Universidade Federal de Uberlândia - UFU
Faculdade de Matemática - FAMAT
Coordenação dos Cursos de Bacharelato e Licenciatura em Matemática

Projeto de Trabalho

Trabalho de Conclusão de Curso

Equivalências e aplicações do Lema de Zorn

Aluno: Maurício Antônio da Costa Neto

Orientador(a): Vinícius Vieira Fávaro

Maurício Antônio da Costa Neto

Equivalências e aplicações do Lema de Zorn

Trabalho apresentado à Faculdade de Matemática, como parte dos requisitos para obtenção do título de licenciado em matemática

Universidade Federal de Uberlândia – UFU

Faculdade de Matemática

Orientador: Prof. Dr. Vinícius Vieira Fávaro

Uberlândia-MG

2021

Maurício Antônio da Costa Neto

Equivalências e aplicações do Lema de Zorn

Trabalho apresentado à Faculdade de Matemática, como parte dos requisitos para obtenção do título de licenciado em matemática

Trabalho aprovado. Uberlândia-MG, ***17 de Junho de 2021***:

Prof. Dr. Vinícius Vieira Fávaro
Orientador

Prof. Dr. Fábio José Bertoloto

Prof. Dra. Laís Bássame Rodrigues

Uberlândia-MG
2021

AGRADECIMENTOS

Meus agradecimentos se estendem a todos aqueles que fizeram com que fosse possível eu chegar a este ponto da minha vida. Agradeço a todos os professores que eu tive, todos aqueles que me estimularam a sempre buscar e querer mais, espero um dia ser como eles. Agradeço também a todas as escolas por onde passei, e sobretudo à UFU, que me transformou enquanto pessoa.

Agradeço à primeira professora que tive, minha mãe, pessoa que me acompanhou em todas as etapas da minha vida, sempre me apoiando e incentivando. Agradeço também àquela que cada dia me ensina a ser uma pessoa melhor e que escolheu estar comigo em todos os momentos, de agora para sempre, Stefânia, a você todo meu amor. Agradeço ao meu pai, meus avós, meus tios, tias, primos e primas que sempre me apoiaram e me fizeram chegar até aqui.

Ao professor Vinícius agradeço pelos seis anos de orientação, agradeço por toda paciência e por todos os conselhos. Agradeço o professor Victor Gonzalo por me trazer ao curso de Matemática, e por toda a disposição de sempre me ajudar.

A OBMEP eu agradeço por ter me proporcionado aventurar no campo da Matemática, enquanto ainda estava na educação básica, agradeço por ter me proporcionado o PICME. Agradeço ao CNPQ pelas bolsas concedidas.

Agradeço ao professor Vlademir Marim por ter me acompanhado em toda a jornada do intercâmbio, agradeço também ao professor Arlindo por sempre me ajudar neste processo. Agradeço ainda à Doutora Fátima Leite por ter me recebido na Universidade de Coimbra, e à Doutora Susana Moura por todos os ensinamentos. Agradeço à CAPES pelo apoio financeiro nesse projeto.

Agradeço também ao professor Douglas, por me permitir aventurar em um projeto diferente de tudo que eu conhecia. Agradeço a ele pela oportunidade de estar durante um ano e meio no Prossiga, e agradeço a UFU pela bolsa concedida.

Não posso deixar de agradecer a todos os professores que tive no curso, alguns deles que se tornaram verdadeiros amigos. Devo agradecer também aos colegas de curso, dentre os quais eu encontrei companhias que levo para a minha vida.

Agradeço ainda aos professores Fábio e Laís, por aceitarem participar dessa banca, e por disporem de seu tempo para analisar o que aqui está apresentado.

"Faça ou não faça, tentativa não há".

RESUMO

O Lema de Zorn é talvez o tópico mais discutido de toda a teoria de conjuntos e, como é de grande conhecimento, vários são os resultados equivalentes a ele. Neste trabalho abordaremos as equivalências clássicas do Lema de Zorn, que são com o Axioma da Escolha e com o Princípio de Boa Ordem. Apresentaremos também a equivalência deste lema com a existência de base de espaços vetoriais. Por fim, apresentaremos duas aplicações do Lema de Zorn na Análise Funcional.

Palavras-chave: Lema de Zorn, Axioma da Escolha, Princípio da Boa Ordem, bases de espaços vetoriais, Teorema de Hahn-Banach, sistemas ortonormais completos.

ABSTRACT

The Zorn's Lemma is perhaps the most discussed topic of the set theory and, since it is well-known, there are several results equivalent to it. In this work we will prove the classical equivalences of Zorn's Lemma, that are with the Axiom of Choice and with the Well-Ordering Principle. We will also present the equivalence of this lemma with the existence of basis of vector spaces. Finally, we will present two applications of the Zorn's Lemma in the Functional Analysis.

Keywords: Zorn's Lemma, Axiom of Choice, Well-Ordering Principle, basis of vector spaces, Hahn-Banach Theorem, complete orthonormal systems.

SUMÁRIO

INTRODUÇÃO

A Teoria de Conjuntos nasce em 1870, quando Georg Cantor começou a desenvolver estudos sobre conjuntos infinitos. Todavia, as noções ingênuas de conjuntos propostas por ele não poderiam compor uma teoria completa. A ideia de que a Teoria de Conjuntos seria a base da matemática parece ser questionável, quando inúmeros paradoxos são elaborados acerca da teoria de Cantor.

O mais conhecido e representativo desses paradoxos foi o paradoxo de Russel, proposto em 1901, que representou um grande furo na teoria, pois mostrava que as bases da Teoria de Conjuntos não estavam bem desenvolvidas. A partir desse momento, viu-se necessária a criação de sistemas axiomáticos completos que fizessem com que as definições "ingênuas" de Cantor fossem formalizadas para dar origem a uma teoria sólida.

A primeira axiomatização foi proposta por Ernst Zermelo, em 1908, e em 1922 esta axiomatização foi alterada por Abraham Fraenkel, dando assim origem ao sistema axiomático Zermelo-Fraenkel (ZF), que embora não seja o único existente, é o mais utilizado nas teorias matemáticas. Segundo o próprio Fraenkel, "a axiomatização é o que legitima a teoria, e a afasta das zonas das contradições" [2].

O conjunto de axiomas de Zermelo-Fraenkel renuncia a necessidade de dar uma definição ao conceito de conjunto, bem como à relação de pertinência entre um conjunto e seus elementos. A axiomática de Zermelo-Fraenkel será em sua totalidade apresentada no Capítulo 1 deste trabalho.

O sistema Zermelo-Fraenkel original não continha o Axioma da Escolha, entretanto era necessário, para que certas teorias fizessem sentido, que se enunciasse que dada uma coleção de conjuntos, é possível formar um novo conjunto contendo exatamente um elemento de cada conjunto da coleção dada. A excitação em enunciar este axioma decorria do fato de que não se sabia se esse axioma era, ou não, uma consequência dos anteriores. Contudo, houve um certo consenso na comunidade matemática na época que levou Zermelo a adicionar este último axioma ao sistema.

Só em 1939, Kurt Gödel conseguiu mostrar que o sistema ZF é um sistema consistente, e finalmente, em 1963, Paul Cohen mostrou que o axioma da escolha não se deriva de nenhum dos demais. Entretanto, a discussão passa a partir de então a ser se existe algum enunciado mais formal que pode ser um substituto plausível do Axioma da Escolha no sistema ZF.

Com tal discussão, que se arrasta até os dias contemporâneos, começam a surgir inúmeros textos, e artigos científicos, que provam a equivalência do Axioma da Escolha a outras proposições.

A mais conhecida destas equivalências é com o Lema de Zorn. Apresentado pela primeira vez em 1935, o Lema de Zorn foi enunciado por Max Zorn. A princípio este lema foi apresentado como uma equivalência do Princípio da Boa Ordem, e Zorn prometeu publicar um texto que comprovasse a equivalência de seu lema com o Axioma da Escolha. Esse resultado nunca foi publicado por Zorn, mas em 1939, Bourbaki formulou o enunciado clássico deste lema e mostrou sua equivalência ao Axioma da Escolha.

Um dos objetivos deste trabalho é apresentar alguns dos resultados equivalentes ao Lema de Zorn. As equivalências mais clássicas, que são entre o Lema de Zorn, o Axioma da Escolha e o Princípio da Boa Ordem, serão apresentadas no Capítulo 1. No Capítulo 2 apresentaremos outra equivalência do Lema de Zorn, que não é tão conhecida. Em cursos avançados de Álgebra Linear é apresentada a prova de que o Lema de Zorn implica a existência de base para espaços vetoriais. Mostraremos que, na verdade, este fato é equivalente ao Lema de Zorn.

Finalmente, no Capítulo 3, provaremos dois resultados da teoria de Análise Funcional em que o Lema de Zorn faz-se necessário, mais precisamente o Teorema de Hahn-Banach que é um dos teoremas centrais da Análise Funcional, e o fato que todo conjunto ortonormal de um espaço com produto interno está contido em um sistema ortonormal completo. Aproveitaremos para tecer comentários a respeito da equivalência ou não-equivalência destes resultados com o Lema de Zorn.

1 TEORIA AXIOMÁTICA DE CONJUNTOS

1.1 RESULTADOS INICIAIS

A base do estudo neste trabalho serão os conjuntos. Os conjuntos são entes matemáticos formados por elementos. A relação entre conjuntos e elementos é a relação de pertinência. Diz-se que x *pertence a* A, e escrevemos $x \in A$, quando x for elemento do conjunto A, e diz-se que x *não pertence a* A, e escrevemos $x \notin A$, quando x não for elemento de A.

Neste capítulo traremos alguns dos axiomas do sistema Zermelo-Fraenkel (ZF). Nesse processo de axiomatização denotaremos conjuntos por letras minúsculas e por letras maiúsculas. Usamos isso para tentar estabelecer entre elas uma relação de hierarquia, e para seguir as notações estabelecidas por [6]. Usaremos também algumas expressões lógicas para abreviar as correspondências em português. Veja a lista dessas expressões, e suas devidas correspondências:

1. $\forall x$ corresponde à "para todo conjunto x".
2. $\exists x$ corresponde à "existe um conjunto x".
3. $\Rightarrow$ corresponde à "implica", é utilizado para abreviar a condição de "se ... então ...".
4. $\Leftrightarrow$ corresponde à "se e somente se".

Na axiomática ZF, são tomados como noções primitivas, a ideia de conjunto e de pertinência descrita acima. A relação de pertinência é uma relação estabelecida entre conjuntos e elementos (que também são conjuntos). Faz-se necessário neste ponto definir uma relação entre conjuntos, neste sentido, temos a seguinte definição:

Definição 1.1. Sejam A e B conjuntos. Se todo elemento de A for elemento de B, diz-se que A é *subconjunto* de B, ou que A *está contido* em B, e escrevemos $A \subseteq B$.

Agora, precisamos introduzir a noção de igualdade entre conjuntos e, para isso, precisaremos do primeiro axioma do sistema ZF.

Axioma 1.2 (Axioma da Extensão)**.** *Dois conjuntos são iguais se, e somente se, eles têm exatamente os mesmos elementos.*

$$\forall A, \forall B[A = B \Leftrightarrow (\forall x(x \in A \Leftrightarrow x \in B))]$$

Este Axioma define a igualdade de conjuntos. A seguir apresentaremos axiomas que garantem a existência de alguns conjuntos.

Axioma 1.3 (Axioma do Conjunto Vazio). *Existe um conjunto que não tem nenhum elemento.*

$$\exists B(\forall x, x \notin B)$$

O conjunto que foi apresentado no Axioma anterior é chamado de *conjunto vazio* e denotado por $\emptyset$. Vejamos então que este conjunto de fato é único.

Proposição 1.4. *O conjunto vazio é único.*

Demonstração. Suponha que $\emptyset$ e $\emptyset'$ sejam conjuntos que cumprem o Axioma 1.3, isto é, para todo x, x não pertence a $\emptyset$ nem a $\emptyset'$. Logo, segue que,

$$\forall x, x \notin \emptyset \Leftrightarrow x \notin \emptyset'$$

Isto é,

$$\forall x, x \in \emptyset \Leftrightarrow x \in \emptyset'$$

Pelo Axioma 1.2, temos então que $\emptyset = \emptyset'$. □

Axioma 1.5 (Axioma dos Pares). *Dados dois conjuntos, existe um outro conjunto cujos elementos são exatamente os dois conjuntos dados.*

$$\forall u, \forall v \ \exists B[\forall x(x \in B \Leftrightarrow x = u \ \ ou \ x = v)]$$

Denotaremos por $\{X, Y\}$. O conjunto que tem como elementos os conjuntos X e Y, esse conjunto existe por conta do Axioma anterior. O Axioma da Extensão, ainda garante que $\{X, Y\} = \{Y, X\}$. É preciso todavia provar que este conjunto é único.

Proposição 1.6. *Sejam X e Y conjuntos. O conjunto que contém exatamente X e Y é único.*

Demonstração. Suponha que os conjuntos B e B' sejam conjuntos cujos elementos são exatamente os conjuntos X e Y. Logo, $\forall b(b \in B \Leftrightarrow b = X$ ou $b = Y)$, por outro lado, $\forall b(b \in B' \Leftrightarrow b = X$ ou $b = Y)$. Portanto, $b \in B \Leftrightarrow b \in B'$, e pelo Axioma 1.2, $B = B'$. □

Com estes axiomas, é possível construir uma série de conjuntos, mas estes têm no máximo dois elementos. Pode-se, por exemplo, construir o conjunto $\{\emptyset\}$, para isso basta tomar o conjunto vazio duas vezes no Axioma 1.5. Entretanto, até aqui, a Teoria dos Conjuntos fica restrita a conjuntos que têm no máximo dois elementos. Para expandir essa teoria é necessário novos axiomas.

Axioma 1.7 (Axioma da União)**.** *Dado um conjunto a, existe um conjunto cujos elementos são exatamente os elementos de a.*

$$\forall a \exists b\ [\forall x(x \in b \Leftrightarrow \exists y(y \in a \ \ e \ x \in y))]$$

O Axioma da Extensão (1.2), garante ainda que o conjunto cuja existência foi dada pelo Axioma anterior é único, e ele será denotado por $\bigcup a$. Este axioma garante condições importantes para a Teoria de Conjuntos, como a proposição seguinte.

Proposição 1.8. *Dados $a_1, a_2, \ldots, a_n$, existe um único conjunto que contém $a_1, a_2, \ldots, a_n$. O conjunto que satisfaz estas condições será denotado por $\{a_1, a_2, \ldots, a_n\}$.*

Demonstração. Do Axioma dos Pares (1.5), segue que o conjunto $\{a_1, a_2\}$ existe, isto é, para $n = 2$ a proposição é válida.

Suponha então que seja válida para n. Ou seja, o conjunto $\{a_1, a_2, \ldots, a_n\}$ existe e é único. Ao tomar este conjunto, juntamente com o conjunto $\{a_{n+1}\}$, pelo Axioma dos Pares (1.5) temos que o conjunto $\{\{a_1, a_2, \ldots, a_n\}, \{a_{n+1}\}\}$, existe. E pelo Axioma da União (1.7) existe $\bigcup\{\{a_1, a_2, \ldots, a_n\}, \{a_{n+1}\}\}$ e os elementos desse conjunto são os elementos de $\{a_1, a_2, \ldots, a_n\}$ e de $\{a_{n+1}\}$. Isto é $\bigcup\{\{a_1, a_2, \ldots, a_n\}, \{a_{n+1}\}\} = \{a_1, \ldots, a_{n+1}\}$, e este conjunto é único. $\square$

Axioma 1.9 (Axioma do Conjunto das Partes)**.** *Dado qualquer conjunto a, existe um conjunto cujos elementos são exatamente os subconjuntos de a.*

$$\forall a \exists B\ [\forall x(x \in B \Leftrightarrow x \subseteq a)]$$

O conjunto formado por todos os subconjuntos de A é chamado de *conjunto das partes de A*, e denotado por $\mathcal{P}(A)$. Como feito anteriormente, pode-se provar que o conjunto das partes é único, utilizando o Axioma 1.2.

O próximo axioma da axiomática ZF é tido como um dos mais importantes de toda a Teoria de Conjuntos. Para o entendimento deste axioma é necessário que o leitor tenha algum conhecimento de lógica proposicional e, para isso, aconselhamos a leitura de [8].

Nosso intuito é formalizar a ideia de fórmula, mas antes disso precisamos dar sentido para o conceito de variável proposicional, que pode ser entendida como uma sentença que pode ser classificada como verdadeira ou falsa. Além disso, precisamos também dos conectivos e quantificadores lógicos. Eles são:

I. Conectivos lógicos;

   - Binários: a conjunção ($\wedge$), designada por *e*, a disjunção ($\vee$), designada por *ou*, e as condicionais ($\Rightarrow$ e $\Leftrightarrow$) designadas por *implica* e *se e somente se*, respectivamente.

- Unitário: a negação ($\sim$), designada *não*.

II. Quantificadores lógicos;

- Existencial denotado por $\exists$, designado por *existe*.
- Universal denotado por $\forall$, designado por *para todo*.

Vamos formalizar o que será uma *fórmula*. Primeiro definimos as fórmulas mais simples, que são as chamadas *fórmulas atômicas*, elas são da forma $a \in B$ e $a = b$. Temos também que, se Ψ e Φ são fórmulas, então $(\sim \Psi), (\Psi \wedge \Phi), (\Psi \vee \Phi), (\Psi \Rightarrow \Phi)$ e $(\Psi \Leftrightarrow \Phi)$ são fórmulas. E ainda, se Ψ é uma fórmula e x uma variável proposicional, então $\forall x \Psi$ e $\exists x \Psi$ são fórmulas. Nesse sentido, as duas últimas fórmulas podem ser entendidas, respectivamente, como "para todo x vale Ψ na variável x" e "existe um x tal que Ψ vale em x".

Uma definição formal pode ser encontrada em [8, pag I-1].

Com o conhecimento das definições de fórmula e da notação lógica, segue o próximo axioma.

Axioma 1.10 (Axioma da Substituição)**.** *Para cada fórmula ϕ, vale o seguinte:*

$$\forall t_1, \ldots, \forall t_k, \forall C \exists B [\forall x (x \in B \Leftrightarrow x \in C \wedge \phi)$$

Em que ϕ é uma fórmula envolvendo $t_1, \ldots, t_k, x$ e que não contém B.

Em outras palavras, o Axioma da Substituição garante que para $t_1, \ldots, t_k$ e C, existe um conjunto B com elementos sendo os conjuntos x em C tais que ocorre ϕ. Dados $t_1, \ldots, t_k$ e C, o conjunto B é unicamente determinado, o que se prova pelo Axioma 1.2.

Este axioma ainda nos garante que,

$$\forall a, \forall c, \exists B, \forall x (x \in B \Leftrightarrow (x \in c) \wedge (x \in a)).$$

Este conjunto B é unicamente determinado, e ele é chamado de *interseção de a e c*. O conjunto B é denotado por $a \cap c$. O Axioma 1.10 ainda garante que

$$\forall A, \forall B, \exists S, \forall t (t \in S \Leftrightarrow (t \in A) \wedge (t \notin B)).$$

O conjunto S é chamado de *complementar de B em A*.

O Axioma 1.10 é responsável ainda por corrigir erros lógicos que poderiam surgir na Teoria de Conjuntos. Um desses erros é o conhecido Paradoxo de Russell [10] que questiona a existência de um conjunto que contenha todos os conjuntos. Na teoria ingênua dos conjuntos, elaborada por Cantor, esse paradoxo aceitava, ao mesmo tempo a existência e a não existência do conjunto de todos os conjuntos. O Paradoxo de Russel foi um dos motivadores para que se criasse uma axiomatização consistente para a Teoria de Conjuntos.

Teorema 1.11. *Não existe o conjunto de todos os conjuntos. Isto é,* $\mathcal{U} := \{A | A \text{ é conjunto}\}$*, não é conjunto.*

Demonstração. Suponha que $\mathcal{U}$ seja conjunto. Seja a fórmula $\phi(x) \Leftrightarrow x \notin x$. Pelo Axioma 1.10, $\exists B = \{x \in \mathcal{U} | \phi(x)\} = \{x \in \mathcal{U} | x \notin x\}$. Pela construção de B sabemos que,

$$B \in B \Leftrightarrow B \in \mathcal{U} \text{ e } B \notin B$$

. Como $B \in \mathcal{U}$, segue que,

$$B \in B \Leftrightarrow B \notin B$$

que é um absurdo. Logo, $\mathcal{U}$ não é conjunto. □

O último axioma da Teoria de Conjuntos que veremos é o Axioma da Escolha. Ele é o objeto central de estudo deste trabalho e será apresentado na próxima seção.

1.1.1 RELAÇÕES

Os axiomas apresentados nos permitem desenvolver o conceito de relações, mas para isso, precisamos primeiro de algumas definições preliminares.

Definição 1.12. Dados x e y, o conjunto $(x, y) = \{\{x\}, \{x, y\}\}$ é chamado de *par ordenado* de x e y.

A partir desta definição, é possível provar que $(x, y) = (u, v) \Leftrightarrow x = u$ e $y = v$. O que justifica o nome de par ordenado. É comum, em um par ordenado (x, y), dizer que x é a primeira coordenada, e que y é a segunda coordenada. Essa demonstração, assim como as demais demonstrações deste capítulo podem ser encontradas em [6].

Dados dois conjuntos A e B, pode-se definir um conjunto C, tal que

$$C = \{(x, y) | x \in A \text{ e } y \in B\}.$$

O conjunto C é chamado de *produto cartesiano de A e B* e é denotado por $A \times B$. Segue de [6, Corollary 3C] que $A \times B$ é conjunto.

Definição 1.13. Um conjunto R de pares ordenados, é chamado de *relação.*

Se $(a, b) \in R$, é usual escrever que aRb, e dizer que *a está relacionado com b.* A definição de relação é tecnicamente simples, mas, a partir dela surgem novos conceitos.

Definição 1.14. Dada uma relação R, dizemos que o conjunto $\{x | (x, y) \in R, \forall y\}$, é o *domínio* de R. E o conjunto $\{x | (y, x) \in R, \forall y\}$ é chamado de *imagem* de R.

Também em [6] é possível encontrar a prova de que o domínio e a imagem de uma relação são, de fato, conjuntos.

Definição 1.15. Chamaremos de *função*, às relações em que cada elemento do domínio estiver relacionado com um, e somente um, elemento do contradomínio.

Definição 1.16. Seja R uma relação em $A \times A$, dizemos que R é,

(i.) *reflexiva*, se $(a, a) \in R, \forall a \in A$.

(ii.) *simétrica*, se $(a, b) \in R \Rightarrow (b, a) \in R, \forall a \in A, b \in A$.

(iii.) *antissimétrica*, se $(a, b) \in R$ e $(b, a) \in R \Rightarrow a = b$.

(iv.) *transitiva*, se $(a, b) \in R$ e $(b, c) \in R \Rightarrow (a, c) \in R \ \ \forall a, b, c \in A$.

Se R for reflexiva, simétrica e transitiva, R é chamada de *relação de equivalência*. Por outro lado, se R for reflexiva, antissimétrica e transitiva, R será chamada de *relação de ordem parcial*.

Dessa forma, a relação menor ou igual ($\leq$) é um exemplo de relação de ordem parcial no conjunto dos reais. Sendo assim, é comum se referir a uma relação de ordem genérica pelo símbolo $\leq$.

Podemos ainda definir uma *relação de ordem total* em um conjunto A, que é uma relação de ordem parcial em A, tal que para todos $x, y \in A$, vale que $x \leq y$ ou $y \leq x$ (neste caso dizemos que A é *totalmente ordenado*). Se um conjunto A tiver uma ordem total e se todo subconjunto de A tiver menor elemento, isto é, se A tiver um elemento x menor ou igual a todos os outros elementos do conjunto, dizemos que A tem uma *boa ordem*, ou que A é *bem ordenado* e dizemos x é o *mínimo* de A, ou simplesmente escrevemos, $\min A = x$.

Com estes conceitos podemos introduzir a seguinte definição que será essencial no decorrer deste trabalho.

Definição 1.17. Seja P um conjunto com ordem parcial e $Q \subseteq P$. Dizemos que um elemento $p \in P$ é *cota superior* de Q se para todo $q \in Q$, valer que $q \leq p$. Dizemos um $m \in P$ é um elemento maximal de P se para todo $p \in P$ tal que $p \geq m$ tivermos que $p = m$.

1.2 EQUIVALÊNCIAS DO LEMA DE ZORN EM TEORIA DE CONJUNTOS

Neste capítulo mostraremos a equivalência mais conhecida do Axioma da Escolha (AE), que se dá com o Princípio de Boa Ordem (PBO) e com o Lema de Zorn (LZ).

O Axioma da Escolha, diz de forma não rigorosa que, dada qualquer coleção não vazia de conjuntos não vazios é possível formar um conjunto com um, e apenas um, elemento de cada um dos conjuntos. Para dar rigor a esta ideia, é preciso definir *função escolha.*

Definição 1.18. Uma função f é uma *função escolha* para o conjunto A, se o domínio de f é $A - \{\emptyset\}$, e $f(a) \in a$ para todo $a \in A - \{\emptyset\}$.

De posse dessa definição, é possível enunciar o Axioma da Escolha.

Axioma 1.19 (Axioma da Escolha)**.** *Todo conjunto admite função escolha.*

Agora apresentaremos o Princípio de Boa Ordem e o Lema de Zorn, a fim de provar a equivalência deles com o Axioma da Escolha.

Teorema 1.20 (Princípio de Boa Ordem)**.** *Todo conjunto não vazio pode ser bem ordenado.*

Definição 1.21. Uma *cadeia* é um subconjunto totalmente ordenado de um conjunto parcialmente ordenado.

Lema 1.22 (Lema de Zorn)**.** *Se A é um conjunto não vazio e parcialmente ordenado, tal que toda cadeia tem cota superior. Então A tem elemento maximal.*

Para provar a equivalência proposta é necessário que ela seja dividida em três partes: primeiro provaremos que o Princípio de Boa Ordem(PBO) implica o Axioma da Escolha (AE), depois que o Lema de Zorn (LZ) implica no Princípio de Boa Ordem, e por fim que o Axioma da Escolha implica no Lema de Zorn.

1.2.1 PBO $\Rightarrow$ AE

Teorema 1.23. *Princípio de Boa Ordem $\Rightarrow$ Axioma da Escolha.*

Demonstração. Seja A um conjunto não vazio. Segue do Princípio de Boa Ordem que $\bigcup A$ pode ser bem ordenado. Como cada elemento de A é um subconjunto de $\bigcup A$, considere a função

$$f : A - \{\emptyset\} \to \bigcup A$$

dada por $f(a) =$"menor elemento de a", para cada $a \in A - \{\emptyset\}$. Então f é uma função escolha. $\square$

1.2.2 LZ $\Rightarrow$ PBO

Provaremos agora a implicação que diz que o Lema de Zorn implica o Princípio de Boa Ordem. O objetivo a partir daqui será então provar que, dado um conjunto não-vazio X, esse conjunto pode ser totalmente ordenado. Antes de prosseguir à demonstração, é preciso definir um conjunto que será essencial nesse processo.

Definição 1.24. Seja $X \neq \emptyset$. Definimos

$$\mathfrak{F} = \{A \subseteq X : A \neq \emptyset \text{ e } A \text{ é bem ordenado}\}.$$

É claro que $\mathfrak{F}$ é não-vazio, visto que se tomarmos $a \in X$, e $A = a$, temos que $A \subseteq X$, e mais, A é bem ordenado (uma vez que é um conjunto unitário).

Seguimos com a estratégia da demonstração que será provar que X é elemento maximal de $\mathfrak{F}$, e dessa forma é bem ordenado. Entretanto, para tal resultado ser válido, é necessário criar uma ordem em $\mathfrak{F}$, e isso será feito a seguir:

Lema 1.25. *Sejam $(A, \leq_A)$ e $(B, \leq_B)$, conjuntos com suas devidas ordens. Definimos em $\mathfrak{F}$ a relação $(A, \leq_A) \leq_{\mathfrak{F}} (B, \leq_B)$, se:*

i. $A \subseteq B$.

ii. *Se $x, y \in A$, então $x \leq_A y \Leftrightarrow x \leq_B y$.*

iii. *Se $x \in A$ e $y \in B - A$, então $x \leq_B y$.*

A relação $\leq_{\mathfrak{F}}$ é uma relação de ordem parcial em $\mathfrak{F}$.

Demonstração. O objetivo é mostrar que esta relação torna $\mathfrak{F}$ parcialmente ordenado. Isto é, a relação $\leq_{\mathfrak{F}}$ é reflexiva, antissimétrica e transitiva.

- Reflexiva:
  $(i.) A \subseteq A$;
  $(ii.)$ Se $x, y \in A$, é evidente que $x \leq_A y \Leftrightarrow x \leq_A y$;
  $(iii.)$ Não acontece, uma vez que $A - A = \emptyset$.

- Antissimétrica: Se $(A, \leq_A) \leq_{\mathfrak{F}} (B, \leq_B)$ e $(B, \leq_B) \leq_{\mathfrak{F}} (A, \leq_A)$, então, de $(i.)$ temos que $A \subseteq B$ e $B \subseteq A$, mas isso implica que $A = B$, o que retorna no caso anterior.

- Transitiva: Se $(A, \leq_A) \leq_{\mathfrak{F}} (B, \leq_B)$ e $(B, \leq_B) \leq_{\mathfrak{F}} (C, \leq_C)$, então:
  $(i.)$ $A \subseteq B$ e $B \subseteq C$, logo pela transitividade de "$\subseteq$", $A \subseteq C$;
  $(ii.)$ Se $x, y \in A$, então, $x \leq_A y \Leftrightarrow x \leq_B y$, mas como $x \leq_B y$, então $x \leq_C y$;
  $(iii.)$ Já temos que $A \subseteq B \subseteq C$. Se $x \in A$ e $y \in C - A$, então $x \in C$ e $y \notin A$. Agora existem dois casos $y \in B$, ou $y \notin B$. Consideremos $y \in B$, logo $y \in B - A$, o que implica que se $x \leq_B y$, então $x \leq_C y$. Consideremos agora que $y \notin B$, então $y \in C - B$, o que implica que $x \leq_C y$.

$\square$

O próximo passo para prosseguir à demonstração é mostrar que toda cadeia em $\mathfrak{F}$ admite cota superior. A partir daí, teremos o conjunto $\mathfrak{F}$ parcialmente ordenado, tal que toda cadeia tem cota superior, assim será possível utilizar o Lema de Zorn (1.22), para provarmos a implicação desejada.

Lema 1.26. *Seja I um conjunto de índices e $C = \{(A_i, \leq_{A_i}) : i \in I\}$ uma cadeia em $\mathfrak{F}$. Então C admite cota superior.*

Demonstração. Seja $A = \bigcup C$. Se $x, y \in A$, então $x \in A_i$ e $y \in A_j$. Como C é totalmente ordenado, segue que, $A_i \leq_{\mathfrak{F}} A_j$, ou $A_j \leq_{\mathfrak{F}} A_i$, sem perda de generalidade, suponha $A_i \leq_{\mathfrak{F}} A_j$. Dessa forma temos que $x, y \in A_j$.

Agora criemos uma ordem para o conjunto A. Dados $x, y \in A$, defina $x \leq_A y$, se $x \leq_{A_i} y$ e $x, y \in A_i$, para algum $i \in I$. É claro que a relação "$\leq_A$" é uma relação de ordem total, uma vez que depende unicamente das relações de ordem dos conjuntos A_k, que estão contidos em A.

Provemos que $(A, \leq_A)$ é bem ordenado, e desse resultado será possível tirar que $(A, \leq_A) \in \mathfrak{F}$. Primeiro é preciso observar que para todo $i, j \in I$, $\min A_i = \min A_j$. Como $A_i \leq_{\mathfrak{F}} A_j$, então $A_i \subseteq A_j$, criando assim dois casos possíveis. No primeiro, $\min A_j \in A_i$, e no segundo, $\min A_j \in A_j - A_i$. Analisemos separadamente cada uma dos casos:

(a) Se $\min A_j \in A_i$, pelo item *ii.* da definição de "$\leq_{\mathfrak{F}}$" segue que $\min A_i \leq_{A_j} \min A_j$, então, $\min A_i = \min A_j$.

(b) Se $\min A_j \in A_j - A_i$,como $\min A_i \in A_j$ pelo item *iii.* da definição de "$\leq_{\mathfrak{F}}$" segue que $\min A_i \leq_{A_j} \min A_j$, então, $\min A_i = \min A_j$.

Segue da definição de "$\leq_A$", que o mínimo de $(A_i, \leq_{A_i})$ é o mínimo de $(A, \leq_A)$.

Vamos mostrar agora que o conjunto $(A, \leq_A)$ é bem ordenado. Seja B um subconjunto não vazio de A e $J = \{i \in I : B \cap A_i \neq \emptyset\}$. Logo, é possível deduzir que $\min(B \cap A_i, \leq_{A_i}) = \min(B \cap A_j, \leq_{A_j})$, para todo $i, j \in J$, podemos deduzir também que $\min(B \cap A_i, \leq_{A_i})$ é o menor elemento de B. Logo, $(A, \leq_A)$ é bem ordenado.

Queremos mostrar agora que $(A, \leq_A)$ é cota superior para C. Seja $k \in I$, segue que:

I. $A_k \subseteq A$, pela definição de A;

II. Se $x, y \in A_k$, então $x \leq_{A_k} y \Leftrightarrow x \leq_A y$, pela definição de "$\leq_A$";

III. Se $x \in A_k$ e $y \in A - A_k$, então existe $l \in I$ tal que $y \in A_l$, e portanto $A_l \not\subseteq A_k$, o que implica que $A_l \not\leq_{\mathfrak{F}} A_k$. Como C é totalmente ordenado, $A_k \leq_{\mathfrak{F}} A_l$, então $A_k \subseteq A_l$.

Como $x \in A_k$ e $y \in A_l - A_k$, pelo item *iii.* da definição de "$\leq_{\mathfrak{F}}$", segue que $x \leq_{A_l} y$, e pela definição de "$\leq_A$", segue que $x \leq_A y$.

Os itens *I.*, *II.* e *III.* correspondem, respectivamente, aos itens *i.*, *ii.* e *iii.* da definição de "$\leq_{\mathfrak{F}}$". Dessa forma, temos que $A_k \leq_{\mathfrak{F}} A$, para todo $k \in I$. Ou seja, $(A, \leq_A)$ é uma cota superior para C, e assim, C admite cota superior. □

Com os lemas enunciados criamos o conjunto $\mathfrak{F}$, parcialmente ordenado (1.25), em que toda cadeia tem cota superior (1.26). Dessa forma, podemos usar o Lema de Zorn (1.22), para provar a implicação desejada, e podermos mostrar que o conjunto X, apresentado no começo desta seção é um conjunto bem ordenado.

Teorema 1.27. *Lema de Zorn $\Rightarrow$ Princípio da Boa Ordem.*

Demonstração. Partindo dos lemas passados, pelo Lema de Zorn, o conjunto $\mathfrak{F}$ tem elemento maximal, digamos que este elemento é $(M, \leq_M)$. Queremos provar que $M = X$. Se $M \neq X$, tome $n \in X - M$, e defina o conjunto $N = M \cup \{n\}$, bem como a relação "$\leq_N$" dada por:

Para $x, y \in M$, $x \leq_N y \Leftrightarrow x \leq_M y$ e para $x \in M$, $x \leq_N n$.

Assim, $(N, \leq_N) \in \mathfrak{F}$, e $(M, \leq_M) \leq_{\mathfrak{F}} (N, \leq_N)$, o que é um absurdo, já que M é maximal.

Logo, $M = X$. Considerando $\leq_X$ como $\leq_M$, temos que o conjunto $(X, \leq_X)$ é bem ordenado. □

1.2.3 AE $\Rightarrow$ LZ

Para terminar a demonstração da equivalência original é necessário provar, que o Axioma da Escolha implica no Lema de Zorn. De outra forma, queremos mostrar que um conjunto não vazio X, parcialmente ordenado, em que toda cadeia admite cota superior, tem pelo menos um elemento maximal.

Para demonstrar a implicação desejada é necessário começarmos definindo o conjunto $\mathfrak{X} = \{C \subseteq X : C \text{ é cadeia}\}$. A partir de então provar que ele tem elemento maximal. Feito isso será simples mostrar que X tem elemento maximal. Para isso, começamos criando uma relação de ordem para $\mathfrak{X}$. É intuitivo que uma relação de ordem que irá funcionar bem neste caso é a relação de inclusão de conjuntos. Formalizemos esta ideia:

Proposição 1.28. *A relação $C_1 \leq_{\mathfrak{X}} C_2 \Leftrightarrow C_1 \subseteq C_2$ é relação de ordem parcial em $\mathfrak{X}$.*

Demonstração. Queremos mostrar que $\leq_{\mathfrak{X}}$ é relação de ordem parcial em $\mathfrak{X}$. Para isso, devemos mostrar que essa relação é reflexiva, antissimétrica e transitiva.

- Reflexiva: Seja $C_1 \in \mathfrak{X}$. Por definição $C_1 \subseteq C_1$, logo $C_1 \leq_{\mathfrak{X}} C_1$.
- Antissimétrica: Sejam $C_1, C_2 \in \mathfrak{X}$, tais que, $C_1 \leq_{\mathfrak{X}} C_2$ e $C_2 \leq_{\mathfrak{X}} C_1$. Então $C_1 \subseteq C_2$, e $C_2 \subseteq C_1$, logo, pela definição de "$\subseteq$", segue que $C_1 = C_2$.
- Transitiva: Sejam C_1, C_2 e $C_3 \in \mathfrak{X}$, tais que $C_1 \leq_{\mathfrak{X}} C_2$, e $C_2 \leq_{\mathfrak{X}} C_3$. Então, temos que $C_1 \subseteq C_2$, e $C_2 \subseteq C_3$, mas pela transitividade de "$\subseteq$", segue que $C_1 \subseteq C_3$. Então, $C_1 \leq_{\mathfrak{X}} C_3$.

□

Proposição 1.29. *Sejam $C \in \mathfrak{X}$ e $C' = \{x \in X : C \cup \{x\} \in \mathfrak{X}\}$. Então C é maximal em $\mathfrak{X}$, se e somente se, $C = C'$.*

Demonstração. Consideramos que C é maximal em $\mathfrak{X}$ como hipótese, e vejamos que $C = C'$. Suponha que $C \neq C'$. Como $C \subseteq C'$, então existe $a \in C'$ tal que $a \notin C$. Pela definição de C' segue que $C \cup \{a\} \in \mathfrak{X}$ e como $C \subset C \cup \{a\}$, segue do fato de C ser maximal em $\mathfrak{X}$ que $C = C \cup \{a\}$, o que é uma contradição.

Suponha como hipótese agora que $C = C'$ e vejamos que C é maximal em $\mathfrak{X}$. Seja $A \in \mathfrak{X}$ tal que $C \leq_{\mathfrak{X}} A$, isto é $C' = C \subseteq A$. Como A é totalmente ordenado, é claro que todo subconjunto de A é totalmente ordenado, em particular, $C \cup \{x\}$ é totalmente ordenado, para todo $x \in A$. Logo, $C \cup \{x\} \in \mathfrak{X}$, para todo $x \in A$. Logo $x \in C'$, para todo $x \in A$. Assim, $A \subseteq C'$ e, portanto, $C' = A$. Logo $C = A$ e isto prova que C é maximal.

□

Pelo Axioma da Escolha, existe uma função escolha para o conjunto $\mathcal{P}(X)$. Chamaremos de f esta função, ou seja, $f : \mathcal{P}(X) - \{\emptyset\} \to X$, e mais, $f(A) \in A$, para todo $A \in \mathcal{P}(X) - \{\emptyset\}$. Vamos agora definir uma função g de $\mathfrak{X}$ em $\mathfrak{X}$, que será usada muitas vezes, da seguinte forma:

$$g(C) = \begin{cases} C, & \text{se } C = C' \\ C \cup f(C' - C), & \text{se } C \neq C' \end{cases}.$$

A função g está bem definida, pois se $C \neq C'$, então $C' - C \neq \emptyset$ e $f(C' - C) \in C' - C$, pois f é função escolha. Logo, pela definição de C', $C \cup f(C' - C) \in \mathfrak{X}$. Além disso, da Proposição 1.29, segue que C é maximal em $\mathfrak{X}$, se e somente se, $g(C) = C$.

Definição 1.30. Diremos que $\mathcal{T} \subseteq \mathfrak{X}$ é uma *torre* se:

i. $\emptyset \in \mathcal{T}$;

ii. Se $C \in \mathcal{T}$, então $g(C) \in \mathcal{T}$;

iii. Se $\mathcal{C}$ é uma cadeia em $\mathcal{T}$, então $\bigcup \mathcal{C} \in \mathcal{T}$.

Proposição 1.31. *(a) O conjunto $\mathfrak{X}$ é uma torre.*

(b) A interseção de uma família de torres em $\mathfrak{X}$ também é uma torre.

Demonstração. (a) Vejamos que $\mathfrak{X}$ satisfaz às condições da definição de torre:

i. É claro que $\emptyset \in \mathfrak{X}$;

ii. Se $C \in \mathfrak{X}$, então C é cadeia em X, logo, $C \cup f(C' - C)$ também é cadeia em X, e portanto $C \cup f(C' - C) \in \mathfrak{X}$, ou seja, $g(C) \in \mathcal{T}$;

iii. Seja $\mathcal{C}$ uma cadeia em $\mathfrak{X}$ e vejamos que $\bigcup \mathcal{C} \in \mathfrak{X}$. Sejam $x, y \in \bigcup \mathcal{C}$. Então existem $C_1, C_2 \in \mathcal{C}$ tais que $x \in C_1$ e $y \in C_2$. Como $\mathcal{C}$ é uma cadeia em $\mathfrak{X}$, segue que $C_1 \subset C_2$ ou $C_2 \subset C_1$. Assim, $x, y \in C_1$ ou $x, y \in C_2$. Como $C_1, C_2 \in \mathfrak{X}$, segue que $x \leq y$ ou $y \leq x$. Logo $\bigcup \mathcal{C}$ é cadeia.

(b) Seja $\{\mathcal{T}_i : i \in I\}$ uma família de torres em $\mathfrak{X}$ e vejamos que a interseção delas, que denotaremos por $\bigcap_{i \in I} \mathcal{T}_i$, ainda é uma torre:

i. Pela definição de torre, temos que $\emptyset \in \mathcal{T}_i$, para $i \in I$. Dessa forma, temos que $\emptyset \in \bigcap_{i \in I} \mathcal{T}_i$.

ii. Seja $C \in \bigcap_{i \in I} \mathcal{T}_i$. Como cada $\mathcal{T}_i$ é torre e $C \in \mathcal{T}_i$ segue que $g(C) \in \mathcal{T}_i$, para todo $i \in I$. Portanto, $g(C) \in \bigcap_{i \in I} \mathcal{T}_i$.

iii. Seja $\mathcal{C}$ uma cadeia em $\bigcap_{i \in I} \mathcal{T}_i$. Como $\mathcal{C} \subset \bigcap_{i \in I} \mathcal{T}_i$, segue que $\mathcal{C} \subset \mathcal{T}_i$ e como cada $\mathcal{T}_i$ é torre, segue que $\bigcup \mathcal{C} \in T_i$, para todo $i \in I$. Logo $\bigcup \mathcal{C} \in \bigcap_{i \in I} \mathcal{T}_i$.

$\square$

Segue diretamente desta proposição que a intersecção de todas as torres de um conjunto qualquer é a menor torre deste conjunto. Denotaremos por $\mathcal{T}_0$ a intersecção de todas as torres de $\mathfrak{X}$. Queremos mostrar que $\mathcal{T}_0$ é uma cadeia em $\mathfrak{X}$. Esta é a parte mais trabalhosa desta demonstração. Começaremos definindo quando que um conjunto em $\mathcal{T}_0$ é comparável:

Definição 1.32. Diremos que $A \in \mathcal{T}_0$ é *comparável* se dado $B \in \mathcal{T}_0$, tem-se $A \subseteq B$ ou $B \subseteq A$.

Com base nesta definição, para mostrar que $\mathcal{T}_0$ é cadeia, basta provar que cada elemento A de $\mathcal{T}_0$ é comparável. Ao mostrar que todos os elementos de $\mathcal{T}_0$ são comparáveis, estaremos estabelecendo uma relação de ordem em $\mathcal{T}_0$, que é dada por $\subseteq$.

Para mostrar que os elementos A de $\mathcal{T}_0$ são comparáveis, basta provar que o conjunto $\mathcal{E}$ dos elementos comparáveis de $\mathcal{T}_0$ formam uma torre, pois, neste caso, essa torre estaria contida em $\mathcal{T}_0$, e assim estaria também contida em $\mathfrak{X}$. Mas, como $\mathcal{T}_0$ é a menor torre de $\mathfrak{X}$, é claro que a torre formada pelos elementos comparáveis de $\mathcal{T}_0$, tem que ser o próprio $\mathcal{T}_0$. Antes de partir a demonstração, mostremos o seguinte lema:

Lema 1.33. *Sejam A um elemento comparável de $\mathcal{T}_0$ fixado e*

$$U = \{B \in \mathcal{T}_0 : B \subseteq A \text{ ou } g(A) \subseteq B\}.$$

Então $U = \mathcal{T}_0$.

Demonstração. Primeiro devemos mostrar que U é torre. Vamos, dessa forma, provar cada um dos itens da Definição 1.30.

i. É fato que $\emptyset \in U$, pois o conjunto vazio está contido em qualquer conjunto A fixado e $\emptyset \in \mathcal{T}_0$;

ii. Queremos mostrar agora que se $\mathcal{C}$ é uma cadeia em U, então $\bigcup \mathcal{C} \in U$. De fato, se $B \subseteq A$, para todo $B \in \mathcal{C}$, então $\bigcup \mathcal{C} \subseteq A$. Agora, se existir $B_0 \in \mathcal{C}$, tal que $B_0 \nsubseteq A$, como $B_0 \in U$ segue que $g(A) \subseteq B_0 \subseteq \bigcup \mathcal{C}$.

iii. Agora falta provar que se $b \in U$, então $g(B) \in U$. Temos três possibilidades:

   (a) Se $B \subsetneq A$, vejamos que $g(B) \subseteq A$. Como $\mathcal{T}_0$ é torre, $g(B) \in \mathcal{T}_0$. Como A é comparável, temos que $g(B) \subseteq A$ ou $A \subsetneq g(B)$. Mas $A \subsetneq g(B)$ não ocorre, pois $B \subsetneq A$ e, pela definição de g, $g(B) = B$ ou $g(B)$ é a união de B com um ponto. Assim, $g(B) \subseteq A$ e daí $g(B) \in U$.

   (b) Se $B = A$, então $g(B) = g(A)$, ou seja, $g(A) \subseteq g(B)$ e, portanto, $g(B) \in U$.

   (c) Se $A \subsetneq B$, então $g(A) \subseteq B$, mas $B \subseteq g(B)$, portanto, $g(B) \in U$.

Portanto, U é torre em $\mathcal{T}_0$ e $U \subseteq \mathcal{T}_0$. Mas definimos $\mathcal{T}_0$ como a menor das torres de $\mathfrak{X}$, logo, $U = \mathcal{T}_0$. □

Teorema 1.34. *Os elementos comparáveis de $\mathcal{T}_0$, formam uma torre.*

Demonstração. Como no lema anterior, provaremos que $\mathcal{T}_0$ cumpre os três itens da definição de torre (1.30).

i. Claramente $\emptyset$ é um conjunto comparável, pois $\emptyset \in \mathcal{T}_0$ e $\emptyset \subseteq A$, para todo A elemento de $\mathcal{T}_0$.

ii. Temos que provar agora que se A é um conjunto comparável de $\mathcal{T}_0$, então $g(A)$ é também comparável. Conforme já fizemos na demonstração do lema anterior, segue que se $B \in \mathcal{T}_0$ e $B \subsetneq A$, então $g(B) \subseteq A$. Todavia, pelo Lema 1.33, $B \subseteq A$ ou $g(A) \subseteq B$. Portanto, $g(A)$ é comparável.

iii. Seja $\mathcal{C}$ uma cadeia de conjuntos comparáveis de $\mathcal{T}_0$ e B um elemento de $\mathcal{T}_0$. Teremos aqui dois subcasos:

(a) Se $C \subseteq B$, para todo $C \in \mathcal{C}$, então $\bigcup \mathcal{C} \subseteq B$.

(b) Se existir $C_0 \in \mathcal{C}$, tal que $C_0 \nsubseteq B$, então $B \subseteq C_0$, logo, $B \subseteq \bigcup \mathcal{C}$.

Desta forma, provamos que se $\mathcal{C}$ é uma cadeia de conjuntos comparáveis de $\mathcal{T}_0$, então $\bigcup \mathcal{C}$ é conjunto comparável de $\mathcal{T}_0$.

□

Como os elementos de $\mathcal{T}_0$ são todos comparáveis, segue que $\mathcal{T}_0$ é uma cadeia em $\mathfrak{X}$. Defina $A_0 := \bigcup \mathcal{T}_0$. Como $\mathcal{T}_0$ é torre, segue que $A_0 \in \mathcal{T}_0$, e mais, $g(A_0) \in \mathcal{T}_0$. Portanto $g(A_0) = A_0$. Criamos então um A_0 que é maximal em $\mathfrak{X}$. Como $\mathcal{T}_0 \subset \mathfrak{X}$ e $A_0 \in \mathcal{T}_0$, então $A_0 \in \mathfrak{X}$. Logo A_0 é cadeia em X.

Teorema 1.35. *$AE \Rightarrow LZ$.*

Demonstração. Relembre que o conjunto X é não vazio, parcialmente ordenado e tal que toda cadeia em X tem cota superior. Vejamos que X tem elemento maximal. Como A_0 é cadeia em X, então A_0 é limitada superiormente, ou seja, existe $m \in X$, tal que $m \leq a$, para todo $a \in A_0$. Como A_0 é um elemento maximal de $\mathfrak{X}$, segue que $m \in A_0$, pois como $A_0 \cup \{m\} \in \mathfrak{X}$, segue que

$$A_0 = A_0' = \{x \in X : A_0 \cup \{x\} \in \mathfrak{X}\}.$$

Afirmamos que m é maximal em X. Seja $n \in X$, tal que $m \leq n$. Como A_0 é cadeia maximal, segue que $n \in A_0$. Logo $n \leq m$, dessa forma, $n = m$. □

2 ÁLGEBRA LINEAR

2.1 RESULTADOS INICIAIS

Para este capítulo são necessárias algumas definições de Álgebra Linear. Na sequência serão, apresentadas as definições centrais deste tópico, as demais definições podem ser encontradas em [5].

Ao longo deste capítulo, usaremos $\mathbb{K}$ para denotar um corpo qualquer.

Definição 2.1. Sejam V um $\mathbb{K}$-espaço vetorial e $B \subset V$. Dizemos que B é um *conjunto gerador* de V (ou que B *gera* V), se cada elemento de V se escreve como combinação linear de um número finito de elementos de B.

Se um espaço vetorial tem como gerador um conjunto finito, dizemos que este espaço é *finitamente gerado*.

Definição 2.2. Sejam V um $\mathbb{K}$-espaço vetorial e B um subconjunto de V. Dizemos que B é *linearmente independente* (LI), se para todo $v_i \in B$ e $\alpha_i \in \mathbb{K}$, com $i = 1, \ldots, n$, vale que

$$\alpha_1 \cdot v_1 + \cdots + \alpha_n \cdot v_n = 0 \Rightarrow \alpha_1 = \cdots = \alpha_n = 0.$$

Definição 2.3. Seja V um $\mathbb{K}$-espaço vetorial. Dizemos que um subconjunto $B \subset V$ é uma *base* de V se for gerador de V e linearmente independente.

Em cursos introdutórios de Álgebra Linear se demonstra que todo espaço vetorial finitamente gerado tem base, mas não se prova que esse resultado vale num contexto mais amplo, ou seja, com base infinita enumerável ou infinita não enumerável.

Neste capítulo provaremos este resultado e, mais ainda, provaremos que a existência de base é equivalente ao Axioma da Escolha. Começamos demonstrando o caso finitamente gerado.

Lema 2.4. *Seja V um $\mathbb{K}$-espaço vetorial. Seja $B = \{v_1, \ldots, v_m\}$ um conjunto LI em V. Se existir $v \in V$ que não seja combinação linear dos elementos de B, então $\{v_1, \ldots, v_m, v\}$ é LI.*

Demonstração. Sejam $\alpha_1, \ldots, \alpha_m, \alpha_{m+1} \in \mathbb{K}$, tais que $\alpha_1 v_1 + \cdots + \alpha_m v_m + \alpha_{m+1} v = 0$.

Se $\alpha_{m+1} \neq 0$, então podemos escrever

$$v = -\frac{\alpha_1}{\alpha_{m+1}} v_1 - \cdots - \frac{\alpha_m}{\alpha_{m+1}} v_m$$

Ou seja, v é uma combinação linear de elementos de B, o que é uma contradição com as hipóteses. Então $\alpha_{m+1} = 0$, e, portanto, $\alpha_1 v_1 + \cdots + \alpha_m v_m = 0$. Como B é LI, segue que $\alpha_1 = \cdots = \alpha_m = 0$. Portanto, $\{v_1, \ldots, v_m, v\}$ é LI. □

Teorema 2.5. *Todo espaço vetorial não nulo finitamente gerado possui uma base.*

Demonstração. Seja V um espaço vetorial não nulo, finitamente gerado sobre $\mathbb{K}$. Então V possui um conjunto gerador finito, digamos com $m \leq 1$ elementos. Agora, seja $v_1 \in V$ um vetor não nulo. Então, $B_1 = \{v_1\}$ é LI. Se B_1 gerar V, então B_1 é uma base de V. Caso contrário, existe $v_2 \in V$ que não é um múltiplo de v_1. Pelo Lema 2.4, $B_2 = \{v_1, v_2\}$ é LI. Novamente, se B_2 gera V, então B_2 é uma base de V. Caso contrário existe $v_3 \in V$ tal que $B_3 = \{v_1, v_2, v_3\}$ é LI.

Podemos então repetir esse procedimento e chegaremos ou a uma base de V, ou construiremos conjuntos LI em V arbitrariamente grandes. O segundo caso não é possível, pois todo conjunto LI de V tem no máximo m elementos. □

A demonstração anterior só foi possível, pois o espaço V era finitamente gerado. Mas o resultado é mais geral, como será provado em seguida.

2.2 EQUIVALÊNCIA DO LEMA DE ZORN EM ÁLGEBRA LINEAR

Começamos provando a condição necessária, isto é, que ao assumirmos o Lema de Zorn, temos por consequência que todo espaço vetorial tem base. Este teorema é geralmente apresentado em cursos de Álgebra Linear avançados.

2.2.1 LZ $\Rightarrow$ TODO ESPAÇO VETORIAL TEM BASE

Teorema 2.6. *LZ $\Rightarrow$ todo espaço vetorial tem base.*

Demonstração. Sejam $V \neq \{0\}$ um espaço vetorial sobre um corpo $\mathbb{K}$ e $X \subseteq V$ um conjunto linearmente independente (LI). Defina $W = \{Y \subseteq V : Y \text{ é LI e } X \subseteq Y\}$. Como $X \in W$, temos que $W \neq \emptyset$. Considere em W a ordem parcial dada pela inclusão ($\subseteq$). Seja $\mathcal{C}$ uma cadeia em W. Vejamos que $Z = \bigcup \mathcal{C}$ é uma cota superior para a cadeia $\mathcal{C}$. Claramente $Z \subseteq V$, $X \subseteq Z$ e $Y \subseteq Z$, para todo $Y \in \mathcal{C}$. Dessa forma, basta provar que Z é LI. Sejam, $v_1, v_2, \ldots, v_n \in Z$. Então existem $Y_1, Y_2, \ldots, Y_n \in \mathcal{C}$ tais que $v_i \in Y_i$, para todo $i = 1, \ldots, n$. Como $\mathcal{C}$ é totalmente ordenado, segue que $\{Y_1, \ldots, Y_n\}$ pelo fato de que em qualquer conjunto finito, toda relação de ordem total é bem ordenada, este conjunto tem um maior elemento, digamos que seja Y_j. E, $\{v_1, \ldots, v_n\}$ é LI. Daí, Z é cota superior para a cadeia $\mathcal{C}$. Pelo Lema de Zorn (1.22), W tem elemento maximal, digamos B. Assim, $B \in W$, isto é, $B \subseteq V$, B é LI e $X \subseteq B$, além disso, B não está contido propriamente em nenhum elemento de W. Suponha que B não gera V. Então existe um $v \in V$ que não é combinação linear de vetores de B. Assim, $B \cup \{v\}$ é um subconjunto de V, $X \subseteq B \cup \{v\}$, e $B \cup \{v\}$ é LI. Então, $B \cup \{v\} \in W$ e $B \subsetneq B \cup \{v\}$, pois $v \notin B$. Essa última afirmação é um absurdo pois B é maximal. Logo B gera V e, portanto, B é base de V contendo X. □

2.2.2 TODO ESPAÇO VETORIAL TEM BASE $\Rightarrow$ LZ

Veremos agora a demonstração da condição suficiente. Ou seja, vamos mostrar que a existência de base em espaços vetoriais implica no Axioma da Escolha (logo implica no Lema de Zorn). Esse resultado foi provado por Blass [3] em 1984. Para essa demonstração precisaremos de alguns fatos preliminares.

Primeiramente, cabe destacarmos que não provaremos diretamente que a existência de base em espaços vetoriais implica no Axioma da Escolha. Na verdade provaremos que a existência de base em espaços vetoriais implica no seguinte axioma:

Axioma 2.7 (Axioma da Múltipla Escolha)**.** *Para toda família S de conjuntos não vazios, existe uma função $f\colon S \to \mathcal{P}(\bigcup S)$ tal que $F(A)$ é um subconjunto finito de A, para cada $A \in S$*

É um resultado conhecido da Teoria de Conjuntos que, em ZF, o Axioma da Múltipla Escolha implica no Axioma da Escolha. Não faremos esta demonstração no texto, mas os detalhes podem ser encontrados em [7, Theorem 9.1].

Agora precisaremos de algumas definições. Dada uma família $\mathcal{X} = \{X_i : i \in I\}$, em que I é um conjunto qualquer e os elementos de $\mathcal{X}$ são conjuntos não vazios e dois a dois disjuntos, definimos o conjunto $X = \bigcup \mathcal{X}$. Considere o conjunto $\mathbb{R}[X]$ dos polinômios com variáveis em X. Para detalhes das definições de corpo, subcorpo e anel de polinômios, veja [11]. Mais precisamente, em [11, p.40] aparecem os detalhes da construção de $\mathbb{R}[X]$. Assim como nessa referência, denotaremos por $\mathbb{R}(X)$ o corpo de frações de $\mathbb{R}[X]$, ou seja, $\mathbb{R}(X) = \left\{r = \frac{p}{q} : p, q \in \mathbb{R}[X], q \neq 0\right\}$. Segue também de [11] que $\mathbb{R}[X]$ é um domínio de fatoração única e isto implica que todo elemento de $\mathbb{R}(X)$ se escreve de maneira única como o quociente de polinômios primos entre si, isto é, sem fatores em comum.

Definição 2.8. Um *monômio* M em $\mathbb{R}[X]$ é dado por, $M = \prod_{x \in X} x^{\alpha_n}$ em que $\alpha_n \in \mathbb{N}$. Um *polinômio* em $\mathbb{R}[X]$ é uma combinação linear finita de monômios com coeficientes em $\mathbb{R}$.

Definição 2.9. Seja $q = x_{j_1}^{n_1} x_{j_2}^{n_2} \dots x_{j_m}^{n_m}$ um monômio real com variáveis em X. Dizemos que q tem *i-grau* N, se as variáveis de q que estiverem em X_i tiverem N como a soma de seus graus.

Para ilustrar tome, por exemplo, os conjuntos $X_i = \{x_1, x_2\}$ e $X_j = \{y_1, y_2\}$. Então $X = \{x_1, x_2, y_1, y_2\}$. Um monômio em $\mathbb{R}[X]$ é da forma $M = x_1^{\alpha_1} x_2^{\alpha_2} y_1^{\beta_1} y_2^{\beta_2}$. O i-grau de M é $\alpha_1 + \alpha_2$ e o j-grau de M é $\beta_1 + \beta_2$.

Definição 2.10. Seja $p \in \mathbb{R}(X)$. Diremos que p é *i-homogêneo de grau d* (aqui d pode ser inclusive um inteiro negativo), se p é o quociente de dois polinômios tais que todos os monômios no denominador têm o mesmo i-grau n enquanto que todos no numerador têm i-grau $n + d$. Denotamos este grau d do polinômio p por i-grau(p).

Lema 2.11. *Seja $\mathcal{K}$ o conjunto das funções racionais $r \in \mathbb{R}(X)$, tais que, $r = \frac{p}{q}$, com $p, q \in \mathbb{R}[X]$, $q \neq 0$ e com p e q sem fatores comuns, tais que para todo $i \in I$, ocorre que i-grau(p)=i-grau(q). Então $\mathcal{K}$ é subcorpo de $\mathbb{R}(X)$.*

Demonstração. Para mostrar que $\mathcal{K}$ é subcorpo, devemos mostrar que:

i. $0 \in \mathcal{K}$ e $1 \in \mathcal{K}$;

ii. Se $\frac{p}{q} \in \mathcal{K}$, então $-\frac{p}{q} \in \mathcal{K}$ e $\frac{q}{p} \in \mathcal{K}$;

iii. Se $\frac{p_1}{q_1}, \frac{p_2}{q_2} \in \mathcal{K}$, então $\frac{p_1}{q_1} + \frac{p_2}{q_2} \in \mathcal{K}$;

iv. Se $\frac{p_1}{q_1}, \frac{p_2}{q_2} \in \mathcal{K}$, então $\frac{p_1}{q_1} \cdot \frac{p_2}{q_2} \in \mathcal{K}$.

A propriedade i. é imediata, uma vez que $0 = \dfrac{0}{1}$ e $1 = \dfrac{1}{1}$, são quocientes de polinômios i-homogêneos de grau 0.

Já a propriedade ii. segue do fato que i-grau(p) = i-grau($-p$).

Agora sejam $\dfrac{p_1}{q_1}, \dfrac{p_2}{q_2} \in \mathcal{K}$ e vamos provar as propriedades iii. e iv. Primeiro veja que:

$$\frac{p_1}{q_1} + \frac{p_2}{q_2} = \frac{p_1 q_2 + p_2 q_1}{q_1 q_2}$$

.

Sejam p e q dois polinômios sem fatores em comum, tais que $p_1 q_2 + p_2 q_1 = fp$, e $q_1 q_2 = fq$, dessa forma, $\dfrac{p_1}{q_1} + \dfrac{p_2}{q_2} = \dfrac{p}{q}$.

Agora, veja que para todo $i \in I$, vale que:

$$i\text{-grau}(p_1 q_2) = i\text{-grau}(p_1) + i\text{-grau}(q_2) = i\text{-grau}(q_1) + i\text{-grau}(p_2) = i\text{-grau}(q_1 p_2).$$

Portanto, $p_1 q_2 + p_2 q_1$ é i-homogêneo. E ainda,

$$i\text{-grau}(p_1 q_2 + p_2 q_1) = i\text{-grau}(p_2 q_1) = i\text{-grau}(q_1 q_2).$$

Logo, i-grau(fp) = i-grau(fq), e portanto, i-grau(p) = i-grau(q). Ou seja, $\dfrac{p}{q} \in \mathcal{K}$.

O mesmo raciocínio pode ser usado para provar que i-grau$(p_1p_2) = i$-grau(q_1q_2), e assim temos a validade da propriedade iv. Portanto, $\mathcal{K}$ é subcorpo de $\mathbb{R}(X)$.

□

Teorema 2.12. *Todo espaço vetorial tem base $\Rightarrow$ AE.*

Demonstração. Considere uma família de conjuntos não vazios $\mathcal{X} = \{X_i : i \in I\}$. Sem perda de generalidade, podemos supor que os X_i's são dois a dois disjuntos (se não for o caso, basta considerar a família $\{X_i \times \{X_i\} : i \in I\}$). Conforme vimos acima, basta mostrar que vale o Axioma da Múltipla Escolha. Para isto é suficiente encontrar uma família $\{Y_i : i \in I\}$ de subconjuntos não vazios finitos $Y_i \subset X_i$ e definir $f(X_i) = Y_i$, para cada $i \in I$. Considere $X = \bigcup \mathcal{X}$ e o corpo de frações $\mathbb{R}(X)$. Seja $V \subset \mathbb{R}(X)$ o $\mathcal{K}$-espaço vetorial gerado pelo conjunto X, onde $\mathcal{K}$ é o corpo introduzido no Lema 2.11. Pela hipótese V tem uma base, digamos $B \subset V$. Seja $i \in I$ e $x \in X_i$. Então x pode ser escrito como uma combinação linear finita de elementos de B, isto é,

$$x = \sum_{b \in B(x)} \alpha_b(x) \cdot b,$$

onde $B(x) := \{b \in B : \alpha_b(x) \neq 0\}$ é um subconjunto finito de B. Se y é outro elemento de X_i, então temos

$$y = \sum_{b' \in B(y)} \alpha_{b'}(y) \cdot b'.$$

Como $y = \left(\dfrac{y}{x}\right) \cdot x$ e, pela definição de $\mathcal{K}$, temos que $\dfrac{y}{x} \in \mathcal{K}$, então

$$y = \sum_{b \in B(x)} \left(\frac{y}{x} \cdot \alpha_b(x)\right) \cdot b.$$

Comparando as duas representações de y e utilizando a unicidade da representação de vetores em uma base, temos que

$$B(x) = B(y) \text{ e } \alpha_b(y) = \frac{y}{x} \cdot \alpha_b(x), \text{ para todo } b \in B(x).$$

Isso significa que o subconjunto finito $B(x) \subset B$ e os elementos $\dfrac{\alpha_b(x)}{x}$ de $\mathbb{R}(X)$ só dependem de i e não da escolha particular de $x \in X_i$. Portanto podemos chamar $B(x)$ simplesmente de B_i e $\dfrac{\alpha_b(x)}{x}$ de α_b^i, qualquer que seja $x \in X_i$. Como α_b^i é i-homogêneo de grau -1 (e j-homogêneo de grau 0 para $j \neq i$), quando escrevemos α_b^i como um quociente de polinômios sem fatores em comum, alguma(s) variável(is) de X_i devem aparecer no denominador. Seja D_b^i o conjunto das variáveis de X_i que aparecem no denominador de α_b^i (quando α_b^i é escrito como quociente de polinômios sem fatores em comum). Assim D_b^i é um subconjunto finito e não vazio de X_i. Tomando $Y_i = \bigcup_{b \in B_i} D_b^i$, segue que Y_i também é um subconjunto finito e não vazio de X_i. Agora, basta definir $f(X_i) = Y_i$, para cada $i \in I$. □

3 ANÁLISE FUNCIONAL

3.1 RESULTADOS INICIAIS

Neste capítulo, apresentaremos dois teoremas da Análise Funcional, que têm as suas demonstrações dependentes do Lema de Zorn. Entretanto, para podermos apresentar estes resultados precisamos de definições preliminares, que serão tratadas a seguir. Nosso texto de referência foi [4], e as demais definições podem ser encontradas lá. A partir de agora $\mathbb{K}$ denota o corpo dos números reais ou complexos.

Definição 3.1. Um *espaço normado* é um espaço vetorial E munido de uma *norma*, isto é, uma função

$$x \in E \to \|x\| \in \mathbb{R}$$

satisfazendo as seguintes propriedades:

(a) $\|x\| \geq 0$ para todo $x \in E$;

(b) $\|x\| = 0$ se e só se $x = 0$;

(c) $\|\lambda x\| = |\lambda|\|x\|$ para todo $\lambda \in \mathbb{K}$ e $x \in E$;

(d) $\|x + y\| \leq \|x\| + \|y\|$ para todo $x, y \in E$.

Diremos que E completo se toda sequência de Cauchy em E for convergente em E. Lembrando que $(x_n) \subset E$ é:

- *de Cauchy* se para todo $\varepsilon > 0$, existe $n_0 \in \mathbb{N}$ tal que $\|x_n - x_m\| < \varepsilon$, para todos $m, n \geq n_0$.
- *convergente* se existe $x \in E$ tal que para todo $\varepsilon > 0$, existe $n_0 \in \mathbb{N}$ tal que $\|x_n - x\| < \varepsilon$, para todo $n \geq n_0$.

A seguinte caracterização de continuidade para operadores lineares entre espaços normados é bastante útil.

Proposição 3.2. *Sejam E e F espaços normados. Dado um operador linear $T \colon E \to F$, as seguintes condições são equivalentes:*

(a) T é contínua.

(b) $\sup\{\|Tx\| : x \in E, \|x\| \leq 1\} < \infty$.

(c) Existe uma constante $c > 0$ tal que $\|Tx\| \leq c\|x\|$ para todo $x \in E$.

Definição 3.3. Dada um operador linear $T : E \to F$, definimos

$$\|T\| := \sup\{\|Tx\| : x \in E, \|x\| \leq 1\}.$$

O caso que estamos particularmente interessados é o de *funcionais lineares*, isto é, operadores lineares da forma $\phi \colon E \to \mathbb{K}$. Neste caso temos

$$\|\phi\| = \sup\{|\phi(x)| : x \in E \text{ e } \|x\| \leq 1\}.$$

O espaço vetorial de todos os funcionais lineares contínuos de E em $\mathbb{K}$ é denotado por E' e é chamado de *dual topológico*, ou simplesmente *dual* de E.

Proposição 3.4. *(a) A função $\phi \to \|\phi\|$ é uma norma em E'.*

(b) $|\phi(x)| \leq \|\phi\|\|x\|$ para todo $\phi \in E'$ e $x \in E$.

(c) $\|\phi\| = \inf\{c > 0 : |\phi(x)| \leq c\|x\|,$ para todo $x \in E\}$

Definição 3.5. Seja E um espaço vetorial sobre $\mathbb{K}$. Um *produto interno* em E é uma função

$$(x, y) \in E \times E \to \langle x, y\rangle \in \mathbb{K}$$

com as seguintes propriedades:

(a) $\langle x_1 + x_2, y\rangle = \langle x_1, y\rangle + \langle x_2, y\rangle$, $\forall x_1, x_2, y \in E$.

(b) $\langle \lambda x, y\rangle = \lambda \langle x, y\rangle$, $\forall x, y \in E$ e $\lambda \in \mathbb{K}$.

(c) $\langle x, y\rangle = \overline{\langle y, x\rangle}$, $\forall x, y \in E$.

(d) $\langle x, x\rangle \geq 0$, e mais $\langle x, x\rangle = 0 \Leftrightarrow x = 0$, $\forall x \in E$.

Proposição 3.6. *Seja E um espaço com produto interno. Então a função $\|x\| = \sqrt{\langle x, x\rangle}$, é uma norma em E, que é chamada de* norma induzida pelo produto interno.

Definição 3.7. Sejam E um espaço com produto interno e $x, y \in E$, com $x \neq y$. Dizemos que x e y são *ortogonais*, e escrevemos $x \perp y$, se $\langle x, y\rangle = 0$.

Definição 3.8. Sejam E um espaço com produto interno e A subconjunto não-vazio de E. Chamamos de *complemento ortogonal* de A o conjunto $A^{\perp} = \{y \in E : \langle x, y\rangle = 0, \forall x \in A\}$.

Dessa forma, dizemos ainda que dois vetores x e y de um espaço E são *ortonormais*, se além de ortogonais, valer que $\langle x, x\rangle = 1$ e $\langle y, y\rangle = 1$. Um conjunto em que todos os elementos são ortonormais entre si é chamado de *conjunto ortonormal.*

Definição 3.9. Um conjunto ortonormal S tal que $S^{\perp} = \{0\}$ é chamado de *sistema ortonormal completo.*

3.2 APLICAÇÕES DO LEMA DE ZORN EM ANÁLISE FUNCIONAL

A partir daqui, apresentaremos e discutiremos duas aplicações do Lema de Zorn em Análise Funcional.

3.2.1 LZ $\Rightarrow$ EM UM ESPAÇO COM PRODUTO INTERNO TODO CONJUNTO ORTONORMAL ESTÁ CONTIDO EM ALGUM SISTEMA ORTONORMAL COMPLETO

Teorema 3.10. *LZ $\Rightarrow$ todo conjunto ortonormal está contido em algum sistema ortonormal completo.*

Demonstração. Sejam E um espaço com produto interno, S_0 um sistema ortonormal em E e $\mathcal{F}$ a família de todos os conjuntos ortonormais em E que contém S_0. Pela inclusão de conjuntos, temos que $\mathcal{F}$ é um conjunto parcialmente ordenado, e mais, $\mathcal{F} \neq \emptyset$, pois $S_0 \in \mathcal{F}$. Seja $\{S_i : i \in I\} \subseteq \mathcal{F}$ totalmente ordenado.

Primeiro, vamos provar que $\bigcup_{i \in I} S_i \in \mathcal{F}$. De fato, se $x, y \in \bigcup S_i$, então existem i_1 e i_2 índices tais que $x \in S_{i_1}$ e $y \in S_{i_2}$. Se $x = y$, então $\langle x, y\rangle = \langle x, x\rangle = 1$, pois S_{i_1} é ortonormal. Se $x \neq y$, como $\{S_i : i \in I\}$ é totalmente ordenado, segue que $S_{i_1} \subseteq S_{i_2}$ ou $S_{i_2} \subseteq S_{i_1}$. Assim, temos que $x, y \in S_{i_1}$ ou $x, y \in S_{i_2}$. Como tanto S_{i_1}, quanto S_{i_2} são conjuntos ortonormais, segue que $x \perp y$. Portanto, como tomamos $x, y \in \bigcup_{i \in I} S_i$, segue que este é um conjunto ortonormal em E que contém S_0. E assim, $\bigcup_{i \in I} S_i \in \mathcal{F}$.

Como $\bigcup_{i \in I} S_i$ é uma cota superior para o conjunto $\{S_i : i \in I\}$, segue que todo subconjunto totalmente ordenado de $\mathcal{F}$ tem cota superior. Sendo assim, podemos usar o Lema de Zorn sobre $\mathcal{F}$, o que nos garante que existe um conjunto ortonormal $S \in \mathcal{F}$ que é maximal. Agora basta mostrar que S é completo.

Suponha que S não é completo. Então existe $s \in E, s \neq 0$ tal que $s \in S^{\perp}$. Assim $S' := S \cup \left\{\frac{s}{\|s\|}\right\}$ é ortonormal, $S' \neq S$ e $S \subseteq S'$, o que é um absurdo, pois S é maximal.

$\square$

Uma pergunta natural é se vale a recíproca deste teorema. Este é um problema em aberto, conforme descrito no livro [9]. Como este livro é de 2007, fizemos uma pesquisa detalhada sobre trabalhos posteriores, e acreditamos que este continua sendo um problema em aberto até a data de publicação deste texto.

3.2.2 O TEOREMA DE HAHN-BANACH

A partir de agora caminharemos no sentido de demonstrar o Teorema de Hahn-Banach. Para demonstrar este teorema, é fundamental a utilização do Lema de Zorn 1.22. Para isso vamos a uma definição:

Definição 3.11. Sejam E um espaço vetorial, G um subespaço de E e $\varphi : G \to \mathbb{K}$ um funcional linear. Dizemos que um funcional linear $\tilde{\varphi} : E \to \mathbb{K}$ *estende* φ se $\tilde{\varphi}(x) = \varphi(x)$ para todo $x \in G$.

Para demonstrar o Teorema de Hanh-Banach, primeiro o restringiremos ao caso real.

Teorema 3.12 (Teorema de Hahn-Banach para espaços vetoriais reais)**.** *Sejam E um espaço vetorial real e $p : E \to \mathbb{R}$ uma função tal que:*

$$p(ax) = ap(x) \text{ para todo } a \in \mathbb{K} \text{ e todo } x \in E$$
$$p(x+y) \leq p(x) + p(y) \text{ para quaisquer } x, y \in E$$

Se $G \subseteq E$ é um subespaço vetorial e $\varphi : G \to \mathbb{R}$ é um funcional linear tal que $\varphi(x) \leq p(x)$ para todo $x \in G$, então existe um funcional linear $\tilde{\varphi} : E \to \mathbb{R}$ que estende φ a E e que satisfaz $\tilde{\varphi}(x) \leq p(x)$ para todo $x \in E$.

Demonstração. Definimos $\mathcal{P}$ como uma família de funcionais lineares definidos em subespaços de E que contém G, tal que:

$$\mathcal{P} = \left\{ \begin{array}{c} \phi : D(\phi) \subseteq E \to \mathbb{R} | D(\phi) \text{ é subespaço vetorial de } E, \phi \text{ é linear, estende } \varphi, \\ G \subseteq D(\phi),\ \phi(x) = \varphi(x) \text{ para todo } x \in G \text{ e } \phi(x) \leq p(x) \text{ para todo } x \text{ em } D(\phi) \end{array} \right\}$$

Na família $\mathcal{P}$ podemos definir a seguinte relação:

$$\phi_1 \leq \phi_2 \Leftrightarrow D(\phi_1) \subseteq D(\phi_2) \text{ e } \phi_2 \text{ estende } \phi_1 \text{ a } D(\phi_2)$$

Vejamos que esta relação é uma ordem parcial. Para isso, devemos provar que esta relação é reflexiva, antissimétrica e transitiva, como apresentado na Definição 1.16.

- A relação é reflexiva, pois, trivialmente, $D(\phi_1) \subseteq D(\phi_1)$ e $\phi_1(x) = \phi_1(x)$ para todo $x \in D(\phi_1)$, ou seja, ϕ_1 estende ϕ_1. Logo, $\phi_1 \leq \phi_1$.
- A relação é antissimétrica, pois, se $\phi_1 \leq \phi_2$ e $\phi_2 \leq \phi_1$, então temos que $D(\phi_1) \subseteq D(\phi_2)$ ao mesmo tempo que $D(\phi_2) \subseteq D(\phi_1)$, disso tiramos que $D(\phi_1) = D(\phi_2)$. Por outro lado, ϕ_2 estende ϕ_1, isto é, $\phi_1(x) = \phi_2(x)$ para todo $x \in D(\phi_2) = D(\phi_1)$. Em resumo, $\phi_1 = \phi_2$.

- A relação é transitiva, pois, se $\phi_1 \leq \phi_2$ e $\phi_2 \leq \phi_3$, então temos que $D(\phi_1) \subseteq D(\phi_2) \subseteq D(\phi_3)$. E como ϕ_2 estende ϕ_1 temos que $\phi_1(x) = \phi_2(x)$ para todo $x \in D(\phi_1)$, temos ainda que ϕ_3 estende ϕ_2, isto é, $\phi_2(x) = \phi_3(x)$ para todo $x \in D(\phi_2)$. Logo, segue que $\phi_1(x) = \phi_2(x) = \phi_3(x)$ para todo $x \in D(\phi_1)$, isto é, ϕ_3 estende ϕ_1. Portanto, temos, $\phi_1 \leq \phi_3$.

Da hipótese do teorema segue trivialmente que $\varphi \in \mathcal{P}$, ou seja, $\mathcal{P}$ é não-vazio.

Para usarmos o Lema de Zorn falta mostrar então que todo subconjunto totalmente ordenado de $\mathcal{P}$ admite uma cota superior.

De fato, seja $\mathcal{Q} \subseteq \mathcal{P}$ totalmente ordenado. A candidata a cota superior de $\mathcal{Q}$ é a função $\psi : D(\psi) \to \mathbb{R}$, definida por:

$$D(\psi) = \bigcup_{\theta \in \mathcal{Q}} D(\theta) \text{ e } \psi(x) = \theta(x) \text{ se } x \in D(\theta)$$

Vejamos que, ψ é cota superior para $\mathcal{Q}$. Seja $\tilde{\psi} \in \mathcal{Q}$. Agora, temos que $D(\tilde{\psi}) \subseteq \bigcup_{\theta \in \mathcal{Q}} = D(\psi)$, e ainda, $\psi(x) = \tilde{\psi}(x)$ para todo $x \in D\tilde{\psi}$. Logo, temos que $\tilde{\psi} \leq \psi$.

Finalmente podemos usar o Lema de Zorn. Dele segue que $\mathcal{P}$ admite um elemento maximal, que denotaremos por $\tilde{\varphi}$. Para provarmos o resultado pretendido basta mostrarmos que $D(\tilde{\varphi}) = E$. Suponha, por absurdo que $D(\tilde{\varphi}) \neq E$. Escolhemos então um $x_0 \in E - D(\tilde{\varphi})$ e definimos $\tilde{\phi} : D(\tilde{\phi}) \to \mathbb{R}$ por:

$$D(\tilde{\phi}) = D(\tilde{\varphi}) + [x_0] \text{ e } \tilde{\phi}(x + tx_0) = \tilde{\varphi}(x) + t\alpha$$

Nesta definição, α é uma constante que será escolhida posteriormente de forma a garantir que $\tilde{\phi} \in \mathcal{P}$. Queremos, por enquanto, que α satisfaça as seguintes desigualdades:

$$\tilde{\varphi}(x) + \alpha = \tilde{\phi}(x + x_0) \leq p(x + x_0) \text{ para todo } x \in D(\tilde{\varphi}) \text{ e}$$
$$\tilde{\varphi}(x) - \alpha = \tilde{\phi}(x - x_0) \leq p(x - x_0) \text{ para todo } x \in D(\tilde{\varphi}).$$

Sendo assim, temos que $\alpha \leq p(x + x_0) - \tilde{\varphi}(x)$, e $\alpha \geq \tilde{\varphi}(x) - p(x - x_0)$ para todo $x \in D(\tilde{\varphi})$. Por isso, devemos escolher α de modo que,

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

Veja que essa escolha é de fato possível, pois para $x, y \in D(\tilde{\varphi})$ temos:

Disso segue que, $\tilde{\varphi}(y) - p(y - x_0) \leq p(x + x_0) - \tilde{\varphi}(x)$, para quaisquer $x, y \in D(\tilde{\varphi})$. Portanto, podemos escolher α dentro de tais condições. Agora, temos que:

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

- Se $t > 0$,

$$\begin{aligned}
\tilde{\phi}(x + tx_0) &= \tilde{\phi}\left(t\left(\frac{x}{t} + x_0\right)\right) \\
&= t\tilde{\phi}\left(\frac{x}{t} + x_0\right) \\
&\leq tp\left(\frac{x}{t} + x_0\right) = p(x + tx_0)
\end{aligned}$$

- Se $t < 0$,

$$\begin{aligned}
\tilde{\phi}(x + tx_0) &= \tilde{\phi}\left(-t\left(\frac{x}{-t} - x_0\right)\right) \\
&= -t\tilde{\phi}\left(\frac{x}{-t} - x_0\right) \\
&\leq -tp\left(-\frac{x}{t} - x_0\right) = p(x + tx_0)
\end{aligned}$$

- Se $t = 0$,

$$\begin{aligned}
\tilde{\phi}(x + tx_0) &= \tilde{\phi}(x) \\
&= \tilde{\varphi}(x) \\
&\leq p(x) = p(x + tx_0)
\end{aligned}$$

Segue então que $\tilde{\phi} \in \mathcal{P}$, $\tilde{\varphi} \leq \tilde{\phi}$ e $\tilde{\varphi} \neq \tilde{\phi}$. Mas $\tilde{\varphi}$ é máximo. Temos portanto, um absurdo. Logo $D(\tilde{\varphi}) = E$. □

Com isso podemos enunciar e demonstrar o teorema no caso geral, em que o corpo $\mathcal{K}$ é os reais ou complexos.

Teorema 3.13 (Teorema de Hahn-Banach)**.** *Sejam E um espaço vetorial sobre o corpo $\mathbb{K}$ e $p : E \to \mathbb{R}$ uma função, tal que:*

$$p(ax) = |a|p(x) \text{ para todo } a \in \mathbb{K} \text{ e todo } x \in E \tag{3.1}$$

$$p(x+y) \leq p(x) + p(y) \text{ para quaisquer } x, y \in E \tag{3.2}$$

Se $G \subseteq E$ é um subespaço vetorial e $\varphi : G \to \mathbb{K}$ é um funcional linear tal que $|\varphi(x)| \leq p(x)$ para todo $x \in G$, então existe um funcional linear $\tilde{\varphi} : E \to \mathbb{K}$ que estende φ a E e que satisfaz $|\tilde{\varphi}(x)| \leq p(x)$ para todo $x \in E$.

Demonstração. Vejamos que $p(x) \geq 0$ para todo $x \in E$. Com efeito, de 3.2, segue que, $p(0) = p(0+0) \leq 2p(0)$, isto é, $p(0) \geq 1$, e consequentemente $p(0) \geq 0$. Por outro lado, de 3.1, fazendo $a = -1$, segue que $p(x) = p(-x)$. Assim sendo, temos, $2p(x) = p(x) + p(-x) \geq p(x + (-x)) = p(0) \geq 0$, logo, $p(x) \geq 0$, para todo $x \in E$.

Trataremos primeiro o caso em que $\mathbb{K} = \mathbb{R}$. Nesse caso, por hipótese, temos que $\varphi(x) \leq p(x)$, para todo $x \in G$. O Teorema 3.12 garante que existe um funcional linear $\tilde{\varphi} : E \to \mathbb{R}$ que estende φ a E e que satisfaz $\tilde{\varphi}(x) \leq p(x)$ para todo $x \in E$. Dessa desigualdade e de 3.1 segue que, $-\tilde{\varphi}(x) = \tilde{\varphi}(-x) \leq p(-x) = |-1|p(x) = p(x)$, para todo $x \in E$. Logo, $|\tilde{\varphi}(x)| \leq p(x)$ para todo $x \in E$.

Vamos então para o caso $\mathbb{K} = \mathbb{C}$. Nesse caso E é um espaço vetorial complexo e φ toma valores em $\mathbb{C}$. Começamos essa demonstração definindo $\varphi_1, \varphi_2 : E \to \mathbb{R}$ por $\varphi_1(x) = \mathrm{Re}(\varphi(x))$ e $\varphi_2(x) = \mathrm{Im}(\varphi(x))$. Temos que φ_1 e φ_2 são funcionais lineares reais, e mais, vale que $\varphi = \varphi_1 + i\varphi_2$.Vamos chamar de $E_{\mathbb{R}}$ e $G_{\mathbb{R}}$ os espaços vetoriais reais subjacentes a E e G, isto é, como conjuntos os espaços são os mesmos, a adição definida também é a mesma, entretanto, a multiplicação por escalar toma somente escalares reais. Então φ_1 e φ_2 são funcionais lineares sobre $G_{\mathbb{R}}$. Vamos então tomar $x \in G_{\mathbb{R}}$. Disso obtemos,

$$\varphi_1(x) \leq |\varphi_1(x)| \leq |\varphi(x)| \leq p(x).$$

Do caso real, isto é do Teorema 3.12, segue que existe um funcional linear $\tilde{\varphi_1} : E_{\mathbb{R}} \to \mathbb{R}$ que estende φ_1 a $E_{\mathbb{R}}$ e que satisfaz $\tilde{\varphi_1}(x) \leq p(x)$ para todo $x \in E_{\mathbb{R}}$. Estudemos agora o caso de φ_2. Note que para $x \in G$,

$$i(\varphi_1(x) + i\varphi_2(x)) = i\varphi(x) = \varphi(ix) = \varphi_1(ix) + i\varphi_2(ix).$$

Portanto $\varphi_2(x) = -\varphi_1(ix)$, para todo $x \in G$. Ao definirmos $\tilde{\varphi} : E \to \mathbb{C}, \tilde{\varphi}(x) = \tilde{\varphi_1}(x) - i\tilde{\varphi_1}(ix)$, segue que $\tilde{\varphi}(x) = \varphi(x)$ para todo $x \in G$. Vejamos que $\tilde{\varphi}$ é um funcional linear no espaço complexo E. É imediato que $\tilde{\varphi}(x+y) = \tilde{\varphi}(x) + \tilde{\varphi}(y)$ para todos $x, y \in E$. Dados $(\alpha + \beta i) \in \mathbb{C}$ e $x \in E$,

$$\begin{aligned}
\tilde{\varphi}((\alpha + \beta i)x) &= \tilde{\varphi_1}(\alpha x + i\beta x) - i\tilde{\varphi_1}((i(\alpha + \beta i)x)) \\
&= \alpha\tilde{\varphi_1}(x) + \beta\tilde{\varphi_1}(ix) - i(\alpha\tilde{\varphi_1}(ix) - \beta\tilde{\varphi_1}(x)) \\
&= (\alpha + \beta i)(\tilde{\varphi_1}(x) - i\tilde{\varphi_1}(ix)) = (\alpha + \beta i)\tilde{\varphi}(x).
\end{aligned}$$

Por fim vamos mostrar que $|\tilde{\varphi}(x)| \leq p(x)$ para todo $x \in E$. Se $\tilde{\varphi}(x) = 0$, então a desigualdade é óbvia pois $p(x) \geq 0$. Tome $x \in E$ tal que $\tilde{\varphi}(x) \neq 0$. Então existe θ tal que $\tilde{\varphi}(x) = |\tilde{\varphi}(x)|e^{i\theta}$. Segue que $|\tilde{\varphi}(x)| = e^{-i\theta}\tilde{\varphi}(x) = \tilde{\varphi}(e^{-i\theta}x)$. Como $|\tilde{\varphi}(x)|$ é real, por 3.2, temos:

$$|\tilde{\varphi}(x)| = \tilde{\varphi}(e^{-i\theta}x) = \tilde{\varphi_1}(e^{-i\theta}x) \leq p(e^{-i\theta}x) = |e^{-i\theta}|p(x) = p(x).$$

□

O próximo corolário é o enunciado clássico do Teorema de Hanh-Banach.

Corolário 3.14 (Teorema de Hanh-Banach)**.** *Seja G um subespaço de um espaço normado E sobre $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ e seja $\varphi : G \to \mathbb{K}$ um funcional linear contínuo. Então existe um funcional linear contínuo $\tilde{\varphi} : E \to \mathbb{K}$ cuja restrição a G coincide com φ e $\|\tilde{\varphi}\| = \|\varphi\|$.*

Demonstração. Definimos:

$$T : E \to \mathbb{R}$$

$$T(x) = \|\varphi\|\|x\|$$

A seguir, vamos provar duas propriedades de T. Se $k \in K$, segue que

i. $T(kx) = \|\varphi\|\|kx\| = \|\varphi\| (|k| \cdot \|x\|) = |k|T(x)$.

ii. $T(x + y) = \|\varphi\|\|x + y\| \leq \|\varphi\| (\|x\| + \|y\|) = \|\varphi\| \cdot \|x\| + \|\varphi\| \cdot \|y\| = T(x) + T(y)$

Como φ é funcional linear contínuo, vale para todo $x \in G$ que

$$|\varphi(x)| \leq \|\varphi\| \cdot \|x\| = T(x)$$

Do teorema anterior, existe um funcional linear $\tilde{\varphi} : E \to \mathbb{R}$, tal que

$$|\tilde{\varphi}(x)| \leq \|\varphi\|\|x\| = T(x).$$

Assim, segue que $\|\varphi\| \leq \|\tilde{\varphi}\|$.

Por outro lado, $|\varphi(x)| = |\tilde{\varphi}(x)| \leq \|\tilde{\varphi}\| \cdot \|x\|$ para todo $x \in G$. Logo, $\|\varphi\| \leq \|\tilde{\varphi}\|$.

Portanto, $\|\varphi\| = \|\tilde{\varphi}\|$.

□

Conforme vimos, o Lema de Zorn implica o Teorema de Hanh-Banach. Um pergunta natural é, se nesse caso, vale a recíproca. A resposta é não, ou seja, o Teorema de Hanh-Banach não implica o Lema de Zorn. Em 1964, Halpern provou que o Teorema do ultrafiltro (um resultado clássico de topologia geral) não implica o Axioma da Escolha e consequentemente, não implica o Lema de Zorn (e já é bem conhecido da teoria básica de topologia geral que o Lema de Zorn implica o Teorema do ultrafiltro). Em 1951, Loś e Ryll-Nardzewski provaram que o Teorema do ultrafiltro implica o Teorema de Hahn-Banach. Finalmente, em 1974, Pincus provou que o Teorema de Hanh-Banach não implica o Teorema do ultrafiltro. Sendo assim, valem as implicações abaixo, mas não valem as implicações reversas:

Axioma da Escolha $\Rightarrow$ Teorema do ultrafiltro $\Rightarrow$ Teorema de Hahn-Banach.

As referências e as descrições destes resultados podem ser encontradas em [1].

REFERÊNCIAS

[1] BECKENSTEIN, E; NARICI, L. *The Hanh-Banach theorem: the life and times*. Topology and its Applications **77** (1997), 193–211.

[2] BERNAYS, P.; FRAENKEL, A. *Axiomatic Set Theory*. North-Holland Publishing Company, Amsterdam, 1968.

[3] BLASS, A. *Existence of bases implies the axiom of choice*. Contemporary Mathematics **31** (1984), 31–33.

[4] BOTELHO, G.; PELLEGRINO, D.; TEIXEIRA, E. *Fundamentos de Análise Funcional*. SBM, Rio de Janeiro, 2012.

[5] COELHO, F. U.; LOURENÇO, M. L. *Um Curso de Álgebra Linear*. EDUSP, São Paulo, 2013.

[6] ENDERTON, H. *Elements of Set Theory*. Elsevier, San Diego, 1977.

[7] JECH, T. *The Axiom of Choice*. North-Holland, 1973.

[8] KOVACEC, A. *Lógica*. Departamento de Matemática da Universidade de Coimbra, Coimbra, 2007.

[9] PIETSCH, A. *History of Banach Spaces and Linear Operators*. Birkhäuser Basel, 2007.

[10] RUSSELL, B. *Principia Mathematica*. Cambridge University Press, Londres, 1927.

[11] ZARISKI, O.; SAMUEL, P. *Comutative Algebra*. vol. 1. D. Van Nostrand Company, Nova York, 1958.