} A_\lambda$, ou seja, $x$ pertence a algum $A_\lambda, \lambda \in \mathbb{N}$. Sendo $A_\lambda$ aberto, existe $V_\varepsilon(x) \subset A_\lambda$ e, é claro que, $A_\lambda \subset \bigcup_{\lambda=1}^{n} A_\lambda$, desta forma temos que $V_\varepsilon(x) \subset \bigcup_{\lambda=1}^{n} A_\lambda$, o que mostra que $x$ é ponto interior ao conjunto $\bigcup_{\lambda=1}^{n} A_\lambda$. □

De modo geral podemos dizer que a reunião de uma família qualquer de conjuntos abertos é um conjunto aberto. Para mostrar esta afirmação, considere $(A_\lambda)_{\lambda \in J}$ uma família qualquer de conjuntos abertos e seja $A = \bigcup_{\lambda \in J} A_\lambda$. Se $x \in A$ então existe $\overline{\lambda} \in J$ tal que $x \in A_{\overline{\lambda}}$ , por hipótese $A_{\overline{\lambda}}$ é aberto, então existe $V_\varepsilon(x) \subset A_{\overline{\lambda}}$ e $A_{\overline{\lambda}} \subset A$, o que mostra que $x$ é ponto interior ao conjunto $A$, logo $A$ é aberto.

**Teorema 1.7.** Se $A_1,\ A_2 \subset \mathbb{R}$ são conjuntos abertos, então $A_1 \cap A_2$ também é um conjunto aberto.

*Demonstração.* Se $x \in A_1 \cap A_2$, então $x \in A_1$ e $x \in A_2$. Sendo $A_1$ um conjunto aberto, existe $V_{\varepsilon_1}(x)$ tal que $V_{\varepsilon_1}(x) \subset A_1$. Analogamente sendo $A_2$ um conjunto aberto, existe $V_{\varepsilon_2}(x)$ tal que $V_{\varepsilon_2}(x) \subset A_2$. Seja $\varepsilon = min\{\varepsilon_1, \varepsilon_2\}$ , então $V_\varepsilon(x) \subset A_1 \cap A_2$ o que mostra que $x$ é ponto interior ao conjunto $A_1 \cap A_2$ . □

Do Teorema 1.7, resulta que a interseção $A_1 \cap A_2 \cap \ldots \cap A_n$ de um número finito de conjuntos abertos é um conjunto aberto. De fato, sejam $A_1, A_2, \ldots, A_n, A_{n+1}$ conjuntos abertos. Suponha que $A_1 \cap A_2 \cap \ldots \cap A_n$ é aberto, então $A_1 \cap A_2 \cap \ldots \cap A_n \cap A_{n+1} = (A_1 \cap A_2 \cap \ldots \cap A_n) \cap (A_{n+1})$ é aberto conforme o Teorema 1.7, e o resultado segue por indução.

**Observação:** Não é sempre verdade que a interseção qualquer de uma família de conjuntos abertos é um conjunto aberto. Tomemos, por exemplo, o conjunto $\bigcap_{n \in \mathbb{N}} (\frac{-1}{n}, \frac{1}{n}) = \{0\}$, a interseção desses conjuntos é o conjunto formado apenas pelo elemento 0 e tal conjunto não é aberto.

**Definição 1.8.** Seja $A \subset \mathbb{R}$. Um ponto $x \in \mathbb{R}$ diz-se aderente a $A$ quando $x = \lim x_n$ com $x_n \in A$.

De posse desta definição, fica claro que se $x \in A$ então $x$ é aderente a $A$, pois basta considerar a sequência de elementos iguais a $x$ que, obviamente, tal sequência converge para $x$.

**Definição 1.9.** Chama-se fecho de $A$ o conjunto $\overline{A}$ formado pelos pontos aderentes ao conjunto A.

Nem todo ponto de $\overline{A}$ pertence ao conjunto $A$, por exemplo, se considerarmos o intervalo $(a,b)$, os pontos $a$ e $b$ são aderentes a $(a,b)$ pois $a = lim(a + \frac{1}{n})$ com $n > \frac{1}{b-a}$ e analogamente $b = lim(b - \frac{1}{n})$ com $n > \frac{1}{b-a}$.

**Teorema 1.8.** Um ponto $x$ é aderente a um conjunto $A \subset \mathbb{R}$ se, e somente se, qualquer vizinhança de $x$ contém pontos de $A$.

*Demonstração.* Seja $x$ aderente a $A$. Temos que $x = \lim x_n$, $x_n \in A$, considere $V(x)$ tal que $V(x) \supset (x - \varepsilon, x + \varepsilon)$ para algum $\varepsilon > 0$. Considerando este $\varepsilon$, pela definição de limite, existe $n_0$ tal que, se $n > n_0$ então $|x - x_n| < \varepsilon$, ou seja, $x_n \in (x - \varepsilon, x + \varepsilon) \subset V(x)$, para $n > n_0$.

Se toda vizinhança de $x$ contém pontos de $A$, tomemos $I_n = (x - \frac{1}{n}, x + \frac{1}{n})$. Então para cada $n$, existe $x_n \in I_n$. Logo, $x_n \in A$ e $\lim x_n = x$. Isto completa a demonstração. □

**Definição 1.10.** Um conjunto $F \subset \mathbb{R}$ é fechado quando $F = \overline{F}$, isto é, quando contém todos os pontos aderentes a $F$.

**Exemplo 1.1.** O conjunto $\left\{1, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \ldots, \frac{n}{n+1}, \ldots\right\}$ é fechado.

**Teorema 1.9.** $F \subset \mathbb{R}$ é fechado se, e somente se, $\mathbb{R} \backslash F$ é aberto.

*Demonstração.* Seja $F$ fechado e $x \in \mathbb{R} \backslash F$, então $x \notin F$ e desta forma $x$ não é aderente a $F$ (pois se fosse aderente a $F$ ele deveria pertencer a $F$), logo existe $V_\varepsilon(x)$ que não possui elementos de $F$ e concluímos que $V_\varepsilon(x) \subset \mathbb{R} \backslash F$, e $\mathbb{R} \backslash F$ é aberto.

Por outro lado, suponha $\mathbb{R} \backslash F$ aberto e seja $x$ aderente a $F$. Se $x \in \mathbb{R} \backslash F$ haveria uma vizinhança $V(x) \subset \mathbb{R} \backslash F$, pois $\mathbb{R} \backslash F$ é aberto. Logo $V(x)$ não conteria pontos de $F$ e assim $x$ não seria aderente a $F$. Logo $x \in F$ e $F$ é fechado. □

**Teorema 1.10.** Se $F_1$ e $F_2$ são conjuntos fechados, então $F_1 \cup F_2$ é fechado.

*Demonstração.* Como $F_1$ é fechado então $\mathbb{R} \backslash F_1$ é aberto, analogamente $\mathbb{R} \backslash F_2$ é aberto. Pela lei de De Morgan temos, $(\mathbb{R} \backslash F_1) \cap (\mathbb{R} \backslash F_2) = \mathbb{R} \backslash (F_1 \cup F_2)$, que pelo Teorema 1.7 é aberto, logo $F_1 \cup F_2$ é fechado. □

Do Teorema 1.10, resulta que a união de um número finito de conjuntos fechados é fechado. De fato, seja $F_1, F_2, \ldots, F_n, F_{n+1}$ conjuntos fechados. Suponha que $F_1 \cup F_2 \cup \ldots \cup F_n$ é fechado, então, $F_1 \cup F_2 \cup \ldots \cup F_n \cup F_{n+1} = (F_1 \cup F_2 \cup \ldots \cup F_n) \cup (F_{n+1})$ é fechado conforme o Teorema 1.10 e o resultado segue por indução.

**Observação:** A união qualquer de conjuntos fechados pode não ser fechado. De fato, como um conjunto formado por um único elemento é fechado, se a união qualquer de fechados fosse fechado, deveríamos ter $(0,1) = \bigcup_{x \in (0,1)} \{x\}$ fechado, o que é um absurdo, pois sabemos que o intervalo $(0,1)$ é aberto.

**Teorema 1.11.** Se $F_1$ e $F_2$ são conjuntos fechados, então $F_1 \cap F_2$ é fechado.

*Demonstração.* Sendo $\mathbb{R} \backslash F_1$ e $\mathbb{R} \backslash F_2$ abertos, pela lei de De Morgan temos que $(\mathbb{R} \backslash F_1) \cup (\mathbb{R} \backslash F_2) = \mathbb{R} \backslash (F_1 \cap F_2)$ que é aberto pelo Teorema 1.6. Logo pelo Teorema 1.9 $F_1 \cap F_2$ é fechado. □

Se $(F_\lambda)_{\lambda \in J}$ é uma família qualquer de conjuntos fechados, então a interseção $F = \bigcap_{\lambda \in J} F_\lambda$ é um conjunto fechado.

De fato, seja $x$ um ponto aderente a $F = \bigcap_{\lambda \in J} F_\lambda$. Se $x \in \mathbb{R} \backslash F$ , então $x \notin \bigcap_{\lambda \in J} F_\lambda \Rightarrow \exists\ \overline{\lambda} \in J$ tal que $x \notin F_{\overline{\lambda}} \Rightarrow x \in \mathbb{R} \backslash F_{\overline{\lambda}}$ que é aberto. Assim, $\exists\ V_\delta(x) \subset \mathbb{R} \backslash F_{\overline{\lambda}}$. Logo, $V_\delta(x) \subset \mathbb{R} \backslash \bigcap_{\lambda \in J} F_\lambda$ o que implica que $x$ não é ponto aderente a $F$, o que é uma contradição. Logo $x \in F$.

**Definição 1.11.** Uma cisão de um conjunto $X \subset \mathbb{R}$ é uma decomposição $X = A \cup B$ tal que: $A \cap \overline{B} = \emptyset$ e $\overline{A} \cap B = \emptyset$, isto é, nenhum ponto de $A$ é aderente a $B$ e nenhum ponto de $B$ é aderente a $A$. A decomposição $X = X \cup \emptyset$ chama-se cisão trivial.

**Exemplo 1.2.** Se $X = \mathbb{R} \backslash \{0\}$, então $X = \mathbb{R}_+ \cup \mathbb{R}_-$ é uma cisão. Se $a < c < b$, então $[a,b] = [a,c] \cup (c,b]$ não é uma cisão.

**Teorema 1.12.** Um intervalo da reta só admite a cisão trivial.

*Demonstração.* Suponhamos por absurdo, que um intervalo $I$ admita uma cisão não trivial $I = A \cup B$. Tomemos $a \in A$, $b \in B$, digamos com $a < b$, logo $[a,b] \subset I$. Seja $c$ o ponto médio do intervalo $[a,b]$. Então $c \in A$ ou $c \in B$. Se $c \in A$, poremos $a_1 = c, b_1 = b$. Se $c \in B$, escreveremos $a_1 = a$ , $b_1 = c$. Em qualquer caso, obteremos um intervalo $[a_1, b_1] \subset [a,b]$, com $b_1 - a_1 = \frac{b-a}{2}$ e $a_1 \in A$, $b_1 \in B$. Por sua vez, o ponto médio de $[a_1, b_1]$ o decompõe em dois intervalos fechados justapostos de comprimento $\frac{b-a}{4}$. Um desses intervalos, que chamaremos $[a_2, b_2]$, tem $a_2 \in A$ e $b_2 \in B$. Seguindo o raciocínio anterior, obteremos uma sequência de intervalos encaixados $[a,b] \supset [a_1, b_1] \supset ... \supset [a_n, b_n] \supset ...$ com $b_n - a_n = \frac{b-a}{2^n}$, $a_n \in A$ e $b_n \in B$ para todo $n \in \mathbb{N}$. Pela propriedade dos intervalos encaixados, existe $d \in \mathbb{R}$ tal que $a_n \leq d \leq b_n$, para todo $n \in \mathbb{N}$. O ponto $d \in I = A \cup B$ não pode estar em $A$ pois $d = \lim b_n \in \overline{B}$, nem em $B$ pois $d = \lim a_n \in \overline{A}$. Contradição. □

**Definição 1.12.** Diz-se que $a \in \mathbb{R}$ é ponto de acumulação do conjunto $X \subset \mathbb{R}$ quando toda vizinhança $V$ de $a$ contém algum ponto de $X$ diferente do próprio $a$. Se $a \in X$ não é ponto de acumulação de $X$ diz-se que $a$ é um ponto isolado de $X$. Chama-se discreto todo conjunto cujos elementos são todos isolados.

**Exemplo 1.3.** O conjunto $A = \left\{\frac{1}{2}, \frac{2}{3}, \frac{3}{4}, ..., \frac{n}{n+1}, ...\right\}$ é discreto, pois seus pontos são todos isolados, e seu único ponto de acumulação é o número 1. Considerando o intervalo $(a,b)$ temos que, todos os pontos do intervalo são seus pontos de acumulação e também os extremos $a$ e $b$.

### 1.1.3 Conjuntos Compactos

**Definição 1.13.** Um conjunto $X \subset \mathbb{R}$ chama-se compacto quando é limitado e fechado.

Todo conjunto finito é compacto. Um intervalo do tipo $[a,b]$ é um conjunto compacto. Por outro lado, $(a,b)$ é limitado mas não é fechado, logo não é compacto.

**Teorema 1.13.** Um conjunto $X \subset \mathbb{R}$ é compacto se, e somente se, toda sequência de pontos em $X$, possui uma subsequência que converge para um ponto de $X$.

*Demonstração.* Se $X$ é compacto então toda sequência de pontos $x_n \in X$ é limitada, logo pelo Teorema de Bolzano-Weierstrass (Teorema 1.5), a sequência $x_n$ possui uma subsequência que converge para um ponto de $X$, pois $X$ é fechado.

Por outro lado, seja $X \subset \mathbb{R}$ um conjunto tal que toda sequência de pontos $x_n \in X$ possui uma subsequência convergente para um ponto de $X$.

**(I)** $X$ é limitado.

Se $X$ não fosse limitado, para cada $n \in \mathbb{N}$, existiria $x_n \in X$ tal que $|x_n| > n$. Assim, a sequência $(x_n)$ obtida não possuiria uma subsequência limitada e, desta forma, a subsequência não seria convergente o que contraria a hipótese, logo $X$ é limitado.

**(II)** $X$ é fechado.

Se $X$ não fosse fechado existiria um ponto $a \notin X$ tal que $a = \lim x_n, x_n \in X$. A sequência $(x_n)$ não possuiria então subsequência alguma convergindo para um ponto de $X$ pois todas suas subsequências teriam limites iguais ao número $a$, o que contraria a hipótese, logo $X$ é fechado. □

**Definição 1.14.** Chama-se cobertura de um conjunto $X$ a uma família $C$ de conjuntos $C_\lambda, \lambda \in L$, cuja reunião contém $X$. Quando todos os conjuntos $C_\lambda$ são abertos, diz-se que $C$ é uma cobertura aberta. Quando $L = \{\lambda_1, ..., \lambda_n\}$ é um conjunto finito, diz-se que $X \subset C_{\lambda_1} \cup ... \cup C_{\lambda_n}$ é uma cobertura finita. Se $L' \subset L$ é tal que, ainda se tem, $X \subset \bigcup_{\lambda' \in L'} C_{\lambda'}$, diz-se que $C' = (C_{\lambda'})_{\lambda' \in L'}$ é uma subcobertura de $C$.

**Teorema 1.14. Borel-Lebesgue:** Toda cobertura aberta de um conjunto compacto da reta possui uma subcobertura finita.

*Demonstração.* Tomemos inicialmente uma cobertura $[a,b] \subset \bigcup_{\lambda \in L} A_\lambda$ do intervalo compacto $[a,b]$. Suponhamos, por absurdo, que $C = (A_\lambda)_{\lambda \in L}$ não admita subcobertura finita. O ponto médio do intervalo $[a,b]$ o decompõe em dois intervalos de comprimento $\frac{b-a}{2}$. Pelo menos um desses intervalos, o qual chamaremos $[a_1, b_1]$, não pode ser coberto por um número finito de conjuntos $A_\lambda$. Por bisseções sucessivas obteremos uma sequência decrescente $[a,b] \supset [a_1,b_1] \supset$

$[a_2,b_2]... \supset ... \supset [a_n,b_n] \supset ...$ de intervalos tais que $b_n - a_n = \frac{b-a}{2^n}$ e nenhum $[a_n,b_n]$ pode estar contido numa reunião finita dos abertos $A_\lambda$.

Existe um número real $c$ que pertence a todos os intervalos $[a_n,b_n]$. Em particular, $c \in [a,b]$. Pela definição de cobertura, existe $\lambda \in L$ tal que $c \in A_\lambda$. Como $A_\lambda$ é aberto, temos $[c-\varepsilon, c+\varepsilon] \subset A_\lambda$ para um certo $\varepsilon > 0$. Tomando $n \in \mathbb{N}$ tal que $\frac{b-a}{2^n} < \varepsilon$ temos, então, $c \in [a_n,b_n] \subset [c-\varepsilon, c+\varepsilon]$, donde $[a_n,b_n] \subset A_\lambda$, logo $[a_n,b_n]$ pode ser coberto por apenas um dos conjuntos $A_\lambda$ o que é uma contradição. No caso geral, temos uma cobertura $X \subset \bigcup_{\lambda \in L} A_\lambda$ do compacto $X$. Tomamos um intervalo compacto $[a,b]$ que contenha $X$ e, acrescentando aos $A_\lambda$ o novo aberto $A_{\lambda_0} = \mathbb{R} - X$, obtemos uma cobertura aberta de $[a,b]$, da qual extraímos, pela parte já provada, uma subcobertura finita $[a,b] \subset A_{\lambda_0} \cup A_{\lambda_1} \cup ... \cup A_{\lambda_n}$. Como nenhum ponto de $X$ pode pertencer a $A_{\lambda_0}$, temos $X \subset A_{\lambda_1} \cup ... \cup A_{\lambda_n}$ e isto completa a demonstração. $\square$

## 1.2 Funções Reais

### *1.2.1 Limites de Funções*

Nesta seção vamos definir o que é limite de uma função e demonstrar alguns teoremas importantes para uso posterior.

**Definição 1.15.** Sejam $X \subset \mathbb{R}$, $f : X \to \mathbb{R}$ e $a$ um ponto de acumulação do conjunto $X$. Diz-se que o número real $L$ é limite de $f(x)$ quando $x$ tende para $a$, e escreve-se $\lim\limits_{x \to a} f(x) = L$, quando, para todo $\varepsilon > 0$ dado, pode-se obter $\delta > 0$ tal que $|f(x) - L| < \varepsilon$, sempre que $x \in X$ e $0 < |x-a| < \delta$. Em símbolos:

$$\lim_{x \to a} f(x) = L \Leftrightarrow \forall \varepsilon > 0, \exists\ \delta > 0; x \in X, 0 < |x-a| < \delta \Rightarrow |f(x) - L| < \varepsilon.$$

A restrição $0 < |x-a|$ significa que $x \neq a$. Portanto, o valor $f(a)$ não tem importância alguma quando se quer determinar $L$, é essencial que $a$ seja um ponto de acumulação do conjunto $X$ mas é irrelevante que $f$ esteja ou não definida no ponto $a$.

**Definição 1.16.** Dizemos que $f(x)$ tem um limite à direita $L$ em a e escrevemos $\lim\limits_{x \to a^+} f(x) = L$ se para todo $\varepsilon > 0$, pode-se obter $\delta > 0$ tal que , para todo $x \in X$, com $a < x < a + \delta \Rightarrow |f(x) - L| < \varepsilon$. Dizemos que $f$ tem um limite à esquerda $L$ em $a$ e escrevemos $\lim\limits_{x \to a^-} f(x) = L$ se para todo $\varepsilon > 0$, pode-se obter $\delta > 0$ tal que, para todo $x \in X$, $a - \delta < x < a \Rightarrow |f(x) - L| < \varepsilon$.

**Definição 1.17.** Dizemos que $f(x)$ tem limite $L$ para $x$ tendendo a $+\infty$ e escrevemos $\lim\limits_{x\to+\infty} f(x) = L$ se, para cada $\varepsilon > 0$, existe um número $N$ correspondente tal que, para todos os valores de $x \in X$, com $x > N \Rightarrow |f(x) - L| < \varepsilon$. Dizemos que $f(x)$ tem limite $L$ para $x$ tendendo a $-\infty$ e escrevemos $\lim\limits_{x\to-\infty} f(x) = L$ se, para cada $\varepsilon > 0$, existe um número $M$ correspondente tal que, para todos os valores de $x \in X$, com $x < M \Rightarrow |f(x) - L| < \varepsilon$.

**Teorema 1.15.** **A unicidade do limite:** Sejam $f : X \to \mathbb{R}$ e $a$ um ponto de acumulação de $X$. Se $\lim\limits_{x\to a} f(x) = L_1$ e $\lim\limits_{x\to a} f(x) = L_2$ então $L_1 = L_2$.

*Demonstração.* Suponha por absurdo que $L_1 \neq L_2$, temos:

$$\lim_{x\to a} f(x) = L_1 \Leftrightarrow \forall\varepsilon > 0, \exists\ \delta_1 > 0; x \in X, 0 < |x-a| < \delta_1 \Rightarrow |f(x) - L_1| < \varepsilon. \text{ (1)}$$

$$\lim_{x\to a} f(x) = L_2 \Leftrightarrow \forall\varepsilon > 0, \exists\ \delta_2 > 0; x \in X, 0 < |x-a| < \delta_2 \Rightarrow |f(x) - L_2| < \varepsilon. \text{ (2)}$$

Escrevendo $L_1 - L_2$ como $L_1 - f(x) + f(x) - L_2$ e aplicando a desigualdade triangular temos:

$$|L_1 - L_2| = |L_1 - f(x) + f(x) - L_2| \leq |L_1 - f(x)| + |f(x) - L_2| = |f(x) - L_1| + |f(x) - L_2|.$$

Considerando $\delta = min\{\delta_1, \delta_2\}$, temos $\delta \leq \delta_1$ e $\delta \leq \delta_2$ assim (1) e (2) estão satisfeitos e $\forall\varepsilon > 0$, considerando este $\delta$ e $x \in X$, $0 < |x-a| < \delta \Rightarrow |L_1 - L_2| < 2\varepsilon$.

Se tomarmos $\varepsilon = \frac{|L_1-L_2|}{2}$ segue que para $\varepsilon = \frac{|L_1-L_2|}{2}$, $\exists\ \delta > 0$; $x \in X$ e $0 < |x-a| < \delta \Rightarrow |L_1 - L_2| < |L_1 - L_2|$ que é uma contradição e, portanto, a nossa suposição é falsa. Logo $L_1 = L_2$. $\square$

**Teorema 1.16.** **Teorema do confronto:** Sejam $f, g, h : X \to \mathbb{R}$, $a$ um ponto de acumulação de $X$ e $\lim\limits_{x\to a} f(x) = \lim\limits_{x\to a} g(x) = L$. Se $f(x) \leq h(x) \leq g(x)$ para todo $x \in X-\ \{a\}$ então $\lim\limits_{x\to a} h(x) = L$.

*Demonstração.* $\forall\varepsilon > 0, \exists\ \delta_1 > 0$ e $\delta_2 > 0; x \in X\ 0 < |x-a| < \delta_1 \Rightarrow L - \varepsilon < f(x) < L + \varepsilon$ e $0 < |x-a| < \delta_2 \Rightarrow L - \varepsilon < g(x) < L + \varepsilon$. Seja $\delta = min\{\delta_1, \delta_2\}$. Então $x \in X, 0 < |x-a| < \delta \Rightarrow L - \varepsilon < f(x) \leq h(x) \leq g(x) < L + \varepsilon \Rightarrow |h(x) - L| < \varepsilon$. Logo $\lim\limits_{x\to a} h(x) = L$. $\square$

**Teorema 1.17.** Seja $f$ uma função monótona e limitada, definida num intervalo $I$, do qual $x = a$ é ponto de acumulação à direita ou à esquerda. Então $f(x)$ tem limite com $x \to a^+$ ou $x \to a^-$, respectivamente.

*Demonstração.* Suponhamos que $f$ seja uma função não-decrescente e $x = a$ seja ponto de acumulação à esquerda. Nesse caso, basta supor que $f$ seja limitada à direita. Seja $L$ o supremo dos valores de $f(x)$, para todo $x \in I$, $x < a$. Provaremos que $\lim_{x\to a^-} f(x) = L$. De fato, dado qualquer $\varepsilon > 0$, existe $\delta > 0$ tal que $L - \varepsilon < f(a-\delta) \leq L$. Mas $f$ é não-decrescente, assim, $f(a-\delta) \leq f(x)$ para $a - \delta < x$ e $x \in I$; logo,

$$x \in I, a-\delta < x < a \Rightarrow L - \varepsilon < f(x) \leq L,$$

que prova o resultado.

As demonstrações dos outros casos são análogas. $\square$

**Teorema 1.18.** Sejam $f : X \to \mathbb{R}$ e $a$ um ponto de acumulação de $X$. A fim de que $\lim_{x\to a} f(x) = L$ é necessário e suficiente que, para toda sequência de pontos $x_n \in X - \{a\}$ com $\lim x_n = a$, tenha-se $\lim f(x_n) = L$.

*Demonstração.* Suponhamos que $\lim_{x\to a} f(x) = L$ e considere uma sequência $x_n \in X - \{a\}$ com $\lim_{x\to a} x_n = a$. Então $\forall \varepsilon > 0, \exists\ \delta > 0; x \in X,\ 0 < |x-a| < \delta \Rightarrow |f(x) - L| < \varepsilon$. Existe também $n_0 \in \mathbb{N}$ tal que $n > n_0 \Rightarrow 0 < |x_n - a| < \delta$ uma vez que $x_n \neq a$, $\forall n \in \mathbb{N}$ e $\lim x_n = a$. Logo se $n > n_0$ então, $|f(x_n) - L| < \varepsilon \Rightarrow \lim_{x\to a} f(x_n) = L$.

Por outro lado, suponhamos que $x_n \in X - \{a\}$ e $\lim x_n = a$ impliquem $\lim f(x_n) = L$ e provemos que $\lim_{x\to a} f(x) = L$.

Negar esta igualdade implicaria afirmar a existência de um número $\varepsilon > 0$ com a seguinte propriedade: $\forall n \in \mathbb{N}$ podemos achar $x_n \in X$ tal que $0 < |x_n - a| < \frac{1}{n}$ mas $|f(x_n) - L| \geq \varepsilon$. Então, teríamos $x_n \in X - \{a\}, \lim x_n = a$ sem que $\lim f(x_n) = L$. Esta contradição completa a demonstração. $\square$

**Teorema 1.19. Operações com limites:** Sejam $f, g : X \to \mathbb{R}$ e $a$ um ponto de acumulação de $X$, com $\lim_{x\to a} f(x) = L_1$ e $\lim_{x\to a} g(x) = L_2$ então:

**(a)** $\lim_{x\to a}[f(x) + g(x)] = \lim_{x\to a} f(x) + \lim_{x\to a} g(x) = L_1 + L_2$;

**(b)** $\lim_{x\to a}[k.f(x)] = k.\lim_{x\to a} f(x) = k.L_1$;

**(c)** $\lim_{x\to a}[f(x).g(x)] = \lim_{x\to a} f(x).\lim_{x\to a} g(x) = L_1.L_2$;

**(d)** $\lim_{x\to a} \dfrac{f(x)}{g(x)} = \dfrac{\lim_{x\to a} f(x)}{\lim_{x\to a} g(x)} = \dfrac{L_1}{L_2}$ desde que $\lim_{x\to a} g(x) \neq 0$.

*Demonstração.* $\lim\limits_{x\to a} f(x) = L_1 \Leftrightarrow \forall\varepsilon > 0, \exists\ \delta_1 > 0;\ \ 0 < |x-a| < \delta_1 \Rightarrow |f(x)-L_1| < \frac{\varepsilon}{2}$.

$\lim\limits_{x\to a} g(x) = L_2 \Leftrightarrow \forall\varepsilon > 0, \exists\ \delta_2 > 0;\ \ 0 < |x-a| < \delta_2 \Rightarrow |g(x)-L_2| < \frac{\varepsilon}{2}$.

Seja $\delta = min\{\delta_1, \delta_2\}$. Assim:

$\forall\varepsilon > 0$ e $x \in X$ temos, $0 < |x-a| < \delta \Rightarrow |f(x)-L_1| + |g(x)-L_2| < \varepsilon$. Pela desigualdade triangular segue-se que:

$|(f(x)-L_1)+(g(x)-L_2)| \leq |f(x)-L_1| + |g(x)-L_2|$ então, $\forall\varepsilon > 0, \exists\ \delta = min\{\delta_1, \delta_2\}$; $0 < |x-a| < \delta \Rightarrow |(f(x)+g(x))-(L_1+L_2)| < \varepsilon$. Logo $\lim\limits_{x\to a}[f(x)+g(x)] = L_1 + L_2$, o que prova (a).

Provemos que, $\lim\limits_{x\to a}[k.f(x)] = k.\lim\limits_{x\to a} f(x)$.

Se $k = 0$, $k.f(x) = 0\ \forall x \in X$, logo $\lim\limits_{x\to a} k.f(x) = 0 = k.\lim\limits_{x\to a} f(x)$.

Se $k \neq 0$, $\forall\varepsilon > 0, \exists\ \delta > 0;\ \ 0 < |x-a| < \delta \Rightarrow |f(x)-L_1| < \dfrac{\varepsilon}{|k|}$. Daí, $0 < |x-a| < \delta \Rightarrow$ $|k.f(x)-k.L_1| < |k|.\dfrac{\varepsilon}{|k|} = \varepsilon$, logo $\lim\limits_{x\to a} k.f(x) = k.L_1$, o que prova (b).

Antes de provarmos que, $\lim\limits_{x\to a}[f(x).g(x)] = \lim\limits_{x\to a} f(x).\lim\limits_{x\to a} g(x)$ , provemos o seguinte lema.

**Lema 1.** Sejam $(a_n)$ e $(b_n)$ duas sequências convergentes com limites $a$ e $b$ respectivamente. Então $\lim(a_n.b_n) = (\lim a_n).(\lim b_n) = a.b$.

*Demonstração.* De fato, suponha $a \neq 0$, $b_n$ é limitada por $M > 0$ pois é convergente (Teorema 1.1). Podemos escrever $|a_n.b_n - ab| = |(a_n-a)b_n + a(b_n-b)| \leq |a_n-a||b_n| + |a||b_n-b| \leq M.|a_n-a| + |a||b_n-b|$.

Tanto $|a_n-a|$ como $|b_n-b|$ podem ficar arbitrariamente pequenos, desde que $n$ seja suficientemente grande. Assim , dado qualquer $\varepsilon > 0$, podemos fazer $|a_n-a| < \frac{\varepsilon}{2M}$ a partir de um índice $N_1$ e $|b_n-b| < \frac{\varepsilon}{2|a|}$ a partir de um índice $N_2$ , então sendo $N$ o maior desses índices , $n > N$ satisfará $n > N_1$ e $n > N_2$ simultaneamente, logo

$$n > N \Rightarrow |a_n.b_n - a.b| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

O caso $a = 0$ é análogo. □

Vamos agora provar (c).

Supondo que $f(x)$ e $g(x)$ tenham limites $L_1$ e $L_2$, respectivamente , com $x \to a$, seja $x_n \in X - \{a\}$ uma sequência convergindo para $a$. Então pela hipótese $f(x_n) \to L_1$ e $g(x_n) \to L_2$ e pelo lema $f(x_n).g(x_n) \to L_1.L_2$, logo $\lim_{x\to a} f(x).g(x) = L_1.L_2$.

Sendo $L_2 \neq 0$, vamos provar que $\lim_{x\to a} \frac{1}{g(x)} = \frac{1}{L_2}$.

Dado $\varepsilon > 0$, sabemos que existe $\delta > 0$ tal que

$$x \in X,\, 0 < |x-a| < \delta \Rightarrow |g(x) - L_2| < \frac{\varepsilon {L_2}^2}{2}.$$

Se necessário, diminuímos o $\delta$ de maneira a termos $x \in X, 0 < |x-a| < \delta \Rightarrow |g(x)| > \frac{|L_2|}{2}$.

Então, se $x \in X$, $0 < |x-a| < \delta$ teremos:

$$\left| \frac{1}{g(x)} - \frac{1}{L_2} \right| = \frac{|g(x) - L_2|}{|L_2.g(x)|} < \frac{\varepsilon {L_2}^2}{2|L_2 g(x)|} < \frac{\varepsilon .{L_2}^2}{2} . \frac{2}{{L_2}^2} = \varepsilon,$$

e isso mostra que $\lim_{x\to a} \frac{1}{g(x)} = \frac{1}{L_2}$.

Pela propriedade (c), $\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} f(x). \lim_{x\to a} \frac{1}{g(x)} = L_1 . \frac{1}{L_2} = \frac{L_1}{L_2}$, o que prova (d). □

### 1.2.2 Funções Contínuas

**Definição 1.18.** Uma função $f : X \to \mathbb{R}$, definida no conjunto $X \subset \mathbb{R}$, diz-se contínua no ponto $a \in X$ se:

$$\forall \varepsilon > 0, \exists\, \delta > 0; x \in X, |x-a| < \delta \Rightarrow |f(x) - f(a)| < \varepsilon.$$

Suponha que $f$ está definida em $a$, e que $a$ seja um ponto de acumulação de $X$. Comparando as definições de limite e continuidade, resulta que $f$ é contínua em $a$ se, e somente se, $\lim_{x\to a} f(x) = f(a)$.

**Teorema 1.20.** Sejam $f, g : X \to \mathbb{R}$ contínuas no ponto $a \in X$, com $f(a) < g(a)$. Existe $\delta > 0$ tal que $f(x) < g(x)$ para todo $x \in X \cap (a-\delta, a+\delta)$.

*Demonstração.* Tomemos $c = \frac{[g(a) + f(a)]}{2}$ e $\varepsilon = g(a) - c = c - f(a)$. Então $\varepsilon > 0$ e $f(a) + \varepsilon = g(a) - \varepsilon = c$. Pela definição de continuidade existem $\delta_1 > 0$ e $\delta_2 > 0$ tais que $x \in X, |x-a| < \delta_1 \Rightarrow f(a) - \varepsilon < f(x) < c$ e $x \in X, |x-a| < \delta_2 \Rightarrow c < g(x) < g(a) + \varepsilon$. Tome $\delta = min\{\delta_1, \delta_2\}$, daí, $x \in X, |x-a| < \delta \Rightarrow f(x) < c < g(x)$, o que prova o teorema. □

**Corolário 1.** Seja $f : X \to \mathbb{R}$ contínua no ponto $a \in X$. Se $f(a) \neq 0$, existe $\delta > 0$ tal que, para todo $x \in X \cap (a-\delta, a+\delta)$, $f(x)$ tem o mesmo sinal de $f(a)$.

*Demonstração.* De fato, suponhamos $f(a) < 0$. Então, basta tomar $g$ identicamente nula no Teorema 1.20. $\square$

**Teorema 1.21.** A função $f : X \to \mathbb{R}$ é contínua no ponto $a \in X$ se e somente se $x_n \in X, \lim x_n = a \Rightarrow \lim f(x_n) = f(a)$.

*Demonstração.* Suponha que $f$ é contínua no ponto $a \in X$ e que $x_n \in X$ com $\lim x_n = a$. Daí, $\forall \varepsilon > 0, \exists\, \delta > 0; x \in X, |x-a| < \delta \Rightarrow |f(x) - f(a)| < \varepsilon$.

Existe também $n_0 \in \mathbb{N}$ tal que se $n > n_0$, então, $0 < |x_n - a| < \delta$. Assim, se $n > n_0$, $0 < |x_n - a| < \delta$ e $|f(x_n) - f(a)| < \varepsilon$, logo, $\lim f(x_n) = f(a)$.

Por outro lado, se $f$ não for contínua no ponto $a$, existe um $\varepsilon > 0$ tal que $\forall\ \delta > 0$, pode-se achar $x_\delta \in X$ de modo que, $|x_\delta - a| < \delta$ mas $|f(x_\delta) - f(a)| \geq \varepsilon$. Tomemos $\delta = \frac{1}{n}$ e façamos $x_n = x_\delta$. Assim, $|x_n - a| < \frac{1}{n}$ mas $|f(x_n) - f(a)| \geq \varepsilon$, o que contradiz a hipótese. $\square$

**Corolário 2.** Se $f, g : X \to \mathbb{R}$ são contínuas no ponto $a \in X$ e $a$ um ponto de acumulação de $X$, então são contínuas nesse mesmo ponto as funções $f+g, f.g : X \to \mathbb{R}$ e $\dfrac{f}{g}$ desde que $g(a) \neq 0$.

O domínio da função $f/g$ é o subconjunto de $X$ formado pelos pontos $x$ tais que $g(x) \neq 0$.

*Demonstração.* De fato, das propriedades operatórias dos limites temos que:

$$\lim_{x\to a}[f(x)+g(x)] = \lim_{x\to a} f(x) + \lim_{x\to a} g(x) = f(a) + g(a);$$

$$\lim_{x\to a}[f(x).g(x)] = \lim_{x\to a} f(x). \lim_{x\to a} g(x) = f(a).g(a);$$

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \frac{\lim_{x\to a} f(x)}{\lim_{x\to a} g(x)} = \frac{f(a)}{g(a)} \text{ com } g(a) \neq 0.$$

$\square$

**Teorema 1.22.** Sejam $f : X \to \mathbb{R}$ contínua no ponto $a \in X, g : Y \to \mathbb{R}$ contínua no ponto $b = f(a) \in Y$ e $f(X) \subset Y$, de modo que a composta $g \circ f : X \to \mathbb{R}$ está bem definida. Então $g \circ f$ é contínua no ponto $a$.

*Demonstração.* Pela continuidade de $g$ no ponto $b$ temos: $\forall \varepsilon > 0, \exists\, \eta > 0; y \in Y, |y-b| < \eta \Rightarrow |g(y) - g(b)| < \varepsilon$. Por sua vez , a continuidade de $f$ no ponto $a$ assegura que existe $\delta > 0$ tal que $x \in X, |x-a| < \delta \Rightarrow |f(x) - b| < \eta$. Consequentemente, $x \in X \cap (a-\delta, a+\delta) \Rightarrow |g(f(x)) - g(b)| = |(g \circ f)(x) - (g \circ f)(a)| < \varepsilon$. $\square$

**Teorema 1.23. Teorema do valor intermediário:** Seja $f : [a,b] \to \mathbb{R}$ contínua. Se $f(a) < d < f(b)$ então existe $c \in (a,b)$ tal que $f(c) = d$.

*Demonstração.* Por absurdo suponha que não exista $c \in (a,b)$ tal que $f(c) = d$. Temos $[a,b] = A \cup B$ onde $A = \{x \in [a,b]; f(x) < d\}$ e $B = \{x \in [a,b]; f(x) > d\}$. $A \cup B$ é uma cisão pois se $x \in A$ , $x$ não pode ser aderente a $B$. Analogamente se $x \in B$ , $x$ não pode ser aderente a $A$. Como um intervalo da reta só admite a cisão trivial então $A = \emptyset$ ou $B = \emptyset$, porém $A \neq \emptyset$ pois $a \in A$ e $B \neq \emptyset$ pois $b \in B$, logo a cisão não é trivial o que é um absurdo. Portanto, existe $c \in (a,b)$ tal que $f(c) = d$. □

**Teorema 1.24.** Se $f$ for contínua no intervalo fechado $[a,b]$, então $f$ será limitada em $[a,b]$.

*Demonstração.* Suponhamos, por absurdo , que $f$ não seja limitada em $[a,b]$. Façamos $a = a_1$ e $b = b_1$, existe então $x_1$ em $[a_1,b_1]$ tal que $|f(x_1)| > 1$. Seja $c_1$ o ponto médio de $[a_1,b_1], f$ não será limitada em um dos intervalos $[a_1,c_1]$ ou $[c_1,b_1]$, suponhamos que não seja limitada em $[c_1,b_1]$ e façamos $a_2 = c_1$ e $b_2 = b_1$. Não sendo $f$ limitada em $[a_2,b_2]$, existirá $x_2 \in [a_2,b_2]$ tal que $|f(x_2)| > 2$. Prosseguindo com este raciocínio, construiremos uma sequência de intervalos $[a_1,b_1] \supset [a_2,b_2] \supset [a_3,b_3] \supset ... \supset [a_n,b_n] \supset ...$ tal que, para todo natural $n > 0$, existe $x_n \in [a_n,b_n]$ com $|f(x_n)| > n$. Segue que $\lim_{n\to\infty} |f(x_n)| = +\infty$. Seja agora $c$ o único real tal que para todo natural $n > 0, c \in [a_n,b_n]$, pelo teorema dos intervalos encaixados. Como a sequência $x_n$ converge para $c$ e $f$ é contínua em $c$, resulta que $\lim_{n\to\infty} |f(x_n)| = |f(c)|$ que está em contradição com $\lim_{n\to\infty} |f(x_n)| = +\infty$. Portanto $f$ é limitada em $[a,b]$. □

**Teorema 1.25. Teorema de Weierstrass:** Seja $f$ contínua no conjunto compacto $X \subset \mathbb{R}$. Existem $x_1$ e $x_2$ em $X$ tais que $f(x_1) \leq f(x) \leq f(x_2)$ para todo $x \in X$.

*Demonstração.* [1] Sendo $f$ contínua em $X$, pelo Teorema 1.24, $f$ será limitada em $X$, daí o conjunto $I_m f = A = \{f(x); x \in X\}$ admitirá supremo e ínfimo. Sejam $M = sup\{f(x); x \in X\}$ e $m = inf\{f(x); x \in X\}$. Assim para todo $x \in X$ , $m \leq f(x) \leq M$. Provemos que $M = f(x_1)$ para algum $x_1$ em $X$. Se tivéssemos $f(x) < M$ para todo $x \in X$, a função $g(x) = \frac{1}{M - f(x)}, x \in X$ seria contínua em $X$ mas não limitada em $X$, pois se $g$ fosse limitada em $X$, existiria $\beta > 0$ tal que, $0 < \frac{1}{M - f(x)} < \beta \Rightarrow f(x).\beta + 1 < \beta M \Rightarrow f(x) < M - \frac{1}{\beta}$. Logo $M$ não seria supremo de $A$ o que é um absurdo. Portanto, deve existir um $x_1 \in X$ tal que $f(x_1) = M$.

[1] Demonstração baseada em [3].

Por outro lado, se $f(x) > m$ para todo $x$ em $[a,b]$ , a função $h(x) = \frac{1}{f(x)-m}$ seria contínua em $X$ mas não limitada em $X$, o que é uma contradição pois se $h$ fosse limitada em $X$, existiria $\alpha > 0$ tal que $0 < \frac{1}{f(x)-m} < \alpha \Rightarrow 1 + m\alpha < \alpha f(x) \Rightarrow \frac{1}{\alpha} + m < f(x)$ o que é um absurdo pois $m$ é ínfimo de $A$, logo deve existir um $x_2 \in X$ tal que $f(x_2) = m$. $\square$

### 1.2.3 Continuidade Uniforme

**Definição 1.19.** Seja $f : X \to \mathbb{R}$ contínua. $f$ diz-se uniformemente contínua se :

$$\forall \varepsilon > 0, \exists\, \delta > 0; x,y \in X, |x-y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon.$$

Nem sempre, dado $\varepsilon > 0$, pode-se encontrar um $\delta > 0$ que sirva em todos os pontos $x \in X$ mesmo sendo $f$ contínua em todos esses pontos.

A noção de função contínua é uma noção local e a noção de função uniformemente contínua é global, isto é, não depende do comportamento da função apenas nas vizinhanças de um ponto, mas depende do comportamento da função em todo domínio onde ela está definida.

**Teorema 1.26.** A fim de que $f : X \to \mathbb{R}$ seja uniformemente contínua é necessário e suficiente que, para todo par de sequências $(x_n), (y_n)$ em $X$ com $\lim(y_n - x_n) = 0$, tenha-se $\lim[f(y_n) - f(x_n)] = 0$.

*Demonstração.* Se $f$ é uniformemente contínua e $lim(y_n - x_n) = 0$ então $\forall \varepsilon > 0, \exists\, \delta > 0; x,y \in X, |y-x| < \delta \Rightarrow |f(y) - f(x)| < \varepsilon$ e, existe, para este $\delta > 0$, $n_0 \in \mathbb{N}$ tal que $n > n_0 \Rightarrow |y_n - x_n| < \delta$. Logo $n > n_0 \Rightarrow |f(y_n) - f(x_n)| < \varepsilon$ e daí $\lim[f(y_n) - f(x_n)] = 0$. Reciprocamente, suponhamos válida a condição do enunciado. Se $f$ não fosse uniformemente contínua, existiria um $\varepsilon > 0$ tal que para todo $n \in \mathbb{N}$ poderíamos achar pontos $x_n, y_n \in X$ tais que $|y_n - x_n| < \delta_n$ e $|f(y_n) - f(x_n)| \geq \varepsilon$. Tomando $\delta_n = \frac{1}{n}$ teríamos $\lim(y_n - x_n) = 0$ sem que $\lim[f(y_n) - f(x_n)] = 0$. Esta contradição conclui a prova. $\square$

**Exemplo 1.4.** Do Teorema 1.26 podemos concluir que a função $f : \mathbb{R} \to \mathbb{R}, f(x) = x^2$ não é uniformemente contínua. Tomando $y_n = n + \frac{1}{n}$ e $x_n = n$ , com $y_n, x_n \in \mathbb{R}$ temos $\lim(y_n - x_n) = \lim(\frac{1}{n}) = 0$. Mas $f(y_n) - f(x_n) = n^2 + 2 + \frac{1}{n^2} - n^2 = 2 + \frac{1}{n^2}$ e portanto $\lim[f(y_n) - f(x_n)] = \lim(2 + \frac{1}{n^2}) = 2$.

**Definição 1.20.** Uma função $f : X \to \mathbb{R}$ chama-se lipschitziana quando existe uma constante $k > 0$ , chamada constante de Lipschitz da função $f$, tal que $|f(x) - f(y)| \leq k|x-y|$ para todos $x, y \in X$. Escrevendo esta condição na forma $\left|\frac{f(x)-f(y)}{x-y}\right| \leq k$ com $x \neq y$ percebemos que tal quociente é limitado por $k$, isto significa que as secantes não ficam próximas da vertical.

Se uma função é lipschitziana então ela é uniformemente contínua. De fato para que uma função $f : X \to \mathbb{R}$ seja uniformemente contínua, dado um $\varepsilon > 0$, deve existir um $\delta > 0$ tal que , $x, y \in X, |x-y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon$. Suponha que uma função $f$ seja lipschitziana, tomando-se $\delta = \frac{\varepsilon}{k}$ e substituindo na condição descrita acima temos $\forall \varepsilon > 0, \exists\, \delta = \frac{\varepsilon}{k} > 0; x, y \in X, |x-y| < \delta \Rightarrow |f(x) - f(y)| \leq k|x-y| < k.\frac{\varepsilon}{k} = \varepsilon$. A recíproca desta implicação não é verdadeira , isto é, existem funções que são uniformemente contínuas que não são lipschitzianas.

**Exemplo 1.5.** A função $f : \mathbb{R} \to \mathbb{R}$ dada por $f(x) = 2x+3$ é lipschitziana com constante $k = 2$ pois $|f(y) - f(x)| = |2||y-x|$. Seja $f : [0,1] \to \mathbb{R}$ dada por $f(x) = \sqrt{x}$. Então, $\frac{f(x)-f(y)}{x-y} = \frac{\sqrt{x}-\sqrt{y}}{x-y}$. Multiplicando o numerador e o denominador por $\sqrt{x}+\sqrt{y}$ ficamos com $\frac{1}{\sqrt{x}+\sqrt{y}}$. Tomando $x \neq y$ suficientemente pequenos, podemos tornar $\sqrt{x}+\sqrt{y}$ tão pequeno quanto se deseje, logo o quociente $\frac{\sqrt{x}-\sqrt{y}}{x-y}$ é ilimitado e a função não é lipschitziana. Porém, como a função é contínua em $[0,1]$, ela é uniformemente contínua, pelo Teorema 1.27.

**Teorema 1.27.** Seja $X \subset \mathbb{R}$ compacto. Toda função contínua $f : X \to \mathbb{R}$ é uniformemente contínua.

*Demonstração.* Seja $f$ contínua. Se $f$ não fosse uniformemente contínua, existiriam $\varepsilon > 0$ e duas sequências $(x_n), (y_n)$ em $X$ satisfazendo $\lim(y_n - x_n) = 0$ e $|f(y_n) - f(x_n)| \geq \varepsilon$ para todo $n \in \mathbb{N}$. Passando a uma subsequência se necessário, podemos supor, em virtude da compacidade de $X$ que $\lim x_n = a$ . Como $y_n = (y_n - x_n) + x_n$ vale também $\lim y_n = a$. Sendo $f$ contínua no ponto $a$, temos $\lim[f(y_n) - f(x_n)] = \lim f(y_n) - \lim f(x_n) = f(a) - f(a) = 0$ contradizendo $|f(y_n) - f(x_n)| \geq \varepsilon$ para todo $n \in \mathbb{N}$. □

A principal referência para a próxima subseção será [11].

### 1.2.4 Derivadas

**Definição 1.21.** A derivada de uma função $f$ em relação à variável $x$ é a função $f'$ cujo valor em $x$ é

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

desde que o limite exista. O domínio de $f'$ é o conjunto de pontos no domínio de $f$ para o qual o limite existe.

**Exemplo 1.6.** Exemplo 1: A derivada de uma função constante é zero. De fato, seja $f$ uma função real tal que $f(x) = k$ para todo $x$, temos $f(x+h) = k$ para todo $x$ e todo $h$, logo, $f'(x) = \lim\limits_{h \to 0} \frac{k-k}{h} = \lim\limits_{h \to 0} 0 = 0$.

**Teorema 1.28.** Se $f : X \to \mathbb{R}$ possui um valor máximo ou mínimo em um ponto $c$ interior de seu domínio e se $f'$ é definida em $c$ , então $f'(c) = 0$.

*Demonstração.* Suponha que $f$ tenha um valor máximo local em $x = c$, de modo que $f(x) - f(c) \leq 0$, para qualquer $x$ próximo de $c$. Então $f'(c)$ é definida por

$$\lim_{x \to c} \frac{f(x) - f(c)}{x - c}$$

o que significa que os limites à direita e à esquerda, existirão quando $x = c$ e serão iguais a $f'(c)$. Examinando separadamente esses limites temos

$$f'(c) = \lim_{x \to c^+} \frac{f(x) - f(c)}{x - c} \leq 0 \text{ , pois } x - c > 0 \text{ e } f(x) \leq f(c).$$

Por outro lado

$$f'(c) = \lim_{x \to c^-} \frac{f(x) - f(c)}{x - c} \geq 0, \text{ pois } x - c < 0 \text{ e } f(x) \leq f(c).$$

Isto nos mostra que $f'(c)$ não pode ser positiva nem negativa, logo $f'(c) = 0$.

O caso em que $f$ tem um valor mínimo local em $x = c$ é provado de modo análogo. $\square$

**Teorema 1.29. Teorema de Rolle:** Suponha que $f : [a,b] \to \mathbb{R}$ é contínua em $[a,b]$ e derivável em todos os pontos de seu interior $(a,b)$. Se $f(a) = f(b)$, então, há pelo menos um número $c$ em $(a,b)$ tal que $f'(c) = 0$.

*Demonstração.* Sendo $f$ contínua em $[a,b]$, pelo Teorema de Weierstrass (Teorema 1.25), $f$ assume mínimo $m$ e máximo $M$ em $[a,b]$. Se o máximo e o mínimo ocorrem nas extremidades $a$ e $b$, então, como $f(a) = f(b)$, $f$ é uma função constante. Assim, $f'(x) = 0$ e o ponto $c$ pode ser tomado em qualquer lugar no interior $(a,b)$. Se o máximo ou o mínimo ocorrem num ponto $c$ entre $a$ e $b$, então, pelo Teorema 1.28 temos $f'(c) = 0$ . $\square$

**Teorema 1.30. Teorema do valor médio:** Se $f : [a,b] \to \mathbb{R}$ é contínua em $[a,b]$ e derivável em $(a.b)$, então há pelo menos um ponto $c$ em $(a,b)$ tal que $\dfrac{f(b)-f(a)}{b-a} = f'(c)$.

*Demonstração.* Consideremos a função $s : [a,b] \to \mathbb{R}$ dada por $s(x) = f(a) + \frac{f(b)-f(a)}{b-a}(x-a)$. Seja $t : [a,b] \to \mathbb{R}$ dada por $t(x) = f(x) - s(x) = f(x) - f(a) - \frac{f(b)-f(a)}{b-a}(x-a)$. Como $t$ é contínua em $[a,b]$, derivável em $(a,b)$ e $t(a) = t(b)$, $t$ é uma função que satisfaz o teorema de Rolle (Teorema 1.29), logo existe $c$ em $(a,b)$ tal que $t'(c) = 0$. Assim $t'(x) = f'(x) - s'(x) \Rightarrow t'(x) = f'(x) - \frac{f(b)-f(a)}{b-a}$, daí $t'(c) = f'(c) - \frac{f(b)-f(a)}{b-a} = 0$ e portanto $f(b) - f(a) = f'(c)(b-a)$. $\square$

**Corolário 3.** Se $f'(x) = 0$ em todos os pontos de um intervalo $(a,b)$, então $f(x) = C$ isto é, funções com derivadas nulas são constantes.

*Demonstração.* Seja $f : [a,b] \to \mathbb{R}$ e $f'(x) = 0$ em todos os pontos do intervalo $(a,b)$. Considere $x_1, x_2$ ambos pertencentes ao intervalo $(a,b)$ tais que $x_1 < x_2$. Observe que $f$ satisfaz a hipótese do teorema do valor médio no intervalo $[x_1,x_2]$. Dessa forma, $\dfrac{f(x_2)-f(x_1)}{x_2-x_1} = f'(c)$ em algum ponto $c$ entre $x_1$ e $x_2$. Como $f' = 0$ ao longo de $(a,b)$, temos $\dfrac{f(x_2)-f(x_1)}{x_2-x_1} = 0$, logo $f(x_1) = f(x_2)$. $\square$

**Corolário 4.** Funções com a mesma derivada diferem por uma constante.

*Demonstração.* Seja $f'(x) = g'(x)$ em cada ponto $x$ de um intervalo $(a,b)$. Em cada ponto $x \in (a,b)$, a derivada da função diferença $h = f - g$ é $h'(x) = f'(x) - g'(x) = 0$. Pelo Corolário 3 temos $h(x) = C$ em $(a,b)$, portanto $f(x) - g(x) = C$. $\square$

### 1.2.5 O conjunto e a função de Cantor

Vamos descrever agora o conjunto de Cantor, um conjunto que possui propriedades interessantes como veremos a seguir. Para a construção do conjunto, considere o intervalo $[0,1]$. Dividimos o intervalo em 3 partes iguais e retiramos o intervalo $(1/3,2/3)$ isto é, retiramos seu terço médio. Na próxima etapa, retiramos o terço médio de cada um dos intervalos restantes, isto é, retiramos os intervalos $(1/9,2/9)$ e $(7/9,8/9)$. Repetimos este processo indefinidamente. O conjunto dos pontos não retirados é o conjunto de Cantor.

Todo extremo de um intervalo omitido pertence ao conjunto de Cantor. O conjunto $E$ dos intervalos omitidos é enumerável, logo o conjunto formado pelos extremos dos intervalos é enumerável. Uma pergunta natural que se pode fazer nesse momento é: “Sobra alguma coisa sem ser os extremos dos intervalos omitidos?". A resposta é , não só sobram mas a maioria dos pontos do conjunto de Cantor não são os extremos dos intervalos omitidos. Para mostrar este fato, mostraremos adiante que o conjunto de Cantor não é enumerável.

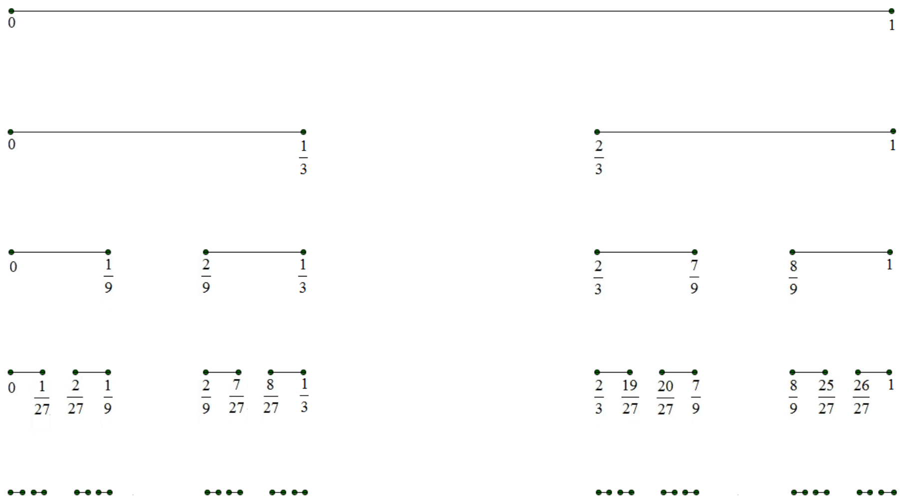



Figura 1: Construindo o conjunto de Cantor.

A seguir listamos as propriedades do conjunto de Cantor.

**1 - É compacto.**

O conjunto de Cantor é um subconjunto do intervalo $[0,1]$, portanto é limitado, e também é fechado pelos teoremas 1.9 e 1.11.

**2 - Tem interior vazio.**

Um conjunto da reta tem interior não vazio se contém intervalos, pois um ponto $a$ é interior ao conjunto $X \subset \mathbb{R}$ quando existe $\varepsilon > 0$ tal que $(a-\varepsilon, a+\varepsilon) \subset X$. Porém o conjunto de Cantor não contém intervalos, pois, na primeira etapa de sua construção sobram dois intervalos de comprimento $\frac{1}{3}$, na segunda etapa de sua construção sobram quatro intervalos de comprimento $\frac{1}{9}$, na terceira intervalos de comprimento $\frac{1}{27}$, na n-ésima etapa intervalos de comprimento $\frac{1}{3^n}$. Portanto, dado qualquer intervalo $J \subset [0,1]$ de comprimento $c > 0$, se tomarmos $n$ tal que $\frac{1}{3^n} < c$, o intervalo $J$ será dividido depois da n-ésima etapa da formação do conjunto de Cantor. Assim o conjunto de Cantor não contém intervalos.

**3 - Não possui pontos isolados.**

Se $x$ é extremo de um intervalo que foi omitido, $x$ sobrevive a todas as etapas seguintes na construção do conjunto de Cantor. Cada ponto extremo desses intervalos é limite de uma sequência de pontos formada pelos extremos dos intervalos omitidos, que por sua vez pertencem ao conjunto de Cantor , ou seja, são pontos de acumulação do conjunto, logo $x$ não é isolado.

Se $x$ pertence ao conjunto de Cantor mas não é extremo de um intervalo omitido na construção do conjunto, então $x$ também não é ponto isolado do conjunto. A justificativa é que $x$ pertence a um intervalo que sobreviveu a n-ésima etapa na construção do conjunto de Cantor. Na etapa seguinte os intervalos que sobram tem comprimento $\frac{1}{3^{n+1}}$ e os extremos destes intervalos pertencem ao conjunto de Cantor, então $x$ é limite de uma sequência de pontos que pertencem ao conjunto de Cantor, logo, $x$ é ponto de acumulação do conjunto e, portanto, não é isolado.

É interessante notar que as somas dos comprimentos dos intervalos removidos converge para 1, que é o mesmo comprimento do intervalo original.

De fato, a soma dos intervalos removidos é:
$\frac{1}{3}+\frac{2}{9}+\frac{4}{27}+...+\frac{2^{n-1}}{3^n}+...=\sum_{n=1}^{\infty}\frac{2^{n-1}}{3^n}$. Esta série geométrica converge para $\frac{\frac{1}{3}}{1-\frac{2}{3}}=1.$

**4 - É não-enumerável.**

Os pontos do conjunto de Cantor têm uma caracterização interessante e útil em termos de sua representação em base 3. Dado $x \in [0,1]$, representar $x$ na base 3 significa escrever $x = 0,x_1x_2x_3...$,onde cada um dos dígitos $x_n$ é igual a $0,1$ ou $2$ , ou seja:

$$x = \sum_{n=1}^{\infty}\frac{x_n}{3^n}.$$

Na primeira etapa da construção do conjunto de Cantor, ao retirar-se o intervalo aberto $(\frac{1}{3},\frac{2}{3})$, ficam excluídos os números $x \in [0,1]$ cuja representação na base 3 tem $x_1$=1, com exceção de $\frac{1}{3}$ que permanece. Na segunda etapa, foram excluídos os números pertencentes aos intervalos $(\frac{1}{9},\frac{2}{9})$ e $(\frac{7}{9},\frac{8}{9})$ ou seja, aqueles da forma $0,01x_3x_4...$ ou da forma $0,21x_3x_4...$ com exceção de $\frac{1}{9}=0,01$ e de $\frac{7}{9}=0,21$, que permanecem. Desta forma, ao construírmos o conjunto de Cantor só restarão os números do intervalo $[0,1]$ cuja representação na base 3 só contém os algarismos 0 e 2, com exceção daqueles que contêm um único algarismo 1 como significativo final, porém estes, podem ser escritos de tal forma que sempre podemos substituir o algarismo final 1 pela sequência $0,02222...=0,1$. Com esta convenção, pode-se afirmar, sem exceções, que os elementos do conjunto de Cantor são os números do intervalo $[0,1]$ cuja representação na base 3 só contém os algarismos 0 e 2. Daí resulta que o conjunto de Cantor é não-enumerável como mostraremos a seguir:

Seja $S$ o conjunto de todas as sequências infinitas, como $s = (02220020202...)$, formadas com os símbolos 0 e 2. Afirmamos que nenhum subconjunto enumerável $X = \{s_1, ..., s_n, ...\}$ $\subset S$ é igual a $S$. De fato, dado $X$, indiquemos com $s_{nm}$ o n-ésimo termo da sequência $s_m \in X$. Formamos uma nova sequência $s^* \in S$ tomando o n-ésimo termo de $s^*$ igual a 0 se for $s_{nn} = 2$, ou igual a 2 se for $s_{nn} = 0$. A sequência $s^*$ não pertence ao conjunto $X$ porque seu n-ésimo termo é diferente do n-ésimo termo de $s_n$, portanto $S$ não é enumerável e concluímos que o conjunto de Cantor não é enumerável.

Vamos construir agora a função de Cantor que será útil mais adiante.

A construção da função de Cantor será feita em etapas.

Etapa 0: Seja $F(x) = 0$ para $x < 0$ e $F(x) = 1$ para $x > 1$. Resta definir $F$ em $[0,1]$.

Etapa 1: Para o intervalo aberto $(\frac{1}{3}, \frac{2}{3})$, referente ao terço central de $[0,1]$, definimos $F(x) = \frac{1}{2}$. Restam sem definição, os valores de $F$ nos intervalos $[0, \frac{1}{3}]$ e $[\frac{2}{3}, 1]$, cujo comprimento total é de $2/3$.

Etapa 2: Definimos $F(x) = \frac{1}{4}$ para $(\frac{1}{9}, \frac{2}{9})$ e $F(x) = \frac{3}{4}$ para $(\frac{7}{9}, \frac{8}{9})$.

Etapa $n+1$: No terço central de cada um dos $2^n$ intervalos restantes após a etapa $n$, seja $F(x)$ igual à média dos valores de $F$, nos dois intervalos vizinhos para os quais $F$ foi definida anteriormente. Restarão $2^{n+1}$ intervalos, de comprimento total $(2/3)^{n+1}$, em que $F$ não está definida. Podemos então definir $F$, por indução para um número enumerável de intervalos abertos, assim, $F$ só não estará definida no complementar desses intervalos, esse complementar é justamente o conjunto de Cantor. Para estender o valor de $F$ para o conjunto de Cantor, denotado por $K$, impomos que $F$ deva ser contínua. Considere $x \in K$ e sejam $a_n$ e $b_n$ os valores de $F$, após a etapa $n$, nos intervalos vizinhos de $x$, à esquerda e à direita, respectivamente. Sabemos por construção, que a diferença entre esses valores é de $(1/2)^n$. Esse limite vai a zero quando $n \to \infty$. Como a função $F$ é monótona não-decrescente no complementar de $K$, os limites de $a_n$ e $b_n$ existirão e serão iguais. Tomamos, então, $F(x) = \lim\limits_{n\to\infty} a_n = \lim\limits_{n\to\infty} b_n$. Logo, a função $F$ tem sua definição concluída em $\mathbb{R}$.

CAPÍTULO

2

# A INTEGRAL DE RIEMANN

Leibniz inventou os símbolos usados nos cálculos diferencial e integral, como os conhecemos hoje. Ele criou o símbolo da integração que era visto por ele como uma soma. O termo integral foi usado pela primeira vez num artigo de um dos irmãos Bernoulli em 1690, e “Cálculo Integral"apareceu como um termo num artigo escrito por Johann Bernoulli juntamente com Leibniz em 1698. A ideia de integral, como área de uma figura plana ou volume, surgiu e alcançou um razoável desenvolvimento com Arquimedes (285-212 a.C), porém naquela época, a matemática era muito geométrica e faltava amadurecimento para o desenvolvimento de um “cálculo integral". No século XVI a simbologia se desenvolveu bastante, sobretudo com François Viéte (1540 - 1603). Depois, com os trabalhos de René Descartes (1596 - 1650), Pierre de Fermat (1601 - 1665) e outros de seus contemporâneos, a moderna notação da Geometria Analítica se difundia e tornava possível o surgimento de métodos sistemáticos e unificados de tratamento do cálculo de áreas e volumes.

A ideia básica de integração é que muitas quantidades podem ser calculadas se são quebradas em pedaços pequenos e, depois, soma-se a contribuição que cada parte dá. Em relação ao cálculo de áreas, se preenchermos interiormente uma região irregular escolhendo retângulos, quanto menor as dimensões das bases dos retângulos escolhidos e consequentemente maior a quantidade desses retângulos no interior da região, melhor será a aproximação por falta para a área da região.

Neste capítulo veremos como é definida a Integral de Riemann, as condições para que uma determinada função seja Riemann integrável, as propriedades da integral e o Teorema

Fundamental do Cálculo que estabelece uma relação surpreendente entre as derivadas e as integrais.

## 2.1 Integral de Riemann

**Definição 2.1.** Uma partição do intervalo $[a,b]$ é um subconjunto finito de pontos $P = \{t_0, t_1, ..., t_n\}$ $\subset [a,b]$ tal que $a \in P$ e $b \in P$, esta notação será utlizada de modo que $a = t_0 < t_1 < ... < t_n = b$. O intervalo $[t_{i-1}, t_i]$ cujo comprimento é $t_i - t_{i-1}$ , será chamado de i-ésimo intervalo da partição $P$.

Sejam $P$ e $Q$ partições do intervalo $[a,b]$. Diz-se que $Q$ refina $P$ quando $P \subset Q$, isto é, todos os pontos de $P$ estão em $Q$.

Dada uma função limitada $f : [a,b] \to \mathbb{R}$, vamos denotar por $m_i$ e $M_i$ o ínfimo e o supremo, respectivamente, de $f$ no i-ésimo intervalo de $P$, isto é,

$$m_i = inf\{f(x); t_{i-1} \leq x \leq t_i\} \text{ e } M_i = sup\{f(x); t_{i-1} \leq x \leq t_i\};$$

e com $\omega_i = M_i - m_i$ a oscilação da função $f$ nesse intervalo. Definimos a soma inferior da função $f$, referente à partição $P$, denotada por $s(f;P)$ como sendo o número

$$s(f;P) = m_1(t_1 - t_0) + ... + m_n(t_n - t_{n-1}) = \textstyle\sum_{i=1}^{n} m_i(t_i - t_{i-1}).$$

Definimos a soma superior de $f$, referente à partição $P$ como sendo o número

$$S(f;P) = M_1(t_1 - t_0) + ... + M_n(t_n - t_{n-1}) = \textstyle\sum_{i=1}^{n} M_i(t_i - t_{i-1}).$$

Sejam $m$ e $M$ o ínfimo e o supremo de $f$, respectivamente no intervalo $[a,b]$. Como $m \leq m_i \leq M_i \leq M$, então

$$m(b-a) \leq s(f;P) \leq S(f;P) \leq M(b-a).$$

Quando $f(x) \geq 0$ para todo $x \in [a,b]$ e contínua neste intervalo, cada soma inferior é um valor aproximado por falta da área da região limitada pelo gráfico de $f$, pelo eixo das abscissas e pelas retas $x = a$ e $x = b$. Analogamente, cada soma superior é um valor aproximado por excesso da mesma área. Quando $f(x) \leq 0$ para todo $x \in [a,b]$ e contínua neste intervalo, essas somas são valores aproximados de tal área, com sinal trocado.

Vemos que o conjunto das somas inferiores é limitado superiormente por $M(b-a)$, de forma que tem supremo finito. Este supremo é chamado de integral inferior da função $f$ sendo denotada por $\underline{\int_a^b} f(x)dx$. Analogamente, o conjunto das somas superiores é limitado inferiormente por $m(b-a)$, logo, tem ínfimo finito, chamado de integral superior de $f$ que será denotado por $\overline{\int_a^b} f(x)dx$.

**Teorema 2.1.** Seja $P$ uma partição qualquer do intervalo $[a,b]$ e $Q$ um refinamento de $P$. Então,

$$s(f;P) \leq s(f;Q) \text{ e } S(f;Q) \leq S(f;P)\ ,$$

isto é, refinando-se uma partição a soma inferior não diminui e a soma superior não aumenta.

*Demonstração.* Suponhamos que $Q = P \cup \{r\}$, isto é, $Q$ contém um só ponto a mais que $P$, digamos com $t_{j-1} < r < t_j$. Sejam $m'$ e $m''$, respectivamente, os ínfimos de $f$ nos intervalos $[t_{j-1},r]$ e $[r,t_j]$. $s(f;Q)$ contém todos os termos de $s(f;P)$ com exceção daquele que corresponde ao $j$-ésimo intervalo que é substituido por $m'(r-t_{j-1})+m''(t_j-r)$. Temos $m_j \leq m', m_j \leq m''$ e $t_j - t_{j-1} = (t_j - r) + (r - t_{j-1})$. Portanto,

$s(f;Q) - s(f;P) = m'(r-t_{j-1}) + m''(t_j - r) - m_j(t_j - t_{j-1}) = (m' - m_j)(r - t_{j-1}) + (m'' - m_j)(t_j - r) \geq 0$, ou seja, $s(f;Q) \geq s(f;P)$.

Analogamente, sejam $M'$ e $M''$ os supremos de $f$ nos intervalos $[t_{j-1},r]$ e $[r,t_j]$. Temos $M_j \geq M'$ e $M_j \geq M''$. Logo,

$S(f;P) - S(f;Q) = M_j(t_j - t_{j-1}) - M'(r - t_{j-1}) - M''(t_j - r) = (M_j - M')(r - t_{j-1}) + (M_j - M'')(t_j - r) \geq 0$ , ou seja , $S(f;P) \geq S(f;Q)$.

O que foi provado faz referência às somas no caso em que $Q$ possui um só ponto a mais que $P$. O caso em que $Q$ resulta de $P$ pelo acréscimo de mais pontos é tratado com o mesmo argumento, aplicado repetidamente, um ponto de cada vez. $\square$

**Corolário 5.** Para quaisquer partições $P$, $Q$ do intervalo $[a,b]$ e qualquer função limitada $f : [a,b] \to \mathbb{R}$ tem-se $s(f;P) \leq S(f;Q)$, ou seja, toda soma inferior é menor ou igual a toda soma superior.

*Demonstração.* A partição $P \cup Q$ refina simultaneamente $P$ e $Q$ uma vez que $P \subset P \cup Q$ e $Q \subset P \cup Q$. Portanto,

$$s(f;P) \leq s(f;P\cup Q) \leq S(f;P\cup Q) \leq S(f;Q).$$

□

**Lema 2.** Sejam $A, B \subset \mathbb{R}$ não vazios, tais que $a \leq b$ para todo $a \in A$ e todo $b \in B$. Então $sup\, A \leq inf\, B$. Temos ainda que $sup\, A = inf\, B$ se, e somente se, para todo $\varepsilon > 0$ dado, existem $a \in A$ e $b \in B$ com $b - a < \varepsilon$.

*Demonstração.* Todo $b \in B$ é cota superior de $A$, logo $sup\, A \leq b$. Isto mostra que $sup\, A$ é cota inferior de $B$, portanto $sup\, A \leq inf\, B$. Se $sup\, A < inf\, B$ então $\varepsilon = inf\, B - sup\, A > 0$ e $b - a \geq \varepsilon$ para quaisquer $a \in A, b \in B$. Reciprocamente, se $sup\, A = inf\, B$ então, para todo $\varepsilon > 0$ dado, $sup\, A - \frac{\varepsilon}{2}$ não é cota superior de $A$ e $inf\, B + \frac{\varepsilon}{2}$ não é cota inferior de $B$, logo existem $a \in A$ e $b \in B$ tais que $sup\, A - \frac{\varepsilon}{2} < a \leq sup\, A = inf\, B \leq b < inf\, B + \frac{\varepsilon}{2}$. Segue-se que $b - a < \varepsilon$. □

**Corolário 6.** Dada $f : [a,b] \to \mathbb{R}$, se $m \leq f(x) \leq M$ para todo $x \in [a,b]$ então

$$m(b-a) \leq \underline{\int_a^b} f(x)dx \leq \overline{\int_a^b} f(x)dx \leq M(b-a).$$

*Demonstração.* Pelo Corolário 5, temos que $s(f;P) \leq S(f;Q)$ quaisquer que sejam as partições $P$ e $Q$. Sejam $A$ o conjunto das somas inferiores de $f$ e $B$ o conjunto das somas superiores de $f$. Pelo Lema 2 $sup\, A \leq inf\, B$ e daí

$$\underline{\int_a^b} f(x)dx \leq \overline{\int_a^b} f(x)dx.$$

Como $m(b-a) \leq s(f;P) \leq S(f;Q) \leq M(b-a)$ o corolário está provado. □

**Definição 2.2.** Diz-se que a função limitada $f : [a,b] \to \mathbb{R}$ é integrável quando sua integral inferior e sua integral superior são iguais. Esse valor comum chama-se integral (de Riemann) de $f$ e é indicado com o símbolo $\int_a^b f$ ou $\int_a^b f(x)dx$.

## 2.2 Condições de Integrabilidade

**Lema 3.** Uma condição necessária e suficiente para que uma função limitada $f$ seja integrável no intervalo $[a,b]$ é que, dado qualquer $\varepsilon > 0$, existam partições $P$ e $Q$ de $[a,b]$ tais que $S(f;P) - s(f;Q) < \varepsilon$.

*Demonstração.* Suponha que a função $f$ seja integrável. Então, dado $\varepsilon > 0$, existem partições $P$ e $Q$, tais que

$$S(f;P) < \int_a^b f + \frac{\varepsilon}{2} \quad \text{e} \quad s(f;Q) > \int_a^b f - \frac{\varepsilon}{2}.$$

Somando membro a membro as desigualdades temos,

$$S(f;P) + \int_a^b f - \frac{\varepsilon}{2} < \int_a^b f + \frac{\varepsilon}{2} + s(f;Q).$$

Portanto, $S(f;P) - s(f;Q) < \varepsilon$. Reciprocamente, considere $A$ o conjunto das somas inferiores de $f$ e $B$ o conjunto das somas superiores de $f$. Se $S(f;P) - s(f;Q) < \varepsilon$, então pelo Lema 2, $sup\, A = inf\, B$, logo

$$\underline{\int_a^b} f = \overline{\int_a^b} f$$

e segue que $f$ é integrável. □

**Teorema 2.2.** Uma condição necessária e suficiente para que uma função limitada $f$ seja integrável no intervalo $[a,b]$ é que, dado qualquer $\varepsilon > 0$, exista uma partição $P$ de $[a,b]$ tal que

$$S(f;P) - s(f;P) = \sum_{i=1}^{n} (M_i - m_i)(t_i - t_{i-1}) < \varepsilon.$$

*Demonstração.* Se $f$ é integrável, então existem partições $P_1$ e $P_2$ tal que $S(f;P_1) - s(f;P_2) < \varepsilon$. Tomemos $P = P_1 \cup P_2$, então $s(f;P_2) \leq s(f;P) \leq S(f;P) \leq S(f;P_1)$ e daí, $S(f;P) - s(f;P) < \varepsilon$. Por outro lado se $S(f;P) - s(f;P) < \varepsilon$ , tomemos $P = Q$ no Lema 3, assim $f$ é integrável. □

**Teorema 2.3.** Toda função contínua $f : [a,b] \to \mathbb{R}$ é integrável.

*Demonstração.* Como $f$ é contínua em $[a,b]$, $f$ é uniformemente contínua. Então, dado qualquer $\varepsilon > 0$, existe $\delta > 0$ tal que $x,y \in [a,b], |y-x| < \delta$ implica que $|f(y) - f(x)| < \dfrac{\varepsilon}{b-a}$. Seja $P = \{t_0, t_1, ..., t_{n-1}, t_n\}$ uma partição de $[a,b]$ cujos intervalos $[t_{i-1}, t_i]$ têm todos comprimento menor do que $\delta$. Em todo intervalo $[t_{i-1},t_i]$ de $P$, existem $x_i, y_i$ tais que $m_i = f(x_i)$ e $M_i = f(y_i)$ onde $m_i$ e $M_i$ são respectivamente o valor mínimo e o valor máximo de $f$ em $[t_{i-1},t_i]$. Temos $0 \leq M_i - m_i = f(y_i) - f(x_i) < \dfrac{\varepsilon}{b-a}$. Em consequência $S(f;P) - s(f;P) = \sum_{i=1}^n (M_i - m_i)(t_i - t_{i-1}) < \dfrac{\varepsilon}{b-a}\sum_{i=1}^n (t_i - t_{i-1}) = \varepsilon$. □

**Teorema 2.4.** Toda função monótona $f : [a,b] \to \mathbb{R}$ é integrável.

*Demonstração.* Seja $f$ não-decrescente e não-constante. Dado $\varepsilon > 0$, seja $P = \{t_0, t_1, \ldots, t_n\}$ uma partição de $[a,b]$ cujos intervalos têm todos comprimento menor que $\dfrac{\varepsilon}{f(b)-f(a)}$. Para cada $i = 1,2,\ldots,n$ temos $M_i - m_i = f(t_i) - f(t_{i-1})$, portanto, $\sum_{i=0}^n f(t_i) - f(t_{i-1}) = f(b) - f(a)$ e

$$\sum_{i=0}^n (f(t_i) - f(t_{i-1}))(t_i - t_{i-1}) < \frac{\varepsilon}{f(b)-f(a)}\sum_{i=0}^n f(t_i) - f(t_{i-1}) = \varepsilon.$$

Logo $f$ é integrável.

Por outro lado, se $f$ for não-crescente e não-constante, dado $\varepsilon > 0$ considere uma partição $Q = \{x_0, x_1, \ldots, x_n\}$ de $[a,b]$ cujos intervalos têm todos comprimento menor que $\dfrac{\varepsilon}{f(a)-f(b)}$, temos $\sum_{i=0}^n M_i - m_i = f(a) - f(b)$ portanto $\sum\limits_{i=0}^n (M_i - m_i)(x_i - x_{i-1}) < \dfrac{\varepsilon}{f(a)-f(b)}\sum\limits_{i=0}^n (M_i - m_i) = \varepsilon$ e $f$ é integrável. □

**Definição 2.3.** Se $a < b$, indicaremos com $|I| = b - a$ o comprimento do intervalo $I$ cujos extremos são $a$ e $b$. Diz-se que o conjunto $X \subset \mathbb{R}$ tem medida nula quando, para todo $\varepsilon > 0$ dado, existe uma cobertura finita ou infinita enumerável $X \subset \bigcup I_k$ de $X$ por intervalos abertos $I_k$ cuja soma dos comprimentos é $\sum |I_k| < \varepsilon$.

Todo conjunto enumerável $X = \{x_1, \ldots, x_k, \ldots\}$ tem medida nula. De fato, dado $\varepsilon > 0$, seja $I_k$ o intervalo aberto de centro $x_k$ e comprimento $\varepsilon/2^{k+1}$. Então $X \subset \bigcup I_k$ e $\sum |I_k| = \varepsilon/2 < \varepsilon$. Um exemplo de conjunto não-enumerável que tem medida nula é o conjunto de Cantor, que foi apresentado no capítulo 1, seção 1.2.5 denotado por $K$. Com efeito, na n-ésima etapa da construção do conjunto, $K$ está contido na união de $2^n$ intervalos de comprimento $1/3^n$, assim, a soma total dos comprimentos é $(\frac{2}{3})^n$. Portanto, para $n$ suficientemente grande, $K \subset \bigcup I_k$ e $\sum |I_k| = (\frac{2}{3})^n < \varepsilon$.

Temos também que, toda união enumerável de conjuntos de medida nula é de medida nula. De fato, seja $A = \cup_{i=1}^\infty A_i$ essa união, onde cada $A_i$ é de medida nula. Dado qualquer $\varepsilon > 0$,

$A_i$ pode ser coberto por uma família enumerável de intervalos abertos $(J_{ij})_{j=1}^{\infty}$ cuja soma dos comprimentos seja menor do que $\varepsilon/2^i$. Então o conjunto $A$ pode ser coberto pela família de intervalos $J_{ij} : i = 1,2,...; j = 1,2,...$, cuja soma dos comprimentos é menor do que $\sum_{i=1}^{\infty} \varepsilon/2^i = \varepsilon$, donde segue o resultado.

**Teorema 2.5.** Se o conjunto $D$ dos pontos de descontinuidade de uma função limitada $f : [a,b] \to \mathbb{R}$ tem medida nula então $f$ é integrável.

*Demonstração.* Dado $\varepsilon > 0$, existem intervalos abertos $I_1,...,I_k,...$ tais que $D \subset \bigcup I_k$ e $\sum |I_k| < \frac{\varepsilon}{2K}$, onde $K = M - m$ é a oscilação de $f$ em $[a,b]$. Para cada $x \in [a,b] \backslash D$, seja $J_x$ um intervalo aberto de centro $x$ tal que a oscilação de $f|_{(J_x \cap [a,b])}$ é menor do que $\frac{\varepsilon}{2(b-a)}$. Pelo teorema de Borel-Lebesgue (Teorema 1.14), a cobertura aberta $[a,b] \subset (\bigcup_k I_k) \cup (\bigcup_x J_x)$ possui uma subcobertura finita $[a,b] \subset I_1 \cup ... \cup I_m \cup J_{x_1} \cup ... \cup J_{x_n}$. Seja $P$ a partição de $[a,b]$ formada pelos pontos $a$ e $b$ e os extremos desses $m+n$ intervalos que pertencem a $[a,b]$. Indiquemos com $[t_{\alpha-1}, t_\alpha]$ os intervalos de $P$ que estão contidos em algum $\bar{I}_k$ e com $[t_{\beta-1}, t_\beta]$ os demais intervalos de $P$, cada um dos quais está contido em algum $J_{x_i}$. Então $\sum (t_\alpha - t_{\alpha-1}) < \varepsilon/2K$ e a oscilação de $f$ em cada intervalo $[t_{\beta-1}, t_\beta]$ é $\omega_\beta < \varepsilon/2(b-a)$. Logo

$$S(f;P) - s(f;P) = \sum \omega_\alpha (t_\alpha - t_{\alpha-1}) + \sum \omega_\beta (t_\beta - t_{\beta-1})$$
$$< \sum K(t_\alpha - t_{\alpha-1}) + \sum \frac{\varepsilon(t_\beta - t_{\beta-1})}{2(b-a)} < \frac{K\varepsilon}{2K} + \frac{\varepsilon(b-a)}{2(b-a)} = \varepsilon.$$

Concluímos que $f$ é integrável. □

**Definição 2.4.** Seja $f : X \to \mathbb{R}$ limitada no conjunto $X \subset \mathbb{R}$. A oscilação $\omega(f;a)$ da função limitada $f : X \to \mathbb{R}$ no ponto $a \in X$ é definida da seguinte forma: para cada $\delta > 0$, seja $\omega(\delta) = M_\delta - m_\delta$, onde $M_\delta$ e $m_\delta$ são respectivamente o *sup* e o *inf* de $f$ em $[a-\delta, a+\delta] \subset X$. A função $\omega(\delta)$ é não-negativa, limitada (pois $f$ é limitada) e não-decrescente. Seja $L$ o ínfimo dos valores de $\omega(\delta)$. Dado $\varepsilon > 0$, existe $\overline{\delta} > 0$ tal que, $\omega(\overline{\delta}) < L + \varepsilon$, com $[a - \overline{\delta}, a + \overline{\delta}] \subset X$. Mas $\omega$ é não-decrescente e daí, $\omega(\delta) \leq \omega(\overline{\delta})$ para $\delta < \overline{\delta}$. Assim, $\delta < \overline{\delta} \Rightarrow L \leq \omega(\delta) \leq \omega(\overline{\delta}) < L + \varepsilon$, logo existe o limite $\omega(f;a) = lim_{\delta \to 0} \omega(\delta)$ que chamaremos de oscilação de $f$ no ponto $a$.

Segue-se que $\omega(f;a) > 0$ se, e somente se, a função $f$ é descontínua no ponto $a$. De fato, suponha que $f$ é descontínua em $a$, então existe $\varepsilon > 0$ tal que, para todo $\delta > 0$; $x \in X, |x-a| < \delta \Rightarrow |f(x) - f(a)| \geq \varepsilon$, mas $|M_\delta - m_\delta| \geq |f(x) - f(a)| \geq \varepsilon$ e daí concluímos que $\omega(f;a) > 0$.

Por outro lado suponha que $f$ é contínua em $a$ isto é, dado $\varepsilon > 0$, existe $\delta > 0$ ; $x \in X$, $|x-a| < \delta \Rightarrow |f(x) - f(a)| < \frac{\varepsilon}{4}$.

Seja $M_\delta = sup(f)|_{X\cap(a-\delta,a+\delta)}$ e $m_\delta = inf(f)|_{X\cap(a-\delta,a+\delta)}$, temos

$M_\delta - \frac{\varepsilon}{4} < f(\overline{x})$ para algum $\overline{x} \in (a-\delta, a+\delta)$ e

$m_\delta + \frac{\varepsilon}{4} > f(\overline{\overline{x}})$ para algum $\overline{\overline{x}} \in (a-\delta, a+\delta)$.

Portanto, $M_\delta - m_\delta - \frac{\varepsilon}{2} < f(\overline{x}) - f(\overline{\overline{x}}) \Rightarrow M_\delta - m_\delta - \frac{\varepsilon}{2} < (f(\overline{x}) - f(a)) + (f(a) - f(\overline{\overline{x}})) < \frac{\varepsilon}{4} + \frac{\varepsilon}{4} \Rightarrow M_\delta - m_\delta - \frac{\varepsilon}{2} < \frac{\varepsilon}{2} \Rightarrow M_\delta - m_\delta < \varepsilon \Rightarrow \omega(f;a) = 0$.

Com esta definição, podemos provar a recíproca do teorema 2.5.

**Teorema 2.6.** O conjunto $D$ dos pontos de descontinuidade da função integrável $f : [a,b] \to \mathbb{R}$ tem medida nula.

*Demonstração.* Para cada $k \in \mathbb{N}$, seja $D_k = \{x \in [a,b]; \omega(f;x) \geq 1/k\}$. Então $D = \cup_{k=1}^{\infty} D_k$. Mostremos agora que cada $D_k$ tem medida nula. Fixemos $k$ e tomemos $\varepsilon > 0$. Sendo $f$ integrável, existe uma partição $P = \{t_0, t_1, ..., t_m\}$ de $[a,b]$ tal que $\sum \omega_i(t_i - t_{i-1}) < \varepsilon/k$, onde $\omega_i$ é a oscilação de $f$ em $(t_{i-1}, t_i)$. Um ponto de $D_k$ pode estar no interior de algum intervalo de $P$, o qual será indicado por $[t_{\alpha-1}, t_\alpha]$ ou pode ser um ponto de $P$, estes últimos formam um conjunto finito $F$. Tem-se $\omega_\alpha \geq 1/k$ para cada $\alpha$ e $D_k \subset [\cup(t_{\alpha-1}, t_\alpha)] \cup F$. Então

$$\frac{1}{k}\sum(t_\alpha - t_{\alpha-1}) \leq \sum \omega_\alpha(t_\alpha - t_{\alpha-1}) \leq \sum \omega_i(t_i - t_{i-1}) < \frac{\varepsilon}{k},$$

logo $\sum(t_\alpha - t_{\alpha-1}) < \varepsilon$. Assim, para todo $\varepsilon > 0$ dado, é possível cobrir $D_k$ com um conjunto finito $F$ mais uma reunião de intervalos cuja soma dos comprimentos é menor que $\varepsilon$. Segue-se que $D_k$ tem medida nula. $\square$

## 2.3 Propriedades da Integral

**Lema 4.** Sejam $A, B \subset \mathbb{R}$ conjuntos limitados e $c \in \mathbb{R}$, tem-se $sup(A+B) = sup\,A + sup\,B, inf(A+B) = inf\,A + inf\,B$. Se $c \geq 0$, então $sup(c.A) = c.sup\,A, inf(c.A) = c.inf\,A$ e caso $c < 0$ então $sup(c.A) = c.inf\,A$ e $inf(c.A) = c.sup\,A$.

*Demonstração.* Considere $a = sup\,A$ e $b = sup\,B$, para todo $x \in A$ e todo $y \in B$ tem-se $x \leq a, y \leq b$, logo $x + y \leq a + b$. Portanto, $a + b$ é cota superior de $A + B = \{x + y; x \in A, y \in B\}$. Dado

$\varepsilon > 0$, existem $x \in A$ e $y \in B$ tais que $a - \varepsilon/2 < x$ e $b - \varepsilon/2 < y$ e daí, $a + b - \varepsilon < x + y$, o que mostra que $a + b$ é a menor cota superior de $A + B$, ou seja, que $sup(A+B) = sup\, A + sup\, B$.

Mostremos que se $c \geq 0$ vale a igualdade $sup(c.A) = c.sup\, A$. De fato, se $c = 0$ então $sup(c.A) = sup(0) = 0.sup\, A$. Se $c > 0$, dado qualquer $x \in A$ tem-se $x \leq a$, logo $cx \leq ca$. Portanto $ca$ é cota superior do conjunto $c.A$. Dado qualquer número $d$ menor do que $ca$, temos $d/c < a$, logo existe $x \in A$ tal que $d/c < x$. Segue-se que $d < cx$, o que mostra que $ca$ é a menor cota superior de $c.A$, ou seja, que $sup(c.A) = c.sup\, A$. Os casos restantes do lema se provam de modo análogo. □

**Corolário 7.** Sejam $f, g : X \to \mathbb{R}$ funções limitadas. Para todo $c \in \mathbb{R}$ são limitadas as funções $f + g, cf : X \to \mathbb{R}$. Tem-se além disso, $sup(f+g) \leq sup(f) + sup(g)$, $inf(f+g) \geq inf(f) + inf(g)$, $sup(cf) = c.sup(f)$, e $inf(cf) = c.inf(f)$ quando $c \geq 0$. Caso $c < 0$, tem-se $sup(cf) = c.inf(f)$ e $inf(cf) = c.sup(f)$.

*Demonstração.* Sejam $A = f(X)$, $B = g(X)$, $C = (f+g)(X) = \{f(x) + g(x); x \in X\}$. Evidentemente $C \subset A + B$, logo $sup(f+g) = sup\, C \leq sup(A+B) = sup\, A + sup\, B = sup(f) + sup(g)$. Além disso, $sup(cf) = sup\{c.f(x); x \in X\} = sup(cA) = c.sup\, A$, quando $c \geq 0$. Os demais casos enunciados no corolário se provam de modo análogo. □

**Teorema 2.7.** Seja $a < c < b$. A função limitada $f : [a,b] \to \mathbb{R}$ é integrável se, e somente se, suas restrições $f|_{[a,c]}$ e $f|_{[c,b]}$ são integráveis. Neste caso, tem-se $\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx$.

*Demonstração.* Seja $A$ o conjunto das somas inferiores de $f$ restrita ao intervalo $[a,c]$ e $B$ o conjunto das somas inferiores de $f$ restrita ao intervalo $[c,b]$. Desta forma $A + B$ é o conjunto das somas inferiores de $f$ relativas as partições de $[a,b]$ que contém o ponto $c$. Daí,

$$\underline{\int_a^b} f(x)dx = sup(A+B) = sup\, A + sup\, B = \underline{\int_a^c} f(x)dx + \underline{\int_c^b} f(x)dx \ \ (I).$$

Analogamente,

$$\overline{\int_a^b} f(x)dx = \overline{\int_a^c} f(x)dx + \overline{\int_c^b} f(x)dx \ \ (II).$$

Fazendo (II) - (I) temos

$$\overline{\int_a^b} f - \underline{\int_a^b} f = \left( \overline{\int_a^c} f - \underline{\int_a^c} f \right) + \left( \overline{\int_c^b} f - \underline{\int_c^b} f \right).$$

As duas parcelas dentro dos parênteses são maiores do que ou iguais a 0, sua soma é zero se, e somente se, elas são ambas nulas. Assim, $f$ é integrável se, e somente se, suas restrições $f|_{[a,c]}$ e $f|_{[c,b]}$ o são. Observando-se (I) ou (II) temos que $\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx$. □

Vamos adotar as seguintes convenções:

(1) $\int_a^a f(x)dx = 0$.

(2) $\int_a^b f(x)dx = -\int_b^a f(x)dx$.

**Teorema 2.8.** Se $f$ e $g$ são funções integráveis em $[a,b]$, o mesmo é verdade de $f+g$ e $\int_a^b (f+g) = \int_a^b f + \int_a^b g$.

*Demonstração.* Dada uma partição arbitrária $P$ de $[a,b]$, sejam $m_i'$, $m_i''$ e $m_i$ respectivamente os ínfimos de $f$, $g$ e $f+g$ no i-ésimo intervalo de $P$, pelo Corolário 7 temos, $m_i' + m_i'' \leq m_i$, logo $s(f;P) + s(g;P) \leq s(f+g;P) \leq \underline{\int}_a^b (f+g)$ para toda partição $P$. Se tomarmos duas partições $P$ e $Q$ teremos ainda

$s(f;P) + s(g;Q) \leq s(f;P\cup Q) + s(g;P\cup Q) \leq s(f+g;P\cup Q) \leq \underline{\int}_a^b (f+g)$ e daí,

$\underline{\int}_a^b f + \underline{\int}_a^b g = sup\ s(f;P) + sup\ s(g;Q) = sup[s(f;P)+s(g;Q)] \leq sup[s(f+g;P\cup Q)] = \underline{\int}_a^b (f+g)$, logo

$$\underline{\int}_a^b f + \underline{\int}_a^b g \leq \underline{\int}_a^b (f+g).$$

Da mesma maneira, se mostra que

$$\overline{\int}_a^b (f+g) \leq \overline{\int}_a^b f + \overline{\int}_a^b g.$$

E pelo Corolário 6

$$\underline{\int}_a^b (f+g) \leq \overline{\int}_a^b (f+g).$$

Temos então as seguintes desigualdades

$$\underline{\int}_a^b f + \underline{\int}_a^b g \leq \underline{\int}_a^b (f+g) \leq \overline{\int}_a^b (f+g) \leq \overline{\int}_a^b f + \overline{\int}_a^b g.$$

Quando $f$ e $g$ são integráveis, as três desigualdades se reduzem à igualdades, o que prova o teorema. $\square$

**Teorema 2.9.** Se $f$ é uma função integrável em $[a,b]$ e $c$ é uma constante qualquer, $cf$ é integrável no mesmo intervalo e $\int_a^b cf = c.\int_a^b f$.

*Demonstração.* Se $c = 0$ então $\int_a^b cf = \int_a^b 0 = 0.\int_a^b f$. Se $c \neq 0$, dado qualquer $\varepsilon > 0$ , seja $P$ uma partição do intervalo $[a,b]$ tal que

$$S(f;P) - s(f;P) < \varepsilon/|c|.$$

Tendo em vista o Corolário 7, sabemos que $S(cf;P) = c.S(f;P)$, $s(cf;P) = c.s(f;P)$ caso $c > 0$ e $S(cf;P) = c.s(f;P)$, $s(cf;P) = c.S(f;P)$ caso $c < 0$. Então,

$S(cf;P) - s(cf;P) = |c|[S(f;P) - s(f;P)] < |c|.\varepsilon/|c| = \varepsilon$, o que mostra que $cf$ é integrável. Se $c > 0$, $S(cf;P) = c.S(f;P)$ e $inf[S(cf;P)] = inf[c.S(f;P)] = c.inf[S(f;P)] = c.\int_a^b f$. Se $c < 0$ , $S(cf;P) = c.s(f;P)$ e $inf[S(cf;P)] = inf[c.s(f;P)] = c.sup[s(f;P)] = c.\int_a^b f$. $\square$

**Teorema 2.10.** Se $f$ é integrável em $[a,b]$, o mesmo é verdade de $f^2$.

*Demonstração.* Suponha inicialmente que $f(x) \geq 0$ para todo $x \in [a,b]$. Dado qualquer $\varepsilon > 0$, como $f$ é integrável, existe uma partição $P = \{t_0,...,t_n\}$ de $[a,b]$ tal que

$$S(f;P) - s(f;P) = \textstyle\sum_{i=1}^n (M_i - m_i)\Delta t_i < \varepsilon.$$

Seja $M$ o supremo de $f$ no intervalo $[a,b]$, de forma que $M_i + m_i \leq 2M$. Então,

$$S(f^2;P) - s(f^2;P) = \textstyle\sum_{i=1}^n (M_i^2 - m_i^2)\Delta t_i =$$

$\sum_{i=1}^n (M_i + m_i)(M_i - m_i)\Delta t_i \leq 2M\sum_{i=1}^n (M_i - m_i)\Delta t_i < 2M\varepsilon$. Logo $f^2$ é integrável pelo Teorema 2.2.

Caso $f$ não seja sempre maior do que ou igual a zero, $m = inf(f) < 0$ e $(f-m)^2$ é integrável, pois $f(x) - m \geq 0$, e é integrável. Como $f^2 = (f-m)^2 + 2mf - m^2$, vemos que $f^2$ é integrável, o que prova o teorema. $\square$

**Teorema 2.11.** Se $f$ e $g$ são funções integráveis em $[a,b]$ , então $fg$ também é integrável no mesmo intervalo.

*Demonstração.* Isso é consequência imediata de

$$fg = \tfrac{1}{2}[(f+g)^2 - f^2 - g^2].$$

□

**Teorema 2.12.** Se $f, g : [a,b] \to \mathbb{R}$ são funções integráveis e $0 < k \leq |g(x)|$ para todo $x \in [a,b]$ então o quociente $f/g$ é integrável.

*Demonstração.* Escrevendo $f/g$ como $f.(1/g)$, basta provar que $1/g$ é integrável. Sejam $x_i$ e $x_i'$ respectivamente as oscilações de $g$ e $1/g$ no i-ésimo intervalo de uma partição $P$. Dado $\varepsilon > 0$, tomemos uma partição $P$ de modo que $\sum x_i(t_i - t_{i-1}) < \varepsilon.k^2$. Para quaisquer $x, y$ no i-ésimo intervalo de $P$ tem-se

$$\left|\frac{1}{g(y)} - \frac{1}{g(x)}\right| = \frac{|g(x) - g(y)|}{|g(y).g(x)|} \leq \frac{x_i}{k^2},$$

portanto $x_i' \leq x_i/k^2$ e daí, $\sum x_i'(t_i - t_{i-1}) \leq \frac{1}{k^2}\sum x_i(t_i - t_{i-1}) < \frac{1}{k^2}.\varepsilon.k^2 = \varepsilon$, o que mostra que $1/g$ é integrável. □

**Teorema 2.13.** Se $f, g : [a,b] \to \mathbb{R}$ são funções integráveis e $f(x) \leq g(x)$ para todo $x \in [a,b]$, então $\int_a^b f(x)dx \leq \int_a^b g(x)dx$.

*Demonstração.* Se $f(x) \leq g(x)$ para todo $x \in [a,b]$ então $s(f;P) \leq s(g;P)$ e $S(f;P) \leq S(g;P)$ para toda partição $P$, logo $\int_a^b f(x)dx \leq \int_a^b g(x)dx$. □

**Teorema 2.14.** Se $f : [a,b] \to \mathbb{R}$ é integrável, então , $|f|$ é integrável e $|\int_a^b f(x)dx| \leq \int_a^b |f(x)|dx$.

*Demonstração.* A integrabilidade de $|f|$ segue da desigualdade $||f(x)| - |f(y)|| \leq |f(x) - f(y)|$, que é válida para todo $x, y \in [a,b]$, assim a oscilação de $|f|$ em qualquer subintervalo de uma partição $P$ de $[a,b]$ será menor ou igual à oscilação de $f$ no mesmo subintervalo. Assim, sendo $M_i' - m_i'$ e $M_i - m_i$, respectivamente, a oscilação de $|f|$ e $f$ no i-ésimo intervalo de $P$ temos:

$S(|f|;P) - s(|f|;P) = \sum(M_i' - m_i')(t_i - t_{i-1}) \leq \sum(M_i - m_i)(t_i - t_{i-1}) = S(f;P) - s(f;P) <$ $\varepsilon$ e pelo Teorema 2.2 $|f|$ é integrável.

Além disso, como $-|f(x)| \leq f(x) \leq |f(x)|$ para todo $x \in [a,b]$ tem-se

$$-\int_a^b |f(x)|dx \leq \int_a^b f(x)dx \leq \int_a^b |f(x)|dx$$

e concluímos que, $\left|\int_a^b f(x)dx\right| \leq \int_a^b |f(x)|dx.$ $\square$

## 2.4 O Teorema Fundamental do Cálculo

O Teorema Fundamental do Cálculo estabelece a conexão entre derivada e integral.

**Teorema 2.15. Teorema Fundamental do Cálculo:** Seja $f : I \to \mathbb{R}$ contínua no intervalo $I$. As seguintes afirmações a respeito de uma função $F : I \to \mathbb{R}$ são equivalentes:

(1) $F$ é uma integral indefinida de $f$, isto é, existe $a \in I$ tal que $F(x) = F(a) + \int_a^x f(t)dt$, para todo $x \in I$.

(2) $F$ é uma primitiva de $f$, isto é, $F'(x) = f(x)$ para todo $x \in I$.

*Demonstração.* (1) $\Rightarrow$ (2). Se $x_0, x_0 + h \in I$ então $F(x_0 + h) - F(x_0) = \int_{x_0}^{x_0+h} f(t)dt$ e $h.f(x_0) = \int_{x_0}^{x_0+h} f(x_0)dt$, portanto

$$\frac{F(x_0 + h) - F(x_0)}{h} - f(x_0) = \frac{1}{h}\int_{x_0}^{x_0+h} [f(t) - f(x_0)]dt.$$

Dado $\varepsilon > 0$, pela continuidade de $f$ no ponto $x_0$, existe $\delta > 0$ tal que se $t \in I$ e $|t - x_0| < \delta$ temos $|f(t) - f(x_0)| < \varepsilon$. Então $0 < |h| < \delta$, $x_0 + h \in I$ implicam

$$\left|\frac{F(x_0 + h) - F(x_0)}{h} - f(x_0)\right| \leq \frac{1}{|h|}\left|\int_{x_0}^{x_0+h} |f(t) - f(x_0)|dt\right| < \frac{1}{|h|}.|h|.\varepsilon = \varepsilon.$$

Isto mostra que $F'(x_0) = f(x_0)$.

(2) $\Rightarrow$ (1). Seja $F' = f$. Como acabamos de ver na demonstração (1) $\Rightarrow$ (2), se $\varphi(x) = \int_a^x f(t)dt$, para $a \in I$ fixado, teremos $\varphi' = f$. As duas funções $F, \varphi : I \to \mathbb{R}$, tendo a mesma derivada, diferem por uma constante ( Capítulo 1, Corolário 4). Como $\varphi(a) = 0$, essa constante é $F(a)$. Portanto $F(x) = F(a) + \varphi(x)$, isto é, $F(x) = F(a) + \int_a^x f(t)dt$ para todo $x \in I$. $\square$

Como $F(x) = F(a) + \int_a^x f(t)dt$ e $F' = f$, se fizermos $F(b)$ teremos $F(b) = F(a) + \int_a^b F'(t)dt$, o que reduz o cálculo da integral $\int_a^b f(x)dx$ à procura de uma primitiva de $f$, isto é $\int_a^b f(x)dx = F(b) - F(a)$.

## 2.5 A integral como limite de somas de Riemann

**Definição 2.5.** Definimos norma de uma partição $P = \{t_0, ..., t_n\} \subset [a,b]$ como sendo o número $|P| = max\{t_i - t_{i-1}, i = 1, 2, ..., n\}$.

**Teorema 2.16.** Seja $f : [a,b] \to \mathbb{R}$ limitada. Para todo $\varepsilon > 0$ dado, existe $\delta > 0$ tal que $|P| < \delta$ implica $\lim\limits_{|P|\to 0} S(f;P) = \overline{\int_a^b} f(x)dx$[1].

*Demonstração.* Devemos provar que $\forall \varepsilon > 0, \exists\, \delta > 0; |P| < \delta \Rightarrow |S(f;P) - \overline{\int_a^b} f(x)dx| < \varepsilon$, que equivale a $-\varepsilon + \overline{\int_a^b} f(x)dx < S(f;P) < \overline{\int_a^b} f(x)dx + \varepsilon$.

A desigualdade da esquerda é evidente pois $\overline{\int_a^b} f(x)dx$ é o ínfimo das somas superiores. Suponha inicialmente que $f(x) \geq 0$ em $[a,b]$. Dado $\varepsilon > 0$, existe uma partição $P_0 = \{t_0, ..., t_n\}$ de $[a,b]$ tal que

$$S(f;P_0) < \overline{\int_a^b} f(x)dx + \frac{\varepsilon}{2}.$$

Tome $\delta$ tal que $0 < \delta < \dfrac{\varepsilon}{2Mn}$ onde $M = supf$. Seja $P = \{r_0, ..., r_n\}$ uma partição qualquer de $[a,b]$, com $|P| < \delta$. Vamos denotar por $[r_{\alpha-1}, r_\alpha]$ os intervalos de $P$ que estão contidos em algum $[t_{i-1}, t_i]$ de $P_0$ e $[r_{\beta-1}, r_\beta]$ os demais intervalos de $P$. Observe que, há no máximo $n$ intervalos do tipo $[r_{\beta-1}, r_\beta]$ uma vez que cada um destes intervalos deve conter pelo menos um $t_i$. Vamos escrever $\alpha \subset i$ para significar $[r_{\alpha-1}, r_\alpha] \subset [t_{i-1}, t_i]$, e neste caso valem as desigualdades $0 \leq M_\alpha \leq M_i$ e $0 \leq \sum_{\alpha\subset i}(r_\alpha - r_{\alpha-1}) \leq \sum t_i - t_{i-1}$. Logo $\sum_{\alpha\subset i} M_\alpha(r_\alpha - r_{\alpha-1}) \leq \sum M_i(t_i - t_{i-1})$ e $M_\beta(r_\beta - r_{\beta-1}) \leq M.\delta$, e daí:

$$S(f;P) = \sum M_\alpha(r_\alpha - r_{\alpha-1}) + \sum M_\beta(r_\beta - r_{\beta-1}) \leq \sum M_i(t_i - t_{i-1}) + M\sum(r_\beta - r_{\beta-1})$$
$$< S(f;P_0) + Mn\frac{\varepsilon}{2Mn} \Rightarrow S(f;P) < \overline{\int_a^b} f(x)dx + \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \overline{\int_a^b} f(x)dx + \varepsilon.$$

No caso geral, como $f$ é limitada, existe uma constante $c$ tal que $f(x) + c \geq 0$ para todo $x \in [a,b]$. Tomando $g(x) = f(x) + c$ temos $S(g;P) = S(f;P) + c(b-a)$ e $\overline{\int_a^b} g(x)dx = \overline{\int_a^b} f(x)dx + c(b-a)$ e recaímos no caso anterior. $\square$

[1] $\lim\limits_{|P|\to 0}$ considera o limite quando $n \to \infty$ de sucessivos refinamentos da partição.

**Teorema 2.17.** Seja $f : [a,b] \to \mathbb{R}$ limitada. Para todo $\varepsilon > 0$ dado, existe $\delta > 0$ tal que $|P| < \delta \Rightarrow \lim\limits_{|P| \to 0} s(f;P) = \underline{\int}_a^b f(x)dx.$

*Demonstração.* Devemos provar que $\forall \varepsilon > 0, \exists\, \delta > 0; |P| < \delta \Rightarrow \left| s(f;P) - \underline{\int}_a^b f(x)dx \right| < \varepsilon$, que equivale a $-\varepsilon + \underline{\int}_a^b f(x)dx < s(f;P) < \underline{\int}_a^b f(x)dx + \varepsilon$.

A desigualdade da direita é evidente uma vez que $\underline{\int}_a^b f(x)dx$ é o supremo das somas inferiores.

Suponha inicialmente que $f(x) \leq 0$ em $[a,b]$. Dado $\varepsilon > 0$, existe uma partição $P_0 = \{t_0, ..., t_n\}$ de $[a,b]$ tal que

$$s(f;P_0) > \underline{\int}_a^b f(x)dx - \frac{\varepsilon}{2}.$$

Tome $\delta$ tal que $0 < \delta$ e $\delta > \dfrac{-\varepsilon}{2mn}$ onde $m = inf(f)$. Seja $P = \{r_0, ..., r_n\}$ uma partição qualquer de $[a,b]$, com $|P| < \delta$. Vamos denotar por $[r_{\alpha-1}, r_\alpha]$ os intervalos de $P$ que estão contidos em algum $[t_{i-1}, t_i]$ de $P_0$ e $[r_{\beta-1}, r_\beta]$ os demais intervalos de $P$. Observe que, há no máximo $n$ intervalos do tipo $[r_{\beta-1}, r_\beta]$ uma vez que cada um destes intervalos deve conter pelo menos um $t_i$. Vamos escrever $\alpha \subset i$ para significar $[r_{\alpha-1}, r_\alpha] \subset [t_{i-1}, t_i]$, e neste caso valem as desigualdades, $m_\alpha \geq m_i$, $\sum_{\alpha \subset i}(r_\alpha - r_{\alpha-1}) \leq \sum t_i - t_{i-1}$ e $m_\beta(r_\beta - r_{\beta-1}) \geq m_\beta \delta \geq m\delta$. Logo $\sum_{\alpha \subset i} m_\alpha(r_\alpha - r_{\alpha-1}) \geq \sum m_i(t_i - t_{i-1})$ e $m_\beta(r_\beta - r_{\beta-1}) \geq m\delta$. Daí:

$$\begin{aligned} s(f;P) &= \sum m_\alpha(r_\alpha - r_{\alpha-1}) + \sum m_\beta(r_\beta - r_{\beta-1}) \geq \sum m_i(t_i - t_{i-1}) + mn\delta \\ &> s(f;P_0) + mn\frac{-\varepsilon}{2mn} > \underline{\int}_a^b f(x)dx - \frac{\varepsilon}{2} - \frac{\varepsilon}{2} = \underline{\int}_a^b f(x)dx - \varepsilon. \end{aligned}$$

No caso geral, como $f$ é limitada, existe uma constante $c$ tal que $f(x) - c \leq 0$ para todo $x \in [a,b]$. Tomando $g(x) = f(x) - c$ temos $s(g;P) = s(f;P) - c(b-a)$ e $\underline{\int}_a^b g(x)dx = \underline{\int}_a^b f(x)dx - c(b-a)$ e recaímos no caso anterior. $\square$

Considere uma partição qualquer $P = \{t_0, ..., t_n\}$ do intervalo $[a,b]$, escolhemos um número $\xi_i$ de forma que $t_{i-1} \leq \xi_i \leq t_i$ para cada $i = 1, 2, ..., n$. O par $P^* = (P, \xi_i)$ é chamado de partição pontilhada do intervalo $[a,b]$.

Dada uma função limitada $f : [a,b] \to \mathbb{R}$ e uma partição pontilhada $P^*$ de $[a,b]$, tem-se a soma de Riemann

$$\sum(f;P^*) = \sum_{i=1}^{n} f(\xi_i)(t_i - t_{i-1}).$$

Seja qual for o modo de pontilhar a partição $P$, tem-se

$$s(f;P) \leq \sum(f;P^*) \leq S(f;P).$$

**Teorema 2.18.** Se $f : [a,b] \to \mathbb{R}$ é integrável então

$$\int_a^b f(x)dx = \lim_{|P|\to 0} \sum(f;P^*).$$

*Demonstração.* Segue-se dos Teoremas 2.16 e 2.17 que se $f$ é integrável então

$$\lim_{|P|\to 0} s(f;P) = \lim_{|P|\to 0} S(f;P) = \int_a^b f(x)dx.$$

Como $s(f;P) \leq \sum(f;P^*) \leq S(f;P)$, resulta do teorema do confronto (Teorema 1.16) que $\lim\limits_{|P|\to 0} \sum(f;P^*) = \int_a^b f(x)dx$. $\square$

Vamos encerrar este capítulo com a definição de um tipo de integral imprópria, isto é, integrais de funções definidas em intervalos ilimitados.

**Definição 2.6.** Dada $f : [a,+\infty) \to \mathbb{R}$, definimos $\int_a^{+\infty} f(x)dx$ como $\int_a^{+\infty} f(x)dx = \lim\limits_{k\to+\infty} \int_a^k f(x)dx$. Se o limite existir, dizemos que a integral é convergente. Se $f : (-\infty,b] \to \mathbb{R}$, então, $\int_{-\infty}^{b} f(x)dx = \lim\limits_{j\to-\infty} \int_j^b f(x)dx$. Para $f : (-\infty,+\infty) \to \mathbb{R}$ tomamos um ponto $a \in \mathbb{R}$ e definimos $\int_{-\infty}^{+\infty} f(x)dx = \int_{-\infty}^{a} f(x)dx + \int_a^{+\infty} f(x)dx$, a integral é convergente se ambas as integrais do 2º membro forem convergentes.

CAPÍTULO

3

# A INTEGRAL DE RIEMANN-STIELTJES

Neste capítulo será apresentada a integral de Riemann-Stieltjes, formulada por Thomas Joannes Stieltjes (1856-1894). Essa integral é uma generalização da integral de Riemann apresentada no capítulo 2. Veremos a importância desta integral no estudo do valor esperado de uma variável aleatória, no sentido em que podemos unificar as expressões do valor esperado para diversos tipos de variáveis. Veremos um resultado importante a respeito do cálculo do valor esperado de uma variável aleatória que nos possibilita obter o valor esperado como a diferença de duas integrais de Riemann, esse resultado decorre diretamente de resultados válidos para a integral de Riemann-Stieltjes. Na seção 3.1 apresentamos alguns conceitos básicos da teoria de probabilidade como experimentos aleatórios e espaço amostral, apresentamos o conceito de variável aleatória e sua classificação em discreta, contínua, singular ou mista. Na seção 3.2 apresentamos o conceito de valor esperado de uma variável aleatória. Grande parte dos exemplos apresentados nessas seções são encontrados em [6] e [9]. A integral de Riemann-Stieltjes, a existência da integral, bem como as suas propriedades são apresentadas respectivamente nas seções 3.3, 3.4 e 3.5. A definição de valor esperado utilizando a integral de Riemann-Stieltjes será feita na seção 3.6. Encerramos o capítulo com o estudo de sequências de funções apresentada na seção 3.7 como uma motivação para a integral de Lebesgue.

## 3.1 Probabilidade e Variáveis Aleatórias

**Definição 3.1.** Chamamos de experimentos aleatórios aqueles que, repetidos em idênticas condições, podem produzir resultados diferentes.

**Exemplo 3.1.** (a) Lançar uma moeda e observar a face de cima.

(b) Lançar um dado e observar o número da face de cima.

(c) Lançar duas moedas e observar a sequência de cara e coroa obtida.

**Definição 3.2.** Chamamos de espaço amostral, e indicamos por $\Omega$, o conjunto formado por todos os resultados possíveis de um experimento aleatório. Cada resultado possível é denominado evento elementar de $\Omega$ e denotado genericamente por $\omega$.

(a) Lançar uma moeda e observar a face de cima.

$\Omega = \{C, C_0\}$ onde $C$ representa cara e $C_0$ coroa.

(b) Lançar um dado e observar o número da face de cima.

$\Omega = \{1,2,3,4,5,6\}$.

(c) Lançar duas moedas e observar a sequência de cara e coroa obtida.

$\Omega = \{(C,C),(C,C_0),(C_0,C),(C_0,C_0)\}$.

**Definição 3.3.** Uma classe de subconjuntos de $\Omega$, representada por $\mathcal{F}$, é denominada uma $\sigma$-álgebra se satisfaz as seguintes propriedades:

**(i)** $\Omega \in \mathcal{F}$;

**(ii)** Se $A \in \mathcal{F}$, então $A^c \in \mathcal{F}$;

**(iii)** Se $A_i \in \mathcal{F}$, $i \geq 1$, então $\displaystyle\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$.

**Exemplo 3.2.** Considere $\Omega = \{1,2,3\}$ e as seguintes coleções de subconjuntos:

$\mathcal{F}_1 = \{\emptyset, \Omega, \{1\}, \{2,3\}\}$;

$\mathcal{F}_2 = \{\emptyset, \Omega, \{1\}, \{2\}, \{1,3\}, \{2,3\}\}$.

Verifica-se que, $\mathcal{F}_1$ é uma $\sigma$-álgebra e $\mathcal{F}_2$ não é uma $\sigma$-álgebra.

**Definição 3.4.** Seja $\Omega$ um espaço amostral. Uma função $P$ definida na $\sigma$-álgebra $\mathcal{F}$ de subconjuntos de $\Omega$ e com valores em $\mathbb{R}$, é uma probabilidade se:

**(i)** $0 \leq P(A) \leq 1$ para todo subconjunto $A \in \mathcal{F}$;

**(ii)** $P(\Omega) = 1$;

**(iii)** Para $A_i \in \mathcal{F}$, $i \geq 1$, mutuamente exclusivos[1], temos $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$.

A trinca formada por $(\Omega, \mathcal{F}, P)$ é um espaço de probabilidade.

**Definição 3.5.** Seja $(\Omega, \mathcal{F}, P)$ um espaço de probabilidade. A função $X : \Omega \to \mathbb{R}$ tal que $X^{-1}(I) = \{\omega \in \Omega : X(\omega) \in I\} \in \mathcal{F}$, para todo intervalo $I \subset \mathbb{R}$, denomina-se variável aleatória.

**Exemplo 3.3.** Suponha que um experimento consista em lançar 3 moedas honestas para cima e observar a sequência de cara e coroa obtida. Para a $\sigma$-álgebra, consideremos o conjunto das partes de $\Omega$ que será denotado por $\Omega_p$. Se $X$ for o número de caras que aparecem, então $X$ é uma variável aleatória que pode assumir um dos valores $0, 1, 2$ ou $3$.

**Definição 3.6.** Seja $X$ uma variável aleatória em $(\Omega, \mathcal{F}, P)$. A função de distribuição de $X$, $F : \mathbb{R} \to \mathbb{R}$, é definida por

$$F(x) = P(X \leq x) = P(\{\omega \in \Omega : X(\omega) \leq x\}), \forall x \in \mathbb{R}.$$

Tal função possui as seguintes propriedades:

**(1)** $0 \leq F(x) \leq 1$;

**(2)** $\lim_{x \to -\infty} F(x) = 0$;

**(3)** $\lim_{x \to +\infty} F(x) = 1$;

**(4)** $F$ é contínua à direita. Isto significa que se considerarmos os eventos $[X \leq x_n] = \{\omega \in \Omega : X(\omega) \leq x_n\}$ e $x \in \mathbb{R}$, $\{x_n : n \geq 1, n \in \mathbb{N}\}$ uma sequência tal que os $x_n's$ se aproximam de $x$ pela direita ou, em outras palavras, por valores maiores que $x$, então, $[X \leq x_n]$ tende a $[X \leq x]$ e, assim, para todo $x$, $P(X \leq x_n)$ tende a $P(X \leq x)$;

**(5)** é não-decrescente.

É importante mencionar que, conhecendo-se a função de distribuição podemos obter qualquer informação sobre a variável aleatória, como veremos mais adiante.

**Definição 3.7.** Uma variável aleatória é classificada como discreta, se assume somente uma quantidade enumerável de valores (finito ou infinito).

[1] Dois conjuntos são mutuamente exclusivos se sua interseção é vazia.

**Exemplo 3.4.** Três bolas são sorteadas de uma urna contendo 3 bolas brancas, 3 bolas vermelhas e 5 bolas pretas. Suponha que ganhamos $R\$1,00$ por cada bola branca sorteada, perdemos $R\$1,00$ para cada bola vermelha sorteada e não perdemos nem ganhamos se uma bola preta for retirada. Para a $\sigma$-álgebra, consideremos $\Omega_p$. Se $X$ representa a quantia final após o experimento, então $X$ é uma variável aleatória discreta que pode assumir os valores $-3,-2,-1,0,1,2,3$.

A função de probabilidade de uma variável aleatória discreta é uma função que atribui probabilidade a cada um dos possíveis valores assumidos pela variável. Isto é, sendo $X$ uma variável com valores $x_1, x_2, \ldots$, temos para $i = 1,2,\ldots$,

$$p(x_i) = P(X = x_i) = P(\{\omega \in \Omega : X(\omega) = x_i\}).$$

A função de probabilidade de $X$ possui as seguintes propriedades:

**(i)** $0 \leq p(x_i) \leq 1, \forall i = 1,2,\ldots;$

**(ii)** $\sum\limits_i p(x_i) = 1$; com a soma percorrendo todos os possíveis valores (eventualmente infinitos).

Considerando o exemplo 3.3, temos

$P(X = 0) = P\{(C_0,C_0,C_0)\} = \frac{1}{8}$.

$P(X = 1) = P\{(C_0,C_0,C),(C_0,C,C_0),(C,C_0,C_0)\} = \frac{3}{8}$.

$P(X = 2) = P\{(C_0,C,C),(C,C_0,C),(C,C,C_0)\} = \frac{3}{8}$.

$P(X = 3) = P\{(C,C,C)\} = \frac{1}{8}$.

Da função de probabilidade obtemos a função de distribuição e vice-versa. De fato, dada a função de probabilidade temos

$$F(x) = \sum_{i \in A_x} p(x_i), \text{ com } A_x = \{i : x_i \leq x\} \text{ para } A_x \neq \emptyset. \text{ Se } A_x = \emptyset, F(x) = 0.$$

Por outro lado, dada a função de distribuição temos

$$p(x_i) = F(x_i) - F(x_i^-)$$

em que $F(x_i^-)$ é o limite de $F$ tendendo a $x_i$ pela esquerda.

Para o exemplo 3.3, a função de probabilidade será dada por:

| $X$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $p(x_i)$ | $\frac{1}{8}$ | $\frac{3}{8}$ | $\frac{3}{8}$ | $\frac{1}{8}$ |

A função de distribuição correspondente será:

$$F(x) = \begin{cases} 0, & se \quad x < 0; \\ \frac{1}{8}, & se \quad 0 \leq x < 1; \\ \frac{1}{8} + \frac{3}{8} = \frac{4}{8}, & se \quad 1 \leq x < 2; \\ \frac{4}{8} + \frac{3}{8} = \frac{7}{8}, & se \quad 2 \leq x < 3; \\ \frac{7}{8} + \frac{1}{8} = 1, & se \quad x \geq 3. \end{cases}$$

Para as variáveis discretas, a função de distribuição tem a forma de escada sendo descontínua nos valores assumidos pela variável, e o tamanho do salto é a probabilidade da váriável assumir aquele determinado valor.

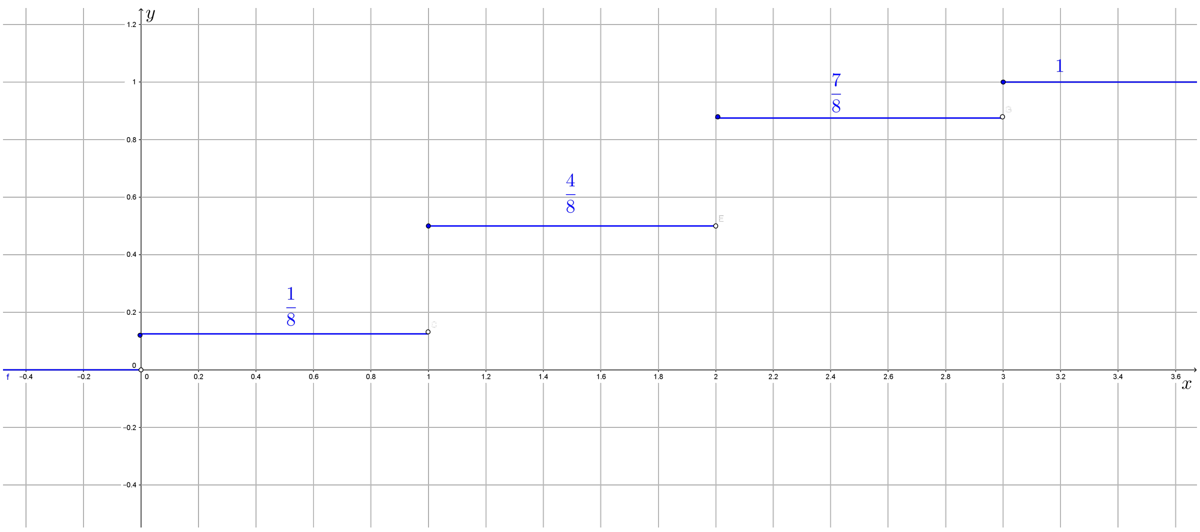



Figura 2: Função de Distribuição para Número de Caras.

**Definição 3.8.** Dizemos que uma função $F : [a,b] \to \mathbb{R}$ é absolutamente contínua se dado $\varepsilon > 0$, existe $\delta > 0$ tal que para qualquer coleção (finita ou não) de subintervalos disjuntos $[a_i, b_i]$ temos

$$\sum_i (b_i - a_i) < \delta \Rightarrow \sum_i |F(b_i) - F(a_i)| < \varepsilon.$$

Se $F$ é absolutamente contínua, então $F$ é a integral indefinida de sua derivada.

**Definição 3.9.** Uma variável aleatória $X$, com função de distribuição $F$, será classificada como contínua (ou absolutamente contínua), se existir uma função não negativa $f$ tal que:

$$F(x) = \int_{-\infty}^{x} f(t)dt, \text{ para todo } x \in \mathbb{R}.$$

A função $f$ é chamada de função densidade de probabilidade da variável aleatória $X$.

A função densidade $f$ da variável aleatória $X$ satisfaz:

(1) $f(x) \geq 0, \forall x \in \mathbb{R}$;

(2) $\displaystyle\int_{-\infty}^{+\infty} f(x)dx = 1$.

Note que, a partir da definição, temos uma relação entre a função de distribuição e a função densidade. Dada a função densidade, a função de distribuição segue por integração. Por outro lado, derivando a função de distribuição, obtemos a densidade.

Assim, classificamos a variável $X$ como contínua, se sua função de distribuição $F$ é absolutamente contínua.

Para obter a probabilidade de a variável estar num certo intervalo $[a,b]$, fazemos a integral da função densidade de probabilidade neste intervalo. Assim,

$$P(a \leq X \leq b) = \int_{a}^{b} f(x)dx.$$

Dessa forma, para as variáveis contínuas, a probabilidade da variável ser igual a um particular valor é zero.

**Exemplo 3.5.** Suponha que, a quantidade de tempo em horas que um computador funciona sem estragar é uma variável aleatória contínua com função densidade de probabilidade dada por:

$$f(x) = \begin{cases} \lambda e^{-x/100}, & se \quad x \geq 0; \\ 0, & se \quad x < 0. \end{cases}$$

Calculemos a probabilidade de que o computador funcione entre 50 e 150 horas.

Como $1 = \displaystyle\int_{-\infty}^{+\infty} f(x)dx = \lambda \lim_{k \to +\infty} \int_{0}^{k} e^{-x/100}dx \Rightarrow 1 = 100\lambda$ e daí, $\lambda = \frac{1}{100}$.

Portanto, a probabilidade de que o computador funcione entre 50 e 150 horas antes de estragar é dada por

$$P\{50 < X < 150\} = \int_{50}^{150} \frac{1}{100} e^{-x/100}dx = -e^{-3/2} + e^{-1/2} \approx 0,384.$$

Se estivermos interessados em calcular a probabilidade de o computador funcionar menos de 100 horas fazemos

$$P(X < 100) = \int_0^{100} \frac{1}{100} e^{-x/100} dx = 1 - e^{-1} \approx 0,633.$$

Como mencionamos anteriormente, dada a função densidade, a função distribuição segue por integração, assim, pelo exemplo, temos a seguinte função densidade:

$$f(t) = \begin{cases} \frac{1}{100} e^{-t/100}, & se \quad t \geq 0; \\ 0, & se \quad t < 0. \end{cases}$$

Como $F(x) = \int_{-\infty}^{x} f(t)dt$ então, para $x < 0$, $F(x) = 0$, pois a função densidade é nula neste intervalo. Para $x \geq 0$, temos $F(x) = \int_0^x \frac{1}{100} e^{-t/100} dt$, logo, $F(x) = 1 - e^{-x/100}$.

Portanto

$$F(x) = \begin{cases} 1 - e^{-x/100}, & se \quad x \geq 0; \\ 0, & se \quad x < 0. \end{cases}$$

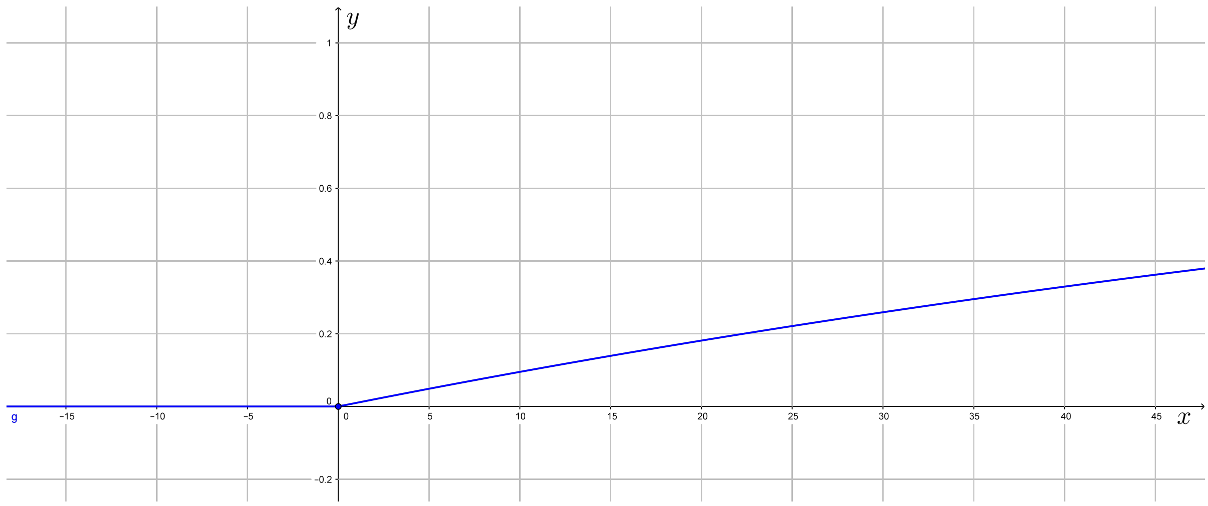



Figura 3: Função de Distribuição para o Exemplo 3.5.

Resta mencionar que uma variável aleatória pode ser classificada como singular, se sua função de distribuição é contínua, mas sua derivada é zero em quase todos os pontos, isto é, esta propriedade só não é válida num conjunto de pontos que tem probabilidade zero ou ainda, num conjunto de medida nula. Um exemplo de função de distribuição singular é a função de Cantor que envolve o conjunto de Cantor (que será denotado por $K$) já apresentado em 1.2.5.

Seja $X$ uma variável aleatória com função de distribuição de Cantor. Então $X$ não é discreta, pois a função de Cantor é contínua em $\mathbb{R}$. Entretanto, a variável aleatória $X$ não é contínua, pois $F'(x)$ é zero em $K^c$, uma vez que, em cada etapa, foi sempre definida como constante nos intervalos respectivos. Logo a variável $X$ será classificada como singular.

Dizemos que uma variável aleatória é mista se tem partes em diferentes classificações. O mais comum é a mistura da parte contínua com a discreta.

A função de distribuição de qualquer variável sempre pode ser escrita como a ponderação de uma função de distribuição discreta, uma absolutamente contínua e uma singular. Isto é,

$$F(x) = \alpha_d F^d(x) + \alpha_{ac} F^{ac}(x) + \alpha_s F^s(x);$$

com $\alpha_d + \alpha_{ac} + \alpha_s = 1$, $\alpha_d$, $\alpha_{ac}$, $\alpha_s \geq 0$ e $F^d$, $F^{ac}$ e $F^s$ funções de distribuição do tipo discreta, absolutamente contínua e singular, respectivamente.

**Exemplo 3.6.** Suponha que uma variável aleatória $X$ tenha função de distribuição dada por

$$F(x) = \begin{cases} 0, & se \quad x < 0; \\ \frac{x}{4}, & se \quad 0 \leq x < 1; \\ \frac{1}{2}, & se \quad 1 \leq x < 2; \\ 1, & se \quad x \geq 2. \end{cases}$$

Observe através do gráfico de $F(x)$ que a variável aleatória $X$ é do tipo mista. Ela é contínua nos intervalos $(-\infty, 1)$, $(1, 2)$ e $(2, +\infty)$ e os pontos de descontinuidade em 1 e 2 indicam os valores de sua parte discreta.

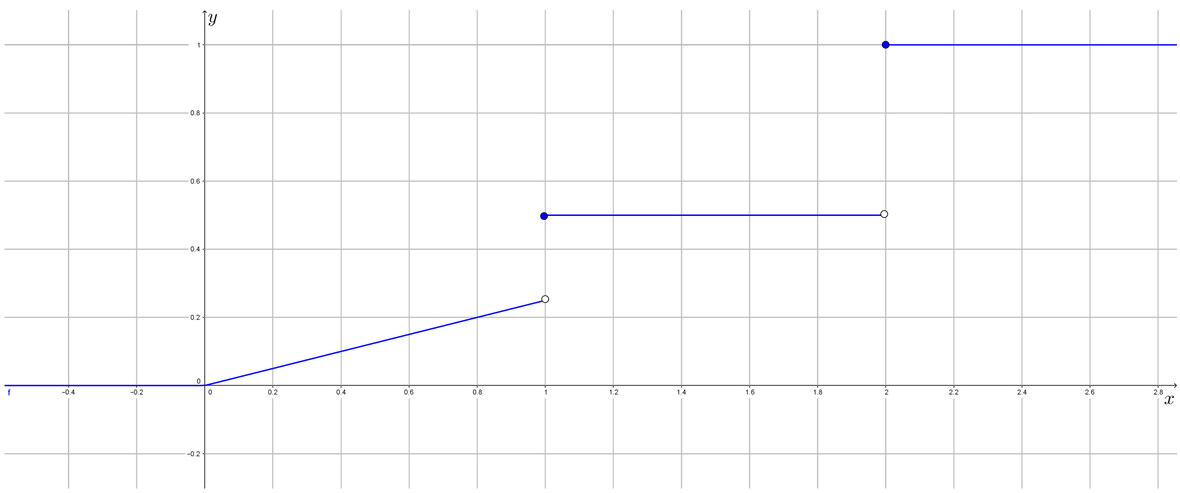



Figura 4: Função de Distribuição Mista.

Existe um grande número de modelos que surgiram a partir do estudo de problemas práticos. Alguns modelos para o caso discreto e contínuo podem ser encontrados em [6] e [9].

## 3.2 Valor Esperado de uma Variável Aleatória

Sendo $X$ uma variável aleatória, o valor esperado $E(X)$ de $X$ desempenha um papel importante na teoria das probabilidades. Historicamente, o valor esperado ou média parece ter sido desenvolvido para avaliar ganhos em jogos com apostas em dinheiro. A questão de interesse era avaliar esse retorno num horizonte de várias jogadas. Seria uma espécie de balanço final de muitas jogadas, após contabilizar perdas e ganhos. Nesta seção, usaremos os nomes valor esperado, esperança e média de forma indistinta.

**Definição 3.10.** Seja $X$ uma variável aleatória discreta com função de probabilidade $p(x)$ e valor $x_i$ para $i$ num certo conjunto de índices $I$. O valor esperado ou esperança ou média de $X$ é definido por

$$E(X) = \sum_{i \in I} x_i p(x_i),$$

desde que a soma esteja determinada.

**Exemplo 3.7.** Um jogador lança 3 moedas não viciadas. Ganha R\$ $6,00$ se 3 caras ocorrerem, ganha R\$ $3,00$ se duas caras ocorrerem, ganha R\$ $1,00$ se somente uma cara ocorrer. Por outro lado, perde R\$ $10,00$ se 3 coroas ocorrerem. Calculemos o valor esperado para este jogo.

Vamos escrever $C$ para cara e $C_0$ para coroa. O espaço amostral do jogo é:

$$\Omega = \{CCC, CCC_0, CC_0C, C_0CC, CC_0C_0, C_0CC_0, C_0C_0C, C_0C_0C_0\}$$

| $x_i$ | $P(X = x_i)$ | $x_i p(x_i)$ |
|---|---|---|
| $6,00$ | $\frac{1}{8}$ | $\frac{6}{8}$ |
| $3,00$ | $\frac{3}{8}$ | $\frac{9}{8}$ |
| $1,00$ | $\frac{3}{8}$ | $\frac{3}{8}$ |
| $-10,00$ | $\frac{1}{8}$ | $-\frac{10}{8}$ |

$$E(X) = \sum x_i p(x_i) = \frac{6}{8} + \frac{9}{8} + \frac{3}{8} - \frac{10}{8} = 1.$$

Portanto, o valor esperado para este jogo é de R\$ $1,00$. Este valor deve ser interpretado como uma média, ou seja, jogando-se este jogo uma grande quantidade de vezes, o jogador ganha em média R\$$1,00$.

**Definição 3.11.** Seja $X$ uma variável aleatória contínua, com função densidade de probabilidade $f$. Definimos o valor esperado $E(X)$ de $X$ por

$$E(X) = \int_{-\infty}^{+\infty} xf(x)dx,$$

desde que a integral esteja bem definida.

**Exemplo 3.8.** Suponha que o tempo de duração de um determinado tipo de bateria seja uma variável aleatória $X$, contínua, com função densidade de probabilidade dada por

$$f(x) = \begin{cases} \frac{1}{3}e^{-x/3}, & se \quad x \geq 0; \\ 0, & se \quad x < 0. \end{cases}$$

sendo o tempo medido em anos. Vamos calcular, quanto tempo em média deve durar essa bateria.

$$E(X) = \int_{-\infty}^{+\infty} xf(x)dx = \frac{1}{3}\int_{0}^{+\infty} xe^{-x/3}dx = \frac{1}{3}\lim_{k\to+\infty}\int_{0}^{k} xe^{-x/3}dx$$
$$= \lim_{k\to+\infty}[-ke^{-k/3} - 3e^{-k/3}] + 3 = \lim_{k\to+\infty}\frac{-k-3}{e^{k/3}} + 3 = 0 + 3 = 3.$$

Portanto a bateria deve durar em média 3 anos.

**Exemplo 3.9.** Seja $X$ uma variável aleatória do tipo mista com função de distribuição dada por

$$F(x) = \begin{cases} 0, & se \quad x < -2; \\ \frac{1}{24}(x^2+4x+4), & se \quad -2 \leq x < 0; \\ \frac{1}{6}(x+2), & se \quad 0 \leq x < 1; \\ \frac{1}{6}(x+3), & se \quad 1 \leq x < 2; \\ \frac{1}{24}(-x^2+8x+8), & se \quad 2 \leq x < 4; \\ 1, & se \quad x \geq 4. \end{cases}$$

Para obter $E(X)$, vamos calcular separadamente o seu valor para a parte discreta e depois para a parte contínua.

Para a parte discreta temos:

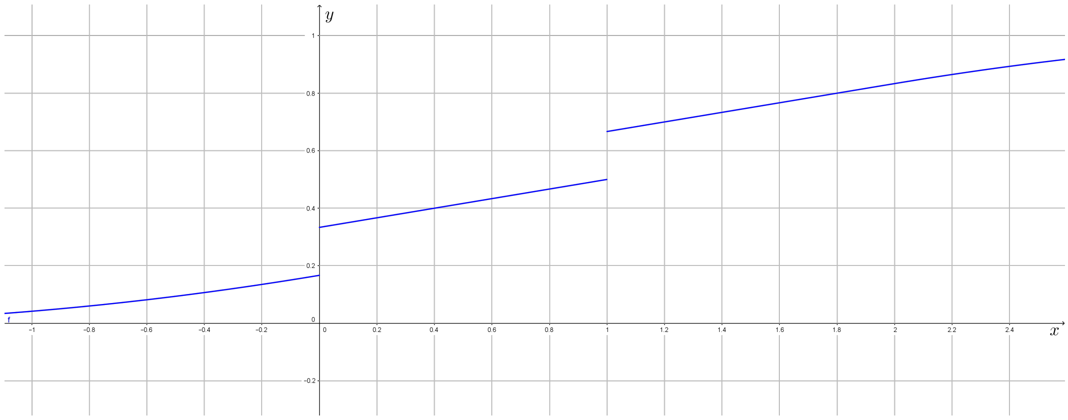



Figura 5: Função de Distribuição Mista.

$$E_1(X) = \sum x_i p(x_i) = 0.\frac{1}{6} + 1.\frac{1}{6} = \frac{1}{6}.$$

Note que, para a parte discreta, $P(X = 0) = \frac{1}{6}$, $P(X = 1) = \frac{1}{6}$, logo $\alpha_d = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$. Para a parte contínua obtemos $P(X < 0) = F(0^-) = \frac{1}{6}$, $P(0 < X < 1) = F(1^-) - F(0) = \frac{1}{2} - \frac{1}{3} = \frac{1}{6}$ e $P(X > 1) = 1 - F(1) = \frac{1}{3}$, logo $\alpha_{ac} = \frac{1}{6} + \frac{1}{6} + \frac{1}{3} = \frac{2}{3}$. Dessa forma, não existe parte singular e $\alpha_s = 0$.

Calculemos a função de distribuição associada à parte contínua denotada por $F^{ac}$. Para isso, derivamos $F(x)$ e integramos o resultado, multiplicando-o por $1/\alpha_{ac}$. Assim,

$$\frac{dF(x)}{dx} = \begin{cases} 0, & se \quad x < -2; \\ \frac{1}{12}(x+2), & se \quad -2 \leq x < 0; \\ \frac{1}{6}, & se \quad 0 \leq x < 2; \\ \frac{1}{12}(-x+4), & se \quad 2 \leq x < 4; \\ 0, & se \quad x \geq 4. \end{cases}$$

Dessa forma, para $x < -2$,

$$F^{ac}(x) = \frac{1}{\alpha_{ac}} \int_{-\infty}^{x} \frac{dF(x)}{dx} dx = \frac{3}{2} \int_{-\infty}^{-2} 0 dx = 0.$$

Para $-2 \leq x < 0$,

$$F^{ac}(x) = \frac{1}{\alpha_{ac}} \int_{-\infty}^{x} \frac{dF(x)}{dx} dx = \frac{3}{2} \int_{-\infty}^{-2} 0 dx + \frac{3}{2} \int_{-2}^{x} \frac{1}{12}(x+2) dx = \frac{1}{16}(x^2 + 4x + 4).$$

Para $0 \leq x < 2$,

$$F^{ac}(x) = \frac{1}{\alpha_{ac}} \int_{-\infty}^{x} \frac{dF(x)}{dx} dx = \frac{3}{2} \int_{-\infty}^{-2} 0dx + \frac{3}{2} \int_{-2}^{0} \frac{1}{12}(x+2)dx + \frac{3}{2} \int_{0}^{x} \frac{1}{6} dx = \frac{1}{4}(x+1).$$

Para $2 \leq x < 4$,

$$F^{ac}(x) = \frac{1}{\alpha_{ac}} \int_{-\infty}^{x} \frac{dF(x)}{dx} dx = \frac{3}{2} \int_{-\infty}^{-2} 0dx + \frac{3}{2} \int_{-2}^{0} \frac{1}{12}(x+2)dx + \frac{3}{2} \int_{0}^{2} \frac{1}{6} dx$$

$$+ \frac{3}{2} \int_{2}^{x} \frac{1}{12}(-x+4)dx = \frac{1}{16}(-x^2+8x).$$

Para $x \geq 4$,

$$F^{ac}(x) = \frac{1}{\alpha_{ac}} \int_{-\infty}^{x} \frac{dF(x)}{dx} dx = 1.$$

Logo

$$F^{ac} = \begin{cases} 0, & se \quad x < -2; \\ \frac{1}{16}(x^2+4x+4), & se \quad -2 \leq x < 0; \\ \frac{1}{4}(x+1), & se \quad 0 \leq x < 2; \\ \frac{1}{16}(-x^2+8x), & se \quad 2 \leq x < 4; \\ 1, & se \quad x \geq 4. \end{cases}$$

Calculamos a densidade derivando $F^{ac}$ e obtemos:

$$f^{ac}(x) = \begin{cases} 0, & se \quad x < -2 \;\; ou \;\; x \geq 4; \\ \frac{1}{8}(x+2), & se \quad -2 \leq x < 0; \\ \frac{1}{4}, & se \quad 0 \leq x < 2; \\ \frac{1}{8}(-x+4), & se \quad 2 \leq x < 4. \end{cases}$$

Finalmente calculamos a esperança para a parte contínua:

$$E_2(X) = \int_{-\infty}^{\infty} \alpha_{ac} x f^{ac}(x)dx = \int_{-\infty}^{-2} \frac{2}{3} x.0dx + \int_{-2}^{0} \frac{2}{3}\frac{x}{8}(x+2)dx + \int_{0}^{2} \frac{2}{3}\frac{x}{4}dx$$
$$+ \int_{2}^{4} \frac{2}{3}\frac{x}{8}(-x+4)dx + \int_{4}^{\infty} \frac{2}{3} 0dx = \frac{2}{3}.$$

Concluímos que

$$E(X) = E_1(X) + E_2(X) = \frac{1}{6} + \frac{2}{3} = \frac{5}{6}.$$

Nas próximas seções, faremos a construção da integral de Riemann-Stieltjes, vamos analisar as condições de existência e descreveremos suas propriedades. Mais adiante, veremos como esta integral nos auxilia no estudo de variáveis aleatórias.

## 3.3 Somas de Riemann-Stieltjes

Sejam $f:[a,b]\to\mathbb{R}$ limitada e $\phi:[a,b]\to\mathbb{R}$ limitada e não-decrescente. Sendo $P=\{x_0,x_1,\ldots,x_n\}$ com $a=x_0<x_1<\ldots<x_n=b$ uma partição do intervalo $[a,b]$ e $C=\{c_1,c_2,\ldots,c_n\}$ uma lista de $n$ números escolhidos de forma que $x_{k-1}\leq c_k\leq x_k$ para $k=1,2,\ldots,n$. Considere a soma

$$f(c_1)[\phi(x_1)-\phi(x_0)]+f(c_2)[\phi(x_2)-\phi(x_1)]+\ldots+f(c_n)[\phi(x_n)-\phi(x_{n-1})]=$$

$$\sum_{k=1}^{n} f(c_k)[\phi(x_k)-\phi(x_{k-1})]=\sum_{P} f.\Delta\phi.$$

Tal soma é chamada de soma de Riemann-Stieltjes.

Obviamente, se $\phi:[a,b]\to\mathbb{R}$ for dada por $\phi(x)=x$, temos a soma de Riemann. Logo, as somas de Riemann são um caso particular das somas de Riemann-Stieltjes.

**Definição 3.12.** Seja $P$ uma partição do intervalo $[a,b]$, $f:[a,b]\to\mathbb{R}$ limitada, $\phi:[a,b]\to\mathbb{R}$ limitada e não-decrescente e $L$ um número real. Dizemos que $L$ é o limite de $\sum\limits_{P} f.\Delta\phi$ quando $|P|\to 0$ e escrevemos

$$\lim_{|P|\to 0}\sum_{P} f.\Delta\phi=L$$

quando para todo $\varepsilon>0$ dado, pode-se obter $\delta>0$ tal que $\left|\sum\limits_{P} f.\Delta\phi-L\right|<\varepsilon$ para toda partição $P$ com $|P|<\delta$ . Neste caso, dizemos que $f$ é Riemann-Stieltjes integrável com respeito a $\phi$ em $[a,b]$ e escrevemos

$$\int_a^b f(x)d\phi(x)=L.$$

## 3.4 Existência da integral de Riemann-Stieltjes

Vamos considerar, inicialmente, algumas situações particulares:

**a)** Suponha $f$ constante em $[a,b]$ com $f(x)=k$. Para toda partição $P$ de $[a,b]$ temos

$\sum_P f.\Delta\phi = k[\phi(x_1)-\phi(x_0)]+k[\phi(x_2)-\phi(x_1)]+...+k[\phi(x_n)-\phi(x_{n-1})] = k[\phi(x_n)-\phi(x_0)] = k[\phi(b)-\phi(a)]$, logo $\int_a^b f(x)d\phi(x) = k[\phi(b)-\phi(a)]$.

**b)** Suponha $\phi$ constante em $[a,b]$ com $\phi(x)=c$. Para toda partição $P$ de $[a,b]$ tem-se $\Delta\phi = [\phi(x_k)-\phi(x_{k-1})] = 0$ para $k=1,...,n$, logo $\sum_P f.\Delta\phi = 0$ e daí $\int_a^b f(x)d\phi(x) = 0$.

Agora, vamos analisar situações onde temos descontinuidades, na $\phi$ ou simultaneamente em $f$ e $\phi$.

**1ºcaso:** Seja $f$ contínua em $[a,b]$, $P=\{x_0,...,x_n\}$ uma partição do intervalo $[a,b]$ e

$$\phi(x) = \begin{cases} \beta_1, & se \quad a \leq x < \overline{x}; \\ \beta, & se \quad x = \overline{x}; \\ \beta_2, & se \quad \overline{x} < x \leq b. \end{cases}$$

Observe que $\phi(x)$ possui uma descontinuidade do "tipo salto"em $\overline{x}$.

**a)** O que acontece quando $\overline{x}$ é um ponto da partição com $\overline{x}=x_k$ para $0<k<n$ ?

Neste caso, $x_{k-1} \leq c_k < \overline{x} = x_k \leq c_{k+1} \leq x_{k+1}$. Assim,

$$\sum_P f.\Delta\phi = f(c_k)(\beta-\beta_1)+f(c_{k+1})(\beta_2-\beta) =$$
$$[f(c_k)-f(\overline{x})](\beta-\beta_1)+[f(c_{k+1})-f(\overline{x})](\beta_2-\beta)+f(\overline{x})(\beta_2-\beta_1).$$

Fazendo $|P|\to 0$, a continuidade de $f$ faz com que as duas primeiras parcelas se aproximem de 0, logo

$$\int_a^b f(x)d\phi(x) = f(\overline{x})[\phi(\overline{x}^+)-\phi(\overline{x}^-)], \text{ onde } \phi(\overline{x}^+) = \lim_{x\to\overline{x}^+}\phi(x) \text{ e } \phi(\overline{x}^-) = \lim_{x\to\overline{x}^-}\phi(x).$$

**b)** O que acontece quando $\overline{x}$ não é um ponto da partição? Neste caso $x_{k-1} < \overline{x} < x_k$ e $x_{k-1} \leq c_k \leq x_k$. Assim

$\sum_P f.\Delta\phi = f(c_k)(\beta_2-\beta_1) = [f(c_k)-f(\overline{x})](\beta_2-\beta_1)+f(\overline{x})(\beta_2-\beta_1)$ e pela continuidade de $f$, quando $|P|\to 0$ resulta:

$$\int_a^b f(x)d\phi(x) = f(\overline{x})[\phi(\overline{x}^+) - \phi(\overline{x}^-)].$$

Nos casos em que $\overline{x} = a$ ou $\overline{x} = b$ a integral de Riemann-Stieltjes tem o mesmo valor.

De modo geral, se $f$ é contínua em $[a,b]$ e $\phi$ é uma função com saltos em $\overline{x_1}, \overline{x_2}, ..., \overline{x_n}$ em $[a,b]$, a integral de Riemann-Stieltjes de $f$ com respeito a $\phi$ existe e

$$\int_a^b f(x)d\phi(x) = \sum_{k=1}^{n} f(\overline{x_k})[\phi(\overline{x_k}^+) - \phi(\overline{x_k}^-)],$$

com $\phi(a^+) - \phi(a^-) = \phi(a^+) - \phi(a)$ e $\phi(b^+) - \phi(b^-) = \phi(b) - \phi(b^-)$.

**Exemplo 3.10.** Suponha que $f(x) = x^2$ e $\phi(x) = [x]$ (função maior inteiro), com $0 \leq x \leq n$. Temos:

$$\int_0^n f(x)d\phi(x) = 1^2 + 2^2 + 3^2 + ... + n^2 = \frac{n(n+1)(2n+1)}{6}.$$

**2º caso:** O que acontece quando $f$ e $\phi$ possuem pontos de descontinuidade em comum?

Sejam

$$f(x) = \begin{cases} 1, & se \quad x \neq \frac{1}{2}; \\ 2, & se \quad x = \frac{1}{2}. \end{cases} \qquad e \qquad \phi(x) = \begin{cases} 1, & se \quad 0 \leq x < \frac{1}{2}; \\ 3, & se \quad \frac{1}{2} \leq x \leq 1. \end{cases}$$

Ambas as funções são descontínuas em $x = \frac{1}{2}$. Suponha que $P$ é uma partição de $[0,1]$ e $\frac{1}{2}$ é um ponto da partição, com $x_k = \frac{1}{2}$. A soma de Riemann-Stieltjes referente a esta partição é:

$$\sum_P f.\Delta\phi = f(c_k)\left[\phi\left(\frac{1}{2}\right) - \phi(x_{k-1})\right] + f(c_{k+1})\left[\phi(x_{k+1}) - \phi\left(\frac{1}{2}\right)\right]$$

$$= f(c_k)[3-1] + f(c_{k+1})[3-3] = 2f(c_k).$$

Se $\frac{1}{2}$ não é ponto da partição o resultado é o mesmo.

Temos problemas uma vez que $f(c_k) = 1$ se $c_k \neq \frac{1}{2}$ ou $f(c_k) = 2$ se $c_k = \frac{1}{2}$. Deste modo a integral de Riemann-Stieltjes não existe. Assim no que segue, assumimos que $f$ e $\phi$ são funções reais limitadas em $[a,b]$ com nenhum ponto de descontinuidade em comum.

**Teorema 3.1.** Suponha que $f : [a,b] \to \mathbb{R}$ e $\phi : [a,b] \to \mathbb{R}$ são funções limitadas sem pontos comuns de descontinuidade no intervalo $[a,b]$ e que a integral de Riemann-Stieltjes de $f$ com relação a $\phi$ existe. Então a integral de Riemann-Stieltjes de $\phi$ com relação a $f$ existe e

$$\int_a^b \phi(x)df(x) = f(b)\phi(b) - f(a)\phi(a) - \int_a^b f(x)d\phi(x).$$

*Demonstração.* Seja $P = \{x_0, x_1, ..., x_n\}$ uma partição de $[a,b]$ tal que $x_k - x_{k-1} < \delta/2$ para $k = 1,2,...,n$. Para obtermos a soma de Riemann-Stieltjes, escolhemos pontos $c_k \in [a,b]$ tais que $x_{k-1} \leq c_k \leq x_k$. Fazendo $a = x_0 = c_0$ e $b = x_n = c_{n+1}$, uma partição $P$ de $[a,b]$ pode ser dada por $P = \{c_0, c_1, ..., c_n, c_{n+1}\}$ com $c_k \leq x_k \leq c_{k+1}$ e $c_{k+1} - c_k < \delta$, $k = 0,...,n$. Portanto,

$$\left|\sum_P \phi\Delta f - \left(f(b)\phi(b) - f(a)\phi(a) - \int_a^b f(x)d\phi(x)\right)\right|$$

$$= \left|\sum_{k=1}^n \phi(c_k)[f(x_k) - f(x_{k-1})] - \left(f(b)\phi(b) - f(a)\phi(a) - \int_a^b f(x)d\phi(x)\right)\right|$$
$$= \left|\sum_{k=1}^n \phi(c_k)[f(x_k) - f(x_{k-1})] - [f(x_n)\phi(c_{n+1}) - f(x_0)\phi(c_0)] + \int_a^b f(x)d\phi(x)\right|$$
$$= \left|\sum_{k=0}^n \phi(c_k)f(x_k) - \sum_{k=1}^{n+1} \phi(c_k)f(x_{k-1}) + \int_a^b f(x)d\phi(x)\right|$$

$$= \left|\int_a^b f(x)d\phi(x) - \sum_{k=0}^n f(x_k)[\phi(c_{k+1}) - \phi(c_k)]\right| < \varepsilon,$$

pois $\sum_{k=0}^n f(x_k)[\phi(c_{k+1}) - \phi(c_k)]$ está associada a uma partição $P = \{c_0, c_1, ..., c_n, c_{n+1}\}$ de $[a,b]$ cuja norma é menor do que $\delta$ e por hipótese a integral de Riemann-Stieltjes de $f$ com relação a $\phi$ existe, isto é, dado $\varepsilon > 0$, existe $\delta > 0$ tal que $\left|\sum_P f\Delta\phi - \int_a^b f d\phi(x)\right| < \varepsilon$ para toda partição de norma menor do que $\delta$ e isto prova o teorema. □

## Somas de Darboux-Stieltjes

Assim como definimos a integral de Riemann utilizando as noções de soma inferior e soma superior, podemos fazer o mesmo para as integrais de Riemann-Stieltjes. Seja $f : [a,b] \to \mathbb{R}$ limitada, $\phi(x) : [a,b] \to \mathbb{R}$ uma função não-decrescente e $P = \{x_0, ..., x_n\}$ uma partição do intervalo $[a,b]$.

A soma inferior de $f$ com respeito a $\phi$ e $P$ é o número

$s(f,\phi;P) = \sum_{i=1}^{n} m_i(f;P)[\phi(x_i) - \phi(x_{i-1})]$, onde $m_i(f;P) = inf\{f(x); x_{i-1} \leq x \leq x_i\}$.

A soma superior de $f$ com respeito a $\phi$ e $P$ é o número

$S(f,\phi;P) = \sum_{i=1}^{n} M_i(f;P)[\phi(x_i) - \phi(x_{i-1})]$, onde $M_i(f;P) = sup\{f(x); x_{i-1} \leq x \leq x_i\}$.

**Lema 5.** Suponha que $P$ e $Q$ são duas partições do intervalo $[a,b]$, $f$ é uma função limitada e $\phi$ é uma função não-decrescente em $[a,b]$. Tem-se

**(a)** $s(f,\phi;P) \leq S \leq S(f,\phi;P)$ para toda soma de Riemann-Stieltjes $S$ de $f$ com respeito a $\phi$ e $P$,

**(b)** $s(f,\phi;P) \leq s(f,\phi;Q)$ e $S(f,\phi;Q) \leq S(f,\phi;P)$ caso $Q$ seja um refinamento de $P$,

**(c)** $s(f,\phi;P) \leq S(f,\phi;Q)$,

**(d)** $sup\ s(f,\phi;P) \leq inf\ S(f,\phi;P)$, onde o ínfimo e o supremo são tomados em relação a todas as partições de $[a,b]$.

A prova deste lema pode ser encontrada em [7].

## A definição de Integral de Riemann-Stieltjes a partir das somas de Darboux-Stieltjes

Considere um intervalo $[a,b]$, uma função não-decrescente $\phi[a,b] \to \mathbb{R}$, a função limitada $f(x)$ definida em $[a,b]$ e suponha que há um número, por exemplo I, que pode ser aproximado arbitrariamente pelas somas inferiores e superiores de Darboux-Stieltjes de $f$ com respeito a $\phi$. Em seguida, pelo Lema 5 (d), temos que

$$sup\ s(f,\phi;P) = I = inf\ S(f,\phi;P),$$

onde o ínfimo e o supremo são tomados em relação a todas as partições $P$ de $[a,b]$. Isto significa que para cada número positivo $\varepsilon$ haverá partições $P$ e $Q$ de $[a,b]$ tal que $0 \leq I - s(f,\phi;P) < \varepsilon$ e $0 \leq S(f,\phi;Q) - I < \varepsilon$, e o Lema 5 (a) e (b) implica que

$$s(f,\phi;P) \leq s(f,\phi;R) \leq I \leq S(f,\phi;R) \leq S(f,\phi;Q),$$

para qualquer refinamento $R$ de $P$ e $Q$ e $|I - S| < \varepsilon$ para toda soma de Riemann-Stieltjes $S$ de $f$ com respeito a $\phi$ e $R$. Portanto, este número $I$ parece ter todas as propriedades de aproximação necessárias para se ter a integral de Riemann-Stieltjes de $f$ em relação a $\phi$ sobre $[a,b]$ e isso motiva a seguinte definição:

**Definição 3.13.** Seja $\phi(x)$ uma função não-decrescente definida no intervalo fechado $[a,b]$. A função limitada $f$ definida em $[a,b]$ é Riemann-Stieltjes integrável em $[a,b]$ com respeito a $\phi$ se $sup\ s(f,\phi;P) = inf\ S(f,\phi;P)$ onde o ínfimo e o supremo são tomados em relação a todas as partições $P$ de $[a,b]$, e neste caso o valor comum do supremo e do ínfimo é chamado de integral de Riemann-Stieltjes de $f$ em $[a,b]$ com respeito a $\phi$ e é indicada por $\int_a^b f(x)d\phi(x)$.

**Teorema 3.2.** Sejam $f$ limitada e $\phi$ uma função não-decrescente, ambas definidas no intervalo $[a,b]$, as seguintes afirmações são equivalentes:

**(a)** Dado $\varepsilon > 0$, existe uma partição $P$ de $[a,b]$ tal que

$$0 \leq S(f,\phi;P) - s(f,\phi;P) < \varepsilon.$$

**(b)** $f$ é Riemann-Stieltjes integrável.

**(c)** Existe um número I com a propriedade de que dado $\varepsilon > 0$, existe uma partição $P$ de $[a,b]$ tal que $|I - S| < \varepsilon$, para toda soma de Riemann-Stieltjes $S$ de $f$ com respeito a $\phi$ e $P$.

Além disso, se estas condições são satisfeitas, o número $I$ em (c) é necessariamente igual a $\int_a^b f(x)d\phi(x)$.

A demonstração deste teorema pode ser encontrada em [7].

O próximo teorema mostra que, em certas condições podemos substituir uma integral de Riemann-Stieltjes por uma integral de Riemann. A demonstração deste teorema foi baseado em [10].

**Teorema 3.3.** Seja $f : [a,b] \to \mathbb{R}$ limitada em $[a,b]$. Considere $\phi$ monótona não-decrescente e $\phi'$ Riemann integrável em $[a,b]$. Temos $f$ Riemann-Stieltjes integrável com relação a $\phi$ se e somente $f\phi'$ é Riemann integrável. Neste caso

$$\int_a^b f d\phi = \int_a^b f(x)\phi'(x)dx.$$

*Demonstração.* Dado $\varepsilon > 0$, sendo $\phi'$ Riemann integrável, existe uma partição $P = \{x_0, \ldots, x_n\}$ de $[a,b]$ tal que

$$S(\phi';P) - s(\phi';P) < \varepsilon.$$

O teorema do valor médio fornece pontos $t_i \in [x_{i-1}, x_i]$ de modo que

$$\Delta\phi_i = \phi'(t_i)\Delta x_i$$

para $i = 1,...,n$. Se $s_i \in [x_{i-1}, x_i]$, temos

$$\sum_{i}^{n} |\phi'(s_i) - \phi'(t_i)|\Delta x_i < \varepsilon.$$

Sendo $M = sup|f(x)|$ e

$$\sum_{i}^{n} f(s_i)\Delta\phi_i = \sum_{i}^{n} f(s_i)\phi'(t_i)\Delta x_i,$$

segue que

$$\left|\sum_{i=1}^{n} f(s_i)\Delta\phi_i - \sum_{i=1}^{n} f(s_i)\phi'(s_i)\Delta x_i\right| \leq M\varepsilon. \tag{1}$$

Em particular,

$$\sum_{i=1}^{n} f(s_i)\Delta\phi_i \leq S(f\phi';P) + M\varepsilon,$$

para todas as escolhas de $s_i \in [x_{i-1}, x_i]$, temos

$$S(f,\phi;P) \leq S(f\phi';P) + M\varepsilon.$$

O mesmo argumento, conduz a partir de (1) que

$$S(f\phi';P) \leq S(f,\phi;P) + M\varepsilon.$$

Assim,

$$|S(f,\phi;P) - S(f\phi';P)| \leq M\varepsilon. \tag{2}$$

Note que $S(\phi';P) - s(\phi';P) < \varepsilon$ se mantém verdadeira para qualquer refinamento da partição $P$ e a desigualdade (2) também se mantém verdadeira. Concluímos que

$$\left|\overline{\int_a^b} f d\phi - \overline{\int_a^b} f(x)\phi'(x)dx\right| \leq M\varepsilon.$$

Sendo $\varepsilon$ arbitrário, consequentemente

$$\overline{\int_a^b} f d\phi = \overline{\int_a^b} f(x)\phi'(x)dx,$$

para toda função limitada $f$. A igualdade das integrais inferiores segue de (1) de modo análogo.

$\square$

**Exemplo 3.11.** Seja $f(x) = x$ e $\phi(x) = x^2$ em $[0,1]$.

$$\int_0^1 f(x)d\phi(x) = \int_0^1 x.2xdx = \frac{2}{3}.$$

## 3.5 Propriedades da integral de Riemann-Stieltjes

**Teorema 3.4.** Sejam $f$ e $g$ funções limitadas e $\phi$ uma função não-decrescente no intervalo $[a,b]$.

**(a)** Se $f$, $g$ são Riemann-Stieltjes integráveis com relação a $\phi$ em $[a,b]$, então $f+g$ é Riemann-Stieltjes integrável com relação a $\phi$ em $[a,b]$ e

$$\int_a^b (f+g)d\phi(x) = \int_a^b fd\phi(x) + \int_a^b gd\phi(x).$$

**(b)** Se $f$ é Riemann-Stieltjes integrável com relação a $\phi$ em $[a,b]$ e $c \in \mathbb{R}$, então $cf$ é Riemann-Stieltjes integrável com relação a $\phi$ em $[a,b]$ e

$$\int_a^b cfd\phi(x) = c\int_a^b fd\phi(x).$$

**(c)** Se $f$, $g$ são Riemann-Stieltjes integráveis com relação a $\phi$ em $[a,b]$ e $f \leq g$ em $[a,b]$, então

$$\int_a^b fd\phi(x) \leq \int_a^b gd\phi(x).$$

**(d)** Se $f$ é Riemann-Stieltjes integrável com relação a $\phi$ em $[a,b]$, então $|f|$ também é Riemann-Stieltjes integrável e

$$\left|\int_a^b fd\phi(x)\right| \leq \int_a^b |f|d\phi(x).$$

As demonstrações destas propriedades podem ser encontradas em [7].

**Teorema 3.5.** Seja $\phi : [a,b] \to \mathbb{R}$ não-decrescente e $c \in (a,b)$. Sendo $f : [a,b] \to \mathbb{R}$ limitada, $f$ é Riemann-Stieltjes integrável em relação a $\phi$ em $[a,b]$ se e somente se suas restrições $f|[a,c]$ e $f|[c,b]$ são integráveis e tem-se

$$\int_a^b f d\phi(x) = \int_a^c f d\phi(x) + \int_c^b f d\phi(x).$$

A demonstração deste teorema pode ser encontrada em [7].

**Teorema 3.6. Teorema Fundamental do Cálculo para as integrais de Riemann-Stieltjes.**

Seja $f$ contínua no intervalo $[a,b]$ e $\phi(x)$ monótona não-decrescente em $[a,b]$, então $\int_a^b f(x)d\phi(x)$ existe.

Definimos uma nova função no intervalo $[a,b]$ por

$$F(x) = \int_a^x f(t)d\phi(t) \text{ então}$$

**(a)** $F$ é contínua em qualquer ponto onde $\phi$ é contínua e

**(b)** $F$ é diferenciável em cada ponto onde $\phi$ é diferenciável e $F' = f.\phi'$ em tal ponto.

*Demonstração.* A existência de $\int_a^b f(x)d\phi(x)$ é óbvia se $\phi(a) = \phi(b)$ pois neste caso para toda partição $P$ de $[a,b]$, $\sum_P f\Delta\phi(x) = 0$ e daí $\int_a^b f(x)d\phi(x) = 0$. Suponha $\phi(b) - \phi(a) > 0$. Temos $s(f,\phi;P) \leq S(f,\phi;P)$. Observe que $S(f,\phi;P) - s(f,\phi;P) = \sum\limits_P (sup(f) - inf(f))\Delta\phi$ pode ficar tão pequena quanto se queira devido à continuidade uniforme de $f$, então pelo Teorema 3.2 a integral de Riemann-Stieltjes existe.

Para mostrarmos o item (a), podemos assumir que $|f(t)| \leq B$, $\forall t \in [a,b]$ com $B \neq 0$ assim

$$|F(y) - F(x)| = \left| \int_x^y f(t)d\phi(t) \right| \text{ e,}$$

para qualquer partição $P = \{x_0, ..., x_n\}$ de $[a,b]$ temos

$\sum\limits_P f\Delta\phi = f(c_1)[\phi(x_1) - \phi(x_0)] + f(c_2)[\phi(x_2) - \phi(x_1)] + ... + f(c_n)[\phi(x_n) - \phi(x_{n-1})]$

$\leq B[\phi(x_1) - \phi(x_0)] + ... + B[\phi(x_n) - \phi(x_{n-1})] = B[\phi(x_n) - \phi(x_0)] = B[\phi(b) - \phi(a)]$

$\Rightarrow \sum\limits_P f\Delta\phi \leq B[\phi(b) - \phi(a)]$. Analogamente $-B[\phi(b) - \phi(a)] \leq \sum\limits_P f\Delta\phi$. Logo

$-B[\phi(b) - \phi(a)] \leq \sum\limits_P f\Delta\phi \leq B[\phi(b) - \phi(a)] \Rightarrow -B[\phi(b) - \phi(a)] \leq \lim\limits_{|P|\to 0} \sum f\Delta\phi \leq B[\phi(b) - \phi(a)]$. Dado que a integral existe temos, $-B[\phi(b) - \phi(a)] \leq \int_a^b f d\phi \leq B[\phi(b) - \phi(a)] \Rightarrow \left| \int_a^b f d\phi \right| \leq B|\phi(b) - \phi(a)|$ e portanto

$$\left| \int_x^y f(t) d\phi(t) \right| \leq B|\phi(y) - \phi(x)|.$$

Se $\phi$ é contínua em $y$, dado $\varepsilon > 0$, existe $\delta > 0$ tal que $|y-x| < \delta \Rightarrow |\phi(y) - \phi(x)| < \varepsilon/B$ e daí $B|\phi(y) - \phi(x)| < \varepsilon$.

Assim $|F(y) - F(x)| \leq B|\phi(y) - \phi(x)| < \varepsilon$ e concluímos que $F$ é contínua em $y$.

Para provar (b), suponha que $a < x < b$ e que $\phi$ seja diferenciável no ponto $x$. Sendo $f$ contínua, dado $\varepsilon > 0$, existe $\delta > 0$ tal que para todo $t \in [a,b]$ com $0 < |x-t| < \delta$ temos $|f(t) - f(x)| < \varepsilon$ e

$$\left| \frac{\phi(t) - \phi(x)}{t - x} - \phi'(x) \right| < \varepsilon.$$

Então, se $y$ é qualquer ponto de $[a,b] \bigcap (x, x+\delta)$, temos

$$\begin{aligned}
&\left| \frac{F(y) - F(x)}{y - x} - f(x)\phi'(x) \right| \leq \left| \frac{1}{y-x} \int_x^y f(t) d\phi(t) - \frac{1}{y-x} \int_x^y f(x) d\phi(t) \right| \\
&+ \left| \frac{1}{y-x} \int_x^y f(x) d\phi(t) - f(x)\phi'(x) \right| \\
&\leq \frac{1}{y-x} \int_x^y |f(t) - f(x)| d\phi(t) + \left| \frac{f(x)}{y-x} \int_x^y d\phi(t) - f(x)\phi'(x) \right| \\
&= \frac{1}{y-x} \int_x^y |f(t) - f(x)| d\phi(t) + \left| f(x). \frac{\phi(y) - \phi(x)}{y-x} - f(x)\phi'(x) \right| \\
&= \frac{1}{y-x} \int_x^y |f(t) - f(x)| d\phi(t) + |f(x)| \left| \frac{\phi(y) - \phi(x)}{y-x} - \phi'(x) \right| \\
&\leq \frac{\varepsilon}{y-x} \int_x^y d\phi(t) + M\varepsilon = \frac{\varepsilon[\phi(y) - \phi(x)]}{y-x} + M\varepsilon \leq \varepsilon[\phi'(x) + \varepsilon] + M\varepsilon,
\end{aligned}$$

onde $M = sup\{|f(x)|; x \in [a,b]\}$. Isto prova que $F$ é diferenciável à direita do ponto $x$ e que neste ponto $F' = f.\phi'$.

O caso em que $y$ é qualquer ponto de $[a,b] \bigcap (x-\delta, x)$ é análogo. □

## 3.6 Valor esperado e Integral de Riemann-Stieltjes

Podemos unificar as definições de valor esperado de uma variável aleatória discreta e variável aleatória contínua através da integral de Riemann-Stieltjes.

**Definição 3.14.** Seja $X$ uma variável aleatória com função de distribuição $F$. Definimos valor esperado ou esperança matemática ou média de $X$ por

$$E(X)=\int_{-\infty}^{+\infty} x dF(x),$$

desde que a integral esteja bem definida.

**Exemplo 3.12.** Vamos calcular o valor esperado para a variável discreta do Exemplo 3.7 utilizando a integral de Riemann-Stieltjes. Um jogador lança 3 moedas não viciadas. Ganha R\$ $6,00$ se 3 caras ocorrerem, ganha R\$ $3,00$ se duas caras ocorrerem, ganha R\$ $1,00$ se somente uma cara ocorrer. Por outro lado, perde R\$ $10,00$ se 3 coroas ocorrerem. Neste caso, temos a seguinte função de distribuição:

$$F(x)=P(X\leq x)\begin{cases} 0, & se \quad x<-10; \\ \frac{1}{8}, & se \quad -10\leq x<1; \\ \frac{1}{8}+\frac{3}{8}=\frac{4}{8}, & se \quad 1\leq x<3; \\ \frac{4}{8}+\frac{3}{8}=\frac{7}{8}, & se \quad 3\leq x<6; \\ \frac{7}{8}+\frac{1}{8}=1, & se \quad x\geq 6. \end{cases}$$

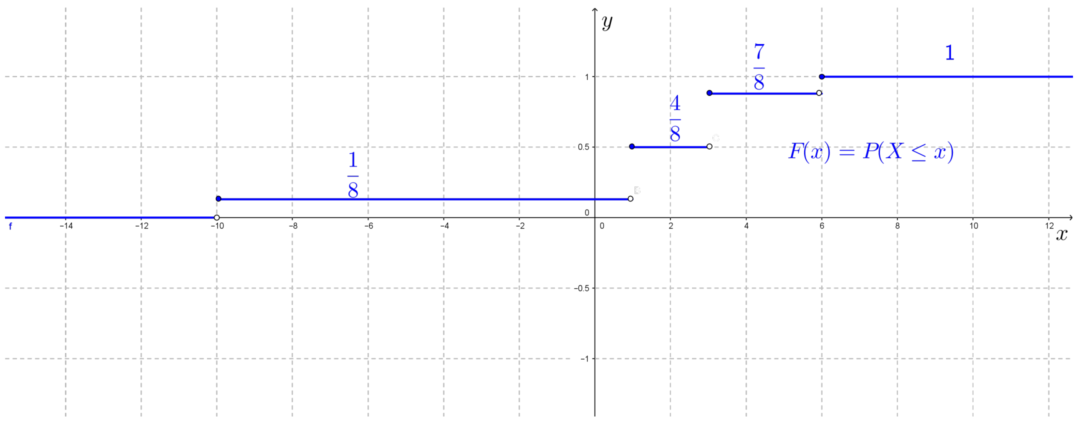



Figura 6: Gráfico da função de distribuição.

Como observado anteriormente na seção 3.3, como $F$ possui descontinuidades do tipo salto, o valor esperado é dado por

$$E(X)=\int_{-\infty}^{+\infty} x dF(x)=\sum_{-\infty<x<+\infty} x[F(x)-F(x^{-})] \text{ e portanto}$$

$$E(X)=\int_{-\infty}^{+\infty} x dF(x)=-10\left(\frac{1}{8}-0\right)+1\left(\frac{4}{8}-\frac{1}{8}\right)+3\left(\frac{7}{8}-\frac{4}{8}\right)+6\left(1-\frac{7}{8}\right)=1.$$

No caso de $X$ ser uma variável aleatória contínua, a função de distribuição $F$ é absolutamente contínua e temos

$$dF(x) = f(x)dx,$$

com $f$ sendo a derivada de $F$. Neste caso

$$E(X) = \int_{-\infty}^{+\infty} x dF(x) = \int_{-\infty}^{+\infty} x f(x) dx.$$

**Exemplo 3.13.** Seja $X$ uma variável contínua com a seguinte função de distribuição:

$$F(x) = \begin{cases} 0, & se \quad x < 2; \\ 1 - e^{4-2x}, & se \quad x \geq 2. \end{cases}$$

Para o cálculo do valor esperado, usaremos que $dF(x) = 2e^{4-2x}dx$ para $x$ em $[2, \infty)$ e, portanto:

$$E(X) = \int_{-\infty}^{+\infty} x dF(x) = \int_{2}^{+\infty} x 2e^{4-2x} dx = \frac{5}{2}.$$

Encerramos a seção, com um teorema que pode tornar mais simples o cálculo do valor esperado para variáveis aleatórias. Tal resultado é muito útil, pois possibilita a obtenção do valor esperado de uma variável aleatória qualquer, como a diferença de duas integrais de Riemann.

**Teorema 3.7.** Seja $X$ uma variável aleatória com função de distribuição $F$ e cujo valor esperado existe. Então,

$$E(X) = \int_{-\infty}^{+\infty} x dF(x) = \int_{0}^{+\infty} [1 - F(x)] dx - \int_{-\infty}^{0} F(x) dx.$$

*Demonstração.* Dividindo o intervalo de integração em duas partes, vamos mostrar que

$$\int_{0}^{+\infty} x dF(x) = \int_{0}^{+\infty} [1 - F(x)] dx \quad \text{e} \quad \int_{-\infty}^{0} x dF(x) = -\int_{-\infty}^{0} F(x) dx.$$

Aplicando integração por partes (Teorema 3.1), temos para $b > 0$,

$$\int_{0}^{b} x dF(x) = bF(b) - 0F(0) - \int_{0}^{b} F(x) dx = \int_{0}^{b} [F(b) - F(x)] dx \leq \int_{0}^{b} [1 - F(x)] dx$$

e passando ao limite para $b \to +\infty$,

$$\int_{0}^{+\infty} x dF(x) \leq \int_{0}^{+\infty} [1 - F(x)] dx. \tag{1}$$

Considere, agora, $c > 0$ tal que $c < b$. Temos

$$\int_{0}^{b} x dF(x) = \int_{0}^{b} [F(b) - F(x)] dx \geq \int_{0}^{c} [F(b) - F(x)] dx$$

$$= \int_0^c [F(b)-1]dx + \int_0^c [1-F(x)]dx$$

$$= c[F(b)-1] + \int_0^c [1-F(x)]dx;$$

então, passando ao limite para $b \to +\infty$, obtemos

$$\int_0^{+\infty} xdF(x) \geq \int_0^c [1-F(x)]dx,$$

e, agora, com $c \to +\infty$ segue que

$$\int_0^{+\infty} xdF(x) \geq \int_0^{+\infty} [1-F(x)]dx. \tag{2}$$

Das desigualdades (1) e (2), concluímos que

$$\int_0^{+\infty} xdF(x) = \int_0^{+\infty} [1-F(x)]dx.$$

Para a parte negativa, seja $b < 0$, aplicando integração por partes temos

$$\int_b^0 xdF(x) = 0F(0) - bF(b) - \int_b^0 F(x)dx;$$

passando ao limite, para $b \to -\infty$, obtemos

$$\int_{-\infty}^0 xdF(x) = -\int_{-\infty}^0 F(x)dx,$$

e isto completa a demonstração.

$\square$

**Exemplo 3.14.** A variável $X$ do Exemplo 3.9 é do tipo mista e sua função de distribuição é reapresentada a seguir:

$$F(x) = \begin{cases} 0, & se \quad x < -2; \\ \frac{1}{24}(x^2+4x+4), & se \quad -2 \leq x < 0; \\ \frac{1}{6}(x+2), & se \quad 0 \leq x < 1; \\ \frac{1}{6}(x+3), & se \quad 1 \leq x < 2; \\ \frac{1}{24}(-x^2+8x+8), & se \quad 2 \leq x < 4; \\ 1, & se \quad x \geq 4. \end{cases}$$

Pelo Teorema 3.7 temos

$$
\begin{aligned}
E(X) &= \int_0^{+\infty} [1-F(x)]dx - \int_{-\infty}^{0} F(x)dx \\
&= \int_0^1 \frac{(4-x)}{6}dx + \int_1^2 \frac{(3-x)}{6}dx \\
&+ \int_2^4 \frac{(16+x^2-8x)}{24}dx - \int_{-2}^{0} \frac{(x^2+4x+4)}{24}dx = \frac{5}{6}.
\end{aligned}
$$

Observe que esse valor já havia sido obtido antes. Entretanto, agora, calculamos o valor esperado sem a necessidade de obtermos a decomposição na parte contínua. Outras propriedades do valor esperado utilizando a integral de Riemann-Stieltjes podem ser encontradas em [6].

## 3.7 Sequências de Funções

**Definição 3.15.** Seja $X \subset \mathbb{R}$. Se para cada $n \in \mathbb{N}$ fizermos corresponder uma função $f_n$, definida em $X$, isto é, $f_n : X \to \mathbb{R}$ então, $f_n$ é uma sequência de funções.

**Definição 3.16.** Uma sequência de funções $f_n : X \to \mathbb{R}$, $n = 1, 2, ...$ converge simplesmente para a função $f : X \to R$ quando para todo $x \in X$, a sequência de números $f_1(x), ..., f_n(x)$ converge para $f(x)$. Isto significa que dado $\varepsilon > 0$ e $x \in X$, existe $n_0 \in \mathbb{N}$ ( dependendo de $\varepsilon$ e de $x$) tal que $n > n_0 \Rightarrow |f_n(x) - f(x)| < \varepsilon$.

Graficamente, dado $x \in X$, traçando-se uma reta vertical pelo ponto $x$, a interseção da reta com os gráficos de cada $f_n$ determinam uma sequência de pontos $(x, f_1(x)), ..., (x, f_n(x)), ...$ que convergem para $(x, f(x))$.

**Exemplo 3.15.** A sequência de funções $f_n : \mathbb{R} \to \mathbb{R}$ definidas por $f_n(x) = x/n$ convergem simplesmente para a função identicamente nula. A Figura 7 ilustra o fato de que, fixado um $\varepsilon > 0$, se $x = 1,5$ temos $n_0 = 2$ e se $x = 3,25$ temos $n_0 = 4$ deixando claro que $n_0$ depende não apenas de $\varepsilon$ mas também de $x$.

**Definição 3.17.** Uma sequência de funções $f_n : X \to \mathbb{R}$ converge uniformemente para a função $f : X \to \mathbb{R}$ quando, para todo $\varepsilon > 0$ dado, existe $n_0 \in \mathbb{N}$ (dependendo apenas de $\varepsilon$) tal que $n > n_0 \Rightarrow |f_n(x) - f(x)| < \varepsilon$ seja qual for $x \in X$.

No plano $\mathbb{R}^2$, dado $\varepsilon > 0$, a faixa de raio $\varepsilon$ em torno do gráfico de $f$ é o conjunto

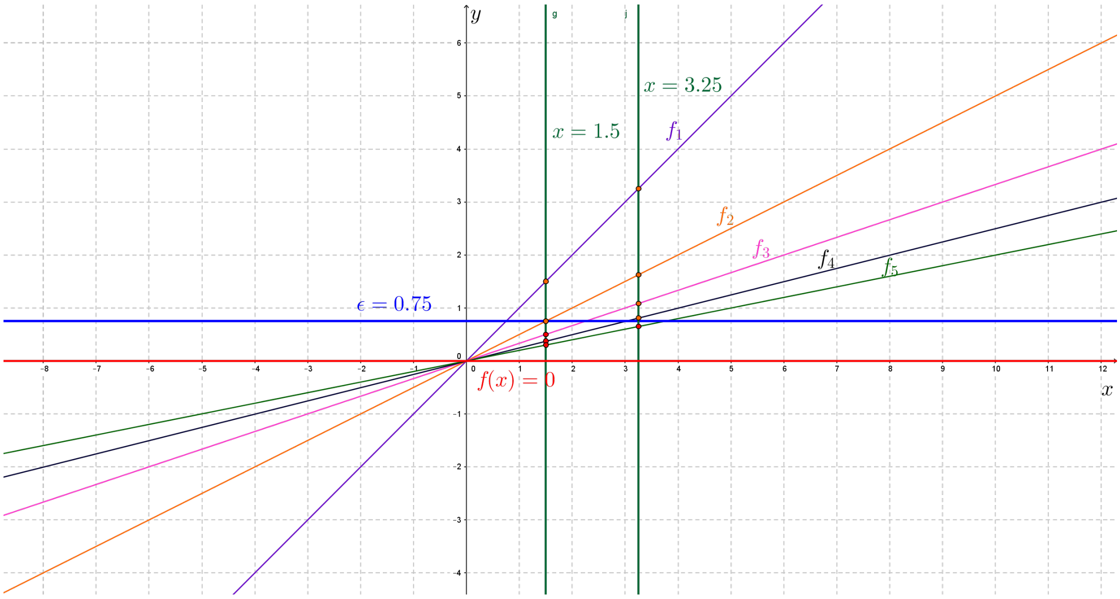



Figura 7: Convergência simples ou pontual.

$$F(f;\varepsilon) = \{(x,y) \in \mathbb{R}^2; x \in X, f(x) - \varepsilon < y < f(x) + \varepsilon\}.$$

Geometricamente, dizer que $f_n \to f$ uniformemente significa que para todo $\varepsilon > 0$, existe $n_0 \in \mathbb{N}$ tal que o gráfico de $f_n$, para todo $n > n_0$, está contido na faixa de raio $\varepsilon$ em torno do gráfico de $f$.

**Exemplo 3.16.** A sequência de funções $f_n : \mathbb{R} \to \mathbb{R}$ definidas por $f_n(x) = x/n$ não converge uniformemente para zero em $\mathbb{R}$ pois, nenhuma faixa de raio $\varepsilon$ em torno do gráfico da função identicamente nula pode conter o gráfico de uma função $f_n(x) = x/n$. Porém se $X \subset \mathbb{R}$ é um conjunto limitado, digamos com $|x| \leq c$ para todo $x \in X$, então $f_n \to 0$ uniformemente em $X$. De fato, dado $\varepsilon > 0$, tome $n_0 > c/\varepsilon$, então para $n > n_0$, $|f_n(x)| = \frac{|x|}{n} \leq \frac{c}{n} < \frac{c}{n_0} < c\frac{\varepsilon}{c} = \varepsilon$. A Figura 8 ilustra o caso em que $|x| < 3$, observe que dado $\varepsilon = 0,8$ tomando $n > 3$, $f_n(x) = x/n$ está contida na faixa de raio $\varepsilon$ em torno do gráfico da função identicamente nula, assim, a convergência é uniforme.

**Teorema 3.8. Passagem ao limite sob o sinal de integral.** Se a sequência de funções integráveis $f_n : [a,b] \to \mathbb{R}$ converge uniformemente para $f : [a,b] \to \mathbb{R}$ então $f$ é integrável e

$$\int_a^b \lim_{n\to\infty} f_n(x)dx = \lim_{n\to\infty} \int_a^b f_n(x)dx.$$

*Demonstração.* Dado $\varepsilon > 0$, existe $n_0 \in \mathbb{N}$ tal que $n > n_0 \Rightarrow |f(x) - f_n(x)| < \varepsilon/4(b-a)$ para todo $x \in [a,b]$. Tome $m > n_0$. Como $f_m$ é integrável, existe uma partição $P$ de $[a,b]$ tal que,

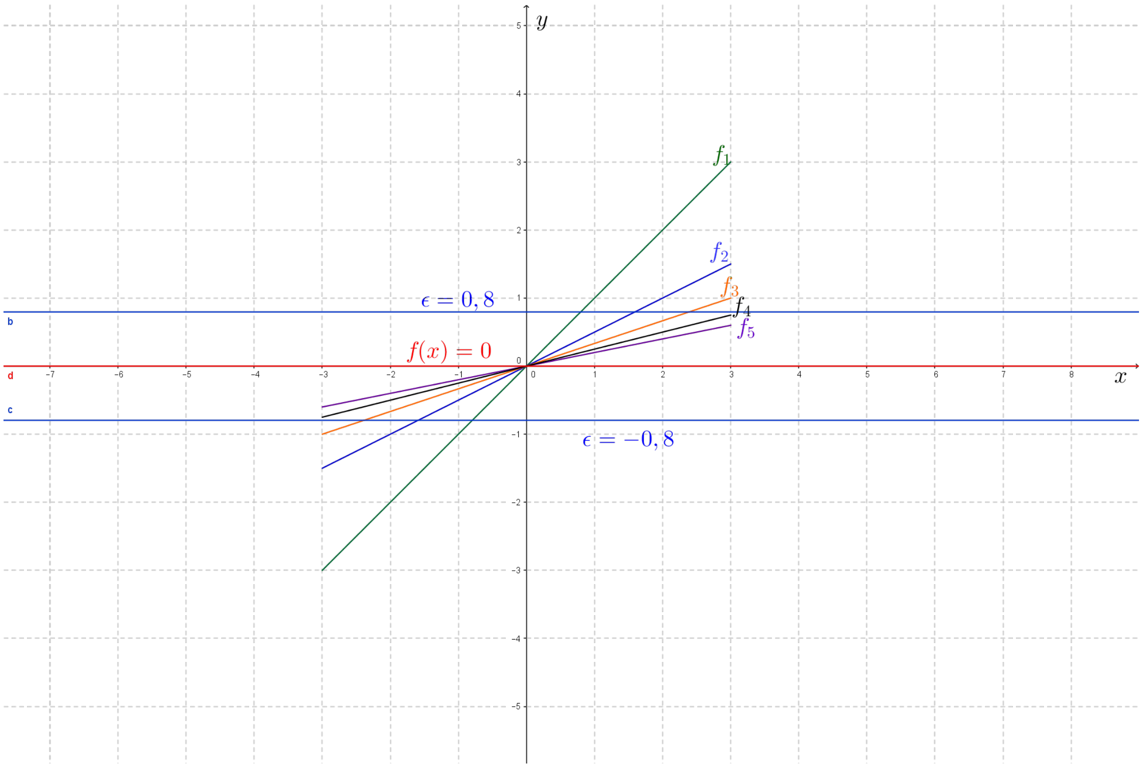



Figura 8: A convergência é uniforme se $f_n : [-3,3] \to \mathbb{R}$, com $f_n(x) = x/n$.

indicando com $\omega_i$ e $\omega_i'$ respectivamente as oscilações de $f$ e $f_m$ no intervalo $[t_{i-1}, t_i]$ de $P$, tem-se $\sum \omega_i'(t_i - t_{i-1}) < \varepsilon/2$. Mas, para $x$, $y \in [t_{i-1}, t_i]$ quaisquer, aplicando a desigualdade triângular sucessivas vezes temos:

$|f(y) - f(x)| = |(f(y) - f_m(y)) + (f_m(y) - f_m(x) + f_m(x) - f(x))| \leq |f(y) - f_m(y)| + |(f_m(y) - f_m(x)) + (f_m(x) - f(x))| \leq |f(y) - f_m(y)| + |f_m(y) - f_m(x)| + |f_m(x) - f(x)| < \varepsilon/4(b - a) + \omega_i' + \varepsilon/4(b-a) = \omega_i' + \varepsilon/2(b-a)$.

Portanto $\omega_i \leq \omega_i' + \varepsilon/2(b-a)$. Segue-se que

$\sum \omega_i(t_i - t_{i-1}) \leq \sum \omega_i'(t_i - t_{i-1}) + [\varepsilon/2(b-a)] \sum (t_i - t_{i-1}) < \varepsilon/2 + \varepsilon/2 = \varepsilon$. Isto mostra que $f$ é integrável. Além disso, se $n > n_0$,

$$\left| \int_a^b f(x)dx - \int_a^b f_n(x)dx \right| = \left| \int_a^b [f(x) - f_n(x)]dx \right|$$

$$\leq \int_a^b |f(x) - f_n(x)|dx \leq \frac{(b-a)\varepsilon}{4(b-a)} < \varepsilon.$$

Logo, $\lim\limits_{n\to\infty} \int_a^b f_n(x)dx = \int_a^b f(x)dx$. $\square$

A seguir vamos listar alguns exemplos que ilustram a necessidade da convergência ser uniforme no Teorema 3.8.

**Exemplo 3.17.** Se uma sequência de funções integráveis $f_n : [a,b] \to \mathbb{R}$ converge simplesmente para $f : [a,b] \to \mathbb{R}$, pode ocorrer que $f$ não seja integrável. Por exemplo, se $\{r_1, r_2, ..., r_n, ...\}$ for uma enumeração dos números racionais de $[a,b]$ e definirmos $f_n$ como a função que assume o valor 1 nos pontos $r_1, ..., r_n$ e é zero nos demais pontos de $[a,b]$ então $(f_n)$ converge simplesmente para uma função $f : [a,b] \to \mathbb{R}$ tal que $f(x) = 1$ se $x \in \mathbb{Q} \cap [a,b]$ e $f(x) = 0$ se $x$ é irracional. Observe que cada $f_n$ é integrável pois $\underline{\int}_a^b f_n(x)dx = \overline{\int}_a^b f_n(x)dx = 0$ mas $f$ não é uma vez que $\underline{\int}_a^b f(x)dx = 0$ e $\overline{\int}_a^b f(x)dx = b-a$.

**Exemplo 3.18.** Mesmo quando a sequência de funções integráveis $f_n : [a,b] \to \mathbb{R}$ converge simplesmente para a função integrável $f : [a,b] \to \mathbb{R}$ e as integrais existem, pode ocorrer que $\lim\limits_{n\to\infty} \int_a^b f_n(x)dx \neq \int_a^b f(x)dx$. Por exemplo, para cada $n \in \mathbb{N}$, seja $f_n : [0,1] \to \mathbb{R}$ definida por $f_n(x) = nx^n(1-x^n)$. Temos $\lim\limits_{n\to\infty} nx^n = 0$ se $0 \leq x < 1$ . Portanto $(f_n)$ converge simplesmente em $[0,1]$ para a função identicamente nula. Entretanto

$$\int_0^1 f_n(x)dx = \int_0^1 (nx^n - nx^{2n})dx = n\int_0^1 (x^n - x^{2n})dx = \frac{n}{n+1} - \frac{n}{2n+1} = \frac{n^2}{2n^2+3n+1}$$

e

$$\lim_{n\to\infty} \frac{n^2}{2n^2+3n+1} = \frac{1}{2}, \text{ ou seja, } \lim_{n\to\infty} \int_0^1 f_n(x)dx = \frac{1}{2} \text{ enquanto } \int_0^1 (\lim_{n\to\infty} f_n(x))dx = 0.$$

Vale mencionar aqui que a hipótese da convergência uniforme é muito forte. O matemático Arzelà conseguiu enfraquecer essa hipótese, porém, seu teorema é apenas um caso particular do teorema que em 1902 o matemático Henri Lebesgue apresentou como resultado central de sua extensão do conceito de integral, o Teorema da Convergência Dominada.

O conceito de integral tem várias extensões, podemos citar a integral de Lebesgue, integral de Henstock-Kurzweil, integral de Wiener, entre outros e, consequentemente, permitem a obtenção de novos resultados, a aplicabilidade de outros, o avanço da Matemática. Detalhes sobre as integrais citadas aqui podem ser encontradas em [2]. Não abordaremos, neste trabalho, a definição de integral de Lebesgue, mas vamos apontar um problema sobre convergência de funções integráveis e um resultado obtido da teoria de integração de Lebesgue que nos auxilia na resolução deste problema, o conhecido Teorema da Convergência Dominada.

**Teorema 3.9. Teorema da Convergência Dominada**. Se uma sequência de funções integráveis $f_n$ é tal que $f_n(x) \to f(x)$ para todo $x$, e se existe alguma função integrável $g$ tal que $g(x) \geq |f_n(x)|$

para todo $n$ e todo $x$, então $f$ é integrável (à Lebesgue) e a integral de $f$ é o limite da integral de $f_n$.

Para funções $f_n \geq 0$ integráveis à Lebesgue vale

$$\int_a^b \sum_{n=1}^{\infty} f_n(x)dx = \sum_{n=1}^{\infty} \int_a^b f_n(x)dx,$$

mas à Riemann, a integral à esquerda pode não estar definida como vimos no exemplo 3.18 desta seção.

Este problema de trocar os limites aparece, por exemplo, na teoria de equações diferenciais, quando queremos mostrar a existência de soluções de uma equação diferencial ordinária.

Vimos que uma função limitada é integrável à Riemann em $[a,b]$ se e somente se o conjunto de seus pontos de descontinuidade tem medida nula (medida de Lebesgue zero). A condição da medida de Lebesgue zero nos mostra que a teoria de Lebesgue veio generalizar a teoria de Riemann.

CAPÍTULO

4

# APLICAÇÕES DA INTEGRAL DE RIEMANN

A soma inferior e superior para integrais de Riemann apresentadas no capítulo 2, podem ser interpretadas como um valor aproximado, respectivamente, por falta e por excesso, da área da região limitada pelo gráfico de uma função contínua e positiva $f : [a,b] \to \mathbb{R}$, pelas retas verticais levantadas nos pontos $a$ e $b$ e pelo intervalo $[a,b]$ do eixo $x$. Neste capítulo, vamos utilizar a ideia de aproximação por retângulos, tomando retângulos cada vez mais estreitos e consequentemente, vamos nos aproximando cada vez mais da área da região. Para uma melhor visualização deste procedimento, vamos utilizar o software GeoGebra (Versão 5.0.166.0) para construir as somas inferiores e superiores.

Com o uso do GeoGebra, podemos dar uma noção intuitiva para os alunos de como calcular a área de uma região com o contorno curvo somando as áreas de um conjunto de retângulos, o software é de fácil acesso, os comandos são simples, e nos permite criar uma animação, preenchendo a área da região com retângulos cada vez mais estreitos.

Vamos utilizar essas aproximações da área de uma região para estudarmos o deslocamento e o espaço percorrido por um móvel em movimento retilíneo uniforme (MRU) e movimento retilíneo uniformemente variado (MRUV). Através das aproximações, estudamos também o deslocamento e o espaço percorrido para movimentos cuja função velocidade não é necessariamente afim.

## 4.1 Fazendo estimativas com somas finitas

Seja $f$ contínua em $[a,b]$, com $f(x) \geq 0$ em $[a,b]$ e $P : a = t_0 < t_1 < ... < t_n = b$ uma partição de $[a,b]$. A medida que a norma da partição $P$ se aproxima de zero, a soma inferior e a soma superior se aproximam cada vez mais da área da região limitada pelo gráfico de $f$, pelo intervalo $[a,b]$ do eixo $x$ e pelas retas $x = a$ e $x = b$. Mas sabemos que, o valor limite é a integral de Riemann como apresentada no capítulo 2, e como descrito na seção 2.5, essa aproximação não necessariamente precisa ser feita por falta ou por excesso, podemos considerar qualquer retângulo considerando as partições pontilhadas. Assim, a área sob a curva $y = f(x)$ em $[a,b]$ será a integral de $f$ de $a$ até $b$:

$$A = \int_a^b f(x)dx.$$

Por meio deste conceito de área de uma região com contorno curvo, podemos propor aos alunos a tarefa de calcular aproximações desta área, utilizando retângulos cada vez mais estreitos. Para isso, mostramos a eles como fazer uma estimativa para a área de uma região somando a área de um conjunto de retângulos. Usando-se uma quantidade maior de retângulos podemos aumentar a precisão dessa estimativa. Tomemos como exemplo a área da região que se encontra acima do eixo $x$, abaixo da curva $y = 1 - x^2$, e entre as retas verticais $x = 0$ e $x = 1$.

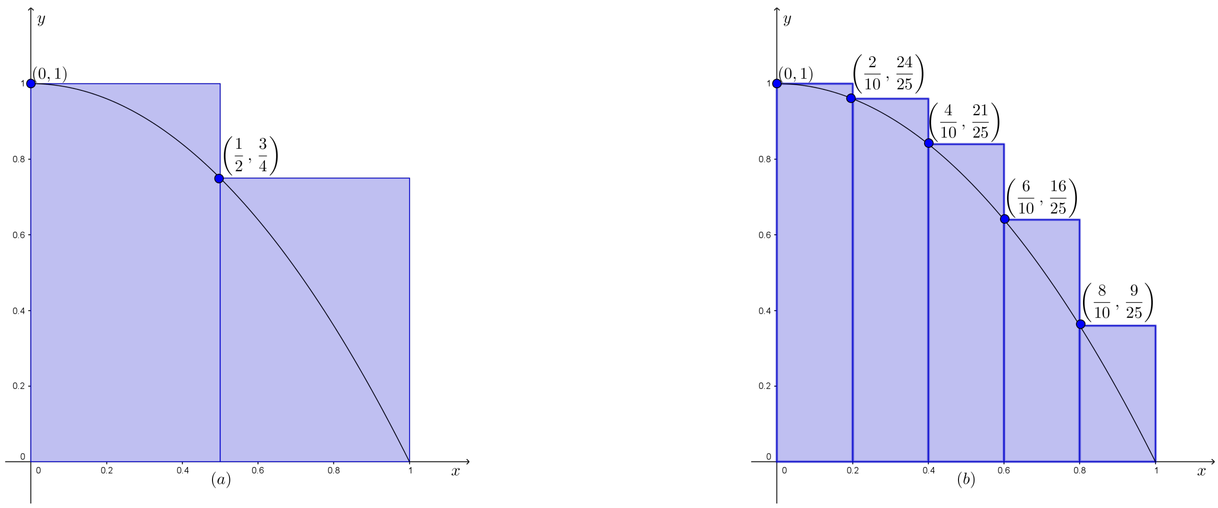



Figura 9: Somas Superiores.

Vamos denotar a área desta região por $S$ e estimar o seu valor através de somas superiores, usando dois retângulos que contêm $S$ e, em seguida, cinco retângulos.

Para isso, consideramos a partição $P_1 = \{0;\ 0,5;\ 1\}$ de $[0,1]$ e calculamos a soma superior, obtendo como estimativa da área da região (Figura 9 (a))

$$S \approx \frac{1}{2}.1 + \frac{1}{2}.\frac{3}{4} = \frac{7}{8} = 0,875.$$

Na Figura 9 (b), melhoramos nossa estimativa ao considerarmos a partição $P_2 = \{0;\ 0,2;\ 0,4;\ 0,6;\ 0,8;\ 1\}$ de $[0,1]$ e, consequentemente obtendo retângulos mais estreitos, cada um de largura $1/5$. A soma da área desses cinco retângulos nos fornece a seguinte aproximação

$$S \approx \frac{1}{5}.1 + \frac{1}{5}.\frac{24}{25} + \frac{1}{5}.\frac{21}{25} + \frac{1}{5}.\frac{16}{25} + \frac{1}{5}.\frac{9}{25} = \frac{95}{125} = 0,76.$$

Ambas as estimativas fornecem valores que ultrapassam o valor real da área.

Vamos agora, estimar a área da região com somas inferiores, primeiro considerando a partição $P_1 = \{0;\ 0,5;\ 1\}$ e depois a partição $P_2 = \{0;\ 0,2;\ 0,4;\ 0,6;\ 0,8;\ 1\}$ de $[0,1]$ como mostra a Figura 10:

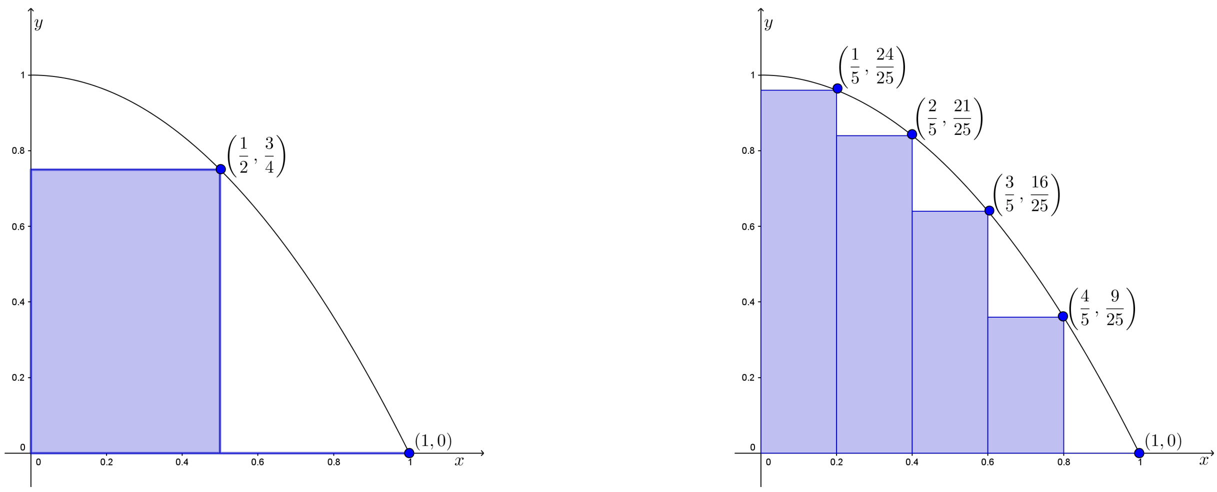



Figura 10: Somas Inferiores.

No primeiro caso, temos a seguinte estimativa para a área da região

$$S \approx \frac{1}{2}.\frac{3}{4} = \frac{3}{8} = 0,375.$$

Já no segundo caso, temos

$$S \approx \frac{1}{5}.\frac{24}{25} + \frac{1}{5}.\frac{21}{25} + \frac{1}{5}.\frac{16}{25} + \frac{1}{5}.\frac{9}{25} = \frac{70}{125} = 0,56.$$

Agora, ambas as estimativas fornecem valores que subestimam o valor real da área.

O verdadeiro valor de $S$ fica em algum ponto entre as somas superior e inferior. Pelos exemplos anteriores

$$0,56 < S < 0,76.$$

Para melhorar as estimativas, podemos considerar retângulos cada vez mais estreitos, vemos isso observando as figuras com 10 retângulos.

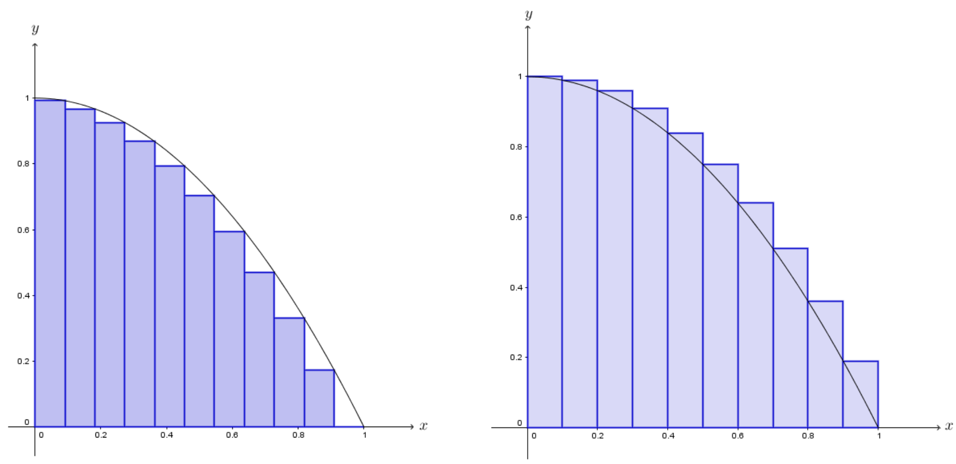


Figura 11: Soma inferior e superior com 10 retângulos.

Para essa visualização, podemos utilizar o software GeoGebra (Versão 5.0.166.0) seguindo os três passos descritos a seguir:

**1º:** Para inserir a função $f : [0,1] \to \mathbb{R}$, levamos o cursor na barra de entrada localizada na parte inferior e inserimos o comando "$Se[0 <= x <= 1, 1 - x^2]$".

**2º:** Para determinar a quantidade de retângulos que desejamos inserir, utilizamos a ferramenta "Controle deslizante". Criamos um controle deslizante com o nome $n$, com incremento 1 e com intervalo $min : 0$ e $max : 50$, este intervalo determina a quantidade mínima e máxima de retângulos.

**3º:** Inserimos na barra de entrada o comando "SomaDeRiemannSuperior$[f, 0, 1, n]$". Neste comando, $f$ é a função determinada no passo 1, os números 0 e 1 indicam o intervalo onde serão inseridos os retângulos e $n$ faz referência ao controle deslizante criado no passo 2. Agora, basta movimentar o botão da barra do controle deslizante para inserir a quantidade de retângulos desejado.

Para a soma inferior o procedimento é o mesmo, bastando inserir o comando “SomaDe-RiemannInferior"ao invés de “SomaDeRiemannSuperior".

Se quisermos criar uma animação, basta clicar com o botão direito do mouse no controle deslizante e selecionar a opção animação, com isso, o botão da barra do controle deslizante vai se movimentar sozinho, percorrendo o intervalo de 0 a 50 como determinamos em sua criação. Assim, o software nos permite criar um “movimento". Colocamos, a seguir, algumas imagens estáticas obtidas durante o processo.

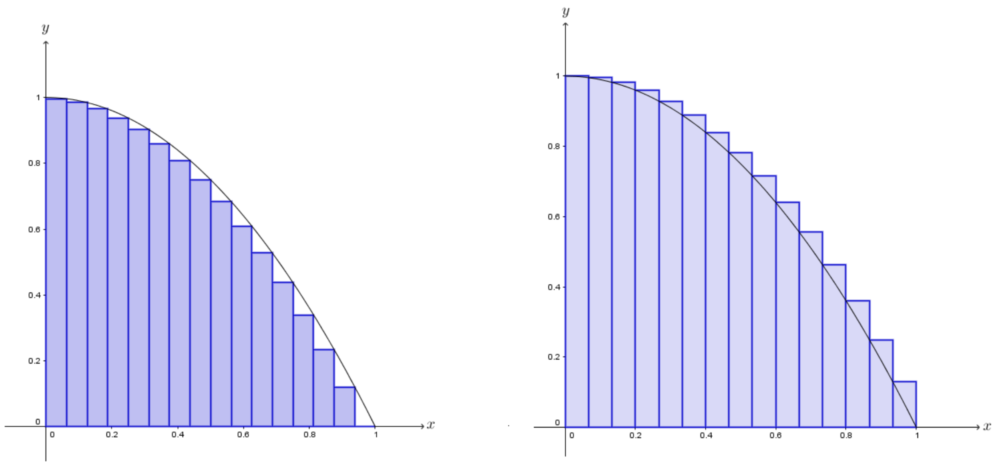


Figura 12: Soma inferior e superior com 15 retângulos.

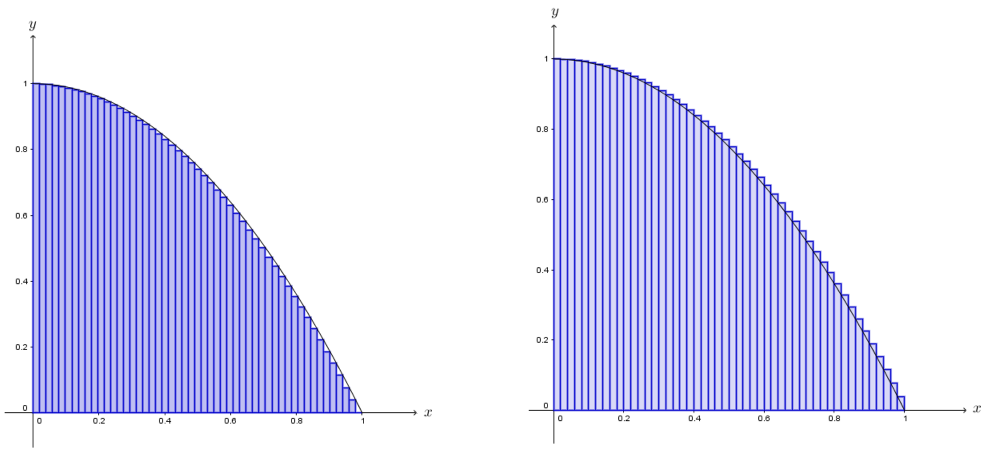


Figura 13: Soma inferior e superior com 50 retângulos.

Observe neste caso que, à medida que tomamos retângulos cada vez mais estreitos, as somas inferiores e superiores ficam cada vez mais próximas. Após a realização destas atividades, espera-se que os alunos tenham adquirido uma noção da interpretação da integral de Riemann como área, mesmo sem mencionar o termo integral de Riemann e o processo de limite.

## 4.2 Deslocamento e espaço percorrido

Nesta seção, nossas principais referências foram [4] e [8]. Localizar um objeto significa determinar a posição do objeto em relação a um ponto de referência, frequentemente a origem (marco zero). Na trajetória de um móvel, escolhemos um marco zero a partir do qual medimos os comprimentos que indicam a posição do móvel. Observe que um móvel pode encontrar-se de um lado ou de outro em relação ao marco zero, sendo portanto conveniente orientar a trajetória, adotando-se um sentido positivo. A uma mudança de uma posição $s_1$ para uma posição $s_2$ é associado um deslocamento $\Delta S$ dado por

$$\Delta S = s_2 - s_1.$$

Se um móvel se desloca a favor da orientação positiva da trajetória, temos um deslocamento positivo, se o móvel se desloca em sentido oposto ao da orientação positiva da trajetória o deslocamento é negativo. Observe que o deslocamento de um móvel não depende da trajetória descrita por ele, sendo necessário saber apenas a posição final e a posição inicial para se determinar o módulo do deslocamento.

O espaço percorrido é a medida referente a trajetória descrita no movimento e o seu valor depende da trajetória.

Quando a velocidade de um móvel é positiva, significa que o móvel está caminhando a favor da orientação positiva da trajetória e se a velocidade é negativa significa que o móvel está caminhando em sentido contrário ao da orientação da trajetória. Em um movimento retilíneo, se a velocidade em um determinado intervalo de tempo não muda de sinal, então, o valor absoluto do módulo do deslocamento é igual ao espaço percorrido. Se em um determinado movimento a velocidade do móvel mudar de sinal, então, há mudança de sentido no movimento e, com isso, o valor absoluto do módulo do deslocamento não será o espaço percorrido.

Vamos usar a ideia de aproximação da área de uma região por retângulos cada vez mais estreitos apresentada na seção 4.1 para estudar o deslocamento e o espaço percorrido por um móvel em movimento retilíneo, ou seja, um movimento em que a trajetória do móvel é uma reta.

No movimento retilíneo uniforme, a velocidade escalar é uma função constante com o tempo e a aceleração é nula. No gráfico da Figura 14, a área da região limitada pela função $v(t) = v, \forall\, t \in [t_1, t_2]$, pelas retas $t = t_1$, $t = t_2$ e pelo eixo $t$, é numericamente igual ao valor absoluto do módulo do deslocamento $\Delta S$ no intervalo de tempo $t_1$ a $t_2$.

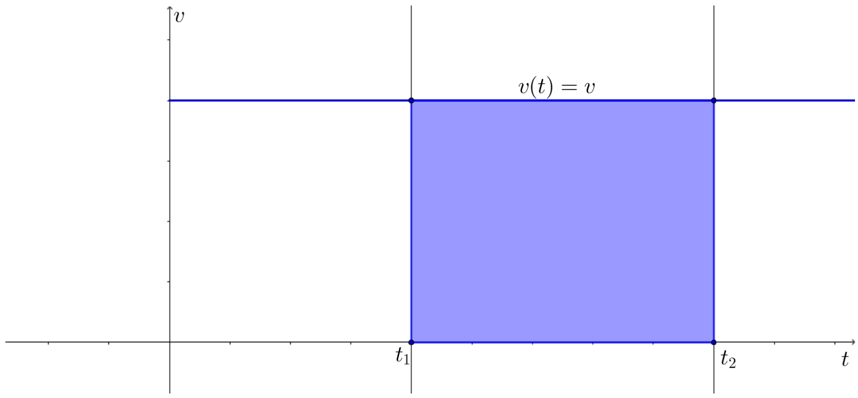



Figura 14: A área do retângulo é numericamente igual ao módulo do deslocamento $\Delta S$.

Se denotarmos a área do retângulo por $A$ temos que $A = (t_2 - t_1)v$. Como $t_2 - t_1 = \Delta t$ e $v = v_m = \frac{\Delta S}{\Delta t}$, vem

$$A = \Delta t.\frac{\Delta S}{\Delta t} \Rightarrow A = \Delta S \text{ (numericamente).}$$

Pelo resultado anterior $\Delta S = v\Delta t$. Fazendo $t_2 = t$ e $t_1 = 0$ temos $s(t) - s(0) = v(t-0)$. Sendo $s_0 = s(0)$ obtemos $s(t) = s_0 + vt$ (função horária do MU).

A seguir, damos alguns exemplos que podem ser aplicados em sala de aula após uma abordagem inicial de cinemática. O objetivo é fazer com que os alunos percebam que os valores do deslocamento e do espaço percorrido pelo móvel em movimento retilíneo, podem ser obtidos pelas funções usuais, já conhecidas por eles, ou através do cálculo de áreas.

**Exemplo 4.1.** Um móvel em movimento retilíneo uniforme tem velocidade de $8m/s$.

(a) Calcule o deslocamento e o espaço percorrido pelo móvel no intervalo de $4s$ a $10s$.

(b) Esboce o gráfico da velocidade em função do tempo para esse movimento.

(c) Calcule a área da região limitada pela função velocidade, pelas retas $t = 4$, $t = 10$ e pelo eixo $t$.

(d) Compare o resultado do item $c$ com o resultado do item $a$. O que podemos concluir a respeito do deslocamento e do espaço percorrido pelo móvel?

**Resolução:**

**(a)** Como o movimento é uniforme, a função dos espaços em relação ao tempo é dada por $s(t) = s_0 + vt$.

No instante $t_1 = 4s$ a posição do móvel é $s_1 = s_0 + 8.4 \Rightarrow s_1 = s_0 + 32$.

No instante $t_2 = 10s$ a posição do móvel é $s_2 = s_0 + 8.10 \Rightarrow s_2 = s_0 + 80$.

O deslocamento entre os instantes $t_1$ e $t_2$ é $\Delta S = s_2 - s_1 = 48m$.

Como o movimento é retilíneo e a velocidade é positiva, o valor do deslocamento coincide com o valor do espaço percorrido $d$ pelo móvel, logo $d = 48m$.

**(b)**

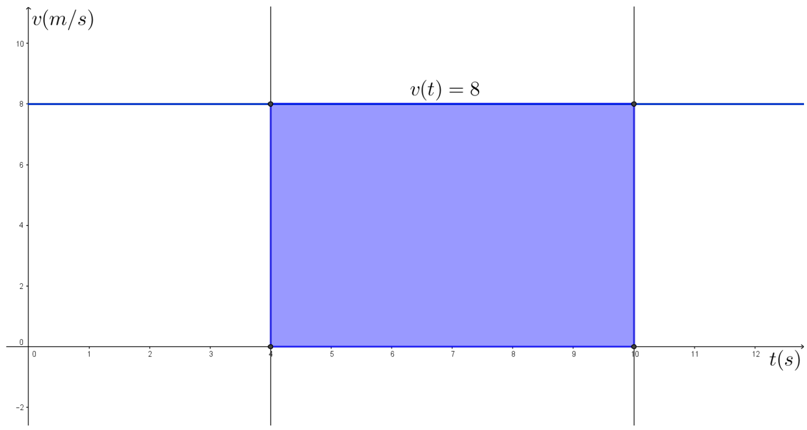



Figura 15: Movimento uniforme.

**(c)** Denotando a área da região por $A$ temos $A = 6.8 = 48$.

**(d)** Podemos concluir que a área da região $A$ é numericamente igual ao módulo do deslocamento e igual a distância percorrida.

**Exemplo 4.2.** Um móvel em movimento retilíneo uniforme tem velocidade de $-6m/s$.

(a) Calcule o deslocamento e o espaço percorrido pelo móvel no intervalo de $3s$ a $8s$.

(b) Esboce o gráfico da velocidade em função do tempo para esse movimento.

(c) Calcule a área da região limitada pela função velocidade, pelas retas $t = 3$, $t = 8$ e pelo eixo $t$.

(d) Compare o resultado do item $c$ com o resultado do item $a$. O que podemos concluir a respeito do deslocamento e do espaço percorrido pelo móvel?

**Resolução:**

**(a)** No instante $t_1 = 3s$ a posição do móvel é $s_1 = s_0 - 6.3 \Rightarrow s_1 = s_0 - 18$.

No instante $t_2 = 8s$ a posição do móvel é $s_2 = s_0 - 6.8 \Rightarrow s_2 = s_0 - 48$.

O deslocamento entre os instantes é $\Delta S = s_2 - s_1 = -48 + 18 = -30m$.

Já o espaço percorrido pelo móvel é $d = 30m$.

**(b)**

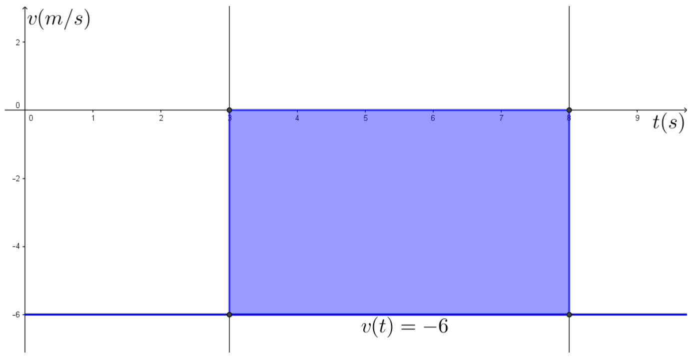



Figura 16: Movimento uniforme.

**(c)** Denotando a área da região por $A$ temos $A = 5.6 = 30$.

**(d)** Podemos concluir que a área da região $A$ é numericamente igual ao valor absoluto do módulo do deslocamento e que ela coincide com o espaço percorrido.

No movimento uniformemente variado (MUV) a velocidade é uma função do tipo $v(t) = at + b$ onde $a$ é a aceleração do móvel e $b$ é a velocidade inicial. Considere um móvel em MUV, com $b > 0$ e $a > 0$ . Calculemos a área sob o gráfico de $v$, limitado pelas retas $t = 0$, $t = \bar{t}$ e pelo eixo $t$.

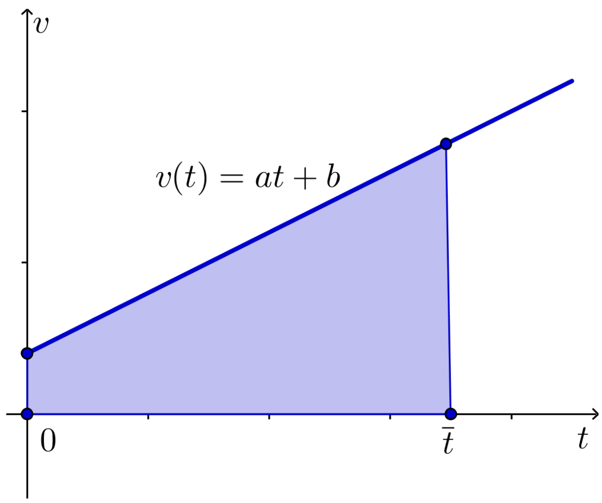



Figura 17: Calculando o deslocamento de um móvel em MUV.

Observe que temos um trapézio e sua área é dada por

$\dfrac{(v(\bar{t})+v(0)).(\bar{t}-0)}{2} = \dfrac{(a\bar{t}+b+b)(\bar{t})}{2} = \dfrac{a\bar{t}^2}{2} + \dfrac{2b\bar{t}}{2} = \dfrac{a\bar{t}^2}{2} + b\bar{t} = \Delta S$, ou seja, a área sob o gráfico de $v$ nos dá, numericamente, o módulo do deslocamento do móvel entre os instantes 0 e $\bar{t}$. Temos $b = v(0) = v_0$, assim, $\Delta S = s(\bar{t}) - s(0) = \dfrac{a\bar{t}^2}{2} + b\bar{t} \Rightarrow s(\bar{t}) = s_0 + v_0\bar{t} + \dfrac{a\bar{t}^2}{2}$, como aprendido pelos alunos.

Vamos utilizar outra estratégia, vamos nos aproximar da área real através de somas superiores tomando retângulos cada vez mais estreitos.

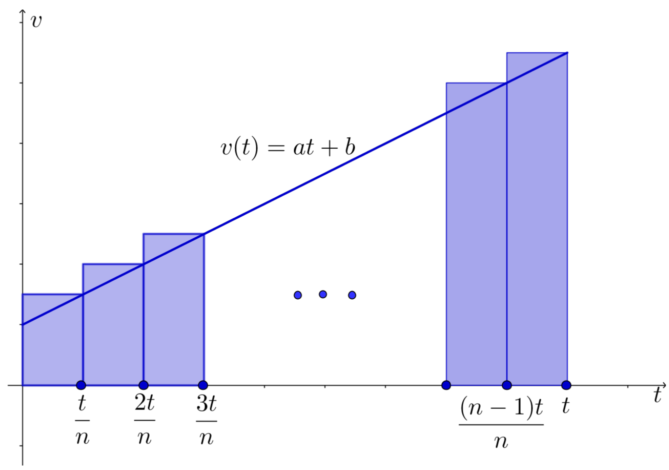



Figura 18: Calculando o deslocamento de um móvel em MUV.

Para isso, dividimos o intervalo $[0,t]$ em $n$ intervalos de mesmo comprimento e calculamos a soma das áreas de todos os retângulos de base $t/n$ e altura $v(\frac{kt}{n})$, $k = 1,2,3,\ldots,n$ como mostrado na Figura 18. Denotando por $A$ a soma dessas áreas temos:

$$A = \frac{t}{n}\left(\frac{at}{n}+b\right) + \frac{t}{n}\left(\frac{2at}{n}+b\right) + \frac{t}{n}\left(\frac{3at}{n}+b\right) + \ldots + \frac{t}{n}\left(\frac{(n-1)at}{n}+b\right) + \frac{t}{n}(at+b)$$

$$= \frac{t}{n}\left(\frac{at}{n} + \frac{2at}{n} + \frac{3at}{n} + \ldots + at\right) + bt = \frac{t}{n}\left[\frac{at}{n}(1+2+3+\ldots+n)\right] + bt$$

$$= \frac{t}{n}\left[\frac{at}{n}\frac{(n+1)n}{2}\right] + bt = \frac{at^2}{2}\frac{(n+1)}{n} + bt.$$

Observe que a medida que tornamos os retângulos mais estreitos, isto é, fazendo $n$ cada vez maior, $A$ se aproxima cada vez mais de $\frac{at^2}{2} + bt$ pois $\lim\limits_{n\to\infty} \frac{n+1}{n} = 1$. Assim, $\Delta S = \frac{at^2}{2} + bt$. Logo a área da região limitada pela função $v(t)$, pelo eixo $t$ e pelas retas $x = 0$ e $x = t$ nos dá, numericamente, o módulo do deslocamento do móvel.

Considere um móvel que se desloca com uma velocidade conforme mostra o gráfico a seguir

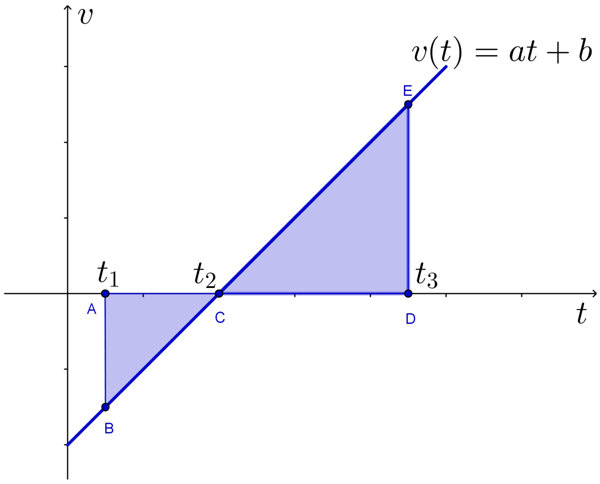



Figura 19: Neste movimento há mudança de sentido.

Sejam $s_1$, $s_2$ e $s_3$ a posição do móvel respectivamente nos instantes $t_1$, $t_2$ e $t_3$ e $s_0$ a posição inicial, isto é, a posição no instante $t = 0$. Mostremos que a diferença das áreas dos triângulos $ABC$ e $CDE$ é numericamente igual ao deslocamento do móvel entre os instantes $t_1$ e $t_3$.

De fato, tal diferença é dada por:
$\dfrac{(t_3-t_2)(at_3+b)}{2} - \left(-\dfrac{(t_2-t_1)(at_1+b)}{2}\right) = \dfrac{at_1t_2}{2} + \dfrac{bt_2}{2} - \dfrac{at_1^2}{2} - \dfrac{bt_1}{2} + \dfrac{at_3^2}{2} + \dfrac{bt_3}{2} - \dfrac{at_2t_3}{2}$
$-\dfrac{bt_2}{2} = \dfrac{at_3^2}{2} + \dfrac{bt_3}{2} - \dfrac{at_1^2}{2} - \dfrac{bt_1}{2} + \dfrac{at_1t_2}{2} - \dfrac{at_2t_3}{2} = \dfrac{at_3^2}{2} + bt_3 - \dfrac{at_1^2}{2} - bt_1 - \dfrac{bt_3}{2} + \dfrac{bt_1}{2} + \dfrac{at_1t_2}{2}$
$-\dfrac{at_2t_3}{2} = \dfrac{at_3^2}{2} + bt_3 - \dfrac{at_1^2}{2} - bt_1 + \dfrac{t_1}{2}(at_2+b) - \dfrac{t_3}{2}(at_2+b) = \dfrac{at_3^2}{2} + bt_3 - \dfrac{at_1^2}{2} - bt_1 + \dfrac{t_1}{2}v(t_2)$
$-\dfrac{t_3}{2}v(t_2)$. Observe que $v(t_2) = 0$ e ficamos com

$$\frac{at_3^2}{2} + bt_3 - \frac{at_1^2}{2} - bt_1 = (s_3 - s_0) - (s_1 - s_0) = s_3 - s_1 = \Delta S.$$

No intervalo $(t_1, t_2)$, o móvel se desloca contra a orientação positiva da trajetória pois neste intervalo $v < 0$. Há uma mudança de sentido no instante $t_2$ e a partir daí, o móvel se desloca a favor da orientação positiva da trajetória. Neste caso o espaço percorrido pelo móvel no intervalo $(t_1, t_3)$ será $(s_3 - s_2) + (s_1 - s_2)$. Podemos calcular o espaço percorrido pelo móvel, somando a área do triângulo $CDE$ com a área do triângulo $ABC$. De fato

$$\frac{(t_3-t_2)(at_3+b)}{2}+\left(-\frac{(t_2-t_1)(at_1+b)}{2}\right)=\frac{at_3^2}{2}+\frac{bt_3}{2}-\frac{at_2t_3}{2}-\frac{bt_2}{2}$$

$$-\left(\frac{at_1t_2}{2}+\frac{bt_2}{2}-\frac{at_1^2}{2}-\frac{bt_1}{2}\right)=\frac{at_3^2}{2}+\frac{bt_3}{2}-\frac{at_2t_3}{2}-\frac{bt_2}{2}-\frac{at_1t_2}{2}-\frac{bt_2}{2}+\frac{at_1^2}{2}+\frac{bt_1}{2}$$

$$=\frac{at_3^2}{2}+\frac{bt_3}{2}+\frac{at_1^2}{2}+\frac{bt_1}{2}-bt_2-\frac{at_2t_3}{2}-\frac{at_1t_2}{2}$$

$$=\frac{at_3^2}{2}+bt_3+\frac{at_1^2}{2}+bt_1-\frac{bt_3}{2}-\frac{bt_1}{2}-bt_2-\frac{at_2t_3}{2}-\frac{at_1t_2}{2}$$

$$=\frac{at_3^2}{2}+bt_3+\frac{at_1^2}{2}+bt_1-\frac{t_1}{2}(at_2+b)-\frac{t_3}{2}(at_2+b)-bt_2$$

$$=\frac{at_3^2}{2}+bt_3+\frac{at_1^2}{2}+bt_1-\frac{t_1}{2}(v(t_2))-\frac{t_3}{2}(v(t_2))-bt_2=\frac{at_3^2}{2}+bt_3+\frac{at_1^2}{2}+bt_1-bt_2$$

$$=s_3-s_0+s_1-s_0-bt_2-\frac{at_2^2}{2}+\frac{at_2^2}{2}=s_3-s_0+s_1-s_0-(s_2-s_0)+\frac{at_2^2}{2}$$

$$=(s_3-s_2)+\left(s_1-s_0-\frac{at_2^2}{2}-bt_2+at_2^2+bt_2\right)=(s_3-s_2)+(s_1-(s_2-s_0)-s_0+at_2^2+bt_2)$$
$$=(s_3-s_2)+(s_1-s_2)+t_2(at_2+b)=(s_3-s_2)+(s_1-s_2)+t_1(v(t_2))=(s_3-s_2)+(s_1-s_2).$$

**Exemplo 4.3.** Um móvel em movimento retilíneo uniformemente variado (MRUV) tem velocidade inicial $v_0=2m/s$ e aceleração $a=3m/s^2$.

(a) Calcule o deslocamento e o espaço percorrido pelo móvel entre os instantes $2s$ e $6s$.

(b) Esboce o gráfico da função velocidade em função do tempo para esse movimento.

(c) Calcule a área da região limitada pela função velocidade, pelas retas $t=2$, $t=6$ e pelo eixo $t$.

(d) Compare o resultado do item $c$ com o resultado do item $a$. O que podemos concluir a respeito do deslocamento e do espaço percorrido pelo móvel?

**Resolução:**

**(a)** Como o movimento é uniformemente variado, a função dos espaços em relação ao tempo é dada por $s(t)=s_0+v_0t+\frac{1}{2}at^2$.

No instante $t_1=2s$, a posição do móvel é $s_1=s_0+2.2+\frac{1}{2}.3.2^2\Rightarrow s_1=s_0+10$.

No instante $t_2=6s$, a posição do móvel é $s_2=s_0+2.6+\frac{1}{2}.3.6^2\Rightarrow s_2=s_0+66$.

Portanto, o deslocamento entre os instantes é $\Delta S = s_2 - s_1 = 56m$.

O espaço percorrido coincide com o deslocamento, portanto, $d = 56m$.

**(b)**

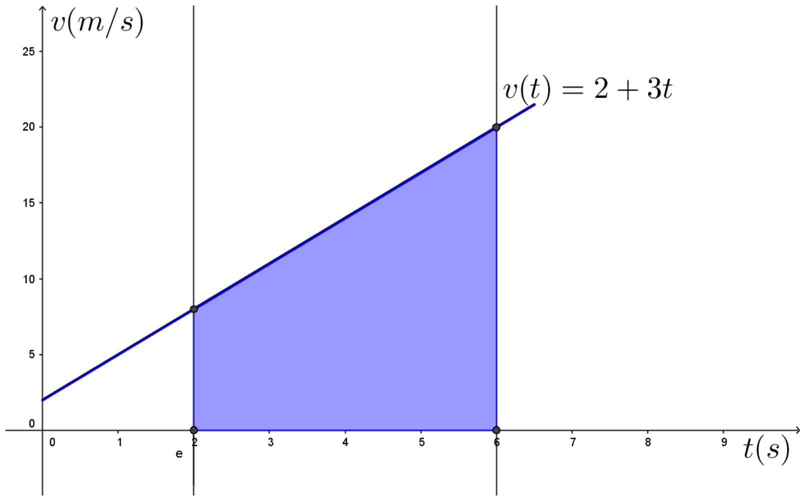



Figura 20: Movimento uniformemente variado.

**(c)** Trata-se da área de um trapézio. Denotando esta área por $A$ temos $A = (20+8).\frac{4}{2} \Rightarrow A = 56$.

**(d)** Concluímos que, a área da região $A$ é numericamente igual ao deslocamento e igual a distância percorrida.

**Exemplo 4.4.** Um móvel em MRUV tem velocidade inicial $v_0 = -4m/s$ e $a = 2m/s^2$.

(a) Calcule o deslocamento e o espaço percorrido pelo móvel entre os instantes $1s$ e $5s$.

(b) Esboce o gráfico da função velocidade em função do tempo para esse movimento.

(c) Calcule a área da região limitada pela função velocidade, pelas retas $t = 1$, $t = 5$ e pelo eixo $t$.

(d) Compare o resultado do item $c$ com o resultado do item $a$. O que podemos concluir a respeito do deslocamento e do espaço percorrido pelo móvel?

**Resolução:**

**(a)** No instante $t_1 = 1s$ a posição do móvel é $s_1 = s_0 - 4.1 + \frac{1}{2}.2.1^2 \Rightarrow s_1 = s_0 - 3$.

No instante $t_2 = 5s$ a posição do móvel é $s_2 = s_0 - 4.5 + \frac{1}{2}.2.5^2 \Rightarrow s_2 = s_0 + 5$.

O deslocamento do móvel é $\Delta S = s_2 - s_1 = 8m$.

Observe que neste movimento há mudança de sentido. A mudança de sentido ocorre em um instante $t$ tal que $v(t) = 0$. Para esse movimento, $v(t) = -4+2t$, fazendo $v(t) = 0$ obtemos $t = 2$. Logo a mudança de sentido ocorre no instante $t = 2s$.

A posição do móvel no instante $t = 2s$ é $s_3 = s_0 - 4.2 + \frac{1}{2}.2.2^2 \Rightarrow s_3 = s_0 - 4$.

O espaço percorrido pelo móvel é $d = 1+9 = 10m$.

**(b)**

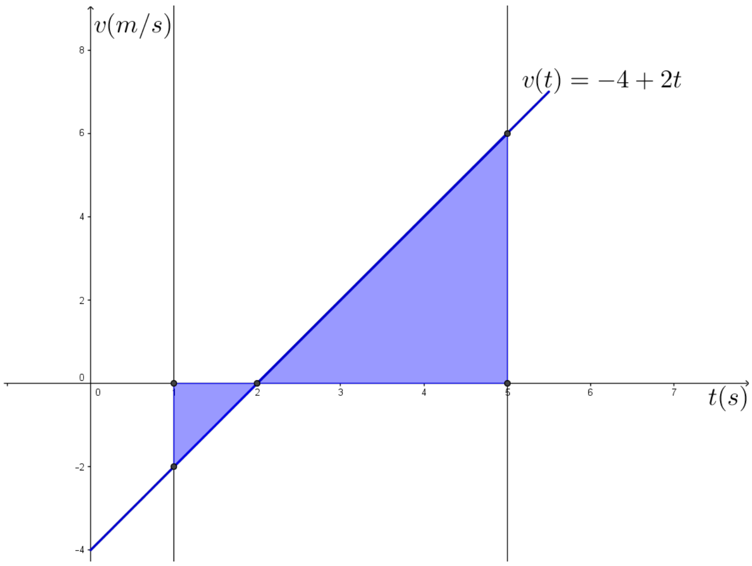



Figura 21: Movimento uniformemente variado.

**(c)** Seja $A_1$ a área limitada pela função velocidade, pelas retas $t = 1$, $t = 2$ e pelo eixo $t$ e $A_2$ a área limitada pela função velocidade, pelas retas $t = 2$, $t = 5$ e pelo eixo $t$. Temos

$A_1 = \frac{1.2}{2} = 1$ e $A_2 = \frac{3.6}{2} = 9$. A área desejada é $A = A_1 + A_2 = 10$.

**(d)** Concluímos que, a soma das áreas das regiões $A_1$ e $A_2$ é numericamente igual ao espaço percorrido e a diferença $A_2 - A_1$ é numericamente igual ao deslocamento.

Vimos que, se a velocidade de um móvel não muda de sinal então, a área limitada pelo gráfico da função velocidade, pelas retas verticais $t = t_a$, $t = t_b$ e pelo eixo horizontal $t$ é numericamente igual ao valor absoluto do módulo do deslocamento. Essa propriedade é válida em qualquer tipo de movimento, como mostraremos a seguir para o caso em que a velocidade entre os instantes $t_a$ e $t_b$ é positiva.

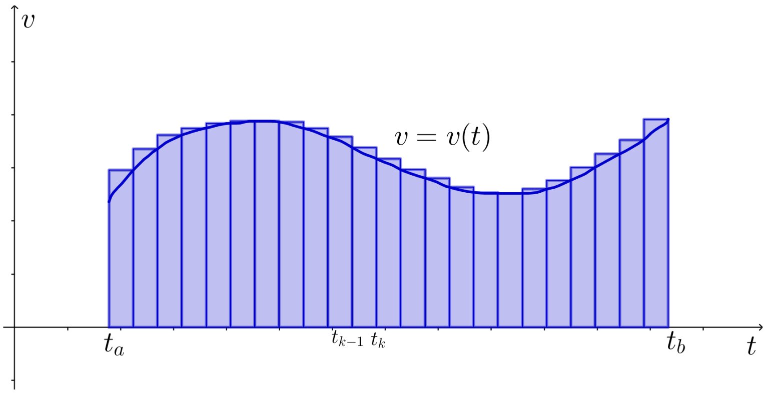



Figura 22: A área é numericamente igual ao deslocamento.

Suponha que entre os instantes de tempo $t_a$ e $t_b$, um determinado movimento tenha uma função velocidade $v(t)$ como mostrado na Figura 22. Seja $S_k = \Delta t.v(t_{k-1})$ a área do retângulo de base $(t_k - t_{k-1})$. Se o comprimento $(t_k - t_{k-1})$ é pequeno, a velocidade do movimento entre os instantes $t_{k-1}$ e $t_k$ é praticamente constante e a área do retângulo $S_k$ é aproximadamente o valor do deslocamento entre esses instantes. Particionando o intervalo $[t_a, t_b]$ como fizemos na Figura 22 e somando a área de cada um dos retângulos temos uma boa aproximação para o valor do deslocamento entre os instantes $t_a$ e $t_b$. Considerando retângulos cada vez mais estreitos, a soma das áreas de todos esses retângulos se aproximam cada vez mais do valor real do deslocamento. Logo

$$\Delta S = \int_{t_a}^{t_b} v(t)dt.$$

O caso em que a velocidade entre os instantes $t_a$ e $t_b$ é negativa se prova de modo análogo.

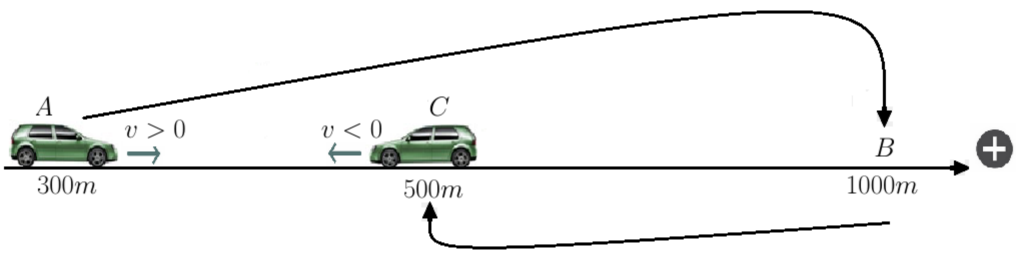



Figura 23: Espaço percorrido e deslocamento.

A Figura 23 ilustra a seguinte situação. Um móvel em movimento retilíneo parte do ponto $A$ no instante $t_a$ com velocidade $v > 0$ e dirigi-se até o ponto $B$ atingindo este ponto no instante $t_b$, depois, retorna atingindo o ponto $C$ no instante $t_c$. Assim o valor do deslocamento do móvel no intervalo de tempo $\Delta t = t_c - t_a$ foi $\Delta S = 700m - 500m = 200m$ enquanto o espaço percorrido foi de $700m + 500m = 1200m$. Isso pode ser interpretado graficamente da seguinte forma. $v(t) \geq 0$ em $[t_a, t_b]$ e $v(t) \leq 0$ em $[t_b, t_c]$.

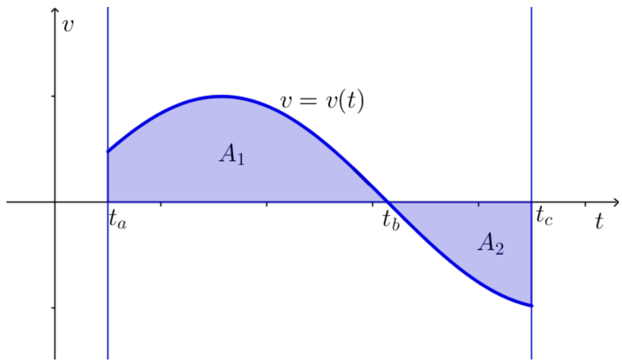



Figura 24: O deslocamento é dado por $A_1 - A_2$ e o espaço percorrido é dado por $A_1 + A_2$.

O valor do deslocamento entre os instantes $t_a$ e $t_c$ será

$$A_1 - A_2 = \int_{t_a}^{t_c} v(t)dt.$$

Já o espaço percorrido entre estes instantes será

$$\int_{t_a}^{t_c} |v(t)|dt = \int_{t_a}^{t_b} v(t)dt - \int_{t_b}^{t_c} v(t)dt = A_1 + A_2.$$

**Exemplo 4.5.** Com o uso do GeoGebra e dos conhecimentos obtidos nas atividades anteriores, vamos calcular aproximações para o deslocamento e o espaço percorrido por um móvel em movimento retilíneo com função velocidade dada por $v(t) = -t^2 + t$, $t \geq 0$, entre os instantes $t = 0s$ e $t = 2s$ (Considere a velocidade em $m/s$).

Vamos seguir os mesmos passos já mencionados anteriormente, porém, ao criar o controle deslizante, determinamos o valor mínimo igual a 1 e máximo igual a 100 para que tenhamos maior precisão nas aproximações.

Na figura a seguir, criamos aproximações que subestimam a área da região limitada pelas retas $t = 0$, $t = 2$, pelo eixo $t$ e pelo gráfico da função $v(t)$. No intervalo $[0,1]$ aproximamos por somas inferiores, no intervalo $[1,2]$ por somas superiores. Escolhendo o valor 20 para o controle deslizante, o software nos dá uma aproximação de $0,15$ para o intervalo $[0,1]$ e de $0,78$ para o intervalo $[1,2]$.

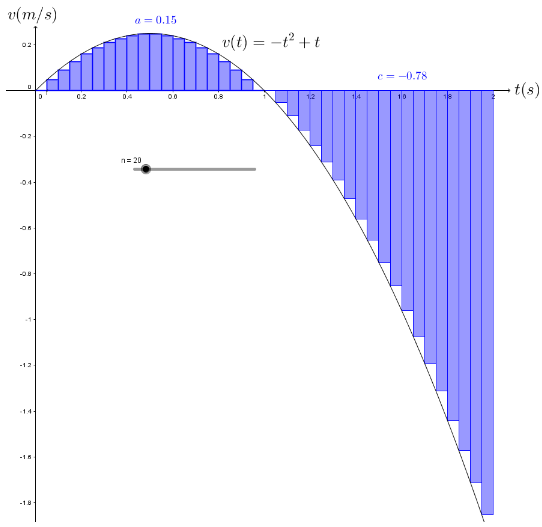



Figura 25: Gráfico da função $v(t) = -t^2 + t$ com valor 20 para o controle deslizante.

Assim, o deslocamento do móvel entre os instantes $t_1 = 0s$ e $t_2 = 2s$ é aproximadamente igual a $0,15 - 0,78 = -0,63m$. Já o espaço percorrido pelo móvel é de aproximadamente $0,78 + 0,15 = 0,93m$.

Escolhendo o valor $n = 100$ no controle deslizante, obtemos para o deslocamento, o valor $-0,66m$ e para o espaço percorrido $0,98m$.

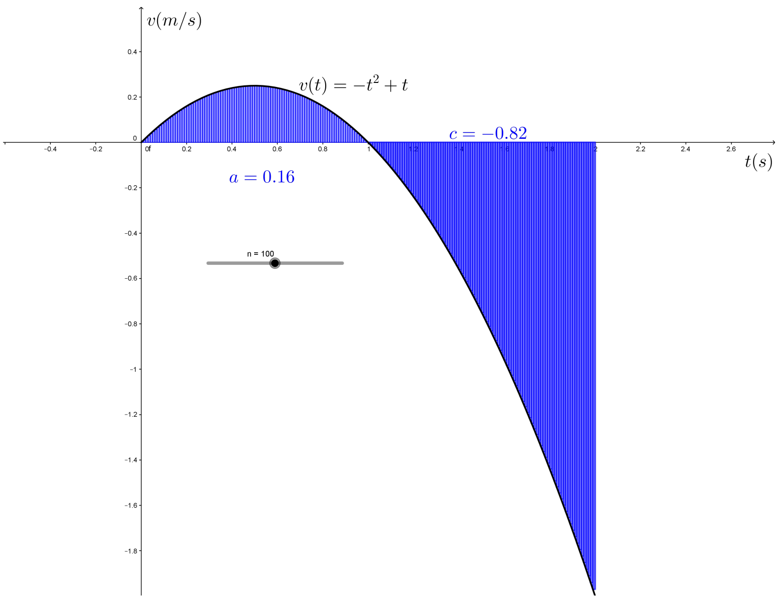



Figura 26: Gráfico da função $v(t) = -t^2 + t$ com valor 100 para o controle deslizante.

A seguir propomos uma série de exercíos que podem ser aplicados em sala de aula para fixação do conteúdo. Os exercícios podem ser usados como modelo para a criação de outros, o professor também pode modificá-los de acordo com sua preferência.

**Exercício proposto 1:** Considere um móvel em MRU e seja $v$ a velocidade desenvolvida pelo móvel. Responda os itens (a), (b), (c) e (d) para cada caso a seguir.

(I) $v(t) = v = 12m/s$, $t_1 = 2s$ e $t_2 = 8s$.

(II) $v(t) = v = -5m/s$ , $t_1 = 3s$ e $t_2 = 7s$.

(III) $v(t) = v = -10m/s$, $t_1 = 4s$ e $t_2 = 6s$.

(IV) $v(t) = v = 15m/s$, $t_1 = 1s$ e $t_2 = 9s$.

(a) Calcule o deslocamento e o espaço percorrido pelo móvel entre os instantes $t_1$ e $t_2$.

(b) Esboce o gráfico da função velocidade em função do tempo para esse movimento.

(c) Calcule a área da região limitada pela função velocidade, pelas retas $t = t_1$, $t = t_2$ e pelo eixo $t$.

(d) Compare o resultado do item $c$ com o resultado do item $a$. O que podemos concluir a respeito do deslocamento e do espaço percorrido pelo móvel?

**Exercício proposto 2:** Considere um móvel em MRUV. Seja $v_0$ a velocidade inicial e $a$ a aceleração do móvel. Responda os itens (a), (b), (c) e (d) para cada caso.

(I) $v_0 = 5m/s$, $a = 2m/s^2$, $t_1 = 3s$ e $t_2 = 5s$.

(II) $v_0 = -6m/s$, $a = 4m/s^2$, $t_1 = 1s$ e $t_2 = 3s$.

(III) $v_0 = 10m/s$, $a = -2m/s^2$, $t_1 = 4s$ e $t_2 = 9s$.

(IV) $v_0 = -3m/s$, $a = -5m/s^2$, $t_1 = 4s$ e $t_2 = 8s$.

(a) Calcule o deslocamento e o espaço percorrido pelo móvel entre os instantes $t_1$ e $t_2$.

(b) Esboce o gráfico da função velocidade em função do tempo para esse movimento.

(c) Calcule a área da região limitada pela função velocidade, pelas retas $t = t_1$, $t = t_2$ e pelo eixo $t$.

(d) Compare o resultado do item $c$ com o resultado do item $a$. O que podemos concluir a respeito do deslocamento e do espaço percorrido pelo móvel?

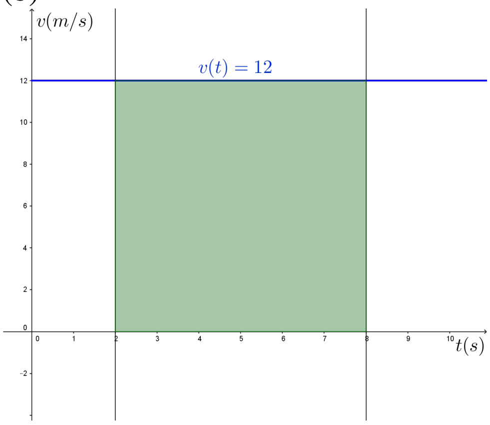
**Exercício proposto 3:** Considere um móvel em MRUV com função velocidade dada por $v(t) = t^2 - 2t - 3$, $t \geq 0$. Calcule aproximações para o deslocamento e para o espaço percorrido pelo móvel entre os instantes $t = 0s$ e $t = 4s$ (Considere a velocidade em $m/s$).

## Respostas dos Exercícios Propostos

**1.**

**(I)** (a) $\Delta S = 72m$, $d = 72m$.

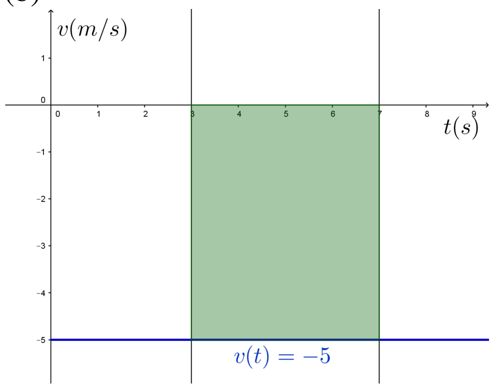
(b)




(c) $A = 72$.
(d) Pode-se concluir que $A = \Delta S = d$.

**(II)** (a) $\Delta S = -20m$, $d = 20m$.

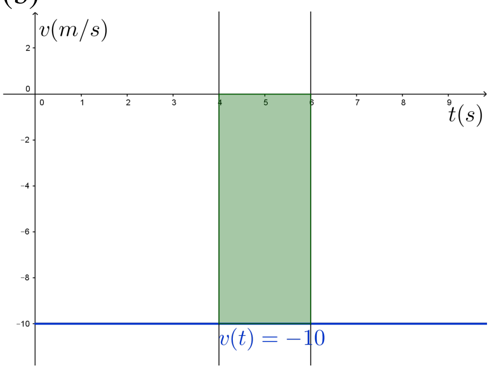
(b)




(c) $A = 20$.
(d) Pode-se concluir que $A = |\Delta S| = d$.

**(III)** (a) $\Delta S = -20m$, $d = 20m$.

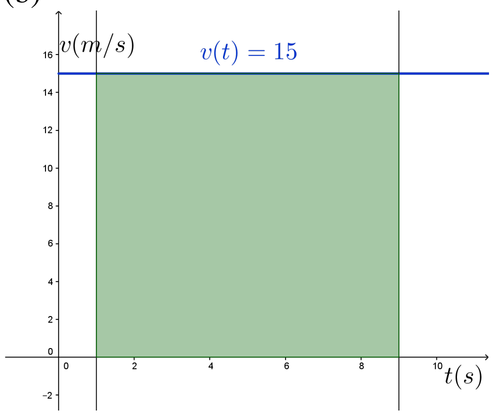
(b)




(c) $A = 20$.
(d) Pode-se concluir que $A = |\Delta S| = d$.

**(IV)** (a) $\Delta S = 120m$, $d = 120m$.

(b)

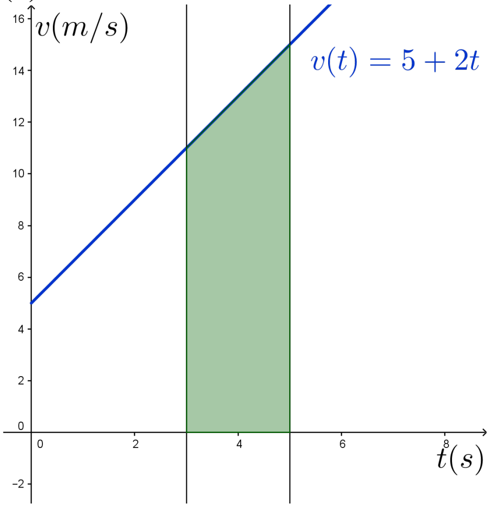



(c) $A = 120$.
(d) Pode-se concluir que $A = \Delta S = d$.

**2.**

**(I)** (a) $\Delta S = 26m$, $d = 26m$.

(b)

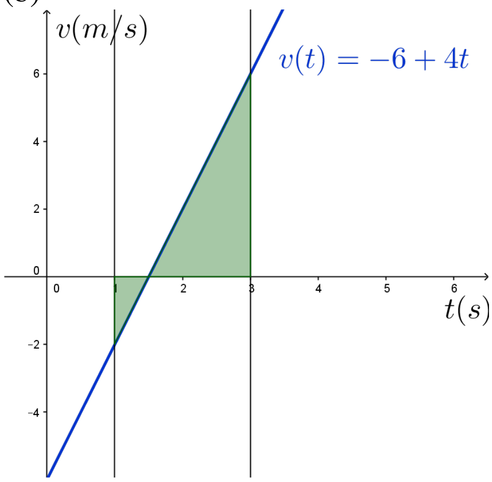



(c) $A = 26$.

(d) Pode-se concluir que $A = \Delta S = d$.

**(II)** (a) $\Delta S = 4m$, $d = 5m$.

(b)

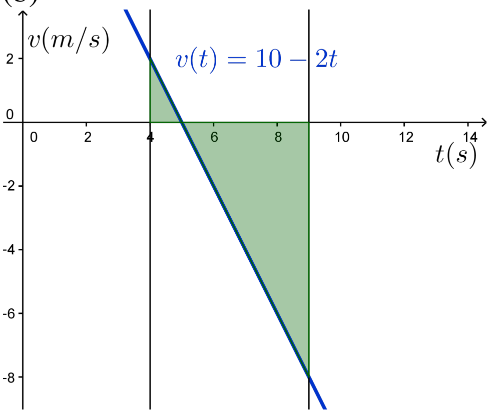



(c) $A = 5$.

(d) Pode-se concluir que $\Delta S = A_2 - A_1$ e $d = A_2 + A_1 = A$.

**(III)** (a) $\Delta S = -15m$, $d = 17m$.

(b)

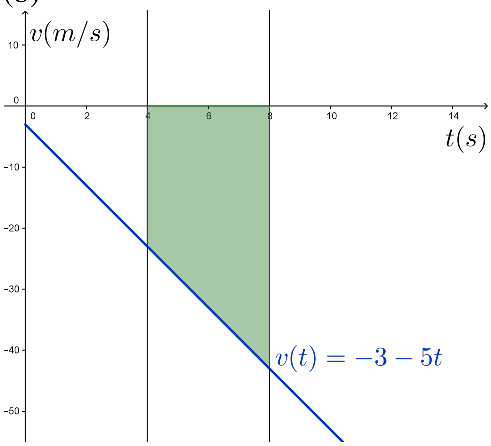



(c) $A = 17$.

(d) Pode-se concluir que $\Delta S = A_1 - A_2$ e $d = A_1 + A_2 = A$.

**(IV)** (a) $\Delta S = -132m$, $d = 132m$.

(b)




(c) $A = 132$.

(d) Pode-se concluir que $A = |\Delta S| = d$.

**3.** Utilizando o software GeoGebra temos $\Delta S \approx -6,67m$ e $d \approx 11,32m$.

# REFERÊNCIAS

[1] ÁVILA, G. *Introdução À Análise Matemática*. 2. ed. revista: Edgard Blucher LTDA, 2003. Citado na página 21.

[2] BURK, F.E. *A Garden of Integrals*. California State University at Chico: The Mathematical Association of America, 2007. Citado na página 89.

[3] GUIDORIZZI, H.L. *Um curso de cálculo*. v. 1, 5. ed. Rio de Janeiro: LTC, 2013. Citado na página 36.

[4] HALLIDAY, D.; RESNICK, R.; WALKER, J. *Fundamentos de Física: Mecânica*. v. 1, 9. ed. : LTC, 2012. Citado na página 96.

[5] LIMA, E.L. *Análise Real: Funções de uma variável*. v. 1, 12. ed. Rio de Janeiro-RJ: Coleção Matemática Universitária: IMPA, 2014. Citado na página 21.

[6] MAGALHÃES, M.N. *Probabilidade e Variáveis Aleatórias*. 3. ed. São Paulo: Universidade de São Paulo, 2013. Citado 3 vezes nas páginas 61, 69 e 86.

[7] NIELSEN, O.A. *An Introduction to Integration and Measury Theory*. New york: Wile, c1997. Citado 4 vezes nas páginas 77, 78, 80 e 81.

[8] RAMALHO, F.; FERRARO, N.G; TOLEDO, P.A. *Os Fundamentos da Física*. v. 1, 9. ed. Moderna, 2007. Citado na página 96.

[9] ROSS, S. *Probabilidade - Um curso moderno com aplicações*. 8. ed. Porto Alegre: Bookman, 2010. Citado 2 vezes nas páginas 61 e 69.

[10] RUDIN, W. *Principles of Mathematical Analysis*. 3. ed. McGrawHill, 1976. Citado na página 78.

[11] THOMAS, G.B.; GIORDANO, W.H. *Cálculo*. v. 1, 11. ed. São Paulo: Addison Wesley, 2009. Citado na página 38.