



UNIVERSIDADE ESTADUAL PAULISTA "JÚLIO DE MESQUITA FILHO"
Instituto de Geociências e Ciências Exatas
Câmpus de Rio Claro

Um Estudo Sobre a Teoria de Sturm-Liouville

Valterlan Atanasio de Souza

Dissertação apresentada ao Programa de Pós-Graduação em Matemática como requisito parcial para a obtenção do grau de Mestre

Orientadora
Profa. Dra. Marta Cilene Gadotti

2016

517.5 Souza, Valterlan Atanasio de
S729e Um estudo sobre a Teoria de Sturm-Liouville / Valterlan
Atanasio de Souza. - Rio Claro, 2016
139 f. : il., gráfs.

Dissertação (mestrado) - Universidade Estadual Paulista,
Instituto de Geociências e Ciências Exatas
Orientadora: Marta Cilene Gadotti

1. Análise funcional. 2. Teoria espectral. 3. Operadores
compactos. 4. Problema de Sturm-Liouville. 5. Equações
diferenciais. 6. Espaços de Hilbert. I. Título.

TERMO DE APROVAÇÃO

Valterlan Atanasio de Souza
UM ESTUDO SOBRE A TEORIA DE STURM-LIOUVILLE

Dissertação APROVADA como requisito parcial para a obtenção do grau de Mestre no Curso de Pós-Graduação em Matemática do Instituto de Geociências e Ciências Exatas da Universidade Estadual Paulista “Júlio de Mesquita Filho”, pela seguinte banca examinadora:

Profa. Dra. Marta Cilene Gadotti
Orientadora

Profa. Dra. Suzete Maria Silva Afonso
Departamento de Matemática - UNESP - Rio Claro (SP)

Profa. Dra. Katia Andreia Gonçalves de Azevedo
Departamento de Computação e Matemática - USP - Ribeirão Preto (SP)

Rio Claro, 12 de Dezembro de 2016

*À minha mãe Vera, à memória de meu pai Lialdo,
à minha esposa Tatiane, à minha filha Júlia e ao Pedro que chegará em 2017.*

Agradecimentos

Agradeço primeiramente a Deus por todas as bênçãos em minha vida e por me dar força nos momentos mais difíceis.

Agradeço minha mãe Vera por todo o sacrifício em me manter na universidade e por seu amor. A minha irmã Erika pelo seu eterno carinho. A minha esposa Tatiane por toda a força, paciência e um amor incondicional. A minha filha Júlia pela alegria nesse tempo.

A minha querida amiga e professora Dra. Marta Cilene Gadotti, por todo o conhecimento e simpatia transmitidos não só nesse projeto como também na graduação. Obrigado pela paciência e compreensão com minha falta de tempo.

Aos professores do Departamento de Matemática que contribuíram para minha formação, em especial ao professor Dr. Vanderlei Marcos do Nascimento por sua grande contribuição e amizade. Aos funcionários do Departamento de Matemática, em especial a Ana e Maria Elisa.

Aos meus eternos amigos de graduação: Glalco, Lucas Mazzi, Givanildo e Vinicius Wasques, por fazerem parte dos momentos mais especiais de minha vida.

E a todos, que de algum modo contribuíram para a realização deste trabalho.

Costumamos dizer que amigos de verdade são os que estão ao seu lado em momentos difíceis... Mas não!

*Amigos verdadeiros são os que suportam a tua felicidade.
Porque em um momento difícil qualquer um se aproxima de você.
Mas o seu inimigo jamais suportaria a sua felicidade.*

Padre Fábio de Melo

Resumo

Este texto aborda os principais resultados sobre a Teoria de Sturm-Liouville assim como os pré-requisitos necessários para construí-los, entre eles o Teorema Espectral para Operadores Compactos e a Teoria de Fredholm. Também são apresentados alguns exemplos e uma aplicação envolvendo uma equação diferencial parcial que modela o problema da corda vibrante.

Palavras-chave: Teoria Espectral, Operadores Compactos, Problema de Sturm-Liouville, Equações Diferenciais, Espaços de Hilbert.

Abstract

This research approaches the main results on the Sturm-Liouville Theory, as well the necessary prerequisites for constructing them, including the Spectral Theorem for Compact Operators and Fredholm Theory. It is also presented some examples and an application involving a partial differential equation that models the vibrating string problem.

Keywords: Spectral Theory, Compact Operators, Sturm-Liouville Problem, Differential Equation, Hilbert Spaces.

Sumário

1	Introdução	15
2	Teoria Básica Sobre Espaços Métricos e Banach	17
2.1	Espaço Vetorial	17
2.2	Espaços Métricos	22
2.2.1	Conjuntos Abertos, Fechados e Limitados	23
2.2.2	Sequências e Completude	25
2.3	Espaços Métricos Compactos	28
2.3.1	O Teorema de Ascoli-Arzelá	31
2.4	Espaços Normados e de Banach	32
2.4.1	O Teorema de Hahn-Banach	34
2.4.2	A Desigualdade de Hölder	42
2.4.3	Somabilidade em Espaços Normados	44
3	Espaços de Hilbert	47
3.1	Complemento Ortogonal e Soma Direta	50
3.2	Conjunto Ortonormal	56
3.3	Representação de Funcionais Lineares em Espaços de Hilbert	64
3.3.1	Forma Sesquilinear	66
4	Teoria dos Operadores	69
4.1	Operadores Lineares Limitados	70
4.2	Operadores Compactos	73
4.3	Operador Hilbert-Adjunto	75
4.4	Operadores Autoadjunto, Unitário e Normal	80
5	Teoria Espectral	89
5.1	Teoria Espectral dos Operadores Compactos	89
5.2	A Equação Integral de Fredholm com Núcleo Hermitiano	97
6	O Problema de Sturm-Liouville	105
	Referências	131

A	Método de Separação de Variáveis e Séries de Fourier	133
A.1	Separação de Variáveis e Séries de Fourier	135
A.1.1	O Método de Separação de Variáveis	135

1 Introdução

Este trabalho tem por objetivo apresentar e aplicar os resultados referentes à Teoria de Sturm-Liouville e, para isso, foi preciso desenvolver a Teoria Espectral dos Operadores Hermitianos Compactos em espaços de dimensão não finita. Esse estudo é importante pois inúmeros problemas que são modelados por equações diferenciais parciais, após o uso do método de separação de variáveis, recaem em problemas de valores de contorno envolvendo equações diferenciais ordinárias de segunda ordem. Deste modo, a teoria de Sturm-Liouville fornece uma solução do problema sob certas condições.

Para tanto, o capítulo 2 aborda assuntos como espaço vetorial, espaços métricos e espaços de Banach. Destaca-se o conceito de base de um espaço vetorial e faz-se um breve estudo sobre espaços métricos, do qual aborda-se o essencial sobre sequências e completude e, não menos importante, um estudo, mesmo que breve, sobre espaços métricos compactos dando destaque para os espaços métricos relativamente compactos, com o objetivo de apresentar o Teorema de Ascoli-Arzelá, que será de utilidade no capítulo de operadores. Por fim, tem-se os Teoremas de Hahn-Banach sobre funcionais lineares e no final do capítulo encontra-se a desigualdade de Hölder e alguns resultados pertinentes sobre somabilidade em espaços normados.

No capítulo 3 tem-se a definição de um espaço de Hilbert e um espaço pré-hilbertiano, mas o objetivo maior do capítulo está em enunciar e demonstrar o Teorema da Representação de Riesz que é usado no capítulo seguinte, no estudo dos operadores Adjuntos. Para tanto, tem-se um breve estudo sobre complemento ortogonal, soma direta e conjunto ortonormal, tendo como destaque a Desigualdade de Bessel e alguns resultados sobre convergências de séries em espaços de Hilbert ou simplesmente em espaços pré-hilbertianos. Conclui-se o capítulo com o Teorema de Riesz e o Teorema da Representação de Riesz.

O capítulo 4, Teoria dos Operadores, aborda alguns resultados sobre operadores lineares limitados e compactos, dando destaque para um operador linear definido sobre um núcleo contínuo, onde tal operador é de grande importância no estudo do Problema de Sturm-Liouville. Ainda neste capítulo, faz-se um estudo sobre operadores Hilbert-

Adjunto (operadores definidos entre dois espaços de Hilbert), Autoadjunto, Unitário e Normal, donde destaca-se a existência do adjunto para operadores lineares limitados definidos em espaços de Hilbert e propriedades destes operadores. O capítulo 5 traz a prova do Teorema Espectral para operadores compactos e apresenta a Equação Integral de Fredholm com núcleo hermitiano como uma das consequências do Teorema Espectral.

Por fim, o capítulo 6 engloba a teoria de Sturm-Liouville, de modo a apresentar o Problema de Sturm-Liouville e resultados que buscam uma solução do problema. Tais resultados têm como consequência final dois importantes teoremas, o Teorema da função de Green e o teorema que caracteriza, por meio de seis itens, o Problema de Sturm-Liouville. Há, ainda, alguns exemplos do problema e uma aplicação na física.

Também, para o leitor interessado, no apêndice A encontra-se uma breve introdução às equações diferenciais parciais e o método de separação de variáveis. O leitor pode perceber que os capítulos 2 e 3, a rigor, podem ser considerados capítulos preliminares, uma vez que a grande maioria dos assuntos abordados nestes capítulos não tem uma ligação direta com o problema de Sturm-Liouville de modo que o objetivo destes capítulos está em apresentar alguns resultados da Análise Funcional.

2 Teoria Básica Sobre Espaços Métricos e Banach

Este capítulo tem por objetivo apresentar alguns resultados sobre Espaços Métricos, Normados e de Banach. Para os resultados de Álgebra Linear utilizou-se das referências [1], [2] e [3]. Já na teoria de Espaços Métricos foram utilizadas as referências [4], [5] e [3], por fim, para os Espaços Normados e de Banach utilizou-se as [3] e [7].

2.1 Espaço Vetorial

Espaços vetoriais desempenham um papel importante em muitos ramos da matemática e suas aplicações. Em diversos problemas práticos (e teóricos) temos um conjunto X cujos elementos podem ser vetores em espaços tridimensionais, ou sequência de números, ou ainda funções; estes elementos podem ser adicionados ou multiplicados por uma constante (número) de uma forma natural e o resultado ainda será um elemento de X . Tais situações sugerem o conceito de espaço vetorial como definido abaixo. Nessa definição envolverá um corpo geral \mathbb{K} mas, em análise funcional, \mathbb{K} será \mathbb{R} ou \mathbb{C} . Os elementos de \mathbb{K} são chamados escalares, portanto, no nosso caso eles serão números reais ou complexos.

Definição 2.1. *Um **espaço vetorial** sobre um corpo \mathbb{K} é um conjunto não vazio X de elementos x, y, \dots (chamados de vetores) munido de duas operações algébricas. Estas operações são chamadas adição de vetores e multiplicação de vetores por escalar, isto é, por um elemento de \mathbb{K} .*

Adição de vetores: *associa a cada par x, y de vetores um vetor $x + y$ chamado a soma de x e y de tal modo que essa operação seja comutativa e associativa, isto é, para todos os vetores tem-se*

$$x + y = y + x$$

$$x + (y + z) = (x + y) + z$$

além do mais, existe um vetor 0 , chamado de vetor nulo e para cada vetor x existe um vetor $-x$, chamado oposto, tal que para todo vetor tem-se

$$x + 0 = x$$

e

$$x + (-x) = 0.$$

Multiplicação por escalar: associa a cada vetor x e um escalar α um vetor αx (também se escreve $x\alpha$), chamado de produto de α por x , de tal modo que para todos vetores x, y e escalares α, β tem-se

$$\alpha(\beta x) = (\alpha\beta)x$$

$$1x = x$$

e a lei distributiva

$$\alpha(x + y) = \alpha x + \alpha y$$

$$(\alpha + \beta)x = \alpha x + \beta x$$

A partir da definição observa-se que adição de vetores é um aplicação de $X \times X$ em X , já a multiplicação por escalar é uma aplicação de $\mathbb{K} \times X$ em X .

O corpo \mathbb{K} é chamado de **corpo de escalares** do espaço vetorial X , e X é chamado um **espaço vetorial real** se $\mathbb{K} = \mathbb{R}$, e um **espaço vetorial complexo** se $\mathbb{K} = \mathbb{C}$.

Exemplo 2.1. O corpo dos números complexos \mathbb{C} pode ser considerado como um espaço vetorial sobre o corpo \mathbb{R} dos números reais. De maneira mais geral, seja \mathbb{R} o corpo dos números reais e seja V o conjunto das n -úplas $x = (x_1, x_2, \dots, x_n)$ onde x_1, \dots, x_n são números complexos. Define-se a adição de vetores e a multiplicação por escalar, respectivamente, por

$$x + y = (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n) \quad \alpha x = (\alpha x_1, \alpha x_2, \dots, \alpha x_n).$$

Desta forma obtém-se um espaço vetorial sobre \mathbb{R} que é diferente do espaço \mathbb{C}^n e do espaço \mathbb{R}^n .

Definição 2.2. Um **subespaço** de um espaço vetorial X é um subconjunto não vazio $Y \subset X$ tal que para todos vetores y_1, y_2 em Y e todos escalares α, β tem-se $\alpha y_1 + \beta y_2$ em Y .

Note que Y é, por si próprio, um espaço vetorial com as operações de adição de vetores e multiplicação por escalar induzidas de X . Com efeito, uma vez que X é um espaço vetorial sobre \mathbb{K} as propriedades comutativa e associativa de vetores, bem como as propriedades da multiplicação restringidas aos elementos de Y são satisfeitas. Como Y é não vazio, existe um vetor w em Y , logo $(-1)w + 1w = 0$ está em Y . Assim, se x é um vetor qualquer em Y e α um escalar qualquer, o vetor $\alpha x = \alpha x + 0$ está em Y . Em particular, $(-1)x = -x$ está em Y . Então se x e y estão em Y , então $x + y = 1x + y$ está em Y .

Se X é um espaço vetorial qualquer, X é um subespaço de X ; o subconjunto $\{0\}$ de X também é um subespaço de X , denominado o **subespaço nulo** de X .

Exemplo 2.2. Em \mathbb{K}^n , o conjunto das n -úplas (x_1, x_2, \dots, x_n) com $x_1 = 0$ é um subespaço; contudo, o conjunto das n -úplas com $x_1 = 1 + x_2$ não é um subespaço ($n \geq 2$).

Uma combinação linear de vetores x_1, x_2, \dots, x_m de um espaço vetorial X é uma expressão da forma

$$\alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_m x_m$$

onde os coeficientes $\alpha_1, \alpha_2, \dots, \alpha_m$ são quaisquer escalares. Assim um vetor x em X é dito uma **combinação linear** dos vetores x_1, x_2, \dots, x_m se existirem escalares $\alpha_1, \alpha_2, \dots, \alpha_m$ em \mathbb{K} tais que

$$x = \alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_m x_m = \sum_{i=1}^m \alpha_i x_i$$

Para qualquer subconjunto não vazio $M \subset X$, consideremos o conjunto de todas as combinações lineares de vetores de M , tal conjunto é um subespaço Y de X , e é chamado de **subespaço gerado** por M . Escrevemos $Y = [M]$. Para mostrar que Y é um subespaço de X , sejam x e y vetores quaisquer de Y e α, β escalares. Se $x = \alpha_1 x_1 + \dots + \alpha_m x_m$ e $y = \beta_1 y_1 + \dots + \beta_n y_n$ são elementos de Y , então $\alpha x + \beta y = \alpha(\alpha_1 x_1 + \dots + \alpha_m x_m) + \beta(\beta_1 y_1 + \dots + \beta_n y_n) = (\alpha\alpha_1)x_1 + \dots + (\alpha\alpha_m)x_m + (\beta\beta_1)y_1 + (\beta\beta_n)y_n$ também é um elemento de Y .

Definição 2.3. *Sejam X um espaço vetorial sobre \mathbb{K} e M um subconjunto de X .*

- (a) *Diz-se que M é **linearmente independente** (ou **l.i.**) se $\alpha_1 x_1 + \dots + \alpha_n x_n = 0$, para $x_i \in M$ e $\alpha_i \in \mathbb{K}$, $i = 1, \dots, n$, implica que $\alpha_1 = \dots = \alpha_n = 0$.*
- (b) *O conjunto M é chamado **linearmente dependente** (ou **l.d.**) se não for linearmente independente.*

Usa-se o conceito de dependência e independência linear para definir a dimensão de um espaço vetorial, começando como a seguir.

Definição 2.4. *Um espaço vetorial X é dito ser de **dimensão finita** se existir um inteiro positivo n tal que X contenha um conjunto linearmente independente de n vetores ao passo que qualquer conjunto de $n+1$ ou mais vetores de X seja linearmente dependente. Nessas condições n é chamado de dimensão de X , escreve-se $n = \dim X$. Por definição, $X = \{0\}$ é de dimensão finita e $\dim X = 0$. Se X não é de dimensão finita, então ele é dito de **dimensão infinita**.*

Se $\dim X = n$, uma n -pla linearmente independente de vetores de X é chamada uma **base** para X . Se $\{e_1, \dots, e_n\}$ é uma base para X , cada elemento x de X tem uma única representação como uma combinação linear de vetores da base, isto é,

$$x = \alpha_1 e_1 + \dots + \alpha_n e_n.$$

De modo geral, se X é um espaço vetorial qualquer, não necessariamente de dimensão finita, e B é um subconjunto linearmente independente de X que gera X , então B é chamado uma base (ou **Base de Hamel**) para X . Portanto, se B é uma base para X então cada elemento não nulo x de X tem uma única representação como uma combinação linear de (quantidade finita) elementos de B com escalares não todos nulos como coeficientes.

Exemplo 2.3. Seja M uma matriz quadrada, de ordem n , inversível com elementos no corpo \mathbb{K} . Então as colunas M_1, M_2, \dots, M_n de M formam uma base do espaço das matrizes colunas $M_{n \times 1}(\mathbb{K})$. De fato, seja $X = (x_{ij})_{n \times 1}$ uma matriz coluna, então

$$MX = x_{11}M_1 + x_{21}M_2 + \dots + x_{n1}M_n.$$

Como M é inversível, a equação $MX = 0$ admite somente a solução trivial $X = 0$, assim $\{M_1, \dots, M_n\}$ é um conjunto linearmente independente. Para mostrar que tal conjunto gera $M_{n \times 1}(\mathbb{K})$ considere Y uma matriz coluna arbitrária. Se $X = M^{-1}Y$, então $Y = MX$, ou seja,

$$Y = x_{11}M_1 + x_{21}M_2 + \dots + x_{n1}M_n.$$

Portanto, $\{M_1, M_2, \dots, M_n\}$ é uma base de $M_{n \times 1}(\mathbb{K})$ concluindo que $\dim M_{n \times 1}(\mathbb{K}) = n$.

Exemplo 2.4. Sejam \mathbb{C} o conjunto dos números complexos e V o espaço das funções polinomiais sobre \mathbb{C} . Vale lembrar que essas funções são as funções de \mathbb{C} em \mathbb{C} da forma

$$f(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n,$$

para algum natural n . Considere $f_k(x) = x^k$, $k \in \{0, 1, 2, \dots\}$. O conjunto $\{f_0, f_1, f_2, \dots\}$ é uma base de V . De fato, note que $\{f_0, f_1, f_2, \dots\}$ gera V porque a função $f(x) = a_0 + a_1x + \dots + a_nx^n$ é

$$f = a_0f_0 + a_1f_1 + \dots + a_nf_n.$$

Para verificar que o conjunto $\{f_0, f_1, f_2, \dots\}$ é linearmente independente basta mostrar que cada um de seus subconjuntos finitos é linearmente independente. Assim, é suficiente mostrar que, para cada n , o conjunto $\{f_{k_0}, f_{k_1}, f_{k_2}, \dots, f_{k_n}\}$ é linearmente independente. Para tanto, suponha que

$$a_{k_0}f_{k_0} + a_{k_1}f_{k_1} + \dots + a_{k_n}f_{k_n} = 0.$$

Isto é,

$$a_{k_0}x^{k_0} + a_{k_1}x^{k_1} + \dots + a_{k_n}x^{k_n} = 0,$$

para todo $x \in \mathbb{C}$. Mas isto significa que para cada x^{k_j} , com $j \in \{0, 1, 2, \dots, n\}$, seu respectivo coeficiente é identicamente nulo. Então, $a_{k_0} = a_{k_1} = \dots = a_{k_n} = 0$. Logo

$\{f_{k_0}, f_{k_1}, f_{k_2}, \dots, f_{k_n}\}$ é uma base infinita de V . Observe que V não pode ser de dimensão finita, pois considerando um número finito de funções polinomiais g_1, g_2, \dots, g_r haverá uma maior potência de x aparecendo (com coeficiente não nulo) em $g_1(x), \dots, g_r(x)$. Se este expoente for k , então $f_{k+1}(x) = x^{k+1}$ não estará no espaço gerado por g_1, \dots, g_r . Logo, V não é de dimensão finita.

O teorema a seguir mostra que todo espaço vetorial, diferente do nulo, possui uma base.

Teorema 2.1. *Todo espaço vetorial $X \neq \{0\}$ tem uma base.*

Para demonstrar o teorema, faz-se uso do **Lema de Zorn**, mas antes considere as seguintes definições.

Definição 2.5. *Um conjunto **parcialmente ordenado** ou **Cadeia** é um conjunto X no qual é definida uma ordem parcial, isto é, uma relação binária, denotada por \preceq , e satisfazendo as seguintes condições:*

- i. $a \preceq a$ para cada $a \in X$. (Reflexiva)
- ii. Se $a \preceq b$ e $b \preceq a$, então $a = b$. (Antissimétrica)
- iii. Se $a \preceq b$ e $b \preceq c$, então $a \preceq c$. (Transitiva)

*Parcialmente enfatiza que M pode conter elementos a e b para os quais nem $a \preceq b$ e nem $b \preceq a$. Nestas condições a e b são ditos **elementos incomparáveis**. Por outro lado, dois elementos a e b são ditos **comparáveis** se eles satisfazem $a \preceq b$ ou $b \preceq a$ (ou ambas).*

*Um conjunto **totalmente ordenado** ou **cadeia** é um conjunto parcialmente ordenado tal que cada dois elementos do conjunto são comparáveis. Em outras palavras, uma cadeia é um conjunto parcialmente ordenado que não tem elementos incomparáveis.*

*Um **limitante superior** de um subconjunto Y de um conjunto parcialmente ordenado X é um elemento $\alpha \in X$ tal que $x \preceq \alpha$ para todo $x \in Y$.*

*Um **elemento maximal** de um conjunto X é um $m \in X$ tal que $m \preceq x$, para $x \in X$, implica $m = x$.*

Exemplo 2.5. O Conjunto dos Números Reais. Seja \mathbb{R} o conjunto de todos os números reais e seja $x \leq y$ com seu significado usual. Nestas condições \mathbb{R} é totalmente ordenado e não possui elemento maximal.

Exemplo 2.6. Conjunto das Partes. Seja $\mathcal{P}(X)$ o conjunto de todos os subconjuntos de um dado conjunto X . Dados A e B dois conjuntos, a relação $A \preceq B$ significa $A \subset B$, isto é, A é um subconjunto de B . Nestas condições $\mathcal{P}(X)$ é um conjunto parcialmente ordenado e o único elemento maximal de $\mathcal{P}(X)$ é X .

Lema 2.1. (Lema de Zorn). *Seja X um conjunto não vazio parcialmente ordenado. Suponha que cada cadeia $Y \subset X$ tenha um limitante superior. Então X tem pelo menos um elemento maximal.*

Demonstração. Ver seção 16 de [6]. □

A demonstração do Teorema 2.1 segue abaixo.

Demonstração. Seja M o conjunto de todos os subconjuntos linearmente independentes de X . Uma vez que $X \neq \{0\}$, existe um elemento $x \neq 0$ em X e como $\{x\}$ é linearmente independente, segue que $M \neq \emptyset$. O conjunto M é parcialmente ordenado com a relação inclusão de conjuntos. Deve-se mostrar que cada cadeia $C \subset M$ possui um limitante superior. Como $C = \{A_\alpha\}_{\alpha \in L}$, seu candidato natural para limitante superior é o conjunto $A = \bigcup_{\alpha \in L} A_\alpha$. O conjunto A é linearmente independente. De fato, seja $\mathcal{A} = \{x_1, \dots, x_n\}$ um subconjunto finito de A . Então, para cada $i = 1, \dots, n$, existe $\alpha_i \in L$ tal que $x_i \in A_{\alpha_i}$. Sendo C totalmente ordenado tem-se que, reordenando os elementos de \mathcal{A} se necessário, $A_{\alpha_1} \subseteq \dots \subseteq A_{\alpha_n}$, e portanto, $x_i \in A_{\alpha_n}$ para cada $i = 1, \dots, n$. Assim, \mathcal{A} é linearmente independente como um subconjunto finito do conjunto linearmente independente A_{α_n} . Como \mathcal{A} é qualquer, segue que A é linearmente independente concluindo que C tem A como limitante superior. Segue do Lema de Zorn que M tem um elemento maximal, por exemplo, B . Tem-se que B gera todo espaço X , pois do contrário existiria $x \in X$ não gerado por B , então $B \cup \{x\}$ seria linearmente independente o que contraria a maximalidade do conjunto B . Portanto B gera X e é de fato uma base para X . □

Teorema 2.2. (Dimensão de um subespaço). *Seja X um espaço vetorial n -dimensional. Então qualquer subespaço próprio Y de X possui dimensão menor que n .*

Demonstração. Se $n = 0$, devemos ter $X = \{0\}$ e então X não possui subespaço próprio. Se $\dim Y = 0$, então $Y = \{0\}$ e como $X \neq Y$ temos $\dim X \geq 1$. Portanto $\dim Y \leq \dim X = n$. Se $\dim Y$ for n , então Y teria uma base de n elementos, que seria também uma base para X desde que $\dim X = n$, então $X = Y$. Isto mostra que qualquer conjunto linearmente independente de vetores em Y deve ter menos que n elementos, e portanto $\dim Y < n$. □

2.2 Espaços Métricos

Definição 2.6. *Um espaço métrico é um par (X, d) , onde X é um conjunto e d é uma métrica em X , isto é, uma função real definida em $X \times X$ tal que para todos x, y, z em X tem-se:*

$$M1 - d(x, y) = 0 \text{ se, e somente se, } x = y.$$

$M2$ - Se $x \neq y$ então $d(x, y) > 0$.

$M3$ - $d(x, y) = d(y, x)$.

$M4$ - $d(x, y) \leq d(x, z) + d(z, y)$.

Exemplo 2.7. (Espaço Funcional). Considere o conjunto de todas as funções reais definidas e contínuas no intervalo fechado $[a, b]$ com imagem em \mathbb{R} . Escolhendo a métrica definida por

$$d(f, g) = \max\{|f(t) - g(t)|, t \in [a, b]\},$$

obtém-se um espaço métrico denotado por $\mathcal{C}([a, b], \mathbb{R})$. Também $\mathcal{C}([a, b], \mathbb{R})$ é um espaço métrico com a métrica definida por

$$d(f, g) = \int_a^b |f(t) - g(t)| dt,$$

para toda função real $f, g \in \mathcal{C}([a, b], \mathbb{R})$.

Exemplo 2.8. (Espaço das Funções Limitadas). Seja X um conjunto arbitrário. Uma função real $f : X \rightarrow \mathbb{R}$ diz-se ser *limitada* quando existe uma constante $k = k_f > 0$ tal que $|f(x)| \leq k_f$ para todo $x \in X$. O conjunto de todas as funções limitadas $f : X \rightarrow \mathbb{R}$, denotado por $\mathcal{B}(X, \mathbb{R})$, é um espaço métrico com a métrica, do supremo, definida por

$$d(f, g) = \sup_{t \in X} |f(t) - g(t)|.$$

No caso em que $X = [a, b]$, escreve-se $\mathcal{B}([a, b], \mathbb{R})$

2.2.1 Conjuntos Abertos, Fechados e Limitados

Abaixo definem-se três subconjuntos importantes de um espaço métrico $X = (X, d)$.

Definição 2.7. Dados um ponto $x_0 \in X$ e um número real $r > 0$, define-se:

(a) **Bola Aberta** - $B(x_0; r) = \{x \in X \mid d(x, x_0) < r\}$;

(b) **Bola Fechada** - $B[x_0; r] = \{x \in X \mid d(x, x_0) \leq r\}$;

(c) **Esfera** - $S(x_0; r) = \{x \in X \mid d(x, x_0) = r\}$;

Em todos os casos, x_0 é chamado de centro e r de raio.

Seja A um subconjunto de um espaço métrico X . Um ponto $a \in A$ chama-se um ponto **interior** de A quando este é centro de uma bola aberta contida em A , isto é, quando existe $r > 0$ tal que $d(x, a) < r \Rightarrow x \in A$. Chama-se, ainda, **interior** de A em X ao conjunto *int* A formado pelos pontos interiores a A .

Um subconjunto A de um espaço métrico X diz-se **aberto** em X quando todos os seus pontos são interiores, isto é, $\text{int } A = A$.

Um espaço métrico, diz-se que o conjunto V é uma **vizinhança** do ponto a quando $a \in \text{int } V$. Assim, V é uma vizinhança de a se, e somente se, V contém um aberto que contém a .

Seja X um espaço métrico. Um ponto $x \in X$ diz-se um **ponto isolado** de X quando ele é uma bola aberta em X , ou seja, quanto existe $r > 0$ tal que $B(x; r) = \{x\}$. Um espaço métrico X chama-se **discreto** quando todo ponto de X é isolado.

Um ponto a diz-se **aderente** a um subconjunto A de um espaço métrico X quando $d(a, A) = 0$. Isto significa que existem pontos de A arbitrariamente próximos de a , ou seja, para cada $\epsilon > 0$, pode-se encontrar $x \in A$ tal que $d(a, x) < \epsilon$.

Outras maneiras equivalentes de dizer que a é aderente a A são:

1. para todo $\epsilon > 0$ tem-se $B(a; \epsilon) \cap A \neq \emptyset$;
2. para todo aberto B contendo a , tem-se $B \cap A \neq \emptyset$;
3. toda vizinhança de a tem pontos em comum com A .

O **fecho** de um conjunto X num espaço métrico E é o conjunto \overline{X} dos pontos de E que são aderentes a X . Portanto, escrever $a \in \overline{X}$ é o mesmo que afirmar que o ponto a é aderente a X em E .

Diz-se que um conjunto $F \subset E$ é **fechado** no espaço métrico E quando seu complementar $E - F$ é aberto em E . De modo imediato, dado $F \subset E$, tem-se $\overline{F} = F$ se, e somente se, $E - F$ é aberto. Isto significa que um conjunto é fechado se, e somente se, contém todos os seus pontos aderentes.

Um subconjunto X de um espaço métrico E chama-se **limitado** quando existe uma constante $c > 0$ tal que $d(x, y) \leq c$ para quaisquer $x, y \in X$. O menor desses números c é chamado de diâmetro de X e representa-se pelo símbolo $\delta(X)$.

Uma aplicação $f : X \rightarrow M$, definida num conjunto arbitrário X e tomando valores num espaço métrico M , chama-se *limitada* quando sua imagem $f(X)$ é um subconjunto limitado de M .

Definição 2.8. *Sejam X e Y espaços métricos. Uma aplicação $T : X \rightarrow Y$ diz-se ser **contínua** no ponto $a \in X$ quando, para todo $\epsilon > 0$ dado arbitrariamente, existe $\delta > 0$ tal que $d(x, y) < \delta$ implica $d(T(x), T(y)) < \epsilon$.*

Diz-se, simplesmente, T é **contínua** se T for contínua em todos os pontos de X .

Definição 2.9. *Seja M um subconjunto do espaço métrico X . Um ponto $a \in X$, que pode ou não pertencer a M , é chamado de **ponto de acumulação** de M quando toda vizinhança V de a em X contém pelo menos um ponto $y \in M$, distinto do ponto a .*

O conjunto dos pontos de acumulação de M chama-se **derivado** de M e é denotado por M' .

Definição 2.10. Um conjunto M de um espaço métrico X é **denso** em X quando seu fecho \overline{M} coincide com o espaço inteiro X .

Diz-se que o espaço métrico X é **separável** se este possui um subconjunto enumerável que é denso em X . Vale lembrar que um conjunto X é dito ser enumerável se este é finito ou existe uma aplicação biunívoca $f : X \rightarrow \mathbb{N}$.

Exemplo 2.9. A reta real \mathbb{R} é separável, pois o conjunto dos números racionais \mathbb{Q} é um conjunto enumerável denso em \mathbb{R} .

2.2.2 Sequências e Completude

Uma **sequência** em um conjunto X é uma aplicação $x : \mathbb{N} \rightarrow X$, definida no conjunto $\mathbb{N} = \{1, 2, \dots, n, \dots\}$. O valor que a sequência x assume no número $n \in \mathbb{N}$ será indicado por x_n , em vez de $x(n)$. Este valor é chamado de *n-ésimo* termo da sequência.

Usam-se as notações $(x_1, x_2, \dots, x_n \dots)$, $(x_n)_{n \in \mathbb{N}}$ ou simplesmente (x_n) para representar uma sequência.

Uma **subsequência** de (x_n) é uma restrição da aplicação $n \mapsto x_n$ a um subconjunto infinito $\mathbb{N}' = \{n_1 < n_2 < \dots < n_k < \dots\}$ de \mathbb{N} . A subsequência é indicada pelas notações $(x_{x_1}, x_{x_2}, \dots, x_{x_k}, \dots)$, $(x_n)_{n \in \mathbb{N}'}$ ou simplesmente (x_{n_k}) .

Definição 2.11. Em um espaço métrico $X = (X, d)$, diz-se que o ponto x é **limite de uma sequência** (x_n) quando, para todo $\epsilon > 0$ dado arbitrariamente, é possível obter um natural $N = N(\epsilon)$ tal que $d(x_n, x) < \epsilon$ para todo $n > N$.

Exemplo 2.10. O limite da sequência $(1/n)$ de números reais é 0, pois dado $\epsilon > 0$ arbitrário, e sendo \mathbb{N} ilimitado superiormente, é possível obter um inteiro $n_0 > 0$ tal que $n_0 > 1/\epsilon$. Assim, $n > n_0$ implica

$$\left| \frac{1}{n} - 0 \right| = \frac{1}{n} < \frac{1}{n_0} < \epsilon.$$

Quando x é limite da sequência (x_n) , diz-se também que (x_n) converge para x , e escreve-se

$$\lim_{n \rightarrow \infty} x_n = x,$$

ou simplesmente $\lim x_n = x$. Se (x_n) não é convergente diz-se que (x_n) é divergente.

Uma sequência de aplicações $f_n : M \rightarrow X$, definidas em um conjunto arbitrário M e tomando valores em um espaço métrico X , converge **simplesmente (ou pontualmente)** em M para a aplicação $f : M \rightarrow X$ quando, para cada $x \in M$, a sequência $(f_1(x), f_2(x), \dots, f_n(x), \dots)$ tem limite $f(x)$ em X . Isto significa que dados arbitrariamente $x \in M$ e $\epsilon > 0$, existe $n_0 \in \mathbb{N}$ (dependendo de x e ϵ) tal que $n > n_0$ implica $d(f_n(x), f(x)) < \epsilon$.

Exemplo 2.11. A sequência de funções $f_n : \mathbb{R} \rightarrow \mathbb{R}$, dadas por $f_n(x) = x/n$, converge simplesmente em \mathbb{R} para a função identicamente nula.

Uma sequência de aplicações $f_n : M \rightarrow X$ converge **uniformemente** em M para a aplicação $f : M \rightarrow X$ quando, para todo número real $\epsilon > 0$ dado, for possível obter $n_0 \in \mathbb{N}$ tal que $n > n_0$ implica $d(f_n(x), f(x)) < \epsilon$, qualquer que seja $x \in M$.

No exemplo anterior, mesmo mantendo $\epsilon > 0$ fixo, não se pode determinar um número natural n_0 que seja satisfatório para todos os $x \in M$. Assim a convergência uniforme das aplicações $f_n : \mathbb{R} \rightarrow \mathbb{R}$ dadas por $f_n(x) = x/n$ só é satisfeita se $M \subset \mathbb{R}$ é subconjunto limitado de \mathbb{R} .

Teorema 2.3. *Sejam M, N espaços métricos. Se uma sequência de aplicações $f_n : M \rightarrow N$, contínuas no ponto $a \in M$, converge uniformemente em M para uma aplicação $f : M \rightarrow N$, então f é contínua no ponto a .*

Demonstração. Seja $\epsilon > 0$ qualquer, da continuidade uniforme, existe um número natural n tal que $d(f_n(x), f(x)) < \epsilon/3$ para todo $x \in M$. Como f_n é contínua no ponto a , existe $\delta > 0$ tal que $d(x, a) < \delta$ em M implica $d(f_n(x), f_n(a)) < \epsilon/3$. Então, para todo $x \in M$ com $d(x, a) < \delta$, tem-se

$$d(f(x), f(a)) \leq d(f(x), f_n(x)) + d(f_n(x), f_n(a)) + d(f_n(a), f(a)) < \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon.$$

□

Definição 2.12. Uma **sequência** (x_n) em um espaço métrico $X = (X, d)$ é dita ser de **Cauchy** se para cada $\epsilon > 0$ existe um natural $N = N(\epsilon)$ tal que $d(x_n, x_m) < \epsilon$ para quaisquer $m, n > N$.

O espaço métrico X é dito ser **completo** se toda sequência de Cauchy é convergente em X , isto é, tem um limite que é elemento de X .

Exemplo 2.12. O espaço funcional $\mathcal{C}([a, b], \mathbb{R})$ do exemplo 2.7 é completo. De fato, seja (f_m) uma sequência de Cauchy qualquer em $\mathcal{C}([a, b], \mathbb{R})$. Então, dado qualquer $\epsilon > 0$, existe $N \in \mathbb{N}$ tal que

$$d(f_m, f_n) = \max |f_m(t) - f_n(t)| < \epsilon, \tag{2.1}$$

para todos $m, n > N$ e todo $t \in [a, b]$. Fixado $t = t_0 \in [a, b]$, tem-se ainda $|f_m(t_0) - f_n(t_0)| < \epsilon$ para todos $n, m > N$. Assim a sequência $(f_m(t_0))$ é uma sequência de Cauchy de números reais. Como \mathbb{R} é completo, tem-se $\lim f_m(t_0) = f(t_0)$, para algum $f(t_0) \in \mathbb{R}$. Deste modo, pode-se associar para cada $t \in [a, b]$ um único número real $f(t)$, uma vez que o limite de uma sequência é único. Assim, fica definida uma função f em $[a, b]$ tomando valores reais, restando apenas mostrar que $f \in \mathcal{C}([a, b], \mathbb{R})$ e que $f_m \rightarrow f$.

De (2.1) com $n \rightarrow \infty$ tem-se

$$\max |f_m(t) - f(t)| \leq \epsilon,$$

para todos $m > N$ e todo $t \in [a, b]$. Consequentemente, para cada $t \in [a, b]$,

$$|f_m(t) - f(t)| \leq \epsilon,$$

sempre que $m > N$. Isto mostra que a sequência $(f_m(t))$ converge uniformemente para $f(t)$ em $[a, b]$. Uma vez que cada aplicação f_m é contínua em $[a, b]$ e que sua convergência é uniforme, o seu limite é uma função contínua em $[a, b]$ pelo Teorema 2.3. Portanto $f \in \mathcal{C}([a, b], \mathbb{R})$ e $f_m \rightarrow f$ concluindo que $\mathcal{C}([a, b], \mathbb{R})$ é completo.

Exemplo 2.13. Seja $\mathcal{C}([0, 1], \mathbb{R})$ o conjunto das funções contínuas em $[0, 1]$ tomando valores reais. Com a métrica

$$d(f, g) = \int_0^1 |f(t) - g(t)| dt$$

$\mathcal{C}([0, 1], \mathbb{R})$ não é completo.

De fato, considere a sequência de funções, Figura 2.1-(a), definidas por

$$f_m(t) = \begin{cases} 0, & \text{se } 0 \leq t \leq \frac{1}{2} \\ m(t - \frac{1}{2}), & \text{se } \frac{1}{2} < t \leq a_m \\ 1, & \text{se } a_m < t \leq 1, \end{cases}$$

onde $a_m = \frac{1}{2} + \frac{1}{m}$. A sequência (f_m) é de Cauchy, pois dadas $f, g \in \mathcal{C}([0, 1], \mathbb{R})$, tem-se

$$d(f_m, f_n) = \int_0^1 |f_m(t) - f_n(t)| dt = \frac{1}{2} \left(\frac{1}{m} - \frac{1}{n} \right),$$

que é a área do triângulo hachurado da Figura 2.1-(b). Logo, para todo $\epsilon > 0$, tem-se $d(f_m, f_n) \rightarrow 0$ quando $m, n \rightarrow \infty$.

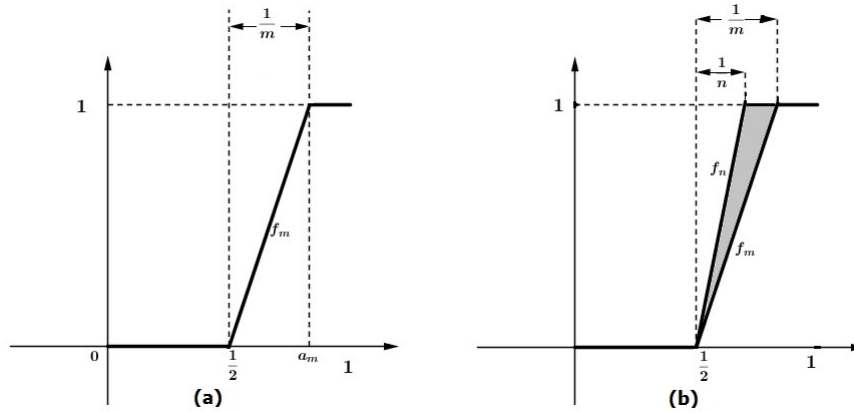
Note que $f_m \rightarrow f$, onde

$$f(t) = \begin{cases} 0, & \text{se } 0 \leq t \leq \frac{1}{2} \\ 1, & \text{se } \frac{1}{2} < t \leq 1, \end{cases}$$

pois

$$d(f_m, f) = \int_0^1 |f_m(t) - f(t)| dt = \frac{1}{2m} \rightarrow 0,$$

quando $m \rightarrow \infty$. Mas f é descontínua o que implica $f \notin \mathcal{C}([0, 1], \mathbb{R})$. Portanto $\mathcal{C}([0, 1], \mathbb{R})$ não é completo.

Figura 2.1: Sequência (f_m)

2.3 Espaços Métricos Compactos

Um espaço métrico E diz-se **totalmente limitado** quando, para todo $\epsilon > 0$, pode-se exprimir $E = S_1 \cup \dots \cup S_n$ como reunião de um número finito de subconjuntos, cada um dos quais com diâmetro menor que ϵ .

Seja X um subconjunto de um espaço métrico E . Uma **cobertura** de X é uma família $\mathcal{C} = (C_\lambda)_{\lambda \in L}$ de subconjuntos de E tal que $X \subset \bigcup_{\lambda \in L} C_\lambda$. Se existe um subconjunto $L' \subset L$ tal que, para cada $x \in X$, ainda se pode obter $\lambda \in L'$ com $x \in C_\lambda$, isto é, $X \subset \bigcup_{\lambda \in L'} C_\lambda$, então a subfamília $\mathcal{C}' = (C_\lambda)_{\lambda \in L'}$ chama-se uma **subcobertura** de \mathcal{C} .

Uma cobertura $X \subset \bigcup_{\lambda \in L} A_\lambda$ diz-se aberta quando cada conjunto A_λ , $\lambda \in L$, é aberto em E . A cobertura $X \subset \bigcup_{\lambda \in L} C_\lambda$ diz-se **finita** quando L é um conjunto finito. Neste caso, tem-se $L = \{\lambda_1, \lambda_2, \dots, \lambda_n\}$ e escreve-se $X \subset C_{\lambda_1} \cup C_{\lambda_2} \cup \dots \cup C_{\lambda_n}$.

Definição 2.13. Um espaço métrico E chama-se **compacto** quando toda cobertura aberta possui uma subcobertura finita. Isto é, se $E \subset \bigcup_{\lambda \in L} A_\lambda$, onde cada A_λ é aberto em E , então existem $\lambda_1, \dots, \lambda_n$ tais que $E \subset A_{\lambda_1} \cup \dots \cup A_{\lambda_n}$.

A seguir alguns resultados sobre espaços métricos compactos.

Teorema 2.4. Um subespaço de um espaço compacto é compacto se, e somente se, for fechado.

Demonstração. Seja X um espaço métrico compacto e $F \subset X$ fechado. Dada uma cobertura aberta $F \subset \bigcup_{\lambda \in L} A_\lambda$, tem-se que $\bigcup_{\lambda \in L} A_\lambda \cup (X - F)$ é uma cobertura aberta de X , que é compacto, logo existe uma subcobertura finita $A_{\lambda_1} \cup \dots \cup A_{\lambda_n} \cup (X - F)$ de X . Como nenhum ponto de F pertence a $X - F$, tem-se $F \subset A_{\lambda_1} \cup \dots \cup A_{\lambda_n}$. Logo F é compacto. Reciprocamente, seja $K \subset X$ um conjunto compacto de um espaço

métrico X . Se K não fosse fechado em X , existiria $x \in \overline{K} - K$. Para cada $n \in \mathbb{N}$, tome $A_n = X - B[x, 1/n]$, onde $B[x, 1/n]$ é a bola fechada de centro em x e raio $1/n$. Note que $K \subset \bigcup_{n=1}^{\infty} A_n$. De fato, como $\bigcap_{n=1}^{\infty} B[x; 1/n] = \{x\}$, tem-se $\bigcup_{n=1}^{\infty} A_n = X - \{x\} \supset K$ pois $x \notin K$. Como $A_1 \subset A_2 \subset A_3 \subset \dots$, a reunião de uma coleção finita de conjuntos A_n é igual ao conjunto de maior índice da coleção. Mas $x \in \overline{K}$, assim cada bola $B[x; 1/n]$ contém algum ponto de K , ou seja, nenhum A_n contém K . Logo a cobertura $\bigcup_{n=1}^{\infty} A_n$ não admite subcobertura finita, contrariando o fato de K ser compacto. □

Teorema 2.5. *A imagem de um conjunto compacto por uma aplicação contínua é um conjunto compacto.*

Demonstração. Sejam $f : M \rightarrow N$ contínua, onde M e N são espaços métricos, e $K \subset M$ compacto. Dada uma cobertura $f(K) \subset \bigcup_{\lambda \in L} A_\lambda$, obtém-se a cobertura $K \subset \bigcup_{\lambda \in L} f^{-1}(A_\lambda)$, da qual possui uma subcobertura finita $K \subset f^{-1}(A_{\lambda_1}) \cup \dots \cup f^{-1}(A_{\lambda_n})$. Logo

$$f(K) \subset ff^{-1}(A_{\lambda_1}) \cup \dots \cup ff^{-1}(A_{\lambda_n}) \subset A_{\lambda_1} \cup \dots \cup A_{\lambda_n}.$$

Portanto, $f(K)$ é compacto. □

Teorema 2.6. *Seja E um espaço métrico. São equivalentes as seguintes propriedades:*

- 1) E é compacto.
- 2) Todo subconjunto infinito de E possui um ponto de acumulação.
- 3) Toda sequência de pontos em E possui uma subsequência convergente.
- 4) E é completo e totalmente limitado, isto é, dado $\epsilon > 0$ arbitrário, existe um recobrimento finito de E formado por conjuntos de diâmetros menores que ϵ .

Demonstração. 1) \Rightarrow 2). Seja $X \subset E$ um subconjunto infinito. Se $X' = \emptyset$, então $\overline{X} = X \cup X' = X$, isto é, X é fechado em E e portanto compacto pelo Teorema 2.4. Como nenhum ponto de X é ponto de acumulação, X é discreto e portanto finito, contradizendo a hipótese.

2) \Rightarrow 3). Seja $(x_n)_{n \in \mathbb{N}}$ uma sequência de pontos de E ; há duas possibilidades: ou o conjunto $X = \{x_1, x_2, \dots, x_n, \dots\}$ é finito ou infinito. No caso em que X é finito, algum valor $x_{n_1} = x_{n_2} = \dots$ deve repetir-se uma infinidade de vezes, o que nos dá uma subsequência constante e, portanto, convergente de (x_n) . No caso em que X é infinito, segue que X possui um ponto de acumulação $x \in E$. Assim, toda vizinhança

de x conterá termos x_n com índices arbitrariamente grandes e portanto x será limite de uma subsequência de (x_n) .

3) \Rightarrow 4). Seja (x_n) uma sequência de Cauchy em E . Por hipótese (x_n) possui uma subsequência x_{n_k} convergente. Logo a própria sequência (x_n) é convergente. Para mostrar que E é totalmente limitado, seja dado $\epsilon > 0$. E pode ser coberto por um número finito de bolas $B(x_i; \epsilon/2)$. De fato, tomado $x_1 \in E$ qualquer. Se $E = B(x_1; \epsilon/2)$, a afirmação está provada. Caso contrário, existirá $x_2 \in E$, com $d(x_2, x_1) > \epsilon/2$. Se $E = B(x_1; \epsilon/2) \cup B(x_2; \epsilon/2)$, estará concluída a demonstração. No caso contrário, existirá $x_3 \in E$, com $d(x_3, x_2) > \epsilon/2$, $d(x_3, x_1) > \epsilon/2$.

Prosseguindo, tem-se que, ou existe um número finito de pontos x_1, x_2, \dots, x_n em E tais que

$$E = B(x_1; \epsilon/2) \cup \dots \cup B(x_n; \epsilon/2),$$

ou então é possível obter uma sequência de pontos $x_n \in E$, com $d(x_m, x_n) > \epsilon/2$ para $m \neq n$ quaisquer. Tal sequência, porém, não possuirá uma subsequência de Cauchy, bem como não apresentará uma subsequência convergente, contradizendo a hipótese. Portanto E é totalmente limitado.

4) \Rightarrow 1). Suponha que E não seja compacto, assim existe uma cobertura aberta $(A_\lambda)_{\lambda \in L}$ de E que não admite subcobertura finita. Sendo totalmente limitado, E pode ser coberto por um número finito de conjuntos com diâmetro menor que 1. Note que pelo menos um desses conjuntos, diga-se S_1 , não pode ser coberto por um número finito dos A_λ , pois, se todos pudessem, E seria coberto por uma quantidade finita dos A_λ .

Tem-se que S_1 também é totalmente limitado, logo é reunião de um número finito de subconjuntos de diâmetro menor que $1/2$. Do mesmo modo, pelo menos um desses conjuntos, diga-se S_2 , não está contido em uma reunião finita dos A_λ , pois, se todos tivessem, sua reunião S_1 também estaria.

Prosseguindo analogamente, obtém-se uma sequência $S_1 \supset S_2 \supset \dots \supset S_n \supset \dots$, onde o diâmetro de S_n é menor que $1/n$ e nenhum dos S_n pode ser coberto por um número finito de conjuntos A_λ . Tome para cada $n \in \mathbb{N}$, um ponto $x_n \in S_n$. Como $\delta(S_n) < 1/n$, segue que a sequência (x_n) é de Cauchy. Com efeito, basta notar que $\lim_{n \rightarrow \infty} \delta(S_n) = 0$, logo para todo $\epsilon > 0$ existe $n_0 \in \mathbb{N}$ tal que $\delta(S_n) < \epsilon$, para todo $n > n_0$, mas isto equivale a dizer que $d(x_n, x_m) < \epsilon$ uma vez que cada S_n é, em particular, limitado.

Sendo E completo, existe $\lim x_n = x$. O ponto x pertence a algum aberto A_λ . Logo, existe um $\epsilon > 0$ tal que $B(x; \epsilon) \subset A_\lambda$. Para n suficientemente grande, de modo que $1/n < \epsilon$, tem-se $S_n \subset B(x; \epsilon) \subset A_\lambda$, contradizendo o fato de que S_n não está contido em uma reunião finita dos conjuntos A_λ .

□

Definição 2.14. Diz-se que um subconjunto X de um espaço métrico E é **relativamente compacto** quando seu fecho \overline{X} é um subconjunto compacto de E .

Corolário 2.1. *Dado um subconjunto X de um espaço métrico completo E , são equivalentes as seguintes propriedades:*

- 1') X é relativamente compacto em E .
- 2') Toda sequência de pontos de X contém uma subsequência convergente em E .
- 3') X é totalmente limitado.

Demonstração. Basta notar que cada uma das propriedades 1'), 2') e 3') de X é equivalente à propriedade correspondente 1), 3) e 4) de \overline{X} . \square

2.3.1 O Teorema de Ascoli-Arzelá

Nesta seção E indica um espaço compacto e F um espaço métrico completo com distância d . $\mathcal{C}(E, F)$ indica o conjunto das funções contínuas de E em F munido da distância

$$d(f, g) = \sup_{x \in E} d(f(x), g(x)).$$

De modo análogo ao Exemplo 2.12 mostra-se que $\mathcal{C}(E, F)$ completo.

Seja H um conjunto de aplicações de E em F . Diz-se que H é **equicontínuo** no ponto $x_0 \in E$ quando, para todo $\epsilon > 0$, existe uma vizinhança V_{x_0} de x_0 em E tal que, para todo $x \in V_{x_0}$ tem-se $d(f(x), f(x_0)) < \epsilon$, qualquer que seja $f \in H$.

Note que se H é equicontínuo no ponto x_0 , então todas as aplicações $f \in H$ são contínuas no ponto x_0 e todo subconjunto de H é equicontínuo no ponto x_0 .

Diz-se que o conjunto H de aplicações $f : E \rightarrow F$ é equicontínuo quando H é equicontínuo em todos os pontos de E . Nestas condições tem-se $H \subset \mathcal{C}(E, F)$.

Exemplo 2.14. Todo conjunto finito $H = \{f_1, \dots, f_n\}$ de aplicações contínuas $f_i : E \rightarrow F$ é equicontínuo. Com efeito, seja $x_0 \in E$, para cada $\epsilon > 0$ e cada $i = 1, 2, \dots, n$, existe uma vizinhança V_i de x_0 em E tal que para todo $x \in V_i$ tem-se $d(f_i(x), f_i(x_0)) < \epsilon$. Para que H seja equicontínuo basta considerar $V = V_1 \cap \dots \cap V_n$. Então, para todo $x \in V$ tem-se $d(f(x), f(x_0)) < \epsilon$ seja qual for $f = f_i \in H$.

Teorema 2.7. O Teorema de Ascoli-Arzelá. *Um subconjunto $H \subset \mathcal{C}(E, F)$ é relativamente compacto se, e somente se, ele satisfaz as condições:*

- 1) H é equicontínuo;
- 2) Para todo $x \in E$, o conjunto $H(x) = \{f(x) | f \in H\}$ é relativamente compacto em F .

Demonstração. Suponha, primeiro, que $H \subset \mathcal{C}(E, F)$ seja relativamente compacto. Fixado $x \in E$, a aplicação $v_x : \mathcal{C}(E, F) \rightarrow F$ definida por $v_x(f) = f(x)$ é contínua. Sendo \overline{H} compacto, segue que $v_x(\overline{H})$ é compacto. Pelo Teorema 2.4, em particular,

$v_x(\overline{H})$ é fechado em F . Como $v_x(H) \subset v_x(\overline{H})$, segue que $\overline{v_x(H)} \subset v_x(\overline{H})$ e portanto $\overline{v_x(H)}$ é compacto. Como $H(x) = v_x(H)$, segue que $H(x)$ é relativamente compacto.

H é equicontínuo. Com efeito, da propriedade 3') do Corolário 2.1, dado $\epsilon > 0$ arbitrário, existe um número finito de conjuntos H_1, \dots, H_n , de diâmetro menor que $\epsilon/3$, tais que $H = H_1 \cup \dots \cup H_n$. Fixados os elementos $f_1 \in H_1, \dots, f_n \in H_n$, da continuidade das funções f_i , segue que dado $x_0 \in E$ existe uma vizinhança V de x_0 tal que para todo $x \in V$ tem-se

$$d(f_i(x), f_i(x_0)) < \frac{\epsilon}{3},$$

para cada $i = 1, 2, \dots, n$. Dado $f \in H$, tem-se que $f \in H_i$ para algum i , assim para $x \in V$, tem-se

$$d(f(x), f(x_0)) \leq d(f(x), f_i(x)) + d(f_i(x), f_i(x_0)) + d(f_i(x_0), f(x_0)) < \epsilon.$$

Portanto H é equicontínuo.

Reciprocamente, suponha que $H \subset \mathcal{C}(E, F)$ seja equicontínuo e tal que para todo $x \in E$ o conjunto $H(x)$ seja relativamente compacto em F . Do Corolário 2.1, segue que para mostrar que H é relativamente compacto basta mostrar que H é totalmente limitado, isto é, dado $\epsilon > 0$ existe um número finito de conjuntos de diâmetro menor que ϵ tais que H se exprime como reunião destes conjuntos. Assim, sendo H equicontínuo, para todo $x \in E$ existe uma vizinhança aberta V_x de x tal que se $x' \in V_x$ tem-se $d(f(x'), f(x)) < \epsilon/3$ para toda $f \in H$. Como E é compacto, E pode ser recoberto por um número finito de abertos V_{x_1}, \dots, V_{x_n} com a propriedade anterior. Por outro lado $H(x_i) = H^{(i)}$, $i = 1, 2, \dots, n$, é relativamente compacto por hipótese, portanto existem conjuntos $H_1^{(i)}, H_2^{(i)}, \dots, H_{m_i}^{(i)}$, de diâmetro menor ou igual a $\epsilon/3$, tais que $H^{(i)} = H_1^{(i)} \cup H_2^{(i)} \cup \dots \cup H_{m_i}^{(i)}$. Para cada sequência de inteiros p_1, \dots, p_n com $1 \leq p_i \leq m_i$, $i = 1, 2, \dots, n$, considere

$$H_{p_i} = \{f \in H \mid f(x_i) \in H_{p_i}^{(i)}, i = 1, 2, \dots, n\}.$$

Estes conjuntos formam um recobrimento finito de H , restando mostrar apenas que cada H_{p_i} tem diâmetro menor que ϵ . Para isto, sejam $f, g \in H_{p_i}$, para todo $x \in E$ existe $i \in \{1, 2, \dots, n\}$ tal que $x \in V_{x_i}$ e portanto

$$d(f(x), g(x)) \leq d(f(x), f(x_i)) + d(f(x_i), g(x_i)) + d(g(x_i), g(x)) < \epsilon.$$

Portanto H é relativamente compacto. □

2.4 Espaços Normados e de Banach

Definição 2.15. Um *espaço normado* X é um espaço vetorial sobre \mathbb{K} com uma norma definida sobre ele. Um *espaço de Banach* é um espaço normado completo (completo na métrica definida pela norma; veja (2.2), abaixo).

Aqui, uma **norma** sobre um espaço vetorial (real ou complexo) X é uma função real $\|\cdot\| : X \rightarrow \mathbb{R}$, que associa a cada vetor $x \in X$ um número real $\|x\|$, chamado a norma de x , que possui as seguintes propriedades:

(N1) - $\|x\| \geq 0$ para todo $x \in X$.

(N2) - $\|x\| = 0$ se, e somente se, $x = 0$.

(N3) - $\|\alpha x\| = |\alpha| \|x\|$.

(N4) - $\|x + y\| \leq \|x\| + \|y\|$.

onde x e y são vetores quaisquer de X e α é qualquer escalar em \mathbb{K} .

A norma sobre X define uma métrica d sobre X que é dada por

$$d(x, y) = \|x - y\|, \quad x, y \in X \quad (2.2)$$

e é chamada de **métrica induzida pela norma**.

O espaço normado definido acima é denotado por $(X, \|\cdot\|)$ ou simplesmente por X .

Proposição 2.1. *Seja X um espaço vetorial normado. Então, para quaisquer x, y em X tem-se $|\|x\| - \|y\|| \leq \|x - y\|$.*

Demonstração. Observe que,

- i. $\|x\| = \|(x - y) + y\| \leq \|x - y\| + \|y\| \Rightarrow \|x\| - \|y\| \leq \|x - y\|$.
- ii. $\|y\| = \|(y - x) + x\| \leq \|y - x\| + \|x\| \Rightarrow \|y\| - \|x\| \leq \|x - y\| \Rightarrow -\|x - y\| \leq \|x\| - \|y\|$.

Portanto, de i. e ii.,

$$|\|x\| - \|y\|| \leq \|x - y\|.$$

□

Lema 2.2. *Seja X um espaço vetorial normado, a função $\|\cdot\| : X \rightarrow \mathbb{R}$, dada por $x \mapsto \|x\|$ é uniformemente contínua.*

Demonstração. Dado $\epsilon > 0$, basta tomar $\delta = \epsilon$ e aplicar a proposição 2.1. □

Exemplo 2.15. O espaço euclidiano \mathbb{R}^n é um espaço de Banach com a norma definida por

$$\|x\| = \left(\sum_{j=1}^n |x_j|^2 \right)^{\frac{1}{2}}$$

onde $x = (x_1, x_2, \dots, x_n)$.

De fato, seja $(x_k)_{k \in \mathbb{N}}$ uma sequência de Cauchy em \mathbb{R}^n , com $x_k = (x_{k1}, x_{k2}, \dots, x_{kn}) = (x_{ki})_{k \in \mathbb{N}}$ ($i = 1, \dots, n$). Como (x_k) é de Cauchy, dado $\epsilon > 0$, existe um natural N tal que

$$\|x_k - x_p\| = \left(\sum_{i=1}^n |x_{ki} - x_{pi}|^2 \right)^{\frac{1}{2}} < \epsilon \quad \text{para todo } k, p > N. \quad (2.3)$$

Note que, para cada $1 \leq i \leq n$ tem-se

$$|x_{ki} - x_{pi}| \leq \|x_k - x_p\| < \epsilon.$$

Assim, para cada $1 \leq i \leq n$ fixo, a sequência $(x_{ki})_{k \in \mathbb{N}}$ é de Cauchy em \mathbb{R} , que é convergente pois \mathbb{R} é completo, ou seja, $\lim_{k \rightarrow \infty} x_{ki} = x_i$ existe e é um número real. Fazendo isto para todo $1 \leq i \leq n$ e $p \rightarrow \infty$, segue que a sequência (x_k) é convergente, com seu limite em \mathbb{R}^n .

Do mesmo modo, mostra-se que o espaço \mathbb{C}^n também é de Banach.

No exemplo 2.13, o espaço $\mathcal{C}([0, 1], \mathbb{R})$ com a norma $\|f\| = \int_0^1 |f(t)| dt$ não é um espaço de Banach.

A seguir, mais alguns exemplos de espaços normados. Tais espaços podem ser retomados durante o texto.

Exemplo 2.16. Dado um espaço compacto K , indica-se por $\mathcal{C}(K, \mathbb{C})$ o conjunto das funções definidas em K tomando valores complexos que são contínuas. A menos que haja menção contrária, $\mathcal{C}(K, \mathbb{C})$ é sempre munido da norma

$$x \in \mathcal{C}(K, \mathbb{C}) \mapsto \|x\| = \sup_{t \in K} |x(t)|.$$

Exemplo 2.17. O espaço normado $\mathcal{C}_{L_1}([a, b], \mathbb{C})$ indica o conjunto $\mathcal{C}([a, b], \mathbb{C})$, (conjunto das funções contínuas em $[a, b]$ tomando valores complexos), munido da norma

$$x \in \mathcal{C}([a, b], \mathbb{C}) \mapsto \|x\|_1 = \int_a^b |x(t)| dt.$$

No caso em que não houver confusão a notação $\|x\|_1$ será omitida, denotando-se apenas por $\|x\|$.

Exemplo 2.18. $\mathcal{C}^{(m)}([a, b], \mathbb{C})$ indica o conjunto das funções definidas em $[a, b]$ tomando valores complexos que são m vezes continuamente diferenciáveis. Com a notação do Exemplo 2.16

$$x \in \mathcal{C}^{(m)}([a, b], \mathbb{C}) \mapsto \|x\|^{(m)} = \sup_{0 \leq i \leq m} \|x^{(i)}\|.$$

2.4.1 O Teorema de Hahn-Banach

Definição 2.16. Seja X um espaço vetorial sobre \mathbb{K} . Um **funcional linear** é uma transformação linear com domínio X e imagem no corpo escalar \mathbb{K} , assim

$$f : X \longrightarrow \mathbb{K}.$$

Definição 2.17. Um funcional linear $f : X \rightarrow \mathbb{K}$ é dito ser **limitado** se existir um número real c tal que para todo $x \in X$, tem-se

$$|f(x)| \leq c\|x\|. \quad (2.4)$$

Assim, define-se a **norma** $\|f\|$, do funcional f , por

$$\|f\| = \sup_{\substack{x \in X \\ x \neq 0}} \frac{|f(x)|}{\|x\|}.$$

Lema 2.3. Seja f um funcional linear limitado. Então:

a) Uma fórmula alternativa para a norma de f é

$$\|f\| = \sup_{\substack{x \in X \\ \|x\|=1}} |f(x)|. \quad (2.5)$$

b) $\|f\| = \sup_{\substack{x \in X \\ x \neq 0}} \frac{|f(x)|}{\|x\|}$ define uma norma.

Demonstração. a) Como f é linear, para todo $x \neq 0$ em X , tem-se

$$\|f\| = \sup_{\substack{x \in X \\ x \neq 0}} \frac{|f(x)|}{\|x\|} = \sup_{\substack{x \in X \\ x \neq 0}} \frac{1}{\|x\|} |f(x)| = \sup_{\substack{x \in X \\ x \neq 0}} \left| f \left(\frac{x}{\|x\|} \right) \right| = \sup_{\substack{y \in X \\ \|y\|=1}} |f(y)|,$$

onde $y = \frac{x}{\|x\|}$.

b) (N1) segue de modo imediato da definição.

(N2) Se $f = 0$ não há o que provar. Reciprocamente, suponha que $f \neq 0$, então existe $x_0 \in X$, $x_0 \neq 0$ tal que $|f(x_0)| > 0$, assim $\|f\| = \sup_{\substack{x \in X \\ x \neq 0}} \frac{|f(x)|}{\|x\|} > 0$.

(N3) Sejam $x \in X$ e α um escalar qualquer, então

$$\|\alpha f\| = \sup_{\|x\|=1} |\alpha f(x)| = \sup_{\|x\|=1} |\alpha| |f(x)| = |\alpha| \sup_{\|x\|=1} |f(x)| = |\alpha| \|f\|.$$

(N4) Seja $x \in X$, então

$$\begin{aligned} \|f_1 + f_2\| &= \sup_{\|x\|=1} |(f_1 + f_2)(x)| = \sup_{\|x\|=1} |f_1(x) + f_2(x)| \\ &\leq \sup_{\|x\|=1} |f_1(x)| + \sup_{\|x\|=1} |f_2(x)| \\ &= \|f_1\| + \|f_2\|. \end{aligned}$$

□

Teorema 2.8. *Um funcional linear $f : X \rightarrow \mathbb{K}$ definido em um espaço normado X é contínuo se, e somente se, f é limitado.*

Demonstração. Se $f = 0$ o resultado segue de imediato. Suponha que $f \neq 0$. Então $\|f\| \neq 0$. Suponha, também, que f seja um funcional linear limitado e x_0 um elemento qualquer de X .

Seja $\epsilon > 0$ arbitrário, uma vez que f é linear, para todo $x \in X$ tal que $\|x - x_0\| < \delta$ com $\delta = \frac{\epsilon}{\|f\|}$, tem-se

$$|f(x) - f(x_0)| = |f(x - x_0)| \leq \|f\| \|x - x_0\| < \|f\| \delta = \epsilon.$$

Portanto f é contínuo.

Reciprocamente, suponha que f seja contínuo em $x_0 \in X$. Então, dado $\epsilon > 0$ arbitrário, existe $\delta > 0$ tal que para todo $x \in X$ com $\|x - x_0\| < \delta$ tem-se $|f(x) - f(x_0)| < \epsilon$.

Para mostrar que f é limitado considere $y \in X$ arbitrário e tome $x = x_0 + \frac{\delta}{\|y\|}y$. Assim $x - x_0 = \frac{\delta}{\|y\|}y$ e da continuidade e linearidade de f obtém-se

$$|f(x) - f(x_0)| = |f(x - x_0)| = \left| f\left(\frac{\delta}{\|y\|}y\right) \right| = \frac{\delta}{\|y\|} \|f(y)\| < \epsilon.$$

Assim $|f(y)| < \frac{\epsilon}{\delta} \|y\|$ concluindo que f é limitado. \square

Para um caso mais geral do que os funcionais lineares, considere o seguinte teorema.

Teorema 2.9. *Sejam X e Y espaços normados e f uma aplicação linear de X em Y . São equivalentes as seguintes propriedades:*

- a) f é contínua na origem;
- b) $\sup_{\substack{x \in X \\ \|x\| \leq 1}} \|f(x)\| = M < \infty$;
- c) Existe $c > 0$ tal que $\|f(x)\| < c\|x\|$ para todo $x \in X$;
- d) f é contínua.

Demonstração.

a) \Rightarrow b) Sendo f contínuo na origem, então dado $\epsilon > 0$ existe $\delta > 0$ tal que $\|x\| \leq \delta$ implica $\|f(x) - 0\| = \|f(x)\| \leq \epsilon$. Portanto $\|x\| \leq 1$ implica $\|f(x)\| \leq \frac{\epsilon}{\delta}$.

b) \Rightarrow c) Para todo $x \in X$, $x \neq 0$, o elemento $\frac{x}{\|x\|}$ tem norma 1 e portanto

$$\left\| f\left(\frac{x}{\|x\|}\right) \right\| \leq \sup \left\| f\left(\frac{x}{\|x\|}\right) \right\| = M,$$

assim, $\|f(x)\| \leq M\|x\|$.

c) \Rightarrow d) Seja x_0 um elemento qualquer de X , dado $\epsilon > 0$ basta tomar $\delta = \frac{\epsilon}{c}$, pois se $\|x - x_0\| < \frac{\epsilon}{c}$ então $\|f(x) - f(x_0)\| = \|f(x - x_0)\| \leq c\|x - x_0\| < \epsilon$.

d) \Rightarrow a) É evidente. \square

Definição 2.18. Dados espaços normados X e Y , $\mathcal{L}(X, Y)$ indica o espaço vetorial das aplicações lineares contínuas de X em Y munido da norma

$$f \in \mathcal{L}(X, Y) \mapsto \|f\| = \sup_{\|x\| \leq 1} \|f(x)\|.$$

Portanto, do Teorema 2.9, para todo $x \in X$ tem-se $\|f(x)\| \leq \|f\|\|x\|$, isto é, $\|f\|$ é a menor constante c tal que $\|f(x)\| \leq c\|x\|$.

Proposição 2.2. Sejam $X = \mathcal{C}([a, b], \mathbb{C})$, $Y = \mathcal{C}([c, d], \mathbb{C})$ e $K : [c, d] \times [a, b] \rightarrow \mathbb{C}$ uma função contínua. Para todo $f \in X$ define-se

$$(kf)(t) = \int_a^b K(t, s)f(s)ds \quad (c \leq t \leq d),$$

onde $k \in \mathcal{L}(X, Y)$ e tem-se $\|k\| \leq \sup_{c \leq t \leq d} \int_a^b |K(t, s)|ds$.

Demonstração. É imediato que k é uma aplicação linear de X em Y . Para a sua continuidade basta notar que

$$|(kf)(t)| = \left| \int_a^b K(t, s)f(s)ds \right| \leq \int_a^b |K(t, s)f(s)|ds \leq \int_a^b |K(t, s)|ds \cdot \|f\|$$

implica que

$$\|kf\| \leq \sup_{c \leq t \leq d} \int_a^b |K(t, s)|ds \cdot \|f\|,$$

e portanto

$$\|k\| \leq \sup_{c \leq t \leq d} \int_a^b |K(t, s)|ds.$$

\square

Definição 2.19. Seja X um espaço vetorial. Uma função $p : X \rightarrow \mathbb{R}$ é chamada de **funcional sublinear** se satisfaz as seguintes propriedades:

- (a) $p(x + y) \leq p(x) + p(y)$, para todo $x, y \in X$,
- (b) $p(\alpha x) = \alpha p(x)$, para todo $\alpha \in \mathbb{R}_+$ e $x \in X$.

Teorema 2.10. Teorema de Hahn-Banach (Extensão de funcionais lineares). Seja X um espaço vetorial real e $p : X \rightarrow \mathbb{R}$ um funcional sublinear. Além disso, seja $f : Z \rightarrow \mathbb{R}$ um funcional linear definido em um subespaço Z de X , distintos dos triviais, tal que:

$$f(x) \leq p(x) \tag{2.6}$$

para todo $x \in Z$. Então existe um funcional linear $\tilde{f} : X \rightarrow \mathbb{R}$ tal que

$$\tilde{f}(x) \leq p(x) \quad (2.7)$$

para todo $x \in X$, e $\tilde{f}(x) = f(x)$ para todo $x \in Z$.

Demonstração. Deve-se provar:

- (a) O conjunto E , de todas as extensões lineares $g : D(g) \subset X \rightarrow \mathbb{R}$ de f , isto é, $g(x) = f(x)$ para todo $x \in Z$, satisfazendo $g(x) \leq p(x)$ para todo $x \in D(g)$, onde $D(g)$ é o domínio de g , pode ser parcialmente ordenado e o Lema de Zorn produz um elemento maximal \tilde{f} de E .
- (b) \tilde{f} é definido em todo o espaço X .

Com efeito,

(a) seja E o conjunto de todas as extensões lineares g de f tais que

$$g(x) \leq p(x)$$

para todo $x \in D(g)$. Uma vez que $f \in E$, tem-se $E \neq \emptyset$. Assim, E cumpre as condições do Lema de Zorn com a seguinte ordem parcial:

$$g_1 \leq g_2$$

se, e somente se, $D(g_2) \supset D(g_1)$ e $g_2(x) = g_1(x)$ para todo $x \in D(g_1)$, ou seja, g_2 é uma extensão de g_1 .

De fato, para qualquer conjunto totalmente ordenado $C = \{g_i | i \in I\} \subset E$, considere $D(g) = \bigcup_{i \in I} D(g_i)$ e defina $g : D(g) \rightarrow \mathbb{R}$ por $g(x) = g_i(x)$ para todo $x \in D(g_i)$.

Afirma-se que g está bem definida. De fato, se $g(x) = g_i(x)$ e $g(x) = g_j(x)$ para algum $x \in D(g)$ e para certos $i, j \in I$, como $g_i, g_j \in C$, tem-se $g_i \leq g_j$ ou $g_j \leq g_i$. Por exemplo, $g_i \leq g_j$. Neste caso, $D(g_j) \supset D(g_i)$ e $g_j(x) = g_i(x)$ para todo $x \in D(g_i)$.

Afirma-se que $D(g)$ é um subespaço de X . De fato, sejam $x, y \in D(g)$, então, $x \in D(g_i)$ e $y \in D(g_j)$ para certos $i, j \in I$. Como $g_i, g_j \in C$, tem-se $g_i \leq g_j$ ou $g_j \leq g_i$. Suponha $g_i \leq g_j$. Então $D(g_j) \supset D(g_i)$ e $g_j(x) = g_i(x)$ para todo $x \in D(g_i)$. Assim, pode-se dizer que $x, y \in D(g_j)$. Sendo $D(g_j)$ um subespaço vetorial, para quaisquer $\alpha, \beta \in \mathbb{R}$, tem-se $\alpha x + \beta y \in D(g_j)$ e, portanto $\alpha x + \beta y \in D(g) = \bigcup_{i \in I} D(g_i)$.

Afirma-se que g é linear. De fato, sejam $x, y \in D(g)$ quaisquer, assim $x \in D(g_i)$ e $y \in D(g_j)$ para certos $i, j \in I$. Como $g_i, g_j \in C$, tem-se $g_i \leq g_j$ ou $g_j \leq g_i$. Suponha $g_i \leq g_j$. Então $D(g_j) \supset D(g_i)$ e $g_j(x) = g_i(x)$ para todo $x \in D(g_i)$.

Portanto, $x, y \in D(g_j)$ e, para $\alpha, \beta \in \mathbb{R}$, tem-se $\alpha x + \beta y \in D(g_j) \subset D(g)$ e $g(\alpha x + \beta y) = g_j(\alpha x + \beta y) = \alpha g_j(x) + \beta g_j(y) = \alpha g(x) + \beta g(y)$.

Por fim, $g(x) = f(x)$ para todo $x \in Z$ e $g(x) \leq p(x)$ para todo $x \in D(g)$. De fato, seja $x \in Z$. Como $Z \subset D(g_i)$ para todo $i \in I$ e $g_i(x) = f(x)$, conclui-se que

$g(x) = g_i(x) = f(x)$. Além disso, dado $x \in D(g)$, tem-se $x \in D(g_i)$ para algum $i \in I$. Logo,

$$g(x) = g_i(x) \leq p(x).$$

Então $g \in E$, $D(g_i) \subset D(g)$ para todo $i \in I$ e $g(x) = g_i(x)$ para todo $x \in D(g_i)$, ou seja, g é um limitante superior de C .

Pelo Lema de Zorn, E possui um elemento maximal. Tal elemento será denotado por $\tilde{f} : D(\tilde{f}) \rightarrow \mathbb{R}$.

(b) Para mostrar que $D(\tilde{f}) = X$ suponha que $D(\tilde{f}) \neq X$, ou seja, existe $x_0 \in X - D(\tilde{f})$. Assim, chame $D(\tilde{g}) = D(\tilde{f}) + [x_0]$. Dessa forma, $D(\tilde{f}) \subset D(\tilde{g})$ e $D(\tilde{g})$ é um subespaço de X . Note que, para quais $x, y \in D(\tilde{f})$, tem-se:

$$\begin{aligned} \tilde{f}(x) + \tilde{f}(y) = \tilde{f}(x+y) &\leq p(x+y) = p(x+x_0+y-x_0) \\ &\leq p(x+x_0) + p(y-x_0) \end{aligned}$$

uma vez que p é sublinear e p está definida em X . Portanto,

$$\tilde{f}(y) - p(y-x_0) \leq p(x+x_0) - \tilde{f}(x)$$

para quaisquer $x, y \in D(\tilde{f})$. Deste modo tem-se,

$$\sup_{y \in D(\tilde{f})} \{\tilde{f}(y) - p(y-x_0)\} \leq \inf_{x \in D(\tilde{f})} \{p(x+x_0) - \tilde{f}(x)\}.$$

Tome $\alpha \in \mathbb{R}$ tal que

$$\sup_{y \in D(\tilde{f})} \{\tilde{f}(y) - p(y-x_0)\} \leq \alpha \leq \inf_{x \in D(\tilde{f})} \{p(x+x_0) - \tilde{f}(x)\} \quad (2.8)$$

e defina $\tilde{g} : D(\tilde{g}) \rightarrow \mathbb{R}$ por $\tilde{g}(x+tx_0) = \tilde{f}(x) + t\alpha$ para $x \in D(\tilde{f})$ e $t \in \mathbb{R}$. Como \tilde{g} é linear, por 2.8, tem-se

1. $\tilde{f}(y) - p(y-x_0) \leq \alpha \Leftrightarrow \tilde{f}(y) - \alpha \leq p(y-x_0)$
2. $\alpha \leq p(x+x_0) - \tilde{f}(x) \Leftrightarrow \tilde{f}(x) + \alpha \leq p(x+x_0)$.

Primeiramente considere $t > 0$. Dado $x \in D(\tilde{f})$, tem-se

$$\tilde{g}(x+tx_0) = \tilde{f}(x) + t\alpha = t \left[\tilde{f}\left(\frac{x}{t}\right) + \alpha \right] \stackrel{2}{\leq} t \cdot p\left(\frac{x}{t} + x_0\right) = p(x+tx_0).$$

E para $t < 0$ e $x \in D(\tilde{f})$, tem-se

$$\tilde{g}(x+tx_0) = \tilde{f}(x) + t\alpha = -t \left[\tilde{f}\left(-\frac{x}{t}\right) - \alpha \right] \stackrel{1}{\leq} -t \cdot p\left(-\frac{x}{t} - x_0\right) = p(x+tx_0).$$

Assim, \tilde{g} pertence a E e esta é um extensão de \tilde{f} contradizendo a maximalidade de \tilde{f} . \square

O Teorema Hahn Banach 2.10 diz respeito a espaços vetoriais reais. Uma generalização para espaços vetoriais complexos será dada no seguinte teorema.

Teorema 2.11. Teorema de Hahn-Banach (Generalizado). *Seja X um espaço vetorial real ou complexo e $p : X \rightarrow \mathbb{R}$ um funcional tal que, para todo $x, y \in X$,*

$$p(x + y) \leq p(x) + p(y) \quad (2.9)$$

como no Teorema 2.10, e para cada escalar α , tem-se

$$p(\alpha x) = |\alpha|p(x). \quad (2.10)$$

Além disso, seja f um funcional linear definido em um subespaço Z de X tal que

$$|f(x)| \leq p(x) \quad (2.11)$$

para todo $x \in Z$. Então f possui uma extensão linear \tilde{f} de Z para X satisfazendo

$$|\tilde{f}(x)| \leq p(x) \quad (2.12)$$

para todo $x \in X$.

Demonstração. Considere, inicialmente, X um espaço vetorial sobre \mathbb{R} . Por hipótese, $|f(x)| \leq p(x)$, para todo $x \in Z$. Mas $f(x) \leq |f(x)| \leq p(x)$ para todo $x \in Z$. Segue então que $f(x) \leq p(x)$, para todo $x \in Z$. Pelo Teorema de Hahn-Banach 2.10, existe um funcional linear $\tilde{f} : X \rightarrow \mathbb{R}$ tal que $\tilde{f}(x) = f(x)$ para todo $x \in Z$ e $\tilde{f}(x) \leq p(x)$ para todo $x \in X$.

Note também que

$$-\tilde{f}(x) = \tilde{f}(-x) \leq p(-x) = |-1|p(x) = p(x),$$

isto é, $\tilde{f}(x) \geq -p(x)$ para todo $x \in X$. Assim, conclui-se que

$$|\tilde{f}(x)| \leq p(x),$$

para todo $x \in X$ o que completa a prova.

Agora, considere X um espaço vetorial sobre \mathbb{C} . Então Z também é um espaço vetorial complexo. Como $f : Z \rightarrow \mathbb{C}$, pode-se escrever

$$f(x) = f_1(x) + if_2(x),$$

para todo $x \in Z$, onde f_1 e f_2 são funções de Z a valores reais. Como f é linear, f_1 e f_2 também o são. Por hipótese, tem-se

$$f_1(x) \leq |f_1(x)| \leq |f(x)| \leq p(x),$$

para todo $x \in Z$. Deste modo, existe um funcional linear $\tilde{f}_1 : X \rightarrow \mathbb{R}$ tal que $\tilde{f}_1(x) = f_1$ para todo $x \in Z$ e $|\tilde{f}_1(x)| \leq p(x)$ para todo $x \in X$.

Observe que,

$$i[f_1(x) + if_2(x)] = if(x) = f(ix) = f_1(ix) + if_2(ix),$$

para todo $x \in Z$. Daí $f_2(x) = -f_1(ix)$ e $f(x) = f_1(x) - if_1(ix)$, para todo $x \in Z$. Define-se $\tilde{f} : X \rightarrow \mathbb{C}$ por $\tilde{f}(x) = \tilde{f}_1(x) - i\tilde{f}_1(ix)$. Tem-se que \tilde{f} é linear. De fato, da linearidade de \tilde{f}_1 , para qualquer escalar $(a + ib) \in \mathbb{C}$, tem-se

$$\begin{aligned}\tilde{f}((a + ib)x) &= \tilde{f}_1(ax + ibx) - i\tilde{f}_1(iax - bx) \\ &= a\tilde{f}_1(x) + b\tilde{f}_1(ix) - i[a\tilde{f}_1(ix) - b\tilde{f}_1(x)] \\ &= (a + ib)[\tilde{f}_1(x) - i\tilde{f}_1(ix)] \\ &= (a + ib)\tilde{f}(x).\end{aligned}$$

Além disso, $\tilde{f}(x) = f(x)$ para todo $x \in Z$. Dado $x \in X$, pode-se escrever $\tilde{f}(x) = |\tilde{f}(x)|e^{i\theta}$ e, portanto, $|\tilde{f}(x)| = \tilde{f}(x)e^{-i\theta}$. Uma vez que $|\tilde{f}(x)| \in \mathbb{R}$, para $x \in X$, tem-se

$$|\tilde{f}(x)| = \tilde{f}(x)e^{-i\theta} = \tilde{f}(e^{-i\theta}x) = \tilde{f}_1(e^{-i\theta}x) \leq p(e^{-i\theta}x) = |e^{-i\theta}|p(x) = p(x).$$

Isto completa a prova. □

Teorema 2.12. Teorema de Hahn-Banach para Espaços Normados. *Seja f um funcional linear limitado em um subespaço Z de um espaço normado X . Então existe um funcional linear limitado \tilde{f} em X que é uma extensão de f para X tal que*

$$\|\tilde{f}\|_X = \|f\|_Z, \quad (2.13)$$

onde

$$\|\tilde{f}\|_X = \sup_{\substack{x \in X \\ \|x\|=1}} |\tilde{f}(x)|, \text{ e } \|f\|_Z = \sup_{\substack{x \in Z \\ \|x\|=1}} |f(x)|.$$

Demonstração. Se $Z = \{0\}$, então $f = 0$, e sua extensão é $\tilde{f} = 0$. Suponha, agora, $Z \neq \{0\}$. Para fazer uso do Teorema 2.11 deve-se primeiro determinar p . Como, para todo $x \in Z$, tem-se

$$|f(x)| \leq \|f\|_Z \|x\|,$$

é conveniente tomar $p(x) = \|f\|_Z \|x\|$, pois desta forma (2.11) fica satisfeita.

Vê-se que p está definida em todo X . Além disso,

$$p(x + y) = \|f\|_Z \|x + y\| \leq \|f\|_Z (\|x\| + \|y\|) = p(x) + p(y),$$

para todo $x, y \in X$, pela desigualdade triangular. Note também que

$$p(\alpha x) = \|f\|_Z \|\alpha x\| = |\alpha| \|f\|_Z \|x\| = |\alpha| p(x),$$

para $\alpha \in \mathbb{K}$. Assim (2.9) e (2.10) estão satisfeitas, concluindo as hipóteses do Teorema 2.11. Portanto existe um funcional linear \tilde{f} em X que é uma extensão de f satisfazendo

$$|\tilde{f}(x)| \leq p(x) = \|f\|_Z \|x\|,$$

para todo $x \in X$. Tomando o supremo sobre todos $x \in X$ de norma 1, obtém-se a desigualdade

$$\|\tilde{f}\|_X = \sup_{\substack{x \in X \\ \|x\|=1}} |\tilde{f}(x)| \leq \|f\|_Z.$$

Tem-se também $\|\tilde{f}\|_X \geq \|f\|_Z$ pois a norma sobre uma extensão não pode diminuir.

Portanto $\|\tilde{f}\|_X = \|f\|_Z$ e o teorema está provado. \square

Teorema 2.13. *Seja X um espaço normado e $x_0 \neq 0$ um elemento qualquer de X . Então existe um funcional linear limitado \tilde{f} em X tal que*

$$\|\tilde{f}\| = 1 \text{ e } \tilde{f}(x_0) = \|x_0\|.$$

Demonstração. Considere um subespaço Z de X constituído por todos os elementos $x = \alpha x_0$ onde α é um escalar. Defina $f : Z \rightarrow \mathbb{K}$ por

$$f(x) = f(\alpha x_0) = \alpha \|x_0\|, \quad (2.14)$$

Tem-se f limitado e $\|f\| = 1$ pois

$$|f(x)| = |f(\alpha x_0)| = |\alpha| \|x_0\| = \|\alpha x_0\| = \|x\|.$$

Assim tomando o supremo tem-se,

$$\|f\| = \sup_{\substack{x \in Z \\ x \neq 0}} \frac{|f(x)|}{\|x\|} = \sup_{\substack{x \in Z \\ x \neq 0}} \frac{\|x\|}{\|x\|} = 1.$$

O Teorema 2.12 implica que f possui uma extensão \tilde{f} de Z para X tal que $\|\tilde{f}\| = \|f\| = 1$. De (2.14) tem-se $\tilde{f}(x_0) = f(x_0) = \|x_0\|$. \square

2.4.2 A Desigualdade de Hölder

Dado $1 \leq p \leq \infty$, indica-se com p' o elemento de $[1, \infty]$, tal que $\frac{1}{p} + \frac{1}{p'} = 1$. Observe que $p'' = p$, $1' = \infty$ e $2' = 2$. Nestas condições diz-se que p e p' são **expoentes conjugados**.

Dado $x \in \mathbb{C}^n$ define-se

$$\|x\|_p = \left[\sum_{j=1}^n |x_j|^p \right]^{\frac{1}{p}},$$

para $p < \infty$.

Dado $f \in \mathcal{C}([a, b], \mathbb{C})$ define-se

$$\|f\|_p = \left[\int_a^b |f(t)|^p dt \right]^{\frac{1}{p}},$$

para $p < \infty$.

Lema 2.4. *Seja $1 < p < \infty$, para quaisquer $a, b \in \mathbb{R}_+$ tem-se*

$$ab \leq \frac{a^p}{p} + \frac{b^{p'}}{p'}$$

Demonstração. Considere a curva

$$b = a^{p-1}, \quad (a = b^{\frac{1}{p-1}} = b^{p'-1})$$

que é estritamente crescente (convexa se $p > 2$ e côncava se $p < 2$).

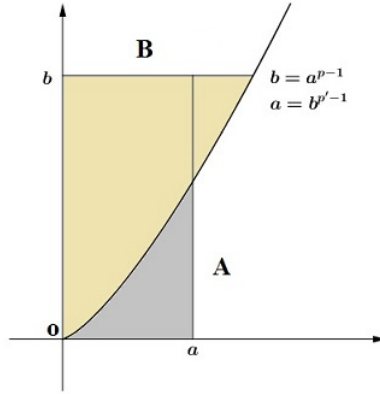


Figura 2.2: Lema 2.4

Observe, a partir da figura acima, que $ab \leq \text{área } A + \text{área } B$, onde A e B são as regiões cinza e bege, respectivamente, destacadas na figura. Assim

$$ab \leq \int_0^a t^{p-1} dt + \int_0^b t^{p'-1} dt = \frac{a^p}{p} + \frac{b^{p'}}{p'}.$$

□

Teorema 2.14. (Desigualdade de Hölder) *Sejam $1 < p, p' < \infty$ tais que $\frac{1}{p} + \frac{1}{p'} = 1$;*

a) *Dados $x, y \in \mathbb{C}^n$, tem-se $\|xy\|_1 \leq \|x\|_p \|y\|_{p'}$, isto é,*

$$\sum_{j=1}^n |x_j y_j| \leq \left[\sum_{j=1}^n |x_j|^p \right]^{\frac{1}{p}} \cdot \left[\sum_{j=1}^n |y_j|^{p'} \right]^{\frac{1}{p'}},$$

quando $1 < p < \infty$.

b) *Dados $f, g \in \mathcal{C}([a, b], \mathbb{C})$ tem-se $\|fg\|_1 \leq \|f\|_p \|g\|_{p'}$, isto é,*

$$\int_a^b |f(t)g(t)| dt \leq \left[\int_a^b |f(t)|^p dt \right]^{\frac{1}{p}} \cdot \left[\int_a^b |g(t)|^{p'} dt \right]^{\frac{1}{p'}},$$

quando $1 < p < \infty$.

Demonstração. a) Para o caso em que $p = 1$ ou $p' = \infty$ não há o que mostrar, pela mesma razão suponha $x \neq 0$ e $y \neq 0$. Assim, aplicando o Lema 2.4 ao pares

$$a_j = \frac{|x_j|}{\|x\|_p} \text{ e } b_j = \frac{|y_j|}{\|y\|_{p'}},$$

com $j = 1, 2, \dots, n$, tem-se

$$a_j b_j = \frac{|x_j|}{\|x\|_p} \cdot \frac{|y_j|}{\|y\|_{p'}} = \frac{|x_j y_j|}{\|x\|_p \|y\|_{p'}} \leq \frac{1}{p} \cdot \left(\frac{|x_j|}{\|x\|_p} \right)^p + \frac{1}{p'} \cdot \left(\frac{|y_j|}{\|y\|_{p'}} \right)^{p'}.$$

Efetuando a soma membro a membro da desigualdade acima obtém-se

$$\sum_{j=1}^n \left(\frac{|x_j y_j|}{\|x\|_p \|y\|_{p'}} \right) \leq \frac{1}{p} \cdot \frac{1}{(\|x\|_p)^p} \sum_{j=1}^n |x_j|^p + \frac{1}{p'} \cdot \frac{1}{(\|y\|_{p'})^{p'}} \sum_{j=1}^n |y_j|^{p'} = \frac{1}{p} + \frac{1}{p'} = 1.$$

Portanto

$$\sum_{j=1}^n \left(\frac{|x_j y_j|}{\|x\|_p \|y\|_{p'}} \right) = \frac{1}{\|x\|_p \|y\|_{p'}} \sum_{j=1}^n |x_j y_j| \leq 1,$$

ou seja,

$$\sum_{j=1}^n |x_j y_j| \leq \|x\|_p \|y\|_{p'} = \left[\sum_{j=1}^n |x_j|^p \right]^{\frac{1}{p}} \cdot \left[\sum_{j=1}^n |y_j|^{p'} \right]^{\frac{1}{p'}}.$$

b) Segue de modo análogo ao item a), aplicando o Lema 2.4 e trocando $\sum_{j=1}^n$ por \int_a^b . \square

Exemplo 2.19. Seja $1 \leq p \leq \infty$; $\mathcal{C}_{L_p}([a, b], \mathbb{C})$ indica o espaço vetorial $\mathcal{C}([a, b], \mathbb{C})$ munido da norma $f \in \mathcal{C}([a, b], \mathbb{C}) \mapsto \|f\|_p$, onde

$$\|f\|_p = \left[\int_a^b |f(x)|^p dx \right]^{\frac{1}{p}},$$

quando $p < \infty$ e

$$\|f\|_\infty = \sup_{a \leq x \leq b} |f(x)|.$$

2.4.3 Somabilidade em Espaços Normados

A definição habitual de série convergente de números reais ou complexos se estende naturalmente aos espaços normados: dada uma sequência $(x_n)_{n \in \mathbb{N}}$ de elementos de um espaço normado X , diz-se que a série

$$\sum_{n=1}^{\infty} x_n$$

é convergente para o elemento $x \in E$, e se escreve

$$\sum_{n=1}^{\infty} x_n = x,$$

e para a sequência das reduzidas

$$s_m = \sum_{n=1}^m x_n$$

tem-se $s_m \rightarrow x$ quando $m \rightarrow \infty$.

Mas em espaços normados e especialmente em espaços de Hilbert tem-se a necessidade de dar um sentido a somatórias da forma $\sum_{i \in I} x_i$ quando o conjunto dos índices I não é necessariamente o conjunto \mathbb{N} dos naturais. Quando $I = \mathbb{Z}, \mathbb{N} \times \mathbb{N}, \mathbb{Z}^n$, etc. pode-se ordenar o conjunto dos índices de modo conveniente procurando recair no caso que nos é familiar. Neste caso a soma da série pode depender da ordem em que se efetuam as somas parciais. Para tanto, segue a seguinte definição.

Definição 2.20. *Dada uma família $(x_i)_{i \in I}$ de elementos de um espaço normado X , diz-se que esta família é **somável** e tem por **soma** $x \in X$, e se escreve*

$$\sum_{i \in I} x_i = x,$$

se dado $\epsilon > 0$ existe um subconjunto finito $F_\epsilon \subset I$ tal que para todo subconjunto finito $F \subset I$ com $F \supset F_\epsilon$ tem-se $\left\| x - \sum_{i \in F} x_i \right\| < \epsilon$.

A definição acima pode ser motivada a partir do seguinte fato: seja $(x_n)_{n \in \mathbb{N}}$ uma seqüência de números reais ou complexos. São equivalentes:

- a) A série $\sum_{n=1}^{\infty} x_n$ é comutativamente convergente para x onde, vale lembrar, uma série $\sum_{n=1}^{\infty} x_n$ diz-se ser comutativamente convergente para x se reordenando-a de um modo qualquer obtém-se sempre uma série convergente cuja soma é x .
- b) Dado $\epsilon > 0$ existe um subconjunto finito $F_\epsilon \subset \mathbb{N}$ tal que para qualquer subconjunto finito $F \subset \mathbb{N}$ com $F \supset F_\epsilon$ tem-se $\left| x - \sum_{n \in F} x_n \right| < \epsilon$.

As duas propriedades acima são equivalentes a dizer que a série $\sum_{x=1}^{\infty} x_n$ é absolutamente convergente, ou seja, $\sum_{x=1}^{\infty} |x_n|$ é convergente.

Proposição 2.3. *Sejam X e Y espaços normados e $f : X \rightarrow Y$ uma aplicação linear contínua. Dada a família $(x_i)_{i \in I}$ de elementos de X tal que $\sum_{i \in I} x_i = x$, então a família $(f(x_i))_{i \in I}$ de Y é tal que $\sum_{i \in I} f(x_i) = f(x)$.*

Demonstração. Segue da definição de família somável e da seguinte desigualdade

$$\left\| f(x) - \sum_{i \in F} f(x_i) \right\| = \left\| f\left(x - \sum_{i \in F} x_i\right) \right\| \leq \|f\| \left\| x - \sum_{i \in F} x_i \right\|.$$

□

Do mesmo modo que para séries de números reais ou complexos tem-se o critério de Cauchy, que permite uma condição necessária e suficiente para que uma série seja convergente sem a necessidade de conhecer a soma da mesma. Deste modo, tem-se também um critério para famílias somáveis.

Definição 2.21. Diz-se que uma família $(x_i)_{i \in I}$ de elementos de um espaço normado X satisfaz a **condição de Cauchy** se dado $\epsilon > 0$ existe um conjunto finito $F_\epsilon \subset I$ tal que para todo subconjunto finito $F' \subset I$ e disjunto de F_ϵ tem-se

$$\left\| \sum_{i \in F'} x_i \right\| < \epsilon.$$

Proposição 2.4. Toda família somável $(x_i)_{i \in I}$ satisfaz a condição de Cauchy.

Demonstração. Seja $x = \sum_{i \in I} x_i$, assim dado $\epsilon > 0$ existe $F_\epsilon \subset I$ tal que para todo $F \supset F_\epsilon$, com $F \subset I$, tem-se

$$\left\| x - \sum_{i \in F} x_i \right\| < \epsilon.$$

Tomando qualquer subconjunto finito $F' \subset I$ com $F' \cap F_\epsilon = \emptyset$ segue

$$\begin{aligned} \left\| \sum_{i \in F'} x_i \right\| &= \left\| \sum_{i \in F' \cup F_\epsilon} x_i - \sum_{i \in F_\epsilon} x_i + x - x \right\| \\ &= \left\| \sum_{i \in F' \cup F_\epsilon} x_i - x - \left(\sum_{i \in F_\epsilon} x_i - x \right) \right\| \\ &\leq \left\| \sum_{i \in F' \cup F_\epsilon} x_i - x \right\| + \left\| \sum_{i \in F_\epsilon} x_i - x \right\| < 2\epsilon. \end{aligned}$$

□

Proposição 2.5. Se a família $(x_i)_{i \in I}$ satisfaz a condição de Cauchy então o subconjunto $I^* = \{i \in I \mid x_i \neq 0\}$ é enumerável.

Demonstração. Considere o conjunto $I_n = \left\{ i \in I \mid \|x_i\| \geq \frac{1}{n} \right\}$, pela definição da condição de Cauchy cada I_n é finito, assim basta notar que $I^* = \bigcup_{n=1}^{\infty} I_n$. □

Teorema 2.15. Seja X um espaço de Banach. Uma condição necessária e suficiente para que uma família $(x_i)_{i \in I}$ seja somável é que ela satisfaça a condição de Cauchy.

Demonstração. A demonstração pode ser encontrada no capítulo I, da referência [7]. □

Definição 2.22. Diz-se que uma família $(x_i)_{i \in I}$ de elementos de um espaço normado X é **absolutamente somável** se a família $(\|x_i\|)_{i \in I}$ (de números reais) for somável.

3 Espaços de Hilbert

Com o texto que se segue aborda-se a teoria referente aos Espaços de Hilbert, o Teorema da Representação de Riez e define-se a Forma Sesquilinear. O texto é baseado na referência [3].

Definição 3.1. *Um espaço com produto interno (ou espaço pré-hilbertiano) é um espaço vetorial X com produto interno definido em X . Um espaço de Hilbert é um espaço com produto interno completo (completo em relação à métrica definida pelo produto interno).*

Aqui, um produto interno sobre X é uma função que associa a cada par ordenado de vetores x e y em X um escalar $\langle x, y \rangle$ no corpo \mathbb{K} de tal maneira que para quaisquer x, y, z em X e todo escalar α tem-se

$$(PI1) - \langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$$

$$(PI2) - \langle \alpha x, y \rangle = \alpha \langle x, y \rangle$$

$$(PI3) - \langle x, y \rangle = \overline{\langle y, x \rangle}, \text{ onde a barra indica conjugação complexa;}$$

$$(PI4) - \begin{aligned} \langle x, x \rangle &\geq 0 \\ \langle x, x \rangle &= 0 \Leftrightarrow x = 0. \end{aligned}$$

Um produto interno em X define uma norma em X dada por

$$\|x\| = \sqrt{\langle x, x \rangle} \tag{3.1}$$

e uma métrica em X é dada por

$$d(x, y) = \|x - y\| = \sqrt{\langle x - y, x - y \rangle}. \tag{3.2}$$

Consequentemente, espaços com produto interno são espaços normados, e espaços de Hilbert são espaços de Banach.

Quando \mathbb{K} é o corpo \mathbb{R} , dos números reais, os complexos conjugados que aparecem em (PI3) são supérfluos; no entanto, no caso de \mathbb{K} ser complexo eles são necessários para a consistência das condições. Sem estes complexos conjugados, haveria uma contradição:

$$\langle x, x \rangle \text{ e } \langle ix, ix \rangle = -1\langle x, x \rangle > 0.$$

Em primeiro lugar, deve-se verificar que (3.1) define uma norma. Para tanto, precisa-se da seguinte definição e do próximo Lema.

Definição 3.2. Um elemento x de um espaço com produto interno X é dito ser **ortogonal** ao elemento $y \in X$ se $\langle x, y \rangle = 0$.

Diz-se, também, que x e y são ortogonais, e escreve-se $x \perp y$. Similarmente, para subconjuntos A, B de X escreve-se $x \perp A$ se $x \perp a$ para todo $a \in A$, e $A \perp B$ se $a \perp b$ para todo $a \in A$ e todo $b \in B$.

Lema 3.1. (Desigualdade de Schwarz). Se X é um espaço com produto interno, então, para quaisquer x e y em X temos $|\langle x, y \rangle| \leq \|x\| \|y\|$.

Demonstração. (a) Se $x = 0$, tem-se

$$\langle x, y \rangle = \langle 0, y \rangle = \langle 0 + 0, y \rangle = \langle 0, y \rangle + \langle 0, y \rangle \Rightarrow \langle 0, y \rangle = 0. \text{ Portanto } |\langle x, y \rangle| = |\langle 0, y \rangle| = |0| \leq \|0\| \|y\| = 0.$$

(b) Se $x \neq 0$, considerar a projeção ortogonal de y na direção de x .

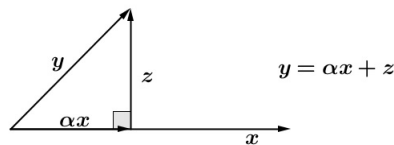


Figura 3.1: Projeção ortogonal de y na direção de x .

Assim,

$$\begin{aligned} \langle y, x \rangle &= \langle \alpha x + z, x \rangle = \langle \alpha x, x \rangle + \langle z, x \rangle = \alpha \langle x, x \rangle + \langle z, x \rangle = \alpha \langle x, x \rangle \\ &\Rightarrow \alpha \langle x, x \rangle = \langle y, x \rangle \\ &\Rightarrow \alpha = \frac{\langle y, x \rangle}{\|x\|^2}, \text{ pois } x \neq 0 \end{aligned}$$

Logo, $z = y - \alpha x = y - \frac{\langle y, x \rangle}{\|x\|^2} x$, e

$$\begin{aligned} 0 \leq \|z\|^2 &= \langle z, z \rangle = \langle z, y - \alpha x \rangle = \langle z, y \rangle - \bar{\alpha} \langle z, x \rangle = \langle z, y \rangle \\ &= \langle y - \alpha x, y \rangle = \langle y, y \rangle - \alpha \langle x, y \rangle \\ &= \|y\|^2 - \frac{\langle y, x \rangle}{\|x\|^2} \langle x, y \rangle = \|y\|^2 - \frac{\overline{\langle x, y \rangle}}{\|x\|^2} \langle x, y \rangle \\ &= \|y\|^2 - \frac{|\langle x, y \rangle|^2}{\|x\|^2} \\ &\Rightarrow \|y\|^2 - \frac{|\langle x, y \rangle|^2}{\|x\|^2} \geq 0. \end{aligned}$$

Portanto, $|\langle x, y \rangle| \leq \|x\| \|y\|$.

□

Prova-se, agora, que 3.1 define uma norma. Com efeito. (N1) e (N2) da definição 2.15 seguem de (PI4). Além disso (N3) é obtida pelo uso de (PI2) e (PI3); de fato

$$\|\alpha x\| = \sqrt{\langle \alpha x, \alpha x \rangle} = \sqrt{|\alpha|^2 \langle x, x \rangle} = |\alpha| \|x\|.$$

Finalmente, para (N4), tem-se

$$\begin{aligned} \|x + y\|^2 &= \langle x + y, x + y \rangle = \langle x, x + y \rangle + \langle y, x + y \rangle \\ &= \langle x, x \rangle + \langle x, y \rangle + \overline{\langle x, y \rangle} + \langle y, y \rangle \\ &= \|x\|^2 + \langle x, y \rangle + \overline{\langle x, y \rangle} + \|y\|^2 \\ &= \|x\|^2 + \operatorname{Re}(\langle x, y \rangle) + \operatorname{Im}(\langle x, y \rangle) + \operatorname{Re}(\langle x, y \rangle) - \operatorname{Im}(\langle x, y \rangle) + \|y\|^2 \\ &= \|x\|^2 + 2\operatorname{Re}(\langle x, y \rangle) + \|y\|^2 \\ &\leq \|x\|^2 + 2|\langle x, y \rangle| + \|y\|^2. \end{aligned}$$

Pela Desigualdade de Schwarz

$$|\langle x, y \rangle| \leq \|x\| \|y\|,$$

tem-se,

$$\begin{aligned} \|x + y\|^2 &\leq \|x\|^2 + 2|\langle x, y \rangle| + \|y\|^2 \\ &\leq \|x\|^2 + 2\|x\| \|y\| + \|y\|^2 \\ &= (\|x\| + \|y\|)^2. \end{aligned}$$

Portanto

$$\|x + y\| \leq \|x\| + \|y\|.$$

Proposição 3.1. *Dado um espaço pré-hilbertiano H , para todo $y \in H$ o funcional linear $f_y : H \rightarrow \mathbb{C}$, dado por $f_y(x) = \langle x, y \rangle$, é linear e contínuo e tem-se $\|f_y\| = \|y\|$.*

Demonstração. A linearidade de f_y segue da definição 3.1. Para a continuidade de f_y , note que se $y = 0$ não há o que provar, assim suponha $y \neq 0$ e $x_0 \in H$ um elemento qualquer, dado $\epsilon > 0$ basta tomar $\delta = \frac{\epsilon}{\|y\|}$, então $\|x - x_0\| < \delta$ implica $|f_y(x) - f_y(x_0)| = |\langle x, y \rangle - \langle x_0, y \rangle| = |\langle x - x_0, y \rangle| \leq \|x - x_0\| \|y\| < \epsilon$. Ainda, da Desigualdade de Cauchy-Schwarz segue que $\|f_y\| \leq \|y\|$ e tomando $x = y$ em $f_y(x) = \langle x, y \rangle$ segue que $\|f_y\| = \|y\|$. \square

Proposição 3.2. *Seja $(x_i)_{i \in I}$ uma família somável de elementos de um espaço pré-hilbertiano H , tendo por soma x . Então para todo $y \in H$ a família $\langle x_i, y \rangle$, $i \in I$, de números complexos é somável e tem por soma $\langle x, y \rangle$.*

Demonstração. Segue das Proposições 2.3 e 3.1. \square

Lema 3.2. *Se em um espaço com produto interno, $\lim_{n \rightarrow \infty} x_n = x$ e $\lim_{n \rightarrow \infty} y_n = y$, então $\lim_{n \rightarrow \infty} \langle x_n, y_n \rangle = \langle x, y \rangle$.*

Demonstração. Das desigualdades de Schwarz e triangular, tem-se

$$\begin{aligned} |\langle x_n, y_n \rangle - \langle x, y \rangle| &= |\langle x_n, y_n \rangle - \langle x_n, y \rangle + \langle x_n, y \rangle - \langle x, y \rangle| \\ &= |\langle x_n, y_n - y \rangle + \langle x_n - x, y \rangle| \leq |\langle x_n, y_n - y \rangle| + |\langle x_n - x, y \rangle| \\ &\leq \|x_n\| \|y_n - y\| + \|x_n - x\| \|y\| \end{aligned}$$

que converge para 0, uma vez $\lim_{n \rightarrow \infty} (x_n - x) = 0$ e $\lim_{n \rightarrow \infty} (y_n - y) = 0$.
 Portanto, $\lim_{n \rightarrow \infty} \langle x_n, y_n \rangle = \langle x, y \rangle$.

□

Proposição 3.3. *Em um espaço pré-hilbertiano H vale a **Lei do Paralelogramo**: para quaisquer $x, y \in H$ tem-se*

$$\|x + y\|^2 + \|x - y\|^2 = 2[\|x\|^2 + \|y\|^2].$$

Demonstração. Basta desenvolver o primeiro membro.

□

Um **subespaço** Y de um espaço com produto interno X é definido como sendo um subespaço vetorial de X munido com o produto interno em X restrito a $Y \times Y$.

Similarmente, um **subespaço** Y de um espaço de Hilbert H é definido como sendo um subespaço de H , considerado como um espaço com produto interno. Note que Y não precisa ser um espaço de Hilbert porque Y pode não ser completo.

3.1 Complemento Ortogonal e Soma Direta

Em um espaço métrico X , a distância δ de um elemento $x \in X$ para um conjunto não vazio $M \subset X$ é definida como sendo

$$\delta = \inf_{\tilde{y} \in M} d(x, \tilde{y}).$$

Em um espaço normado isto torna-se

$$\delta = \inf_{\tilde{y} \in M} \|x - \tilde{y}\|. \quad (3.3)$$

É de grande importância saber se existe $y \in M$ de tal modo que

$$\delta = \|x - y\|, \quad (3.4)$$

isto é, intuitivamente falando, um ponto $y \in M$ que está mais próximo de um dado ponto x , e se este é único. Neste caso tem-se um problema de existência e unicidade.

A figura abaixo ilustra que mesmo em um espaço simples, tal como o plano Euclidiano \mathbb{R}^2 , pode não haver y satisfazendo (3.4), ou precisamente um único y , ou mais que um y . Assim, pode-se esperar que em espaços de dimensão infinita tenha-se algo mais complicado a esse respeito. Para espaços normados gerais esse é o caso, mas para espaços de Hilbert, a situação permanece relativamente simples.

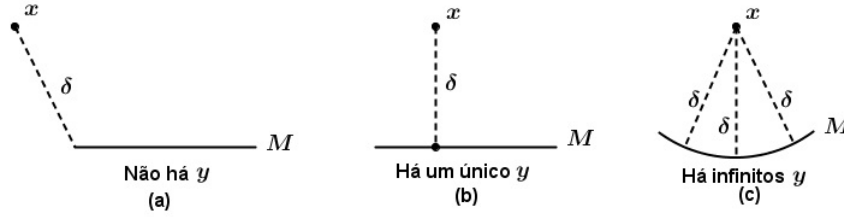


Figura 3.2: Existência e unicidade de pontos $y \in M$ satisfazendo 3.4, onde $M \subset \mathbb{R}^2$ é um segmento aberto [em (a) e (b)] e um arco circular [em (c)]

Para considerar o problema de existência e unicidade em espaços de Hilbert e para formular o teorema abaixo precisa-se de dois conceitos relacionados, que são de interesse geral, como segue.

O **segmento** que une dois elementos dados x e y de um espaço vetorial X é definido como sendo o conjunto de todos os $z \in X$ da forma

$$z = \alpha x + (1 - \alpha)y, \quad \alpha \in \mathbb{R}, 0 \leq \alpha \leq 1.$$

Um subconjunto M de X é dito ser **convexo** se para cada $x, y \in M$ o segmento que une x e y está contido em M .

Por exemplo, todo subespaço Y de X é convexo, e a interseção de conjuntos convexos é um conjunto convexo.

Teorema 3.1. *Sejam X um espaço com produto interno e $M \neq \emptyset$ um subconjunto convexo que é completo (na métrica induzida pelo produto interno). Então, para cada $x \in X$ dado, existe um único $y \in M$ tal que,*

$$\delta = \inf_{\tilde{y} \in M} \|x - \tilde{y}\| = \|x - y\|. \quad (3.5)$$

Demonstração. (a) *Existência.* Seja $\delta = \inf_{\tilde{y} \in M} \|x - \tilde{y}\|$, pela definição de ínfimo, para cada ϵ da forma $\frac{1}{n}$ existe um $y_n \in M$ de modo que $\delta < \delta_n + \frac{1}{n}$, onde $\delta_n = \|x - y_n\|$. Logo existe uma sequência $(y_n) \in M$ tal que

$$\delta_n \rightarrow \delta, \quad (3.6)$$

quando $n \rightarrow \infty$. A sequência (y_n) , assim construída, é de Cauchy. Com efeito, escrevendo $y_n - x = v_n$ tem-se $\|v_n\| = \delta_n$ e

$$\|v_n + v_m\| = \|y_n + y_m - 2x\| = 2\left\|\frac{1}{2}(y_n + y_m) - x\right\| \geq 2\delta,$$

pois $\frac{1}{2}(y_n + y_m) \in M$, já que M é convexo. Além disso $y_n - y_m = v_n - v_m$. Consequentemente pela Lei do Paralelogramo, Proposição 3.3,

$$\begin{aligned} \|y_n - y_m\|^2 &= \|v_n - v_m\|^2 = -\|v_n + v_m\|^2 + 2(\|v_n\|^2 + \|v_m\|^2) \\ &\leq -(2\delta)^2 + 2(\delta_n^2 + \delta_m^2), \end{aligned}$$

e (3.6), implica que (y_n) é de Cauchy.

Como M é completo, (y_n) converge para algum $y \in M$. Assim, tem-se $\|x-y\| \geq \delta$. Também, por (3.6),

$$\|x-y\| = \|x-y_n + y_n - y\| \leq \|x-y_n\| + \|y_n - y\| = \delta_n + \|y_n - y\| \rightarrow \delta,$$

quando $n \rightarrow \infty$. Logo $\|x-y\| \leq \delta$ e como $\|x-y\| \geq \delta$, conclui-se que $\|x-y\| = \delta$.

(b) *Unicidade.* Sejam $y \in M$ e $y_0 \in M$ tais que $\|x-y\| = \delta$ e $\|x-y_0\| = \delta$. Pela igualdade do paralelogramo

$$\begin{aligned} \|y-y_0\|^2 &= \|(y-x) - (y_0-x)\|^2 \\ &= 2\|y-x\|^2 + 2\|y_0-x\|^2 - \|(y-x) + (y_0-x)\|^2 \\ &= 2\delta^2 + 2\delta^2 - 2 \cdot 2\left\|\frac{1}{2}(y+y_0) - x\right\|^2. \end{aligned}$$

Como $\frac{1}{2}(y+y_0) \in M$, segue que

$$\left\|\frac{1}{2}(y+y_0) - x\right\| \geq \delta.$$

Isto implica que o lado direito é menor ou igual a $2\delta^2 + 2\delta^2 - 4\delta^2 = 0$. Assim, $\|y-y_0\| \leq 0$. Claramente, $\|y-y_0\| \geq 0$, concluindo que $y_0 = y$. □

Lema 3.3. (Ortogonalidade). No Teorema 3.1, seja M um subespaço completo Y de X e $x \in X$ fixado. Então $z = x - y$ é ortogonal a Y .

Demonstração. Suponha que z não seja ortogonal a Y , então existe um $y_1 \in Y$ tal que

$$\langle z, y_1 \rangle = \beta \neq 0. \quad (3.7)$$

Claramente, $y_1 \neq 0$, pois do contrário $\langle z, y_1 \rangle = 0$. Além disso, para qualquer escalar α ,

$$\begin{aligned} \|z - \alpha y_1\|^2 &= \langle z - \alpha y_1, z - \alpha y_1 \rangle = \langle z, z - \alpha y_1 \rangle - \langle \alpha y_1, z - \alpha y_1 \rangle \\ &= \langle z, z \rangle - \bar{\alpha} \langle z, y_1 \rangle - \alpha \langle y_1, z \rangle + \alpha \bar{\alpha} \langle y_1, y_1 \rangle \\ &= \langle z, z \rangle - \bar{\alpha} \langle z, y_1 \rangle - \alpha [\langle y_1, z \rangle - \bar{\alpha} \langle y_1, y_1 \rangle] \\ &= \langle z, z \rangle - \bar{\alpha} \beta - \alpha [\beta - \bar{\alpha} \langle y_1, y_1 \rangle]. \end{aligned}$$

Para $\bar{\alpha}$ da forma

$$\bar{\alpha} = \frac{\bar{\beta}}{\langle y_1, y_1 \rangle},$$

a expressão em colchetes $[\dots]$ é igual a zero. De (3.5), tem-se $\|z\| = \|x-y\| = \delta$, de modo que nossa equação produz agora

$$\|z - \alpha y_1\|^2 = \|z\|^2 - \frac{|\beta|^2}{\langle y_1, y_1 \rangle} < \delta^2.$$

Mas isto é impossível, pois

$$z - \alpha y_1 = x - y_2,$$

onde $y_2 = y + \alpha y_1 \in Y$, contradizendo $\|z - \alpha y_1\| \geq \delta$ pela definição do δ . Portanto, (3.7) não pode ocorrer, e o Lema está provado. \square

Definição 3.3. (Soma Direta). Um espaço vetorial X é dito ser uma soma direta de dois subespaços Y e Z de X , escreve-se

$$X = Y \oplus Z,$$

se cada $x \in X$ tem uma única representação

$$x = y + z$$

com $y \in Y$ e $z \in Z$.

Então Z é chamado complemento algébrico de Y em X e vice-versa, e Y e Z são subespaços complementares em X .

Por exemplo, $Y = \mathbb{R}$ é um subespaço do plano Euclidiano \mathbb{R}^2 . Claramente, Y tem infinitos complementos algébricos em \mathbb{R}^2 , cada qual é uma reta real. Mas o mais conveniente é o complemento que é perpendicular. Tal fato é usado quando se escolhe o sistema de coordenadas Cartesianas.

Similarmente, no caso de um espaço de Hilbert H qualquer, o interesse principal se diz respeito a representação de H como soma direta de um subespaço fechado Y e o seu **complemento ortogonal**, definido como

$$Y^\perp = \{z \in H | z \perp Y\}$$

que é o conjunto de todos os vetores ortogonais a Y .

Lema 3.4. Seja X um espaço vetorial e Y, Z dois subespaços de X tais que $X = Y \oplus Z$. Então, $Y \cap Z = \{0\}$.

Demonstração. Como cada elemento de X se escreve como soma de elementos de Y e Z , então $X = Y + Z$. Suponha agora que $Y \cap Z$ tenha um elemento não nulo w . Observe então que w pode ser escrito como $w = 0 + w$ para $0 \in Y$ e $w \in Z$ e também como $w = w + 0$ para $w \in Y$ e $0 \in Z$, o que contradiz a hipótese da unicidade. Logo $Y \cap Z = \{0\}$ e o resultado está provado. \square

Teorema 3.2. (Soma Direta). Seja Y um subespaço fechado qualquer de um espaço de Hilbert H . Então

$$H = Y \oplus Y^\perp \tag{3.8}$$

Demonstração. Como H é completo e Y é fechado, Y é completo pelo Teorema ???. Sendo Y convexo, o Teorema 3.1 e o Lema 3.2 implicam que para cada $x \in H$ existe $y \in Y$ tal que

$$x = y + z \tag{3.9}$$

onde z é ortogonal a Y , isto é, $z \in Y^\perp$.

Para provar a unicidade, suponha que

$$x = y + z = y_1 + z_1$$

onde $y, y_1 \in Y$ e $z, z_1 \in Y^\perp$. Então $y - y_1 = z - z_1$. Como Y é um subespaço de H , $y - y_1 \in Y$ ao passo que $z_1 - z \in Y^\perp$ pois Y^\perp também é um subespaço de H . Pelo Lema 3.4 e o fato de que $y - y_1 \in Y \cap Y^\perp$, conclui-se que $y = y_1$. Consequentemente $z = z_1$. \square

O elemento y em (3.9) é chamado de **projeção ortogonal** de x em Y . Este termo pode ser motivado pela geometria elementar. Por exemplo, para $H = \mathbb{R}^2$ a projeção de qualquer ponto $x = (x_0, y_0)$ no eixo das abcissas, que passa a desempenhar o papel de Y , é o ponto $y = (x_0, 0)$.

A equação (3.9) define uma aplicação

$$\begin{aligned} P : H &\longrightarrow Y \\ x &\longmapsto y = Px. \end{aligned}$$

A aplicação P é chamada de **projeção ortogonal** (ou *operador projeção*) de H em Y .

Proposição 3.4. *Sejam H um espaço de Hilbert e Y um subespaço fechado de H . O operador projeção $P : H \longrightarrow Y$ é tal que,*

- (a) P é sobrejetiva.
- (b) P leva Y em si mesmo.
- (c) P é idempotente, isto é, $P^2 = P$.

Demonstração. (a) e (b) seguem de imediato da definição de P . Para provar (c), considere $x \in H$, pelo Teorema 3.2, $x = y + z$ com $y \in Y$ e $z \in Y^\perp$. Assim $P^2(x) = P(Px) = P(y) = y = Px$. \square

Definição 3.4. *Sejam X e Y dois espaços vetoriais sobre o corpo \mathbb{K} e seja $f : \mathcal{D}(f) \subset X \longrightarrow Y$ uma aplicação linear. O **núcleo** $\mathcal{N}(f)$, ou $\ker(f)$, de f é o conjunto de todos $x \in \mathcal{D}(f)$ tais que $f(x) = 0$.*

Lema 3.5. *O complemento ortogonal Y^\perp de um subespaço fechado Y de um espaço de Hilbert H é o núcleo $\mathcal{N}(P)$ da projeção ortogonal P de H em Y .*

Demonstração. Seja $x \in Y^\perp$, como $x \in H = Y \oplus Y^\perp$ segue que $x = 0 + x$ com $0 \in Y$. Portanto $Px = 0$ e $x \in \mathcal{N}(P)$. Reciprocamente, se $x \in \mathcal{N}(P)$ tem-se $Px = 0 = P(0 + x)$, com $x \in Y^\perp$. \square

Um complemento ortogonal é um anulador especial, onde, por definição, o *anulador* M^\perp de um conjunto $M \neq \emptyset$ em um espaço com produto interno X é o conjunto

$$M^\perp = \{x \in X \mid x \perp M\}$$

Assim, $x \in M^\perp$ se, e somente se, $\langle x, v \rangle = 0$ para todo $v \in M$. Isto explica o nome.

Note que M^\perp é um subespaço vetorial de X . Com efeito, sejam $x, y \in M^\perp$, para todo $v \in M$ e quaisquer escalares α, β

$$\begin{aligned} \langle \alpha x + \beta y, v \rangle &= \langle \alpha x, v \rangle + \langle \beta y, v \rangle \\ &= \alpha \langle x, v \rangle + \beta \langle y, v \rangle \\ &= \alpha 0 + \beta 0 = 0, \end{aligned}$$

portanto $\alpha x + \beta y \in M^\perp$.

M^\perp é fechado. Para provar que M^\perp é fechado, seja $x \in \overline{M^\perp}$, então existe uma sequência de pontos $(x_n) \in M^\perp$ que converge para x . Para cada $v \in M$ tem-se $\langle x_n, v \rangle = 0$ para todo n . Observe que

$$|\langle x_n, v \rangle - \langle x, v \rangle| = |\langle x_n - x, v \rangle| \leq \|x_n - x\| \|v\|,$$

que converge para 0, uma vez que $\lim_{n \rightarrow \infty} x_n = x$. Assim $\lim_{n \rightarrow \infty} \langle x_n, v \rangle = \langle x, v \rangle$ e

$$\langle x_n, v \rangle = 0 \Rightarrow \lim_{n \rightarrow \infty} \langle x_n, v \rangle = 0 \Rightarrow \langle x, v \rangle = 0.$$

Portanto $x \in M^\perp$ e M^\perp é fechado.

O conjunto $(M^\perp)^\perp$ é também denotado por $M^{\perp\perp}$. Para espaços vetoriais de dimensão finita tem-se sempre $M = M^{\perp\perp}$, porém em espaços vetoriais de dimensão infinita pode-se garantir

$$M \subset M^{\perp\perp} \tag{3.10}$$

pois

$$x \in M \Rightarrow x \perp M^\perp \Rightarrow x \in (M^\perp)^\perp.$$

Quando o subespaço for fechado pode-se garantir a outra inclusão, como segue no teorema abaixo.

Lema 3.6. (Subespaço Fechado.) *Se Y é um subespaço fechado de um espaço de Hilbert H , então*

$$Y = (Y^\perp)^\perp. \tag{3.11}$$

Demonstração. Tem-se que $Y \subset (Y^\perp)^\perp$ por (3.10). Para mostrar que $Y \supset (Y^\perp)^\perp$, considere $x \in (Y^\perp)^\perp$. Então $x = y + z$, pelo Teorema 3.2, onde $y \in Y \subset (Y^\perp)^\perp$. Sendo $(Y^\perp)^\perp$ um subespaço vetorial e $x \in (Y^\perp)^\perp$, por hipótese, tem-se $z = x - y \in (Y^\perp)^\perp$, logo $z \in Y^{\perp\perp}$. Mas, também, $z \in Y^\perp$ pelo Teorema 3.2. Portanto $z \perp z$, o que implica $z = 0$, e então $x = y$, concluindo-se que $x \in Y$. Como $x \in (Y^\perp)^\perp$ é arbitrário, isto prova que $Y \supset (Y^\perp)^\perp$. \square

No Lema 3.6 a igualdade (3.11) é a principal razão para o uso de subespaços fechados no presente contexto. Como $Z^\perp = Y^{\perp\perp} = Y$, (3.8) pode ser escrita como

$$H = Z \oplus Z^\perp.$$

Disto, segue que $x \mapsto z$ define uma projeção

$$P_Z : H \longrightarrow Z \tag{3.12}$$

de H em Z , cujas propriedades são bastante semelhantes com as propriedades de P consideradas anteriormente.

Lema 3.7. (Conjunto Denso). *Para qualquer subconjunto $M \neq \emptyset$ de um espaço de Hilbert H , o subespaço $[M]$ é denso em H se, e somente se $M^\perp = \{0\}$.*

Demonstração. (a). Seja $x \in M^\perp$ e suponha $V = [M]$ denso em H . Em particular $x \in H = \overline{V}$. Então existe uma sequência (x_n) em V tal que $x_n \longrightarrow x$. Note que para cada $v \in V$, existem escalares $\alpha_1, \dots, \alpha_k$ tais que

$$v = \alpha_1 m_1 + \alpha_2 m_2 + \dots + \alpha_k m_k$$

para certos m_1, \dots, m_k em M , uma vez que $V = [M]$. Note que se $y \in M^\perp$, segue que

$$\begin{aligned} \langle y, v \rangle &= \langle y, \alpha_1 m_1 + \alpha_2 m_2 + \dots + \alpha_k m_k \rangle \\ &= \overline{\alpha_1} \langle y, m_1 \rangle + \dots + \overline{\alpha_k} \langle y, m_k \rangle = 0. \end{aligned}$$

Como y e v são arbitrários segue que $M^\perp \perp V$. Deste modo, tem-se $\langle x_n, x \rangle = 0$. A continuidade do produto interno (ver Lema 3.2) implica que $\langle x_n, x \rangle \longrightarrow \langle x, x \rangle$. Consequentemente $\langle x, x \rangle = \|x\|^2 = 0$, de modo que $x = 0$. Como $x \in M^\perp$ é arbitrário, isto mostra que $M^\perp \subset \{0\}$. Obviamente $\{0\} \subset M^\perp$. Portanto $M^\perp = \{0\}$.

(b) Reciprocamente, suponha que $M^\perp = \{0\}$. Se $x \perp V$, ou seja, $x \in V^\perp$, tem-se $x \perp M$, de modo que $x \in M^\perp$ e $x = 0$. Consequentemente $V^\perp = \{0\}$. Notando que V é um subespaço de H , tem-se $\overline{V} = H$ do Teorema 3.2 com $Y = V^\perp$.

□

3.2 Conjunto Ortonormal

Definição 3.5. *Um conjunto ortogonal M em um espaço com produto interno X é um subconjunto $M \subset X$ cujos elementos são dois a dois ortogonais. Um conjunto ortonormal $M \subset X$ é um conjunto ortogonal em X tal que $\|x\| = 1$ para todo $x \in M$.*

Se um conjunto ortogonal ou ortonormal M é enumerável, pode-se organizá-lo em uma sequência (x_n) e chamá-lo de **sequência ortogonal** ou **sequência ortonormal** respectivamente.

De modo geral, um conjunto indexado, ou família, (x_α) , $\alpha \in I$, é chamada ortogonal se $x_\alpha \perp x_\beta$ para todo $\alpha, \beta \in I$, $\alpha \neq \beta$. A família é chamada ortonormal se esta é ortogonal e para todo x_α tem-se $\|x_\alpha\| = 1$.

Exemplo 3.1. (Relação de Pitágoras). Para quaisquer x, y ortogonais, isto é, $\langle x, y \rangle = 0$, tem-se

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2.$$

Com efeito,

$$\|x + y\|^2 = \langle x + y, x + y \rangle = \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle = \|x\|^2 + \|y\|^2.$$

Lema 3.8. (Independência Linear). Um conjunto ortonormal é linearmente independente.

Demonstração. Seja $\{x_1, x_2, \dots, x_n\}$ um conjunto ortonormal tal que

$$\alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_n x_n = 0$$

Note que para qualquer x_j fixado, tem-se

$$\left\langle \sum_{k=1}^n \alpha_k x_k, x_j \right\rangle = \sum_{k=1}^n \alpha_k \langle x_k, x_j \rangle = \alpha_j \langle x_j, x_j \rangle = \alpha_j = 0$$

e portanto $\{x_1, x_2, \dots, x_n\}$ é linearmente independente. Isto também implica independência linear se o conjunto ortonormal for infinito, pela definição de independência linear. \square

Uma grande vantagem de sequências ortonormais sobre sequências linearmente independentes arbitrárias é a seguinte: se um dado x pode ser representado como uma combinação linear de elementos de uma sequência ortonormal, então a ortonormalidade faz com que a determinação dos coeficientes seja simples. De fato, se (x_n) é uma sequência ortonormal em um espaço com produto interno X e $x \in [x_1, \dots, x_n]$, onde n é fixo, então pela definição de conjunto gerado,

$$x = \sum_{k=1}^n \alpha_k x_k, \tag{3.13}$$

e tomando o produto interno de x por um x_j fixo, obtém-se

$$\langle x, x_j \rangle = \left\langle \sum_{k=1}^n \alpha_k x_k, x_j \right\rangle = \sum_{k=1}^n \alpha_k \langle x_k, x_j \rangle = \alpha_j.$$

Com estes coeficientes, (3.13) torna-se

$$x = \sum_{k=1}^n \langle x, x_k \rangle x_k. \quad (3.14)$$

Isto mostra que a determinação dos coeficientes desconhecidos em (3.13) é simples.

De modo geral, considere um elemento $x \in X$ arbitrário, não necessariamente em $Y_n = [x_1, x_2, \dots, x_n]$, e defina $y \in Y_n$ por

$$y = \sum_{k=1}^n \langle x, x_k \rangle x_k, \quad (3.15)$$

onde n é fixo, como anteriormente, e então defina $z \in X$ por

$$x = y + z, \quad (3.16)$$

isto é, $z = x - y$. Nestas condições, $z \perp y$. Para entender o que está acontecendo, observe o seguinte. Cada $y \in Y_n$ é uma combinação linear

$$y = \sum_{k=1}^n \alpha_k x_k,$$

onde $\alpha_k = \langle y, x_k \rangle$, como visto anteriormente. A alegação é que para a escolha particular de $\alpha_k = \langle x, x_k \rangle$, $k = 1, 2, \dots, n$, tenha-se que obter um y tal que $x - y = z \perp y$.

Para provar este fato, note que, pela ortonormalidade,

$$\|y\|^2 = \left\langle \sum_{k=1}^n \langle x, x_k \rangle x_k, \sum_{k=1}^n \langle x, x_k \rangle x_k \right\rangle = \sum_{k=1}^n \langle x, x_k \rangle \overline{\langle x, x_k \rangle} \langle x_k, x_k \rangle = \sum_{k=1}^n |\langle x, x_k \rangle|^2. \quad (3.17)$$

Usando isto, prova-se agora que $z \perp y$:

$$\begin{aligned} \langle z, y \rangle &= \langle x - y, y \rangle = \langle x, y \rangle - \langle y, y \rangle \\ &= \left\langle x, \sum_{k=1}^n \langle x, x_k \rangle x_k \right\rangle - \|y\|^2 \\ &= \sum_{k=1}^n \langle x, x_k \rangle \overline{\langle x, x_k \rangle} - \|y\|^2 \\ &= \sum_{k=1}^n |\langle x, x_k \rangle|^2 - \sum_{k=1}^n |\langle x, x_k \rangle|^2 \\ &= 0. \end{aligned}$$

Como $z = x - y$ e $z \perp y$, do Teorema de Pitágoras, Exemplo (3.1), tem-se

$$\|x\|^2 = \|y + z\|^2 = \|y\|^2 + \|z\|^2.$$

De (3.17) segue que

$$0 \leq \|z\|^2 = \|x\|^2 - \|y\|^2 = \|x\|^2 - \sum_{k=1}^n |\langle x, x_k \rangle|^2,$$

ou seja,

$$\sum_{k=1}^n |\langle x, x_k \rangle|^2 \leq \|x\|^2, \quad (3.18)$$

para todo $n = 1, 2, \dots$.

Note que esta soma possui termos não negativos, de modo que eles formam uma sequência monótona crescente. Esta sequência é convergente pois é, também, limitada por $\|x\|^2$. Além disso esta é uma sequência das somas parciais de uma série infinita, que assim converge. Logo (3.18) prova o Teorema a seguir.

Teorema 3.3. (Desigualdade de Bessel). *Seja (x_k) uma sequência ortonormal em um espaço com produto interno X . Então para todo $x \in X$*

$$\sum_{k=1}^{\infty} |\langle x, x_k \rangle|^2 \leq \|x\|^2. \quad (3.19)$$

O produto interno $\langle x, x_k \rangle$ em (3.19) é chamado de **coeficientes de Fourier** de x com respeito a sequência ortonormal (x_k) .

Observação 3.1. Note que se X for de dimensão finita, então cada conjunto ortonormal em X deve ser finito, pois este é linearmente independente pelo Lema 3.8. Portanto em (3.19) tem-se uma soma finita. Além disso a desigualdade de Bessel segue também, de maneira natural, quando o conjunto de índices é não enumerável, conforme definição 2.20.

Vê-se que as sequências ortonormais são mais convenientes para trabalhar. O problema agora é como obter uma sequência ortonormal se uma sequência arbitrária linearmente independente é dada. Isto é conseguido por um processo construtivo, o **processo de ortonormalização de Gram-Schmidt**, usado para ortonormalizar uma sequência (v_k) linearmente independente em um espaço com produto interno. A sequência ortonormal (x_k) obtida a partir da sequência (v_k) , deve ter a propriedade de que para cada n ,

$$[x_1, \dots, x_n] = [v_1, \dots, v_n].$$

Isto é, deve-se construir uma sequência ortonormal (x_j) tal que para todo n , estas sequências gerem o mesmo subespaço de X .

Teorema 3.4. *Seja (v_j) uma sequência de vetores linearmente independentes em um espaço vetorial X com produto interno. Então existe uma sequência ortonormal (x_j) em X tal que*

$$[x_1, x_2, \dots, x_j] = [v_1, v_2, \dots, v_j].$$

Demonstração. A construção dos elementos x_j é feita por recorrência. Onde a sequência (x_j) precisa ser ortogonal e seus elementos ter norma 1.

Inicialmente o primeiro passo é construir um conjunto ortogonal $\{y_1, y_2, \dots, y_n\}$ a partir do conjunto $\{v_1, v_2, \dots, v_n\}$. Para isso, defina $y_1 = v_1$. O vetor y_2 é obtido considerando $y_2 = v_2 - \alpha v_1$, onde o escalar α é escolhido de forma que $\langle y_2, v_1 \rangle = 0$, isto é, $\langle v_2 - \alpha v_1, v_1 \rangle = 0$, assim

$$0 = \langle v_2, v_1 \rangle - \alpha \langle v_1, v_1 \rangle \Rightarrow \alpha = \frac{\langle v_2, v_1 \rangle}{\|v_1\|^2},$$

ou seja,

$$y_2 = v_2 - \frac{\langle v_2, v_1 \rangle}{\|v_1\|^2} v_1 = v_2 - \frac{\langle v_2, y_1 \rangle}{\|y_1\|^2} y_1.$$

Para obter o vetor y_3 , considere $y_3 = v_3 - \beta_1 y_1 - \beta_2 y_2$ onde β_1 e β_2 são escolhidos de forma que $\langle y_3, y_1 \rangle = \langle y_3, y_2 \rangle = 0$, assim

$$\langle v_3 - \beta_1 y_1 - \beta_2 y_2, y_1 \rangle = 0 \text{ e } \langle v_3 - \beta_1 y_1 - \beta_2 y_2, y_2 \rangle = 0,$$

o que implica

$$0 = \langle v_3, y_1 \rangle - \beta_1 \langle y_1, y_1 \rangle - \beta_2 \langle y_2, y_1 \rangle \Rightarrow \beta_1 = \frac{\langle v_3, y_1 \rangle}{\|y_1\|^2}$$

e

$$0 = \langle v_3, y_2 \rangle - \beta_1 \langle y_1, y_2 \rangle - \beta_2 \langle y_2, y_2 \rangle \Rightarrow \beta_2 = \frac{\langle v_3, y_2 \rangle}{\|y_2\|^2}.$$

Portanto,

$$y_3 = v_3 - \frac{\langle v_3, y_1 \rangle}{\|y_1\|^2} y_1 - \frac{\langle v_3, y_2 \rangle}{\|y_2\|^2} y_2.$$

De modo geral, na n -ésima etapa deve-se ter

$$y_n = v_n - \sum_{j=1}^{n-1} \frac{\langle v_n, y_j \rangle}{\|y_j\|^2} y_j.$$

Logo, o conjunto $\{y_1, \dots, y_n\}$ é ortogonal, com o vetor y_n não nulo para todo n , pois do contrário v_n seria combinação linear, contradizendo o fato de $\{v_1, \dots, v_n\}$ é linearmente independente. Portanto, definindo $x_j = \frac{y_j}{\|y_j\|}$, com $1 \leq j \leq n$, tem-se que o conjunto $\{x_1, \dots, x_n\}$ é ortonormal e por construção $[x_1, \dots, x_n] = [v_1, \dots, v_n]$. \square

Dada qualquer sequência ortonormal (x_j) em um espaço de Hilbert H , considere a série da forma

$$\sum_{k=1}^{\infty} \alpha_k x_k, \tag{3.20}$$

onde $\alpha_1, \alpha_2, \dots$ são quaisquer escalares.

Teorema 3.5. (Convergência). *Seja (x_k) uma sequência ortonormal em um espaço de Hilbert H . Então:*

(a) A série $\sum_{k=1}^{\infty} \alpha_k x_k$ converge (na norma em H) se, e somente se, a série $\sum_{k=1}^{\infty} |\alpha_k|^2$ é convergente.

(b) Se $\sum_{k=1}^{\infty} \alpha_k x_k$ converge, então os coeficientes α_k são os coeficientes de Fourier $\langle x, x_k \rangle$, onde x denota a soma de $\sum_{k=1}^{\infty} \alpha_k x_k$; portanto, neste caso, $\sum_{k=1}^{\infty} \alpha_k x_k$ pode ser escrita como

$$x = \sum_{k=1}^{\infty} \langle x, x_k \rangle x_k.$$

(c) Para qualquer $x \in H$, a série $\sum_{k=1}^{\infty} \alpha_k x_k$ com $\alpha_k = \langle x, x_k \rangle$ converge na norma em H .

Demonstração. (a) Sejam $s_n = \alpha_1 x_1 + \dots + \alpha_n x_n$ e $\sigma_n = |\alpha_1|^2 + \dots + |\alpha_n|^2$. Como (x_k) é ortonormal, para todo m e $n > m$,

$$\begin{aligned} \|s_n - s_m\|^2 &= \|\alpha_{m+1} x_{m+1} + \dots + \alpha_n x_n\|^2 \\ &= \langle \alpha_{m+1} x_{m+1} + \dots + \alpha_n x_n, \alpha_{m+1} x_{m+1} + \dots + \alpha_n x_n \rangle \\ &= \langle \alpha_{m+1} x_{m+1}, \alpha_{m+1} x_{m+1} \rangle + \dots + \langle \alpha_n x_n, \alpha_n x_n \rangle \\ &= \alpha_{m+1} \overline{\alpha_{m+1}} \langle x_{m+1}, x_{m+1} \rangle + \dots + \alpha_n \overline{\alpha_n} \langle x_n, x_n \rangle \\ &= |\alpha_{m+1}|^2 + \dots + |\alpha_n|^2 \\ &= \sigma_n - \sigma_m. \end{aligned}$$

Portanto, (s_n) é de Cauchy em H se, e somente se, (σ_n) é de Cauchy em \mathbb{R} . Como H e \mathbb{R} são completos, segue o resultado.

(b) Tomando o produto interno de s_n com x_j e usando a ortonormalidade, tem-se

$$\langle s_n, x_j \rangle = \langle \alpha_1 x_1 + \dots + \alpha_n x_n, x_j \rangle = \alpha_j \langle x_j, x_j \rangle = \alpha_j$$

para todo $j \in \{1, 2, \dots, k\}$ e $k \leq n$ com n fixo. Por hipótese, $s_n \rightarrow x$. Uma vez que o produto interno é contínuo, ver Lema 3.2,

$$\alpha_j = \langle s_n, x_j \rangle \rightarrow \langle x, x_j \rangle,$$

para $j \leq k$. Para $n \rightarrow \infty$, pode-se tomar $k(k \leq n)$ tão grande quanto se queira, logo $\alpha_j = \langle x, x_j \rangle$ para todo $j = 1, 2, \dots$.

(c) Da desigualdade de Bessel, Teorema 3.19, tem-se

$$\sum_{k=1}^{\infty} |\langle x, x_k \rangle|^2$$

é convergente. Deste fato e de (a) e (b) segue o resultado. □

Proposição 3.5. *Sejam $(x_i)_{i \in I}$ uma família ortonormal de um espaço pré-hilbertiano H e $\alpha_i = \langle x, x_i \rangle$, $i \in I$ os coeficientes de Fourier do elemento $x \in H$ com respeito a família ortonormal $(x_i)_{i \in I}$.*

- a) *Se a família $(\alpha_i x_i)_{i \in I}$ for somável, então $\langle \sum_{i \in I} \alpha_i x_i, x_j \rangle = \alpha_j$ para todo $j \in I$.*
- b) *Sejam $x = \sum_{i \in I} \alpha_i x_i$ e $y = \sum_{j \in I} \beta_j x_j$, então $\langle x, y \rangle = \sum_{j \in I} \alpha_j \bar{\beta}_j$.*
- c) *Seja $x = \sum_{i \in I} \alpha_i x_i$, então $\|x\|^2 = \sum_{i \in I} |\alpha_i|^2$.*

Demonstração.

- a) Da proposição 3.2, se $(\alpha_i x_i)_{i \in I}$ é uma família somável com soma x , então, para todo $x_j \in H$ com $j \in I$, a família $(\langle \alpha_i x_i, x_j \rangle)_{i \in I}$ é somável com soma $\langle x, x_j \rangle$. Portanto

$$\left\langle \sum_{i \in I} \alpha_i x_i, x_j \right\rangle = \langle x, x_j \rangle = \alpha_j,$$

para todo $j \in I$.

- b) Sejam $x = \sum_{i \in I} \alpha_i x_i$ e $y = \sum_{j \in I} \beta_j x_j$, da linearidade do produto interno tem-se

$$\langle x, y \rangle = \left\langle \sum_{i \in I} \alpha_i x_i, \sum_{j \in I} \beta_j x_j \right\rangle = \sum_{j \in I} \bar{\beta}_j \left\langle \sum_{i \in I} \alpha_i x_i, x_j \right\rangle,$$

do item a) desta proposição, segue-se que

$$\sum_{j \in I} \bar{\beta}_j \left\langle \sum_{i \in I} \alpha_i x_i, x_j \right\rangle = \sum_{j \in I} \alpha_j \bar{\beta}_j.$$

Portanto

$$\langle x, y \rangle = \sum_{j \in I} \alpha_j \bar{\beta}_j.$$

- c) Fazendo $x = y$ em b), tem-se $\|x\|^2 = \langle x, x \rangle = \sum_{i \in I} \alpha_i \bar{\alpha}_i = \sum_{i \in I} |\alpha_i|^2$. □

Teorema 3.6. *Sejam H um espaço pré-hilbertiano, $\{x_i | i \in I\}$, uma família ortonormal formada por elementos de H e $\alpha_i = \langle x, x_i \rangle$, $i \in I$ os coeficientes de Fourier do elemento $x \in H$ com respeito a família ortonormal $(x_i)_{i \in I}$. São equivalentes as seguintes propriedades:*

1. *Para todo $x \in H$ tem-se $\sum_{i \in I} \alpha_i x_i$, isto é, a família $(\alpha_i x_i)_{i \in I}$ é somável e tem por soma x .*
2. *Para quaisquer $x, y \in H$ tem-se*

$$\langle x, y \rangle = \sum_{i \in I} \alpha_i \bar{\beta}_i,$$

denominada identidade de Parseval.

3. Para todo $x \in H$ tem-se

$$\|x\|^2 = \sum_{i \in I} |\alpha_i|^2,$$

denominada igualdade de Bessel.

4. O conjunto das combinações lineares finitas dos x_i , $i \in I$, é denso em H , isto é, dado $x \in H$ e $\epsilon > 0$ existe uma combinação linear finita $\sum_{i \in F} \alpha_i x_i$ tal que

$$\|x - \sum_{i \in F} \alpha_i x_i\| < \epsilon.$$

5. Todo funcional linear contínuo $f : H \rightarrow \mathbb{C}$ que é nulo sobre todos x_i , $i \in I$, é identicamente nulo.

Demonstração.

1) \Rightarrow 2): Foi demonstrado na parte b) da Proposição 3.5.

2) \Rightarrow 3): Basta tomar $x = y$.

3) \Rightarrow 4): Sejam $x \in H$ um elemento qualquer e $\epsilon > 0$ arbitrário. Por hipótese existe um subconjunto finito F_ϵ de índices tal que para todo $F \supset F_\epsilon$, tem-se

$$\left| \|x\|^2 - \sum_{i \in F} |\alpha_i|^2 \right| < \epsilon.$$

Assim, como

$$\begin{aligned} \left\| x - \sum_{x \in F} \alpha_i x_i \right\|^2 &= \left\langle x - \sum_{x \in F} \alpha_i x_i, x - \sum_{x \in F} \alpha_i x_i \right\rangle \\ &= \langle x, x \rangle - \left\langle x, \sum_{x \in F} \alpha_i x_i \right\rangle - \left\langle \sum_{x \in F} \alpha_i x_i, x \right\rangle + \left\langle \sum_{x \in F} \alpha_i x_i, \sum_{x \in F} \alpha_i x_i \right\rangle \\ &= \|x\|^2 - \sum_{i \in F} |\alpha_i|^2 - \sum_{i \in F} |\alpha_i|^2 + \sum_{i \in F} |\alpha_i|^2 \\ &= \|x\|^2 - \sum_{i \in F} |\alpha_i|^2, \end{aligned}$$

segue-se que

$$\left\| x - \sum_{x \in F} \alpha_i x_i \right\|^2 = \|x\|^2 - \sum_{i \in F} |\alpha_i|^2 \leq \left| \|x\|^2 - \sum_{i \in F} |\alpha_i|^2 \right| < \epsilon.$$

4) \Rightarrow 5): Seja X o conjunto das combinações lineares finitas dos x_i , $i \in I$. Tem-se que X é um subespaço vetorial de H . Tem-se também que se o funcional linear f é nulo sobre x_i , isto é, $f(x_i) = 0$ para todo x_i , $i \in I$, então f é nulo para todos os elementos de X . Para cada $p \in H$, como X é denso em H , existe uma sequência $(y_k)_{k \in \mathbb{N}}$ de elementos de X tal que $\lim_{k \rightarrow \infty} y_k = p$. Assim, da continuidade da f , segue que $\lim_{k \rightarrow \infty} f(y_k) = f(p)$, mas $f(y_k) = 0$ para todo k , portanto $f(p) = 0$.

5) \Rightarrow 1): Seja $x \in H$, suponha que $x \neq \sum_{i \in I} \alpha_i x_i$, da desigualdade de Bessel, segue

$$\text{que } \|x\|^2 - \sum_{i \in I} |\alpha_i|^2 > 0.$$

Considere o funcional linear contínuo $f : H \rightarrow \mathbb{C}$ definido por

$$f(y) = \langle y, x \rangle - \sum_{i \in I} \beta_i \overline{\alpha_i},$$

onde $\beta_i = \langle y, x_i \rangle$ e $\alpha_i = \langle x, x_i \rangle$.

Para cada $k \in I$, tem-se

$$f(x_k) = \langle x_k, x \rangle - \sum_{i \in I} \langle x_k, x_i \rangle \overline{\alpha_i} = \overline{\langle x, x_k \rangle} - \overline{\alpha_k} = \overline{\alpha_k} - \overline{\alpha_k} = 0,$$

mas

$$f(x) = \langle x, x \rangle - \sum_{i \in I} \alpha_i \overline{\alpha_i} = \|x\|^2 - \sum_{i \in I} |\alpha_i|^2 > 0,$$

contradizendo 4). □

Definição 3.6. *Se uma família ortonormal $\{x_i | i \in I\}$ satisfaz as propriedades equivalentes de 1. a 5., diz-se que ela é uma família ortonormal completa.*

3.3 Representação de Funcionais Lineares em Espaços de Hilbert

Teorema 3.7. Teorema de Riesz (Funcionais em Espaços de Hilbert). *Todo funcional linear limitado f , em um espaço de Hilbert H pode ser representado em termos de produto interno, a saber,*

$$f(x) = \langle x, z \rangle, \tag{3.21}$$

onde z depende de f e é unicamente determinado por f e tem norma

$$\|z\| = \|f\|. \tag{3.22}$$

Demonstração. Deve-se provar:

- (a) f tem a representação $f(x) = \langle x, z \rangle$, para todo $x \in H$,
- (b) z é unicamente determinado por f ,
- (c) $\|z\| = \|f\|$.

Com efeito:

- (a) Note que se $f = 0$, então basta tomar $z = 0$. Logo $f(x) = \langle x, z \rangle = 0$ para todo $x \in H$. Suponha $f \neq 0$ e seja $\mathcal{N}(f)$ o núcleo da f . Sejam $x, y \in \mathcal{N}(f)$ e α um escalar, como $f(x + \alpha y) = f(x) + \alpha f(y) = 0 + \alpha 0 = 0$, segue que $\mathcal{N}(f)$ é um subespaço vetorial de H . Note também que $\mathcal{N}(f)$ é fechado, pois para todo $x \in \overline{\mathcal{N}(f)}$ existe uma sequência de pontos $x_n \in \mathcal{N}(f)$ tal que $x_n \rightarrow x$. Sendo f um funcional linear limitado, segue que f é contínua. Então $f(x_n) \rightarrow f(x)$. Logo $f(x) = 0$ uma vez que $f(x_n) = 0$. Portanto $x \in \mathcal{N}(f)$. Como x é arbitrário, $\mathcal{N}(f)$ é fechado. Pelo Teorema 3.2,

$$H = \mathcal{N}(f) \oplus \mathcal{N}(f)^\perp,$$

além disso, sendo $f \neq 0$, existe $w \in H$ tal que $f(w) \neq 0$, isto é, $w \notin \mathcal{N}(f)$, então $\mathcal{N}(f) \neq H$. Deste modo $\mathcal{N}(f)^\perp \neq \{0\}$. Consequentemente $\mathcal{N}(f)^\perp$ contém $z_0 \neq 0$. Para cada $x \in H$ arbitrário considere

$$v = f(x)z_0 - f(z_0)x.$$

Aplicando f , tem-se

$$f(v) = f(f(x)z_0 - f(z_0)x) = f(x)f(z_0) - f(z_0)f(x) = 0.$$

Isto mostra que $v \in \mathcal{N}(f)$. Como $z_0 \in \mathcal{N}(f)^\perp \Rightarrow z_0 \perp \mathcal{N}(f)$, tem-se

$$\begin{aligned} 0 = \langle v, z_0 \rangle &= \langle f(x)z_0 - f(z_0)x, z_0 \rangle \\ &= f(x)\langle z_0, z_0 \rangle - f(z_0)\langle x, z_0 \rangle. \end{aligned}$$

Uma vez que $\langle z_0, z_0 \rangle = \|z_0\|^2 \neq 0$, obtém-se f da forma,

$$f(x) = \frac{f(z_0)}{\langle z_0, z_0 \rangle} \langle x, z_0 \rangle.$$

Este pode ser escrito como

$$f(x) = \frac{f(z_0)}{\langle z_0, z_0 \rangle} \langle x, z_0 \rangle = \langle x, \frac{\overline{f(z_0)}}{\langle z_0, z_0 \rangle} z_0 \rangle.$$

Portando basta tomar

$$z = \frac{\overline{f(z_0)}}{\langle z_0, z_0 \rangle} z_0.$$

Como x é arbitrário, (3.21) está provado.

- (b) Para provar que z é único. Suponha que para todo $x \in H$,

$$f(x) = \langle x, z_1 \rangle = \langle x, z_2 \rangle.$$

Então $\langle x, z_1 \rangle - \langle x, z_2 \rangle = 0$, isto implica que $\langle x, z_1 - z_2 \rangle = 0$ para todo x . Em particular, escolhendo $x = z_1 - z_2$, tem-se

$$\langle x, z_1 - z_2 \rangle = \langle z_1 - z_2, z_1 - z_2 \rangle = \|z_1 - z_2\|^2 = 0.$$

Portanto $z_1 - z_2 = 0$, o que mostra $z_1 = z_2$.

(c) Se $f = 0$, então $z = 0$ e (3.22) está provado. Suponha $f \neq 0$. Então $z \neq 0$. Para $x = z$ (3.21), tem-se

$$\|z\|^2 = \langle z, z \rangle = f(z) \leq \|f\| \|z\|,$$

uma vez que f é um funcional linear limitado. Dividindo por $\|z\| \neq 0$, obtém-se $\|z\| \leq \|f\|$. Resta mostrar que $\|f\| = \|z\|$. De (3.21) e aplicando a Desigualdade de Schwarz, tem-se

$$|f(x)| = |\langle x, z \rangle| \leq \|x\| \|z\|.$$

Considerando o supremo para todo $x \in H$ com $\|x\| = 1$. A desigualdade acima fica

$$\|f\| = \sup_{\|x\|=1} |f(x)| = \sup_{\|x\|=1} |\langle x, z \rangle| \leq \|z\|.$$

Logo $\|f\| \leq \|z\|$ e portanto $\|f\| = \|z\|$.

□

A ideia da prova da unicidade do item (b) do teorema acima parte do seguinte lema.

Lema 3.9. (Igualdade). *Se $\langle v_1, w \rangle = \langle v_2, w \rangle$ para todo w em um espaço com produto interno X , então $v_1 = v_2$. Em particular, $\langle v_1, w \rangle = 0$ para todo $w \in X$ implica $v_1 = 0$.*

Demonstração. Suponha, para todo w ,

$$\langle v_1 - v_2, w \rangle = \langle v_1, w \rangle - \langle v_2, w \rangle = 0.$$

Para $w = v_1 - v_2$ tem-se $\|v_1 - v_2\|^2 = 0$. Logo $v_1 - v_2 = 0$, isto implica $v_1 = v_2$. Em particular, se $\langle v_1, w \rangle = 0$ para todo w , tomando $w = v_1$, tem-se $\|v_1\|^2 = 0$, isto implica que $v_1 = 0$. □

3.3.1 Forma Sesquilinear

Definição 3.7. *Sejam X e Y espaços vetoriais sobre o mesmo corpo \mathbb{K} (\mathbb{R} ou \mathbb{C}). Então uma **forma sesquilinear** (ou **funcional sesquilinear**) h , em $X \times Y$, é uma aplicação*

$$h : X \times Y \longrightarrow \mathbb{K}$$

tal que para todos $x, x_1, x_2 \in X$ e $y, y_1, y_2 \in Y$ e todos escalares α, β

- (a) $h(x_1 + x_2, y) = h(x_1, y) + h(x_2, y)$
- (b) $h(x, y_1 + y_2) = h(x, y_1) + h(x, y_2)$
- (c) $h(\alpha x, y) = \alpha h(x, y)$
- (d) $h(x, \beta y) = \bar{\beta} h(x, y)$.

Note que nos itens (a) e (c) da definição 3.7 h é linear em relação a primeira componente, já nos itens (b) e (d) h é linear conjugada em relação a segunda componente. Se X e Y são reais, isto é, $K = \mathbb{R}$, então (d) é simplesmente

$$h(x, \beta y) = \beta h(x, y),$$

e h é chamada de **bilinear** uma vez que esta é linear em relação as duas componentes.

Se X e Y são espaços normados e se existir um número real c tal que para todos x, y

$$|h(x, y)| \leq c \|x\| \|y\|, \tag{3.23}$$

então h é dito **limitado**, e o número

$$\|h\| = \sup_{\substack{x \in X - \{0\} \\ y \in Y - \{0\}}} \frac{|h(x, y)|}{\|x\| \|y\|} = \sup_{\substack{\|x\|=1 \\ \|y\|=1}} |h(x, y)| \tag{3.24}$$

é chamado **norma** de h .

É de grande importância que a partir do Teorema 3.7 possa obter uma representação geral de formas sesquilineares em espaços de Hilbert como segue.

Teorema 3.8. (Representação de Riesz). *Sejam H_1, H_2 espaços de Hilbert e*

$$h : H_1 \times H_2 \longrightarrow \mathbb{K}$$

uma forma sesquilinear limitada. Então h tem uma representação

$$h(x, y) = \langle Sx, y \rangle, \tag{3.25}$$

onde $S : H_1 \longrightarrow H_2$ é um operador linear limitado. S é unicamente determinado por h e tem norma

$$\|S\| = \|h\|. \tag{3.26}$$

Demonstração. Considere $\overline{h(x, y)}$. Note que este é linear em y , pois se $w, t \in H_2$ e α um escalar, tem-se

$$\overline{h(x, \alpha w + t)} \stackrel{(b)}{=} \overline{h(x, \alpha w)} + \overline{h(x, t)} = \alpha \overline{h(x, w)} + \overline{h(x, t)}.$$

Para fazer uso do Teorema 3.7 deve-se manter x fixo. Então o teorema produz uma representação em que y é variável, por exemplo,

$$\overline{h(x, y)} = \langle y, z \rangle.$$

Portanto,

$$h(x, y) = \langle z, y \rangle, \tag{3.27}$$

onde $z \in H_2$ é único mas, depende do $x \in H_1$ fixo. Deste modo (3.27) na variável x define um operador

$$S : H_1 \rightarrow H_2$$

dados por $z = Sx$. Substituindo $z = Sx$ em (3.27), tem-se

$$h(x, y) = \langle Sx, y \rangle.$$

S é linear. De fato, seu domínio é o espaço vetorial H_1 , de (3.25) e da sesquilinearidade de h , obtém-se

$$\begin{aligned} \langle S(\alpha x_1 + \beta x_2), y \rangle &= h(\alpha x_1 + \beta x_2, y) \\ &= \alpha h(x_1, y) + \beta h(x_2, y) \\ &= \alpha \langle Sx_1, y \rangle + \beta \langle Sx_2, y \rangle \\ &= \langle \alpha Sx_1 + \beta Sx_2, y \rangle \end{aligned}$$

para todo $y \in H_2$. Então, do Lema (3.9),

$$S(\alpha x_1 + \beta x_2) = \alpha Sx_1 + \beta Sx_2.$$

S é limitado e $\|S\| = \|h\|$. De fato, deixando de lado o caso trivial, $S = 0$, tem-se de (3.24) e (3.25) que

$$\|h\| = \sup_{\substack{x \in H_1 - \{0\} \\ y \in H_2 - \{0\}}} \frac{|\langle Sx, y \rangle|}{\|x\| \|y\|} \geq \sup_{\substack{x \in H_1 - \{0\} \\ Sx \in H_2 - \{0\}}} \frac{|\langle Sx, Sx \rangle|}{\|x\| \|Sx\|} = \sup_{x \in H_1 - \{0\}} \frac{\|Sx\|}{\|x\|} = \|S\|.$$

Portanto $\|h\| \geq \|S\|$. Para mostrar $\|h\| = \|S\|$, note que $\|h\| \leq \|S\|$ segue da Desigualdade de Schwarz:

$$\|h\| = \sup_{\substack{x \in H_1 - \{0\} \\ y \in H_2 - \{0\}}} \frac{|\langle Sx, y \rangle|}{\|x\| \|y\|} \leq \sup_{x \in H_1 - \{0\}} \frac{\|Sx\| \|y\|}{\|x\| \|y\|} = \|S\|.$$

S é único. Com efeito, suponha que exista um operador linear $T : H_1 \rightarrow H_2$ tal que para todo $x \in H_1$ e $y \in H_2$ tem-se

$$h(x, y) = \langle Sx, y \rangle = \langle Tx, y \rangle,$$

o que implica $Sx = Tx$, pelo Lema 3.9, para todo $x \in H_1$. Portanto $S = T$. \square

4 Teoria dos Operadores

Este capítulo inicia-se com o estudo da teoria dos Operadores Lineares com o objetivo de utilizar seus resultados na construção da Teoria Espectral a ser desenvolvida no capítulo seguinte. As principais referências utilizadas são [3] e [7].

No estudo de funções definidas em espaços vetoriais e, em particular, em espaços normados, uma aplicação com certas propriedades é chamada de **operador**. O principal interesse são os operadores que "preservam" as duas operações algébricas de um espaço vetorial, como na definição que segue.

Definição 4.1. *Sejam X e Y espaços vetoriais sobre um mesmo corpo. Um **operador linear** $T : \mathcal{D}(T) \subset X \rightarrow Y$ é uma aplicação tal que para todo $x, y \in \mathcal{D}(T)$ e escalar α , tem-se*

$$\begin{aligned} T(x + y) &= Tx + Ty, \\ T(\alpha x) &= \alpha Tx. \end{aligned} \tag{4.1}$$

Abaixo seguem dois exemplos de operadores lineares que são facilmente verificados a partir das propriedades vistas no Cálculo Diferencial e Integral.

Exemplo 4.1. Seja X o espaço vetorial de todos os polinômios definidos em $[a, b] \subset \mathbb{R}$. O operador $T : X \rightarrow X$ definido por

$$Tp(t) = p'(t), \quad \forall p(t) \in X,$$

é linear. T representa a primeira derivada de p na variável t .

Exemplo 4.2. $T : \mathcal{C}[a, b] \rightarrow \mathcal{C}[a, b]$, definido por

$$Tx(t) = \int_a^t x(s) ds, \quad \forall t \in [a, b]$$

é um operador linear.

A notação Tx em vez de $T(x)$ será usada em todo o texto, salvo em alguns momentos onde a notação tradicional é de melhor clareza. Tal notação é padrão dentro do estudo da análise funcional.

Definição 4.2. *Sejam X e Y dois espaços vetoriais sobre o corpo \mathbb{K} e $T : \mathcal{D}(T) \subset X \rightarrow Y$ um operador linear. O **núcleo** de T é o conjunto de todos $x \in \mathcal{D}(T)$ tais que $Tx = 0$.*

Observação 4.1. Vale observar que o núcleo de um operador T , denotado por $\ker(T) = \{x \in \mathcal{D}(T); Tx = 0\}$ é um subespaço vetorial.

4.1 Operadores Lineares Limitados

Definição 4.3. *Sejam X e Y espaços normados e $T : \mathcal{D}(T) \rightarrow Y$ um operador linear, onde $\mathcal{D}(T) \subset X$. T é dito ser **limitado** se existir um número real c tal que para todo $x \in \mathcal{D}(T)$,*

$$\|Tx\|_Y \leq c\|x\|_X. \quad (4.2)$$

Observação 4.2. Note que em (4.2) a norma no primeiro membro da desigualdade é a do espaço Y enquanto que a norma no segundo membro é a de X . Por simplicidade, ambas as normas serão denotadas pelo mesmo símbolo $\|\cdot\|$. Nos casos em que houver necessidade de distinção as normas serão indexadas de acordo com os espaços nos quais elas estão definidas, por exemplo, $(\|x\|_X, \|Tx\|_Y)$.

Em (4.2), para $x \neq 0$, tem-se

$$\frac{\|Tx\|}{\|x\|} \leq c,$$

de modo que o conjunto $\left\{ \frac{\|Tx\|}{\|x\|}, x \in \mathcal{D}(T) - \{0\} \right\}$ é limitado superiormente. Assim, define-se a **norma** $\|T\|$, do operador T , por

$$\|T\| = \sup_{\substack{x \in \mathcal{D}(T) \\ x \neq 0}} \frac{\|Tx\|}{\|x\|}. \quad (4.3)$$

Se $\mathcal{D}(T) = \{0\}$, define-se $\|T\| = 0$.

Note que (4.2), com $c = \|T\|$ se torna

$$\|Tx\| \leq \|T\|\|x\|, \forall x \in \mathcal{D}(T). \quad (4.4)$$

Lema 4.1. *Seja T um operador linear limitado. Então:*

a) *Uma fórmula alternativa para a norma de T é*

$$\|T\| = \sup_{\substack{x \in \mathcal{D}(T) \\ \|x\|=1}} \|Tx\|. \quad (4.5)$$

b) $\|T\| = \sup_{\substack{x \in \mathcal{D}(T) \\ x \neq 0}} \frac{\|Tx\|}{\|x\|}$ *define uma norma.*

Demonstração. a) Como T é linear, para todo $x \neq 0$ em $\mathcal{D}(T)$, tem-se

$$\|T\| = \sup_{\substack{x \in \mathcal{D}(T) \\ x \neq 0}} \frac{\|Tx\|}{\|x\|} = \sup_{\substack{x \in \mathcal{D}(T) \\ x \neq 0}} \frac{1}{\|x\|} \|Tx\| = \sup_{\substack{x \in \mathcal{D}(T) \\ x \neq 0}} \left\| T \left(\frac{x}{\|x\|} \right) \right\| = \sup_{\substack{y \in \mathcal{D}(T) \\ \|y\|=1}} \|Ty\|,$$

onde $y = \frac{x}{\|x\|}$.

b) (N1) segue de modo imediato da definição.

(N2) Se $T = 0$ então $\|T\| = 0$. Por outro lado, se $T \neq 0$, existe $x \neq 0$ tal que $Tx \neq 0$, ou seja, $\|T\| \geq \|Tx\| > 0$.

(N3) Sejam $x \in \mathcal{D}(T)$ e α um escalar qualquer, então

$$\|\alpha T\| = \sup_{\|x\|=1} \|\alpha Tx\| = \sup_{\|x\|=1} |\alpha| \|Tx\| = |\alpha| \sup_{\|x\|=1} \|Tx\| = |\alpha| \|T\|.$$

(N4) Sejam $T_1, T_2 : \mathcal{D}(T) \subset X \rightarrow Y$ operadores lineares, então

$$\begin{aligned} \|T_1 + T_2\| &= \sup_{\substack{\|x\|=1 \\ x \in \mathcal{D}(T)}} \|(T_1 + T_2)x\| = \sup_{\substack{\|x\|=1 \\ x \in \mathcal{D}(T)}} \|T_1x + T_2x\| \\ &\leq \sup_{\substack{\|x\|=1 \\ x \in \mathcal{D}(T)}} \|T_1x\| + \sup_{\substack{\|x\|=1 \\ x \in \mathcal{D}(T)}} \|T_2x\| \\ &= \|T_1\| + \|T_2\|. \end{aligned}$$

□

Observação 4.3. O conjunto de todos os operadores lineares contínuos ou limitados $T : X \rightarrow Y$ será denotado por $\mathcal{B}(X, Y)$.

Exemplo 4.3. Seja $k : \mathcal{C}([0, 1], \mathbb{C}) \rightarrow \mathcal{C}([0, 1], \mathbb{C})$ um operador definido por

$$(kf)(t) = \int_0^1 K(t, s)f(s)ds, \quad \forall f \in \mathcal{C}([0, 1], \mathbb{C}),$$

onde K é uma função contínua, chamada núcleo de k , definida em $[0, 1] \times [0, 1]$ no plano- ts .

Nas condições definidas acima k é um operador linear. De fato, sejam α um escalar qualquer e $f, g \in \mathcal{C}([0, 1], \mathbb{C})$, então

$$\begin{aligned} k(\alpha f + g)(t) &= \int_0^1 K(t, s)(\alpha f + g)(s)ds = \int_0^1 [K(t, s)\alpha f(s) + K(t, s)g(s)]ds \\ &= \int_0^1 K(t, s)\alpha f(s)ds + \int_0^1 K(t, s)g(s)ds = \\ &= \alpha(kf)(t) + (kg)(t). \end{aligned}$$

k é limitado. De fato, como K é contínua em $[0, 1] \times [0, 1]$ segue que K é limitada, isto é, existe um k_0 real tal que $|K(t, s)| \leq k_0$ para todo $(t, s) \in [1, 0] \times [1, 0]$. Além disso,

$$|f(t)| \leq \max_{t \in [0, 1]} |f(t)| = \|f\|$$

em $\mathcal{C}([0, 1], \mathbb{C})$.

Daí,

$$\begin{aligned} \|kf\| &= \max_{t \in [0,1]} \left| \int_0^1 K(t, s) f(s) ds \right| \\ &\leq \max_{t \in [0,1]} \int_0^1 |K(t, s)| |f(s)| ds \\ &\leq k_0 \|f\|, \quad \forall f \in \mathcal{C}([0, 1], \mathbb{C}). \end{aligned}$$

Portanto $\|kf\| \leq k_0 \|f\|$. Isto é, vale (4.2) com $c = k_0$.

Teorema 4.1. *Seja $T : \mathcal{D}(T) \rightarrow Y$ um operador linear, onde $\mathcal{D}(T) \subset X$ e X, Y são espaços normados. Então,*

- a) *T é contínuo se, e somente se, T é limitado.*
- b) *Se T é contínuo em um único ponto, então T é contínuo.*

Demonstração. a) Se $T = 0$ o resultado segue de imediato. Suponha que $T \neq 0$. Então $\|T\| \neq 0$. Suponha, também, que T seja um operador linear limitado e x_0 um elemento qualquer de $\mathcal{D}(T)$.

Seja $\epsilon > 0$ arbitrário, uma vez que T é linear, para todo $x \in \mathcal{D}(T)$ tal que $\|x - x_0\| < \delta$ onde $\delta = \frac{\epsilon}{\|T\|}$, tem-se

$$\|Tx - Tx_0\| = \|T(x - x_0)\| \leq \|T\| \|x - x_0\| < \|T\| \delta = \epsilon.$$

Portanto T é contínuo.

Reciprocamente, suponha que T seja contínuo em $x_0 \in \mathcal{D}(T)$. Então, dado $\epsilon > 0$ arbitrário, existe $\delta > 0$ tal que para todo $x \in \mathcal{D}(T)$ com $\|x - x_0\| < \delta$ tem-se $\|Tx - Tx_0\| < \epsilon$.

Para mostrar que T é limitado considere $y \in \mathcal{D}(T)$, $y \neq 0$, arbitrário e tome $x = x_0 + \frac{\delta}{\|y\|} y$. Assim $x - x_0 = \frac{\delta}{\|y\|} y$. Da continuidade e linearidade de T obtém-se

$$\|Tx - Tx_0\| = \|T(x - x_0)\| = \left\| T \left(\frac{\delta}{\|y\|} y \right) \right\| = \frac{\delta}{\|y\|} \|Ty\| < \epsilon.$$

Assim $\|Ty\| < \frac{\epsilon}{\delta} \|y\|$ concluindo que T é limitado.

- b) Se T é contínuo em um ponto do seu domínio, então pela prova do item a) T é limitado, que por sua vez implica na continuidade de T também por a).

□

4.2 Operadores Compactos

Proposição 4.1. *Sejam X e Y espaços normados e $T : X \rightarrow Y$ uma aplicação linear. São equivalentes as seguintes propriedades:*

- a) T leva a bola unitária B de X em um conjunto relativamente compacto de Y .
- b) T leva os conjuntos limitados de X em conjuntos relativamente compactos de Y .
- c) Toda sequência limitada de pontos (x_n) de X contém uma subsequência (x_{r_n}) tal que a sequência $(T(x_{r_n}))$ é convergente em Y .

Demonstração. a) \Rightarrow b). Seja $L \subset X$ um subconjunto limitado qualquer. Então existe $a > 0$ tal que L está contido na bola $B(0, a)$. Mas $B(0, a) = aB(0, 1)$. Assim, $L \subset B(0, a) = aB(0, 1) \Rightarrow T(L) \subset T(aB(0, 1))$. Como T é linear tem-se

$$T(L) \subset aT(B(0, 1)).$$

Por hipótese $T(B(0, 1))$ é um conjunto relativamente compacto de F . Isto implica $aT(B(0, 1))$ um conjunto relativamente compacto. Portanto, $T(L)$ é relativamente compacto pelo Teorema 2.4.

b) \Rightarrow c). Seja (x_n) uma sequência limitada de pontos de X . Sendo $L = \{x_n \mid n \in \mathbb{N}\}$ um subconjunto limitado de X , então $T(L)$ é, por b), um subconjunto relativamente compacto de Y e do Corolário 2.1 segue que a sequência de pontos $(T(x_n))$ possui uma subsequência convergente.

c) \Rightarrow a). Como a bola unitária B é limitada, toda sequência $(x_n) \subset B$ é limitada. Por hipótese esta sequência contém uma subsequência (x_{r_n}) tal que $T(x_{r_n})$ é convergente em Y . Assim $T(B)$ é relativamente compacto pelo Corolário 2.1. \square

Definição 4.4. *Diz-se que uma aplicação linear $T : X \rightarrow Y$ é **compacta** ou **completamente contínua** se ela satisfaz as condições equivalentes da Proposição 4.1.*

Exemplo 4.4. Sejam $X = \mathcal{C}([a, b], \mathbb{C})$, $Y = \mathcal{C}([c, d], \mathbb{C})$ e $K : (t, s) \in [c, d] \times [a, b] \mapsto K(t, s) \in \mathbb{C}$ uma função contínua. Para toda $f \in \mathcal{C}([a, b], \mathbb{C})$ define-se $y = k(f) \in \mathcal{C}([c, d], \mathbb{C})$ por

$$y(t) = (kf)(t) = \int_a^b K(t, s)f(s)ds,$$

onde $t \in [c, d]$. Então k é um operador linear compacto. De fato, para mostrar que k é compacto basta provar que $k(B)$ é um subconjunto relativamente compacto de $\mathcal{C}([c, d], \mathbb{C})$ e para isto, pelo Teorema de Ascoli-Arzelá, é suficiente demonstrar que

1. $k(B)$ é equicontínuo;
2. para todo $t_0 \in [c, d]$ o conjunto

$$k(B)(t_0) = \{(kf)(t_0) \in \mathbb{C} \mid f \in B\}$$

é limitado em \mathbb{C} .

Demonstração de 1.: Sendo K contínua em $[c, d] \times [a, b]$ segue que K é também uniformemente contínua, assim dado $\epsilon > 0$ existe $\delta > 0$ tal que para todo $s_1, s_2 \in [a, b]$, em particular, para $s_1 = s_2 = s \in [a, b]$ e $t_1, t_2 \in [c, d]$ com $\|(t_1, s) - (t_2, s)\| = \|(t_1 - t_2, s - s)\| = |t_1 - t_2| < \delta$ tem-se $|K(t_1, s) - K(t_2, s)| < \epsilon$. Para toda $f \in \mathcal{C}([a, b], \mathbb{C})$, tem-se

$$\begin{aligned} |y(t_1) - y(t_2)| &= |(kf)(t_1) - (kf)(t_2)| = \left| \int_a^b K(t_1, s)f(s)ds - \int_a^b K(t_2, s)f(s)ds \right| \\ &= \left| \int_a^b (K(t_1, s) - K(t_2, s))f(s)ds \right| \\ &\leq \int_a^b |(K(t_1, s) - K(t_2, s))f(s)|ds. \end{aligned}$$

Logo, para $|t - t_0| < \delta$ e $f \in B$, tem-se $|f(s)| \leq \|f\| \leq 1$ e

$$\begin{aligned} |(kf)(t) - (kf)(t_0)| &\leq \int_a^b |(K(t, s) - K(t_0, s))f(s)|ds \\ &\leq \epsilon \int_a^b |f(s)|ds \\ &\leq \epsilon \int_a^b \|f\|ds \\ &< \epsilon \int_a^b ds = \epsilon(b - a). \end{aligned}$$

Portanto $k(B)$ é equicontínuo.

Demonstração de 2.: Seja $t_0 \in [c, d]$, para toda $f \in B$ tem-se

$$|(kf)(t_0)| = \left| \int_a^b K(t_0, s)f(s)ds \right| \leq \int_a^b |K(t_0, s)| |f(s)|ds.$$

Uma vez que K é limitada (contínua em um compacto) e $|f(s)| \leq \|f\| < 1$, segue que $k(B)(t_0)$ é limitado em \mathbb{C} e portanto k é um operador linear compacto.

Exemplo 4.5. Dado $1 \leq p < \infty$, sejam $X = \mathcal{C}_{L_p}([a, b], \mathbb{C})$ e $Y = \mathcal{C}([a, b], \mathbb{C})$. Considere o operador k definido como no Exemplo 4.4. Dado, ainda,

$$B = \{f \in \mathcal{C}_{L_p}([a, b], \mathbb{C}) \mid \|f\|_p \leq 1\},$$

então k é um operador linear compacto. De fato, basta proceder como no exemplo anterior utilizando o Teorema de Ascoli-Arzelá. Com as mesmas notações do Exemplo

4.4, tem-se para todo $f \in B$ e $t, t_0 \in [c, d]$ com $|t - t_0| < \delta$, da Desigualdade de Hölder,

$$\begin{aligned} |(kf)(t) - (kf)(t_0)| &\leq \int_a^b |(K(t, s) - K(t_0, s))| |f(s)| ds \\ &\leq \left[\int_a^b |K(t, s) - K(t_0, s)|^{p'} ds \right]^{\frac{1}{p'}} \cdot \left[\int_a^b |f(s)|^p ds \right]^{\frac{1}{p}} \\ &= \left[\int_a^b |K(t, s) - K(t_0, s)|^{p'} ds \right]^{\frac{1}{p'}} \cdot \|f\|_p \\ &\leq \left[\epsilon^{p'} \int_a^b ds \right]^{\frac{1}{p'}} \cdot 1 \\ &= \epsilon(b - a)^{\frac{1}{p'}}. \end{aligned}$$

Portanto $k(B)$ é equicontínuo. Para todo $f \in B$, tem-se também da Desigualdade de Hölder

$$|(kf)(t_0)| \leq \int_a^b |K(t_0, s)| |f(s)| ds \leq \left[\int_a^b |K(t_0, s)|^{p'} ds \right]^{\frac{1}{p'}}.$$

Sendo K limitado, segue o resultado.

4.3 Operador Hilbert-Adjunto

Os resultados da seção anterior permitem, agora, introduzir o operador Hilbert-Adjunto em um espaço de Hilbert. Tal operador ajudará a definir outros três operadores, a saber, *operador autoadjunto*, *operador unitário* e *operador normal*.

Definição 4.5. *Seja $T : H_1 \rightarrow H_2$ um operador linear limitado, onde H_1 e H_2 são espaços de Hilbert. Então o **operador Hilbert-Adjunto** T^* de T é o operador*

$$T^* : H_2 \rightarrow H_1$$

tal que para todos $x \in H_1$ e $y \in H_2$,

$$\langle Tx, y \rangle = \langle x, T^*y \rangle. \tag{4.6}$$

Deve-se primeiro mostrar que esta definição faz sentido, isto é, deve-se provar que para um dado T tem-se a existência de T^* .

Teorema 4.2. (Existência). *O operador Hilbert-Adjunto T^* de T , na definição 4.5, existe, é único e é um operador linear limitado com norma*

$$\|T^*\| = \|T\|. \tag{4.7}$$

Demonstração. Note que

$$h(y, x) = \langle y, Tx \rangle \tag{4.8}$$

define uma forma sesquilinear em $H_2 \times H_1$ pois o produto interno é sesquilinear e T é linear. De fato,

$$\begin{aligned} h(\alpha y_1 + \beta y_2, x) &= \langle \alpha y_1 + \beta y_2, Tx \rangle = \alpha \langle y_1, Tx \rangle + \beta \langle y_2, Tx \rangle \\ &= \alpha h(y_1, x) + \beta h(y_2, x) \end{aligned}$$

e

$$\begin{aligned} h(y, \alpha x_1 + \beta x_2) &= \langle y, T(\alpha x_1 + \beta x_2) \rangle \\ &= \langle y, \alpha T x_1 + \beta T x_2 \rangle \\ &= \bar{\alpha} \langle y, T x_1 \rangle + \bar{\beta} \langle y, T x_2 \rangle \\ &= \bar{\alpha} h(y, x_1) + \bar{\beta} h(y, x_2). \end{aligned}$$

A aplicação h é limitada. De fato, pela Desigualdade de Schwarz,

$$|h(x, y)| = |\langle y, Tx \rangle| \leq \|y\| \|Tx\| \leq \|T\| \|x\| \|y\|.$$

Isto também implica $\|h\| \leq \|T\|$. Além disso, tem-se $\|h\| \geq \|T\|$ pois

$$\|h\| = \sup_{\substack{x \neq 0 \\ y \neq 0}} \frac{|\langle y, Tx \rangle|}{\|y\| \|x\|} \geq \sup_{\substack{x \neq 0 \\ Tx \neq 0}} \frac{|\langle Tx, Tx \rangle|}{\|Tx\| \|x\|} = \|T\|.$$

Portanto,

$$\|h\| = \|T\|. \quad (4.9)$$

Do Teorema 3.8, h tem uma representação de Riesz, isto é,

$$h(y, x) = \langle T^* y, x \rangle, \quad (4.10)$$

onde T^* é um operador linear limitado unicamente determinado com norma

$$\|T^*\| = \|h\| = \|T\|.$$

Isto prova (4.7). Tem-se também $\langle y, Tx \rangle = \langle T^* y, x \rangle$ de (4.8) e (4.10), de modo que (4.6) é satisfeita tomando o conjugado e o teorema está provado. \square

Para provar algumas propriedades de operadores Hilbert-Adjuntos é de grande utilidade o seguinte lema.

Lema 4.2. (Operador Nulo). *Sejam X e Y espaços com produto interno e $Q : X \rightarrow Y$ um operador linear limitado. Então:*

- (a) $Q = 0$ se, e somente se, $\langle Qx, y \rangle = 0$ para todos $x \in X$ e $y \in Y$.
- (b) Se $Q : X \rightarrow X$, onde X é complexo, e $\langle Qx, y \rangle = 0$ para todo $x \in X$, então $Q = 0$.

Demonstração. (a) Suponha $Q = 0$, isto é, $Qx = 0$ para todo $x \in X$. Assim

$$\langle Qx, y \rangle = \langle 0, y \rangle = 0.$$

Reciprocamente, suponha $\langle Qx, y \rangle = 0$ para todos $x \in X$ e $y \in Y$. Assim $Qx = 0$ para todo x pelo Lema 3.9. Portanto $Q = 0$.

(b) Suponha $\langle Qx, x \rangle = 0$ para todo $x \in X$. Em particular $\langle Qv, v \rangle = 0$ para cada $v = \alpha x + y \in X$, assim

$$\begin{aligned} 0 = \langle Q(\alpha x + y), \alpha x + y \rangle &= \langle \alpha Qx + Qy, \alpha x + y \rangle \\ &= \alpha \bar{\alpha} \langle Qx, x \rangle + \langle Qy, y \rangle + \alpha \langle Qx, y \rangle + \bar{\alpha} \langle Qy, x \rangle \\ &= |\alpha|^2 \langle Qx, x \rangle + \langle Qy, y \rangle + \alpha \langle Qx, y \rangle + \bar{\alpha} \langle Qy, x \rangle. \end{aligned}$$

Por hipótese os dois primeiros termos do segundo membro são iguais a zero. Logo

$$0 = \alpha \langle Qx, y \rangle + \bar{\alpha} \langle Qy, x \rangle.$$

Para $\alpha = 1$, tem-se

$$\langle Qx, y \rangle + \langle Qy, x \rangle = 0.$$

Para $\alpha = i$, tem-se $\bar{\alpha} = -i$ e

$$\langle Qx, y \rangle - \langle Qy, x \rangle = 0.$$

Logo

$$\begin{cases} \langle Qx, y \rangle + \langle Qy, x \rangle = 0 \\ \langle Qx, y \rangle - \langle Qy, x \rangle = 0 \end{cases} \Rightarrow 2\langle Qx, y \rangle = 0,$$

concluindo $Q = 0$ pelo item (a). □

Na parte (b) deste lema, é essencial que X seja complexo. De fato, a conclusão pode não ser verdadeira se X é real. Um contra-exemplo é a rotação $Q : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ dada por $Q(x, y) = (-y, x)$. Q é linear, e $Qz \perp z$, portanto $\langle Qz, z \rangle = 0$ para todo $z \in \mathbb{R}^2$, mas $Q \neq 0$.

Teorema 4.3. (Propriedades de operadores Hilbert-Adjunto). *Sejam H_1, H_2 espaços de Hilbert, $S : H_1 \rightarrow H_2$ e $T : H_1 \rightarrow H_2$ operadores lineares limitados e α qualquer escalar. Então:*

- (a) $\langle T^*y, x \rangle = \langle y, Tx \rangle$, para todo $x \in H_1$ e $y \in H_2$;
- (b) $(S + T)^* = S^* + T^*$;
- (c) $(\alpha T)^* = \bar{\alpha} T^*$;
- (d) $(T^*)^* = T$;
- (e) $\|T^*T\| = \|TT^*\| = \|T\|^2$;
- (f) $T^*T = 0 \Leftrightarrow T = 0$;
- (g) $(ST)^* = T^*S^*$, assumindo $H_2 = H_1$;
- (h) Se existe $(T^{-1})^*$ ou $(T^*)^{-1}$ então $(T^{-1})^* = (T^*)^{-1}$.

Demonstração.

(a) Sabe-se que $\langle Tx, y \rangle = \langle x, T^*y \rangle$, assim

$$\langle T^*y, x \rangle = \overline{\langle Tx, y \rangle} = \langle y, Tx \rangle.$$

(b) De (4.6), para todo x e y ,

$$\begin{aligned}\langle x, (S + T)^*y \rangle &= \langle (S + T)x, y \rangle = \langle Sx + Tx, y \rangle \\ &= \langle Sx, y \rangle + \langle Tx, y \rangle = \langle x, S^*y \rangle + \langle x, T^*y \rangle \\ &= \langle x, S^*y + T^*y \rangle = \langle x, (S^* + T^*)y \rangle.\end{aligned}$$

Pelo Lema 3.9 segue o resultado. Portanto $(S + T)^*y = (S^* + T^*)y$, para todo y .

(c) Note que

$$\begin{aligned}\langle (\alpha T)^*y, x \rangle &= \langle y, (\alpha T)x \rangle = \langle y, \alpha(Tx) \rangle \\ &= \bar{\alpha}\langle x, Tx \rangle = \bar{\alpha}\langle T^*y, x \rangle \\ &= \langle \bar{\alpha}T^*y, x \rangle.\end{aligned}$$

Aplicando o Lema 4.2(a) para $Q = (\alpha T)^* - \bar{\alpha}T^*$ tem-se,

$$\begin{aligned}\langle ((\alpha T)^* - \bar{\alpha}T^*)y, x \rangle &= \langle (\alpha T)^*y, x \rangle - \langle \bar{\alpha}T^*y, x \rangle \\ &= \langle \bar{\alpha}T^*y, x \rangle - \langle \bar{\alpha}T^*y, x \rangle \\ &= 0.\end{aligned}$$

Portanto $Q = 0$ e o item está provado.

(d) $(T^*)^*$ se escreve como T^{**} . Do item (a) e (4.6) tem-se

$$\langle (T^*)^*x, y \rangle = \langle x, T^*y \rangle = \langle Tx, y \rangle.$$

Assim, de modo análogo ao item anterior, (d) segue do Lema 4.2(a) para $Q = (T^*)^* - T$.

(e) Observe que, em geral, $T^*T \neq TT^*$ uma vez que $T^*T : H_1 \rightarrow H_1$ e $TT^* : H_2 \rightarrow H_2$ para $H_1 \neq H_2$. Pela Desigualdade de Schwarz,

$$\|Tx\|^2 = \langle Tx, Tx \rangle = \langle T^*Tx, x \rangle \leq \|T^*Tx\| \|x\| \leq \|T^*T\| \|x\|^2.$$

Tomando o supremo sobre todos os x de norma 1, obtém-se $\|T\|^2 \leq \|T^*T\|$. Como $\|T_1T_2\| \leq \|T_1\| \|T_2\|$, para todos operadores limitados T_1 e T_2 , deste fato e de (4.7) tem-se

$$\|T\|^2 \leq \|T^*T\| \leq \|T^*\| \|T\| = \|T\|^2.$$

Portanto $\|T^*T\| = \|T\|^2$. Substituindo T por T^* e usando (4.7), tem-se também

$$\|T^{**}T^*\| = \|T^*\|^2 = \|T\|^2,$$

onde $T^{**} = T$ pelo item (d). Logo (e) está provado.

(f) Do item anterior segue de imediato $T^*T = 0 \Leftrightarrow T = 0$.

(g) De (4.6) tem-se

$$\langle x, (ST)^*y \rangle = \langle (ST)x, y \rangle = \langle Tx, S^*y \rangle = \langle x, T^*S^*y \rangle.$$

Assim, $(ST)^*y = T^*S^*y$ pelo Lema 3.9, concluindo resultado.

(h) Suponha que exista $(A^*)^{-1}$, assim $A^*(A^{-1})^* = (A^{-1}A)^* = I^* = I$. Logo $(A^{-1})^* = (A^*)^{-1}$. De modo análogo $(A^{-1})^*A = I$.

□

Observação 4.4. No Teorema 4.3, os itens (b), (c), (d), (g) e (h) são satisfeitos com a hipótese de T e S serem transformações lineares, não havendo necessidade de serem limitados.

Proposição 4.2. *Sejam H um espaço de Hilbert e $T : H \rightarrow H$ um operador linear limitado. Então:*

1. *Se F é um subespaço vetorial de H invariante por T , isto é, $T(F) \subset F$, então F^\perp é invariante por T^* .*
2. *Se F é um subespaço invariante por T e T^* , então F^\perp é invariante por T e T^* .*

Demonstração. 1. Seja $x \in T^*(F^\perp)$, então existe $y \in F^\perp$ tal que $T^*y = x$. Uma vez que $y \in F^\perp$, tem-se $\langle y, w \rangle = 0$ para todo $w \in F$. Assim $\langle x, w \rangle = \langle T^*y, w \rangle = \langle y, Tw \rangle = 0$ para todo $w \in F$ pois $Tw \in F$. Logo $x \in F^\perp$.

2. Segue-se do item (d) do Teorema 4.3 e do item 1. desta proposição.

□

Exemplo 4.6. Sejam $H = \mathcal{C}_{L_2(\rho)}([a, b], \mathbb{C})$ e $K : [a, b] \times [a, b] \rightarrow \mathbb{C}$ uma função contínua. Considere o operador linear $k : H \rightarrow H$ definido por

$$(kf)(t) = \int_a^b K(t, s)f(s)\rho(s)ds,$$

onde $\rho \in \mathcal{C}([a, b], \mathbb{C})$ e $t \in [a, b]$. Então $(k^*f)(t) = \int_a^b K^*(t, s)f(s)\rho(s)ds$, onde $K^*(t, s) = \overline{K(s, t)}$.

De fato,

$$\begin{aligned} \langle kf, g \rangle_\rho &= \int_a^b \left[\int_a^b K(t, s)f(s)\rho(s)ds \right] \overline{g(t)\rho(t)}dt \\ &= \int_a^b f(s) \left[\overline{\int_a^b K^*(s, t)g(t)\rho(t)dt} \right] \rho(s)ds = \langle f, k^*g \rangle_\rho. \end{aligned}$$

4.4 Operadores Autoadjunto, Unitário e Normal

Vale lembrar, na Definição 3.7, que dado um espaço vetorial X , sobre \mathbb{K} , diz-se que uma forma (ou funcional) $f : X \times X \rightarrow \mathbb{K}$ é sesquilinear (bilinear se $\mathbb{K} = \mathbb{R}$) se para todos $x, x_1, x_2, y, y_1, y_2 \in X$ e todos escalares $\alpha, \beta \in \mathbb{K}$, tem-se

$$\begin{aligned} (a) \quad & f(x_1 + x_2, y) = f(x_1, y) + f(x_2, y) \\ (b) \quad & f(x, y_1 + y_2) = f(x, y_1) + f(x, y_2) \\ (c) \quad & f(\alpha x, y) = \alpha f(x, y) \\ (d) \quad & f(x, \beta y) = \overline{\beta} f(x, y). \end{aligned}$$

Exemplo 4.7. Se f é uma forma sesquilinear sobre X então $f^*(x, y) = \overline{f(y, x)}$ define uma forma sesquilinear sobre X e $f^{**} = f$.

Exemplo 4.8. Se f é uma forma sesquilinear sobre X e $T : X \rightarrow X$ é uma transformação linear, então $f_T(x, y) = f(T(x), y)$ define uma forma sesquilinear sobre X . No caso em que f é o produto interno de um espaço pré-hilbertiano X , tem-se

$$\varphi_T(x, y) = \langle T(x), y \rangle$$

e

$$\psi_T(x, y) = \langle x, T(y) \rangle.$$

Proposição 4.3. Identidade de Polarização. Se f é uma forma sesquilinear definida sobre X , então

- I. $2[f(x, y) + f(y, x)] = f(x + y, x + y) - f(x - y, x - y)$, no caso real;
- II. $4f(x, y) = f(x + y, x + y) - f(x - y, x - y) + if(x + iy, x + iy) - if(x - iy, x - iy)$, no caso complexo.

Demonstração. I.

$$\begin{aligned} f(x + y, x + y) - f(x - y, x - y) &= f(x, x + y) + f(y, x + y) - \\ &\quad - f(x, x - y) - f(-y, x - y) = \\ &= 2f(x, y) + 2f(y, x) = \\ &= 2[f(x, y) + f(y, x)]. \end{aligned}$$

II. Segue de modo análogo, desenvolvendo o segundo membro. \square

Proposição 4.4. Sejam X um espaço pré-hilbertiano e $T \in L(X)$, onde $L(X)$ denota o espaço vetorial das aplicações lineares contínuas de X em X munido da norma $\|T\| = \sup_{\|x\| \leq 1} \|T(x)\|$. Considere a forma sesquilinear $\phi : X \times X \rightarrow \mathbb{C}$ definida por $\phi_T(x, y) = \langle T(x), y \rangle$. Então $\|\phi_T\| = \|T\|$.

Demonstração. Da Desigualdade de Cauchy-Schwarz,

$$\begin{aligned} |\phi_T(x, y)| &= |\langle T(x), y \rangle| \leq \|T(x)\| \|y\| \leq \|T\| \|x\| \|y\| \Rightarrow \\ &\Rightarrow \sup_{\substack{\|x\| \leq 1 \\ \|y\| \leq 1}} |\langle T(x), y \rangle| \leq \sup_{\substack{\|x\| \leq 1 \\ \|y\| \leq 1}} \|T\| \Rightarrow \\ &\Rightarrow \|\phi_T\| \leq \|T\|. \end{aligned}$$

Por outro lado

$$0 \leq \|T(x)\|^2 = \langle T(x), T(x) \rangle = \phi_T(x, T(x)) \leq \|\phi_T\| \|x\| \|T(x)\|, \forall x \in X$$

concluindo que

$$\|T(x)\| \leq \|\phi_T\| \|x\|, \forall x \in X,$$

ou seja, $\|T\| \leq \|\phi_T\|$. Portanto segue que $\|\phi_T\| = \|T\|$. □

Corolário 4.1. *Sejam X um espaço pré-hilbertiano e $T : X \rightarrow X$ um operador linear limitado. Considerando a forma sesquilinear $\psi_T(x, y) = \langle x, Ty \rangle$, tem-se $\|\psi_T\| = \|T\| = \|\phi_T\|$.*

Definição 4.6. *Diz-se que um forma sesquilinear f definida sobre um espaço vetorial X é **autoadjunta** se para quaisquer $x, y \in X$ tem-se*

$$f(y, x) = \overline{f(x, y)} = f^*(y, x).$$

Diz-se simplesmente que f é hermitiana se X for um espaço vetorial sobre \mathbb{C} , e que f é simétrica, no caso real.

Exemplo 4.9. No Exemplo 4.6 segue que o operador k definido pelo núcleo K é hermitiano se, e somente se, o núcleo for hermitiano, isto é, se $K(s, t) = \overline{K(t, s)}$ para quaisquer $s, t \in [a, b]$.

Teorema 4.4. *Seja f uma forma sesquilinear sobre o espaço vetorial **complexo** X . São equivalentes as seguintes propriedades:*

1. f é hermitiana;
2. $f(x, x) \in \mathbb{R}$ para todo $x \in X$.

Demonstração. 1 \Rightarrow 2. Como $\overline{f(x, x)} = f^*(x, x) = f(x, x)$ segue que $f(x, x)$ é real, uma vez $f(x, x)$ é igual ao seu conjugado.

2 \Rightarrow 1. Segue da fórmula de polarização. □

Definição 4.7. (Operadores Autoadjunto, Unitário e Normal).

Seja $T : H \rightarrow H$ um operador linear limitado em um espaço de Hilbert H . Diz-se:

- (a) T é autoadjunto se $T^* = T$;
- (b) T é unitário se T é bijetor e $T^* = T^{-1}$;
- (c) T é normal se $TT^* = T^*T$.

No caso em que o corpo $\mathbb{K} = \mathbb{C}$ o operador autoadjunto é denominado **hermitiano**. E no caso em que o corpo $\mathbb{K} = \mathbb{R}$ denomina-se **simétrico**.

O operador Hilbert-Adjunto T^* de T como definido na seção anterior é tal que

$$\langle Tx, y \rangle = \langle x, T^*y \rangle.$$

Se T é autoadjunto, a igualdade acima torna-se

$$\langle Tx, y \rangle = \langle x, Ty \rangle. \quad (4.11)$$

Da definição acima, todo operador unitário ou autoadjunto é normal. Mas note que se T é um operador normal pode este não ser autoadjunto ou unitário. Por exemplo, se $I : H \rightarrow H$ é o operador identidade, então o operador $T = 2iI$ é normal. De fato, note que $\langle Tx, y \rangle = \langle 2ix, y \rangle = 2i\langle x, y \rangle = \langle x, -2iy \rangle = \langle x, -2iIy \rangle$, concluindo que $T^* = -2iI$. Como $TT^* = T^*T = 4I$ segue que T é normal com $T^* \neq T$ assim como $T^* \neq T^{-1} = -\frac{1}{2}iI$.

Teorema 4.5. (Autoadjunto).

Seja $T : H \rightarrow H$ um operador linear limitado em espaço de Hilbert H . Então:

- (a) Se T é autoadjunto, $\langle Tx, x \rangle$ é real para todo $x \in H$.
- (b) Se H é complexo e $\langle Tx, x \rangle$ é real para todo $x \in H$, o operador T é autoadjunto.

Demonstração. (a) Se T é autoadjunto, então para todo x ,

$$\overline{\langle Tx, x \rangle} = \langle x, Tx \rangle \stackrel{T=T^*}{=} \langle Tx, x \rangle.$$

Portanto $\langle Tx, x \rangle$ é um número real uma vez que ele é igual ao seu conjugado.

- (b) Se $\langle Tx, x \rangle$ é real para todo x , então

$$\langle Tx, x \rangle = \overline{\langle Tx, x \rangle} = \overline{\langle x, T^*x \rangle} = \langle T^*x, x \rangle.$$

Portanto,

$$0 = \langle Tx, x \rangle - \langle T^*x, x \rangle = \langle (T - T^*)x, x \rangle$$

e $T - T^* = 0$ pelo Lema 4.2(b) uma vez que H é complexo. □

Na parte (b) do teorema anterior é essencial que H seja complexo. Isto é claro uma vez que para um H real o produto interno é um número real, o que torna $\langle Tx, x \rangle$ real sem quaisquer suposições sobre o operador linear.

Definição 4.8. Seja X um espaço vetorial sobre o corpo \mathbb{K} e seja $T : X \rightarrow X$ um operador linear. Um vetor $x \neq 0$ em X chama-se **autovetor** do operador T quando existe $\lambda \in \mathbb{K}$ tal que

$$Tx = \lambda x.$$

O número λ , por sua vez, denomina-se um **autovalor** do operador T nas condições acima. Diz-se então que o autovalor λ corresponde ao autovetor x e vice-versa, que o autovetor x também corresponde ao autovalor λ .

Observação 4.5. Fixado λ , o conjunto $X(\lambda) = \{x \in X \mid Tx = \lambda x\}$ é um subespaço vetorial de X pois

$$Tx = \lambda x \Leftrightarrow (T - \lambda I)x = 0 \Leftrightarrow x \in \ker(T - \lambda I),$$

ou seja, o subconjunto definido coincide com $\ker(T - \lambda I)$ que é um subespaço vetorial de X .

Exemplo 4.10. Seja $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ dado por $T(x, y) = (y, x)$. Note que a aplicação T é a reflexão dos vetores em torno da reta $y = x$. Assim, se o vetor está no eixo \vec{Ox} , sua imagem está no eixo \vec{Oy} . Não há portanto, autovetores no eixo \vec{Ox} . No entanto se (x, y) está na reta $y = x$, tem-se $T(x, y) = (x, y)$ e assim todo vetor da reta é um autovetor cujo autovalor é igual a 1.

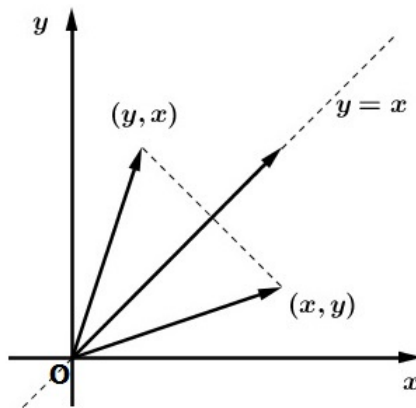


Figura 4.1: Reflexão de vetores com respeito a diagonal.

Proposição 4.5. 1. Os autovalores de um operador hermitiano são reais.

2. Os autovetores correspondentes a autovalores distintos de um operador hermitiano são ortogonais.

Demonstração. Sejam $Tx = \lambda x$ e $Ty = \alpha y$ com $\lambda \neq \alpha$ e $x \neq 0$ e $y \neq 0$.

1. $\lambda \|x\|^2 = \lambda \langle x, x \rangle = \langle \lambda x, x \rangle = \langle Tx, x \rangle = \langle x, Tx \rangle = \langle x, \lambda x \rangle = \bar{\lambda} \|x\|^2 \Rightarrow \lambda = \bar{\lambda}$.
Portanto λ é real.

2. $\lambda \langle x, y \rangle = \langle Tx, y \rangle = \langle x, Ty \rangle = \langle x, \alpha y \rangle = \alpha \langle x, y \rangle \Rightarrow \langle x, y \rangle (\lambda - \alpha) = 0 \Rightarrow \langle x, y \rangle = 0$.
Portanto x, y são ortogonais. \square

Proposição 4.6. Sejam H um espaço pré-hilbertiano e f uma forma hermitiana contínua definida em H . Tem-se

$$\|f\| = \sup_{\|x\| \leq 1} |f(x, x)|,$$

isto é,

$$\sup\{|f(x, y)| \mid \|x\| \leq 1, \|y\| \leq 1\} = \sup_{\|x\| \leq 1} |f(x, x)|.$$

Demonstração. Seja $s_f = \sup_{\|x\| \leq 1} |f(x, x)|$, como $|f(x, x)| \leq \|f\| \|x\| \|x\|$, é evidente que $s_f \leq \|f\|$. Para que $\|f\| \leq s_f$ é suficiente mostrar que dados $x, y \in H$ com $\|x\| \leq 1$, $\|y\| \leq 1$ tem-se $|f(x, y)| \leq s_f$. Se $f(x, y)$ é real então pela fórmula de polarização (quando $\mathbb{K} = \mathbb{R}$ a parte imaginária do segundo membro é necessariamente nula) tem-se

$$f(x, y) = \frac{1}{4} [f(x + y, x + y) - f(x - y, x - y)],$$

logo,

$$\begin{aligned} |f(x, y)| &= \left| \frac{1}{4} [f(x + y, x + y) - f(x - y, x - y)] \right| \\ &\leq \frac{1}{4} |f(x + y, x + y)| + \frac{1}{4} |f(x - y, x - y)| \\ &\leq \frac{1}{4} s_f \|x + y\|^2 + \frac{1}{4} s_f \|x - y\|^2 = \frac{1}{4} [s_f \|x + y\|^2 + s_f \|x - y\|^2] \\ &= \frac{1}{2} [s_f (\|x\|^2 + \|y\|^2)] \leq s_f. \end{aligned}$$

Portanto, no caso em que $f(x, y)$ é real tem-se $|f(x, y)| \leq s_f$.

Se $f(x, y)$ não é real, considere μ tal que

$$|f(x, y)| = \mu f(x, y) = f(\mu x, y),$$

onde μ é um número complexo e $|\mu| = 1$. Assim $\|\mu x\| = \|x\| \leq 1$ e de modo análogo a primeira parte da demonstração segue que $|f(x, y)| = |f(\mu x, y)| \leq s_f$. \square

Corolário 4.2. Seja $T \in L(E)$ um operador hermitiano, então

$$\|T\| = \sup_{\|x\| \leq 1} |\langle Tx, x \rangle|.$$

Demonstração. Sendo T hermitiano, pelo Teorema 4.5, tem-se $\langle Tx, x \rangle$ real para todo $x \in E$. Ainda, sendo $\phi_T(x, y) = \langle Tx, y \rangle$ uma forma sesquilinear, segue que $\phi_T(x, y)$ é hermitiana pelo Teorema 4.4. Por conseguinte o resultado segue da Proposição 4.6. \square

Teorema 4.6. (Produto de operadores Autoadjuntos). *Sejam S e T dois operadores lineares limitados autoadjuntos em um espaço de Hilbert H . Então o produto de S por T é autoadjunto se, e somente se,*

$$ST = TS.$$

Demonstração. Pelo item (g) do Teorema 4.3 da seção anterior e pela hipótese, tem-se

$$(ST)^* = T^*S^* = TS.$$

Portanto,

$$ST = (ST)^* \Leftrightarrow ST = TS.$$

\square

Teorema 4.7. (Sequências de operadores Autoadjuntos). *Seja (T_n) uma sequência de operadores lineares limitados autoadjuntos $T_n : H \rightarrow H$ em um espaço de Hilbert H . Suponha que (T_n) seja convergente, por exemplo, $T_n \rightarrow T$, ou equivalentemente, $\|T_n - T\| \rightarrow 0$, onde $\|\cdot\|$ é a norma no espaço $\mathcal{B}(H, H)$. Nestas condições T é um operador linear limitado autoadjunto em H .*

Demonstração. Observe que

$$\|T_n^* - T^*\| = \|(T_n - T)^*\| = \|T_n - T\|,$$

pelos Teorema 4.2 e 4.3. Da desigualdade triangular no espaço $\mathcal{B}(H, H)$, tem-se

$$\begin{aligned} \|T - T^*\| &= \|T - T_n + T_n - T_n^* + T_n^* - T^*\| \\ &\leq \|T - T_n\| + \|T_n - T_n^*\| + \|T_n^* - T^*\| \\ &= \|T - T_n\| + 0 + \|T_n - T\| \\ &= 2\|T_n - T\|, \end{aligned}$$

que converge para zero quando n tende ao infinito.

Portanto $\|T - T^*\| = 0$ e $T = T^*$. \square

Teorema 4.8. (Operador Unitário). *Sejam $U : H \rightarrow H$ e $V : H \rightarrow H$ dois operadores unitários, onde H é um espaço de Hilbert. Então:*

(a) U é isométrico, isto é, $\|Ux\| = \|x\|$ para todo $x \in H$,

(b) $\|U\| = 1$, provida de $H \neq 0$,

(c) $U^{-1}(= U^*)$ é unitário,

(d) UV é unitário,

(e) U é normal.

Além disso:

(f) Um operador linear T em um espaço de Hilbert complexo H é unitário se, e somente se, T é isométrico e sobrejetivo.

Demonstração. (a) Como

$$\|Ux\|^2 = \langle Ux, Ux \rangle = \langle x, U^*Ux \rangle = \langle x, U^{-1}Ux \rangle = \langle x, Ix \rangle = \|x\|^2 \quad \forall x \in H,$$

tem-se $\|Ux\| = \|x\|$.

(b) Segue de imediato do item anterior.

(c) Uma vez que U é bijetor, U^{-1} também o é, e pelo Teorema 4.3,

$$(U^{-1})^* = U^{**} = U = (U^{-1})^{-1}.$$

Portanto U^{-1} é unitário.

(d) Sendo U e V bijetores, segue que UV também o é, e pelo Teorema 4.3,

$$(UV)^* = (V^*U^*) = V^{-1}U^{-1} = (UV)^{-1}.$$

Portanto UV é unitário.

(e) Basta notar que $UU^* = UU^{-1} = I = U^{-1}U = U^*U$.

(f) Suponha que T seja isométrico e sobrejetor. Como isometria implica em injetividade, segue que T é bijetor. Deve-se ter $T^* = T^{-1}$. Pela isometria,

$$\langle T^*Tx, x \rangle = \langle Tx, Tx \rangle = \langle x, x \rangle = \langle Ix, x \rangle.$$

Assim ,

$$\langle (T^*T - I)x, x \rangle = 0$$

e $(T^*T - I) = 0$ pelo Lema 4.2(b), logo $T^* = T$. Disto

$$TT^* = TT^*I = TT^*(TT^{-1}) = T(T^*T)T^{-1} = TIT^{-1} = I.$$

Portanto, $T^*T = TT^* = T$. Deste modo $T^* = T$, concluindo que T é unitário. A recíproca é imediata pois T é isométrico pelo item (a) e sobrejetor pela própria definição. □

Teorema 4.9. *Seja $T : H \rightarrow H$ um operador linear limitado em um espaço de Hilbert H e tal que T^* existe. O operador T é normal se, e somente se, para todo $x \in H$ tem-se $\|T^*x\| = \|Tx\|$.*

Demonstração. Se T é normal,

$$\|Tx\|^2 = \langle Tx, Tx \rangle = \langle T^*Tx, x \rangle = \langle TT^*x, x \rangle = \langle T^*x, T^*x \rangle = \|T^*x\|^2,$$

ou seja, $\|Tx\| = \|T^*x\|$, $\forall x \in H$.

Reciprocamente, se $\|Tx\| = \|T^*x\|$, então

$$\begin{aligned} \|Tx\|^2 = \|T^*x\|^2 &\Rightarrow \langle TT^*x, x \rangle = \langle T^*Tx, x \rangle \Rightarrow \\ &\Rightarrow \langle TT^*x, x \rangle - \langle T^*Tx, x \rangle = 0 \Rightarrow \\ &\Rightarrow \langle (TT^* - T^*T)x, x \rangle = 0, \end{aligned}$$

para todo $x \in H$. Pelo Lema 4.2 $TT^* - T^*T = 0$, isto é, T é normal. \square

Corolário 4.3. *Se $T : H \rightarrow H$ é normal então $\|T^2\| = \|T\|^2$.*

Demonstração. Como T é normal, $\|Tx\| = \|T^*x\|$, tomando $x = Ty$ tem-se $\|TTY\| = \|T^*Ty\|$, para todo $y \in H$, isto é, $\|T^2y\| = \|T^*Ty\|$. Do item (e) do Teorema 4.3 segue que $\|T^2\| = \|T^*T\| = \|T\|^2$. \square

Corolário 4.4. *Seja $T : H \rightarrow H$ um operador normal. x é um autovetor de T correspondente ao autovalor λ se, e somente se, x é um autovetor de T^* correspondente ao autovalor $\bar{\lambda}$.*

Demonstração. Sendo T um operador normal, segue que $U = T - \lambda I$ é normal, basta notar que $(T - \lambda I)^* = T^* - \bar{\lambda}I$. Logo

$$Tx = \lambda x \Leftrightarrow \|(T - \lambda I)x\| = 0 \Leftrightarrow \|(T - \lambda I)^*x\| = 0 \Leftrightarrow T^*x - \bar{\lambda}Ix = 0.$$

Portanto $T^*x = \bar{\lambda}x$. \square

Corolário 4.5. *Seja $T : H \rightarrow H$ um operador normal. Se x e y são autovetores de T correspondentes a autovalores distintos λ e μ , então x e y são ortogonais, isto é, $\langle x, y \rangle = 0$.*

Demonstração. Do Corolário anterior,

$$\lambda \langle x, y \rangle = \langle \lambda x, y \rangle = \langle Tx, y \rangle = \langle x, T^*y \rangle = \langle x, \bar{\mu}y \rangle = \bar{\mu} \langle x, y \rangle.$$

Assim $\lambda \langle x, y \rangle - \bar{\mu} \langle x, y \rangle = (\lambda - \bar{\mu}) \langle x, y \rangle = 0 \Rightarrow \langle x, y \rangle = 0$, uma vez que $\lambda \neq \bar{\mu}$. \square

Corolário 4.6. *Seja $T : H \rightarrow H$ um operador normal. Se x é um autovetor de T então $[x]^\perp$ é invariante por T e por T^* .*

Demonstração. Como x é um autovetor de T existe $\lambda \in \mathbb{K}$ tal que $Tx = \lambda x$ e, sendo T normal, pelo Corolário 4.4, $T^*x = \bar{\lambda}x$. Sendo $[x]$ um subespaço de H invariante por T e T^* , do item 2. da Proposição 4.2 segue o resultado, isto é, $[x]^\perp$ é invariante por T e T^* . \square

5 Teoria Espectral

Este capítulo apresenta um resultado muito importante que generaliza o Teorema Espectral em dimensão finita para operadores definidos em espaços de dimensão infinita. E será uma ferramenta essencial para a prova do principal resultado do próximo capítulo, sobre a teoria de Sturm-Liouville. As principais referências deste capítulo são [7] e [8].

5.1 Teoria Espectral dos Operadores Compactos

Nesta seção H será um espaço pré-hilbertiano.

Proposição 5.1. *Dado um autovalor λ de um operador $T \in \mathcal{L}(H)$, onde $\mathcal{L}(H)$ indica o espaço das aplicações lineares contínuas de H em H . Então $|\lambda| \leq \|T\|$.*

Demonstração. Sendo $Tx = \lambda x$, tem-se

$$|\lambda|\|x\| = \|\lambda x\| = \|Tx\| \leq \|T\|\|x\|.$$

Portanto $|\lambda| \leq \|T\|$. □

Seja $x \in H$, com $\|x\| = 1$, pelo Lema 3.1 tem-se $|\langle Tx, x \rangle| \leq \|Tx\|\|x\| \leq \|T\|$. Entretanto a igualdade só é satisfeita nas condições da proposição que se segue.

Proposição 5.2. *Seja $x \in H$ com $\|x\| = 1$, tem-se $|\langle Tx, x \rangle| = \|T\|$ se, e somente se, x é um autovetor de T de autovalor $\lambda = \langle Tx, x \rangle$ com $|\lambda| = \|T\|$.*

Demonstração. Suponha que $x \in H$ com $\|x\| = 1$ tal que $|\langle Tx, x \rangle| = \|T\|$. Tomando $\lambda = \langle Tx, x \rangle$ tem-se $|\lambda| = |\langle Tx, x \rangle|$ e

$$\begin{aligned} 0 &\leq \|Tx - \lambda x\|^2 = \langle Tx - \lambda x, Tx - \lambda x \rangle \\ &= \langle Tx, Tx \rangle - \langle Tx, \lambda x \rangle - \langle \lambda x, Tx \rangle + \langle \lambda x, \lambda x \rangle \\ &= \|Tx\|^2 - \bar{\lambda}\langle Tx, x \rangle - \lambda\langle x, Tx \rangle + |\lambda|^2 \\ &= \|T\|^2 - \overline{\langle Tx, x \rangle}\langle Tx, x \rangle - \langle Tx, x \rangle\overline{\langle Tx, x \rangle} + \|T\|^2 \\ &= 0. \end{aligned}$$

Assim, $\|Tx - \lambda x\| = 0$ o que implica $Tx - \lambda x = 0$, isto é, $Tx = \lambda x$. Portanto x é um autovetor de T .

Reciprocamente, $|\langle Tx, x \rangle| = |\langle \lambda x, x \rangle| = |\lambda\langle x, x \rangle| = |\lambda| = \|T\|$. □

Proposição 5.3. *Seja T um operador hermitiano contínuo. A função $g : S \rightarrow \mathbb{R}$, definida por $g(x) = |\langle Tx, x \rangle|$, onde $S = \{x \in H \mid \|x\| = 1\}$, atinge seu máximo num ponto $x_1 \in S$ se, e somente se, x_1 é um autovetor de T correspondente ao autovalor $\|T\|$ ou $-\|T\|$.*

Demonstração. Suponha que x_1 seja um autovetor de T correspondente ao autolavor $\|T\|$. Como $\|x_1\| = 1$, pela Proposição 5.2 tem-se $\|T\| = |\langle Tx_1, x_1 \rangle|$, do Corolário 4.2,

$$g(x_1) = |\langle Tx_1, x_1 \rangle| = \|T\| = \sup_{\|x\| \leq 1} |\langle Tx, x \rangle| = \sup_{\|x\| \leq 1} g.$$

Portanto g atinge seu ponto máximo em $x_1 \in S$. Reciprocamente, suponha que $g(x_1) = \sup_{\|x\| \leq 1} g$, do Teorema 4.6 segue que $g(x_1) = |\langle Tx_1, x_1 \rangle| = \|T\|$. Portanto o resultado segue da Proposição 5.2. \square

Vale lembrar, da Proposição 4.5, que se T é um operador hermitiano sobre H , todos os autovalores de T são reais e que autovetores correspondentes a autovalores distintos são ortogonais. Além disso segue do item 2. da Proposição 4.2 que se F é um subespaço de H invariante por T e por T^* então F^\perp também é invariante por T .

O próximo teorema vem nos garantir a existência de um autovalor de um operador. Vale observar que a existência nem sempre é garantida.

Teorema 5.1. *Seja $T : H \rightarrow H$ um operador hermitiano compacto, $T \neq 0$, então $\|T\|$ ou $-\|T\|$ é um autovalor de T .*

Demonstração. Deve-se mostrar que existe $y \neq 0$ com $Ty = \lambda y$ onde $|\lambda| = \|T\|$. Sendo T um operador hermitiano, pelo Corolário 4.2 tem-se $\|T\| = \sup_{\|x\| \leq 1} |\langle Tx, x \rangle|$. Assim, dado $\epsilon = 1$ existe $x_1 \in H$ com $\|x_1\| = 1$ tal que $\|T\| - 1 < |\langle Tx_1, x_1 \rangle|$, tomando $\epsilon = 1/2$, existe $x_2 \in H$ com $\|x_2\| = 1$ tal que $\|T\| - 1/2 < |\langle Tx_2, x_2 \rangle|$. Prosseguindo de modo análogo, tomando $\epsilon = 1/n$ é possível obter $x_n \in H$ com $\|x_n\| = 1$ tal que $\|T\| - 1/n < |\langle Tx_n, x_n \rangle|$. Logo, fazendo $n \rightarrow \infty$ tem-se $|\langle Tx_n, x_n \rangle| \rightarrow \|T\|$.

A sequência $(\langle Tx_n, x_n \rangle)_{n \in \mathbb{N}}$ de números reais, limitada, admite uma subsequência $(\langle Tx_k, x_k \rangle)_{k \in \mathbb{N}'}$ tal que $\langle Tx_k, x_k \rangle \rightarrow \lambda$ onde $\lambda = \|T\|$ ou $\lambda = -\|T\|$ e portanto $|\lambda| = \|T\|$. Assim

$$\begin{aligned} 0 &\leq \|Tx_k - \lambda x_k\|^2 = \langle Tx_k - \lambda x_k, Tx_k - \lambda x_k \rangle \\ &= \|Tx_k\|^2 - \langle Tx_k, \lambda x_k \rangle - \langle \lambda x_k, Tx_k \rangle + |\lambda|^2 \\ &= \|T\|^2 - 2\lambda \langle Tx_k, x_k \rangle + \|T\|^2, \end{aligned}$$

que converge para zero quando k tende ao infinito pois $\langle Tx_k, x_k \rangle \rightarrow \lambda$. Portanto $\|Tx_k - \lambda x_k\| \rightarrow 0$. Como a sequência (x_n) é limitada e T um operador compacto, pela Proposição 4.1 existe uma subsequência (x_{r_n}) tal que (Tx_{r_n}) converge para um elemento $y \in H$.

Considerando a sequência $(\lambda x_{r_n}) = (Tx_{r_n}) - [(Tx_{r_n}) - (\lambda x_{r_n})]$, tem-se que $\lambda x_{r_n} \rightarrow y$, quando $n \rightarrow \infty$, uma vez que $Tx_{r_n} \rightarrow y$ e $Tx_n \rightarrow \lambda x_n$. Tem-se, também, que T é um

operador contínuo e linear, assim $T(\lambda x_{r_n}) \rightarrow Ty$ e $T(\lambda x_{r_n}) = \lambda T(x_{r_n})$. Logo $T(\lambda x_{r_n}) \rightarrow \lambda y$, pela unicidade do limite, tem-se $Ty = \lambda y$. Observe que $\|y\| = \|\lambda x_{r_n}\| = |\lambda| \neq 0$, o que completa a demonstração. \square

Teorema 5.2. (Teorema espectral dos operadores hermitianos compactos.)

Sejam H um espaço pré-hilbertiano (real ou complexo) e $T : H \rightarrow H$ um operador hermitiano compacto, $T \neq 0$. Existe uma seqüência $(\lambda_n) \in \mathbb{R}$ (finita ou infinita) de autovalores não nulos de T e uma seqüência (x_n) de autovetores correspondentes que formam um conjunto ortonormal tal que cada elemento $x \in H$ tem-se

$$Tx = \sum_n \lambda_n \beta_n x_n \quad (5.1)$$

onde $\beta_n = \langle x, x_n \rangle$.

Tem-se $|\lambda_i| \geq |\lambda_{i+1}|$; a seqüência contém todos os autovalores não nulos de T e se ela for infinita tem-se $|\lambda_n| \rightarrow 0$.

Dado um particular $\lambda = \lambda_m$ a dimensão do subespaço T_λ , gerado pelos autovetores correspondentes ao autovalor λ , é finita e é igual ao número de vezes que o autovalor λ aparece na seqüência (λ_n) .

Demonstração. Considere o autovalor λ_1 e o autovetor unitário x_1 correspondente de T , garantidos pelo Teorema 5.1. Tome $H_1 = H$ e $T_1 = T$, tem-se $|\lambda_1| = \|T_1\|$.

Fazendo $H_2 = [x_1]^\perp$, note que H_2 é um subespaço de H_1 . Com efeito, para quaisquer y e w em H_2 e α um escalar qualquer de \mathbb{K} , tem-se

$$\langle \alpha y + w, x_1 \rangle = \langle \alpha y, x_1 \rangle + \langle w, x_1 \rangle = \alpha \langle y, x_1 \rangle + \langle w, x_1 \rangle = 0,$$

ou seja, $\alpha y + w \in H_2$ concluindo que H_2 é um subespaço vetorial de H_1 . Pelo Corolário 4.6, $H_2 = [x_1]^\perp$ é um subespaço de H_1 invariante por T_1 . Chame de T_2 a restrição de T_1 a H_2 , tem-se então que T_2 é um operador hermitiano compacto e novamente pelo Teorema 5.1 existe um autovalor λ_2 e um autovetor unitário correspondente x_2 de T_2 , que em particular é, também, um autovetor de T , tal que $|\lambda_2| = \|T_2\|$. Assim, da Proposição 5.1, segue que $|\lambda_2| = \|T_2\| \leq \|T_1\|$, ou seja, $|\lambda_2| \leq |\lambda_1|$.

Prosseguindo de modo análogo obtém-se sucessivamente autolavores $\lambda_1, \dots, \lambda_n$ de T não nulos tais que

$$|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_n|,$$

e autovetores unitários correspondentes x_1, \dots, x_n . Obtém-se então uma seqüência ortonormal x_1, \dots, x_n e subespaços H_2, H_3, \dots, H_{n+1} , onde H_{i+1} indica o subespaço de H_i (ou de H) formado pelos vetores ortogonais a x_1, \dots, x_i .

1. Se a restrição T_{n+1} de T a H_{n+1} for nula tem-se para cada $x \in H$

$$Tx = \sum_{i=1}^n \lambda_i \beta_i x_i,$$

onde $\beta_i = \langle x, x_i \rangle$, isto é, $T(H)$ é o subespaço de H gerado pelos vetores x_1, \dots, x_n .

De fato, seja $\tilde{x} = x - \sum_{i=1}^n \beta_i x_i$, assim

$$\begin{aligned} \langle \tilde{x}, x_i \rangle &= \langle x - (\beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n), x_i \rangle \\ &= \langle x - \beta_1 x_1 - \beta_2 x_2 - \dots - \beta_n x_n, x_i \rangle \\ &= \langle x, x_i \rangle - \beta_1 \langle x_1, x_i \rangle - \beta_2 \langle x_2, x_i \rangle - \dots - \beta_n \langle x_n, x_i \rangle \\ &= \langle x, x_i \rangle - \beta_i \langle x_i, x_i \rangle = \beta_i - \beta_i = 0, \end{aligned}$$

ou seja, $\langle \tilde{x}, x_i \rangle = 0$ para $i = 1, \dots, n$ e, portanto, $\tilde{x} \in H_{n+1}$. Por hipótese a restrição T_{n+1} de T a H_{n+1} é identicamente nula, isto é, $T\tilde{x} = 0$, logo

$$Tx = T(\tilde{x} + \sum_{i=1}^n \beta_i x_i) = T\tilde{x} + T\left(\sum_{i=1}^n \beta_i x_i\right) = \sum_{i=1}^n \beta_i T x_i = \sum_{i=1}^n \lambda_i \beta_i x_i.$$

2. Se para todo inteiro positivo n , a restrição T_{n+1} de T a H_{n+1} for sempre não nula então o processo acima nos dá uma seqüência infinita λ_n de autovalores não nulos de T com

$$|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_n| \geq \dots$$

e um sistema ortonormal $(x_n)_{n \in \mathbb{N}}$ formado pelos autovetores correspondentes.

- (a) A seqüência decrescente $|\lambda_n|$ tende para zero, pois, caso contrário, existiria um $\epsilon > 0$ tal que $|\lambda_n| \geq \epsilon$ para todo $n \in \mathbb{N}$, e ainda, a seqüência $\left(\frac{1}{\lambda_n} x_n\right)_{n \in \mathbb{N}}$ seria limitada, uma vez que $\left\|\frac{x_n}{\lambda_n}\right\| = \frac{\|x_n\|}{\|\lambda_n\|} \leq \frac{1}{\epsilon}$, sem que a seqüência $T\left(\frac{1}{\lambda_n} x_n\right) = \frac{1}{\lambda_n} T x_n = \frac{1}{\lambda_n} \lambda_n x_n = x_n$ contenha uma subsequência convergente, pois, $(x_n)_{n \in \mathbb{N}}$ é formada por vetores ortonormais. Contrariando o fato de T ser um operador compacto.

- (b) Para cada $x \in H$ tem-se que a série $\sum_{n=1}^{\infty} \lambda_n \beta_n x_n$ converge para Tx . Para isso, basta mostrar que dado $x \in H$ e $\epsilon > 0$ existe um natural m_0 tal que para todo $m \geq m_0$ tem-se

$$\left\|Tx - \sum_{n=1}^m \lambda_n \beta_n x_n\right\| < \epsilon.$$

Seja $x^{(m+1)} = x - \sum_{n=1}^m \beta_n x_n$ assim, de modo análogo ao que foi feito em 1, $x^{(m+1)} \in H_{m+1}$, pois $\langle x^{(m+1)}, x_i \rangle = 0$ para todo $x = i, \dots, n$. Então

$$\begin{aligned}
x^{(m+1)} = x - \sum_{n=1}^m \beta_n x_n &\Rightarrow x^{(m+1)} + \sum_{n=1}^m \beta_n x_n = x \Rightarrow \\
&\Rightarrow \|x^{(m+1)} + \sum_{n=1}^m \beta_n x_n\|^2 = \|x\|^2 \Rightarrow \\
&\Rightarrow \|x^{(m+1)}\|^2 = \|x\|^2 - \sum_{n=1}^m |\beta_n|^2 \\
&\|x^{(m+1)}\|^2 \leq \|x\|^2.
\end{aligned}$$

De $\|T_{m+1}\| = |\lambda_{m+1}|$ tem-se

$$\|Tx^{(m+1)}\| \leq \|T\| \|x^{(m+1)}\| \leq \|T_{m+1}\| \|x^{(m+1)}\| \leq |\lambda_{m+1}| \|x\|$$

e como a sequência $|\lambda_n|$ converge monotonicamente para 0, basta tomar m_0 tal que $|\lambda_{m_0}| \leq \frac{\epsilon}{\|x\|}$ para que se tenha

$$\|Tx - \sum_{n=1}^m \lambda_n \beta_n x_n\| = \|Tx^{(m+1)}\| \leq \epsilon,$$

se $m \geq m_0$.

- (c) Todo autovalor $\lambda \neq 0$ de T se encontra na sequência (λ_n) , pois caso contrário o autovetor correspondente x seria ortogonal a todos os x_n e de (b) seguiria que $Tx = 0$ contradizendo o fato de $Tx = \lambda x \neq 0$.
- (d) Dado um autovalor $\lambda \neq 0$ que aparece p vezes na sequência λ_n , então o subespaço gerado pelos autovetores correspondentes ao autovalor λ tem dimensão maior ou igual a p , pois existem pelo menos p autovetores ortonormais correspondentes a λ . O subespaço não pode ter dimensão maior que p , pois caso contrário existiria ainda um autovetor x correspondente a λ , ortogonal aos anteriores e a todos os x_n . Como em (c) seguiria que $Tx = 0$.

□

Corolário 5.1. Para quaisquer $x, y \in H$ tem-se

$$\langle Tx, y \rangle = \sum_{n=1}^{\infty} \lambda_n \beta_n \overline{\alpha_n},$$

onde $\overline{\alpha_n} = \langle x_n, y \rangle$.

Demonstração. Por (5.1), tem-se

$$\begin{aligned} \langle Tx, y \rangle &= \left\langle \sum_{n=1}^{\infty} \lambda_n \beta_n x_n, y \right\rangle = \langle \lambda_1 \beta_1 x_1, y \rangle + \langle \lambda_2 \beta_2 x_2, y \rangle + \cdots + \langle \lambda_k \beta_k x_k, y \rangle + \cdots = \\ &= \lambda_1 \beta_1 \langle x_1, y \rangle + \lambda_2 \beta_2 \langle x_2, y \rangle + \cdots + \lambda_k \beta_k \langle x_k, y \rangle + \cdots = \\ &= \sum_{n=1}^{\infty} \lambda_n \beta_n \langle x_n, y \rangle = \sum_{n=1}^{\infty} \lambda_n \beta_n \overline{\alpha_n}, \end{aligned}$$

onde $\beta_n = \langle x, x_n \rangle$. □

Definição 5.1. Um operador linear T de um espaço pré-hilbertiano H é **positivo** se $\langle Tx, x \rangle \geq 0$ para todo $x \in H$.

Observação 5.1. Do Teorema 4.5 tem-se que todo operador positivo T é hermitiano. Já, do Teorema Espectral, se T é um operador hermitiano compacto, T é positivo se, e somente se, todos os seus autovalores são positivos.

Teorema 5.3. Sejam H um espaço pré-hilbertiano (real ou complexo) e $T : H \rightarrow H$ um operador hermitiano compacto, $T \neq 0$. Dado $\lambda \in \mathbb{K}$ ($\mathbb{K} = \mathbb{R}$ ou \mathbb{C}), $\lambda \neq 0$ tal que $\lambda \neq \lambda_n$ para todo n . Então o operador $\lambda I - T$ tem inverso contínuo definido em H ; tal inverso é indicado por $(\lambda I - T)^{-1}$ e $x = (\lambda I - T)^{-1}y$ é dado por

$$x = \frac{1}{\lambda}y + \frac{1}{\lambda} \sum_{n=1}^{\infty} \lambda_n \frac{\alpha_n}{\lambda - \lambda_n} x_n, \quad (5.2)$$

onde $\alpha_n = \langle y, x_n \rangle$.

Demonstração. A demonstração se dará por partes.

1. Dado $y \in H$ e considerando a equação $\lambda x - Tx = y$, do Teorema Espectral 5.2 tem-se que

$$\lambda x - y = Tx = \sum_{n=1}^{\infty} \lambda_n \beta_n x_n. \text{ Aplicando e efetuando o produto interno por } x_m \text{ nos dois membros da equação obtém-se}$$

$$\langle \lambda x - y, x_m \rangle = \left\langle \sum_{n=1}^{\infty} \lambda_n \beta_n x_n, x_m \right\rangle$$

$$\langle \lambda x, x_m \rangle - \langle y, x_m \rangle = \sum_{n=1}^{\infty} \lambda_n \beta_n \langle x_n, x_m \rangle$$

$$\lambda \beta_m - \alpha_m = \lambda_m \beta_m$$

$$\beta_m = \frac{\alpha_m}{\lambda - \lambda_m}.$$

Assim $\lambda x - y = \sum_{n=1}^{\infty} \lambda_n \frac{\alpha_n}{\lambda - \lambda_n} x_n$, que é equivalente a (5.2) e a equação $\lambda x - Tx = y$ tem uma solução x .

2. Note que se a série $\sum_{n=1}^{\infty} \lambda_n \frac{\alpha_n}{\lambda - \lambda_n} x_n$ for convergente, o elemento x definido por $x = \frac{1}{\lambda} y + \frac{1}{\lambda} \sum_{n=1}^{\infty} \lambda_n \frac{\alpha_n}{\lambda - \lambda_n} x_n$ satisfaz a equação $(\lambda - T)x = y$.
3. A série de (5.2) satisfaz o critério de Cauchy. De fato, sejam

$$\alpha = \sup_n \left| \frac{\lambda_n}{\lambda - \lambda_n} \right| \text{ e } \beta = \sup_n \left| \frac{1}{\lambda - \lambda_n} \right|$$

que são números finitos uma vez que $\lambda \neq \lambda_n$, $\lambda \neq 0$ e $|\lambda_n| \rightarrow 0$. Considere agora $v_m = \sum_{n=1}^m \frac{\alpha_n}{\lambda - \lambda_n} x_n$, assim $u_m = Tv_m = \sum_{n=1}^m \lambda_n \frac{\alpha_n}{\lambda - \lambda_n} x_n$. Então

$$\begin{aligned} \|u_{m+p} - u_m\|^2 &= \left\| \sum_{n=1}^{m+p} \lambda_n \frac{\alpha_n}{\lambda - \lambda_n} x_n - \sum_{n=1}^m \lambda_n \frac{\alpha_n}{\lambda - \lambda_n} x_n \right\|^2 = \left\| \sum_{n=m+1}^{m+p} \lambda_n \frac{\alpha_n}{\lambda - \lambda_n} x_n \right\|^2 \\ &= \sum_{n=m+1}^{m+p} \left\langle \lambda_n \frac{\alpha_n}{\lambda - \lambda_n} x_n, \lambda_n \frac{\alpha_n}{\lambda - \lambda_n} x_n \right\rangle = \sum_{n=m+1}^{m+p} \left| \frac{\lambda_n}{\lambda - \lambda_n} \right|^2 |\alpha_n|^2 \\ &\leq \alpha^2 \sum_{n=m+1}^{m+p} |\alpha_n|^2. \end{aligned}$$

A série $\sum_{n=1}^{\infty} |\alpha_n|^2 = \sum_{n=1}^{\infty} |\langle y, x_n \rangle|^2 \leq \|y\|^2$ uma vez convergente, pela Desigualdade de Bessel (3.19), satisfaz a critério de Cauchy e portanto a série de (5.2) também o satisfaz.

4. A série de (5.2) que define x é convergente. De fato, como

$$\|v_m\|^2 = \sum_{n=1}^m \left| \frac{\alpha_n}{\lambda - \lambda_n} \right|^2 \leq \beta^2 \sum_{n=1}^m |\alpha_n|^2 \leq \beta^2 \|y\|^2,$$

tem-se que a sequência (v_m) é limitada em H e sendo T um operador compacto existe uma subsequência (v_{m_r}) tal que $T(v_{m_r}) = (u_{m_r})$ é convergente. Porém pelo item 3. desta demonstração a sequência (u_m) é de Cauchy e tendo ela uma subsequência convergente segue que a própria sequência (u_m) é convergente, completando a demonstração de que a série de (5.2) é convergente.

5. De $x = \frac{1}{\lambda} y + \frac{1}{\lambda} \sum_{n=1}^{\infty} \lambda_n \frac{\alpha_n}{\lambda - \lambda_n} x_n$, tem-se que

$$\begin{aligned} \|x\| &= \left\| \frac{1}{\lambda} y + \frac{1}{\lambda} \sum_{n=1}^{\infty} \lambda_n \frac{\alpha_n}{\lambda - \lambda_n} x_n \right\| \leq \left\| \frac{1}{\lambda} y \right\| + \left\| \frac{1}{\lambda} \sum_{n=1}^{\infty} \lambda_n \frac{\alpha_n}{\lambda - \lambda_n} x_n \right\| \\ &\leq \frac{1}{|\lambda|} \|y\| + \frac{1}{|\lambda|} \sum_{n=1}^{\infty} \left| \frac{\lambda_n}{\lambda - \lambda_n} \right| |\alpha_n| \\ &\leq \frac{1}{|\lambda|} \|y\| + \frac{1}{|\lambda|} \alpha \|y\|, \end{aligned}$$

onde, vale lembrar, que $\alpha = \sup_n \left| \frac{\lambda_n}{\lambda - \lambda_n} \right|$.

Logo $\|(\lambda - T)^{-1}y\| = \|x\| \leq \frac{1}{|\lambda|}\|y\| + \frac{1}{|\lambda|}\alpha\|y\| = \left(\frac{1}{|\lambda|} + \frac{1}{|\lambda|}\alpha \right) \|y\|$, isto é, $(\lambda - T)^{-1}$ é limitado e portanto contínuo pelo Teorema 4.1.

□

Teorema 5.4. *Sejam H um espaço pré-hilbertiano (real ou complexo) e $T : H \rightarrow H$ um operador hermitiano compacto, $T \neq 0$. Dado um autovalor $\lambda \neq 0$ de T , a equação*

$$\lambda x - Tx = y$$

admite uma solução se, e somente se, y é ortogonal a todo autovetor de T associado a λ . As soluções x da equação acima são da forma

$$x = \frac{1}{\lambda}y + \frac{1}{\lambda} \sum_{\lambda_n \neq \lambda} \lambda_n \frac{\alpha_n}{\lambda - \lambda_n} x_n + z \quad (5.3)$$

onde z é qualquer autovetor associado a λ , isto é, $Tz = \lambda z$.

Demonstração. Seja $\lambda x - Tx = y$, e como $\lambda = \bar{\lambda}$, então

$$\begin{aligned} \langle y, z \rangle &= \langle \lambda x - Tx, z \rangle = \langle \lambda x, z \rangle - \langle Tx, z \rangle \\ &= \lambda \langle x, z \rangle - \langle x, Tz \rangle \\ &= \langle x, \bar{\lambda} z \rangle - \langle x, \lambda z \rangle \\ &= \langle x, \lambda z \rangle - \langle x, \lambda z \rangle \\ &= 0. \end{aligned}$$

Portanto para todo z tal que $Tz = \lambda z$ tem-se $\langle y, z \rangle = 0$. Como no Teorema 5.3, demonstra-se que x é da forma (5.3) bastando lembrar que $\alpha_n = \langle y, x_n \rangle = 0$ se $\lambda_n = \lambda$.

Reciprocamente, se $\langle y, z \rangle = 0$ para todo z tal que $Tz = \lambda z$, é imediato que todo elemento x da forma (5.3) é solução da equação $\lambda x - Tx = y$.

□

Os Teoremas 5.3 e 5.4 dão lugar à **Alternativa de Fredholm** que será enunciada a seguir.

Sejam H um espaço pré-hilbertiano (real ou complexo), $T : H \rightarrow H$ um operador hermitiano compacto e $\lambda \in \mathbb{K}$ ($\mathbb{K} = \mathbb{R}$ ou \mathbb{C}), $\lambda \neq 0$. Vale a seguinte alternativa:

- i. ou a equação $\lambda x - Tx = y$ tem solução para todo $y \in H$; então esta solução é única e é dada por

$$x = \frac{1}{\lambda}y + \frac{1}{\lambda} \sum_{n=1}^{\infty} \lambda_n \frac{\alpha_n}{\lambda - \lambda_n} x_n,$$

onde $\alpha_n = \langle y, x_n \rangle$, e λ não é um autovalor de T .

- ii. ou a equação $\lambda x - Tx = 0$ tem solução não trivial; então o conjunto dessas soluções forma um espaço vetorial de dimensão finita e a equação $\lambda x - Tx = y$ tem solução se, e somente se, y for ortogonal a toda solução z da equação $\lambda z - Tz = 0$; então as soluções x são da forma

$$x = \frac{1}{\lambda}y + \frac{1}{\lambda} \sum_{\lambda_n \neq \lambda} \lambda_n \frac{\alpha_n}{\lambda - \lambda_n} x_n + z$$

e λ é um autovalor de T .

5.2 A Equação Integral de Fredholm com Núcleo Hermitiano

O objetivo desta seção é aplicar os resultados obtidos a partir do Teorema Espectral da seção anterior na resolução da equação integral de Fredholm de segunda espécie, dada por

$$(F) \quad \lambda x(t) = y(t) + \int_a^b K(t, s)x(s)ds$$

com núcleo hermitiano (isto é, $K(s, t) = \overline{K(t, s)}$) e contínuo.

Tais resultados aqui obtidos serão aplicados no Problema de Sturm-Liouville no próximo capítulo.

Para tanto, considere o espaço pré-hilbertiano $H = \mathcal{C}_{L_2}([a, b], \mathbb{C})$. Tem-se, no Exemplo 4.5 que o operador $k : f \in H \mapsto kf \in H$, onde $(kf)(t) = \int_a^b K(t, s)f(s)ds$ é compacto e no Exemplo 4.9 tem-se que k é hermitiano se, e somente se, o núcleo contínuo K o for.

Indique por $(\lambda_n)_{n \in \mathbb{N}}$ a sequência de autovalores não nulos de k , e por $(x_n)_{n \in \mathbb{N}}$ o sistema ortonormal correspondente de autofunções, cuja existência é garantida pelo Teorema Espectral (5.2). Assim, com as notações do Teorema Espectral, tem-se

$$kx_m = \sum_n \lambda_n \beta_n x_n = \sum_n \lambda_n \langle x_m, x_n \rangle x_n = \lambda_m x_m,$$

isto é,

$$(kx_n)(t) = \int_a^b K(t, s)x_n(s)ds = \lambda_n x_n(t), \quad t \in [a, b]. \quad (5.4)$$

Proposição 5.4. *Sejam $H = \mathcal{C}_{L_2}([a, b], \mathbb{C})$ o espaço pré-hilbertiano e $k : H \rightarrow H$ o operador definido por $(kf)(t) = \int_a^b K(t, s)f(s)ds$. Então*

$$\sum_{n=1}^{\infty} \lambda_n^2 \leq \int_a^b \int_a^b |K(t, s)|^2 ds dt,$$

onde λ_n são os autovalores não nulos de k .

Demonstração. Para cada $t \in [a, b]$ considere a aplicação $h : [a, b] \rightarrow \mathbb{C}$ dada por $h(s) = K(t, s)$. A Desigualdade de Bessel aplicada à função h e ao sistema ortonormal de funções $\overline{x_n}$ resulta em

$$\sum_{n=1}^{\infty} |\langle h, \overline{x_n} \rangle|^2 = \sum_{n=1}^{\infty} \left| \int_a^b h(s) x_n(s) ds \right|^2 = \sum_{n=1}^{\infty} \left| \int_a^b K(t, s) x_n(s) ds \right|^2 \leq \int_a^b |K(t, s)|^2 ds,$$

isto é,

$$\sum_{n=1}^{\infty} \lambda_n^2 |x_n(t)|^2 \leq \int_a^b |K(t, s)|^2 ds \quad (5.5)$$

donde segue-se que

$$\sum_{n=1}^{\infty} \lambda_n^2 \int_a^b |x_n(t)|^2 dt \leq \int_a^b \int_a^b |K(t, s)|^2 ds dt,$$

ou seja,

$$\sum_{n=1}^{\infty} \lambda_n^2 \leq \int_a^b \int_a^b |K(t, s)|^2 ds dt.$$

□

Teorema 5.5. Para todo $x \in \mathcal{C}([a, b], \mathbb{C})$ tem-se

$$y(t) = \int_a^b K(t, s) x(s) ds = \sum_{n=1}^{\infty} \lambda_n \alpha_n x_n(t),$$

com esta série sendo uniformemente e absolutamente convergente em $[a, b]$.

Demonstração. Do Teorema Espectral segue-se que para toda função $x \in \mathcal{C}_{L_2}([a, b], \mathbb{C})$ tem-se $y = kx = \sum_n \lambda_n \alpha_n x_n$, onde $\alpha_n = \langle x, x_n \rangle = \int_a^b x(s) \overline{x_n(s)} ds$, isto é,

$$y(t) = \int_a^b K(t, s) x(s) ds = \sum_{n=1}^{\infty} \lambda_n \alpha_n x_n(t), \quad t \in [a, b] \quad (5.6)$$

e esta série sendo convergente em $H = \mathcal{C}_{L_2}([a, b], \mathbb{C})$. Para todo $x \in \mathcal{C}([a, b], \mathbb{C})$ a série (5.6) é absolutamente e uniformemente convergente no intervalo $[a, b]$. De fato, aplicando primeiro a Desigualdade de Hölder, depois a Desigualdade de Bessel e, por fim, de (5.5) tem-se

$$\begin{aligned} \sum_{n=1}^{\infty} |\lambda_n \alpha_n x_n(t)| &\leq \left[\sum_{n=1}^{\infty} |\alpha_n|^2 \right]^{\frac{1}{2}} \left[\sum_{n=1}^{\infty} |\lambda_n x_n(t)|^2 \right]^{\frac{1}{2}} \\ &\leq \int_a^b [|K(t, s)|^2 ds]^{\frac{1}{2}} \|x\| = (b-a)^{\frac{1}{2}} M \|x\|, \end{aligned}$$

onde $M = \sup_{a \leq s, t \leq b} |K(t, s)|$.

□

Teorema 5.6. *Sejam $K : [a, b] \times [a, b] \rightarrow \mathbb{C}$ um núcleo hermitiano contínuo, (λ_n) a sequência dos autovalores não nulos k e (x_n) a sequência ortonormal das autofunções correspondentes. Para todo $\lambda \neq 0$, $\lambda \neq \lambda_n$ a equação integral*

$$\lambda x(t) = y(t) + \int_a^b K(t, s)x(s)ds$$

tem uma única solução dada por

$$x(t) = \frac{1}{\lambda}y(t) + \frac{1}{\lambda} \sum_{n=1}^{\infty} \lambda_n \frac{\alpha_n}{\lambda - \lambda_n} x_n(t),$$

onde $\alpha_n = \langle y, x_n \rangle$ e com esta série sendo uniformemente e absolutamente convergente em $[a, b]$.

Demonstração. O Teorema 5.3 garante que dado $\lambda \neq 0$ tal que $\lambda \neq \lambda_n$ para todo n então dado $y \in H = \mathcal{C}_{L_2}([a, b])$ a equação integral (F), isto é, $y = (\lambda - k)^{-1}x$, tem uma e somente uma solução $x = (\lambda - k)^{-1}y \in H$ dada por

$$x = \frac{1}{\lambda}y + \frac{1}{\lambda} \sum_{n=1}^{\infty} \lambda_n \frac{\alpha_n}{\lambda - \lambda_n} x_n,$$

onde $y_n = \langle y, x_n \rangle = \int_a^b y(s)\overline{x_n(s)}ds$, isto é

$$x(t) = \frac{1}{\lambda}y(t) + \frac{1}{\lambda} \sum_{n=1}^{\infty} \lambda_n \frac{\alpha_n}{\lambda - \lambda_n} x_n(t), \quad (t \in [a, b]) \quad (5.7)$$

com esta série sendo convergente em $H = \mathcal{C}_{L_2}([a, b])$. Para todo $y \in H = \mathcal{C}_{L_2}([a, b], \mathbb{C})$ a série (5.7) é absolutamente e uniformemente convergente no intervalo $[a, b]$. De fato, se x satisfaz $x = \frac{1}{\lambda}y + \frac{1}{\lambda}kx$ que comparada com (5.7) mostra que

$$(kx)(t) = \int_a^b K(t, s)x(s)ds = \sum_{n=1}^{\infty} \lambda_n \frac{\alpha_n}{\lambda - \lambda_n} x_n(t),$$

e do Teorema 5.5 segue-se que esta série é absolutamente e uniformemente convergente. \square

Observação 5.2. Quando os x_n formar um sistema ortonormal completo de $\mathcal{C}_{L_2}([a, b], \mathbb{C})$

então substituindo $y(t) = \sum_{n=1}^{\infty} \alpha_n x_n(t)$ em (5.7) vem

$$x(t) = \sum_{n=1}^{\infty} \frac{\alpha_n}{\lambda_n - \lambda} x_n,$$

onde vale lembrar que $\alpha_n = \langle y, x_n \rangle$, com esta série sendo apenas convergente em $\mathcal{C}_{L_2}([a, b])$ e não uniformemente absolutamente (se o desenvolvimento ortonormal de y não o for).

Analogamente, segue do Teorema 5.4 o seguinte resultado.

Teorema 5.7. *Dado um núcleo hermitiano contínuo K e um autovalor $\lambda \neq 0$ do operador compacto k associado a K , então a equação integral*

$$\lambda x(t) = y(t) + \int_a^b K(t, s)x(s)ds$$

admite solução se, e somente se,

$$\int_a^b y(t)\overline{z(t)}dt = 0,$$

para toda função contínua $z \in \mathcal{C}([a, b], \mathbb{C})$ tal que

$$\int_a^b K(t, s)z(s)ds = \lambda z(t). \quad (5.8)$$

Ainda, as soluções são da forma

$$x(t) = \frac{1}{\lambda}y(t) + \frac{1}{\lambda} \sum_{\lambda_n \neq \lambda} \lambda_n \frac{\alpha_n}{\lambda - \lambda_n} x_n(t) + z(t), \quad (5.9)$$

onde z é um elemento de $\mathcal{C}([a, b], \mathbb{C})$ que satisfaz (5.8) e $\alpha_n = \langle y, x_n \rangle = \int_a^b y(s)\overline{x_n(s)}ds$.

A série (5.9) é absolutamente e uniformemente convergente.

Proposição 5.5. *A série*

$$\sum_{n=1}^{\infty} \lambda_n^2 |x_n(t)|^2$$

é absolutamente e uniformemente convergente em $[a, b]$.

Demonstração. Considere

$$K_2(t, s) = \int_a^b K(t, u)K(u, s)du.$$

Para cada $t \in [a, b]$ fixado, aplicando o Teorema 5.5 a $K_2(t, s)$ considerado como uma função de $s \in [a, b]$, tem-se

$$\begin{aligned} K_2(t, s) &= \int_a^b K(t, u)K(u, s)du = \sum_{n=1}^{\infty} \lambda_n \langle K, x_n \rangle x_n(t) \\ &= \sum_{n=1}^{\infty} \lambda_n \left[\int_a^b K(u, s)\overline{x_n(u)}du \right] x_n(t) \end{aligned}$$

De (5.4), tem-se que $\int_a^b K(u, s)\overline{x_n(u)}du = \lambda_n \overline{x_n(s)}$, assim

$$K_2(t, s) = \sum_{n=1}^{\infty} \lambda_n \left[\int_a^b K(u, s)\overline{x_n(u)}du \right] x_n(t) = \sum_{n=1}^{\infty} \lambda_n^2 x_n(t)\overline{x_n(s)},$$

com esta série sendo absolutamente e uniformemente convergente para $s \in [a, b]$ e t fixado. Tem-se em particular

$$K_2(t, t) = \sum_{n=1}^{\infty} \lambda_n x_n(t) \overline{x_n(t)} = \sum_{n=1}^{\infty} \lambda_n^2 |x_n(t)|^2$$

e K_2 sendo contínua no compacto segue-se do Teorema de Dini que esta série é uniformemente convergente em $[a, b]$. \square

Observação 5.3. A Proposição 5.5 mostra que

$$\lim_{n \rightarrow \infty} \left[K_2(t, t) - \sum_{i=1}^n \lambda_n^2 x_i(t) \overline{x_i(t)} \right] = 0 \tag{5.10}$$

uniformemente para $t \in [a, b]$; mas

$$\int_a^b |K(t, s) - \sum_{i=1}^n \lambda_n x_i(t) \overline{x_i(s)}|^2 ds = \dots = K_2(t, t) - \sum_{i=1}^n \lambda_i^2 |x_i(t)|^2 \tag{5.11}$$

mostra que quanto $n \rightarrow \infty$,

$$\int_a^b |K(t, s) - \sum_{i=1}^n \lambda_n x_i(t) \overline{x_i(s)}|^2 dt$$

tende para zero uniformemente para $s \in [a, b]$.

Portanto,

$$\lim_{n \rightarrow \infty} \int_a^b \left[\int_a^b |K(t, s) - \sum_{i=1}^n x_i(t) \overline{x_i(s)}|^2 ds \right] dt = 0.$$

O teorema que segue faz uso do Teorema de Dini que será apenas enunciado, mas sua demonstração se encontra em [8].

Teorema 5.8 (Teorema de Dini). *Seja X um espaço compacto e $f_n \in \mathcal{C}(X, \mathbb{R})$ uma sequência crescente ou decrescente de funções contínuas que converge para uma função contínua f ; então esta convergência é uniforme.*

Teorema 5.9. (Teorema de Mercer) *Se o núcleo hermitiano contínuo K é tal que o operador k definido por ele é positivo, isto é, $\langle kx, x \rangle \geq 0$ para todo x , então*

$$K(t, s) = \sum_{n=1}^{\infty} \lambda_n x_n(t) \overline{x_n(s)},$$

com esta série sendo absolutamente e uniformemente convergente em $[a, b] \times [a, b]$.

Demonstração. A demonstração se dará por partes.

- a) Para todo $t \in [a, b]$ tem-se $K(t, t) \geq 0$. De fato, seja $t_0 \in [a, b]$ tal que $K(t_0, t_0) < -\delta < 0$. Sendo K contínua, pelo teorema da conservação do sinal, existe uma vizinhança $V \times V$ de (t_0, t_0) tal que para todo (t, s) em $V \times V$ tem-se $\operatorname{Re}K(t, s) \leq -\delta$. Seja $x \in \mathcal{C}([a, b], [0, 1])$ tal que $x(t_0) = 1$ e $x(t) = 0$ para $t \notin V$, como k é hermitiano, tem-se

$$\begin{aligned} \langle kx, x \rangle &= \operatorname{Re}\langle kx, x \rangle \\ &= \int_a^b \left[\int_a^b \operatorname{Re}K(t, s)x(s)ds \right] x(t)dt \\ &\leq -\delta \int_a^b \left[\int_a^b x(s)ds \right] x(t)dt \\ &= -\delta(\|x\|)^2 < 0 \end{aligned}$$

contra o fato de k ser positivo.

- b) Dado um número finito $\lambda_1, \dots, \lambda_n$ de autovalores de k então

$$K_n(t, s) = K(t, s) - \sum_{i=1}^n \lambda_i x_i(t) \overline{x_i(s)}$$

é o núcleo de um operador positivo k_n . De fato,

$$\begin{aligned} \langle k_n x, x \rangle &= \int_a^b (k_n x)(t) \overline{x(t)} dt \\ &= \int_a^b \left[\int_a^b K_n(t, s)x(s)ds \right] \overline{x(t)} dt \\ &= \int_a^b \left[\int_a^b \left(K(t, s) - \sum_{i=1}^n \lambda_i x_i(t) \overline{x_i(s)} \right) x(s)ds \right] \overline{x(t)} dt \\ &= \int_a^b \left[\int_a^b K(t, s)x(s)ds - \sum_{i=1}^n \lambda_i \int_a^b x_i(t) \overline{x_i(s)} x(s)ds \right] \overline{x(t)} dt \\ &= \int_a^b (kx)(t) \overline{x(t)} dt - \int_a^b \left[\sum_{i=1}^n \lambda_i \int_a^b x_i(t) \overline{x_i(s)} x(s)ds \right] \overline{x(t)} dt \\ &= \langle kx, x \rangle - \sum_{i=1}^n \lambda_i \int_a^b \left[\int_a^b x_i(t) \overline{x_i(s)} x(s)ds \right] \overline{x(t)} dt \\ &= \langle kx, x \rangle - \sum_{i=1}^n \lambda_i \int_a^b \left[x_i(t) \int_a^b \overline{x_i(s)} x(s)ds \right] \overline{x(t)} dt \\ &= \langle kx, x \rangle - \sum_{i=1}^n \lambda_i \int_a^b x_i(t) \langle x, x_i \rangle \overline{x(t)} dt \\ &= \langle kx, x \rangle - \sum_{i=1}^n \lambda_i \langle x, x_i \rangle \langle x_i, x \rangle \\ &= \langle kx, x \rangle - \sum_{i=1}^n \lambda_i |\alpha_i|^2. \end{aligned}$$

Note que

$$\begin{aligned}
\langle k(x - \sum_{i=1}^n \alpha_i x_i), x - \sum_{i=1}^n \alpha_i x_i \rangle &= \langle kx - \sum_{i=1}^n \alpha_i kx_i, x - \sum_{i=1}^n \alpha_i x_i \rangle \\
&= \langle kx, x \rangle - \langle kx, \sum_{i=1}^n \alpha_i x_i \rangle - \\
&\quad - \langle \sum_{i=1}^n \alpha_i \lambda_i x_i, x \rangle + \langle \sum_{i=1}^n \alpha_i \lambda_i x_i, \sum_{i=1}^n \alpha_i x_i \rangle \\
&= \langle kx, x \rangle - \langle kx, \sum_{i=1}^n \alpha_i x_i \rangle - \langle \sum_{i=1}^n \alpha_i \lambda_i x_i, x \rangle \\
&\quad + \sum_{i=1}^n \lambda_i |\alpha_i|^2 \\
&= \langle kx, x \rangle - \langle kx, \sum_{i=1}^n \alpha_i x_i \rangle - \sum_{i=1}^n \lambda_i |\alpha_i|^2 + \sum_{i=1}^n \lambda_i |\alpha_i|^2 \\
&= \langle kx, x \rangle - \langle kx, \sum_{i=1}^n \alpha_i x_i \rangle = \langle kx, x \rangle - \sum_{i=1}^n \lambda_i |\alpha_i|^2.
\end{aligned}$$

Portanto

$$\begin{aligned}
\langle k_n x, x \rangle &= \langle kx, x \rangle - \sum_{i=1}^n \lambda_i |\alpha_i|^2 \\
&= \langle k(x - \sum_{i=1}^n \alpha_i x_i), x - \sum_{i=1}^n \alpha_i x_i \rangle \geq 0.
\end{aligned}$$

c) Da parte a) e b) segue-se que $K_n(s, s) \geq 0$ e fazendo $t = s$ em b) segue-se que

$$\sum_{n=1}^{\infty} \lambda_n |x_n(t)|^2 \leq K(t, t)$$

para todo $t \in [a, b]$. Da Desigualdade de Cauchy-Schwarz segue-se portanto que

$$\begin{aligned}
\sum_{n=p}^q \lambda_n |x_n(t)x_n(s)| &\leq \left[\sum_{n=p}^q \lambda_n |x_n(t)|^2 \right]^{\frac{1}{2}} \cdot \left[\sum_{n=p}^q \lambda_n |x_n(s)|^2 \right]^{\frac{1}{2}} \\
&\leq \left[\sum_{n=p}^q \lambda_n |x_n(t)|^2 \right]^{\frac{1}{2}} \cdot K(s, s)^{\frac{1}{2}} \quad (5.12)
\end{aligned}$$

Da Proposição 5.5, a série (5.12) é absolutamente e uniformemente convergente em $[a, b]$, assim para cada $t \in [a, b]$ fixado, dado $\epsilon > 0$ existe p tal que para todo $q \geq p$ tem-se

$$\left[\sum_{n=p}^q \lambda_n |x_n(t)|^2 \right]^{\frac{1}{2}} < \epsilon,$$

e como $K(s, s)$ é limitada em $[a, b]$ segue que a série

$$\sum_{n=1}^{\infty} \lambda_n x_n(t) \overline{x_n(s)}$$

é absolutamente e uniformemente convergente para $s \in [a, b]$. Da observação 5.3 segue-se que

$$\sum_{n=1}^{\infty} \lambda_n x_n(t) \overline{x_n(s)} = K(s, t)$$

para todo $(t, s) \in [a, b] \times [a, b]$ uma vez que a função

$$s \mapsto \left| K(t, s) - \sum_{n=1}^{\infty} \lambda_n x_n(t) \overline{x_n(s)} \right|$$

é contínua em $[a, b]$ e sua integral é nula. Em particular tem-se

$$K(t, t) = \sum_{n=1}^{\infty} \lambda_n |x_n(t)|^2 \tag{5.13}$$

e do Teorema de Dini segue-se que a série em (5.13) é uniformemente convergente em $[a, b]$. Por fim, de (5.12) a série

$$\sum_{n=1}^{\infty} \lambda_n x_n(t) \overline{x_n(s)}$$

é absolutamente e uniformemente convergente em $[a, b] \times [a, b]$.

□

Como consequência do Teorema de Mercer tem-se o seguinte corolário.

Corolário 5.2. *Nas condições do Teorema de Mercer, tem-se*

$$\sum_{n=1}^{\infty} \lambda_n = \int_a^b K(t, t) dt.$$

6 O Problema de Sturm-Liouville

Na teoria das equações diferenciais o Problema de Sturm-Liouville ¹ aparece de modo natural quando se busca resolver equações diferenciais parciais a partir do método de separação de variáveis. Este capítulo busca fornecer uma solução para o problema a partir de certas condições iniciais e para este desenvolvimento são utilizadas as referências [7], [8] e [9].

Como ponto de partida, considere o operador diferencial linear de segunda ordem

$$L_\lambda[y] \equiv -(p(t)y')' + [q(t) - \lambda\rho(t)]y \quad (6.1)$$

no intervalo $[a, b]$ onde $p \in \mathcal{C}^1([a, b], \mathbb{R})$ com $p(t) > 0$ para $t \in [a, b]$, $\rho \in \mathcal{C}([a, b], \mathbb{R})$ com $\rho(t) > 0$ para $t \in [a, b]$ e $q \in \mathcal{C}([a, b], \mathbb{R})$, e condições de fronteira

$$F_1[y] \equiv \alpha_0 y(a) + \alpha_1 y'(a)$$

$$F_2[y] \equiv \beta_0 y(b) + \beta_1 y'(b)$$

onde $\alpha_0, \alpha_1, \beta_0, \beta_1 \in \mathbb{R}$ com

$$|\alpha_0| + |\alpha_1| \neq 0 \text{ e } |\beta_0| + |\beta_1| \neq 0.$$

O **Problema de Sturm-Liouville** consiste em achar uma função y que seja solução do sistema

$$\begin{cases} (S_\lambda) & -(p(t)y')' + [q(t) - \lambda\rho(t)]y = f(t) \\ (F) & F_1[y] = 0, F_2[y] = 0 \end{cases},$$

para $t \in [a, b]$.

Os exemplos mais comuns de condições de fronteira são

$$y(a) = y(b) = 0 \text{ e } y'(a) = y'(b) = 0.$$

Definição 6.1. Diz-se que λ é um **autovalor** do sistema $(S_\lambda) + (F)$ ou do problema de Sturm-Liouville se a equação homogênea

¹Nome dado em homenagem aos matemáticos Jacques Charles François Sturm (1803-1855) e Joseph Liouville (1809-1882).

$$(S^* \lambda) \quad L_\lambda[y] = 0,$$

isto é, $L_0[y] = \lambda \rho(t)y$, ou ainda

$$-(p(t)y')' + q(t)y = \lambda \rho(t)y$$

tem solução $y \neq 0$ que satisfaz as condições de fronteira (F). A solução y se chama **autofunção** (correspondente ao autovalor λ).

Exemplo 6.1. Considere o seguinte problema de Sturm-Liouville

$$\begin{aligned} y''(t) + \lambda y(t) &= 0, \\ y'(0) = y'(\pi) &= 0. \end{aligned}$$

Para $\lambda \leq 0$ a equação característica $r^2 + \lambda = 0$, da equação $y''(t) + \lambda y(t) = 0$, possui raízes reais. Assim o problema só admite a solução trivial ($y(t) \equiv 0$). Já, para $\lambda > 0$, tem-se que a equação característica $r^2 + \lambda = 0$, da equação $y''(t) + \lambda y(t) = 0$, possui raízes complexas. Logo a solução geral da equação $y''(t) + \lambda y(t) = 0$ é da forma $y(t) = c_1 e^{\alpha t} \cos(\mu t) + c_2 e^{\alpha t} \sen(\mu t)$, onde c_1 e c_2 são constantes arbitrárias, e α e μ são as partes real e imaginária das raízes da equação característica. Assim, a solução do problema fica da forma

$$y(t) = c_1 \cos(\sqrt{\lambda} t) + c_2 \sen(\sqrt{\lambda} t).$$

Das condições de fronteira $y'(0) = y'(\pi) = 0$, segue que

$$\begin{aligned} c_2 \sqrt{\lambda} &= 0, \\ c_2 \sqrt{\lambda} \cos(\sqrt{\lambda} \pi) - c_1 \lambda \sen(\sqrt{\lambda} \pi) &= 0. \end{aligned}$$

Logo $c_2 = 0$ e $c_1 \sen(\sqrt{\lambda} \pi) = 0$. Se $c_1 = 0$, segue que $y(t) = 0$. Para $c_1 \neq 0$, tem-se $\sqrt{\lambda} \pi = k\pi$, $k \in \{1, 2, \dots\}$. Assim $\lambda_k = k^2$ são os autovalores e $y_k(t) = \cos(kt)$ são as autofunções do problema.

Considere agora alguns resultados importantes sobre o problema de Sturm-Liouville.

Proposição 6.1. Considere o operador

$$L[y] = -(p(t)y')' + q(t)y,$$

definido em $[a, b]$ com $q \in \mathcal{C}([a, b], \mathbb{R})$ e sejam dadas duas funções $g, f \in \mathcal{C}^{(2)}([a, b], \mathbb{C})$. Então vale a Identidade de Lagrange:

$$\int_a^b (\overline{f(t)} L[g] - g(t) \overline{L[f]}) dt = M[g, f](b) - M[g, f](a),$$

onde $M[g, f](t) = -p(t)[g'(t)\overline{f(t)} - g(t)\overline{f'(t)}]$.

Demonstração. Sendo $q(t) = \overline{q(t)}$. Então

$$\begin{aligned}
\int_a^b [\bar{f}L[g] - g\overline{L[f]}]dt &= \int_a^b [\bar{f}(-(pg)') + qg) - g(\overline{-(pf)'} + qf)]dt \\
&= \int_a^b [-\bar{f}(pg)') + \bar{f}qg + g\overline{(pf)'} - gq\bar{f}]dt \\
&= \int_a^b [-\bar{f}(pg)') + g\overline{(pf)'}]dt \\
&= -\int_a^b \bar{f}(pg)')dt + \int_a^b g\overline{(pf)'}dt \\
&= -\left[\bar{f}(pg)'\Big|_a^b - \int_a^b \bar{f}'(pg)')dt\right] + \left[g\overline{(pf)'}\Big|_a^b - \int_a^b g'(\overline{(pf)'}))dt\right] \\
&= -(\bar{f}(pg)') - g(\overline{(pf)'}))\Big|_a^b + \int_a^b [\bar{f}'(pg)') - g'(\overline{(pf)'})]dt \\
&= -p(b)[g'(b)\overline{f(b)} - g(b)\overline{f'(b)}] + p(a)[g'(a)\overline{f(a)} - g(a)\overline{f'(a)}] \\
&= M[g, f](b) - M[g, f](a).
\end{aligned}$$

□

Proposição 6.2. *Sejam y_1 e y_2 soluções de $(S_\lambda^*) + F$. Então $M[y_1, y_2](a) = M[y_1, y_2](b)$ e portanto*

$$\int_a^b (\bar{y}_1 L_0[y_2] - y_2 \overline{L_0[y_1]})dt = 0.$$

Demonstração. Tem-se $M[y_1, y_2](a) = -p(a)[y_1'(a)\overline{y_2(a)} - y_1(a)\overline{y_2'(a)}]$, das condições de fronteira segue que:

$$\begin{cases} \alpha_0 y_1(a) + \alpha_1 y_1'(a) = 0 \\ \alpha_0 y_2(a) + \alpha_1 y_2'(a) = 0 \end{cases} \sim \begin{cases} \alpha_0 \overline{y_1(a)} + \alpha_1 \overline{y_1'(a)} = 0 \\ \alpha_0 \overline{y_2(a)} + \alpha_1 \overline{y_2'(a)} = 0 \end{cases}.$$

Uma vez que $|\alpha_0| + |\alpha_1| \neq 0$ tem-se

$$\begin{vmatrix} y_1(a) & y_1'(a) \\ \overline{y_2(a)} & \overline{y_2'(a)} \end{vmatrix} = 0,$$

ou seja, $y_1(a)\overline{y_2'(a)} - y_1'(a)\overline{y_2(a)} = 0$ e portanto $M[y_1, y_2](a) = 0$. De modo análogo tem-se $M[y_1, y_2](b) = 0$. Assim $M[y_1, y_2](a) = M[y_1, y_2](b) = 0$.

A segunda parte da demonstração segue de modo imediato da Proposição 6.1 e da primeira parte. Logo

$$\int_a^b (\bar{y}_1 L_0[y_2] - y_2 \overline{L_0[y_1]})dt = M[y_1, y_2](b) - M[y_1, y_2](a) = 0.$$

□

Proposição 6.3. *Todos os autovalores do problema de Sturm-Liouville são reais.*

Demonstração. Seja λ um autovalor do problema de Sturm-Liouville. Logo existe uma função $y \neq 0$ tal que $L_0[y] = \lambda\rho(t)y(t)$. Note que

$$\overline{L_0[y]} = \overline{-(p(t)y'(t))' + q(t)y(t)} = -(p(t)\overline{y'(t)})' + q(t)\overline{y(t)} = L_0[\overline{y}],$$

assim $L_0[\overline{y}] = \bar{\lambda}\rho(t)\overline{y}(t)$.

Da Proposição 6.2, tem-se

$$\begin{aligned} 0 &= \int_a^b (\overline{y}L_0[y] - y\overline{L_0[\overline{y}]})dt = \int_a^b (\overline{y}L_0[y] - yL_0[\overline{y}])dt \\ &= \int_a^b (\overline{y(t)}\lambda\rho(t)y(t) - y(t)\bar{\lambda}\rho(t)\overline{y(t)})dt = \int_a^b (\lambda - \bar{\lambda})\rho(t)\overline{y(t)}y(t)dt \\ &= \int_a^b (\lambda - \bar{\lambda})\rho(t)|y(t)|^2dt = (\lambda - \bar{\lambda}) \int_a^b \rho(t)|y(t)|^2dt. \end{aligned}$$

Como $y \neq 0$, segue que $\lambda = \bar{\lambda}$, isto é, λ é real. □

Proposição 6.4. *Sejam y_1 e y_2 autofunções do problema de Sturm-Liouville correspondentes a autovalores distintos λ_1 e λ_2 respectivamente. Então y_1 e y_2 são ortogonais relativamente a ρ , isto é, satisfazem*

$$\langle y_1, y_2 \rangle_\rho = \int_a^b y_1(t)\overline{y_2(t)}\rho(t)dt = 0.$$

Demonstração. Das Proposições 6.2 e 6.3, tem-se

$$\begin{aligned} 0 &= \int_a^b (\overline{y_2}L_0[y_1] - y_1\overline{L_0[y_2]})dt \\ &= \int_a^b (\overline{y_2(t)}\lambda_1\rho(t)y_1(t) - y_1(t)\overline{\lambda_2}\rho(t)\overline{y_2(t)})dt \\ &= \int_a^b ((\lambda_1 - \lambda_2)\overline{y_2(t)}\rho(t)y_1(t))dt \\ &= (\lambda_1 - \lambda_2) \int_a^b \overline{y_2(t)}\rho(t)y_1(t)dt. \end{aligned}$$

Uma vez que $\lambda_1 \neq \lambda_2$, então $\int_a^b \rho(t)y_1(t)\overline{y_2(t)}dt = 0$. □

Proposição 6.5. *Toda autofunção do problema de Sturm-Liouville é uma combinação linear de autofunções reais, correspondentes ao mesmo autovalor.*

Demonstração. Seja λ um autovalor do problema. Da Proposição 6.3, tem-se

$$L_0[\overline{y}] = \bar{\lambda}\rho\overline{y} = \lambda\rho\overline{y},$$

onde y é uma autofunção correspondente a λ . Pode-se escrever

$$y(t) = \frac{1}{2}(y(t) + \overline{y(t)}) + \frac{i}{2}(i\overline{y(t)} - iy(t)).$$

Fazendo $y_1(t) = \frac{y(t) + \overline{y(t)}}{2}$ e $y_2(t) = \frac{i(\overline{y(t)} - y(t))}{2}$, tem-se

$$y(t) = y_1(t) + iy_2(t).$$

Assim,

$$\begin{aligned} \lambda\rho(t)y_1(t) &= \lambda\rho(t) \left(\frac{y(t) + \overline{y(t)}}{2} \right) \\ &= \frac{1}{2}\lambda\rho(t)y(t) + \frac{1}{2}\lambda\rho(t)\overline{y(t)} \\ &= \frac{1}{2}L_0[y] + \frac{1}{2}L_0[\overline{y}] \\ &= \frac{1}{2}L_0[y + \overline{y}] \\ &= L_0[y_1], \end{aligned}$$

e, de modo análogo, tem-se também

$$\lambda\rho(t)y_2(t) = L_0[y_2].$$

Portanto as funções y_1 e y_2 são autofunções reais. \square

O resultado do teorema precedente permite considerar apenas as autofunções do problema de Sturm-Liouville como sendo funções reais.

Definição 6.2. Diz-se que f_1, \dots, f_n são funções **linearmente dependentes** em um intervalo, se existirem constantes k_1, \dots, k_n não todas nulas, tais que

$$k_1f_1(t) + \dots + k_nf_n(t) = 0 \tag{6.2}$$

para todo t no intervalo $a[a, b]$. As funções f_1, \dots, f_n são **linearmente independentes** no intervalo se não forem linearmente dependentes, isto é, a equação (6.2) só vale para todo t no intervalo se $k_1 = \dots = k_n = 0$.

Considere, agora, o operador diferencial linear de ordem n

$$L[y] \equiv a_0(t)y^{(n)} + a_1(t)y^{(n-1)} + \dots + a_n(t)y,$$

onde $a_i \in \mathcal{C}([a, b], \mathbb{R})$, $i \in \{1, 2, \dots, n\}$ e $a_0(t) \neq 0$ para todo $t \in [a, b]$.

A seguir alguns resultados da teoria das equações diferenciais lineares.

Teorema 6.1. O conjunto das funções $y \in \mathcal{C}^{(n)}([a, b], \mathbb{R})$ que são soluções da equação homogênea $L[y] = 0$ formam um espaço vetorial E_0 de dimensão n .

Teorema 6.2. $y_1, \dots, y_n \in E_0$ formam uma base de E_0 se, e somente se, o seu determinante *wronskiano*, dado por

$$W(t) = W[y_1, \dots, y_n](t) = \begin{vmatrix} y_1(t) & \cdots & y_n(t) \\ y_1'(t) & \cdots & y_n'(t) \\ \vdots & \ddots & \vdots \\ y_1^{(n-1)}(t) & \cdots & y_n^{(n-1)}(t) \end{vmatrix},$$

for não nulo em todo ponto $t \in [a, b]$.

O teorema seguinte relaciona a independência e a dependência linear ao Wronskiano.

Teorema 6.3. Se f e g forem funções deriváveis em um intervalo aberto I , e se $W[f, g](t_0) \neq 0$ em um ponto t_0 de I , então f e g são linearmente independentes sobre I . De outra forma, se f e g forem linearmente dependentes sobre I , então $W[f, g](t_0) = 0$ para todo x em I .

Demonstração. Primeiro, considere a combinação linear $k_1 f(x) + k_2 g(x)$, e suponha que esta expressão seja nula sobre todo o intervalo. Calculando a expressão, e a sua derivada, num ponto t_0 , temos

$$\begin{aligned} k_1 f(x_0) + k_2 g(x_0) &= 0 \\ k_1 f'(x_0) + k_2 g'(x_0) &= 0 \end{aligned} \tag{6.3}$$

O determinante formado pelos coeficientes das equações (6.3) é

$$\begin{vmatrix} f(x_0) & g(x_0) \\ f'(x_0) & g'(x_0) \end{vmatrix} = f(x_0)g'(x_0) - f'(x_0)g(x_0) = W[f, g](x_0),$$

que não é, por hipótese, igual a zero. Então a única solução das equações (6.3) é $k_1 = k_2 = 0$ e portanto f e g são linearmente independentes.

A segunda parte do teorema é consequência imediata da primeira. Suponha que f e g sejam linearmente dependentes e que a conclusão seja falsa, isto é, $W[f, g]$ não é nulo em todo ponto de I . Então, há um ponto t_0 tal que $W[f, g](t_0) \neq 0$; pela primeira parte do teorema isto determina serem f e g linearmente independentes, o que é uma contradição. \square

Teorema 6.4 (Teorema de Abel). Se y_1 e y_2 forem soluções da equação diferencial

$$L[y] = y'' + p(t)y' + q(t)y = 0, \tag{6.4}$$

onde p e q são contínuas no intervalo aberto I , então o Wronskiano $W[y_1, y_2](t)$ é dado por

$$W[y_1, y_2](t) = c \exp\left[-\int p(t)dt\right],$$

onde c é uma constante que depende de y_1 e de y_2 , mas não de t . Além disso, $W[y_1, y_2](t)$ ou é nulo para todo t em I (se $c = 0$) ou nunca é nulo em I (se $c \neq 0$).

Demonstração. Sendo y_1 e y_2 soluções de (6.4), tem-se

$$\begin{aligned}y_1'' + p(t)y_1' + q(t)y_1 &= 0 \\y_2'' + p(t)y_2' + q(t)y_2 &= 0\end{aligned}$$

Multiplicando a primeira linha por $-y_2$, a segunda por y_1 , segue que

$$\begin{aligned}-y_1''y_2 - p(t)y_1'y_2 - q(t)y_1y_2 &= 0 \\y_2''y_1 + p(t)y_2'y_1 + q(t)y_2y_1 &= 0,\end{aligned}$$

e somando as equações, resulta

$$-y_1'y_2 + y_2''y_1 - p(t)y_1'y_2 + p(t)y_2'y_1 = (y_1y_2'' - y_1''y_2) + p(t)(y_1y_2' - y_1'y_2) = 0 \quad (6.5)$$

Fazendo $W(t) = W[y_1, y_2](t)$ tem-se

$$W = y_1y_2' - y_1'y_2, \quad W' = y_1y_2'' - y_1''y_2.$$

Logo a equação (6.5) pode ser escrita na forma

$$W' + pW = 0. \quad (6.6)$$

Note que a equação (6.6) é uma equação diferencial linear de primeira ordem com variáveis separáveis, logo ela pode ser resolvida com facilidade. Então

$$\begin{aligned}\frac{W'(t)}{W(t)} = -p(t) &\Rightarrow W'(t) \frac{1}{W(t)} = -p(t) \\ \frac{dW}{dt} \frac{1}{W(t)} = -p(t) &\Rightarrow \frac{dW}{W(t)} = -p(t)dt,\end{aligned}$$

integrando ambos os membros, tem-se

$$\int \frac{dW}{W(t)} dt = \int -p(t)dt \Rightarrow \ln W(t) = - \int p(t)dt + k,$$

e, por fim, aplicando a exponencial em ambos os membros, segue que

$$e^{\ln W(t)} = \exp\left[- \int p(t)dt\right]e^k \Rightarrow e^{\ln W(t)} = c \exp\left[- \int p(t)dt\right],$$

onde $c = e^k$ é uma constante. O valor de c depende do par de soluções da equação (6.4). Porém, como a função exponencial nunca é nula, $W(t)$ não é nulo a menos que $c = 0$, e neste caso $W(t)$ é nulo para todo t , o que completa a prova. \square

De modo geral tem-se, também, o seguinte teorema.

Teorema 6.5. *Considere a equação diferencial homogênea*

$$L[y] \equiv a_0(t)y^{(n)} + a_1(t)y^{(n-1)} + \dots + a_n(t)y = 0.$$

Fixado um ponto $t_0 \in [a, b]$, tem-se

$$W(t) = W(t_0) \exp\left[- \int_{t_0}^t \frac{a_1(s)}{a_0(s)} ds\right], \quad \forall t \in [a, b].$$

Demonstração. A demonstração geral é idêntica, exceto que ela utiliza a fórmula para a derivada de um determinante de ordem n . Ou seja, se f_{ij} está em $\mathcal{C}(I)$, $1 \leq i, j \leq n$, então a função F definida por

$$F(x) = \begin{vmatrix} f_{11}(x) & f_{12}(x) & \cdots & f_{1n}(x) \\ f_{21}(x) & f_{22}(x) & \cdots & f_{2n}(x) \\ \vdots & & \ddots & \vdots \\ f_{n1}(x) & f_{n2}(x) & \cdots & f_{nn}(x) \end{vmatrix}$$

é tal que

$$\frac{d}{dx} \begin{vmatrix} f_{11}(x) & f_{12}(x) & \cdots & f_{1n}(x) \\ f_{21}(x) & f_{22}(x) & \cdots & f_{2n}(x) \\ \vdots & & \ddots & \vdots \\ f_{n1}(x) & f_{n2}(x) & \cdots & f_{nn}(x) \end{vmatrix}$$

pode ser expressa como soma de n determinantes, o i -ésimo dos quais se obtém de $|f_{ij}(x)|$, derivando-se as funções da i -ésima coluna. Mais detalhes podem ser encontrados em [10]. \square

Teorema 6.6. *Sejam y_1 e y_2 as soluções da equação*

$$L[y] = y'' + p(t)y' + q(t)y = 0,$$

onde p e q são contínuas no intervalo aberto I . Então y_1 e y_2 são linearmente dependentes em I se e somente se $W[y_1, y_2](t)$ for nulo para todo t em I . Ou, de outra forma, y_1 e y_2 são linearmente independentes em I se e somente se $W[y_1, y_2](t)$ nunca for nulo em I .

Demonstração. Suponha primeiro que f e g sejam linearmente dependentes, pelo Teorema 6.3 tem-se $W[y_1, y_2](t) = 0$ para todo $t \in I$.

Reciprocamente, seja $t_0 \in I$ um ponto qualquer, então, por hipótese, $W(t_0) = 0$. Logo, o sistema de equações

$$\begin{cases} c_1 y_1(t_0) + c_2 y_2(t_0) = 0 \\ c_1 y_1'(t_0) + c_2 y_2'(t_0) = 0 \end{cases} \quad (6.7)$$

em c_1 e c_2 tem uma solução não-trivial. Com estes valores de c_1 e c_2 , considere a função $\phi(t) = c_1 y_1(t) + c_2 y_2(t)$. Assim ϕ é uma solução da equação

$$y'' + p(t)y' + q(t)y = 0,$$

e pelas equações (6.7), ϕ também satisfaz as condições iniciais

$$\phi(t_0) = 0, \quad \phi'(t_0) = 0 \quad (6.8)$$

Portanto, pelo teorema de existência e unicidade² para equações diferenciais ordinárias de segunda ordem, $\phi(t) = 0$ para todo t em I . Uma vez que $\phi(t) = c_1 y_1(t) + c_2 y_2(t)$, onde c_1 e c_2 não são ambas nulas, isto significa que y_1 e y_2 são linearmente dependentes. Segue-se, então, e imediatamente, a outra forma do enunciado do teorema. \square

Teorema 6.7. *Todo autovalor do problema de Sturm-Liouville tem multiplicidade 1, isto é, o espaço vetorial das autofunções correspondentes tem dimensão 1.*

Demonstração. Sejam y_1 e y_2 duas soluções linearmente independentes de $L_\lambda[y] = 0$ e satisfazendo $F_1[y] = F_2[y] = 0$. Considerando o wronskiano

$$W(t) = W[y_1, y_2](t) = \begin{vmatrix} y_1(t) & y_2(t) \\ y_1'(t) & y_2'(t) \end{vmatrix},$$

do Teorema 6.2 segue que

$$W[y_1, y_2](t) \neq 0 \forall t \in [a, b],$$

pois y_1 e y_2 são soluções linearmente independentes e, que do Teorema 6.5,

$$W(t) = W(a) \exp \left[- \int_a^t \frac{p'(s)}{p(s)} ds \right] = W(a) \frac{p(a)}{p(t)},$$

isto é, $p(t)W(t) = p(a)W(a)$, concluindo que $p(t)W(t)$ é constante.

Por outro lado, das condições de fronteira,

$$\begin{aligned} F_1[y] &= \alpha_0 y_1(a) + \alpha_1 y_1'(a) = 0 \\ F_1[y] &= \alpha_0 y_2(a) + \alpha_1 y_2'(a) = 0, \end{aligned}$$

com α_0 e α_1 não simultaneamente nulos tem-se que

$$\begin{vmatrix} y_1(a) & y_1'(a) \\ y_2(a) & y_2'(a) \end{vmatrix} = 0,$$

ou seja, $W[y_1, y_2](a) = 0$ e portanto $W[y_1, y_2] \equiv 0$, o que contradiz a hipótese da independência linear de y_1 e y_2 . Assim o espaço vetorial das autofunções correspondentes a λ possui dimensão 1. \square

Definição 6.3. *Indica-se por $\mathcal{C}_{L^2}^{(1)}([a, b], \mathbb{C})$ o espaço das funções contínuas x definidas no intervalo $[a, b]$ e que tem a derivada primeira contínua. Neste espaço considere a norma $x \in \mathcal{C}^{(1)}([a, b]) \mapsto \|x\|_2$, onde*

$$\|x\|_2 = \left(\int_a^b |x(t)|^2 dt \right)^{1/2}.$$

²Considere o problema de valor inicial $y'' + p(t)y' + q(t)y = g(t)$, com $y(t_0) = y_0$ e $y'(t_0) = y_0'$, onde p, q e g são contínuas em um intervalo aberto I . Então, existe exatamente uma solução $y = \phi(t)$ desse problema e a solução existe em todo o intervalo I .

Lema 6.1. *Seja $x \in \mathcal{C}_{L_2}^{(1)}([a, b], \mathbb{C})$, então para todo $t \in [a, b]$, tem-se*

$$|x(t)|^2 \leq \frac{1}{(b-a)} (\|x\|_2)^2 + 2\|x\|_2 + \|x'\|_2.$$

Demonstração. Seja $\bar{t} \in [a, b]$ o ponto de máximo de $|x(t)|$. Assim

$$\begin{aligned} |x(\bar{t})|^2 &= |x(\bar{t})|^2 + |x(t)|^2 - |x(t)|^2 \\ &= |x(t)|^2 + |x(\bar{t})|^2 - |x(t)|^2 \\ &= |x(t)|^2 + \int_t^{\bar{t}} [x(s)\overline{x(s)}]' ds \\ &= |x(t)|^2 + \int_t^{\bar{t}} [x'(s)\overline{x(s)} + x(s)\overline{x'(s)}] ds \\ &= |x(t)|^2 + \int_t^{\bar{t}} [x'(s)\overline{x(s)}] ds + \int_t^{\bar{t}} [x(s)\overline{x'(s)}] ds \\ &\leq |x(t)|^2 + \int_a^b [x'(s)\overline{x(s)}] ds + \int_a^b [x(s)\overline{x'(s)}] ds \\ &= |x(t)|^2 + \langle x'(s), x(s) \rangle + \langle x(s), x'(s) \rangle \\ &\leq |x(t)|^2 + |\langle x'(s), x(s) \rangle| + |\langle x(s), x'(s) \rangle| \\ &\leq |x(t)|^2 + \|x'\| \|x\| + \|x\| \|x'\| \\ &= |x(t)|^2 + 2\|x\|_2 \|x'\|_2. \end{aligned}$$

Portanto $|x(\bar{t})|^2 \leq |x(t)|^2 + 2\|x\|_2 \|x'\|_2$. Integrando esta desigualdade em $[a, b]$ tem-se

$$(b-a)|x(\bar{t})|^2 \leq (\|x\|_2)^2 + 2(b-a)\|x\|_2 \|x'\|_2.$$

□

Corolário 6.1. *Seja $x \in \mathcal{C}_{L_2}^{(1)}([a, b], \mathbb{C})$ com $\|x\|_2 = 1$ então, para todo $t \in [a, b]$, tem-se*

$$|x(t)|^2 \leq \frac{1}{b-a} + 2\|x'\|_2.$$

Teorema 6.8. *Existe $c_0 \in \mathbb{R}$ tal que para $\lambda \in \mathbb{R}$, autovalor do problema de Sturm-Liouville, tem-se $\lambda \geq c_0$.*

Demonstração. Seja $y = y(t)$ uma autofunção real correspondente ao autovalor λ , tem-se que $L_\lambda[y] \equiv 0$. Note que é suficiente provar que existe $c_0 < 0$ tal que para $\lambda < c_0$ e $y \in \mathcal{C}^{(2)}([a, b], \mathbb{R})$ com $\|y\|_2 = 1$ tem-se sempre

$$\int_a^b L_\lambda[y](t)y(t) dt > 0,$$

isto é,

$$\int_a^b [-(p(t)y'(t))' + q(t)y(t) - \lambda\rho(t)y(t)]y(t) dt > 0.$$

Tem-se

$$\begin{aligned}
 \int_a^b [-(py')' + qy - \lambda\rho y]y dy &= \int_a^b [-(py')'y + qy^2 - \lambda\rho y^2] dt \\
 &= - \int_a^b [(py')'y] dt + \int_a^b qy^2 dt - \lambda \int_a^b \rho y^2 dt \\
 &= - \int_a^b [(py')'y + py'^2 - py'^2] dt + \int_a^b qy^2 dt - \lambda \int_a^b \rho y^2 dt \\
 &= - \int_a^b [(py')y]' dt + \int_a^b py'^2 dt + \int_a^b qy^2 dt - \lambda \int_a^b \rho y^2 dt \\
 &= -(py')y|_a^b + \int_a^b py'^2 dt + \int_a^b qy^2 dt - \lambda \int_a^b \rho y^2 dt.
 \end{aligned}$$

De $-(p(t)y(t)')y(t)|_a^b = -p(b)y'(b)y(b) + p(a)y'(a)y(a)$. Se $y(a) = 0$ então $p(a)y'(a)y(a) = 0$; se $y(a) \neq 0$ de $F_1[y] = \alpha_0 y(a) + \alpha_1 y'(a) = 0$ e do Corolário 6.1 segue que

$$|p(a)y'(a)y(a)| = |p(a)\frac{\alpha_0}{\alpha_1}y(a)^2| \leq \left| \frac{\alpha_0}{\alpha_1} \right| |p(a)| \left(\frac{1}{b-a} + 2\|y'\|_2 \right).$$

Fazendo $\left| \frac{\alpha_0}{\alpha_1} \frac{p(a)}{b-a} \right| = c_1$ e $\left| \frac{\alpha_0}{\alpha_1} 2p(a) \right| = c_2$, tem-se

$$|p(a)y'(a)y(a)| \leq c_1 + c_2\|y'\|_2$$

e portanto

$$p(a)y'(a)y(a) \geq -c_1 - c_2\|y'\|_2.$$

Analogamente, existem c_3 e c_4 tais que

$$-p(b)y'(b)y(b) \geq c_3 + c_4\|y'\|_2.$$

Para $\int_a^b p(t)y'(t)^2 dt$. Primeiramente, tem-se que $p(t) > 0$ para todo $t \in [a, b]$, logo existe p_1 tal que $p_1 = \inf_{a \leq t \leq b} p(t) > 0$. Assim

$$\int_a^b p(t)y'(t)^2 dt \geq \int_a^b p_1|y'(t)|^2 dt = p_1(\|y'\|_2)^2.$$

Para $\int_a^b q(t)y(t)^2 dt$. Sendo q contínua em $[a, b]$ existe q_1 tal que $q_1 = \inf_{a \leq t \leq b} q(t)$. Assim

$$\int_a^b q(t)y(t)^2 dt \geq \int_a^b q_1|y(t)|^2 dt = q_1(\|y\|_2)^2 = q_1.$$

Por fim, para $-\lambda \int_a^b \rho(t)y(t)^2 dt$. Tem-se

$$-\lambda \int_a^b \rho(t)y(t)^2 dt \geq -\lambda \int_a^b \rho_1|y(t)|^2 dt = -\lambda\rho_1(\|y\|_2)^2 = -\lambda\rho_1,$$

onde $\rho = \inf_{a \leq t \leq b} \rho(t) > 0$. Vale lembrar que $\lambda < 0$.

Efetuada portanto a soma, tem-se

$$\begin{aligned} \int_a^b [-(py)'] + qy - \lambda \rho y] y dy &= -(py)y|_a^b + \int_a^b py'^2 dt + \int_a^b qy^2 dt - \lambda \int_a^b \rho y^2 dt \\ &\geq c_3 + c_4 \|y'\|_2 - c_1 - c_2 \|y'\|_2 + p_1 (\|y'\|_2)^2 + q_1 - \lambda \rho_1. \\ &= c_5 + c_6 \|y'\|_2 + p_1 (\|y'\|_2)^2 - \lambda \rho_1 \\ &= \left(\sqrt{p_1} \|y'\|_2 + \frac{1}{2} \frac{c_6}{\sqrt{p_1}} \right)^2 - \frac{1}{4} \frac{c_6^2}{p_1} + c_5 - \lambda \rho_1 \\ &\geq c_7 - \lambda \rho_1, \end{aligned}$$

onde $c_5 = c_3 - c_1 + q_1$, $c_6 = c_4 - c_2$ e, por fim, $c_7 = c_5 - \frac{1}{4} \frac{c_6^2}{p_1}$. Assim

$$c_7 - \lambda \rho_1 > 0 \Leftrightarrow \lambda < \frac{c_7}{\rho_1} = c_8$$

e basta pois tomar $c_0 = \min\{0, c_8\}$.

□

Observação 6.1. O teorema precedente diz que existe um limitante inferior para os autovalores de um problema de Sturm-Liouville, ou seja, os autovalores podem até ser eventualmente negativos, mas não arbitrariamente negativos.

Exemplo 6.2. Considere o problema de Sturm-Liouville

$$\begin{aligned} y''(t) + \lambda y(t) &= 0, \text{ em } [0, 1], \quad y(0) = 0, \\ \beta_0 y(1) + \beta_1 y'(1) &= 0. \end{aligned}$$

Neste problema, $p(t) = 1$, $q(t) = 1$, $\rho(t) = 1$, $\alpha_0 = 1$ e, por fim, $a_1 = 0$. Se $y = y(t)$ é uma autofunção para o problema, então $-y''(t) = \lambda y(t)$. Multiplicando essa igualdade por $y(t)$, tem-se

$$\lambda y^2(t) = -y''(t)y(t).$$

Integrando por partes, no intervalo $[0, 1]$, obtém-se

$$\begin{aligned} \lambda \int_0^1 y^2(t) dt &= - \int_0^1 y''(t)y(t) dt \\ &= - \left[y(t)y'(t) \Big|_0^1 - \int_0^1 [y'(t)]^2 dt \right] \\ &= -y(1)y'(1) + y(0)y'(0) + \int_0^1 [y'(t)]^2 dt \\ &= -y(1)y'(1) + \int_0^1 [y'(t)]^2 dt, \end{aligned}$$

isto é,

$$\lambda \int_0^1 y^2(t) dt = \int_0^1 [y'(t)]^2 dt - y(1)y'(1). \quad (6.9)$$

No caso em que $\beta_0 = 0$ tem-se $y'(1) = 0$, e no caso em que $\beta_1 = 0$ tem-se $y(1) = 0$, ou seja, em ambos os casos tem-se

$$\lambda \int_0^1 y^2(t) dt = \int_0^1 [y'(t)]^2 dt,$$

garantindo $\lambda > 0$.

Se $\beta_0\beta_1 > 0$, tem-se

$$\beta_0 y(1) + \beta_1 y'(1) = 0 \Leftrightarrow \beta_0 y(1)y'(1) + \beta_1 [y'(1)]^2 = 0 \Leftrightarrow -y(1)y'(1) = \frac{\beta_1}{\beta_0} [y'(1)]^2.$$

Assim

$$\lambda \int_0^1 y^2(t) dt = \int_0^1 [y'(t)]^2 dt + \frac{\beta_1}{\beta_0} [y'(1)]^2 > 0,$$

concluindo que $\lambda > 0$.

Por fim, se $\beta_0\beta_1 < 0$, pode-se ter autovalores negativos.

Observação 6.2. Até que se faça menção contrária, $\lambda = 0$ não é autovalor do problema de Sturm-Liouville.

Considere a equação não homogênea

$$L[y] = y'' + p(t)y' + q(t)y = g(t), \quad (6.10)$$

onde p, q e g são funções contínuas dadas em um intervalo I . Considere também a equação homogênea associada a (6.10),

$$L[y] = y'' + p(t)y' + q(t)y = 0. \quad (6.11)$$

O resultado a seguir descreve a estrutura de soluções da equação não homogênea (6.10) e fornece uma base para se construir sua solução geral. Sua demonstração pode ser encontrada na seção 3.6 de [11].

Teorema 6.9. *A solução geral da equação não-homogênea (6.10) pode ser escrita na forma*

$$y(t) = c_1 y_1(t) + c_2 y_2(t) + Y(t), \quad (6.12)$$

onde $y_1(t)$ e $y_2(t)$ formam um conjunto fundamental de soluções da equação homogênea associada (6.11), c_1 e c_2 são constantes arbitrárias e Y é alguma solução específica da equação não homogênea (6.10).

Teorema 6.10. *Existe uma função contínua*

$$G : [a, b] \times [a, b] \rightarrow \mathbb{R}$$

tal que dado $f \in \mathcal{C}([a, b], \mathbb{C})$, $y \in \mathcal{C}^{(2)}([a, b], \mathbb{C})$ é solução de

$$\begin{aligned}
(S_0) \quad L[y] &= L_0[y] \equiv -(py')' + qy = f \text{ em } [a, b] \\
(F) \quad F_1[y] &\equiv \alpha_0 y(a) + \alpha_1 y'(a) = 0 \\
F_2[y] &\equiv \beta_0 y(b) + \beta_1 y'(b) = 0
\end{aligned}$$

se, e somente se,

$$y(t) = \int_a^b G(t, s) f(s) ds. \quad (6.13)$$

A função G é chamada de **função de Green** e é definida por

$$G(t, s) = \begin{cases} \frac{y_1(t)y_2(s)}{p(s)W[y_1, y_2](s)} = G_1(t, s), & \text{se } a \leq t < s \\ \frac{y_2(t)y_1(s)}{p(s)W[y_1, y_2](s)} = G_2(t, s), & \text{se } s \leq t \leq b, \end{cases}$$

onde as funções y_1 e y_2 são soluções linearmente independentes da equação homogênea $L_0[y] = 0$, com as condições de fronteira $F_1[y] = F_2[y] = 0$.

Demonstração. Pelo Teorema 6.9, a solução geral da equação

$$L_0[y] = -[p(t)y'(t)]' + q(t)y(t) = f(t) \quad (6.14)$$

é da forma $y(t) = c_1 y_1(t) + c_2 y_2(t) + Y(t)$, onde c_1 e c_2 são constantes, y_1 e y_2 são soluções linearmente independentes da equação homogênea

$$L_0[y] = -[p(t)y'(t)]' + q(t)y(t) = 0 \quad (6.15)$$

e $Y(t)$ é uma solução particular de (6.14). Tal solução particular pode ser determinada a partir do método de variação de parâmetros, ver seção 3.7 de [11]. O método de variação de parâmetro tem como objetivo substituir as constantes c_1 e c_2 acima por funções $v_1(t)$ e $v_2(t)$, respectivamente, de modo que $Y(t) = v_1(t)y_1(t) + v_2(t)y_2(t)$ seja solução da equação (6.14), em vez da equação (6.15). Assim, suponha que uma solução particular seja da forma

$$Y(t) = v_1(t)y_1(t) + v_2(t)y_2(t), \quad (6.16)$$

onde v_1 e v_2 são funções a serem determinadas. Derivando (6.16), obtém-se

$$Y'(t) = v_1'(t)y_1(t) + v_1(t)y_1'(t) + v_2'(t)y_2(t) + v_2(t)y_2'(t). \quad (6.17)$$

Como há apenas uma equação e duas funções a serem determinadas, existem muitas escolhas possíveis para v_1 e v_2 . Assim, impondo

$$v_1'(t)y_1(t) + v_2'(t)y_2(t) = 0 \quad (6.18)$$

como uma segunda condição de nossa escolha, obtém-se, assim, duas equações para as duas funções desconhecidas v_1 e v_2 .

Então, da equação (6.17), tem-se

$$Y'(t) = v_1(t)y_1'(t) + v_2(t)y_2'(t).$$

Derivando mais uma vez, obtém-se

$$Y''(t) = v_1'(t)y_1'(t) + v_1(t)y_1''(t) + v_2'(t)y_2'(t) + v_2(t)y_2''(t). \quad (6.19)$$

Agora, substituindo Y , Y' e Y'' na equação $L_0[y] = -[p(t)y'(t)]' + q(t)y(t) = f(t)$ e rearrumando os termos na equação resultante tem-se

$$\begin{aligned} & v_1(t)[-p(t)y_1''(t) - p'(t)y_1'(t) + q(t)y_1(t)] \\ & + v_2(t)[-p(t)y_2''(t) - p'(t)y_2'(t) + q(t)y_2(t)] \\ & - p(t)[v_1'(t)y_1'(t) + v_2'(t)y_2'(t)] = f(t) \end{aligned} \quad (6.20)$$

Cada uma das expressões entre colchetes na equação (6.20) é nula, pois ambas as funções y_1 e y_2 são soluções da equação homogênea (6.15). Portanto, a equação (6.20) se reduz a

$$-p(t)[v_1'(t)y_1'(t) + v_2'(t)y_2'(t)] = f(t), \quad (6.21)$$

de modo que as equações (6.18) e (6.21) formam o sistema,

$$\begin{cases} v_1'(t)y_1(t) + v_2'(t)y_2(t) = 0 \\ v_1'(t)y_1'(t) + v_2'(t)y_2'(t) = -\frac{f(t)}{p(t)} \end{cases},$$

de duas equações lineares algébricas para as derivadas $v_1'(t)$ e $v_2'(t)$ das funções desconhecidas.

Resolvendo o sistema, obtém-se

$$v_1'(t) = \frac{f(t)y_2(t)}{p(t)W[y_1, y_2](t)}, \quad (6.22)$$

$$v_2'(t) = -\frac{f(t)y_1(t)}{p(t)W[y_1, y_2](t)}. \quad (6.23)$$

Sendo y_1 e y_2 soluções linearmente independentes, do Teorema 6.6, $W[y_1, y_2](t) \neq 0$ para todo $t \in [a, b]$.

Integrando as igualdades (6.22) e (6.23), segue que

$$v_1(s) = \int_a^s \frac{y_2(t)f(t)}{p(t)W[y_1, y_2](t)} dt$$

e

$$v_2(s) = -\int_b^s \frac{y_1(t)f(t)}{p(t)W[y_1, y_2](t)} dt,$$

e finalmente

$$\begin{aligned} Y(s) &= v_1(s)y_1(s) + v_2(s)y_2(s) \\ &= y_1(s) \int_a^s \frac{y_2(t)f(t)}{p(t)W[y_1, y_2](t)} dt - y_2(s) \int_b^s \frac{y_1(t)f(t)}{p(t)W[y_1, y_2](t)} dt \\ &= \int_a^s \frac{y_1(s)y_2(t)f(t)}{p(t)W[y_1, y_2](t)} dt + \int_s^b \frac{y_2(s)y_1(t)f(t)}{p(t)W[y_1, y_2](t)} dt. \end{aligned}$$

Logo, define-se a função de Green como

$$G(s, t) = \begin{cases} \frac{y_1(s)y_2(t)}{p(s)W[y_1, y_2](s)} = G_1(s, t), & \text{se } a \leq t < s \\ \frac{y_2(s)y_1(t)}{p(s)W[y_1, y_2](s)} = G_2(s, t), & \text{se } s \leq t \leq b. \end{cases}$$

Do Teorema 6.7, segue que

$$p(s)W[y_1, y_2](s) = p(a)W[y_1, y_2](a) = p(t)W[y_1, y_2](t),$$

logo $G(s, t) = G(t, s)$ e então

$$G(t, s) = \begin{cases} \frac{y_1(t)y_2(s)}{p(s)W[y_1, y_2](s)} = G_1(t, s), & \text{se } a \leq t < s \\ \frac{y_2(t)y_1(s)}{p(s)W[y_1, y_2](s)} = G_2(t, s), & \text{se } s \leq t \leq b. \end{cases}$$

Portanto

$$Y(s) = \int_a^s G_1(t, s)f(t)dt + \int_s^b G_2(t, s)f(t)dt = \int_a^b G(t, s)f(t)dt.$$

A demonstração do teorema mostra que a aplicação

$$f \in \mathcal{C}([a, b], \mathbb{C}) \mapsto y = \mathcal{G}f \in \mathcal{C}^{(2)}([a, b], \mathbb{C}),$$

onde

$$(\mathcal{G}f)(t) = \int_a^b G(t, s)f(s)ds$$

é contínua, e

$$\lim_{\epsilon \rightarrow 0^+} \frac{\partial G_2}{\partial s}(t - \epsilon, t) - \lim_{\epsilon \rightarrow 0^+} \frac{\partial G_1}{\partial s}(t + \epsilon, t) = -\frac{1}{p(t)}.$$

□

Exemplo 6.3. Considere a equação $y''(t) = f(t)$ em que f é uma função contínua em $[a, b]$, com as condições de fronteira $y(a) = y(b) = 0$.

As soluções para a equação homogênea $y''(t) = 0$ satisfazendo as condições de fronteira são $y_1(t) = t - a$ e $y_2(t) = t - b$. Tem-se

$$W(t) = W[y_1, y_2](t) = \begin{vmatrix} t - a & t - b \\ 1 & 1 \end{vmatrix} = t - a - (t - b) = b - a.$$

Assim, a função de Green para o problema é

$$G(t, s) = \begin{cases} \frac{y_1(t)y_2(s)}{p(s)W[y_1, y_2](s)} = G_1(t, s), & \text{se } a \leq t < s \\ \frac{y_2(t)y_1(s)}{p(s)W[y_1, y_2](s)} = G_2(t, s), & \text{se } s \leq t \leq b, \end{cases}$$

ou seja,

$$G(t, s) = \begin{cases} \frac{(t-a)(s-b)}{b-a} = G_1(t, s), & \text{se } a \leq t < s \\ \frac{(t-b)(s-a)}{b-a} = G_2(t, s), & \text{se } s \leq t \leq b, \end{cases}$$

donde segue-se que a solução y do problema é dada por

$$\begin{aligned} y(s) &= \int_a^b G(t, s) dt \\ &= \int_a^s G_1(t, s) dt + \int_s^b G_2(t, s) dt \\ &= \frac{s-b}{b-a} \int_a^s (t-a) f(t) dt + \frac{s-a}{b-a} \int_s^b (t-b) f(t) dt \end{aligned}$$

Por exemplo, para $f(t) = -6t$, $a = 0$ e $b = 1$, a solução do problema $y''(t) = -6t$, com $y(0) = y(1) = 0$ é

$$y(t) = (t-1) \int_0^t s(-6t) ds + t \int_t^1 (s-1)(-6s) ds = t(1-t^2).$$

Exemplo 6.4. Considere a equação

$$y''(t) + y(t) = f(t) \text{ em } [0, \pi],$$

com as condições de fronteira $y(0) = y'(\pi) = 0$. Observe que $\lambda = 0$ não é um autovalor deste problema, pois a solução geral da equação homogênea $y''(t) + y(t) = 0$, dada por $y(t) = c_1 \cos t + c_2 \sin t$, com as condições de fronteira $y(0) = y'(\pi) = 0$, implica $c_1 = c_2 = 0$. Mas, note que $y_1(t) = \sin t$ e $y_2(t) = \cos t$ são soluções da equação homogênea $y''(t) + y(t) = 0$ satisfazendo respectivamente $y_1(0) = 0$ e $y_2'(\pi) = 0$.

Tem-se

$$W[y_1, y_2](t) = \begin{vmatrix} \sin t & \cos t \\ \cos t & \sin t \end{vmatrix} = -1.$$

Portanto a função de Green do problema é dada por

$$G(t, s) = \begin{cases} -\sin t \cos s, & \text{para } 0 \leq t \leq s, \\ -\cos t \sin s, & \text{para } s \leq t \leq \pi, \end{cases}$$

e a solução do problema é, portanto

$$y(t) = -\cos t \int_0^t \sin(s) f(s) ds - \sin t \int_t^\pi \cos(s) f(s) ds.$$

Corolário 6.2. Tem-se que y é solução do problema de Sturm-Liouville

$$\begin{cases} (S_\lambda) & -(p(t)y')' + [q(t) - \lambda\rho(t)]y = f(t) \\ (F) & F_1[y] = 0, F_2[y] = 0 \end{cases},$$

se, e somente se,

$$y(t) - \lambda \int_a^b G(t, s) y(s) \rho(s) ds = g(t),$$

onde $g(t) = \int_a^b G(t, s) f(s) ds$.

Demonstração. Uma vez que

$$\begin{aligned} L_\lambda[y] &= -(p(t)y'(t))' + (q(t) - \lambda\rho(t))y(t) \\ &= L_0[y] - \lambda\rho(t)y(t), \end{aligned}$$

então $L_\lambda[y] = f(t)$ se, e somente se, $L_0[y] = \lambda\rho y(t) + f(t)$. Pelo Teorema 6.10 tem-se $L_0[y] = \lambda\rho(t)y(t) + f(t)$ com $F_1[y] = F_2[y]$ se, e somente se,

$$y(t) = \int_a^b G(t, s)[\lambda\rho(s)y(s) + f(s)]ds,$$

isto é, se, e somente se,

$$y(t) - \lambda \int_a^b G(t, s)y(s)\rho(s)ds = g(t) = \int_a^b G(t, s)f(s)ds.$$

□

Indica-se por \mathcal{G} o operador integral definido em $\mathcal{C}_{L_2(\rho)}([a, b], \mathbb{C})$ pelo núcleo G , isto é,

$$\mathcal{G}[y](t) = \int_a^b G(t, s)y(s)\rho(s)ds, \quad t \in [a, b].$$

Vale lembrar que $\mathcal{C}_{L_2(\rho)}([a, b], \mathbb{C})$ indica o espaço vetorial $\mathcal{C}([a, b], \mathbb{C})$ munido do produto interno

$$x, y \in \mathcal{C}([a, b], \mathbb{C}) \mapsto \langle x, y \rangle_\rho = \int_a^b x(t)\overline{y(t)}\rho(t)dt.$$

Corolário 6.3. a) λ é um autovalor do problema de Sturm-Liouville se, e somente se, $\frac{1}{\lambda}$ é um autovalor de \mathcal{G} .

b) y é autofunção do problema de Sturm-Liouville correspondente ao autovalor λ se, e somente se, y é autofunção do operador \mathcal{G} correspondente ao autovalor $\frac{1}{\lambda}$.

Demonstração. Do Teorema 6.10, tem-se $L_0[y] = \lambda\rho y$ com $F_1[y] = F_2[y] = 0$ se, e somente se,

$$y(t) - \lambda \int_a^b G(t, s)y(s)\rho(s)ds = 0,$$

isto é, se, e somente se, $\frac{1}{\lambda}y(t) = \mathcal{G}[y](t)$, ou ainda, $\frac{1}{\lambda}$ é um autovalor de \mathcal{G} . □

Exemplo 6.5. Considere o problema

$$\begin{aligned} (1 + t^2)y''(t) + 2ty'(t) - \lambda y(t) &= f(t), \\ y(0) &= 0, \\ y'(1) &= 0, \end{aligned}$$

onde $f \in \mathcal{C}([0, 1], \mathbb{R})$. Escrevendo $(1 + t^2)y''(t) + 2ty'(t) = [(1 + t^2)y'(t)]'$ tem-se a equação na forma

$$[(1 + t^2)y'(t)]' - \lambda y(t) = f(t).$$

Se $L_\lambda[y] = [(1 + t^2)y'(t)]' - \lambda y(t)$, então $L_0[y] = [(1 + t^2)y'(t)]'$.

Para encontrar duas soluções, basta observar que se $y'(t) = 0$ então $y_1(t) = C_1$ e se $(1 + t^2)y'(t) = C_2$, obtém-se $y_2(t) = C_2 \operatorname{arctg}(t)$. Assim, $y_1'(1) = 0$ e $y_2(0) = 0$. Além disso,

$$W[y_1, y_2](t) = \begin{vmatrix} C_1 & C_2 \operatorname{arctg}(t) \\ 0 & \frac{C_2}{1 + t^2} \end{vmatrix} = \frac{C_1 C_2}{1 + t^2}.$$

Como $p(t) = 1 + t^2$, então $p(s)W[y_1, y_2](s) = C_1 C_2$, e a função de Green fica

$$G(t, s) = \begin{cases} \frac{y_1(t)y_2(s)}{p(s)W[y_1, y_2](s)}, & \text{se } 0 \leq t \leq s, \\ \frac{y_2(t)y_1(s)}{p(s)W[y_1, y_2](s)}, & \text{se } s < t \leq 1, \end{cases}$$

ou seja,

$$G(t, s) = \begin{cases} \operatorname{arctg}(t) = G_1(t, s), & \text{se } 0 \leq t \leq s, \\ \operatorname{arctg}(s) = G_2(t, s), & \text{se } s \leq t \leq 1. \end{cases}$$

Portanto, do Corolário 6.2, uma função $y = y(t)$ é solução do problema de Sturm-Liouville se, e somente se,

$$y(t) - \lambda \int_0^1 G(t, s)y(s)\rho(s)ds = \int_0^1 G(t, s)f(s)ds,$$

ou seja, se, e somente se,

$$\begin{aligned} y(t) - \lambda \int_0^t \operatorname{arctg}(s)y(s)ds + \int_t^1 \operatorname{arctg}(s)y(s)ds \\ = \int_0^t \operatorname{arctg}(t)f(s)ds + \int_t^1 \operatorname{arctg}(s)f(s)ds. \end{aligned}$$

Por exemplo, para $f(x) = 0$ tem-se que $y = y(t)$ é solução para o problema se, e somente se,

$$y(t) = \lambda \int_0^1 G(t, s)ds.$$

O núcleo G sendo real e simétrico segue que o operador \mathcal{G} é hermitiano compacto e todos os seus autovalores são reais. Pode-se aplicar a \mathcal{G} todos os resultados do capítulo anterior. Estes resultados, junto com o do presente capítulo permite enunciar as principais propriedades do problema de Sturm-Liouville no teorema que se segue.

Observação 6.3. Para a prova do item c) do teorema que segue, usa-se o seguinte resultado, do qual a demonstração será aqui omitida, mas encontrada no apêndice de [7].

Lema 6.2. O espaço $\mathcal{D}(]a, b[, \mathbb{C})$ indica o conjunto das funções definidas em $]a, b[$ e tomando valores complexos que são infinitamente deriváveis e nulas fora de um intervalo fechado contido em $]a, b[$. Nestas condições, $\mathcal{D}(]a, b[, \mathbb{C})$ é denso em $\mathcal{C}_{L_p(\rho)}([a, b], \mathbb{C})$, $1 \leq p < \infty$, onde $\mathcal{C}_{L_p(\rho)}([a, b], \mathbb{C})$ indica o espaço $\mathcal{C}([a, b], \mathbb{C})$ munido da norma

$$x \in \mathcal{C}([a, b], \mathbb{C}) \mapsto \|x\|_{p,\rho} = \left[\int_a^b |x(t)|^p \rho(t) dt \right]^{\frac{1}{p}}.$$

Teorema 6.11. Considere o problema de Sturm-Liouville

$$(S_\lambda) \quad L_\lambda[y] = -(p(t)y')' + [q(t) - \lambda\rho(t)]y = f(t) \text{ em } [a, b];$$

$$(F) \quad \begin{cases} F_1[y] = \alpha_0 y(a) + \alpha_1 y'(a) = 0, \\ F_2[y] = \beta_0 y(b) + \beta_1 y'(b) = 0. \end{cases}$$

onde L_λ , F_1 e F_2 satisfazem as condições mencionadas no começo deste capítulo.

- a) Os valores $\lambda \in \mathbb{C}$ tais que existe uma solução $y \neq 0$ de $L_\lambda[y]$ satisfazendo $F_1[y] = F_2[y] = 0$, isto é, os autovalores do problema de Sturm-Liouville, formam uma sequência infinita crescente λ_n de números reais tais que

$$\lim_{n \rightarrow \infty} \lambda_n = +\infty \text{ e } \sum_{n=1}^{\infty} \frac{1}{\lambda_n} < \infty.$$

- b) Cada autovalor λ_n tem multiplicidade 1, isto é, o espaço vetorial das autofunções correspondentes tem dimensão 1; fixando uma autofunção real ϕ_n tal que

$$\int_a^b \phi_n(t)^2 \rho(t) dt = 1,$$

então qualquer outra autofunção correspondente a λ_n é múltipla de ϕ_n .

- c) A sequência ϕ_n , de autofunções, é uma família ortonormal completa, (ver Teorema 3.6), do espaço pré-hilbertiano $\mathcal{C}_{L_2(\rho)}([a, b], \mathbb{C})$.

- d) Para toda função $x \in \mathcal{C}^{(2)}([a, b], \mathbb{C})$ tal que

$$F_1[x] = F_2[x] = 0,$$

tem-se

$$x(t) = \sum_{n=1}^{\infty} \beta_n \phi_n(t),$$

onde

$$\beta_n = \langle x, \phi_n \rangle_\rho = \int_a^b x(t) \phi_n(t) \rho(t) dt$$

com a série sendo uniforme e absolutamente convergente em $[a, b]$.

e) Seja $\lambda \neq \lambda_n$ para todo n e $f \in \mathcal{C}([a, b], \mathbb{R})$. O sistema

$$L_\lambda[y] = f \text{ com } F_1[y] = F_2[y] = 0$$

tem uma única solução y ,

$$y(t) = \sum_{n=1}^{\infty} \frac{\langle f, \phi_n \rangle}{\lambda - \lambda_n} \phi_n(t),$$

com esta série sendo uniformemente e absolutamente convergente em $[a, b]$.

f) Se $\lambda = \lambda_n$ dado $f \in \mathcal{C}([a, b], \mathbb{C})$, o sistema

$$L_\lambda[y] = f \text{ com } F_1[y] = F_2[y] = 0$$

tem solução se, e somente se, $\langle f, \phi_n \rangle = 0$, isto é,

$$\int_a^b f(t) \phi_m(t) dt = 0.$$

Demonstração. b) Segue do Teorema 6.7 e da Proposição 6.5.

d) Seja $x \in \mathcal{C}([a, b], \mathbb{C})$ com $F_1[x] = F_2[x] = 0$, fazendo $h(s) = L[x](s)$, do Teorema 6.10, segue-se que

$$\begin{aligned} x(t) &= \int_a^b G(t, s) h(s) ds = \\ &= \int_a^b G(t, s) \frac{h(s)}{\rho(s)} \rho(s) ds \\ &= \mathcal{G} \left[\frac{h}{\rho} \right] (t), \end{aligned}$$

onde, vale lembrar, que \mathcal{G} é o operador integral definido em $\mathcal{C}_{L_2(\rho)}([a, b], \mathbb{C})$ pelo núcleo G , com $\mathcal{G}[y](t) = \int_a^b G(t, s) y(s) \rho(s) ds$, $t \in [a, b]$. Sendo \mathcal{G} , também, um operador hermitiano compacto, pelo Teorema Espectral 5.2 e pelo Teorema 6.7, tem-se

$$x(t) = \mathcal{G} \left[\frac{h}{\rho} \right] (t) = \sum_{n=1}^{\infty} \frac{1}{\lambda_n} \left(\frac{h}{\rho} \right)_n \phi_n(t). \quad (6.24)$$

Considere $\beta_n = \langle x, \phi_n \rangle_\rho$, assim

$$\begin{aligned}
 \beta_n &= \langle x, \phi_n \rangle_\rho = \int_a^b x(t) \phi_n(t) \rho(t) dt \\
 &= \int_a^b \left[\int_a^b G(t, s) h(s) ds \right] \phi_n(t) \rho(t) dt \\
 &= \int_a^b \left[\int_a^b G(t, s) \phi_n(t) \rho(t) dt \right] h(s) ds \\
 &= \int_a^b \mathcal{G}[\phi_n](s) h(s) ds \\
 &= \int_a^b \frac{1}{\lambda_n} \phi_n(s) h(s) ds \\
 &= \frac{1}{\lambda_n} \int_a^b \phi_n(s) \frac{h(s)}{\rho(s)} \rho(s) ds \\
 &= \frac{1}{\lambda_n} \left\langle \frac{h}{\rho}, \phi \right\rangle_\rho \\
 &= \frac{1}{\lambda_n} \left(\frac{h}{\rho} \right)_n.
 \end{aligned}$$

Portanto,

$$x(t) = \sum_{n=1}^{\infty} \frac{1}{\lambda_n} \left(\frac{h}{\rho} \right)_n \phi_n(t) = \sum_{n=1}^{\infty} \beta_n \phi_n(t)$$

com a série, dada em (6.24), sendo uniforme e absolutamente convergente.

- c) Observe que todo $x \in \mathcal{D}(]a, b[, \mathbb{C})$ satisfaz as condições de fronteira $F_1[x] = F_2[x] = 0$, assim o item d) se aplica às funções de $\mathcal{D}(]a, b[, \mathbb{C})$, isto é,

$$x(t) = \sum_{n=1}^{\infty} \beta_n \phi_n(t),$$

com convergência uniforme e absoluta em $]a, b[$. Do (Lema 6.2) $\mathcal{D}(]a, b[, \mathbb{C})$ é denso em $\mathcal{C}_{L_2(\rho)}([a, b], \mathbb{C})$, logo, dados $y \in \mathcal{C}_{L_2(\rho)}([a, b], \mathbb{C})$ e $\epsilon > 0$, existe $x \in \mathcal{D}(]a, b[, \mathbb{C})$ tal que

$$\|y - x\| = \left\| y - \sum_{n=1}^{\infty} \beta_n \phi_n \right\| < \epsilon.$$

Logo, se considerar o conjunto de todas as combinações lineares finitas de ϕ_n , tem-se que este é denso em $\mathcal{C}_{L_2(\rho)}([a, b], \mathbb{C})$ e assim fica satisfeita a propriedade 4 do Teorema 3.6, concluindo que os ϕ_n são uma família ortonormal completa de $\mathcal{C}_{L_2(\rho)}([a, b], \mathbb{C})$ e portanto existem infinitas autofunções ϕ_n .

- a) Como se sabe, da Proposição 6.3, todos os autovalores do problema de Sturm-Liouville são reais e do Teorema 6.8 segue-se que quase todos são positivos. Do item c) segue que a sequência desses autovalores, (λ_n) , é infinita e do Teorema

Espectral (5.2) e do Corolário 6.3 segue que $\frac{1}{\lambda_n} \rightarrow 0$ e portanto $\lambda_n \rightarrow +\infty$. Ainda do Corolário 6.3, λ_n é autovalor do problema de Sturm-Liouville se, e somente se, $\frac{1}{\lambda_n}$ é autovalor de \mathcal{G} , assim para a sequência $\left(\frac{1}{\lambda_n}\right)$, a Proposição 5.4 nos garante que $\sum_{n=1}^{\infty} \frac{1}{\lambda_n^2} < \infty$. Por fim, do Corolário 5.2, segue-se mesmo que $\sum_{n=1}^{\infty} \frac{1}{\lambda_n} < \infty$.

e) Do corolário 6.2 segue-se que a solução y do problema de Sturm-Liouville, isto é, de

$$L_0[y] = \lambda \rho y + f,$$

com $F_1[y] = F_2[y] = 0$, satisfaz

$$\begin{aligned} y(t) &= \lambda \int_a^b G(t, s) y(s) \rho(s) ds + \int_a^b G(t, s) f(s) ds \\ &= \int_a^b G(t, s) [\lambda y(s) \rho(s) + f(s)] ds \\ &= \int_a^b G(t, s) \left[\lambda y(s) + \frac{f(s)}{\rho(s)} \right] \rho(s) ds, \end{aligned}$$

ou seja,

$$\frac{1}{\lambda} \mathcal{G}\left[\frac{f}{\rho}\right] = \frac{1}{\lambda} y - \mathcal{G}[y].$$

Do Teorema 5.6, substituindo λ por $\frac{1}{\lambda}$, λ_n por $\frac{1}{\lambda_n}$, x por y e y por $\frac{1}{\lambda} \mathcal{G}\left[\frac{f}{\rho}\right]$, tem-se

$$y(t) = \lambda \frac{1}{\lambda} \mathcal{G}\left[\frac{f}{\rho}\right](t) + \lambda \sum_{n=1}^{\infty} \frac{1}{\lambda_n} \frac{\frac{1}{\lambda} \mathcal{G}\left[\frac{f}{\rho}\right]_n}{\frac{1}{\lambda} - \frac{1}{\lambda_n}} \phi_n(t),$$

com esta série sendo uniformemente e absolutamente convergente em $[a, b]$. Usando (6.24) tem-se

$$\begin{aligned} y(t) &= \sum_{n=1}^{\infty} \frac{1}{\lambda_n} \left(\frac{f}{\rho}\right)_n \phi_n(t) + \lambda \sum_{n=1}^{\infty} \frac{1}{\lambda_n} \frac{\frac{1}{\lambda} \mathcal{G}\left[\frac{f}{\rho}\right]_n}{\frac{1}{\lambda} - \frac{1}{\lambda_n}} \phi_n(t) \\ &= \sum_{n=1}^{\infty} \frac{\langle f, \phi \rangle}{\lambda_n} \phi_n(t) + \lambda \sum_{n=1}^{\infty} \frac{\frac{1}{\lambda_n} \langle f, \phi_n \rangle}{\lambda_n - \lambda} \phi_n(t) \\ &= \sum_{n=1}^{\infty} \frac{\langle f, \phi_n \rangle}{\lambda_n - \lambda} \phi_n(t). \end{aligned}$$

f) Segue analogamente do Teorema 5.7. □

Exemplo 6.6. Considere o sistema

$$\begin{aligned} y''(t) + \lambda y(t) &= f(t), \text{ em } [0, \pi] \\ y(0) &= y(\pi) = 0. \end{aligned}$$

Do mesmo modo que no exemplo 6.1, observa-se que a solução geral da equação homogênea $y''(t) + \lambda y(t) = 0$, com $\lambda > 0$ é $y(t) = C_1 \cos(\sqrt{\lambda}t) + C_2 \sin(\sqrt{\lambda}t)$. Para que $y(0) = 0$ deve-se ter $C_1 = 0$, e para ter $y(\pi) = 0$ deve-se ter $\sqrt{\lambda}t = n\pi$, com $n \in \{1, 2, 3, \dots\}$, ou seja, os autovalores são $\lambda_n = n^2$ e as autofunções são $y_n(t) = \sin(nt)$. Note que a sequência de autovalores (λ_n) , satisfaz $\lambda_1 < \lambda_2 < \lambda_3 < \dots$ com $\lim_{n \rightarrow \infty} \lambda_n = \infty$ e $\sum_{n=1}^{\infty} \frac{1}{\lambda_n} = \sum_{n=1}^{\infty} \frac{1}{n^2} < \infty$. Tem-se também que

$$\langle y_n, y_n \rangle = \int_0^{\pi} \sin^2(nt) dt = \frac{\pi}{2},$$

e para $m \neq n$ tem-se

$$\begin{aligned} \langle y_n, y_m \rangle &= \int_0^{\pi} \sin(nt) \sin(mt) dt \\ &= \int_0^{\pi} \frac{\cos((m-n)t) - \cos((n+m)t)}{2} dt \\ &= \frac{1}{2} \left(\frac{\sin((n-m)t)}{n-m} - \frac{\sin((n+m)t)}{n+m} \right) \Big|_0^{\pi} = 0. \end{aligned}$$

Isto significa que as autofunções $y_n = \sin(nt)$ formam um sistema ortogonal em $\mathcal{C}([a, b], \mathbb{R})$ com relação ao produto interno $\langle f, g \rangle = \int_0^{\pi} f(t)g(t)dt$. Logo, as funções

$\phi_n(t) = \sqrt{\frac{2}{\pi}} \sin(nt)$ para $n \in \{1, 2, 3, \dots\}$ formam um conjunto ortonormal em $\mathcal{C}([a, b], \mathbb{R})$.

Como $\lambda = 0$ não é um autovalor do problema homogêneo, pelo Teorema 6.11(e), a solução do problema

$$\begin{aligned} y''(t) &= f(t), \text{ em } [0, \pi] \\ y(0) &= y(\pi) = 0. \end{aligned}$$

pode ser escrita na forma

$$\begin{aligned} y(t) &= \sum_{n=1}^{\infty} \frac{\langle f, \phi_n \rangle}{\lambda_n - 0} \phi_n(t) \\ &= \sum_{n=1}^{\infty} \frac{2}{\pi} \cdot \frac{1}{n^2} \left[\int_0^{\pi} f(s) \sin(ns) ds \right] \sin(nt). \end{aligned}$$

Por exemplo, se $f(t) = \cos(t^2 + t)$, a solução para o problema $y''(t) = f(t)$, $y(0) = y(\pi) = 0$ é dada por

$$\sum_{n=1}^{\infty} \frac{2}{\pi} \cdot \frac{1}{n^2} \left[\int_0^{\pi} \cos(s^2 + s) \operatorname{sen}(ns) ds \right] \operatorname{sen}(nt).$$

Exemplo 6.7. Os problemas de Sturm-Liouville aparecem naturalmente quando se aplica o método de separação de variáveis, ver Apêndice A, ao estudo de certas equações diferenciais parciais lineares de segunda ordem. Considere o problema da corda vibrante que tem a equação dada por

$$\frac{\partial}{\partial x} \left[p(x) \frac{\partial u}{\partial x}(x, t) \right] - q(x)u(x, t) = \rho(x) \frac{\partial^2 u}{\partial t^2}(x, t), \quad (6.25)$$

com condições iniciais $u(x, 0) = f(x)$ e $\frac{\partial u}{\partial t}(x, 0) = g(x)$ com $x \in [a, b]$ e condições de fronteira $u(a, t) = 0$ e $u(b, t) = 0$ para $t \in [0, \infty[$. As funções p e ρ representam a tensão e a densidade da corda, respectivamente, sendo portanto funções contínuas e positivas em $[a, b]$. Usando o método de separação de variáveis, procura-se soluções da equação diferencial parcial (6.25) que sejam da forma

$$u(x, t) = X(x)T(t). \quad (6.26)$$

Assim, substituindo (6.26) na equação (6.25) obtém-se

$$\begin{aligned} \frac{\partial}{\partial x} \left[p(x) \frac{\partial u}{\partial x}(x, t) \right] - q(x)u(x, t) &= \rho(x) \frac{\partial^2 u}{\partial t^2}(x, t) \\ \frac{\partial}{\partial x} [p(x)X'(x)T(t)] - q(x)X(x)T(t) &= \rho(x)X(x)T''(t) \\ p'(x)X'(x)T(t) + p(x)X''(x)T(t) - q(x)X(x)T(t) &= \rho(x)X(x)T''(t) \\ T(t)[p'(x)X'(x) + p(x)X''(x) - q(x)X(x)] &= \rho(x)X(x)T''(t) \\ \frac{[p(x)X'(x)]' - q(x)X(x)}{\rho(x)X(x)} &= \frac{T''(t)}{T(t)}. \end{aligned}$$

Como cada um dos lados é função de uma variável diferente segue-se que ambos devem ser iguais a uma constante, que será denotado por $-\lambda$. Daí X e T devem satisfazer

- i.* $T''(t) = -\lambda T(t)$
- ii.* $[p(x)X'(x)]' - q(x)X(x) = -\lambda \rho(x)X(x),$

com $X(a) = X(b) = 0$.

Note que *ii.* é um problema de Sturm-Liouville e que $\lambda = 0$ não é um autovalor. Pelo Teorema 6.11, existe uma sequência infinita (λ_n) de autovalores, com $\lim_{n \rightarrow \infty} \lambda_n = \infty$, e que *ii.* possui uma única solução não nula, a menos de uma constante multiplicativa. Chame de ϕ_n as autofunções normalizadas, isto é, $\int_a^b |\phi_n(t)|^2 \rho(t) dt = 1$, correspondentes aos autovalores λ_n .

Para cada n a solução geral da equação $T''(t) = -\lambda_n T(t)$ é da forma

$$T_n(t) = a_n \cos(\sqrt{\lambda_n}t) + b_n \operatorname{sen}(\sqrt{\lambda_n}t).$$

Deste modo, tem-se que, para cada $n \in \{1, 2, 3, \dots\}$,

$$\begin{aligned} u_n(x, t) &= \phi_n(x)T_n(t) \\ &= \phi_n(x)[a_n \cos(\sqrt{\lambda_n}t) + b_n \operatorname{sen}(\sqrt{\lambda_n}t)], \end{aligned}$$

é uma solução da equação (6.25). Procurando então a solução do problema original sob a forma

$$\begin{aligned} u(x, t) &= \sum_{n=1}^{\infty} u_n(x, t) \\ &= \sum_{n=1}^{\infty} \phi_n(x)[a_n \cos(\sqrt{\lambda_n}t) + b_n \operatorname{sen}(\sqrt{\lambda_n}t)] \end{aligned}$$

Para satisfazer as condições de fronteira deve-se ter

$$f(x) = u(x, 0) = \sum_{n=1}^{\infty} a_n \phi_n(x)$$

$$g(x) = \frac{\partial u}{\partial t}(x, 0) = \sum_{n=1}^{\infty} \sqrt{\lambda_n} b_n \phi_n(x).$$

Se f e g admitem este desenvolvimento em série, os coeficientes a_n e b_n são determinados por

$$a_n = \langle f, \phi_n \rangle_{\rho} = \int_a^b f(x) \overline{\phi_n(x)} \rho(x) dx,$$

$$b_n = \frac{1}{\sqrt{\lambda_n}} \langle g, \phi_n \rangle_{\rho} = \frac{1}{\sqrt{\lambda_n}} \int_a^b g(x) \overline{\phi_n(x)} \rho(x) dx.$$

Note a importância do estudo sobre a possibilidade de desenvolvimento de uma função f em série das autofunções ϕ_n , bem como a importância de condições que assegurem que esta série é absolutamente e uniformemente convergente em $[a, b]$, e não apenas em $\mathcal{C}_{L_2(\rho)}([a, b], \mathbb{C})$. Ainda subsiste a questão de saber como devem ser as funções f e g para que a série de $u(x, t)$ seja convergente, para que ela represente de fato uma solução da equação diferencial parcial (6.25), para que se tenha $u(x, t) \rightarrow f(x)$ quando $t \rightarrow 0^+$, entre outras. Essas últimas questões, em geral, têm de ser estudadas particularmente para cada tipo de equação.

Para uma modelagem de um problema da Corda Vibrante pode-se consultar [12].

Referências

- [1] HOFFMAN, K.; KUNZE, R. *Álgebra Linear*. 2. ed. Rio de Janeiro: Livros Técnicos e Científicos Editora S.A., 1979.
- [2] COELHO, F. U.; LOURENCO, M. L. *Um Curso de Álgebra Linear*. 2. ed. São Paulo: EDUSP., 2007.
- [3] KREYSZIG, E. *Introductory Functional Analysis With Applications*. 1. ed. New York: John Wiley and Sons, 1978.
- [4] LIMA, E. *Espaços Métricos*. 2. ed. Rio de Janeiro: Livros Técnicos e Científicos Editora S.A., 1978.
- [5] LIMA, E. *Elementos de Topologia Geral*. 1. ed. Rio de Janeiro: Ao Livro Técnico S.A, Editora da Universidade de São Paulo, 1970.
- [6] HALMOS, P. R. *Teoria Ingênua dos Conjuntos*. 1. ed. São Paulo: EDUSP, Editora Polígono., 1970.
- [7] HÖNIG, C. S. *Análise Funcional e o Problema de Sturm-Liouville*. 1. ed. São Paulo: IME-USP, 1970.
- [8] HÖNIG, C. S. *Análise Funcional e Aplicações. V.II*. 1. ed. São Paulo: IME-USP, 1970.
- [9] CAVALHEIRO, A. C. *Minicurso: O Problema de Sturm-Liouville*. [S.l.]: II Colóquio de Matemática da Região Sul, Universidade Estadual de Londrina, 2012.
- [10] KREIDER, D. L. et al. *Introdução À Análise Linear*. 1. ed. Rio de Janeiro: Ao Livro Técnico S.A, Rio de Janeiro, 1972.
- [11] BOYCE, W. E.; DIPRIMA, R. C. *Equações Diferenciais Elementares e Problemas de Valores de Contorno*. 7. ed. Rio de Janeiro: LTC, 2002.
- [12] SOTOMAYOR, J. *Lições de Equações Diferenciais Ordinárias*. 1. ed. Rio de Janeiro: IMPA, CNPq, 1979.
- [13] IÓRIO, V. *EDP: Um Curso de Graduação, Coleção Matemática Universitária*. 1. ed. Rio de Janeiro: IMPA, CNPq, 1991.

- [14] FIGUEIREDO, D. G. d. *Análise de Fourier e Equações Diferenciais Parciais*. 4. ed. Rio de Janeiro: IMPA, CNPq, 2003.

A Método de Separação de Variáveis e Séries de Fourier

Seja Ω um subconjunto aberto de \mathbb{R}^n . O leitor deve estar familiarizado com as propriedades elementares de funções de $u : \Omega \rightarrow \mathbb{R}$, em particular, com conceitos de continuidade, diferenciabilidade e derivadas parciais.

Uma **equação diferencial parcial** (EDP) é uma equação envolvendo duas ou mais variáveis independentes x, y, z, t, \dots e derivadas parciais de uma função (variável dependente) $u = u(x, y, z, t, \dots)$. De maneira mais precisa, uma EDP em n variáveis independentes x_1, x_2, \dots, x_n é uma equação da forma

$$F(x_1, x_2, \dots, x_n, u, \frac{\partial u}{\partial x_1}, \dots, \frac{\partial u}{\partial x_n}, \frac{\partial^2 u}{\partial x_1^2}, \dots, \frac{\partial^2 u}{\partial x_1 \partial x_n}, \dots, \frac{\partial^k u}{\partial x_n^k}) = 0, \quad (\text{A.1})$$

onde $x = (x_1, x_2, \dots, x_n) \in \Omega$, F é uma função dada e $u = u(x)$ é uma função que queremos determinar.

A classificação de EDP's segundo ordem e linearidade é semelhante à classificação das equações diferenciais ordinárias. A **ordem** de uma EDP é dada pela derivada parcial de maior ordem que ocorre na equação. Uma EDP é dita **linear** se é de primeiro grau em u e em todas as suas derivadas parciais que ocorrem na equação; caso contrário a EDP é dita **não linear**. A forma mais geral de uma EDP linear de primeira ordem é

$$\sum_{j=1}^n a_j(x) \frac{\partial u}{\partial x_j} u + b(x)u + c(x) = 0,$$

onde algum dos coeficientes a_i não é identicamente nulo. Para equações de segunda ordem, a forma mais geral de uma EDP linear é

$$\sum_{i,j=1}^n a_{ij}(x) \frac{\partial^2 u}{\partial x_i \partial x_j} u + \sum_{j=1}^n a_j(x) \frac{\partial u}{\partial x_j} u + c(x) + d(x) = 0,$$

onde algum dos coeficientes a_{ij} não é identicamente nulo.

Diz-se que uma EDP é **homogênea** se o termo que não contém a variável dependente é identicamente nulo.

Exemplo A.1. A equação

$$x \frac{\partial u}{\partial x} - y \frac{\partial u}{\partial y} = \text{sen}(xy)$$

é uma EDP linear não homogênea de primeira ordem.

Exemplo A.2. A equação de Poisson, para o caso em duas dimensões,

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = h(x, y),$$

é uma equação linear de segunda ordem não homogênea se a função h não for identicamente nula. No caso em que $h \equiv 0$, a equação é homogênea e chamada de *equação de Laplace*. A saber, a equação de Poisson está associada a fenômenos físicos estacionários, isto é, independentes do tempo.

Exemplo A.3.

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2}.$$

A equação da onda é também linear, homogênea e de segunda ordem. A variável $t > 0$ representa o tempo, $x \in \mathbb{R}$ é a variável espacial e $c > 0$ é uma constante (velocidade de propagação da onda).

Considere uma EDP de primeira ou segunda ordem com n variáveis independentes x_1, \dots, x_n . Ou seja, considere então uma equação do tipo

$$\sum_{i,j=1}^n a_{ij}(x) \frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_{j=1}^n b_j(x) \frac{\partial u}{\partial x_j} + c(x)u + d(x) = 0. \quad (\text{A.2})$$

Denote por k a ordem da equação com, $k = 1$ ou $k = 2$. Pode-se escrever a equação (A.2) na forma

$$Lu = f, \quad (\text{A.3})$$

onde $f(x) = -d(x)$ e

$$(Lu)(x) = \sum_{i,j=1}^n a_{ij}(x) \frac{\partial^2 u}{\partial x_i \partial x_j}(x) + \sum_{j=1}^n b_j(x) \frac{\partial u}{\partial x_j}(x) + c(x)u(x). \quad (\text{A.4})$$

A cada função u (suficientemente diferenciável) corresponde uma única função Lu , dessa maneira define-se um operador linear L ou operador diferencial parcial.

Como em EDO's, pode-se associar à EDP não homogênea (A.3) a EDP linear homogênea

$$Lu = 0, \quad (\text{A.5})$$

que é chamada *equação homogênea associada* à equação (A.3). Usando a linearidade do operador L e indução, nota-se que qualquer combinação linear de soluções da equação (A.5) é também solução de (A.5).

Proposição A.1. *Seja L um operador diferencial parcial linear de ordem k cujos coeficientes estão definidos no aberto Ω de \mathbb{R}^n . Suponha que $(u_m)_{m \in \mathbb{N}}$ seja uma sequência de funções de classe \mathcal{C}^k em Ω satisfazendo a EDP linear homogênea (A.5). Se $(\alpha_m)_{m \in \mathbb{N}}$ é uma sequência de escalares tal que a série*

$$u(x) = \sum_{m=1}^{\infty} \alpha_m u_m(x) \quad (\text{A.6})$$

é convergente e k vezes diferenciável termo a termo no aberto Ω , então u satisfaz a EDP linear homogênea (A.5).

Demonstração. Segue de imediato, aplicando $u(x)$, $\frac{\partial u}{\partial x_i}(x)$ e $\frac{\partial^2 u}{\partial x_i \partial x_j}(x)$ na definição (A.4). \square

A.1 Separação de Variáveis e Séries de Fourier

A.1.1 O Método de Separação de Variáveis

Considere um problema de condução de calor em uma barra de seção reta uniforme feita com material homogêneo. Escolha o eixo dos x de modo a formar o eixo da barra e de modo que $x = 0$ e $x = L$ correspondem às extremidades da barra. Suponha, ainda, que os lados da barra estão perfeitamente isolados, de modo que não há transmissão de calor aí. Pode-se supor, também, que as dimensões da seção reta são tão pequenas que a temperatura u pode ser considerada constante em qualquer seção reta. Então, u só depende da coordenada axial x e do instante t .

A variação da temperatura na barra é governada por uma equação diferencial parcial de segunda ordem cuja dedução pode ser encontrada no Apêndice A de [11] e tem a forma

$$\frac{\partial u}{\partial t} = \alpha^2 \frac{\partial^2 u}{\partial x^2}, \quad 0 < x < L, \quad t > 0, \quad (\text{A.7})$$

onde α^2 é uma constante conhecida como **difusividade térmica**. Além disso, suponha que a distribuição inicial de temperatura na barra é dada, ou seja,

$$u(x, 0) = f(x), \quad 0 \leq x \leq L, \quad (\text{A.8})$$

onde f é uma função conhecida. Finalmente, suponha que as extremidades da barra são mantidas a temperaturas fixas e iguais. Assim, suponha que u é sempre zero quando $x = 0$ ou $x = L$:

$$u(0, t) = 0, \quad u(L, t) = 0, \quad t > 0. \quad (\text{A.9})$$

Do ponto de vista físico é razoável procurar soluções $u(x, t)$ que satisfazem a equação (A.7), para $0 < x < L$ e para $t > 0$, a condição inicial (A.8) quando $t = 0$ e as condições de contorno (A.9) em $x = 0$ e $x = L$. A EDP em (A.7) é a equação de calor na sua forma

mais simples. Joseph B. Fourier (1768-1830) foi o primeiro a estudar sistematicamente o problema de condução de calor e seu nome está intrinsecamente ligado ao método de separação de variáveis, que é o precursor do método de expansão em autofunções.

O problema de valores de contorno

$$\begin{aligned}\frac{\partial u}{\partial t} &= \alpha^2 \frac{\partial^2 u}{\partial x^2}, & 0 < x < L, \quad t > 0, \\ u(0, t) &= 0 = u(L, t), & t \geq 0,\end{aligned}\tag{A.10}$$

é um problema linear homogêneo, isto é, o espaço das soluções do problema (A.10) é um subespaço vetorial de $\mathcal{C}^2((0, L) \times (0, +\infty)) \cap \mathcal{C}([0, L] \times [0, +\infty))$. Nesse caso vale o princípio da superposição e a ideia então é procurar uma família de soluções $(u_n)_{n \in \mathbb{N}}$ do problema (A.10) de modo que uma solução qualquer possa ser expressa na forma

$$u = \sum_{n=1}^{\infty} \alpha_n u_n.\tag{A.11}$$

Desse modo a solução do problema (A.7), junto com as condições (A.8) e (A.9), seria dada por (A.11), onde a sequência numérica $(\alpha_n)_{n \in \mathbb{N}}$ seria determinada pela condição inicial, isto é, impondo-se que

$$f(x) = \sum_{n=1}^{\infty} \alpha_n u_n(x, 0).\tag{A.12}$$

Este método consiste em procurar soluções de (A.10), da forma

$$u(x, t) = X(x)T(t).\tag{A.13}$$

Substituindo (A.13) na EDP, obtém-se

$$X(x)T'(t) = \alpha^2 X''(x)T(t)$$

e, dividindo por $\alpha^2 X(x)T(t)$ (nos pontos onde X e T não se anulam), resulta

$$\frac{X''(x)}{X(x)} = \frac{1}{\alpha^2} \cdot \frac{T'(t)}{T(t)}.\tag{A.14}$$

O primeiro membro da equação (A.14) é uma função que só depende de x enquanto o segundo membro da equação é uma função que só depende de t , logo ambos os lados devem ser iguais a uma mesma constante, chamada de $-\lambda$. Obtém-se então duas EDO's

$$\begin{aligned}-X''(x) &= \lambda X(x), \\ T'(t) &= -\alpha^2 \lambda T(t).\end{aligned}$$

Procurando soluções $u \in \mathcal{C}^2((0, L) \times (0, +\infty)) \cap \mathcal{C}([0, L] \times [0, +\infty))$, logo deve-se ter $X \in \mathcal{C}^2((0, L)) \times \mathcal{C}([0, L])$ e $T \in \mathcal{C}^2((0, +\infty)) \times \mathcal{C}([0, +\infty))$. Impondo a condição de contorno,

$$X(0)T(0) = 0 = X(L)T(t), \quad \forall t \geq 0,$$

e descartando a solução identicamente nula, X é solução do problema

$$X''(x) + \lambda X(x) = 0, \quad 0 < x < L, \quad (\text{A.15})$$

com $X(0) = 0 = X(L)$, enquanto T é qualquer solução da EDO

$$T'(t) + \alpha^2 \lambda T(t) = 0. \quad (\text{A.16})$$

Um valor λ para o qual (A.15) tem solução não trivial é chamado *autovalor* do problema (A.15) e as soluções não triviais correspondentes são as autofunções correspondentes ao autovalor λ .

Como procura-se soluções reais e $-\lambda$ é igual a ambos os membros da equação (A.14), tem-se assim como visto no Capítulo 6 que λ é real. Deste modo, considere a existência de autovalores e autofunções reais para o problema (A.15). Se $\lambda > 0$, a solução geral (real) da EDO em (A.15) é dada por

$$X(x) = a \cos(\sqrt{\lambda}x) + b \operatorname{sen}(\sqrt{\lambda}x),$$

onde a e b são constantes reais arbitrárias, que impondo as condições de contorno, tem-se

$$\begin{aligned} X(0) &= a = 0, \\ X(L) &= b \operatorname{sen}(\sqrt{\lambda}L) = 0. \end{aligned}$$

Uma vez que X não é identicamente nula, segue que $b \neq 0$ e portanto $\operatorname{sen}(\sqrt{\lambda}L) = 0$, ou seja, $\sqrt{\lambda}L = n\pi$ para algum $n \in \mathbb{N}$, $k \neq 0$, logo

$$\lambda_n = \frac{n^2 \pi^2}{L^2}, \quad n \in \mathbb{N} \quad (\text{A.17})$$

são os autovalores de (A.15) e

$$X_n(x) = \operatorname{sen}\left(\frac{n\pi x}{L}\right), \quad x \in [0, L], \quad (\text{A.18})$$

são as autofunções associadas.

Para resolver a EDO (A.16), basta notar que $e^{-\alpha^2 \lambda t}$ é um fator integrante e a solução geral é, portanto,

$$T(t) = ce^{-\alpha^2 \lambda t},$$

onde c é uma constante arbitrária.

Assim, obtém-se as soluções

$$u_n(x, t) = X(x)T(t) = \operatorname{sen}\left(\frac{n\pi x}{L}\right) \cdot \exp\left(-\frac{\alpha^2 n^2 \pi^2}{L^2} t\right), \quad x \in [0, L], \quad t \geq 0, \quad n \in \mathbb{N}$$

e, usando o princípio da superposição, procura-se uma solução da forma

$$u(x, t) = \sum_{n=1}^{\infty} b_n \operatorname{sen}\left(\frac{n\pi x}{L}\right) \cdot \exp\left(-\frac{\alpha^2 n^2 \pi^2}{L^2} t\right), \quad x \in [0, L], \quad t \geq 0. \quad (\text{A.19})$$

Observe que deixando de lado os problemas de convergência e diferenciabilidade termo a termo, u é solução (A.10). Para resolver a EDP da forma (A.19) e impondo a condição inicial, tem-se

$$f(x) = \sum_{n=1}^{\infty} b_n \operatorname{sen} \left(\frac{n\pi x}{L} \right), \quad x \in [0, L]. \quad (\text{A.20})$$

Portanto, para resolver a EDP por esse método, é preciso saber quais funções $f \in \mathcal{C}([0, L], \mathbb{C})$ podem ser expressas como uma série em senos da forma (A.20), no sentido em que as séries (A.19) e (A.20) convergem, e como calcular os coeficientes b_n conhecendo a função f .

Para o cálculo dos coeficientes, sabe-se que as autofunções X_n e X_m com $n \neq m$ são ortogonais em relação ao produto interno $\langle g, h \rangle = \int_a^b g(x) \overline{h(x)} dx$ definido no espaço $\mathcal{C}_{L^2}([a, b], \mathbb{C})$. Assim basta calcular o produto interno de f com X_n , obtendo

$$\begin{aligned} \langle f, X_n \rangle &= \left\langle \sum_{m=1}^{\infty} b_m X_m, X_n \right\rangle = \sum_{m=1}^{\infty} b_m \langle X_m, X_n \rangle \\ &= b_n \langle X_n, X_n \rangle = b_n \int_0^L \operatorname{sen}^2 \left(\frac{n\pi x}{L} \right) dx = b_n \frac{L}{2}, \end{aligned}$$

e portanto

$$b_n = \frac{2}{L} \int_0^L f(x) \operatorname{sen} \left(\frac{n\pi x}{L} \right) dx. \quad (\text{A.21})$$

Note que os coeficientes b_n , $n \in \mathbb{N}$, dados por (A.21) estão bem definidos se $f \in \mathcal{C}([0, L], \mathbb{C})$.

No caso em que o fluxo de calor em cada extremidade da barra é proporcional à temperatura na extremidade, tem-se um problema um pouco diferente mas que também pode ser tratado pelo método de separação de variáveis. Neste caso tem-se um outro tipo de condições de contorno. Um exemplo desse caso é o problema

$$\begin{aligned} \frac{\partial u}{\partial t} &= \alpha^2 \frac{\partial^2 u}{\partial x^2} \quad \text{em } (0, L) \times (0, +\infty), \\ \frac{\partial u}{\partial x}(0, t) + u(0, t) &= 0 = \frac{\partial u}{\partial x}(L, t) + u(L, t), \quad t \geq 0, \\ u(x, 0) &= f(x), \quad x \in [0, L]. \end{aligned} \quad (\text{A.22})$$

Observe que o problema de contorno obtido de (A.22) retirando-se a condição inicial ainda é um problema linear homogêneo que, neste caso, utilizando o método de separação de variáveis obtém-se

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} \left[a_n \cos \left(\frac{n\pi x}{L} \right) + b_n \operatorname{sen} \left(\frac{n\pi x}{L} \right) \right]. \quad (\text{A.23})$$

O método de separação de variáveis leva então, de modo natural, ao estudo de séries do tipo (A.23), as **séries de Fourier**, onde a_n e b_n denomina-se **coeficientes de Fourier**.

Para saber os detalhes desta teoria e os resultados de convergência da série, o leitor pode consultar as referências [13] e [14].