

# O axioma da escolha na prática matemática

Fernando Ferreira  
Universidade de Lisboa

O axioma da escolha é frequentemente usado na matemática, por vezes de modo quase imperceptível. Se bem que muitas vezes apenas se usem formas enfraquecidas do axioma (p. ex.,  $AC_\omega$  ou DC), outras vezes ele é usado em todo (ou quase todo) o seu poder. Nesta secção, vamos apresentar alguns exemplos de argumentos que ocorrem na prática matemática e que utilizam o axioma da escolha.

Um primeiro exemplo, já discutido anteriormente, é o da equivalência entre as noções de conjunto infinito e infinito à Dedekind. A primeira propriedade implica a segunda com a ajuda do axioma da escolha. O argumento que usámos na demonstração da Proposição 30 apenas necessita do axioma DC das escolhas dependentes. De facto, pode mesmo argumentar-se que  $AC_\omega$  já é suficiente:

**Proposição 1.** *Em  $Z + AC_\omega$ , todo o conjunto infinito é infinito à Dedekind.*

**Demonstração.** Dado que  $X$  é um conjunto infinito, tem-se

$$\forall n \in \omega \exists Z \in \mathcal{P}(X) (\text{card}(Z) = 2^n).$$

Por  $AC_\omega$ , existe uma sucessão  $n \rightsquigarrow Z_n$  de subconjuntos de  $X$  cada qual com  $2^n$  elementos. Ora,

$$\text{card}(\bigcup_{i=0}^n Z_i) \leq \sum_{i=0}^n \text{card}(Z_i) = \sum_{i=0}^n 2^i = 2^{n+1} - 1 < 2^{n+1}.$$

Tem-se, pois, que  $\forall n \in \omega \exists x \in X (x \in Z^{n+1} \setminus (\bigcup_{i=0}^n Z_i))$ . Novamente por  $AC_\omega$ , existe uma sucessão  $n \rightsquigarrow x_n$  de elementos de  $X$  tais que, para todo  $n \in \omega$ ,  $x_n \in Z^{n+1} \setminus (\bigcup_{i=0}^n Z_i)$ . Claro que esta sucessão é uma injeção de  $\omega$  em  $X$ . Logo,  $X$  é infinito à Dedekind.  $\square$

Dado  $w$  um número real e  $X$  um subconjunto de  $\mathbb{R}$ , diz-se que  $w$  é *ponto aderente* de  $X$  se  $\forall \varepsilon > 0 \exists x \in X |x - w| < \varepsilon$ . Na presença de  $AC_\omega$ , esta noção é equivalente a dizer que existe uma sucessão  $(x_n)_{n \in \omega}$  de elementos de  $X$  a convergir para  $w$ . Com efeito, admitamos que  $w$  é ponto aderente de  $X$ . Então,  $\forall n \in \omega \exists x \in X |x - w| < \frac{1}{n+1}$ . Por  $AC_\omega$ , existe  $f : \omega \rightarrow X$  tal que, para todo  $n \in \omega$ ,  $|f(n) - w| < \frac{1}{n+1}$ . Claro que a sucessão  $x_n = f(n)$ , de elementos de  $X$ , converge para  $w$ . A implicação contrária é imediata e não necessita do axioma da escolha.

**Exercício 1.** Uma função  $f : \mathbb{R} \mapsto \mathbb{R}$  diz-se sequencialmente contínua no ponto  $x \in \mathbb{R}$  se, sempre que  $(x_n)_{n \in \omega}$  é uma sucessão de números reais a convergir para  $x$ , então a sucessão  $(f(x_n))_{n \in \omega}$  converge para  $f(x)$ . Mostre em  $Z + AC_\omega$  que uma função real de variável real  $f$  é contínua num ponto  $x \in \mathbb{R}$  se, e somente se, é sequencialmente contínua em  $x$ .

Dado um espaço vectorial  $V$  sobre um corpo  $K$ , um conjunto  $X$  de vectores de  $V$  diz-se *linearmente independente* se, para toda a sequência finita  $x_1, \dots, x_n$  de elementos de  $X$ , se tem:

$$\forall \lambda_1, \dots, \lambda_n \in K (\lambda_1 x_1 + \dots + \lambda_n x_n = 0 \rightarrow \lambda_1 = \dots = \lambda_n = 0).$$

Um conjunto de vectores  $X$  diz-se uma *base de Hamel* do espaço vectorial  $V$  se for linearmente independente e gerar  $V$ , no sentido em que, para todo  $x \in V$ , existem elementos  $x_1, \dots, x_n \in X$  e escalares  $\lambda_1, \dots, \lambda_n \in K$  tais que  $x = \lambda_1 x_1 + \dots + \lambda_n x_n$ . Um conjunto  $X$  de vectores linearmente independente diz-se *maximal* se, sempre que  $Y$  é um conjunto de vectores que contenha propriamente  $X$ , então  $Y$  já não é linearmente independente. O seguinte lema é fundamental:

**Lema 1.** *Todo o conjunto de vectores linearmente independente maximal é uma base.*

**Demonstração.** Seja  $X$  um conjunto linearmente independente maximal. Considere-se  $x$  um elemento de  $V$ . Por maximalidade,  $X \cup \{x\}$  não é linearmente independente. Então existe uma combinação linear nula de elementos de  $X \cup \{x\}$  sem que todos os coeficientes sejam nulos. Dado que  $X$  é linearmente independente, esta combinação linear nula é necessariamente da forma  $\lambda x + \lambda_1 x_1 + \dots + \lambda_n x_n$ , com  $\lambda \neq 0$ . Sai  $x = -\lambda_1^{-1} \lambda x_1 - \dots - \lambda_n^{-1} \lambda x_n$ .  $\square$

**Teorema 1.** *Todo o espaço vectorial tem uma base de Hamel.*

**Demonstração.** Seja  $V$  um espaço vectorial. Seja  $\mathcal{I}$  o conjunto de todos os subconjuntos  $X$  de  $V$  linearmente independentes (note que  $\emptyset \in \mathcal{I}$ ). Vamos ver que  $\mathcal{I}$  munido da ordem parcial “estar contido” está nas condições de aplicação do lema de Zorn. Com efeito, seja  $\mathcal{C}$  uma cadeia (não vazia) de elementos de  $\mathcal{I}$ . Vamos ver que  $\bigcup \mathcal{C}$  é um conjunto linearmente independente. Sejam  $x_1, \dots, x_n \in \bigcup \mathcal{C}$ . Dado que  $\bigcup \mathcal{C}$  é uma cadeia, existe  $X \in \mathcal{C}$  tal que  $\{x_1, \dots, x_n\} \subseteq X$ . Daqui sai que qualquer combinação linear nula de  $x_1, \dots, x_n$  tem necessariamente todos os coeficientes nulos. Pelo lema de Zorn, existe  $B \in \mathcal{I}$  maximal. Pelo lema anterior,  $B$  é base de  $V$ .  $\square$

O axioma da escolha tem, por vezes, consequências desagradáveis. O axioma permite mostrar a existência de certos objectos “patológicos” que, de outro modo, não existiriam. Curiosamente, as formas enfraquecidas  $AC_\omega$  ou  $DC$  da axioma da escolha não são geralmente suficientes para mostrar a existência das patologias.

Um problema importante da análise matemática consiste em estender a noção de comprimento dum intervalo a subconjuntos mais complicados de  $\mathbb{R}$ . Idealmente, gostaríamos de estender esta noção a todos os subconjuntos de  $\mathbb{R}$ .

Põe-se o problema: será que existe uma função  $\mu : \mathcal{P}(\mathbb{R}) \rightarrow [0, +\infty]$  de tal sorte que:

1.  $\mu(\emptyset) = 0$  e  $\mu(\mathbb{R}) = +\infty$ .
2. se  $(X_n)_{n \in \omega}$  é uma sucessão de subconjuntos mutuamente disjuntos de  $\mathbb{R}$ , então  $\mu(\bigcup_{n \in \omega} X_n) = \sum_{n \in \omega} \mu(X_n)$ .
3. se  $a, b \in \mathbb{R}$  com  $a \leq b$ , então  $\mu([a, b]) = b - a$ .
4. se  $a \in \mathbb{R}$  e  $X \subseteq \mathbb{R}$ , então  $\mu(X) = \mu(a + X)$ , onde  $a + X = \{a + x : x \in X\}$ .

Os dois primeiros requisitos garantem que  $\mu$  é uma *medida* (definida em todos os subconjuntos de  $\mathbb{R}$ ). A segunda propriedade é conhecida como  $\sigma$ -aditividade. O terceiro requisito diz que a medida dum intervalo é o seu comprimento. O último requisito diz que a medida é invariante para translações.

**Lema 2.** *Seja  $\mu$  como acima. Têm-se as seguintes propriedades:*

- i. se  $X, Y \subseteq \mathbb{R}$  são conjuntos disjuntos, então  $\mu(X \cup Y) = \mu(X) + \mu(Y)$ .
- ii. se  $X \subseteq Y \subseteq \mathbb{R}$ , então  $\mu(X) \leq \mu(Y)$ .
- iii. se  $(X_n)_{n \in \omega}$  é uma sucessão crescente de subconjuntos de  $\mathbb{R}$  ( $X_n \subseteq X_{n+1} \subseteq \mathbb{R}$ , para todo  $n \in \omega$ ), então  $\mu(\bigcup_{n \in \omega} X_n) = \sup_{n \in \omega} \mu(X_n)$ .

**Demonstração.** A primeira propriedade é consequência da  $\sigma$ -aditividade de  $\mu$  (e do facto de que  $\mu(\emptyset) = 0$ ). A segunda propriedade sai da anterior: com efeito, se  $X \subseteq Y$ , então  $\mu(Y) = \mu(X \cup (Y \setminus X)) = \mu(X) + \mu(Y \setminus X) \geq \mu(X)$ . Finalmente, argumentemos (iii). Seja  $(X_n)_{n \in \omega}$  uma sucessão crescente de subconjuntos de  $\mathbb{R}$ . Defina-se  $Y_0 = X_0$  e  $Y_{n+1} = X_{n+1} \setminus X_n$ . Note-se que  $\bigcup_{n \in \omega} X_n = \bigcup_{n \in \omega} Y_n$  e que esta última união é disjunta. Vem:

$$\mu(\bigcup_{n \in \omega} X_n) = \mu(\bigcup_{n \in \omega} Y_n) = \sum_{n \in \omega} \mu(Y_n) = \sup_{n \in \omega} \left( \sum_{i=0}^n \mu(Y_i) \right) = \sup_{n \in \omega} \mu\left(\bigcup_{i=0}^n Y_i\right) = \sup_{n \in \omega} \mu(X_n),$$

onde se utiliza o facto de que cada  $X_n$  é a união dos  $Y_i$ , com  $i \leq n$ . □

Vamos ver que, na presença do axioma da escolha, não há uma função  $\mu$  que satisfaça (1)-(4). Mais especificamente, com a ajuda do axioma da escolha, vamos exibir um conjunto  $V$  – o chamado *conjunto de Vitali* – cuja existência de medida leva a uma contradição. Defina-se a seguinte relação de equivalência em  $\mathbb{R}$ :  $x \sim y$  sse  $x - y \in \mathbb{Q}$ . Pelo axioma da escolha, seja  $V$  um conjunto de representantes para esta relação de equivalência. Claramente,  $\mathbb{R} = \bigcup_{q \in \mathbb{Q}} (q + V)$ . Além disso, esta união numerável é disjunta. Com efeito, suponhamos que  $(q + V) \cap (q' + V) \neq \emptyset$ , com  $q, q' \in \mathbb{Q}$ . Seja  $x \in (q + V) \cap (q' + V)$ . Então existem  $r, r' \in V$  tais que  $x = q + r = q' + r'$ . Daqui sai que  $r - r' = q' - q \in \mathbb{Q}$ . Logo  $r \sim r'$  e, portanto,  $r = r'$  (já que ambos  $r$  e  $r'$  estão num conjunto de representantes  $V$ ). Por  $\sigma$ -aditividade,  $\sum_{q \in \mathbb{Q}} \mu(q + V) = +\infty$ . Pela invariância da translação,  $\mu(q + V) = \mu(V)$ , para todo  $q \in \mathbb{Q}$ . Conclui-se que  $\mu(V) > 0$ . Por (iii) do

lema anterior,  $\mu(V) = \sup_{n \in \omega} \mu(V_n)$ , onde  $V_n = [-n, n] \cap V$ . Fixe-se, então,  $k$  suficientemente grande tal que  $\mu(V_k) > 0$ . Ora,  $\bigcup_{q \in [0,1] \cap \mathbb{Q}} (q + V_k) \subseteq [-k, k+1]$ . Por  $\sigma$ -aditividade e as propriedades (3) e (4),  $\sum_{q \in [0,1] \cap \mathbb{Q}} \mu(V) \leq 2k+1$ . Isto é absurdo, pois a série à esquerda da desigualdade é soma infinita de um mesmo elemento não nulo.

Perante o exemplo de Giuseppe Vitali, os analistas escolhem manter as propriedades (1)-(4) mas admitir que a função  $\mu$  não esteja definida em *todos* os subconjuntos de  $\mathbb{R}$ . Em análise matemática desenvolve-se a *medida de Lebesgue* com as propriedades (1)-(4), porém apenas definida nos denominados conjuntos *mensuráveis à Lebesgue*. Os conjuntos mensuráveis à Lebesgue contêm os intervalos e são fechados para complementações e uniões numeráveis. Na presença do axioma da escolha, o conjunto de Vitali é um exemplo dum conjunto que não é mensurável à Lebesgue. No início dos anos setenta do século passado, Robert Solovay exibiu um modelo de  $ZF + DC$  em que todo o conjunto é mensurável à Lebesgue. O resultado de Solovay mostra que a existência de conjuntos não mensuráveis à Lebesgue necessita do axioma da escolha (nem sequer bastando formas enfraquecidas deste).

O exemplo de Vitali pode ser refinado de forma mais ou menos dramática. Um resultado de Stefan Banach e Alfred Tarski publicado em 1924 mostra que é possível particionar a bola unitária de  $\mathbb{R}^3$  num número finito de pedaços de tal modo que seja possível re-arranjar esses pedaços, por meio de rotações e translações, e obter no final duas cópias da bola unitária. Este resultado é conhecido por *paradoxo de Banach-Tarski*. Não se trata realmente dum paradoxo, mas tão-somente dum resultado extremamente contra-intuitivo. Os pedaços de que o resultado de Banach-Tarski fala não são mensuráveis: o axioma da escolha desempenha um papel essencial nesta decomposição da bola unitária, à semelhança do que sucede no exemplo de Vitali.