14. O lema de Zorn e o teorema de Zermelo

14.1. Definição. Chamaremos de *relação de ordem parcial* num conjunto X uma relação $\leq$ em X com as seguintes propriedades:
(a) $x \leq x$ para todo $x \in X$ ($\leq$ é reflexiva);
(b) se $x \leq y$ e $y \leq x$, então $x = y$ ($\leq$ é antisimétrica);
(c) se $x \leq y$ e $y \leq z$, então $x \leq z$ ($\leq$ é transitiva).
Neste caso diremos que X é um *conjunto parcialmente ordenado.*

Diremos que $\leq$ é uma *relação de ordem total* se além de verificar (a), (b) e (c), também verifica
(d) dados $x, y \in X$, tem-se que $x \leq y$ ou $y \leq x$.
Neste caso diremos que X é um *conjunto totalmente ordenado.*

14.2. Exemplos.
(a) Se X é um conjunto, então a relação de inclusão é uma relação de ordem parcial em $\mathcal{P}(X)$.

(b) A relação de ordem usual em $\mathbf{R}$ é uma relação de ordem total.

14.3. Definição. Seja X um conjunto parcialmente ordenado, e seja $A \subset X$.
(a) Se existir $a_0 \in A$ tal que $a_0 \leq a$ para todo $a \in A$, diremos que a_0 é o *elemento mínimo* de A. De maneira análoga definimos *elemento máximo.*

(b) Se existir $a_0 \in A$ tal que $a = a_0$ sempre que $a \in A$ e $a \leq a_0$, diremos que a_0 é um *elemento minimal* de A. De maneira análoga definimos *elemento maximal.*

(c) Se existir $c \in X$ tal que $c \leq a$ para todo $a \in A$, diremos que A é *limitado inferiormente* e que c é uma *cota inferior* de A. De maneira análoga definimos conjunto *limitado superiormente* e *cota superior.*

(d) Diremos que A é uma *cadeia* em X se A é totalmente ordenado sob a relação de ordem parcial induzida por X.

(e) Diremos que A é *bem ordenado* se cada subconjunto não vazio de A possui um elemento mínimo.

14.4. Exemplos.
(a) $\mathbf{N}$, com a ordem usual, é um conjunto bem ordenado.
(b) $\mathbf{R}$, com a ordem usual, é um conjunto totalmente ordenado, que não é bem ordenado: o intervalo aberto (a, b) não possui elemento mínimo.

14.5. Lema de Zorn. *Seja X um conjunto parcialmente ordenado não vazio tal que cada cadeia em X é limitada superiormente. Então X possui pelo menos um elemento maximal.*

14.6. Teorema de Zermelo. *Cada conjunto não vazio pode ser bem ordenado.*

14.7. Teorema. *As seguintes afirmações são equivalentes:*
(a) O axioma da escolha.
(b) O lema de Zorn.
(c) O teorema de Zermelo.

Demonstração. $(b) \Rightarrow (c)$: Seja X um conjunto não vazio. Seja $\mathcal{F}$ a família de todos os pares $(A, \leq_A)$ tais que $\emptyset \neq A \subset X$ e $(A, \leq_A)$ é um conjunto bem ordenado. É fácil verificar que $\mathcal{F}$ é um conjunto parcialmente ordenado não vazio se definimos $(A, \leq_A) \leq (B, \leq_B)$ quando:
(i) $A \subset B$;
(ii) se $x, y \in A$, então $x \leq_A y$ se e só se $x \leq_B y$;
(iii) se $x \in A$ e $y \in B \setminus A$, então $x \leq_B y$.

Provaremos que cada cadeia em $\mathcal{F}$ é limitada superiormente. De fato, seja $\{(A_i, \leq_{A_i}) : i \in I\}$ uma cadeia em $\mathcal{F}$, e seja $A = \bigcup_{i \in I} A_i$. Dados $x, y \in A$ definamos $x \leq_A y$ se $x, y \in A_i$ e $x \leq_{A_i} y$. É fácil verificar que a relação $\leq_A$ está bem definida, e é uma relação de ordem parcial em A. Afirmamos que $(A, \leq_A)$ é um conjunto bem ordenado. Seja $\emptyset \neq B \subset A$, e seja

$$J = \{j \in I : B \cap A_j \neq \emptyset\}.$$

Notemos que $\leq_A$ coincide com $\leq_{A_i}$ em A_i para cada $i \in I$. Como $(A_i, \leq_{A_i})$ é bem ordenado para cada $i \in I$, segue que todos os conjuntos $B \cap A_j$, com $j \in J$, tem o mesmo elemento mínimo, que denotaremos por b_0. Segue que b_0 é o elemento mínimo de B. Logo $(A, \leq_A)$ é bem ordenado, ou seja pertence a $\mathcal{F}$. Agora é claro que $(A, \leq_A)$ é uma cota superior da cadeia $\{(A_i, \leq_{A_i}) : i \in I\}$.

Pelo lema de Zorn, $\mathcal{F}$ possui pelo menos um elemento maximal $(A, \leq_A)$. Segue da maximalidade de $(A, \leq_A)$ que $A = X$. Logo $(X, \leq_X)$ é um conjunto bem ordenado.

$(c) \Rightarrow (a)$: Seja $\{X_i : i \in I\}$ uma família não vazia de conjuntos não vazios. Pelo teorema de Zermelo, existe uma boa ordenação para $\bigcup_{i \in I} X_i$. Para cada $i \in I$ seja $f(i)$ o elemento mínimo de X_i. Então $f \in \prod_{i \in I} X_i$.

$(a) \Rightarrow (b)$: Esta é a implicação mais difícil de provar. Seja X um conjunto parcialmente ordenado não vazio no qual cada cadeia é limitada superiormente.

Seja $\mathcal{X}$ a família de todas as cadeias de X. Então $\mathcal{X}$ é um conjunto parcialmente ordenado não vazio, por inclusão de conjuntos.

A estratégia da demonstração é trabalhar com a família de conjuntos $\mathcal{X}$, que é parcialmente ordenada por inclusão, em lugar de trabalhar com o conjunto parcialmente ordenado abstrato X. Depois de provar que $\mathcal{X}$ possui um elemento maximal, será fácil provar que X possui um elemento maximal.

O primeiro passo é caracterizar os elementos maximais de $\mathcal{X}$. Para cada $C \in \mathcal{X}$ seja

$$\hat{C} = \{x \in X : C \cup \{x\} \in \mathcal{X}\}.$$

É claro que $C \subset \hat{C}$. Além disso, C é maximal em $\mathcal{X}$ se e só se $C = \hat{C}$.

Pelo axioma da escolha, existe uma função $f : \mathcal{P}(X) \setminus \{\emptyset\} \to X$ tal que $f(A) \in A$ para cada $A \in \mathcal{P}(X) \setminus \{\emptyset\}$.

Seja $g : \mathcal{X} \to \mathcal{X}$ definida por:

$$g(C) = C \qquad \text{se} \quad C = \hat{C},$$

$$g(C) = C \cup \{f(\hat{C} \setminus C)\} \qquad \text{se} \quad C \neq \hat{C}.$$

A função g está bem definida, pois se $C \neq \hat{C}$, então $f(\hat{C} \setminus C) \in \hat{C} \setminus C$, e portanto $C \cup \{f(\hat{C} \setminus C)\} \in \mathcal{X}$. Além disso, C é maximal em $\mathcal{X}$ se e só se $g(C) = C$.

Diremos que uma família $\mathcal{T} \subset \mathcal{X}$ 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 cadéia em $\mathcal{T}$, então $\bigcup \mathcal{C} \in \mathcal{T}$.

É claro que $\mathcal{X}$ é uma torre. É claro que a interseção de uma família de torres é uma torre. Seja $\mathcal{T}_0$ a interseção de todas as torres de $\mathcal{X}$. Então $\mathcal{T}_0$ é a menor torre de $\mathcal{X}$. Nosso próximo objetivo é provar que $\mathcal{T}_0$ é uma cadeia em $\mathcal{X}$. Isto vai nos dar muito trabalho.

Diremos que $C \in \mathcal{T}_0$ é *comparável* se dado $D \in \mathcal{T}_0$, tem-se que $C \subset D$ ou $D \subset C$.

Para provar que $\mathcal{T}_0$ é cadeia, basta provar que cada $C \in \mathcal{T}_0$ é comparável.

Para provar que cada $C \in \mathcal{T}_0$ é comparável, basta provar que os conjuntos comparáveis em $\mathcal{T}_0$ formam uma torre.

É claro que $\emptyset$ é comparável. É claro também que se $\mathcal{C}$ é uma cadeia de conjuntos comparáveis, então $\bigcup \mathcal{C}$ é comparável. O mais difícil vai ser provar que se C é comparável, então $g(C)$ é comparável também.

Fixemos $C \in \mathcal{T}_0$, C comparável.

Afirmamos que se $D \in \mathcal{T}_0$ e $D \subset C$, $D \neq C$, então $g(D) \subset C$. Como $\mathcal{T}_0$ é torre, $g(D) \in \mathcal{T}_0$. Como C é comparável, tem-se que $g(D) \subset C$ ou $C \subset g(D)$, $C \neq g(D)$. Mas $C \subset g(D)$, $C \neq g(D)$ é impossível, pois $D \subset C$, $D \neq C$ e $g(D) = D$ ou $g(D) = D \cup \{x\}$.

Seja

$$\mathcal{U} = \{D \in \mathcal{T}_0 : D \subset C \quad \text{ou} \quad g(C) \subset D\}.$$

Afirmamos que $\mathcal{U}$ é uma torre. É claro que $\emptyset \in \mathcal{U}$. É claro também que se $\mathcal{D}$ é uma cadeia em $\mathcal{U}$, então $\bigcup \mathcal{D} \in \mathcal{U}$. Falta provar que se $D \in \mathcal{U}$, então $g(D) \in \mathcal{U}$. Há três possibilidades:

(i) $D \subset C$, $D \neq C$. Neste caso já sabemos que $g(D) \subset C$, e portanto $g(D) \in \mathcal{U}$.
(ii) $D = C$. Neste caso $g(D) = g(C)$, e portanto $g(D) \in \mathcal{U}$.
(iii) $g(C) \subset D$. Neste caso $g(D) \supset D \supset g(C)$, e portanto $g(D) \in \mathcal{U}$.

Como $\mathcal{U}$ é torre e $\mathcal{U} \subset \mathcal{T}_0$, segue que $\mathcal{U} = \mathcal{T}_0$. Logo, dado $D \in \mathcal{T}_0 = \mathcal{U}$, tem-se que $D \subset C \subset g(C)$ ou $g(C) \subset D$. Logo $g(C)$ é comparável.

Temos provado assim que os conjuntos comparáveis de $\mathcal{T}_0$ formam uma torre. Segue que cada $C \in \mathcal{T}_0$ é comparável, e daí $\mathcal{T}_0$ é uma cadeia em $\mathcal{X}$.

Como $\mathcal{T}_0$ é torre, temos que $C_0 := \bigcup \mathcal{T}_0 \in \mathcal{T}_0$. Como $\mathcal{T}_0$ é torre, temos que $g(C_0) \in \mathcal{T}_0$, e portanto $g(C_0) = C_0$. Logo C_0 é maximal em $\mathcal{X}$.

Por hipótese existe $m \in X$ tal que $c \leq m$ para todo $c \in C_0$. Como C_0 é uma cadeia maximal, é claro que $m \in C_0$.

Afirmamos que m é um elemento maximal em X. De fato seja $n \in X$, com $m \leq n$. Como C_0 é uma cadeia maximal, segue que $n \in C_0$. Logo $n \leq m$, e portanto $n = m$. Isto completa a demonstração.

Exercícios.

14.A. Seja $X = \{n \in \mathbf{N} : n \geq 2\}$. Dados $m, n \in X$, definamos $m \leq n$ se m divide n.

(a) Prove que $\leq$ é uma relação de ordem parcial em X.

(b) Prove que, dada uma cadeia $C \subset X$ e um elemento $n \in C$, existe apenas um número finito de elementos $n_1, ..., n_k \in C$ que dividem n.

(c) Prove que cada cadeia $C \subset X$ é limitada inferiormente.

(d) Identifique os elementos minimais de X.

14.B. Seja $(X, \leq)$ um conjunto totalmente ordenado com pelo menos dois elementos. Dados $x, y \in X$, escreveremos $x < y$ se $x \leq y$ e $x \neq y$.

(a) Prove que os conjuntos $\{x \in X : a < x\}$, com $a \in X$, junto com os conjuntos $\{x \in X : x < b\}$, com $b \in X$, formam uma sub-base para uma topologia em X, chamada de *topologia da ordem*.

(b) Prove que a topologia usual em $\mathbf{R}$ coincide com a topologia da ordem usual em $\mathbf{R}$.

14.C. Seja E um espaço vetorial, $E \neq \{0\}$. Usando o lema de Zorn prove que cada subconjunto linearmente independente de E está contido em alguma base de E.

14.D. Sejam E e F espaços vetoriais sobre o mesmo corpo, seja E_0 um subespaço vetorial de E, e seja $T_0 : E_0 \to F$ uma aplicação linear. Use o lema de Zorn para provar a existência de uma aplicação linear $T : E \to F$ tal que $Tx = T_0 x$ para todo $x \in E_0$.

14.E. Seja A um anel comutativo com elemento unidade. Um conjunto $I \subset A$ é chamado de *ideal* se verifica as seguintes condições:

(a) $x - y \in I$ para todo $x, y \in I$;

(b) $xy \in I$ para todo $x \in I$, $y \in A$.

Um ideal $I \neq A$ é chamado de *ideal próprio*. Um ideal próprio que não está contido em nenhum outro ideal próprio é chamado de *ideal maximal*. Use o lema de Zorn para provar que cada ideal próprio de A está contido em algum ideal maximal.