Um curso fechado e limitado de Análise Real

© 2026 por Renan M. Mezabarba[†]

Última atualização: 18 de janeiro de 2026.

Quis custodiet ipsos custodes?[†]

Juvenal, Sátiras, VI, linha 347.

[†] Quem vigia os vigilantes?

Sumário

Prefácio

O presente material é a reformulação de um primeiro texto de mesmo título [29], que comecei a escrever por volta de 2021, mas cuja narrativa não se mostrou apropriada para, efetivamente, acompanhar cursos de Análise: a versão original continha uma abordagem bastante generalista, o que dificultaria tanto o trabalho de quem por ventura fosse usá-lo para lecionar (eu) quanto a vida de quem se propusesse a estudá-lo (*você*[†]). Por isso, a presente versão distribui os assuntos de forma diferente e mais organizada (a fim de facilitar a vida de eventuais docentes) com uma escrita mais amigável (para não assustar demais possíveis estudantes, que constituem o *público alvo* do texto).

Em resumo, os três primeiros capítulos se destinam a cobrir *com alguma folga* a ementa do curso Análise I do Bacharelado em Matemática da Universidade Estadual de Santa Cruz (Ilhéus/BA), com a grade em vigor no ano de 2024. Explicitamente: “números reais, topologia da reta, limite e continuidade de funções na reta e derivação”. No entanto, a *folga* é grande o suficiente para permitir *adaptações*, de modo a incluir tópicos de Análise II[‡] e mesmo Análise III[††], tornando assim o material utilizável em cursos de outras instituições.

“Mas por que fazer isso? O Brasil já tem seus cânones da Análise!”.

É justamente pela existência dos cânones que me sinto confortável em propor uma abordagem alternativa, mais alinhada ao panorama matemático de quem se interessa por Fundamentos da Matemática, Topologia Geral e Teoria dos Conjuntos. Se o texto parecer generalista, filosófico ou excessivamente teórico para algumas pessoas, os cânones permanecerão onde sempre estiveram, prontos para guiar quem preferir seguir o percurso tradicional e a receita consagrada.

Para conciliar a generalidade que eu realmente gostaria de empregar com as limitações burocrático-culturais intrínsecas ao contexto em que este material será usado, os conteúdos que extrapolam a ementa de Análise I supracitada foram distribuídos ao longo de seções *extras*, a fim de não atrapalhar o fluxo de estudo de quem prefere o caminho menos acelerado adotado aqui. Isto nos traz à discussão da organização dos capítulos.

Cada um dos três primeiros *capítulos* corresponde a um terço do curso de Análise I. Porém, num cenário ideal, os conteúdos do Capítulo 0 (conjuntos, noções de cardinalidade, corpos ordenados, supremos e ínfimos) podem ser omitidos por já terem sido tratados em outras disciplinas, o que permite dar mais atenção aos capítulos seguintes[‡‡].

[†]Espero que não se importe com a intimidade de me referir diretamente a *você* em vez de “ao leitor”. Esta é a forma gramaticalmente aceitável que encontrei de escrever de maneira mais *neutra*.

[‡]Explicitamente: “integração de Riemann (própria e imprópria), sequências e séries de funções e topologia de espaços euclidianos”.

[††]Explicitamente: “caminhos no espaço euclidiano; continuidade e diferenciabilidade de funções reais de várias variáveis; integrais curvilíneas; continuidade e diferenciabilidade de funções vetoriais; teorema da função inversa e consequências”.

[‡‡]Convém destacar que as cinco primeiras seções do Capítulo 0 são *fortemente* baseadas no texto exposto em [30], embora sem a mesma profundidade. Mas não se preocupe: o autor autorizou.

As *seções* $(N.M)$, por sua vez, correspondem *aproximadamente* aos conteúdos individuais das aulas. Por exemplo: a Seção 1.7 *deveria* corresponder à 7-ésima aula do 1-ésimo terço do curso — o que em linguagem natural seria a *oitava* aula da *segunda parte do curso*[†]. Porém, algumas seções maiores podem requerer um pouco mais de tempo a depender da profundidade desejada, do ânimo da turma, dos feriados na semana, etc. E justamente para calibrar a profundidade desejada, cada seção tem duas subseções: a "Essencial" e a "Extras". De modo geral, a primeira traz os assuntos indispensáveis para serem abordados no curso para o qual este texto oficialmente se destina, enquanto a segunda sugere discussões de aprofundamento (cuja leitura é opcional aqui, mas que pode ser obrigatória a depender do seu contexto).

De volta à descrição dos capítulos, o Capítulo 1 discute limites, continuidade e topologia, exatamente nesta ordem. Embora seja ruim de um ponto de vista bourbakista, este ordenamento busca partir dos assuntos já dominados pelo público alvo a fim de fundamentá-los e generalizá-los. Mas há duas pegadinhas.

(I) Muitas vezes, intervalos abertos serão usados no lugar de épsilons e deltas, como preparação para os argumentos topológicos por meio de *abertos*.

(II) Os limites considerados serão os limites de *redes*, e não *apenas* os de sequências e funções. Isto permitirá tratar, simultaneamente, de diversas noções de limites já nas primeiras seções, inclusive integrais de Riemann!

As seções extras desse capítulo, por sua vez, trazem a perigosa sugestão de substituir expressões do tipo "$|x-y|<\varepsilon$" por outras como "$\|x-y\|<\varepsilon$" ou "$d(x,y)<\varepsilon$", o que na prática significa expandir as discussões na reta para espaços normados e métricos, respectivamente. Derivadas são apresentadas no final do capítulo do modo tradicional, enquanto a formulação de Carathéodory, mais apropriada para dimensões maiores, é sugerida como conteúdo extra[‡].

O Capítulo 2 se aprofunda nos principais teoremas da Análise em contexto básico. As três primeiras seções trazem os aspectos topológicos da compacidade e conexidade, manifestados nos Teoremas de Heine-Borel, Heine-Cantor e do Valor Intermediário na reta real, enquanto as seções extras generalizam esses resultados para outras classes de espaços. Integrais de Riemann e o Teorema Fundamental do Cálculo ficam por conta da última seção desse capítulo — e aqui, como tudo é *Extra* segundo a ementa de Análise I adotada na UESC, a distinção "Essencial Vs. Extras" se torna apenas uma sugestão sobre o que fazer se sobrarem aulas ao final do período letivo.

Evidentemente, os assuntos nas seções extras dos Capítulos 1 e 2 podem/devem ser tratados em disciplinas de Análise II ou III. Se for o seu caso, o Capítulo 3 compila mais algumas discussões que podem ser úteis nesses cenários — por tal razão, a distinção "Essencial vs. Extras" é abandonada nesse capítulo. Eu gostaria de ter acrescentado mais coisas, mas minha paciência é finita.

Assim, quem decidir *enfrentar todas* as seções e subseções dos três primeiros capítulos terá feito, ao final da jornada, um curso de Análise I *robusto*, com muitos acenos a tópicos mais avançados de Análise II, Análise III, Análise Funcional, Topologia, etc. Ainda assim, nada impede que você *enfrente* "apenas" o Essencial desses capítulos — se for o seu caso, ao final da jornada você terá feito um curso de Análise I *honesto*, com a vantagem de ter se guardado para muitas surpresas nos cursos seguintes. E em ambos os casos, a escolha do verbo "enfrentar" é séria: será trabalhoso.

[†]Neste texto, zero é natural. Meu texto, minhas regras.

[‡]E se você nunca ouviu falar nisso, por favor, dê uma chance, você não irá se arrepender.

Sem brincadeira: há muitos exercícios distribuídos ao longo do texto, para serem feitos, preferencialmente, durante a "leitura", bem como exercícios adicionais, agrupados em seções especiais. São pelo menos 500 exercícios, desde verificações simples que você deveria fazer sem que ~~eu~~ o texto precisasse pedir, mas que você achou que não precisava, até alguns resultados mais específicos garimpados nos confins do `math.stackexchange`. Para te ajudar (ou não), eles serão classificados subjetivamente em três tipos principais ($\star$, ${}^{\star}_{\star}$ e ${}^{\star}_{\star\star}$) descritos a seguir.

($\star$) Verificações (que deveriam ser) corriqueiras. Embora nem todos os exercícios nesta classe sejam imediatos, a maioria consiste em "abrir" as definições e fazer as "contas" naturais ao contexto em questão. Não é o tipo de problema que "cai em prova", mas são bons aquecimentos caso você ainda não tenha prática em escrever demonstrações mais complicadas.

(${}^{\star}_{\star}$) Exercícios típicos de listas e provas. Consistem em aplicar resultados previamente demonstrados (tanto no texto quanto em questões anteriores), ou "reciclar" ideias e argumentos a fim de testar o seu entendimento do assunto: quanto maior sua familiaridade, mais fácil será para você fazer os malabarismos necessários.

(${}^{\star}_{\star\star}$) Estes geralmente são mais difíceis, elaborados ou "*filosóficos*". Na prática, eles "testam" não apenas a familiaridade com o assunto, mas também dão a oportunidade de exercitar um pouco mais a criatividade e a reflexão.

Comentários do tipo "(verifique!)${}^{\star}_{\star}$", "(por quê?)*"e afins também devem ser encarados como propostas implícitas de exercício — e os expoentes apenas classificam os problemas de acordo com os critérios acima.

Tente fazer exercícios de todos os tipos: lidar com os "fáceis" te ajudará a ganhar ritmo para enfrentar os "médios", e ao resolver alguns "difíceis" você *pode* aprender técnicas para transformar os exercícios "médios" em "fáceis". Outra dica: se já souber EXATAMENTE como fazer um exercício, pule para o próximo. Mais uma: se ficar muito tempo num exercício sem conseguir resolvê-lo, pule para o próximo e volte mais tarde![†] A última: é altamente recomendável que você já tenha feito pelo menos algum curso de Cálculo e, se possível, Introdução à Teoria dos Números (para ter familiaridade com indução) e Álgebra Linear também (para aproveitar as discussões em espaços normados).

Por fim, agradecimentos às minhas turmas de Análise: desde a primeira turma de ~~cobaias~~ Análise II, ainda na UFES em 2019, que me fez acreditar que os *livros do Elon* não precisam ser a única opção para lecionar Análise no Brasil, até a mais recente, na UESC, que me fez perceber a necessidade de reformular meu texto original — e aqui estendo o "muito obrigado" à Pryscilla Silva, que ao usar este material em suas aulas, teve a paciência de compartilhar comigo muitas observações pertinentes, entre outras coisas.

Renan M. Mezabarba

Ilhéus-BA, 18 de janeiro de 2026.

P.S.: Sempre que possível, leia as notas de rodapé.

P.P.S.: Traçados ~~deste tipo~~ são intencionais — resquícios dos meus tempos de entusiasta da Desciclopédia.

[†]E aqui uma sugestão para estudantes iniciantes: sempre que possível, estude (ou ao menos leia) o "essencial" da aula antes da aula, pois isto tornará ~~possivelmente~~ as aulas bem mais proveitosas. Evidentemente, isto não se aplica apenas às subseções "essenciais" deste texto, mas ao essencial de qualquer disciplina.

Capítulo 0

A reta real

Intuitivamente, a *reta real* é um objeto geométrico: um *segmento retilíneo sem saltos*, i.e., *contínuo*. Por outro lado, como segmentos dessa reta podem ser *somados* (copiados e justapostos) e *multiplicados*[‡], segue que a reta também é um objeto *algébrico*. Daí, uma pergunta natural a se fazer é: como conciliar as duas noções a fim de *descrever*, matematicamente, a reta real?

A resposta usual faz uso da *linguagem de conjuntos*: a reta real será definida como *um conjunto* (de pontos), cujos *aspectos geométricos* desejados serão abstraídos por uma *relação de ordem* que deverá capturar, de alguma forma, as noções de *linearidade* e *continuidade*; os *aspectos algébricos*, por sua vez, serão descritos por meio de *operações binárias* que imitarão as operações usuais que aprendemos na *escola*. Nesta parte do curso lidaremos com todos esses aspectos, a fim de entender a definição matemática da reta real.

0.0 Linguagem: revisão de conjuntos e funções

0.0.0 Essencial

§0 Conjuntos

A palavra "conjunto" é uma daquelas típicas expressões (*atômicas*) que não se explicam por meio de outras expressões mais simples, como *tempo*, *espaço*, *ser*, etc. Costuma ficar a cargo da (vida em) sociedade ensinar o significado dessas coisas: no caso, conjuntos podem ser entendidos como *agrupamentos de objetos* ou *coleções de indivíduos* que partilham algum tipo de característica comum num certo contexto. EXEMPLOS: conjunto das pessoas numa sala, conjunto dos torcedores de um time, conjunto dos times de algum esporte coletivo num campeonato, conjunto dos campeonatos desse esporte coletivo, etc. Porém, usaremos tais noções em contexto matemático, e trataremos apenas de conjuntos formados por objetos de natureza matemática.

Dados *objetos matemáticos* x e A, vamos assumir que apenas dois casos podem ocorrer (e necessariamente um deles ocorrerá): "$x \in A$" (lido como "x pertence a A", "x é elemento de A" ou "x é membro de A") ou o contrário, indicado por "$x \notin A$" (lido como "x não pertence a A", "x não é elemento de A" ou "x não é membro de A"). **Porém, como não se define explicitamente o que significa *ser* conjunto** e, muito menos, o que é a relação de pertinência "$\in$", algumas coisas precisam ser "*convencionadas*".

[‡]Como feito por Descartes na infância da Geometria Analítica. Para saber mais, confira a obra de Tatiana Roque [34].

Definição 0.0.0. Para conjuntos A e B:

(i) escreveremos "$A \subseteq B$" para abreviar a afirmação "para todo x, se $x \in A$, então $x \in B$", lida como "A é **subconjunto** de B", ou "A **está contido em** B";

(ii) escreveremos "$A \nsubseteq B$" para abreviar a negação de "$A \subseteq B$", i.e., para indicar que "existe $x \in A$ tal que $x \notin B$";

(iii) escreveremos "$A \subsetneq B$" para abreviar "$A \subseteq B$ e $A \neq B$" e, em tais situações, diremos que A é **subconjunto próprio** de B. ¶

Axioma da Extensão. *Dois conjuntos são iguais se, e somente se, têm os mesmos elementos. Em notação mais econômica:* $A = B \Leftrightarrow (A \subseteq B)$ *e* $(B \subseteq A)$.

Exemplo 0.0.1 (Requer traquejo com Aritmética Básica)**.** Sejam X o conjunto dos números inteiros múltiplos de 10 e Y o conjunto dos números inteiros simultaneamente múltiplos de 2 e 5. Como mostrar que $X = Y$?

- Primeiro, deve-se verificar que $X \subseteq Y$. Isto ocorre pois se $x \in X$, então $x = 10k$ para algum k inteiro, e isso permite escrever tanto $x = 2 \cdot (5k)$ quanto $x = 5 \cdot (2k)$, i.e., x é simultaneamente múltiplo de 2 e de 5.

- Em seguida, mostra-se que $Y \subseteq X$. Neste caso, se $x = 2k = 5k'$ para certos inteiros k e k', então $5k'$ deve ser divisível por 2 (por quê?)*, o que só é possível se k' for múltiplo de 2, digamos $k' = 2s$, e daí $x = 5k' = 10s$.

Encerrada a demonstração, fica um aviso: não se acostume com esse nível de detalhes para argumentos que você já deveria saber. ▲

Exercício 0.0 ($\star$)**.** Sejam A o conjunto dos números inteiros pares, B o conjunto dos números inteiros múltiplos de 3 e C o conjunto dos números inteiros múltiplos de 6. Determine as relações de inclusão entre A, B e C. ■

Entre outras coisas, o exercício acima indica que precisamos de um modo mais prático para *descrever* conjuntos. A seguir, listam-se modos comuns de descrever o conjunto A do exercício anterior:

- $A = \{x : x \in \mathbb{Z} \text{ e } x \text{ é par}\}$;
- $A = \{x \in \mathbb{Z} : x \text{ é par}\}$;
- $A = \{x \in \mathbb{Z} : \text{existe } n \in \mathbb{Z} \text{ tal que } x = 2n\}$;
- $A = \{2z : z \in \mathbb{Z}\}$.

Em geral, ao descrever um conjunto no formato

$$\underbrace{\{\ldots}_{1^{\text{a}}\text{ parte}} : \underbrace{\ldots\}}_{2^{\text{a}}\text{ parte}}$$

entende-se que o conjunto considerado é formado pelos elementos que satisfazem o que se escreve na primeira parte, **tais que** as condições impostas na segunda parte são satisfeitas. Assim, na primeira descrição de A, entende-se que seus elementos são todas as coisas (já que não há restrições) que são números inteiros pares (a condição imposta na segunda parte). Analogamente, na segunda descrição, A é descrito como a coleção de todos os números inteiros (como imposto na primeira parte) que são pares (já que está é a condição da segunda parte).

Exercício 0.1 ($\star$)**.** Descreva, como acima, os demais conjuntos do exercício anterior. ■

Observação 0.0.2. É comum encontrar a notação "$\{\ldots|\ldots\}$" em vez de "$\{\ldots : \ldots\}$". Você pode usá-la em seus exercícios, desde que não implique com a minha forma de escrever, padrão nos textos de Teoria dos Conjuntos, e que será adotada aqui. △

Outra alternativa prática de descrição, geralmente utilizada para conjuntos *finitos*[†], consiste em listar os elementos do conjunto. Por exemplo: para $S := \{a, 4, \triangle\}$, temos $a \in S$, $4 \in S$ e $\triangle \in S$, enquanto $0 \notin S$, $\nabla \notin S$, etc. Importante destacar o uso do símbolo "$:=$" acima: a ideia é indicar que S foi "definido" ou "declarado" como o conjunto escrito após o símbolo "$:=$". Para ilustrar, observe que no próximo exercício faz sentido usar "$=$" pois S já foi declarado acima.

Exercício 0.2 ($\star$)**.** Mostre que $S = \{\triangle, 4, a\}$. Dica: Axioma da Extensão. ■

Preciosismos à parte[‡], o exercício acima indica que a ordem em que os elementos de um conjunto são descritos é irrelevante. É por essa razão que conjuntos do tipo $\{a, b\}$ são chamados de **pares não ordenados**.

Exercício 0.3 ($\star$)**.** Mostre que $\{x, x\} = \{x\}$ para qualquer x. ■

OBSERVAÇÃO: por conta deste resultado, é de "bom tom" não grafar duas vezes o mesmo elemento ao descrever um conjunto com a notação "$\{\ldots\}$".

Exercício 0.4 ($\star$)**.** Descreva o conjunto $\{0, 1\}$ por meio da notação $\{\ldots : \ldots\}$. ■

A fim de encerrar esta primeira parte da revisão, a próxima definição trás o restante das notações e operações usuais entre conjuntos.

Definição 0.0.3.

(i) Denota-se $\emptyset := \{x : x \neq x\}$, *coleção* que será chamada de **conjunto vazio** por razões *óbvias*: não existe x com $x \in \emptyset$, já que o contrário daria $x \neq x$.

(ii) Para X e Y conjuntos, considera-se $X \setminus Y := \{x \in X : x \notin Y\}$, que denota a **diferença** entre X e Y, também chamada de **complementar de** Y **em** X.

(iii) Para conjuntos A e B, $A \cap B := \{x : x \in A \text{ e } x \in B\}$ e $A \cup B := \{x : x \in A \text{ ou } x \in B\}$ denotam os conjuntos chamados, respectivamente, de **interseção** e (**re**) **união** dos conjuntos A e B. Em particular, A e B são **disjuntos** se $A \cap B = \emptyset$. ¶

Observação 0.0.4. Lembre-se: na linguagem comum, "e" funciona como um *agregador*, enquanto "ou" indica *alternativa*: por exemplo, os pontos da reta nos intervalos $(-\infty, 1)$ *e* $(3, +\infty)$ constituem a solução da inequação $x^2 - 4x + 3 > 0$, o que em linguagem de conjuntos se expressa por meio da reunião $(-\infty, 1) \cup (3, +\infty)$. No entanto, **não** faz sentido dizer que tal reunião é composta por todo x tal que $x \in (-\infty, 1)$ **e** $x \in (3, +\infty)$, pois este "e" (da linguagem matemática usual) indica *simultaneidade* — e não há x com as duas propriedades *ao mesmo tempo*: o correto é dizer que $x \in (-\infty, 1)$ **ou** $x \in (3, +\infty)$.

Cabe ainda destacar que o "ou" matemático não é exclusivo: sempre que dissermos "$x \in A$ ou $x \in B$", deve-se entender que *pelo menos* um dos casos deve ocorrer, o que não inviabiliza a ocorrência de ambos. △

[†]Finitude é um tópico que será discutido ainda neste capítulo.
[‡]Ninguém costuma reprovar por escrever "$=$" em vez de "$:=$".

Proposição 0.0.5. *Para todo conjunto A ocorre $\emptyset \subseteq A$.*

Demonstração. Dado x qualquer, a implicação "$x \in \emptyset \Rightarrow x \in A$" é verdadeira por *vacuidade*, já que "$x \in \emptyset$" é falso. Alternativamente: se a *inclusão* fosse falsa, deveria existir $x \in \emptyset$ com $x \notin A$, mas não existe $x \in \emptyset$, absurdo[†]. □

Observação 0.0.6 (Contido vs. pertence)**.** Cuidado para não confundir *pertinência* e *continência*:

- "$x \in y$" significa que "x" é um dos elementos de "y";
- "$x \subseteq y$" significa "todo elemento de x é também elemento de y".

Assim, embora $\emptyset \subseteq A$ ocorra para qualquer conjunto A, nem sempre ocorre $\emptyset \in A$. Na verdade, fora de contextos mais formais, é raro que se tenha $\emptyset \in A$. Veja que, por exemplo, $\emptyset \notin \emptyset$, já que o contrário é dizer que $\emptyset$ tem um elemento. Mesmo assim, $\emptyset \subseteq \emptyset$. A raiz dessa confusão é, possivelmente, oriunda do fato de que muitas vezes se diz "y contém x" a fim de expressar "$x \in y$". △

Exercício 0.5 ($\star$)**.** Convença-se de que $\emptyset \neq \{\emptyset\}$. ■

Observação 0.0.7. Não escreva "$\{\quad\}$" para representar o conjunto vazio. Em um mundo onde impressoras falham e cada vez mais pessoas lidam com TDAH, quem lê pode se perguntar se foi erro de impressão ou distração de quem escreveu. Já "$\emptyset$" é claro e sem margem para dúvida. "*Ah, mas no Ensino Médio eu escrevia assim!*". Não importa. △

Exercício 0.6 ($\star$)**.** Sejam A, B e C conjuntos. Mostre as identidades, inclusões, equivalências e implicações a seguir.

a) $A \cup B = B \cup A$ e $A \cap B = B \cap A$.

b) $A \cup (B \cup C) = (A \cup B) \cup C$ e $A \cap (B \cap C) = (A \cap B) \cap C$.

c) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ e $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.

d) $A \subseteq A$, $A \subseteq B$ e $B \subseteq C \Rightarrow A \subseteq C$.

e) $A \subseteq B \Leftrightarrow A \cap B = A \Leftrightarrow A \cup B = B$.

f) $A \subseteq B \Rightarrow C \setminus B \subseteq C \setminus A$.

g) $A \setminus B = \emptyset \Leftrightarrow A \subseteq B$.

h) $A \setminus A = \emptyset$, $A \setminus \emptyset = A$ e $A \setminus (A \setminus B) = A \cap B$.

i) $A \cup B = (A \setminus B) \cup (B \setminus A) \cup (A \cap B)$. ■

§1 Funções

O advento das *funções*, seja por invenção ou descoberta, deflagrou uma das maiores mudanças de paradigma na Matemática por permitir incorporar as noções de movimento e variação aos *modelos* que, até então, tratavam apenas de situações estáticas e posicionais.

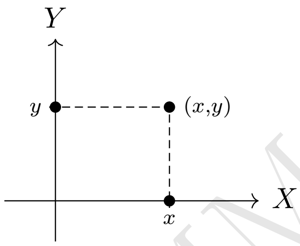
Embora hoje se apresente como um conceito simples, alguns séculos separam as primeiras menções explícitas às funções da "definição" apresentada por Dedekind na segunda metade do Século XIX:

[†]Por exemplo: a sentença "todas as piscinas da minha casa são olímpicas" é verdadeira se a minha casa não tiver piscinas.

Conceito de função. *Uma função é uma regra f que associa cada elemento x de um conjunto X a um único elemento $f(x)$ de um conjunto Y.*

Se, por um lado, a conceituação acima parece englobar os casos clássicos aprendidos na infância (funções *polinomiais*, *trigonométricas*, etc.), por outro lado ela empurra para debaixo do tapete a definição de “regra”. Um modo mais honesto consiste em apelar a *pares ordenados*.

Diferente do que ocorre no caso não ordenado, em que $\{a, b\} = \{b, a\}$ mesmo com $a \neq b$, pode-se *convencionar* escrever (a, b) com o intuito de ter o seguinte comportamento: $(a, b) = (c, d)$ se, e somente se, $a = c$ e $b = d$. Com tal *dispositivo*, usualmente xingado de **par ordenado**, passa a fazer sentido definir o **produto cartesiano** $X \times Y := \{(x, y) : x \in X \text{ e } y \in Y\}$ entre os conjuntos X e Y, cujos membros são todos os pares ordenados da forma (x, y) com $x \in X$ e $y \in Y$: noutras palavras, $X \times Y$ apenas abstrai um típico plano *cartesiano*.



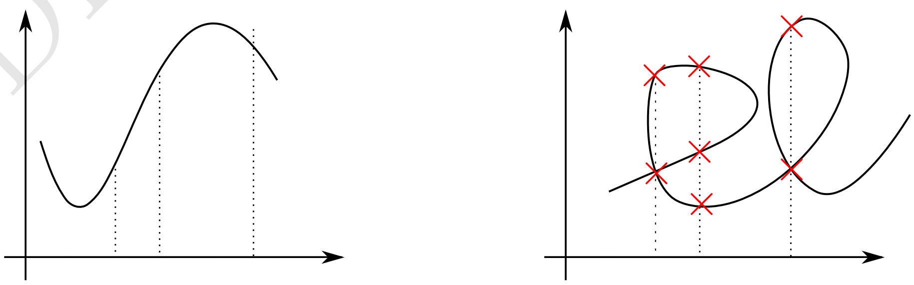
Dado que um par ordenado (x, y) pode ser interpretado como uma *mini-regra* que faz sua *primeira coordenada* x corresponder à sua *segunda coordenada* y, é natural pensar em *regras* que associam elementos de X a Y como subconjuntos de $X \times Y$.

Definição 0.0.8 (Bourbaki, 1939)**.** Uma **função** (ou **mapa** ou **aplicação**) **de** X **em** Y é um subconjunto $f \subseteq X \times Y$ tal que:

(i) para todo $x \in X$ existe $y \in Y$ com $(x, y) \in f$ (cada x se associa a pelo menos um y);

(ii) se $(x, y), (x, z) \in f$, então $y = z$ (o y associado a x é único).

Escreve-se $f : X \to Y$ ou $X \xrightarrow{f} Y$ para indicar que f é uma função de X em Y. ¶



Resumindo: a definição acima apenas formaliza, por meio de conjuntos, o critério que você aprendeu na Escola (ou em Pré-Cálculo, Cálculo I, etc.) para reconhecer *gráficos* de funções: se uma reta vertical interceptar o gráfico em mais de um ponto, então não é função. Não lembra o que são gráficos e muito menos como esboçá-los? Pois relembre!

Acima, o conjunto X costuma ser chamado de **domínio** da função f, enquanto Y é o seu **codomínio** (também chamado de *contradomínio*). Uma vez que a cada $x \in X$ corresponde um único $y \in Y$ com $(x, y) \in f$, faz sentido atribuir a y uma notação que remeta ao elemento x: no caso, faz-se $y := f(x)$ (ou ainda $x \overset{f}{\mapsto} f(x)$), e xinga-se $f(x)$ de **imagem de** x pela função f. Por sua vez, o subconjunto de Y formado por todos os elementos da forma $f(x)$, conforme x *varia* em X, é chamado de **imagem da função**, e denotado por $\mathrm{im}\,(f)$. Em símbolos: $\mathrm{im}\,(f) := \{f(x) : x \in X\}$.

Exemplo 0.0.9 (Funções polinomiais)**.** Dado um *polinômio na indeterminada* t, i.e., uma *expressão* da forma $a_0 + a_1 t + \ldots + a_n t^n$, que abreviaremos como $p(t)$, onde $n \in \mathbb{N}$, $a_0, \ldots, a_n$ são *números reais*[†] e t indica apenas um símbolo *indeterminado*, passa a fazer sentido substituir cada ocorrência de "t" na expressão $p(t)$ por um *número real* x fixado, o que *produz* o *número real* $p(x) := a_0 + a_1 x + \ldots + a_n x^n$.

Dessa forma, pode-se dizer que $p := \{(x, p(x)) : x \in \mathbb{R}\}$ relaciona cada $x \in \mathbb{R}$ ao número $p(x) \in \mathbb{R}$. Uma vez que tal associação é claramente *funcional*[‡], ganha-se uma função $p \colon \mathbb{R} \to \mathbb{R}$, que faz $x \mapsto p(x)$ para cada $x \in \mathbb{R}$. Funções desse tipo são chamadas de (funções) **polinomiais**. ▲

Exemplo 0.0.10 (Funções racionais)**.** Mais geralmente, consideram-se expressões da forma $r(t) := \frac{p(t)}{q(t)}$ em que ambos $p(t)$ e $q(t)$ são polinômios na indeterminada t. Desta vez, só faz *sentido* substituir "t" em $r(t)$ por um número real x fixado nas situações em que se garantir $q(x) \neq 0$, pois a divisão por 0 não é realizável em *corpos*. Assim, a expressão $r(t)$ induz uma função r cujo domínio é $\{x \in \mathbb{R} : q(x) \neq 0\}$, e que faz $r(x) := \frac{p(x)}{q(x)}$ para cada x no domínio de r. Funções assim costumam ser chamadas de (funções) **racionais**. ▲

Exercício 0.7 ($\star$)**.** Para funções f e g de X em Y, mostre que $f = g$ se, e somente se, $f(x) = g(x)$ para todo $x \in X$. ■

Definição 0.0.11. Se $f \colon X \to Y$ e $g \colon Y \to Z$ são funções, então fica *bem definida* uma função $g \circ f \colon X \to Z$ dada pela identidade

$$(g \circ f)\,(x) := g \circ f(x) := g(f(x)),$$

para todo $x \in X$, chamada de (função) **composta** (ou **composição**) entre f e g, em geral denotada apenas por $g \circ f$. ¶

$$\begin{tikzcd}
x \arrow[r, mapsto] & f(x) & y \arrow[r, mapsto] & g(y) \\
X \arrow[rr, "f"] \arrow[rrr, bend right, "g \circ f"'] & & Y \arrow[r, "g"] & Z \\
x \arrow[rrr, mapsto, bend right] & & & g(f(x))
\end{tikzcd}$$

Figura 0.0: Esquematicamente, f vem antes de g, já que as setas costumam sair da direita para esquerda, no sentido de nossa escrita. Por outro lado, para descrever a regra da composição, precisamos escrever g primeiro, já que os cálculos são feitos "de dentro para fora": primeiro, calcula-se $f(x)$, para daí calcular $g(f(x))$. Note que, apesar disso, as duas notações indicam a mesma coisa.

[†]Formalmente ainda não definimos o que é um número real. No entanto, você já fez Cálculo I. De modo geral, os exemplos farão uso de informações que nós já *temos*, embora elas ainda não estejam implementadas formalmente na disciplina/no texto. O Elon *faz* isso e ninguém reclama, então não é agora que isso se tornará um problema, certo?

[‡]No sentido de que se $(x, y), (x, z) \in p$, então $y = z$.

A definição acima faz sentido pois $f(x) \in Y$ para todo $x \in X$ e $Y = \mathrm{dom}\,(g)$, de modo que g *sabe* o que *fazer* com $f(x)$. Um pouco mais formalmente, podemos definir

$$g \circ f := \{(x, z) \in X \times Z : g(f(x)) = z\},$$

que satisfaz as condições para ser uma função do tipo $X \to Z$:

✓ para cada $x \in X$ existe $z \in Z$ tal que $(x, z) \in g \circ f$, basta tomar $z := g(f(x))$;

✓ se $(x, z), (x, z') \in g \circ f$, então $z = z'$ já que $z = g(f(x)) = z'$.

Exercício 0.8 ($\star$)**.** Sejam $f\colon W \to X$, $g\colon X \to Y$ e $h\colon Y \to Z$ funções. Mostre que as funções $h \circ (g \circ f)$ e $(h \circ g) \circ f$ são iguais. ■

Exercício 0.9 ($\star$)**.** Para um conjunto X, escreve-se $\mathrm{Id}_X\colon X \to X$ para denotar a **função identidade de** X, definida por $\mathrm{Id}_X(x) := x$ para cada x em X. Mostre que $f \circ \mathrm{Id}_X = f$ para qualquer função f cujo domínio é X e $\mathrm{Id}_X \circ g = g$ para qualquer função g cujo codomínio é X. ■

0.0.1 Extras

§0 Análise para quê?

Os números naturais foram criados por Deus, todo o resto é trabalho da humanidade.

Leopold Kronecker (1886).

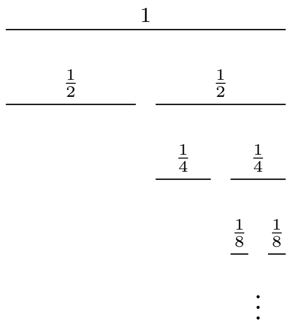
Faz parte do folclore matemático atribuir a frase anterior ao algebrista Leopold Kronecker (1823-1891), como uma síntese de seu ceticismo perante os diversos métodos *infinitários* e não construtivos que se propagaram na Matemática a partir da segunda metade do Século XIX [12]. Longe de ser uma mera declaração teológica, ela expressa a confiabilidade que temos diante dos *métodos aritméticos* usuais, explicitamente ancorados na (aparente?) realidade imediata, em contraponto aos argumentos do *Cálculo*, que frequentemente dependem de suposições incomuns ao nosso cotidiano intrinsecamente *finito*, como no caso das *séries*.

Considere, por exemplo, a série

$$\sum_{n=1}^{\infty} \frac{1}{2^n} = \frac{1}{2} + \frac{1}{4} + \ldots + \frac{1}{2^n} + \ldots \qquad (\dagger)$$

que manifesta a ideia de somar iteradamente parcelas da forma $\frac{1}{2^n}$ conforme se toma n cada vez maior. Embora possa parecer artificial, esse tipo de animal surge naturalmente em problemas que envolvem o *cálculo* de áreas *curvilíneas*, por meio do chamado *método da exaustão*[†]. O ponto a chamar atenção, porém, é o seguinte: embora cada estágio finito desse processo seja facilmente calculável, não é completamente óbvio o que poderia *significar* realizar uma soma com infinitas parcelas.

[†]É o que está por trás das ideias de *integração* que você estudou em Cálculo I.



Com argumentos aritmético-geométricos do tipo ilustrado acima, é razoável *convencionar* que o valor para (†), seja lá qual for, deve corresponder ao *número* para o qual as *somas parciais* se *dirigem*: no caso, tal número *deveria* ser 1, já que $1 = \frac{1}{2} + \frac{1}{2} = \frac{1}{2} + \frac{1}{4} + \frac{1}{4} = \cdots$ Ainda assim, trata-se de uma *convenção*: não há vida suficiente para efetuar *todas* as somas e verificar uma igualdade *empírica*.

Agora, o que significa dizer que tais somas parciais se dirigem para algum valor? Além disso, o que impediria que certas somas se dirigissem para números diferentes? Mais ainda, como determinar os valores de tais somas infinitas?

Evidentemente, nada impede que tais perguntas sejam respondidas de forma vaga e intuitiva — ou apenas ignoradas. No caso da série proposta, por exemplo, um argumento muito comum para justificar as estimativas ~~abanando as mãos~~ sem muito esforço é o seguinte: ao escrever $S := \sum_{n=1}^{\infty} \frac{1}{2^n}$, chega-se a

$$2S = 2 \cdot \left(\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \ldots\right) = 1 + \left(\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \ldots\right) = 1 + S$$

e, consequentemente, $S = 1$, justamente o que o raciocínio intuitivo geométrico sugeria!

Não é difícil adaptar o *método* para mostrar que $\sum_{n=1}^{\infty} \frac{1}{3^n} = \frac{1}{2}$, $\sum_{n=1}^{\infty} \frac{1}{4^n} = \frac{1}{3}$ e, mais geralmente, $\sum_{n=1}^{\infty} \frac{1}{m^n} = \frac{m}{1-m}$ sempre que $m > 1$, identidades compatíveis com suas respectivas interpretações geométricas. O que acontece, porém, ao aplicar tal "metodologia" na determinação do valor da série $\sum_{n=1}^{\infty} 2^n$? Como antes, ao escrever $T := \sum_{n=1}^{\infty} 2^n$, chega-se a

$$2T = 4 + 8 + 16 + \ldots + 2^{n+1} + \ldots \Rightarrow 2T + 2 = 2 + 4 + 8 + \ldots = T \Rightarrow T = -2,$$

resultado que, desta vez, não parece *certo*, por discordar do que a interpretação geométrica sugere.

Isto nos conduz a uma pergunta que não pode ser evitada — nem respondida de modo vago sem que se torne explícita certa desonestidade intelectual: *por que o truque algébrico preguiçoso pareceu funcionar nos primeiros casos mas falhou no último?*

Responder a esse tipo de pergunta é, pelo menos historicamente, um dos papéis da Análise. Pode-se dizer que ela surgiu do processo natural de revisão metodológica, típico do *modus operandi* científico, no contexto do *Cálculo* de Leibniz-Newton. Embora, a princípio, tratasse-se de um movimento voltado a justificar (e expandir) os resultados da área com base em conceitos geométricos menos vagos, seus praticantes não tardaram a empregar a *linguagem de conjuntos* no processo de retraduzir e *sintetizar* o Cálculo a partir de noções aparentemente tão sólidas quanto os números criados pelo deus de Kronecker.

§1 Fundamentos debaixo do tapete

Objetos matemáticos não existem da mesma forma que as entidades físicas existem. Por mais que procure na sua casa ou em qualquer outro lugar, você nunca encontrará *o número 2*, por exemplo. É claro que você pode encontrar símbolos, desenhos ou pinturas que remetam ao seu *significado*, mas nunca *o* próprio número 2, já que este é uma ideia, a abstração de um conceito. Isso é diferente do que ocorre nas Ciências Biológicas, por exemplo, cujos objetos de estudo são palpáveis e *detectáveis*.

Ao longo dos milênios, chegou-se ao aparente consenso de utilizar *o método axiomático* para lidar com o *problema ontológico*[†] na Matemática. Na prática, isto consiste em assumir como válidas algumas afirmações acerca de objetos que não definimos, mas que parecem razoáveis o bastante para serem entendidas como "verdadeiras" para, a partir delas, deduzir e definir "todo o resto". Como a discussão anterior indicou, isto é relativamente desnecessário para questões matemáticas corriqueiras, mas se torna indispensável quando problemas menos usuais entram em cena. Nesse sentido, a adaptabilidade dos conjuntos favoreceu o seu uso no papel dos "objetos não definidos" em praticamente todas as áreas da Matemática. Mas, novamente: conjuntos não existem *materialmente*, por mais que pareçam intuitivos. Logo, precisamos assumir axiomaticamente como tais animais se comportam.

Há basicamente dois axiomas sobre conjuntos que as pessoas utilizam em seus primeiros contatos com o tema: o Axioma da Extensão, já postulado no começo da seção, e o Axioma da Abstração. Este último é usado implicitamente sempre que se emprega a notação $\{\ldots : \ldots\}$. Grosso modo, postula-se o seguinte: dada uma *propriedade* $\mathcal{P}$, existe o conjunto $\{x : x \text{ tem a propriedade } \mathcal{P}\}$. É claro que isto deixa o problema de explicar o que é *propriedade*, mas a intuição basta para a discussão: tudo o que se definiu na seção anterior poderia ser justificado por meio desses dois axiomas, até mesmo pares ordenados!

Exercício 0.10 $(\substack{\star \\ \star})$**.** Para x e y quaisquer, <u>defina</u> $(x, y) := \{\{x\}, \{x, y\}\}$. Mostre que para a, b, c e d quaisquer, $(a, b) = (c, d)$ se, e somente se, $a = c$ e $b = d$. ■

Com um pouco mais de paciência, não seria difícil ver que funções podem ser interpretadas como conjuntos. Mais adiante, veremos que números também podem ser "definidos" via conjuntos, de modo que se chega à conclusão inevitável: é razoável assumir, para propósitos formais, que tudo é conjunto. Do ponto de vista metodológico mencionado na subsubseção anterior, esta *seria* uma vitória enorme: a partir de dois axiomas básicos, *seríamos* capazes de justificar e reconstruir uma quantidade monumental de *fatos* matemáticos, o que tornaria bastante razoável aceitar as consequências menos intuitivas que encontrássemos pelo caminho. Contudo, há um problema — como sugerido pelo tempo verbal empregado.

[†]Exercício: pesquise o significado da palavra `ontologia`, mas não faça isso no Youtube, por favor!

Exercício 0.11 ($^{\star}_{\star}$)**.** Considere $R := \{x : x \notin x\}$. Mostre que $R \in R$ se, e somente se, $R \notin R$. ■

O Axioma da Abstração diz que R deveria *existir*. Porém, se R existir, chega-se a uma contradição, pois tanto $R \in R$ quanto sua negação devem ocorrer simultaneamente! Uma das soluções encontradas para resolver tal problema (conhecido como Paradoxo de Russell) foi abandonar o Axioma da Abstração e, em seu lugar, assumir o Axioma da Separação que, grosso modo, postula o seguinte: dada uma *propriedade* $\mathcal{P}$ e um conjunto A, existe o conjunto $\{x \in A : x \text{ tem a propriedade } \mathcal{P}\}$. Assim, em vez de assumir a habilidade irrestrita de definir conjuntos, supõe-se apenas a capacidade de determinar *subconjuntos* de conjuntos previamente conhecidos. O preço disso é que a lista de axiomas precisou ser estendida. Contudo, tais axiomas não serão abordados aqui — com exceção do Axioma da Escolha. Para saber mais, você pode conferir as referências [10, 30].

0.1 Funções e cardinalidade

0.1.0 Essencial

§0 Injeções, sobrejeções e bijeções

Definição 0.1.0. Uma função $f : X \to Y$ será dita:

(i) **injetora** (ou *injetiva*, *injeção*, etc.) se para quaisquer $x, x' \in X$, a ocorrência de $f(x) = f(x')$ acarretar $x = x'$;

(ii) **sobrejetora** (ou *sobrejetiva*, *sobrejeção*, *sobre* Y, etc.) se $\operatorname{im}(f) = Y$ e

(iii) **bijetora** (ou *bijetiva*, *bijeção*, etc.) se $f : X \to Y$ for injetora e sobrejetora. ¶

Para aquecer, você pode começar com o próximo

Exercício 0.12 ($^{\star}_{\star}$)**.** Sejam $f : X \to Y$ e $g : Y \to Z$ funções.

a) Mostre que se g e f são injetoras, então $g \circ f$ é injetora.

b) Mostre que se g e f são sobrejetoras, então $g \circ f$ é sobrejetora.

c) Conclua que se g e f são bijetoras, então $g \circ f$ é bijetora.

Bijeções estão intimamente ligadas com a noção de *invertibilidade*[†] para funções.

Definição 0.1.1. Dizemos que uma função $g : Y \to X$ é *uma* **inversa** de $f : X \to Y$ se $f \circ g = \mathrm{Id}_Y$ e $g \circ f = \mathrm{Id}_X$. A função f é **invertível** se tem uma inversa. ¶

Observação 0.1.2. Na prática, dizer que g é *uma* inversa de f equivale a dizer que $g : Y \to X$ satisfaz o seguinte:

$$\text{para quaisquer } x \in X \text{ e } y \in Y,\ g(y) = x \Leftrightarrow f(x) = y. \tag{0.0}$$

De fato, se g é *uma* inversa de f, então a condição acima é satisfeita: por um lado, se $g(y) = x$, então $f(g(y)) = f(x)$, com $f(g(y)) = y$ pois $f \circ g = \mathrm{Id}_Y$ vale por hipótese; por outro lado, se $f(x) = y$, então $g(f(x)) = g(y)$, com $g(f(x)) = x$ pois $g \circ f = \mathrm{Id}_X$ vale por hipótese. Reciprocamente, se $g : Y \to X$ satisfaz a condição (0.0), então valem as identidades $g \circ f = \mathrm{Id}_X$ e $f \circ g = \mathrm{Id}_Y$: para a primeira, note que se $f(x) = y$, então (0.0) assegura $g(y) = g(f(x)) = x$; a segunda é *análoga*[‡]. △

[†] O verbo é “inverter” e não “inverser”.

[‡] Afirmações dessa natureza devem ser encaradas como exercícios do tipo ($\star$).

Em outras palavras, *uma* inversa de g desfaz tudo o que f faz. Como consequência:

Exercício 0.13 ($\star$)**.** Mostre que uma inversa, se existir, é única, ou seja: se g e g' são inversas de f, então $g = g'$. ■

Em certo sentido, o resultado acima diz que a inversa de uma função f fica completamente determinada por f (caso exista). Por isso, é comum indicá-la de modo a fazer alusão à função f: neste caso, a notação mais comum para a inversa de f é f^{-1}.

Exemplo 0.1.3. A função $f\colon \mathbb{Z} \to \mathbb{Z}$ dada por $f(z) := z + 1$ tem inversa dada por $g(z) := z - 1$. De fato,

$$g(f(z)) = f(z) - 1 = z + 1 - 1 = z \quad \text{e} \quad f(g(z)) = g(z) + 1 = z - 1 + 1 = z.$$

Futuramente, veremos que a função exponencial $\exp\colon \mathbb{R} \to (0, +\infty)$ tem a função logaritmo $\ln\colon (0, +\infty) \to \mathbb{R}$ como inversa: de suas aulas de Cálculo I, você deve se lembrar de que $e^x = y$ se, e somente se, $\ln(y) = x$. Compare isso com a condição (0.0). ▲

Exemplo 0.1.4. A função $h\colon \mathbb{Z} \to \mathbb{Z}$ dada por $h(z) := z^2$ não é invertível: por valer $h(-1) = h(1) = 1$, não pode haver g satisfazendo (0.0), pois $g(1)$ seria simultaneamente igual a -1 e 1. Em particular, note que h não é injetora e nem sobrejetora. ▲

O que tudo isso tem a ver com bijeções?

Proposição 0.1.5. *Uma função $f\colon X \to Y$ é bijetora se, e somente se, é invertível.*

Demonstração. Assumindo que f é bijetora, vamos definir uma função $g\colon Y \to X$ que provaremos ser a inversa de f. Não há muito o que fazer:

$$g := \{(y, x) \in Y \times X : f(x) = y\}.$$

Por f ser sobrejetora, para todo $y \in Y$ existe $x \in X$ tal que $f(x) = y$. Logo, para todo $y \in Y$ existe $x \in X$ tal que $(y, x) \in g$. Agora, se $(y, x), (y, x') \in g$, então $f(x) = f(x')$ (por quê?)*, donde a injetividade de f assegura $x = x'$. Isto mostra que g é função de $Y \to X$. Para verificar que g é a inversa de f, basta notar que g satisfaz a condição (0.0) da Observação 0.1.2 por construção.

A recíproca, isto é, "se $f\colon X \to Y$ é invertível, então f é bijetora", é consequência do próximo exercício. □

Exercício 0.14 ($\substack{\star\\\star}$)**.** Sejam $f\colon X \to Y$ e $g\colon Y \to Z$ funções.

a) Mostre que se $g \circ f$ é injetiva, então f é injetiva.

b) Mostre que se $g \circ f$ é sobrejetora, então g é sobrejetora.

c) Conclua que se $Z = X$ e $g = f^{-1}$, então f é bijetora. ■

Corolário 0.1.6. *Se $f\colon X \to Y$ é bijeção, então $f^{-1}\colon Y \to X$ é bijeção.*

Demonstração. A função f^{-1} é invertível por ter f como inversa. □

§1 A noção de cardinalidade

Os tipos de função discutidos no começo da seção permitem comparar o *tamanho* dos conjuntos de modo bastante prático.

Definição 0.1.7. Diremos que dois conjuntos **têm a mesma cardinalidade** ("quantidade de elementos") se existir uma bijeção entre os dois. ¶

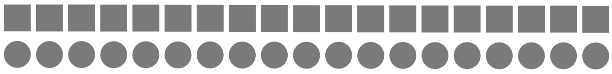


Figura 0.1: Há tantos círculos quanto quadrados, mesmo sem saber *quantos*.

Proposição 0.1.8. *Sejam A, B e C conjuntos.*

(i) *Existe uma bijeção de A para A.*

(ii) *Se existe uma bijeção de A para B, então existe uma bijeção de B para A.*

(iii) *Se existe uma bijeção de A para B e outra bijeção de B para C, então existe uma bijeção de A para C.*

Demonstração. O primeiro item segue por Id_A ser bijeção, enquanto o terceiro item decorre do fato de que a composição de bijeções é bijeção (Exercício 0.12). O segundo item é o Corolário 0.1.6. □

Intuitivamente, a proposição acima diz que ao escrever

$$A \approx B \Leftrightarrow \text{existe bijeção } A \to B,$$

define-se uma *relação binária*[†] entre conjuntos que se comporta, essencialmente, como a relação de igualdade: "$A \approx A$", "$B \approx A$ sempre que $A \approx B$" e "$A \approx C$ sempre que $A \approx B$ e $B \approx C$". Esse tipo de coisa tem um nome: trata-se de uma *relação de equivalência*[‡].

Em vez de determinar quando dois conjuntos *têm* a mesma cardinalidade, podemos usar funções para detectar quando um conjunto tem *mais elementos* do que outro.

Definição 0.1.9. Para conjuntos X e Y, vamos fixar as seguintes notações:

(i) "$X \precsim Y$" será usada para indicar a existência de injeção $X \to Y$;

(ii) "$X \prec Y$" será usada para indicar "$X \precsim Y$" e "$X \not\approx Y$";

(iii) "$Y \succsim X$" será usada para indicar a existência de sobrejeção $Y \to X$;

(iv) "$Y \succ X$" será usada para indicar "$Y \succsim X$" e "$X \not\approx Y$". ¶

A semelhança com os símbolos "$\leq$" e "$<$", usualmente empregados no contexto de ordenação numérica, é intencional: ela serve para nos lembrar que as propriedades de $\precsim$ e $\prec$ se parecem com as propriedades de $\leq$ e $<$, respectivamente. E de fato, tais relações realmente comparam as *cardinalidades* dos conjuntos.

Exercício 0.15 ($\star\star$). Para conjuntos X, Y e Z, mostre que

[†]Subseção 0.1.1 §0.
[‡]Subseção 0.1.1 §1.

a) $X \precsim X$;

b) se $X \precsim Y$ e $Y \precsim Z$, então $X \precsim Z$;

c) se $X \approx Y$, então $X \precsim Z$ se, e somente se, $Y \precsim Z$.

Faça o mesmo trocando $\precsim$ por $\prec$, $\succsim$ e $\succ$. ■

Formalmente, tais relações remetem às ordens que serão abordadas na próxima seção. Todavia, com a terminologia que será utilizada, $\precsim$ não pode satisfazer a *antissimetria* para conjuntos: existem conjuntos X e Y com $X \neq Y$ tais que $X \precsim Y$ e $Y \precsim X$ (dê exemplos)*. Isto ocorre pois $\precsim$ é uma ordem não sobre os conjuntos, mas sobre suas *cardinalidades*.

Teorema 0.1.10 (Cantor-Bernstein)**.** *Se $X \precsim Y$ e $Y \precsim X$, então $X \approx Y$, i.e., se existem injeções da forma $X \to Y$ e $Y \to X$, então existe bijeção $X \to Y$.*

O resultado acima é uma banalidade nas situações em que X e Y são *finitos*: como veremos, em tais situações, $X \precsim Y$ se, e somente se, o *número de elementos de* X, digamos m, é menor do que ou igual ao *número de elementos de* Y, digamos n, de modo que a ocorrência simultânea de $Y \precsim X$ acarreta a desigualdade oposta $n \leq m$, resultando em $m = n$. Com isso dito, note que o enunciado não menciona números ou *finitude*, que ainda não foram formalmente introduzidos. O fato de ser possível demonstrá-lo no contexto que se desenrola pode ser interpretado como um indicativo de que as definições adotadas cumprem bem o papel de abstrair a noção de cardinalidade. Sua demonstração, porém, será postergada (cf. Subseção 0.4.1 §0): há questões mais urgentes a serem discutidas.

Exercício 0.16 $({}^{\star}_{\star\star})$**.** Usando o Teorema de Cantor-Bernstein, mostre que $\mathbb{N} \approx \mathbb{N} \times \mathbb{N}$.■

0.1.1 Extras

§0 Relações binárias e inversas

Dados conjuntos X e Y, uma **relação** (**binária**) R entre X e Y é um subconjunto de $X \times Y$.

(i) Para um par $(x, y) \in X \times Y$, escreve-se $x \, R \, y$ para indicar que o par (x, y) é membro da relação R, enquanto x e y são ditos R**-relacionados**. A ocorrência de $(x, y) \notin R$ será indicada por $x \not R \, y$.

(ii) O **domínio** da relação R é o conjunto $\operatorname{dom}(R) := \{x : \text{existe } y \in Y \text{ com } x \, R \, y\}$.

(iii) A **imagem** da relação R é o conjunto $\operatorname{im}(R) := \{y : \text{existe } x \in X \text{ com } x \, R \, y\}$.

Quando $X = Y$, diz-se apenas que R é uma relação em X.

Exemplo 0.1.11 (Relação de igualdade)**.** Fixado um conjunto X, $\Delta_X := \{(x, y) \in X \times X : x = y\}$ é chamada de *relação de igualdade* em X. Daí, de acordo com a definição anterior, pode-se escrever $x \, \Delta_X \, y$ para indicar que $(x, y) \in \Delta_X$, i.e., $x = y$. ▲

Exemplo 0.1.12 (Partes e inclusão)**.** Fixado um conjunto X, faz sentido considerar a *coleção* de *todos os subconjuntos de* X, denotada por $\wp(X)$ e chamada de **conjunto das partes de** X, simbolicamente: $\wp(X) := \{A : A \subseteq X\}$. Por exemplo:

(i) para $X := \emptyset$, $\wp(\emptyset) = \{\emptyset\}$, já que $\emptyset$ é o único subconjunto de $\emptyset$;

(ii) para $X \coloneqq \{0,2\}$, $\wp(X) = \{\emptyset, \{0\}, \{2\}, \{0,2\}\}$, pois $\emptyset$ e X sempre são subconjuntos de X e, no caso, os demais subconjuntos possíveis são $\{0\}$ e $\{2\}$;

(iii) para $X \coloneqq \mathbb{N}$, ocorre $\{0\}, \{1\}, \{2\}, \ldots, \{0,1\}, \{0,2\}, \ldots \in \wp(\mathbb{N})$, bem como $\{n : n$ é par$\}$, $\{n : n$ é ímpar$\}$, $\{n : n$ é primo$\}, \ldots \in \wp(\mathbb{N})$, além dos típicos $\emptyset, \mathbb{N} \in \wp(\mathbb{N})$. TODO subconjunto de $\mathbb{N}$ é, por definição, membro de $\wp(\mathbb{N})$; oportunamente, veremos que se trata de um conjunto bem grande.

De qualquer forma, para X fixado, a relação de inclusão entre subconjuntos de X define, como a frase sugere, uma relação binária $\subseteq$ na *família*† $\wp(X)$ de todos os subconjuntos de X: explicitamente, $\subseteq \coloneqq \{(A,B) : A \subseteq B \subseteq X\}$. ▲

Exemplo 0.1.13 (*Curvas* e gráficos)**.** Em posse de *estruturas algébricas*, é possível utilizar expressões algébricas a fim de *relacionar variáveis*. Por exemplo, a *expressão polinomial* $x^2 + y^2 = 1$ induz a relação binária $S \coloneqq \{(x,y) \in \mathbb{R} \times \mathbb{R} : x^2 + y^2 = 1\}$. Quando se representa $\mathbb{R} \times \mathbb{R}$ graficamente como o plano cartesiano usual, o subconjunto S *passa a corresponder* aos pontos do plano que *distam* precisamente 1 da origem $(0,0)$.

Em particular, S *não* determina uma função da forma $\mathbb{R} \to \mathbb{R}$ pois para um mesmo $x \in \operatorname{dom}(S)$ existem $y, y' \in \operatorname{im}(S)$ distintos e relacionados a x: explicitamente, se $x^2 + y^2 = 1$, então $y^2 = x^2 - 1$ e, como veremos, em tal situação pode-se concluir apenas que $|y| = \sqrt{x^2 - 1}$, o que dá margem a $y = \sqrt{x^2 - 1}$ e $y' = -\sqrt{x^2 - 1}$, ambos relacionados ao mesmo x. ▲

Definição 0.1.14. Dada uma relação binária R, a **relação inversa** de R, denotada por R^{-1}, é a relação $R^{-1} \coloneqq \{(y,x) : x \mathrel{R} y\}$. ¶

Exercício 0.17 ($^{\star}_{\star}$)**.** Para uma relação binária R, mostre que:

a) $x \mathrel{R} y \Leftrightarrow y \mathrel{R^{-1}} x$ para quaisquer x e y;

b) $\operatorname{dom}(R) = \operatorname{im}(R^{-1})$;

c) $\operatorname{im}(R) = \operatorname{dom}(R^{-1})$; e

d) $(R^{-1})^{-1} = R$. ■

Exercício 0.18 ($^{\star}_{\star\star}$)**.** Supondo que $X = Y$ e $R \subseteq X \times X$, interprete geometricamente a definição de R^{-1} como sendo a “rotação” de R em torno da “diagonal” Δ_X. ■

Esta discussão permite revisar a noção de função inversa. Como toda função é uma relação binária, vemos que toda função $f : X \to Y$ admite uma *relação inversa* f^{-1}. Assim, a Proposição 0.1.5 estabelece critérios para que tal relação inversa seja uma função.

§1 Equivalências, partições e cardinais

Definição 0.1.15. Uma relação binária $\sim$ num conjunto X é dita uma **relação de equivalência** se $\sim$ for

(i) **reflexiva**, i.e., se para todo $x \in X$ ocorrer $x \sim x$,

(ii) **simétrica**, i.e., se para quaisquer $x, y \in X$, a ocorrência de $x \sim y$ acarretar $y \sim x$, e

(iii) **transitiva**, i.e., se para quaisquer $x, y, z \in X$, a ocorrência simultânea de $x \sim y$ e $y \sim z$ acarretar $x \sim z$.

†Não custa frisar: neste texto, “**conjunto**”, “**coleção**” e “**família**” são tratados como sinônimos!

Diremos também que x e y são **$\sim$-equivalentes** sempre que ocorrer $x \sim y$, com a omissão do sufixo "$\sim$" quando a relação estiver clara pelo contexto. ¶

Uma relação de equivalência estabelece um critério por meio do qual objetos a princípio distintos podem ser vistos como iguais, ao mesmo tempo em que separa outros objetos distintos pelo mesmo critério. Dessa forma, não espanta que a *relação de igualdade* ($x \sim y$ se, e somente se, $x = y$) seja o exemplo óbvio de equivalência.

Exemplo 0.1.16 (Horóscopo)**.** Frequentemente, praticantes da (pseudociência chamada de) `Astrologia` fazem uso implícito das relações de equivalência. De fato, de um ponto de vista informal, signos determinam uma "relação de equivalência" no "conjunto" de todas as pessoas:

- ✓ toda pessoa tem o mesmo signo de si mesma;
- ✓ se P tem o mesmo signo de P', então P' tem o mesmo signo de P;
- ✓ se P tem o mesmo signo de P' e esta tem o mesmo signo de P'', então P e P'' têm o mesmo signo.

Se a coisa parasse por aí, a `Astrologia` seria inofensiva. No entanto, é comum encontrar asserções do tipo "o comportamento X é característico do signo Y", o que sugere duas alternativas: ou a afirmação é falsa, ou *toda pessoa* do signo Y apresenta o comportamento X. Esse tipo de máxima ajuda a entender um dos usos mais comuns das relações de equivalência: a simplificação. Com efeito, se tais afirmações *astrológicas* fossem verdadeiras, então para entender os padrões comportamentais de *toda a humanidade* bastaria estudar os comportamentos de doze *representantes*, um de cada signo, algo bem mais simples do que estimar o comportamento individual dos oito bilhões de habitantes do planeta. Por *sorte*, signos estimam tão somente as datas de nascimento de seus portadores. ▲

Exemplo 0.1.17 (Paridade)**.** Recordemo-nos de que os números naturais podem ser classificados como *pares* ou *ímpares*: **pares** são os múltiplos de dois, **ímpares** são os outros. Isso pode ser usado para determinar uma relação de equivalência $\sim$ em $\mathbb{N}$: $m, n \in \mathbb{N}$ serão ditos $\sim$-equivalentes se tiverem a mesma *paridade*. Assim, $0, 2, 4, 6, \ldots$ são $\sim$-equivalentes entre si, $1, 3, 5, 7 \ldots$ são $\sim$-equivalentes entre si, enquanto 0 e 1 não são $\sim$-equivalentes, por exemplo. Note que há certos comportamentos *algébricos* que não dependem dos números escolhidos, e sim de suas paridades: a soma de *quaisquer* dois *ímpares* é *par*, o produto entre quaisquer *ímpares* é *ímpar*, etc. Isto sugere a possibilidade de realizar operações diretamente com as *classes* dos pares e dos ímpares em vez de lidar com seus *infinitos* representantes. ▲

Exemplo 0.1.18 (Restos da divisão por n)**.** Para generalizar o exemplo anterior, pode-se considerar a seguinte relação binária: para $n \in \mathbb{N}$ fixado e $x, y \in \mathbb{N}$, escreveremos $x \sim_n y$ a fim de indicar que x e y têm o mesmo *resto* na *divisão* por n. Verificar que tal relação $\sim_n$ é reflexiva, simétrica e transitiva é um bom exercício para quem se lembra de como fazer divisões. Ocorre que, como antes, certos comportamentos algébricos não dependem dos representantes escolhidos: por exemplo, a soma de quaisquer dois números com resto 2 na divisão por 3 terá resto 1, enquanto o produto de quaisquer dois números com resto 1 na divisão por 3 ainda terá resto 1. ▲

Um efeito colateral inevitável das relações de equivalência é a segregação dos elementos do conjunto em *classes de equivalência*. Mais precisamente:

Definição 0.1.19. Para uma relação de equivalência $\sim$ sobre um conjunto X, diremos que o conjunto $\{y : x \sim y\}$ é a **$\sim$-classe de equivalência de** x. A notação varia com o contexto: é comum escrever $[x]$, $[x]_\sim$, $\pi(x)$, $\overline{x}$, etc. Na dúvida, convêm explicitar a notação que será usada no começo das discussões. ¶

Com *relação* aos exemplos anteriores:

(i) a classe de equivalência de uma pessoa P com respeito aos signos astrológicos seria a coleção de todas as pessoas que têm o mesmo signo de P, consequentemente, existem apenas doze classes de equivalência (correspondentes aos signos possíveis);

(ii) a classe de equivalência de um número $n \in \mathbb{N}$ com respeito à paridade é a coleção dos naturais que têm a mesma paridade de n; logo, existem apenas duas classes, a dos pares e a dos ímpares;

(iii) a classe de equivalência de um número $p \in \mathbb{N}$ com respeito aos restos da divisão por n é a coleção dos números naturais que têm o mesmo resto na divisão, o que leva à conclusão de que existem precisamente n classes de equivalência (correspondentes aos restos possíveis na divisão por n).

Para facilitar a discussão de como uma relação de equivalência *particiona* o seu domínio, vamos introduzir brevemente a generalização do conceito de reunião.

Definição 0.1.20. Para um conjunto $\mathcal{S}$, define-se $\bigcup \mathcal{S} := \{x : \text{existe } S \in \mathcal{S} \text{ com } x \in S\}$, a **reunião da família** $\mathcal{S}$. Nas ocasiões em que $\mathcal{S} := \{S_i : i \in \mathcal{I}\}$ para algum conjunto $\mathcal{I}$ fixado, também é comum escrever $\displaystyle\bigcup_{i\in\mathcal{I}} S_i$ ou $\bigcup_{i\in\mathcal{I}} S_i$. ¶

O dispositivo acima apenas cria um modo bastante esperto de evitar abominações notacionais como "$S_0 \cup S_1 \cup \ldots$" quando se quer expressar uma reunião (possivelmente) *infinita* de conjuntos. Fora isso, ela não traz novidades: note, por exemplo, que para $\mathcal{S} := \{X, Y\}$, tem-se $\bigcup \mathcal{S} = X \cup Y$ (verifique!)*.

Proposição 0.1.21. *Sejam X um conjunto e $\sim$ uma relação de equivalência sobre X. Ao denotar por C_x a $\sim$-classe de equivalência de x para cada $x \in X$, valem as afirmações:*

(i) para todo $y \in X$, existe $x \in X$ com $y \in C_x$ (i.e., $X = \bigcup_{x\in X} C_x$);

(ii) para quaisquer $x, y \in X$ ocorre $C_x = C_y$ ou $C_x \cap C_y = \emptyset$;

(iii) para quaisquer $x, y \in X$, $C_x = C_y$ se, e somente se, $x \sim y$.

Demonstração. O primeiro item decorre da reflexividade de $\sim$: como $y \sim y$, tem-se $y \in C_y$.[†] Os dois itens seguintes seguem do próximo exercício. □

Exercício 0.19 ($^{\star}_{\star\star}$)**.** Sejam $\sim$ uma relação binária em X e $x, y \in X$ elementos quaisquer.

[†] "*Mas não era para ser x?! Por que as letras trocam tanto???!!*". Leia além das letras: nas sentenças do tipo "para todo $y \in X$, existe $x \in X$ tal que...", expressa-se que para qualquer elemento escolhido no conjunto X, existe *um* elemento em X que satisfaz o que se pede com respeito ao primeiro. No presente caso, calhou de que o próprio objeto escolhido no começo tinha a propriedade desejada! Percebeu que tem problemas com esse tipo de linguagem? Parabéns! Perceber o problema é sempre o primeiro passo! O segundo passo é tomar a iniciativa de revisar isso por conta própria nos diversos materiais disponíveis em bibliotecas, Internet, etc. "*Vídeos aleatórios do Youtube?*" Evite.

a) Mostre que se $\sim$ é transitiva, então "$x \sim y \Rightarrow C_y \subseteq C_x$".

b) Mostre que se $\sim$ é simétrica e transitiva, então "$x \sim y \Rightarrow C_x = C_y$".

c) Mostre que se $\sim$ é reflexiva, então "$C_x \subseteq C_y \Rightarrow x \sim y$".

d) Conclua que valem as condições (ii) e (iii) da proposição anterior. Dica: para (ii), o que ocorre com C_z se $z \in C_x \cap C_y$? ■

A última proposição mostra que $X/\sim := \{C_x : x \in X\}$, chamado de **quociente** de X por $\sim$, é uma família de subconjuntos de X que se enquadra como exemplo de *partição*.

Definição 0.1.22. Uma família $\mathcal{P}$ de subconjuntos não vazios de X é uma **partição** de X se valerem as seguintes condições:

(i) para todo $x \in X$ existe $P \in \mathcal{P}$ com $x \in P$ (i.e., $X = \bigcup \mathcal{P}$); e

(ii)[†] se $P, Q \in \mathcal{P}$ e $P \neq Q$, então $P \cap Q = \emptyset$. ¶

Exercício 0.20 ($\star\atop\star$). Mostre que se $\sim$ é uma relação de equivalência sobre X, então $X/\sim$ é uma partição de X. ■

Como o nome sugere, uma partição de X *particiona* o conjunto X em *partes* ou blocos *dois a dois disjuntos*, de modo que cada elemento de X está precisamente em apenas um membro de $\mathcal{P}$. Assim, faz sentido dizer que dois elementos de X são $\mathcal{P}$-*equivalentes* se pertencerem ao mesmo membro de $\mathcal{P}$. Como você deve ter suspeitado, isto define uma relação de equivalência legítima.

Proposição 0.1.23. *Se $\mathcal{P}$ for uma partição de X, então a relação $\sim_{\mathcal{P}}$ definida por*

$$u \sim_{\mathcal{P}} v \Leftrightarrow \exists P \in \mathcal{P} \text{ tal que } \{u, v\} \subseteq P$$

é uma relação de equivalência em X. Além disso, $\mathcal{P} = X/\sim_{\mathcal{P}}$.

Demonstração. A relação $\sim_{\mathcal{P}}$ é

✓ reflexiva, pois dado $x \in X$ existe $P \in \mathcal{P}$ com $x \in P$, e daí $\{x\} = \{x, x\} \subseteq P$,

✓ simétrica, pois se $\{x, y\} \subseteq P \in \mathcal{P}$, então $\{x, y\} = \{y, x\} \subseteq P \in \mathcal{P}$, e

✓ transitiva, pois se $\{x, y\} \subseteq P \in \mathcal{P}$ e $\{y, z\} \subseteq P' \in \mathcal{P}$, então $P \cap P' \neq \emptyset$, acarretando $P = P'$ e, por conseguinte, $\{x, z\} \subseteq \{x, y\} \cup \{y, z\} \subseteq P \in \mathcal{P}$.

A igualdade $\mathcal{P} = X/\sim_{\mathcal{P}}$ segue pois $P = [x]_{\sim_{\mathcal{P}}}$ para quaisquer x e P com $x \in P \in \mathcal{P}$ (verifique!)*. □

Exemplo 0.1.24. Para a relação $\sim_n$ do Exemplo 0.1.18, as classes de equivalência correspondem precisamente a todos os possíveis restos pela divisão por n. Assim, chamando por R_i a coleção dos naturais que têm resto i na divisão por n, segue que $\mathbb{N}/\sim_n = \{R_0, R_1, \ldots, R_{n-1}\}$. Dito isso, observe que ao chamar por $\bar{i}$ a *classe* de equivalência de i, verifica-se $\bar{i} = R_i$. Desse modo, seria lícito escrever, por exemplo, $\mathbb{N}/\sim_2 = \left\{\overline{0}, \overline{1}\right\}$, ou ainda $\mathbb{N}/\sim_2 = \left\{\overline{12}, \overline{13}\right\}$, já que $\overline{0} = \overline{12}$ na relação $\sim_2$ (0 e 12 são divisíveis por 2) e $\overline{1} = \overline{13}$ (1 e 13 têm resto 1 na divisão por 2). Como você já deve ter imaginado, trata-se de um fenômeno mais geral. ▲

[†] Costuma-se expressar a condição (ii) como "os membros de $\mathcal{P}$ são dois a dois disjuntos".

Definição 0.1.25. Um subconjunto $R \subseteq X$ é uma **classe de representantes** de uma

(i) relação de equivalência $\sim$ se para todo $x \in X$ existe um único $r \in R$ tal que $x \sim r$,

(ii) partição $\mathcal{P}$ de X se R for classe de representantes da relação $\sim_{\mathcal{P}}$, i.e., se para cada $P \in \mathcal{P}$ existe um único $r \in R$ tal que $r \in P$. ¶

Exercício 0.21 $(\star)$**.** Nas condições anteriores, mostre que $X/\!\sim\, = \{C_r : r \in R\}$, onde R é uma classe de representantes de $\sim$ e C_r indica a $\sim$-classe de equivalência de $r \in R$. ■

Agora parece um bom momento para uma pergunta ardilosa: quais partições (ou relações de equivalência) sobre *um conjunto* não vazio admitem classes de representantes? Certamente, se X é tal conjunto e $\mathcal{P}$ é uma de suas partições, então cada $P \in \mathcal{P}$ é um subconjunto não vazio de X, o que permite *escolher* um desses elementos $x_P \in P$ para então considerar o conjunto $\{x_P : P \in \mathcal{P}\}$. Dado que para $P, P' \in \mathcal{P}$ distintos ocorre $P \cap P' = \emptyset$, deve-se ter $x_P \neq x_{P'}$ sempre que $P \neq P'$. Em outras palavras, $\mathcal{R} := \{x_P : P \in \mathcal{P}\}$ é uma classe de representantes para $\mathcal{P}$. Se tal argumentação for honesta, significa que provamos o

Teorema 0.1.26 (©)**.** *Se $\sim$ é uma relação de equivalência sobre um conjunto, então existe uma classe de representantes para $\sim$.*

Exercício 0.22 ((?))**.** A argumentação acima foi *honesta*? Se esta for a primeira vez que você se deparou com tal pergunta, pense nela pelo resto do seu dia. ■

As considerações acima ajudam a elucidar o que se *quis dizer* na discussão que sucedeu a Proposição 0.1.8. Para fixar notações:

Definição 0.1.27 (®)**.** Vamos denotar por $\mathbb{V}$ o *conjunto de todos os conjuntos*[†], (provisoriamente) chamado de **conjunto universo**. ¶

Com a terminologia acima, a Proposição 0.1.8 *demonstra* que ao escrever "$A \approx B$" para indicar que existe bijeção da forma $A \to B$, obtém-se uma relação de equivalência em $\mathbb{V}$.

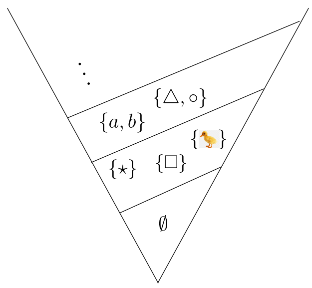


Figura 0.2: Cada $\approx$-classe de equivalência corresponde a uma noção de **cardinalidade**.

Portanto, na prática, uma das noções mais corriqueiras da matemática cotidiana é, na verdade, um dispositivo muito sofisticado: números são representantes de classes de equivalência! Ao dizer que $A := \{x, y\}$ e $B := \{a, b\}$ têm 2 elementos, por exemplo, o símbolo "2" codifica a classe de *todas as coisas com a mesma quantidade de elementos de A* (ou de B, ou de qualquer coisa que tenha... *dois elementos*). Nesse sentido, o que faremos nas próximas seções é encontrar um conjunto "canônico" para representar as cardinalidades dos conjuntos finitos: estes serão os *números naturais*.

[†]Poderíamos defini-lo como $\mathbb{V} := \{x : x = x\}$, já que *todas as coisas são iguais a si mesmas*. Sim, isto é problemático de um ponto de vista formal, mas vamos ignorar essa treta por enquanto.

0.2 Ordens, boas ordens e indução

0.2.0 Essencial

§0 Ordens

O aparato das relações binárias (Subseção 0.1.1 §0) permite abstrair o nosso entendimento de *ordenação*, que será usado posteriormente tanto na descrição da reta real quanto no tratamento das redes. Intuitivamente, os pontos da reta são *ordenados*, no sentido de que há uma *noção* de *antes* e *depois*, como em 3 que antecede 7 e 7 que antecede 9 (note que de nossa experiência diária, 3 também antecede 9). Mais do que isso, a *reta* está ordenada em forma de linha, no sentido de que dados dois pontos nela, algum deles deve anteceder o outro. Tais ideias se formalizam na próxima

Definição 0.2.0 (cf. Definição 0.1.15, itens (i) e (iii))**.** Uma relação binária R num conjunto $\mathbb{X}$ é dita uma **relação de ordem** (**parcial**) se R for reflexiva, transitiva e, além disso, **antissimétrica**, i.e., se para quaisquer $x, y \in \mathbb{X}$, a ocorrência simultânea de $x \; R \; y$ e $y \; R \; x$ acarretar $x = y$, e se escreve $(\mathbb{X}, R)$ quando se busca enfatizar que o conjunto $\mathbb{X}$ é considerado com a ordem R, caso em que $\mathbb{X}$ é dito ser (**parcialmente**) **ordenado** pela ordem (parcial) R. ¶

Dada a óbvia inspiração nas ordenações usuais entre números, costuma-se utilizar símbolos como "$\preceq$", "$\sqsubseteq$" ou mesmo "$\leq$" para denotar ordens parciais — o que sugere uma generalização alternativa, desta vez com base em "$<$".

Definição 0.2.1. Diz-se que $\prec$ é uma **relação de ordem estrita** em $\mathbb{X}$ se $\prec$ for transitiva mas, em vez de reflexiva e antissimétrica, for

(i) **irreflexiva**, i.e., se para todo $x \in \mathbb{X}$ ocorrer $x \not\prec x$, e

(ii) **assimétrica**, i.e., se para quaisquer $x, y \in \mathbb{X}$, a ocorrência de $x \prec y$ acarretar $y \not\prec x$.

Como no caso parcial, escreve-se $(\mathbb{X}, \prec)$ para indicar que $\mathbb{X}$ *está* (**estritamente**) **ordenado** pela ordem (estrita) $\prec$. ¶

Observação 0.2.2. Na prática, podemos chamar ordens parciais e ordens estritas simplesmente de *ordens*. De fato:

✓ se $(\mathbb{S}, \prec)$ é uma ordem estrita, então a relação $\preceq$ definida por

$$x \preceq y \Leftrightarrow (x \neq y \text{ e } x \prec y) \text{ ou } x = y$$

é uma relação de ordem parcial em $\mathbb{S}$;

✓ se $(\mathbb{P}, \sqsubseteq)$ é uma ordem parcial, então a relação $\sqsubset$ definida por

$$x \sqsubset y \Leftrightarrow x \neq y \text{ e } x \sqsubseteq y$$

é uma relação de ordem estrita em $\mathbb{P}$.

Exercício 0.23 ($\star$)**.** Verifique as afirmações anteriores. ■

É claro que ao aplicar o primeiro procedimento à ordem estrita $\sqsubset$, retorna-se à ordem parcial original $\sqsubseteq$, enquanto o segundo procedimento aplicado à ordem parcial $\preceq$ resulta na ordem estrita original $\prec$. Assim, tem-se o direito de chamar tanto $(\mathbb{S}, \prec)$ quanto $(\mathbb{P}, \sqsubseteq)$ de **ordens**. Em tais situações, ficam implicitamente definidas a ordem parcial $\preceq$ e a ordem estrita $\sqsubset$ induzidas por $\prec$ e $\sqsubseteq$, respectivamente[†]. △

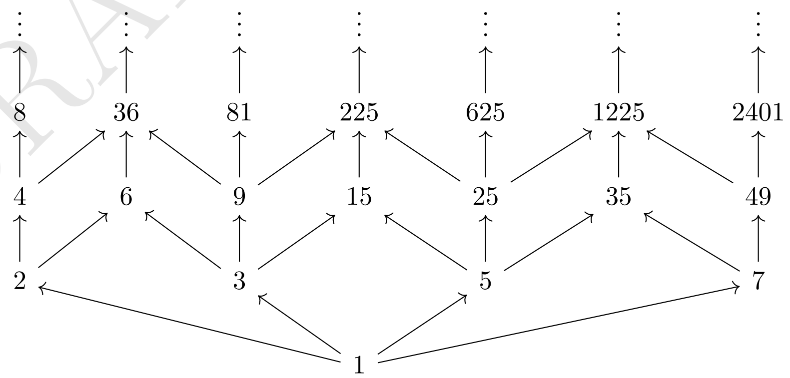
Exemplo 0.2.3. Quem já tem familiaridade com conjuntos numéricos ($\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, etc.) deve ter em mente as formas usuais de ordenação em tais cenários como exemplos de ordens. Apesar disso, é importante saber que mesmo conjuntos *ordinários* admitem ordenações incomuns.

Por exemplo, em $\mathbb{N}^* := \mathbb{N} \setminus \{0\}$ (que será discutido na próxima seção), pode-se declarar $m \preceq n$ sempre que m for um divisor de n. Embora tal relação defina uma ordem parcial sobre $\mathbb{N}^*$ (verifique?)*, ela é bastante diferente de sua ordenação usual: note que $2 \not\preceq 3$ e $3 \not\preceq 2$, já que ambos são primos. Portanto, $(\mathbb{N}^*, \preceq)$ é uma ordem em que podem haver elementos *não comparáveis* entre si, comportamento bem mais comum do que parece. ▲

Exemplo 0.2.4 (Confira o Exemplo 0.1.12)**.** A relação de inclusão $\subseteq$ sobre os membros de $\wp(X)$, para algum conjunto X fixado, faz de $(\wp(X), \subseteq)$ uma ordem, já que: $A \subseteq A$, $A \subseteq B$ e $B \subseteq A$ implicam $A = B$ e $A \subseteq B$ e $B \subseteq C$ implicam $A \subseteq C$, para quaisquer $A, B, C \in \wp(X)$. Como no caso anterior, podem existir $A, B \in \wp(X)$ não comparáveis: para $X := \mathbb{N}$ por exemplo, $A := \{0, 1, 2\}$ e $B := \{0, 1, 3\}$ não são comparáveis, já que $A \not\subseteq B$ (pois $2 \in A \setminus B$) e $B \not\subseteq A$ (pois $3 \in B \setminus A$). ▲

Observação 0.2.5 (Diagramas de Hasse e ordens totais)**.** Um modo bastante prático de *entender* certas ordens (ou *partes* delas) consiste em considerar seus *diagramas de Hasse*. A ideia é muito simples: a ocorrência de $x < y$ em $\mathbb{P}$ é representada por uma seta $(x \to y)$ que liga o vértice anterior (*menor*) x ao posterior (*maior*) y; quando $y < z$ e, por transitividade, $x < z$, não se grafa uma seta entre x e z, pois subentende-se que as duas setas (entre x e y e entre y e z) *compõem* a seta entre x e z.

Assim, por exemplo, o diagrama de Hasse de $(\mathbb{N}^*, \preceq)$ com a ordem do Exemplo 0.2.3 poderia *começar* com



Evidentemente, *diversos* (*infinitos*!) números foram omitidos, como 10 (que deveria estar acima de 5 e 2), 11 (que deveria estar acima de 1 apenas), 12 (que deveria estar acima de 3, 4 e 6), 13... Já para o caso de $X := \{a, b, c\}$ e $(\wp(X), \subseteq)$, o diagrama é mais

[†]Há uma exceção que será adotada neste texto: a inclusão estrita ainda será denotada por "$\subsetneq$" e não por "$\subset$". O motivo é muito simples: o uso do símbolo "$\subset$" para indicar a inclusão parcial está demasiado difundido, o que poderia causar confusão.

simples, embora ainda intrincado.

$$\begin{tikzcd}
& \{a,b,c\} & \\
\{a,b\} \arrow[ur] & \{a,c\} \arrow[u] & \{b,c\} \arrow[ul] \\
\{a\} \arrow[u] \arrow[ur] & \{b\} \arrow[ul] \arrow[ur] & \{c\} \arrow[u] \arrow[ul] \\
& \emptyset \arrow[ul] \arrow[u] \arrow[ur] &
\end{tikzcd}$$

Diagramas desse tipo ajudam a perceber que a ocorrência de elementos não comparáveis se traduz em bifurcações. Uma vez que pretendemos usar ordens para capturar o comportamento das retas, é razoável considerar aquelas em que quaisquer dois elementos sejam comparáveis, condição usualmente chamada de *tricotomia*.

Definição 0.2.6. Uma ordem $(\mathbb{X}, <)$ é **total** se vale a **tricotomia**, i.e., se para quaisquer $x, y \in \mathbb{X}$ ocorrer somente um dos três casos a seguir: $x = y$, $x < y$ ou $y < x$. Se a ordem de $\mathbb{X}$ for parcial, basta dizer que para quaisquer $x, y \in \mathbb{X}$ ocorre $x \leq y$ ou $y \leq x$. ¶

$$\cdots \longrightarrow \bullet \longrightarrow \bullet \longrightarrow \bullet \longrightarrow \bullet \longrightarrow \bullet \longrightarrow \bullet \longrightarrow \bullet \longrightarrow \bullet \longrightarrow \cdots$$

Como o diagrama acima sugere, ordens totais[†] se comportam como linhas precisamente por não terem elementos incomparáveis (bifurcações). Os conjuntos $\mathbb{N}$ e $\mathbb{Z}$ com suas ordenações usuais são exemplos típicos de ordens totais. Os conjuntos $\mathbb{Q}$ e $\mathbb{R}$ também, mas seus diagramas são mais desonestos em virtude da *densidade* de suas ordens: dado que entre quaisquer $x, y \in \mathbb{Q}$ com $x < y$ existe $z \in \mathbb{Q}$ com $x < z < y$, torna-se *impossível* representar *fielmente*, por meio de um diagrama de Hasse, o comportamento linear de tais ordens. Na prática, a alternativa honesta de representação gráfica nesses casos é, justamente... uma linha reta. △

Em geral, costuma-se ler uma expressão do tipo "$x \preceq y$" como "x é **menor do que ou igual** a y", enquanto "$x \prec y$" é lida como "x é (**estritamente**) **menor** do que y" — a menos que o contexto sugira uma terminologia própria para os símbolos. *Alternativamente*, lê-se "$x \preceq y$" como "y é **maior do que ou igual** a x", o que esconde um fato que será importante: escrevendo "$a \succeq b$" para indicar "$b \preceq a$", segue que $\succeq$ também é uma relação de ordem sobre o conjunto em questão: explicitamente, $\succeq$ é apenas a *relação inversa* de $\preceq$ (Definição 0.1.14). Embora pareça banal, tal observação pode ser útil para quem se aventurar na exploração dos princípios de dualidade (cf. Subseção 0.2.1 §1).

Exercício 0.24 ($\star$)**.** Mostre que $(\mathbb{P}, \succeq)$ é ordem parcial sempre que $(\mathbb{P}, \preceq)$ também é. ■

[†]Que com muita razão também são chamadas de ordens *lineares*.

§1 Boas ordens e indução

Definição 0.2.7. Fixada uma *ordem* $(\mathbb{P}, \leq)$, um subconjunto A de $\mathbb{P}$ e um elemento $a \in A$, diremos que a é *um* **menor elemento** (ou **mínimo**) de A se $a \leq x$ ocorrer para todo $x \in A$. No caso de uma ordem estrita $<$, pede-se $a < x$ para todo $x \in A$ tal que $x \neq a$. ¶

Acima, o uso do artigo indefinido "um" foi puro preciosismo, já que um mínimo, quando existe, é único: se $a, a' \in A$ são mínimos de A, então ocorre $a \leq a'$ e $a' \leq a$, donde a antissimetria de $\leq$ acarreta $a = a'$. Isto justifica introduzir uma notação para indicar os mínimos: caso exista, o menor elemento de A será denotado $\min\limits_{a \in A} a$, $\min_{\leq} A$ ou apenas $\min A$.

Definição 0.2.8. Uma ordem $\leq$ (ou $<$) sobre um conjunto $\mathbb{B}$ é chamada de **boa ordem** se todo subconjunto não vazio de $\mathbb{B}$ admite menor elemento. Diz-se também que $\mathbb{B}$ está **bem ordenado** pela (boa) ordem $\leq$, ou ainda que $(\mathbb{B}, \leq)$ é uma boa ordem. ¶

Exemplo 0.2.9. Alguns conjuntos numéricos clássicos que você já viu na escola (ou em Cálculo I) não são bem ordenados: é o caso de $\mathbb{Z}$, $\mathbb{Q}$ e $\mathbb{R}$ (por quê?)*. A situação de $\mathbb{C}$ é um pouco pior, mas ainda é cedo para discutir isso. Por outro lado, $\mathbb{N}$ é o exemplo típico de boa ordem, o que não é mero acidente, como veremos em breve. ▲

Moralmente, um conjunto está bem ordenado quando seus elementos podem ser *enfileirados* por meio da ordem $\leq$: há o *primeiro* elemento, digamos $b_0 := \min \mathbb{B}$, em seguida o seu *sucessor*, digamos $b_1 := \min(\mathbb{B} \setminus \{b_0\})$, em seguida... Como os índices "0" e "1" sugerem, parece haver alguma ligação com os *números naturais* que já conhecemos de longa data, o que suscita uma pergunta deliberadamente evitada até agora (exceto por quem não ignorou o final da Subseção 0.1.1 §1): *o que são números* (*e o que poderiam ser*)*?*[†]

$$b_0 \longrightarrow b_1 \longrightarrow b_2 \longrightarrow b_3 \longrightarrow b_4 \longrightarrow \cdots$$

Evidentemente, esse tipo de pergunta não se refere ao símbolo utilizado para *denotar* um número, mas sim ao próprio número: por exemplo, qual o significado de "três" nas sentenças "A Argentina venceu três Copas do Mundo" e "Neymar rolou por três metros ao simular uma falta"? Quais as diferenças de significado, e quais as semelhanças?

Exercício 0.25 ((?))**.** Reflita (por pelo menos *três* minutos) sobre as questões acima. ■

Apesar da sugestão numérica anterior, é possível evitar o emprego explícito de números no entendimento das boas ordens — o que inclusive será útil quando voltarmos a discutir a *natureza* dos números. A grande sacada para fazer isso está escondida na seguinte

Proposição 0.2.10. *Sejam* $(\mathbb{B}, \leq)$ *uma boa ordem e* $b \in \mathbb{B}$. *Se existir* $c \in \mathbb{B}$ *com* $c > b$, *então existe* $b' \in \mathbb{B}$ *com as seguintes propriedades:*

(i) $b < b'$;

(ii) se $d \in \mathbb{B}$ *e* $d > b$, *então* $b' \leq d$.

[†] Referência ao clássico "*Was Sind Und Was Sollen Die Zahlen?*", de Richard Dedekind, em que *Ele* apresenta Sua construção para (o que hoje chamamos de) um corpo ordenado completo [12].

Demonstração. É mais simples do que parece: a existência de c com $c > b$ garante que $\mathbb{B}_{>b} := \{d \in \mathbb{B} : d > b\}$ é um subconjunto não vazio de $\mathbb{B}$, justamente por ter c como elemento. Logo, a boa ordenação garante a existência de $\min \mathbb{B}_{>b}$, de modo que basta tomar $b' := \min \mathbb{B}_{>b}$. □

Definição 0.2.11. Nas condições da proposição anterior, vamos denotar b' por $\operatorname{suc}_{\mathbb{B}}(b)$, o **sucessor** de b na boa ordem $\mathbb{B}$. ¶

Assim como ocorre com sucessores nos diversos campos da vida real, o sucessor de b numa boa ordem $\mathbb{B}$, caso exista, é o primeiro elemento da ordem a ser maior do que b, o que em particular impede a existência de elementos *intermediários*. O próximo exercício deve esclarecer a ideia.

Exercício 0.26 ($\star$)**.** Seja $(\mathbb{B}, \leq)$ uma boa ordem. Mostre que se $b \in \mathbb{B}$ e existe $\operatorname{suc}_{\mathbb{B}}(b)$, então não existe $c \in \mathbb{B}$ tal que $b < c < \operatorname{suc}_{\mathbb{B}}(b)$. Dica: releia a definição de $\operatorname{suc}_{\mathbb{B}}(b)$. ■

Exercício 0.27 ($^{\star}_{\star\star}$)**.** Seja $(\mathbb{B}, \leq)$ uma boa ordem. Mostre que se $x, y \in \mathbb{B}$ têm sucessores e $x \leq y$, então $\operatorname{suc}_{\mathbb{B}}(x) \leq \operatorname{suc}_{\mathbb{B}}(y)$. Dica: mostre antes que se $\emptyset \neq Y \subseteq X \subseteq \mathbb{B}$, então $\min X \leq \min Y$. ■

Observação 0.2.12 (Sucessores nem sempre existem)**.** Para quem tem familiaridade com os números racionais, por exemplo, note que não faz sentido perguntar qual o sucessor de 0 em $\mathbb{Q}$, já que não existe o *primeiro racional* maior do que 0: sempre que $q > 0$, existe outro q' com $0 < q' < q$. Em particular, a relação de ordem usual sobre $\mathbb{Q}$ não é uma boa ordem. △

Exemplo 0.2.13. Por mais sem graça que pareça, $\mathbb{B} := \emptyset$ pode ser considerado como um conjunto bem ordenado: no caso, sua boa ordem $\leq$ é o único subconjunto de $\emptyset \times \emptyset = \emptyset$, a saber, $\emptyset$! Apesar de sua simplicidade, $\emptyset$ é o gatilho de uma importante reação em cadeia, como veremos adiante. ▲

Definição 0.2.14. Fixada uma boa ordem $(\mathbb{B}, \leq)$, diremos que $u \in \mathbb{B}$ é o **último elemento** de $\mathbb{B}$ se u é o *maior elemento de* $\mathbb{B}$,[†] denotado por $\max \mathbb{B}$. ¶

Proposição 0.2.15. *Se* $(\mathbb{B}, \leq)$ *é uma boa ordem e* $u' \notin \mathbb{B}$*, então existe uma boa ordem* $(\mathbb{B}', \leq')$ *tal que*

(*i*) $\mathbb{B} \subsetneq \mathbb{B}'$,

(*ii*) $x \leq y \Leftrightarrow x \leq' y$ *para quaisquer* $x, y \in \mathbb{B}$, *e*

(*iii*) u' *é o último elemento de* $\mathbb{B}'$.

Demonstração. Basta definir $\mathbb{B}' := \mathbb{B} \cup \{u'\}$ e, para quaisquer $x, y \in \mathbb{B}'$, escrever $x \leq' y$ para indicar a ocorrência de "$x, y \in \mathbb{B}$ e $x \leq y$" ou "$y = u'$". Na prática, $\leq'$ apenas *estende* a definição de $\leq$ sobre $\mathbb{B}' \times \mathbb{B}'$ ao declarar $x \leq' u'$ para todo $x \in \mathbb{B}'$, o que torna quase imediata a verificação das propriedades desejadas. □

Exercício 0.28 ($^{\star}_{\star}$)**.** Complete a demonstração. ■

[†]Adivinhe a definição ou confira a Subseção 0.2.1 §1.

Exemplo 0.2.16. Fixado qualquer objeto u', tem-se por definição que $u' \notin \emptyset$, o que permite empregar a última proposição a fim de estender a boa ordem $\mathbb{B}$ do Exemplo 0.2.13: faz-se $\mathbb{B}' := \mathbb{B} \cup \{u'\} = \{u'\}$, que tem u' como seu último (e único!) elemento. Ora, por que parar? Certamente existe $u'' \neq u'$, donde a proposição anterior garante a boa ordem $\mathbb{B}'' := \mathbb{B}' \cup \{u''\}$, com $u' < u''$ e u'' como último elemento. Em particular, u'' é sucessor de u', mas u'' não tem sucessores em $\mathbb{B}''$. Ora, por que parar? Certamente existe $u''' \notin \{u', u''\}$, donde a proposição anterior garante... ▲

Como os exemplos acima sugerem, a noção de sucessor numa boa ordem torna supérfluo o uso explícito de números *alienígenas* para descrever o seu enfileiramento. Na verdade, mais do que isso, o comportamento dos sucessores é tão parecido com o da *progressão* esperada dos números naturais que chega a ser tentador utilizar a noção de boa ordenação para *descrever* o que os números *poderiam ser*. Para agravar ainda mais o sentimento:

Teorema 0.2.17 (Indução numa boa ordem)**.** *Seja $(\mathbb{B}, \leq)$ uma boa ordem. Suponha que X seja um subconjunto de $\mathbb{B}$ com a seguinte propriedade: para qualquer $c \in \mathbb{B}$,*

sempre que ocorre $b \in X$ para todo $b < c$, também ocorre $c \in X$.

Em tais condições, $\mathbb{B} = X$.

Demonstração. Nada precisa ser feito se ocorrer $\mathbb{B} = \emptyset$. Agora, se $\mathbb{B} \neq \emptyset$ e existir $\widetilde{b} \in \mathbb{B}$ com $\widetilde{b} \notin X$, então o conjunto $T := \{b \in \mathbb{B} : b \notin X\}$ é não vazio e, pela boa ordenação, deve existir $t := \min T$. Em particular, $t \in T$. Ora, isto impede que X tenha a propriedade do enunciado: com efeito, por t ser o menor elemento em T, todo $b < t$ deve ser membro de X, de modo que se X tivesse a propriedade, concluiríamos que $t \in X$, mas $t \in T = \mathbb{B} \setminus X$. □

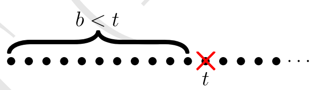


Figura 0.3: Se t fosse o *primeiro* tal que $t \notin X$, então todo $b < t$ pertenceria a X. Logo, $t \in X$!

No dia a dia, o subconjunto X costuma ser formado pelos elementos de $\mathbb{B}$ que têm alguma propriedade $\mathcal{P}$ fixada, i.e., $X := \{x \in \mathbb{B} : \text{vale } \mathcal{P}(x)\}$, de modo que a exigência feita sobre X se traduz na seguinte condição:

sempre que vale $\mathcal{P}(b)$ ***para todo*** $b < c$, *também vale* $\mathcal{P}(c)$.

Muito provavelmente você já se deparou com tal formulação do *Princípio da Indução* em alguma disciplina introdutória de Teoria dos Números ou Álgebra[†]— neste caso, é comum dizer que esta é a "segunda forma da indução" [24, 25]. Como sempre, o intuito é provar que todo elemento c de $\mathbb{B}$ tem a propriedade $\mathcal{P}(c)$ e, nesse sentido, o Teorema 0.2.17 assegura que não precisamos provar $\mathcal{P}(c)$ *diretamente*, mas podemos fazer isso com o auxílio de uma *hipótese indutiva* (a parte em negrito na frase destacada acima): se conseguirmos garantir que $\mathcal{P}(c)$ vale sempre que se assumir a validade de $\mathcal{P}(b)$ para todo $b < c$, então todo c da boa ordem tem a propriedade em questão[‡].

[†] Se não for o seu caso (... ou se isso for um assunto que escapou da sua memória...), confira a Subseção 0.10.1 §3.

[‡] "*Ah, mas não estaríamos supondo o que queremos provar?!*" Calma: ninguém está supondo que vale $\mathcal{P}(c)$, mas sim que $\mathcal{P}(b)$ vale para todo $b < c$!

Note que o raciocínio da demonstração pode ser usado para provar que *cada* elemento de uma boa ordem não vazia $\mathbb{B}$ pertence a X se a hipótese for satisfeita: como não existe $b < \min \mathbb{B}$, a hipótese indutiva é verdadeira (todo $b < \min \mathbb{B}$ pertence a X, por vacuidade!), logo $\min \mathbb{B} \in X$; se $\min \mathbb{B}$ tem sucessor, digamos c, então $c \in X$, pois no passo anterior verificou-se que todo $b < c$ pertence a X; se c tem sucessor...

Esse arquétipo de efeito dominó é equivalente ao que foi apresentado no teorema anterior, mas somente nas boas ordens em que todos os elementos, exceto o menor, são sucessores de *alguém*[†].

Naturalmente, isto será discutido depois. Para encerrar, faça o seguinte

Exercício 0.29 $(\star)$**.** Mostre que se $(\mathbb{B}, \leq)$ é boa ordem, então $(\mathbb{B}, \leq)$ é ordem total. Dica: dados $x, y \in \mathbb{B}$, encare $\{x, y\}$ até que ele te encare de volta. ■

0.2.1 Extras

§0 Elementos minimais e maximais

Definição 0.2.18. Fixada uma *ordem* $(\mathbb{P}, \leq)$, um subconjunto A de $\mathbb{P}$ e um elemento $a \in A$, diremos que $a \in A$ é **elemento minimal de** A se não existe $x \in A$ com $x < a$. De modo *dual*, diremos que $a \in A$ é **elemento maximal** de A se não existe $x \in A$ com $a < x$. ¶

Elementos minimais e maximais não costumam ser explorados em contextos introdutórios de Análise pois as ordens consideradas *geralmente* são totais — e, em tais casos, as definições coincidem com as noções de mínimos e *máximos*.

Proposição 0.2.19. *Sejam* $(\mathbb{T}, \leq)$ *uma ordem,* $A \subseteq \mathbb{T}$ *um subconjunto e* $a \in A$ *um elemento qualquer.*

(i) *Se* $a = \min A$*, então* a *é minimal.*

(ii) *Se* $(\mathbb{T}, \leq)$ *é uma ordem total, então vale a recíproca do item anterior.*

Demonstração. Para a primeira parte, não pode existir $x \neq a$ com $x < a$ e $x \in A$, pois por a ser mínimo deve-se ter $a \leq x$ (lembre-se: ordens estritas são assimétricas!). Para a segunda parte: a princípio, por a ser minimal, para nenhum $x \in A$ *ocorre* $x < a$ e, como x e a são comparáveis pela hipótese de tricotomia, *resta* apenas $a \leq x$. Logo, $a = \min A$. □

Exercício 0.30 $({}^{\star}_{\star\star})$**.** Enuncie e demonstre a versão da proposição anterior para elementos maximais. **Observação:** neste caso, você também precisará da definição de *maior elemento* (ou *máximo*)[‡], apresentada um pouco mais abaixo (mas não é difícil adivinhar qual é). ■

Com isso dito, retorne para as ordens não totais da Observação 0.2.5: no caso de $\wp(X)$, por exemplo, $A := \{\{a\}, \{b\}, \{c\}\}$ é tal que todos os seus elementos são minimais em $\wp(X)$ (e nenhum deles é mínimo!), comportamento similar ao dos números primos em $(\mathbb{N}^*, \preceq)$. Alguma delas apresenta subconjuntos com elementos maximais que não são máximos? $({}^{\star}_{\star})$

[†]Cuidado: ter sucessor $\neq$ ser sucessor. O exemplo óbvio é o menor elemento de uma boa ordem com pelo menos *dois* elementos.

[‡]Note que máximos, assim como mínimos, quando existem, são únicos $(\star)$.

§1 Dualidade

Considere $(\mathbb{P}, \leq)$ uma ordem (parcial). O fato de $(\mathbb{P}, \geq)$ ainda ser uma ordem parcial traz uma consequência curiosa: sempre que você tiver um resultado acerca de um conceito definível em ordens parciais $(\leq)$, você ganha automaticamente outro resultado acerca da versão *dual* do conceito, isto é, onde o conceito é reescrito com a ordem inversa $(\geq)$. Por exemplo: caso não tenha *adivinhado* a definição de máximo, ela se faz assim.

Definição 0.2.20. Fixada uma *ordem* $(\mathbb{P}, \leq)$, um subconjunto A de $\mathbb{P}$ e um elemento $a \in A$, diremos que a é *um* **maior elemento** (ou **máximo**) de A se $x \leq a$ ocorrer para todo $x \in A$. Para uma ordem estrita $<$, pede-se $x < a$ para todo $x \in A$ tal que $x \neq a$. ¶

Agora, esqueça a definição acima e considere $(\mathbb{P}, \geq)$, onde $\geq$ é a inversa de $\leq$. O que significa dizer que $a \in A$ é menor elemento de A com respeito à ordem $\geq$? Checando a Definição 0.2.7, deve-se ter o seguinte: $a \geq x$ ocorre para todo $x \in A$. Como "$a \geq x$" significa "$x \leq a$", resulta que a é máximo de A com respeito à ordem $\leq$. Neste caso, as definições de máximo e mínimo são ditas duais.

Exercício 0.31 ($^{\star}_{\star\star}$)**.** Mostre que as definições de elementos minimais e maximais são duais. ■

Uma das vantagens desse tipo de abordagem é reciclar demonstrações. Por exemplo: como já sabemos que mínimos em ordens parciais são únicos quando existem, resulta que máximos em ordens parciais também são únicos quando existem, simplesmente por eles poderem ser expressos como mínimos nas ordens inversas. Se você duvida, faça o próximo exercício, e perceba que, na prática, você precisa apenas trocar todas as ocorrências de "$<$" e "$\leq$" na demonstração da Proposição 0.2.19 por "$>$" e "$\geq$", respectivamente. Abaixo, $\max A$ indica o máximo de A, caso exista.

Exercício 0.32 ($^{\star}_{\star\star}$)**.** Sejam $(\mathbb{T}, \leq)$ uma ordem, $A \subseteq \mathbb{T}$ um subconjunto e $a \in A$ um elemento qualquer.

(i) Mostre que se $a = \max A$, então a é maximal.

(ii) Mostre que se $(\mathbb{T}, \leq)$ é uma ordem total, então vale a recíproca do item anterior. ■

0.3 Os axiomas de Dedekind-Peano

0.3.0 Essencial

§0 Boas ordens naturais

Na última seção, discutiu-se uma forma de indução válida em boas ordens quaisquer. Agora, vamos nos restringir às boas ordens que remetem diretamente ao entendimento que temos dos *números naturais*, a começar com o

Corolário 0.3.0 (Indução "clássica")**.** *Sejam* $(\mathbb{B}, \leq)$ *uma boa ordem com* $\mathbb{B} \neq \emptyset$*, considere* $p := \min \mathbb{B}$ *e suponha que para todo* $b' \in \mathbb{B} \setminus \{p\}$ *exista* $b \in \mathbb{B}$ *tal que* $b' = \mathrm{suc}_{\mathbb{B}}(b)$*. Em tais condições, se* $X \subseteq \mathbb{B}$ *for tal que*

✓ $p \in X$, *e*

✓ $\mathrm{suc}_{\mathbb{B}}(b) \in X$ *sempre que* $b \in X$,

então $\mathbb{B} = X$.

Demonstração. Basta verificar que X tem a propriedade do Teorema 0.2.17, i.e., que para qualquer $c \in \mathbb{B}$, tenha-se a ocorrência de $c \in X$ sempre que $b \in X$ para todo $b < c$: ora, se $c := p$, então $p \in X$ por hipótese; se $c > p$ e $b \in X$ para todo $b < c$, então em particular para $b \in \mathbb{B}$ com $c = \mathrm{suc}_{\mathbb{B}}(b)$ (que existe pela hipótese sobre $\mathbb{B}$), deve-se ter $b < c$, donde segue que $b \in X$ e, pela hipótese sobre X, $c = \mathrm{suc}_{\mathbb{B}}(b) \in X$, como desejado. □

Exercício 0.33 $({}^{\star}_{\star\star})$**.** Prove o corolário anterior sem apelar para o Teorema 0.2.17. Dica: suponha $\mathbb{B} \setminus X \neq \emptyset$ e note que seu menor elemento *deveria* ser da forma $\mathrm{suc}_{\mathbb{B}}(b)$ para algum $b \in X$. ■

O corolário acima mostra que se $\mathbb{B}$ é uma boa ordem em que "todo elemento maior do que $\min \mathbb{B}$ é sucessor de *alguém*", então vale o *princípio da indução* em sua forma convencional. Tal propriedade entre aspas, de fato, remete aos naturais, mas não completamente: com $\mathbb{B} := \{a, a', a'', a'''\}$, por exemplo, obtemos uma boa ordem ao declarar $a < a' < a'' < a'''$ onde $a = \min \mathbb{B}$, $a' = \mathrm{suc}_{\mathbb{B}}(a)$, $a'' = \mathrm{suc}_{\mathbb{B}}(a')$ e $a''' = \mathrm{suc}_{\mathbb{B}}(a'')$, ou seja, todo elemento diferente de $\min \mathbb{B}$ é sucessor de alguém. Porém, $\mathbb{B}$ não é o que esperaríamos de $\mathbb{N}$, pois ainda faltam sucessores: deveria haver $a'''' := \mathrm{suc}_{\mathbb{B}}(a''')$, $a''''' := \mathrm{suc}_{\mathbb{B}}(a'''')$ e *assim por diante*. Pois bem:

Definição 0.3.1. Diremos que uma boa ordem $(\mathbb{B}, \leq)$ é **natural** se as seguintes condições forem satisfeitas:

(i) $\mathrm{suc}_{\mathbb{B}}(b)$ existe para todo $b \in \mathbb{B}$,

(ii) para todo $b' \in \mathbb{B} \setminus \{\min \mathbb{B}\}$ existe $b \in \mathbb{B}$ com $b' = \mathrm{suc}_{\mathbb{B}}(b)$. ¶

A expressão "natural" acima faz, finalmente, alusão ao *conjunto dos números naturais*, frequentemente denotado por $\mathbb{N}$, que você certamente conhece de sua vida fora da disciplina. Veremos que, para efeitos práticos, *qualquer* boa ordem natural pode fazer o *papel* de $\mathbb{N}$.

Definição 0.3.2. Diremos que $(\mathcal{N}, i, s)$ é um **sistema natural**[†] se $\mathcal{N}$ for um conjunto, i for um elemento de $\mathcal{N}$ e $s \colon \mathcal{N} \to \mathcal{N}$ for uma função satisfazendo os seguintes axiomas:

(DP_i) $\mathcal{N} \setminus \mathrm{im}(s) = \{i\}$;

(DP_{ii}) s é injetora; e

(DP_{iii}) (*indutivo*) se um subconjunto $X \subseteq \mathcal{N}$ for tal que

(C.I.) $i \in X$, e

(H.I.) $s(n) \in X$ sempre que $n \in X$,

então $X = \mathcal{N}$. ¶

Os axiomas (DP_i), (DP_{ii}) e (DP_{iii}), elaborados independentemente (?) por Dedekind e Peano, buscam capturar o *mínimo* que se espera dos *números naturais* dentro de um cenário regido por conjuntos[‡]. Nesse sentido, boas ordens naturais cumprem bem o papel.

[†]Nos cânones brasileiros de Análise Real redigidos por Lima [24, 25], diz-se apenas que "$\mathbb{N}$ é um conjunto dotado de uma função s tal que...", onde as reticências repetem as condições (i), (ii) e (iii) apresentadas para sistemas naturais. Aqui, o emprego do artigo indefinido (um) em vez do definido (o) visa chamar sua atenção para o seguinte: a princípio, nada impediria que existissem diversos sistemas naturais *diferentes*.

[‡]Em certo sentido, eles funcionam como os axiomas que descrevem os *grupos*, que por sua vez buscam capturar as propriedades básicas das noções de simetria. A ideia é que se tal mínimo for satisfeito, então todos os resultados de Aritmética Básica podem ser recuperados via dedução, paciência e um pouco de conjuntos.

Teorema 0.3.3. *Se* $(\mathbb{B}, \leq)$ *é uma boa ordem natural, então* $(\mathbb{B}, \min \mathbb{B}, \operatorname{suc}_{\mathbb{B}})$ *é um sistema natural, ou seja:*

(i) $\mathbb{B} \setminus \operatorname{im}(\operatorname{suc}_{\mathbb{B}}) = \{\min \mathbb{B}\}$;

(ii) *a função* $\operatorname{suc}_{\mathbb{B}} \colon \mathbb{B} \to \mathbb{B}$ *é injetora; e*

(iii) *se* $X \subseteq \mathbb{B}$ *for tal que* $\min \mathbb{B} \in X$ *e* $\operatorname{suc}_{\mathbb{B}}(b) \in X$ *sempre que* $b \in X$*, então* $X = \mathbb{B}$.

Demonstração. A verificação do axioma (DP$_i$) fica por sua conta ($\star$). O axioma indutivo (DP$_{iii}$), por sua vez, já foi demonstrado (confira o Corolário 0.3.0 em caso de dúvida). Resta apenas verificar a validade de (DP$_{ii}$), i.e., que a função $\operatorname{suc}_{\mathbb{B}} \colon \mathbb{B} \to \mathbb{B}$ é injetora.

Explicitamente, deve-se mostrar que se $x, y \in \mathbb{B}$ são distintos, então $\operatorname{suc}_{\mathbb{B}}(x) \neq \operatorname{suc}_{\mathbb{B}}(y)$. Como toda boa ordem é total (Exercício 0.29), a ocorrência de $x \neq y$ acarreta $x < y$ ou $y < x$, de modo que basta mostrar o seguinte: se $x < y$, então $\operatorname{suc}_{\mathbb{B}}(x) < \operatorname{suc}_{\mathbb{B}}(y)$. Ora, pelo Exercício 0.27, já sabemos que $\operatorname{suc}_{\mathbb{B}}(x) \leq \operatorname{suc}_{\mathbb{B}}(y)$. Se tal desigualdade não fosse estrita, teríamos $\operatorname{suc}_{\mathbb{B}}(x) = \operatorname{suc}_{\mathbb{B}}(y)$, resultando em

$$x < y < \operatorname{suc}_{\mathbb{B}}(y) = \operatorname{suc}_{\mathbb{B}}(x),$$

o que viola o Exercício 0.26. $\square$

§1 Axiomatizando o infinito

Há quem não goste da abordagem via boas ordens pois ela demanda *mais estrutura* para capturar parte do comportamento dos sistemas naturais. Todavia, um sistema natural $(\mathcal{N}, i, s)$ vem de fábrica com uma boa ordem natural[†] que tem i como menor elemento e cuja função sucessor é, precisamente, a função s. Caso tenha se interessado, confira o Exercício 0.77. Assim, toda essa discussão culmina na seguinte conclusão:

existe uma boa ordem natural se, e somente se, existe um sistema natural.

E daí a grande questão: existe?

Intuitivamente, a resposta é sim. Porém, *demonstrar* que a resposta é sim envolveria construir uma boa ordem natural *infinita*, o que poderia ser problemático num contexto em que as demonstrações precisam ser *finitas*. O que se faz então é *postular* a existência de ao menos uma boa ordem natural (ou sistema natural, se preferir):

Axioma de Dedekind-Peano. *Existe um sistema* (*ou boa ordem*) *natural.*

Deste ponto em diante, é prática comum fixar *algum* sistema natural e, por meio de *argumentações indutivas*, construir *recursivamente* as operações de *adição* e *multiplicação* de forma precisa e verificar todas as propriedades operatórias esperadas, num árduo, doloroso e ~~gratificante?~~ demorado processo que, na prática, *recria* a Aritmética Básica. É nesse sentido que se costuma dizer que os Axiomas de Dedekind-Peano capturam o básico dos *naturais*: a partir deles e das construções conjuntistas (cf. Observação 0.3.4), recuperam-se todos os dispositivos aritméticos usuais e, *a posteriori*, todos os outros conjuntos numéricos!

[†]Que inclusive precisa ser explicitada em algum momento mesmo por quem opta pelos sistemas naturais, vide [24, 25] por exemplo.

Observação 0.3.4. É relativamente comum encontrar quem propague máximas como "os Axiomas de (Dedekind-) Peano permitem construir a Matemática!", num indicativo claro de amnésia, já que os axiomas usados na descrição de sistemas naturais descrevem apenas tais sistemas. Por exemplo, se $(\mathcal{N}, i, s)$ é um sistema natural, não são os seus axiomas que *permitem* construir o conjunto $\mathcal{N} \times (\mathcal{N} \setminus \{i\})$ a partir do qual se obtém os *inteiros*, mas sim as suposições tácitas (axiomas acerca de conjuntos!) de que tais procedimentos podem ser realizados. Para mais detalhes, confira [10, 30]. $\triangle$

Há, porém, um elefante na sala: *o vermelho que eu vejo é tão vermelho quanto o que você vê?* Explicitamente, o Axioma de Dedekind-Peano não assegura *o* sistema natural, mas apenas *um* sistema natural. Poderia haver vários? Se sim, então os processos descritos acima dependem do sistema natural escolhido? Cada sistema natural tem sua própria Aritmética? Será que essas preocupações realmente fazem sentido?

Comecemos pela última: a rigor, os questionamentos são pertinentes. Por exemplo: assim como um *grupo* é um conjunto dotado de funções que satisfazem certos axiomas, sistemas naturais também são conjuntos dotados de funções que satisfazem certos axiomas. Dado que existem grupos definitivamente *incompatíveis* entre si, há precedente para questionar a possibilidade de sistemas naturais *incompatíveis* em algum sentido. Feita a ressalva, vem a boa notícia: embora existam (*infinitos*!) sistemas naturais distintos, todos eles são *rigorosamente* compatíveis entre si, o que na prática garante que a Aritmética desenvolvida em um seja *indistinguível* da Aritmética desenvolvida em outro.

Teorema 0.3.5 (Dedekind)**.** *Se $(\mathcal{N}, i, s)$ e $(\mathcal{M}, j, t)$ são sistemas naturais, então existe uma única bijeção $\varphi \colon \mathcal{N} \to \mathcal{M}$ tal que $\varphi(i) = j$ e $\varphi(s(n)) = t(\varphi(n))$ para todo $n \in \mathcal{N}$.*

O teorema acima[†] garante, entre outras coisas (cf. Exercício 0.72), a *irrelevância* de se apegar a um começo particular na descrição de um sistema natural: o importante é que exista *algum* começo. Também fica justificada a tranquilidade com que profissionais em Matemática *escolhem* um sistema natural arbitrário para desenvolver Aritmética com a certeza de que os resultados obtidos valerão em qualquer outro sistema[‡].

$$
\begin{array}{ll}
(A) & 1 \xrightarrow{+1} 2 \xrightarrow{+1} 3 \xrightarrow{+1} 4 \xrightarrow{+1} \cdots \\
(B) & \bullet \xrightarrow{+\bullet} \bullet\bullet \xrightarrow{+\bullet} \bullet\bullet\bullet \xrightarrow{+\bullet} \bullet\bullet\bullet\bullet \xrightarrow{+\bullet} \cdots \\
(C) & | \xrightarrow{+|} || \xrightarrow{+|} ||| \xrightarrow{+|} |||| \xrightarrow{+|} \cdots \\
(D) & i \xrightarrow{s} s(i) \xrightarrow{s} s(s(i)) \xrightarrow{s} s(s(s(i))) \xrightarrow{s} \cdots \\
(E) & \emptyset \xrightarrow{\{\}} \{\emptyset\} \xrightarrow{\{\}} \{\{\emptyset\}\} \xrightarrow{\{\}} \{\{\{\emptyset\}\}\} \xrightarrow{\{\}} \cdots \\
(F) & \emptyset \xrightarrow{x \mapsto x \cup \{x\}} \{\emptyset\} \xrightarrow{x \mapsto x \cup \{x\}} \{\emptyset, \{\emptyset\}\} \xrightarrow{x \mapsto x \cup \{x\}} \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\} \xrightarrow{x \mapsto x \cup \{x\}} \cdots \\
 & \ \| \qquad\quad\ \ \| \qquad\qquad\quad \| \qquad\qquad\qquad\qquad \| \\
(G) & 0 \xrightarrow{+1} 1 \xrightarrow{+1} 2 \xrightarrow{+1} 3 \xrightarrow{+1} \cdots
\end{array}
$$

Figura 0.4: Vários sistemas naturais.

[†] Cuja demonstração será apresentada na Subseção 0.3.1 §1.

[‡] Por exemplo, fixados sistemas naturais $(\mathcal{N}, i, s)$ e $(\mathcal{M}, j, t)$, suponha que seja verdadeira a asserção "não existe $n \in \mathcal{N}$ tal que $s(n) = n$". Neste caso, também deverá valer que "não existe $m \in \mathcal{M}$ tal que $t(m) = m$": de fato, se existisse m com $t(m) = m$, então ao tomar o único $n \in \mathcal{N}$ com $\varphi(n) = m$, teria-se $t(m) = t(\varphi(n)) = \varphi(s(n))$ e $t(m) = m = \varphi(n)$, donde a injetividade de φ garantiria $s(n) = n$.

É lícito, portanto, fazer a seguinte

Definição 0.3.6. Vamos denotar por $\mathbb{N}$ *um* sistema natural dado pelo Axioma de Dedekind-Peano, cujos elementos serão chamados de *números naturais*. Além disso:

- $0 := \min \mathbb{N}$;
- $1 := \mathrm{suc}_{\mathbb{N}}(0)$;
- $2 := \mathrm{suc}_{\mathbb{N}}(1)$;
- $3 := \mathrm{suc}_{\mathbb{N}}(2)$;
- $4 := \mathrm{suc}_{\mathbb{N}}(3)$;
- $5 := \mathrm{suc}_{\mathbb{N}}(4)$;
- $6 := \mathrm{suc}_{\mathbb{N}}(5)$;
- $7 := \mathrm{suc}_{\mathbb{N}}(6)$;
- $8 := \mathrm{suc}_{\mathbb{N}}(7)$;
- $9 := \mathrm{suc}_{\mathbb{N}}(8)$;

e assim por diante[†]. ¶

O método usado para justificar o Teorema 0.3.5 (chamado de recursão) permite definir de maneira precisa as operações usuais de adição e multiplicação em $\mathbb{N}$, bem como verificar suas propriedades principais por indução. Embora seja edificante investir algum período da vida na checagem rigorosa dessas coisas, trata-se de algo mais pertinente a um curso de Aritmética ou de (Introdução a) Teoria dos Números do que a um curso de Análise *Real*. Por isso, todo o arcabouço básico de Aritmética será assumido como conhecido — e, apenas por preciosismo, a verificação de alguma*s* propriedade*s* será sugerida na Subseção 0.3.1 §0. Em particular: $\mathrm{suc}_{\mathbb{N}}(n) := n + 1$ de agora em diante.

0.3.1 Extras

§0 Recursão

De acordo com a suposição feita no final da subseção anterior, nós já *sabemos* somar e multiplicar números naturais, o que formalmente consiste em assumir conhecidas funções $(+): \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ e $(\cdot): \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ com as propriedades *operatórias* que aprendemos na escola. Com isso em mente, suponha que agora quiséssemos definir a famosa função **fatorial** $F: \mathbb{N} \to \mathbb{N}$, que associa cada $n \in \mathbb{N}$ ao número $n!$, de acordo com os seguintes critérios: $0! := 1$ e $(n+1)! := (n+1) \cdot n!$. Da forma como está posta, é como se a função F fosse usada em sua própria definição, já que $F(n+1) = (n+1) \cdot F(n)$, algo circular e que poderia trazer dor de cabeça.

Porém, secretamente, a função F acima é a *colagem* de uma *família de funções* cuja *existência* se demonstra por indução: prova-se que para cada $n \in \mathbb{N}$ existe uma função *adequada*

$$F_n: \{m : m \in \mathbb{N} \text{ e } m < n\} \to \mathbb{N}$$

de tal forma que F_{n+1} *estende* F_n para cada $n \in \mathbb{N}$. Daí, para *colar* todas as F_n's numa única F, faz-se $F(m) := F_n(m)$ para qualquer $n \in \mathbb{N}$ com $m < n$, o que torna F uma função pois, nas inevitáveis ocorrências de $n, n' > m$ com $n \neq n'$, ou F_n estende $F_{n'}$ (caso $n > n'$) ou $F_{n'}$ estende F_n (caso $n' > n$) e, portanto, $F_n(m) = F_{n'}(m)$. Implicitamente, F é a *reunião* de *todas* as F_n's.

[†]Em posse das operações de adição, multiplicação e *potenciação*, definem-se rigorosamente os sistemas de representação numérica posicional. Em particular, o sistema em base 10 permite descrever todos os números naturais a partir dos números fixados acima.

Seguindo a Definição 0.0.8 à risca, a descrição de uma função exige que sejam informados domínio e codomínio. No entanto, pode-se relaxar isso: é lícito dizer que uma **função** é meramente um conjunto de pares ordenados com a seguinte propriedade: $y = y'$ sempre que $(x, y), (x, y') \in f$. Em posse de uma função segundo tais critérios, recupera-se uma função no sentido da Definição 0.0.8 fazendo $\operatorname{dom}(f) := \{x : \text{existe } y \text{ tal que } (x, y) \in f\}$ e $\operatorname{im}(f) := \{y : \text{existe } x \text{ tal que } (x, y) \in f\}$, pois assim f se revela uma função do tipo $\operatorname{dom}(f) \to \operatorname{im}(f)$. A vantagem dessa reformulação é puramente técnica: ela permite falar de conjuntos de funções sem que precisemos explicitar o domínio de cada uma delas, o que deixa a escrita mais limpa.

Definição 0.3.7. Dadas funções f e g, diz-se que g **estende** f, ou g é uma **extensão** de f, se ocorrer $\operatorname{dom}(f) \subseteq \operatorname{dom}(g)$ e $g(x) = f(x)$ para todo $x \in \operatorname{dom}(f)$. ¶

A seguir, a notação $\bigcup \mathcal{F}$ indica a *reunião* da *família* $\mathcal{F}$, introduzida na Definição 0.1.20.

Lema 0.3.8. *Seja $\mathcal{F}$ uma família de funções. Se para quaisquer $f, g \in \mathcal{F}$ valer que $f(x) = g(x)$ sempre que $x \in \operatorname{dom}(f) \cap \operatorname{dom}(g)$, então $F := \bigcup \mathcal{F}$ é uma função cujo domínio é $\bigcup_{f \in \mathcal{F}} \operatorname{dom}(f)$. Em particular, $F(x) = f(x)$ para qualquer $f \in \mathcal{F}$ com $x \in \operatorname{dom}(f)$.*

Exercício 0.34 ($\star$)**.** Demonstre o lema acima. Dica: perceba que todo elemento de $\bigcup \mathcal{F}$ é um par ordenado; depois, para $(x, y), (x, z) \in \bigcup \mathcal{F}$, note que devem existir $f, g \in \mathcal{F}$ com $(x, y) \in f$ e $(x, z) \in g$, acarretando $x \in \operatorname{dom}(f) \cap \operatorname{dom}(g)$; conclua que $y = z$. ■

Teorema 0.3.9 (Recursão)**.** *Sejam $(\mathbb{B}, \leq)$ uma boa ordem natural, X um conjunto e $R \colon \mathbb{B} \times X \to X$ uma função*[†]. *Para cada $x \in X$ fixado, existe uma única função $R_x \colon \mathbb{B} \to X$ tal que $R_x(\min \mathbb{B}) = x$ e $R_x(\operatorname{suc}_{\mathbb{B}}(b)) = R(b, R_x(b))$ para cada $b \in \mathbb{B}$.*

Antes de provar o teorema acima, convém observar como ele permite construir funções *na prática*. Para o caso do fatorial, uma vez em posse dos naturais $\mathbb{N}$ e da operação de multiplicação $\cdot \colon \mathbb{N} \times \mathbb{N} \to \mathbb{N}$, tomam-se $\mathbb{B} := X := \mathbb{N}$, $R \colon \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ dada por $R(n, m) := (n + 1) \cdot m$ e $x := 1$. Note então que a função $R_1 \colon \mathbb{N} \to \mathbb{N}$ correspondente é tal que $R_1(0) = 1$ e $R_1(n + 1) = (n + 1) \cdot R_1(n)$ para todo $n \in \mathbb{N}$, i.e., R_1 é a função que faz $n \mapsto n!$, como desejado.

Demonstração. Para cada $b \in \mathbb{B}$, seja $\mathbb{B}_{<b} := \{c \in \mathbb{B} : c < b\}$. Vamos dizer que uma função $f \colon \mathbb{B}_{<b} \to X$ é *R_x-recursiva em b* se:

- $b = \min \mathbb{B}$, ou
- $b > \min \mathbb{B}$, $f(\min \mathbb{B}) = x$ e $f(\operatorname{suc}_{\mathbb{B}}(c)) = R(c, f(c))$ para cada $c < b$ tal que $\operatorname{suc}_{\mathbb{B}}(c) < b$.

Observe que para cada elemento $b > \min \mathbb{B}$, existe no máximo uma função R_x-recursiva em b: se tanto f quanto g são R_x-recursivas em b e $c < b$ é tal que $f(c) \neq g(c)$, então existe

$$c'' := \min\{c : c < b \text{ e } f(c) \neq g(c)\},$$

com $c'' > \min \mathbb{B}$ (pois $f(\min \mathbb{B}) = g(\min \mathbb{B})$); logo, existe $c' < c''$ com $\operatorname{suc}_{\mathbb{B}}(c') = c''$ e, consequentemente,

$$f(c'') = f(\operatorname{suc}_{\mathbb{B}}(c')) = R(c', f(c')) = R(c', g(c')) = g(\operatorname{suc}_{\mathbb{B}}(c')) = g(c'').$$

Em particular, como uma função R_x-recursiva em b' é também R_x-recursiva em $b \leq b'$ (por quê?!)$\overset{\star}{\star}$, segue que se f e g forem R_x-recursivas em b e b', respectivamente, então g estende f. Agora, provaremos que para cada $b \in \mathbb{B}$ *existe* uma função R_x-recursiva em b.

[†]Nesse tipo de situação, escreveremos $R(b, x)$ em vez de $R((b, x))$.

✓ (C.I.) Para $b := \min \mathbb{B}$, a função R_x-recursiva em b é $\emptyset$.

✓ (H.I.) Supondo que existe uma função R_x-recursiva em b, mostraremos que existe função R_x-recursiva em $\operatorname{suc}_{\mathbb{B}}(b)$:

  ✓ se $b := \min \mathbb{B}$, então $\mathbb{B}_{<\operatorname{suc}_{\mathbb{B}}(b)} = \{b\}$ e, assim, basta definir $f(b) := x$;

  ✓ se $b > \min\{\mathbb{B}\}$, então existe $c < b$ com $\operatorname{suc}_{\mathbb{B}}(c) = b$, de modo que $\mathbb{B}_{<b} = \mathbb{B}_{<c} \cup \{c\}$ e $\mathbb{B}_{<\operatorname{suc}_{\mathbb{B}}(b)} = \mathbb{B}_{<c} \cup \{c, b\}$; daí, se $f \colon \mathbb{B}_{<b} \to X$ é a função R_x-recursiva existente por hipótese, basta definir $g \colon \mathbb{B}_{<\operatorname{suc}_{\mathbb{B}}(b)} \to X$ fazendo $g(d) := f(d)$ se $d < b$ (i.e., $d \leq c$) e $g(b) := R(c, f(c))$.

Ao aliar a indução acima com a argumentação do primeiro parágrafo, resulta que para todo $b \in \mathbb{B}$ existe uma única função R_x-recursiva em b, digamos f_b, de tal forma que $f_{b'}$ estende f_b sempre que $b \leq b'$. Isso *permite considerar a família de funções* $\mathcal{F} := \{f_b : b \in \mathbb{B}\}$, que satisfaz as exigências do Lema 0.3.8. Logo, $R_x := \bigcup_{b \in \mathbb{B}} f_b$ é uma função da forma $\mathbb{B} \to X$ tal que $R_x(\min \mathbb{B}) = x$ e $R_x(\operatorname{suc}_{\mathbb{B}}(b)) = f_c(\operatorname{suc}_{\mathbb{B}}(b))$ para *qualquer* $c > b$, acarretando

$$R_x(\operatorname{suc}_{\mathbb{B}}(b)) = R(b, f_c(b)) = R(b, R_x(b)),$$

como desejado. Nesta altura do campeonato, você já deve saber como provar que R_x é a única com tal propriedade, certo? □

Exercício 0.35 ($^{\star}_{\star\star}$). Complete os detalhes da demonstração. Dica: será útil notar que $\bigcup_{b \in \mathbb{B}} \mathbb{B}_{<b} = \mathbb{B}$. ■

Para ilustrar um pouco mais o método da recursão (e tornar este texto mais honesto), vamos ver como definir as operações de adição e multiplicação precisamente — e sem as abanações de mão ~~cometidas em...~~ *canônicas*.

(+) Para cada $m \in \mathbb{N}$, seja $+_m \colon \mathbb{N} \to \mathbb{N}$ a função que faz $+_m(0) = m$ e $+_m(\operatorname{suc}_{\mathbb{N}}(n)) = \operatorname{suc}_{\mathbb{N}}(+_m(n))$ para cada $n \in \mathbb{N}$. Daí, defina $+ \colon \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ por $+(m, n) := +_m(n)$, que por simplicidade será denotado por $m + n$.

($\cdot$) Para cada $m \in \mathbb{N}$, seja $\cdot_m \colon \mathbb{N} \to \mathbb{N}$ a função que faz $\cdot_m(0) = 0$ e $\cdot_m(\operatorname{suc}_{\mathbb{N}}(n)) = \cdot_m(n) + m$. Daí, defina $\cdot \colon \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ por $\cdot(m, n) := \cdot_m(n)$, que por simplicidade será denotado por $m \cdot n$.

Exercício 0.36 ($^{\star}_{\star}$). Use o Teorema 0.3.9 para garantir a existência das funções acima. Dica: para $+_m$, tome $\mathbb{B} = X = \mathbb{N}$ no enunciado original, com $R \colon \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ dada por $R(n, y) := \operatorname{suc}_{\mathbb{B}}(y)$ e $x := m$; para $\cdot_m$, faça $R(n, y) := y + m$ (que existe pelo passo anterior!) e tome $x := 0$. ■

A rigor, as funções anteriores são regras arbitrárias que descrevem diferentes formas de iterar a função sucessor de $\mathbb{N}$. Porém, uma vez investigadas as propriedades de tais operações, percebe-se que elas agem de acordo com a experiência empírica. Por exemplo:

Proposição 0.3.10. *Para quaisquer $m, n \in \mathbb{N}$, tem-se $m + n = n + m$.*

Demonstração. Primeiro, tem-se $0 + n = n + 0$ para todo $n \in \mathbb{N}$: por definição, $n + 0 = n$ para todo $n \in \mathbb{N}$; por outro lado, $0 + 0 = 0$ e, se ocorrer $0 + n = n$, então

$$0 + \operatorname{suc}_{\mathbb{N}}(n) = \operatorname{suc}_{\mathbb{N}}(0 + n) = \operatorname{suc}_{\mathbb{N}}(n),$$

donde segue, por indução, que $0 + n = n$ para todo $n \in \mathbb{N}$. Agora, se para $m \in \mathbb{N}$ fixado ocorrer $m + n = n + m$ para todo $n \in \mathbb{N}$, então o mesmo valerá para $\operatorname{suc}_{\mathbb{N}}(m)$, pois... □

Exercício 0.37 $({}^{\star}_{\star\star})$**.** Complete a demonstração anterior. ■

Exercício 0.38 $({}^{\star}_{\star}\star{}^{\star}_{\star})$**.** Se estiver com paciência, reconstrua a Aritmética. ■

Sim, foram cinco estrelas: estas se reservam a exercícios que devem ser feitos no máximo uma vez na vida. Em outras palavras: foi uma piada.

§1 Recursão mais uma vez: demonstração do Teorema 0.3.5

Há pelo menos duas formas de demonstrar o Teorema 0.3.5 (de Dedekind): a primeira, braçal e honesta, e a segunda, rápida e malandra.

Demonstração braçal. Primeiro, note que *se existir* uma função com as propriedades impostas, então ela é única: com efeito, se $\psi\colon \mathcal{N} \to \mathcal{M}$ satisfaz as mesmas condições, então $X \coloneqq \{n \in \mathcal{N} : \varphi(n) = \psi(n)\}$ é tal que

- ✓ $i \in X$, pois $\varphi(i) = j = \psi(i)$, e
- ✓ se $n \in X$, então $s(n) \in X$, já que $\varphi(s(n)) = t(\varphi(n)) = t(\psi(n)) = \psi(s(n))$,

donde a suposição de $(\mathcal{N}, i, s)$ ser um sistema natural assegura $X = \mathcal{N}$.[†] O restante da prova consiste em fazer uma série de argumentações semelhantes.

Ainda supondo que existe uma função φ satisfazendo $\varphi(i) = j$ e $\varphi(s(n)) = t(\varphi(n))$ para todo $n \in \mathcal{N}$, provaremos que ela deve ser bijetora.

- ✓ É sobrejetora pois $Y \coloneqq \{m \in \mathcal{M} : \text{existe } n \in \mathcal{N} \text{ com } \varphi(n) = m\}$ é tal que $j \in Y$ (pois $\varphi(i) = j$) e $t(m) \in Y$ sempre que $m \in Y$ (pois se $n \in \mathcal{N}$ é tal que $\varphi(n) = m$, então $s(n) \in \mathcal{N}$ e $\varphi(s(n)) = t(\varphi(n)) = t(m)$), donde segue que $Y = \mathcal{M}$ (já que $(\mathcal{M}, j, t)$ é um sistema natural).
- ✓ Para verificar a injetividade, a ideia é mostrar que se $n \neq n'$, então $\varphi(n) \neq \varphi(n')$.

Em outras palavras, para $n \in \mathcal{N}$ fixado, busca-se provar que o conjunto $D_n \coloneqq \{n' \in \mathcal{N} : n' \neq n \Rightarrow \varphi(n') \neq \varphi(n)\}$ satisfaz $D_n = \mathcal{N}$. Tem início a *primeira indução*: mostraremos que $D_i = \mathcal{N}$ (Caso Inicial) bem como $D_{s(n)} = \mathcal{N}$ sempre que $D_n = \mathcal{N}$ (Hipótese Indutiva). Ocorre que para mostrar $D_i = \mathcal{N}$, também precisa-se argumentar por indução! Tem início a segunda indução:

- ✓ $i \in D_i$ (já que a implicação "$i \neq i \Rightarrow \varphi(i) \neq \varphi(i)$" é verdadeira[‡]);
- ✓ se $n \in D_i$ para algum $n \in \mathcal{N}$, pode-se ter $n = i$ ou $n \neq i$; no primeiro caso, $s(i) \neq i$ (certo?) e $\varphi(s(i)) = t(\varphi(i)) = t(j) \neq j$ (por quê?!), mostrando que $s(i) \in D_i$; no segundo caso, existe $n' \in \mathcal{N}$ com $s(n') = n$ (pois $(\mathcal{N}, i, s)$ é sistema natural) e $\varphi(s(n)) = t(\varphi(n)) = t(\varphi(s(n'))) = t(t(\varphi(n'))) \neq j$ (pois $t(m) \neq j$ para todo $m \in \mathcal{M} \setminus \{j\}$), mostrando que $s(n) \in D_i$.

O argumento acima encerra a segunda indução, mas não a primeira: mostrou-se apenas que $D_i = \mathcal{N}$; falta provar a segunda parte, i.e., que $D_{s(n)} = \mathcal{N}$ sempre que $D_n = \mathcal{N}$! E para surpresa de ninguém, tal igualdade será mostrada... na terceira indução:

[†] Na prática, o que se fez foi um argumento indutivo nos moldes da indução clássica descrita no Corolário 0.3.0, porém adaptada para sistemas naturais: provou-se o caso inicial ($i \in X$) e, supondo-se que $n \in X$ (hipótese indutiva), concluiu-se que $s(n) \in X$.

[‡] Posto que o seu *antecedente* é falso. Lembre-se de que afirmações do tipo "se P, então Q" só são falsas na ocorrência de *premissas* (P) verdadeiras com *conclusões* (Q) falsas.

✓ como antes, tem-se $s(n) \in D_{s(n)}$ por conta da igualdade $s(n) = s(n)$;

✓ agora, se $k \in D_{s(n)}$, deve-se provar que $s(k) \in D_{s(n)}$; se ocorrer $s(k) = s(n)$, nada precisa ser feito; se $s(k) \neq s(n)$, então $k \neq n$ (pois s é função!), enquanto o modo como tomamos n, i.e., satisfazendo $D_n = \mathcal{N}$, garante que $\varphi(k) \neq \varphi(n)$, donde finalmente a injetividade de t atesta $t(\varphi(k)) \neq t(\varphi(n))$, com a suposição sobre φ encerrando o trabalho, já que $\varphi(s(k)) = t(\varphi(k))$ e $\varphi(s(n)) = t(\varphi(n))$.

Portanto, mostrou-se que $D_i = \mathcal{N}$ (C.I.) e $D_{s(n)} = \mathcal{N}$ sempre que $D_n = \mathcal{N}$ (H.I.), exatamente o que se desejava. Falta apenas provar que *existe* uma função $\varphi \colon \mathcal{N} \to \mathcal{M}$ satisfazendo $\varphi(i) = j$ e $\varphi(s(n)) = t(\varphi(n))$ para todo $n \in \mathcal{N}$. Isto se faz via recursão: no enunciado do Teorema da Recursão, substitui-se $\mathbb{B}$ pelo sistema natural $(\mathcal{N}, i, s)$ com a boa ordem descrita no Exercício 0.77, toma-se $X := \mathcal{M}$, $x := j$ e faz-se $R_j \colon \mathcal{N} \times \mathcal{M} \to \mathcal{M}$ a função dada por $R(n, m) := t(m)$. Logo, $R_j(i) = j$ e

$$R_j(s(n)) = R(n, R_j(n)) = t(R_j(n))$$

para todo $n \in \mathcal{N}$, mostrando que a função φ procurada é, tão somente, R_j. □

A demonstração "rápida e malandra", por sua vez, apenas extrapola o uso do Exercício 0.77 na demonstração anterior: já que sistemas naturais podem ser substituídos por boas ordens naturais, basta provar a "versão bem ordenada" do Teorema de Dedekind.

Teorema 0.3.11 (de Dedekind, para boas ordens)**.** *Quaisquer duas boas ordens naturais são isomorfas. Além disso, o isomorfismo é único.*

As expressões "isomorfas" e "isomorfismo", que aparecerão no texto com frequência, têm definições bem gerais. Porém, por ora, basta saber que, no contexto acima, significam *bijeção crescente*. Mais precisamente, um **isomorfismo entre boas ordens** $(\mathbb{A}, \leq)$ e $(\mathbb{B}, \leq)$ é uma bijeção $\varphi \colon \mathbb{A} \to \mathbb{B}$ tal que $x \leq y \Rightarrow \varphi(x) \leq \varphi(y)$. Moralmente, a única diferença entre boas ordens isomorfas são os nomes de seus elementos, e o isomorfismo faz justamente o processo de *transliteração*[†].

Demonstração. Como de costume, convém começar a prova pela unicidade: se $\varphi, \psi \colon \mathbb{A} \to \mathbb{B}$ fossem isomorfismos distintos, então o conjunto

$$T := \{a \in \mathbb{A} : \varphi(a) \neq \psi(a)\}$$

seria não vazio e, portanto, existiria $t := \min T$. A pergunta é: quem é t? Note que t não pode ser o menor elemento de $\mathbb{A}$, pois $\varphi(\min \mathbb{A}) = \min \mathbb{B}$ e $\psi(\min \mathbb{A}) = \min \mathbb{B}$ (por quê?)*. Logo, por $\mathbb{A}$ ser boa ordem natural, existe $s \in \mathbb{A}$ tal que $t = \mathrm{suc}_{\mathbb{A}}(s)$. Agora, por minimalidade, deve-se ter $s \notin T$ e, por conseguinte, $\varphi(s) = \psi(s)$. Qual o problema? Este: por φ ser isomorfismo, $\varphi(\mathrm{suc}_{\mathbb{A}}(s)) = \mathrm{suc}_{\mathbb{B}}(\varphi(s))$ (verifique!)$^{\star}_{\star}$ e, pela mesma razão, $\psi(\mathrm{suc}_{\mathbb{A}}(s)) = \mathrm{suc}_{\mathbb{B}}(\psi(s))$, ou seja, $\varphi(t) = \psi(t)$.

Resta assim exibir um isomorfismo, o que é bem simples: pelo Teorema da Recursão, a função

$$\begin{aligned} \Phi \colon \mathbb{A} \times \mathbb{B} &\to \mathbb{B} \\ (a, b) &\mapsto \mathrm{suc}_{\mathbb{B}}(b) \end{aligned}$$

induz uma única função $\varphi \colon \mathbb{A} \to \mathbb{B}$ tal que $\varphi(\min \mathbb{A}) = \min \mathbb{B}$ e $\varphi(\mathrm{suc}_{\mathbb{A}}(a)) = \mathrm{suc}_{\mathbb{B}}(\varphi(a))$ para todo $a \in \mathbb{A}$. Onde está a rapidez e malandragem? `Resposta`: veja abaixo. □

Exercício 0.39 $(^{\star}_{\star\star})$**.** Mostre que a função φ acima é uma bijeção crescente. ■

[†]Como na Figura 0.4 da página 40.

§2 Boas ordens não naturais

Exercício 0.40 $(^{\star}_{\star\star})$**.** Mostre que a função sucessor $\mathrm{suc}_{\mathbb{B}} \colon \mathbb{B} \to \mathbb{B}$ é injetora em qualquer boa ordem $(\mathbb{B}, \leq)$ em que todo elemento tem sucessor. ■

A sutileza do exercício anterior é notar que na parte final da demonstração do Teorema 0.3.3, utilizou-se apenas o fato de todo elemento de $\mathbb{B}$ ter sucessor — e não a suposição de que todo elemento $\neq \min \mathbb{B}$ é sucessor de alguém. Isto levanta a pergunta: existem boas ordens satisfazendo a primeira condição mas não a segunda? Certamente!

Exemplo 0.3.12. Embora tenhamos pouca informação acerca de $\mathbb{N}$ (assumimos apenas que se trata de uma boa ordem/sistema natural), parace razoável dizer que existe $u' \notin \mathbb{N}$. Acontece que em posse disso, a Proposição 0.2.15 nos diz que $\mathbb{N}' := \mathbb{N} \cup \{u'\}$ admite uma boa ordem que tem u' como último elemento! É claro que, agora, nem todo elemento de $\mathbb{N}'$ tem sucessor: justamente por u' ser o último elemento, *ninguém* está acima dele. Porém, note que curioso: u' também não é sucessor (imediato) de ninguém! De fato, se $x \in \mathbb{N}'$ é tal que $x < u'$, então $x \in \mathbb{N}$ e, dessa forma, $\mathrm{suc}_{\mathbb{N}}(x) := x + 1 \in \mathbb{N}$ e, por isso, $x + 1 < u'$ (reveja a Definição 0.2.11 ou o Exercício 0.26 caso tenha se esquecido do que *significa* ser sucessor). ▲

$$0 \longrightarrow 1 \longrightarrow 2 \longrightarrow 3 \longrightarrow 4 \longrightarrow \cdots \longrightarrow n+1 \longrightarrow \cdots \qquad u'$$

Figura 0.5: Se uma boa ordem não tem último elemento, basta acrescentá-lo.

Embora pareça artificial, o cenário acima é bem comum na verdade: ao encontrar *sequências convergentes em* $\mathbb{R}$, por exemplo, podemos pensar nos termos da sequência *ordenados* por $\mathbb{N}$, de modo que o ponto adicional acrescentado acima faz meramente o papel de *limite*.

Por que parar? Chamando $\mathbb{N}_0 := \mathbb{N}$ e $\mathbb{N}_1 := \mathbb{N} \cup \{u'\}$, ainda parece razoável a existência de algum objeto $u'' \notin \mathbb{N}_1$, o que permite definir $\mathbb{N}_2 := \mathbb{N}_1 \cup \{u''\}$. Analogamente, parecece razoável definir $\mathbb{N}_3$, $\mathbb{N}_4$... Ao repetir esse processo para todo $n \in \mathbb{N}$, chega-se a uma boa ordem $\mathbb{B}$ em que todo elemento tem sucessor, mas tal que $\mathbb{B}$ não é natural! Caso tenha achado interessante, confira [30] para mais detalhes.

§3 Afinal, zero é natural?

É comum haver certo debate acerca da corretude de se assumir que $0 \in \mathbb{N}$. Embora se trate apenas de um símbolo utilizado para indicar o menor elemento de $\mathbb{N}$, as definições recursivas adotadas para a soma e a multiplicação escondem um *viés* ideológico: *naturalizar* o *vazio* (confira o Exercício 0.79).

Mais precisamente, na próxima seção, em que a noção de *cardinalidade* entre conjuntos arbitrários será, finalmente, apresentada, os elementos de $\mathbb{N}$ serão usados como *parâmetros* de finitude, num procedimento formal que apenas imita a noção de contagem. E é justamente aí que começa o problema: como contar os elementos do vazio?

✗ Quem defende "$0 \notin \mathbb{N}$" argumenta que ao *contar* os elementos em $\{a, b, c\}$, escolhe-se o *primeiro*, o *segundo* e, finalmente, o *terceiro*, de modo que ao término do processo se chega ao *número de elementos do conjunto*: 3. Há também quem apele a fatores históricos e etimológicos, já que a "noção do zero" ocorreu de maneira relativamente tardia, justificando assim que não se trate de uma ideia *natural*. Moral da história: não se contam os elementos de $\emptyset$.

✓ Já quem defende “$0 \in \mathbb{N}$” costuma pensar nos números naturais mais como *registradores* do processo de contagem: ao se iniciar a indexação a partir do 0, o *número de elementos do conjunto* será o menor número maior que os índices utilizados na indexação. Assim, por exemplo, como nem se começa a contagem dos elementos do vazio, o conjunto vazio deve ter $\min\{n \in \mathbb{N} : n > x \text{ para todo } x \in \emptyset\} = 0$ elementos (pense a respeito!). Já no caso de $\{a, b, c\}$, tem-se o 0-ésimo elemento, o 1-ésimo elemento e, finalmente, o 2-ésimo elemento, de modo que $\{a, b, c\}$ terá $\min\{n \in \mathbb{N} : n > 2\} = 3$ elementos.

Certamente, a postura “$0 \notin \mathbb{N}$” tem a vantagem de ser conceitualmente mais simples, talvez por sua proximidade com a experiência empírica. No entanto, ela é tecnicamente limitada: fica relativamente mais custoso descrever, por exemplo, o sistema numérico posicional utilizado corriqueiramente para dar sentido a coisas como “109”, essencialmente pela *ausência* de um *neutro* aditivo que *anule* a multiplicação; também ao expressar fórmulas de *cardinalidade*, como $|A \cup B| = |A| + |B| - |A \cap B|$ para conjuntos finitos A e B, por exemplo, precisa-se excluir os casos que envolvem o conjunto vazio, já que não se define $|\emptyset|$ como um número *digno* de ser operado com os *naturais*.

Porém, acredito que o argumento mais forte em favor de se adotar $0 \in \mathbb{N}$ seja o da *azeitona*[†]. Como você pode verificar por indução, para todo $n \in \mathbb{N}$ (com 0 incluso), existe bijeção entre $\mathbb{N}_{<n} := \{m : m \in \mathbb{N} \text{ e } m < n\}$ e $\mathbb{N}_{<n+1} \setminus \{0\}$, o que na prática significa o que você sempre soube: “contar de 1 até n e gritar n” é equivalente a “contar de 0 até $n - 1$ e gritar n”. Em outras palavras: se não quiser a azeitona na pizza, basta tirá-la da sua fatia e, se quiser depois, basta pegá-la de volta! Inclusive, isto será feito com frequência ao longo do texto.

0.4 Ao infinito e além

0.4.0 Essencial

§0 Conjuntos finitos

O final da discussão realizada na Subseção 0.1.1 §1 sugeriu definir *número cardinal* como um tipo de *representante* da relação $\approx$ que declara dois conjuntos A e B como equivalentes quando existe uma bijeção entre eles. Aqui, vamos executar esta ideia — pelo menos para conjuntos *finitos*. Por falar neles:

Definição 0.4.0. Diremos que um conjunto X é **finito** se para algum $n \in \mathbb{N}$ existir bijeção entre X e $\mathbb{N}_{<n} := \{m \in \mathbb{N} : m < n\}$. Conjuntos *não*-finitos serão xingados de **infinitos**. ¶

Exercício 0.41 ($\star$)**.** Sejam X e Y conjuntos tais que $X \approx Y$. Mostre que X é finito se, e somente se, Y é finito. ■

Note que pela definição acima, $\emptyset$ é finito, posto que $\mathbb{N}_{<0} = \emptyset$. Também são finitos os conjuntos unitários, i.e., da forma $\{x\}$, por estarem em bijeção com $\mathbb{N}_{<1} = \{0\}$, bem como os conjuntos da forma $\{x, y\}$ com $x \neq y$, por estarem em bijeção com $\mathbb{N}_{<2} = \{0, 1\}$, etc. Como os casos iniciais sugerem, se X é finito, então o número $n \in \mathbb{N}$ na definição de *finitude* é único, precisamente o tipo de comportamento esperado dos representantes de uma relação de equivalência (cf. Definição 0.1.25).

[†] Há outros, relativamente mais abstratos, que fogem ao escopo deste material.

Teorema 0.4.1. *Se X é finito, então existe um único $n \in \mathbb{N}$ tal que $X \approx \mathbb{N}_{<n}$.*

Demonstração. A definição garante a existência do número natural n. Para provar sua unicidade, considere $m, n \in \mathbb{N}$ com $X \approx \mathbb{N}_{<m}$ e $X \approx \mathbb{N}_{<n}$. Como a *relação* $\approx$ é simétrica e transitiva, resulta que $\mathbb{N}_{<m} \approx \mathbb{N}_{<n}$. Logo, para demonstrar o teorema, basta argumentar pela contrapositiva: vamos supor $m \neq n$ a fim de concluir que não existe bijeção entre $\mathbb{N}_{<m}$ e $\mathbb{N}_{<n}$. Isto encerrará a prova, posto que pela observação inicial, a ocorrência de $X \approx \mathbb{N}_{<m}$ e $X \approx \mathbb{N}_{<n}$ *acarretará* $m = n$, como desejado.

⌜**Afirmação.** *Se $Y \subsetneq \mathbb{N}_{<n}$, então não existe bijeção entre $\mathbb{N}_{<n}$ e Y.*

Demonstração. O argumento será feito por indução em n, com o caso $n := 0$ imediato (certo?)*. Supondo a afirmação verdadeira para $n \in \mathbb{N}$ fixado, veremos que a existência de uma bijeção φ entre $\mathbb{N}_{<n+1} = \mathbb{N}_{<n} \cup \{n\}$ e $Y \subsetneq \mathbb{N}_{<n+1}$ resultaria numa bijeção entre $\mathbb{N}_{<n}$ e um subconjunto próprio de $\mathbb{N}_{<n}$, contrariando a *hipótese de indução*[†]:

- ✓ se $n \notin Y$, então $Y \subseteq \mathbb{N}_{<n}$ e, portanto, a restrição de φ a $\mathbb{N}_{<n}$ seria uma bijeção entre $\mathbb{N}_{<n}$ e $Y' := Y \setminus \{\varphi(n)\}$, um subconjunto próprio de $\mathbb{N}_{<n}$ (percebeu?)*;

- ✓ se $n \in Y$, então existe $k \leq n$ com $\varphi(k) = n$, de modo que a função $g \colon \mathbb{N}_{<n} \to Y \setminus \{n\}$, que faz $g(i) := \varphi(i)$ para $i \neq k$ e $g(k) := \varphi(n)$ caso $k < n$, é uma bijeção — e $Y \setminus \{n\} \subsetneq \mathbb{N}_{<n}$, pois existe $b \in \mathbb{N}_{<n}$ tal que $b \notin Y$ (certo?!)*. ⌟

Agora, note que por $(\mathbb{N}, \leq)$ ser uma boa ordem, não há perda de generalidade em supor $m < n$, já que a ordem é total. Daí, $\mathbb{N}_{<m} \subsetneq \mathbb{N}_{<n}$, o que segue pois $m \in \mathbb{N}_{<n} \setminus \mathbb{N}_{<m}$ (e todo $k < m$ também satisfaz $k < n$, pela transitividade da ordem). Logo, pela afirmação, não existe bijeção entre $\mathbb{N}_{<m}$ e $\mathbb{N}_{<n}$, como queríamos. □

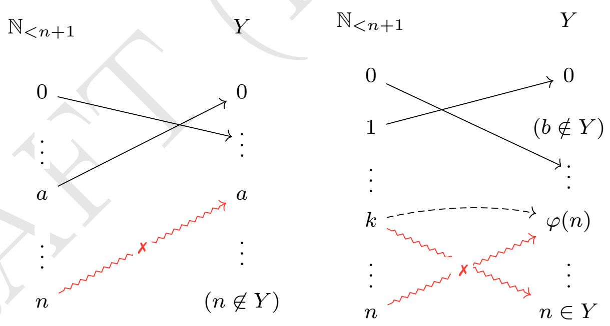


Figura 0.6: A demonstração anterior, em cores.

Definição 0.4.2. Dado um conjunto X finito, indicaremos por $|X|$ o único número natural n tal que $\mathbb{N}_{<n}$ está em bijeção com X, que será chamado de **número cardinal de** X. ¶

ALERTA É comum encontrar obras que utilizam "$\#X$" para indicar o número cardinal de X. Porém, isto não será feito aqui.

Portanto, como prometido, os números naturais cumprem o papel de representantes (das cardinalidades) dos conjuntos finitos por meio de bijeções que abstraem os processos de contagem. Como efeito colateral, todas as *ferramentas* típicas de $\mathbb{N}$, desde suas operações até as argumentações por indução, ficam disponíveis para analisar questões que envolvam a cardinalidade de conjuntos finitos. Para praticar, confira o Exercício 0.63.

[†] Em outras palavras, trata-se do passo indutivo usual, porém escrito na contrapositiva: "se **não** vale o caso $n+1$, então também **não** vale o caso n".

§1 Dois tipos de infinito

Além de firmar a posição dos números naturais como os cardinais de conjuntos finitos, a demonstração do último teorema também sugere, sem alarde, um critério para decidir se um conjunto é infinito. A seguir, $\varphi[Y] := \{\varphi(y) : y \in Y\}$ indica a *imagem de um subconjunto* $Y \subseteq X$ por uma função φ que tem X como domínio[†].

Corolário 0.4.3 (da demonstração do Teorema 0.4.1)**.** *Se X admite bijeção com um subconjunto próprio $Y \subsetneq X$, então X é infinito.*

Demonstração. O diagrama a seguir resume a ideia.

$$\begin{tikzcd} Y \arrow[rr, "\text{bijeção}"] \arrow[d, "\varphi|_Y"', shift right] & & X \arrow[d, "\varphi"'] \\ \varphi[Y] \arrow[u, bend right, "(\varphi|_Y)^{-1}"'] & \subsetneq & \mathbb{N}_{<n} \end{tikzcd}$$

Se existisse uma bijeção $\varphi\colon X \to \mathbb{N}_{<n}$ para *algum* $n \in \mathbb{N}$, então existiria bijeção entre $\varphi[Y]$ e $\mathbb{N}_{<n}$, contrariando a afirmação provada ao longo da demonstração do Teorema 0.4.1 (já que $\varphi[Y]$ é subconjunto próprio de $\mathbb{N}_{<n}$). Os detalhes ficam por sua conta $(^{\star}_{\star})$. $\square$

Corolário 0.4.4. $\mathbb{N}$ *é infinito.*

Demonstração. A correspondência $n \mapsto n+1$ define uma bijeção entre $\mathbb{N}$ e $\mathbb{N} \setminus \{0\}$. $\square$

A constatação de que existe um conjunto infinito cria um problema *momentâneo* na definição dos números cardinais: como nenhum número natural serve para representar a *cardinalidade* de $\mathbb{N}$, não parece haver uma noção "natural" de número que mereça ser xingada de $|\mathbb{N}|$. *Mas isto não deveria ser um problema, já que apenas os conjuntos infinitos ficaram sem cardinais e todos eles têm a mesma cardinalidade, justamente por serem infinitos... certo? Não bastaria fazer* $|\mathbb{N}| := \infty$*?* A situação não é tão simples.

Teorema 0.4.5 (Cantor)**.** *Dado um conjunto X, não existe sobrejeção $X \to \wp(X)$.*

Demonstração. Para uma função $\varphi\colon X \to \wp(X)$, o conjunto $T := \{x \in X : x \notin \varphi(x)\}$ atesta a não sobrejetividade de φ: se ocorresse $\varphi(t) = T$ para algum $t \in X$, a definição de T daria "$t \in T \Leftrightarrow t \notin \varphi(t) = T$", uma contradição. Logo, não há sobrejeção $X \to \wp(X)$. $\square$

Explicitamente, a demonstração acima *cozinha* um subconjunto de X que não pertence à imagem da função φ. Evidentemente, a depender da função φ, o conjunto construído muda — e pode, inclusive, ser vazio (*e tá tudo bem*). Experimente implementar a demonstração para um conjunto X pequeno, digamos $X := \{a, b, c\}$, com uma função $\varphi\colon X \to \wp(X)$ de sua escolha, e perceba que o conjunto T da demonstração, necessariamente, não pertence à imagem da função que você escolheu.

Exercício 0.42 $(^{\star}_{\star\star})$**.** Mostre que se $X \subseteq Y$ e Y é finito, então X é finito. Observação: este é o terceiro item do Exercício 0.63. ■

Corolário 0.4.6. *Se X é infinito, então $\wp(X)$ é infinito e $X \not\approx \wp(X)$.*

[†]Caso precise revisar esse assunto, confira o Exercício 0.55.

Demonstração. Como a correspondência $x \mapsto \{x\}$ define uma função injetora $X \to \wp(X)$, segue que se $\wp(X)$ fosse finito, então X estaria em bijeção com um subconjunto (finito!) de $\wp(X)$ e, portanto, X seria finito. O restante segue do Teorema de Cantor. $\square$

Em particular, $\mathbb{N}$ e $\wp(\mathbb{N})$ são conjuntos infinitos que *não* têm a mesma cardinalidade, ou seja: existem cardinalidades infinitas distintas entre si, o que mostra a inefetividade de se escrever "$|X| = \infty$" para representar ~~decentemente~~ os números cardinais de conjuntos infinitos. Para contornar o problema, costuma-se fazer o seguinte.

Definição 0.4.7. Diremos que X é **enumerável** se existir função injetora da forma $X \to \mathbb{N}$. Nas situações em que se puder garantir a existência de bijeção $X \to \mathbb{N}$ (e isto precisar ser explicitado), X será dito **infinito enumerável**. Por fim, diremos que X é **não enumerável** se X *não for* enumerável, i.e., se não existir injeção $X \to \mathbb{N}$. ¶

Exercício 0.43 ($\star$)**.** Mostre que se X é não enumerável, então X é infinito. ■

Exercício 0.44 ($^\star_\star$)**.** Sejam X e Y conjuntos com $X \subseteq Y$. Mostre que se X é não enumerável, então Y é não enumerável. Dica: contrapositiva. ■

Proposição 0.4.8. $\mathbb{N} \times \mathbb{N}$ *é (infinito!) enumerável.*

Demonstração. Como você deve saber de suas aulas de Aritmética (ou afins), a correspondência $(m, n) \mapsto 2^m 3^n$ define uma função injetora $\mathbb{N} \times \mathbb{N} \to \mathbb{N}$. Dado que existem funções injetoras (óbvias?)* da forma $\mathbb{N} \to \mathbb{N} \times \mathbb{N}$, o resultado segue por Cantor-Bernstein. $\square$

Os primeiros contatos com esse tipo de resultado costumam ser estranhos, já que parece ser claro que $\mathbb{N} \times \mathbb{N}$ deveria ter bem mais elementos do que $\mathbb{N}$: com efeito, $\mathbb{N} \times \mathbb{N} = \bigcup_{n \in \mathbb{N}} \mathbb{N} \times \{n\}$, i.e., é uma *reunião* enumerável de conjuntos enumeráveis dois a dois disjuntos e, mesmo assim, sua cardinalidade não excede a cardinalidade de $\mathbb{N}$. Contudo, uma simples ilustração ajuda a tornar a coisa toda mais palatável.

$$
\begin{tikzcd}
 & 0 & 1 & 2 & 3 & 4 & \cdots \\
0 & (0,0) \arrow[r] & (1,0) \arrow[dl] & (2,0) \arrow[d] & (3,0) \arrow[r] & (4,0) \arrow[d] \\
1 & (0,1) \arrow[r] & (1,1) \arrow[ur] & (2,1) \arrow[dl] & (3,1) \arrow[u] & (4,1) \arrow[dl] \\
2 & (0,2) \arrow[d] & (1,2) \arrow[l] & (2,2) \arrow[ur] & (3,2) \arrow[dl] & \\
3 & (0,3) \arrow[r] & (1,3) \arrow[ur] & (2,3) \arrow[dl] & & \\
4 & (0,4) \arrow[d] & (1,4) \arrow[l] & & & \\
\vdots & \vdots & & & &
\end{tikzcd}
$$

Como o diagrama acima sugere, pode-se determinar uma bijeção da forma $\mathbb{N} \to \mathbb{N} \times \mathbb{N}$ ao *percorrer* $\mathbb{N} \times \mathbb{N}$ num zigue-zague maroto. O mesmo tipo de *percurso* também ajuda a perceber que o *conjunto dos números racionais* é infinito enumerável: a única diferença é que precisa-se evitar representações *equivalentes* de um mesmo número racional já *percorrido*, pois sem isso a injetividade se perderia. Quer uma boa notícia? Isto não traria problemas:

Teorema 0.4.9 (©)**.** *Uma função $f\colon X \to Y$ é sobrejetora se, e somente se, existe uma injeção $g\colon Y \to X$ tal que $f \circ g = \mathrm{Id}_Y$. Em particular, $Y \precsim X$ se, e somente se, $X \succsim Y$.*

Demonstração. Se a função g existe como no enunciado, então a sobrejetividade de f segue do Exercício 0.14 (a menos da ordem das letras). A parte delicada é a recíproca: como cozinhar uma função $g\colon Y \to X$ satisfazendo $f \circ g = \mathrm{Id}_Y$? Note que a identidade procurada se traduz em pedir $f(g(y)) = y$ para todo $y \in Y$. Na prática, isto significa dizer que $g(y)$ é *algum* elemento de X que *foi levado* até y por f.

Algum $x \in X$ satisfaz $f(x) = y$, posto que f é sobrejetora por hipótese, o que garante $P_y := f^{-1}[\{y\}] \neq \emptyset$ para todo $y \in Y$. Ocorre que $\mathcal{P} := \{P_y : y \in Y\}$ é uma partição de X (certo?!)*, de tal maneira que o *insuspeito* Teorema 0.1.26 assegura uma classe de representantes para $\mathcal{P}$, digamos $\mathcal{R}$. Com isso, basta definir $g(y) := r$, onde r é o único elemento de $\mathcal{R}$ pertencente a $\mathcal{P}_y$, que satisfaz a identidade $f(g(y)) = y$ por *construção*. □

Corolário 0.4.10 (©)**.** *Se $f\colon X \to Y$ é sobrejetora, então $Y \precsim X$.*

Corolário 0.4.11 (©)**.** *Seja $\mathcal{X} := \{X_n : n \in \mathbb{N}\}$ uma família de conjuntos. Se cada X_n é enumerável, então $\bigcup \mathcal{X}$ é enumerável.*

Demonstração. Para cada $n \in \mathbb{N}$, seja $\mathrm{Sob}(\mathbb{N}, X_n)$ o conjunto das funções sobrejetoras da forma $\mathbb{N} \to X_n$. Como $\mathrm{Sob}(\mathbb{N}, X_n) \neq \emptyset$, podemos *escolher* $f_n \in \mathrm{Sob}(\mathbb{N}, X_n)$ para cada $n \in \mathbb{N}$ (como na demonstração do Teorema 0.1.26). Com isso, pode-se definir $f\colon \mathbb{N} \times \mathbb{N} \to \bigcup_{n\in\mathbb{N}} X_n$ a função que faz $(m, n) \mapsto f_m(n)$. Note que se $x \in \bigcup_{n\in\mathbb{N}} X_n$, então existe $m \in \mathbb{N}$ com $x \in X_m$, bem como $n \in \mathbb{N}$ tal que $f_m(n) = x$, i.e., $f(m, n) = x$. Logo, f é sobrejetora, acarretando $\bigcup \mathcal{X} \precsim \mathbb{N} \times \mathbb{N} \approx \mathbb{N}$. □

Tais resultados podem ser usados para mostrar que os típicos conjuntos $\mathbb{Z}$ (dos números inteiros) e $\mathbb{Q}$ (dos números racionais), embora infinitos, têm a mesma caradinalidade de $\mathbb{N}$, o que ainda pode parecer muito estranho. Por um lado, isto se deve ao fato de a *construção* dos conjuntos em questão ser feita por meio de métodos que não *aumentam* cardinalidades infinitas (vide Proposição 0.4.8). Por outro lado, a *menor* cardinalidade infinita é a de $\mathbb{N}$, no seguinte sentido:

Teorema 0.4.12 (©)**.** *Se X é infinito, então $\mathbb{N} \precsim X$.*

Demonstração. A ideia é simples: por X ser infinito, não existe $n \in \mathbb{N}$ com uma bijeção $\mathbb{N}_{<n} \to X$; logo, se $n \in \mathbb{N}$ e $\varphi\colon \mathbb{N}_{<n} \to X$ for uma função injetora, então $X \setminus \mathrm{im}\,(\varphi)$ é não vazio, o que permite *escolher*, *recursivamente*, sequências finitas *cada vez maiores* de modo a obter uma injeção $\mathbb{N} \to X$. Parece honesto, certo? □

Exercício 0.45 $(^{\star}_{\star})$**.** Para um conjunto X, mostre que as seguintes afirmações são equivalentes:

a) o conjunto X é infinito (não existe $n \in \mathbb{N}$ com uma bijeção $\mathbb{N}_{<n} \to X$);

b) existe uma função injetora $\mathbb{N} \to X$;

c) existe uma função bijetora $X \to Y$ para algum $Y \subsetneq X$. ■

Exercício 0.46 $(^{\star}_{\star})$**.** Mostre que X é não enumerável se, somente se, $\mathbb{N} \prec X$ (cf. Definição 0.1.9, item (ii)). Dica: Exercício 0.43 + Cantor-Bernstein. ■

Saber que $\mathbb{N}$ tem "*a menor cardinalidade infinita*" sugere a pergunta *natural*: qual é a *próxima* cardinalidade? Certamente, aplicações sucessivas do Teorema de Cantor resultam em diversas cardinalidades não enumeráveis: $\mathbb{N} \prec \wp(\mathbb{N}) \prec \wp(\wp(\mathbb{N})) \prec \ldots$, o que não responde a pergunta, já que *poderia existir* X com $\mathbb{N} \prec X \prec \wp(\mathbb{N})$, caso em que a cardinalidade de $\wp(\mathbb{N})$ não seria a "*próxima*"[†]. Voltaremos a isso, mas não agora. Em todo caso, resta apenas um fato inócuo sobre conjuntos não enumeráveis para ser apresentado — e, devido à sua simplicidade, a verificação ficará por sua conta.

Exercício 0.47 (Princípio da Casa dos Pombos ($^\star_\star$))**.** Sejam X um conjunto não enumerável e $A, B \subseteq X$ subconjuntos disjuntos satisfazendo $X = A \cup B$. Mostre que deve ocorrer $\mathbb{N} \prec A$ ou $\mathbb{N} \prec B$, i.e., (pelo menos) um deles deve ser não enumerável. Dica: suponha que não. ■

0.4.1 Extras

§0 Uma demonstração para o Teorema de Cantor-Bernstein

É importante ter em mente que a demonstração que veremos não é terrivelmente difícil e muito menos feia — ela é belíssima. No entanto, ela requer atenção e tempo, e o segundo recurso costuma ser escasso nas disciplinas de Análise. Recordemo-nos de seu enunciado:

Teorema de Cantor-Bernstein. Se existem funções injetoras $X \to Y$ e $Y \to X$, então existe bijeção $X \to Y$.

Você não deve esperar que a demonstração do Teorema de Cantor-Bernstein envolva a prova de que as injeções são sobrejetivas[‡]: a ideia é *construir* uma "nova" função (bijetora) a partir das injeções dadas. Embora a coisa seja mais simples do que parece, alguns ingredientes preliminares devem ser introduzidos.

Definição 0.4.13. Para um conjunto $\mathcal{S}$, define-se $\bigcap \mathcal{S} := \{x : x \in S \text{ para todo } S \in \mathcal{S}\}$, a **interseção da família** $\mathcal{S}$. Nas ocasiões em que $\mathcal{S} := \{S_i : i \in \mathcal{I}\}$ para algum conjunto $\mathcal{I}$ fixado, também é comum escrever $\displaystyle\bigcap_{i \in \mathcal{I}} S_i$ ou $\bigcap_{i \in \mathcal{I}} S_i$. ¶

O dispositivo acima apenas cria um modo bastante esperto de evitar abominações notacionais como "$S_0 \cap S_1 \cap \ldots$" quando se quer expressar uma interseção (possivelmente) infinita de conjuntos. Fora isso, ela *quase* não traz novidades: note, por exemplo, que para $\mathcal{S} := \{X, Y\}$, tem-se $\bigcap \mathcal{S} = X \cap Y$.

Lema 0.4.14 (Leis de De Morgan)**.** *Sejam X e $\mathcal{Y}$ conjuntos, com $\mathcal{Y} \neq \emptyset$.*

(i) $\displaystyle\bigcap_{B \in \mathcal{Y}} (X \setminus B) = X \setminus \bigcup_{B \in \mathcal{Y}} B.$

(ii) $\displaystyle\bigcup_{B \in \mathcal{Y}} (X \setminus B) = X \setminus \bigcap_{B \in \mathcal{Y}} B.$

[†]Em particular fica o **alerta**: afirmações do tipo "X e Y são não enumeráveis" não significam, *a priori*, que X e Y tenham a mesma cardinalidade, mas apenas que ambas as cardinalidades são *maiores* do que a cardinalidade de $\mathbb{N}$.

[‡]Lembre-se da função sucessor!

Demonstração. Se $x \in \bigcap_{B\in\mathcal{Y}}(X \setminus B)$, então para todo $B \in \mathcal{Y}$ tem-se $x \in X \setminus B$, donde segue que $x \in X$ e não existe $B \in \mathcal{Y}$ tal que $x \in B$, i.e., $x \in X$ e $x \notin \bigcup_{B\in\mathcal{Y}} B$, precisamente $x \in X \setminus \bigcup_{B\in\mathcal{Y}} B$. Pela arbitrariedade do x tomado, segue que $\bigcap_{B\in\mathcal{Y}}(X \setminus B) \subseteq X \setminus \bigcup_{B\in\mathcal{Y}} B$. A recíproca é análoga (faça!)$^{\star}_{\star}$.

Para provar o segundo item, note que se $x \in \bigcup_{B\in\mathcal{Y}}(X \setminus B)$, então existe $B' \in \mathcal{Y}$ com $x \in X \setminus B'$, i.e., $x \in X$ e $x \notin B'$ para algum $B' \in \mathcal{Y}$, donde se infere que $x \notin \bigcap_{B\in\mathcal{Y}} B$ e, por conseguinte, $x \in X \setminus \bigcap_{B\in\mathcal{Y}} B$. Logo, $\bigcup_{B\in\mathcal{Y}}(X \setminus \mathcal{Y}) \subseteq X \setminus \bigcap_{B\in\mathcal{Y}} B$. Novamente, a recíproca é análoga (já sabe né?)$^{\star}_{\star}$. □

Exercício 0.48 ($\star$)**.** Para conjuntos A, B e C, mostre que $C \setminus (A \cup B) = (C \setminus A) \cap (C \setminus B)$ e $C \setminus (A \cap B) = (C \setminus A) \cup (C \setminus B)$. ■

Lema 0.4.15. *Sejam $f\colon X \to Y$ uma função e considere famílias $\mathcal{U}$ e $\mathcal{V}$ de subconjuntos de X e de Y, respectivamente. Então:*

(*i*) $f\left[\bigcup \mathcal{U}\right] = \bigcup_{U\in\mathcal{U}} f[U]$ *e* $f\left[\bigcap \mathcal{U}\right] \subseteq \bigcap_{U\in\mathcal{U}} f[U]$*, com igualdade garantida no último caso se f for injetora;*

(*ii*) $f^{-1}\left[\bigcup \mathcal{V}\right] = \bigcup_{V\in\mathcal{V}} f^{-1}[V]$ *e* $f^{-1}\left[\bigcap \mathcal{V}\right] = \bigcap_{V\in\mathcal{V}} f^{-1}[V]$.

Exercício 0.49 ($^{\star}_{\star}$)**.** Demonstre o lema anterior. ■

Demonstração do Teorema de Cantor-Bernstein. Fixadas funções injetoras $f\colon X \to Y$ e $g\colon Y \to X$, vamos obter partições $\{S, S'\}$ e $\{T, T'\}$ de X e Y, respectivamente, tais que $f[S] = T$ e $g[T'] = S'$ pois, se isso for feito, o teorema estará demonstrado (confira o Exercício 0.50). Para obter tal partição, note que qualquer subconjunto $S \subseteq X$ induz tanto uma partição em X quanto uma partição em Y: basta definir $S' := X \setminus S$, $T := f[S]$ e $T' := Y \setminus T$. Isso resolve *metade* do problema, dado que ainda é preciso garantir a identidade $g[T'] = S'$, essencial para definir a bijeção h. Ao se reescrever a igualdade $g[T'] = S'$ em *função* de S, obtém-se a identidade $g\left[Y \setminus f[S]\right] = X \setminus S$, equivalentemente exprimível como $S = X \setminus g\left[Y \setminus f[S]\right]$. Como obter tal S?

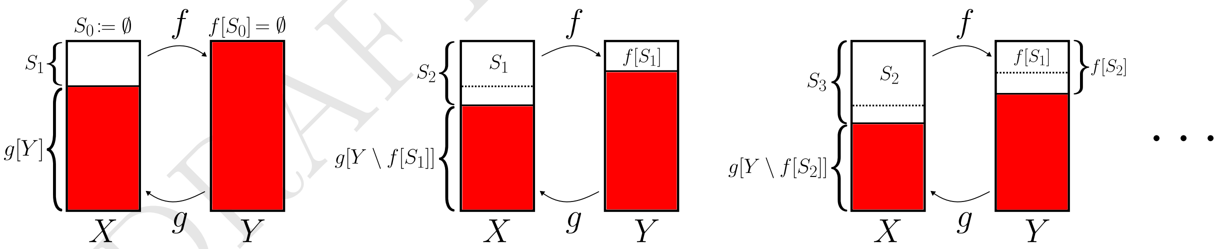


Figura 0.7: Heurística da construção recursiva do conjunto S.

Seja $S_0 := \emptyset$ e, para $n \in \mathbb{N}$ qualquer, $S_{n+1} := X \setminus g\left[Y \setminus f[S_n]\right]$. O subconjunto $S := \bigcup_{n\in\mathbb{N}} S_n$ satisfaz a identidade desejada, pois

$$\begin{aligned}
S &:= \bigcup_{n\in\mathbb{N}} S_n = \bigcup_{n\in\mathbb{N}} S_{n+1} = \bigcup_{n\in\mathbb{N}} \left(X \setminus g\left[Y \setminus f[S_n]\right]\right) = X \setminus \bigcap_{n\in\mathbb{N}} g\left[Y \setminus f[S_n]\right] \\
&\overset{*}{=} X \setminus g\left[\bigcap_{n\in\mathbb{N}} (Y \setminus f[S_n])\right] = X \setminus g\left[Y \setminus \bigcup_{n\in\mathbb{N}} f[S_n]\right] = X \setminus g\left[Y \setminus f\left[\bigcup_{n\in\mathbb{N}} S_n\right]\right] \\
&= X \setminus g\left[Y \setminus f\left[S\right]\right],
\end{aligned}$$

onde a igualdade ($*$) decorre da injetividade de g (lema anterior), enquanto as outras seguem das leis de De Morgan. □

Exercício 0.50 ($\star\atop\star$)**.** Mostre que se $F\colon A \to B$ e $G\colon C \to D$ são bijeções com $A \cap C = B \cap D = \emptyset$, então $F \cup G$ é uma bijeção da forma $A \cup C \to B \cup D$. Use isto para obter a bijeção procurada no teorema anterior. Dica: para a segunda parte, e com as notações da demonstração anterior, faça $A := S$, $B := T$, $C := S'$ e $D := T'$. ■

Portanto, a fim de estabelecer a *existência* de uma bijeção entre conjuntos X e Y, basta exibir injeções $X \to Y$ e $Y \to X$, sem a necessidade de explicitar uma bijeção particular. Isto será feito, por exemplo, quando provarmos a *não enumerabilidade* de $\mathbb{R}$ por meio de uma bijeção entre $\mathbb{R}$ e $\wp(\mathbb{N})$, cuja *existência* será garantida por funções injetoras da forma $\mathbb{R} \to \wp(\mathbb{N})$ e $\wp(\mathbb{N}) \to \mathbb{R}$.

§1 Construção dos inteiros e dos racionais

Um dos modos de *motivar* a *introdução* dos números inteiros é dar *significado* a expressões como "$5 - 7$", que não têm sentido no contexto natural: tipicamente, uma expressão da forma "$m - n$" em $\mathbb{N}$ corresponde à cardinalidade resultante de um conjunto com m elementos após a exclusão de n elementos; logo, se $m < n$, então não se pode realizar tal procedimento. Todavia, como os Bancos nos ensinam desde tempos imemoriais, é possível registrar "quanto falta": no caso, faz-se "$7 - 5 = 2$", e escreve-se "$5 - 7 = -2$" para indicar a natureza "negativa" do resultado†.

Para *descrever* o nosso entendimento do que os inteiros *deveriam* ser dentro do cenário formal que se desenrola, pode-se observar o seguinte: embora coisas como "$5 - 7$" e "$3 - 5$" (que deveriam resultar em -2) não tenham significado em $\mathbb{N}$, a *informação* transmitida por tais expressões pode ser *codificada* em $\mathbb{N}$: em vez de escrever $5 - 7 = 3 - 5$, redistribuem-se as parcelas de modo a se obter uma igualdade entre somas positivas, i.e., $5 + 5 = 7 + 3$. Em outras palavras, duas *expressões* $a - b$ e $c - d$ são iguais (no que gostaríamos que fosse $\mathbb{Z}$) se, e somente se, $a + d$ e $b + c$ são iguais em $\mathbb{N}$. Em linguajar técnico:

Exercício 0.51 ($\star\atop\star\star$)**.** Sobre $Z := \mathbb{N} \times \mathbb{N}$, considere a relação $\sim$ que declara $(a, b) \sim (c, d)$ se, e somente se, $a + d = b + c$.

a) Mostre que $\sim$ é uma relação de equivalência. Dica: para a transitividade, use a *lei do cancelamento*, i.e., "$a = b$ sempre que $a + c = b + c$, para quaisquer $a, b, c \in \mathbb{N}$".

b) Mostre que se $(a, b) \sim (a', b')$ e $(c, d) \sim (c', d')$, então $(a + c, b + d) \sim (a' + c', b' + d')$.

c) Mostre que ao definir $\overline{(a, b)} +' \overline{(c, d)} := \overline{(a + c, b + d)}$, o resultado *independe da escolha de representantes*. Dica: encare o item anterior até que ele te encare de volta‡. ■

Definição 0.4.16. O quociente $Z/{\sim}$ do exercício acima será denotado por $\mathbb{Z}$ e xingado de **conjunto dos números inteiros**. ¶

No caso, a classe de equivalência $\overline{(a, b)}$ representa o que gostaríamos de escrever como $a - b$, razão pela qual vamos escrever assim! Desse modo, a *operação* $+'$ definida no exercício anterior apenas estipula

$$(a - b) +' (c - d) := (a + c) - (b + d),$$

tal qual aprendemos na escola.

†A aceitação dos números negativos pela comunidade matemática foi relativamente tardia — ainda no Século XIX havia quem encrencasse com eles. Quem gostar de aspectos históricos pode conferir [34].

‡Por exemplo: por um lado, $\overline{(1, 2)} +' \overline{(2, 5)} = \overline{(3, 7)}$, enquanto $\overline{(2, 3)} +' \overline{(3, 6)} = \overline{(5, 9)}$ e, de fato, $\overline{(3, 7)} = \overline{(5, 9)}$, já que $3 + 9 = 7 + 5$. Elaborar exemplos particulares para entender afirmações aparentemente abstratas é algo que você deve se acostumar a fazer por conta própria.

Exercício 0.52 $({}^{\star}_{\star}\star{}^{\star}_{\star})$**.** Verifique as propriedades usuais da soma de números inteiros. ■

O "exercício" acima empurra para debaixo do tapete a verificação de diversas propriedades importantes do conjunto $\mathbb{Z}$ como construído aqui, mas que essencialmente permitem tratá-lo como já estamos acostumados — em particular, o apóstrofo empregado no símbolo de adição já pode ser abandonado. De maneira similar, ao definir

$$(a-b)\cdot'(c-d) := (ac+bd)-(ad+bc),$$

verifica-se que o lado direito da expressão acima (a rigor, uma classe de equivalência em $Z/{\sim}$), não depende da escolha dos representantes das classes no lado esquerdo, essencialmente como no caso da adição "$+'$". Após verificar que tal operação $\cdot'$ tem o comportamento que se esperaria da multiplicação em $\mathbb{Z}$, a construção de $\mathbb{Q}$ se torna inevitável.

Exercício 0.53 $({}^{\star}_{\star\star})$**.** Sobre $Q := \mathbb{Z}\times(\mathbb{Z}\setminus\{0\})$, considere a relação $\sim$ que declara $(a,b)\sim(c,d)$ se, e somente se, $ad = bc$.

a) Mostre que $\sim$ é uma relação de equivalência.

b) Chamando por $\frac{a}{b}$ a classe de equivalência de um par (a,b), mostre que $Q/{\sim}$ se comporta, essencialmente, como o ***conjunto*** **dos números racionais** que conhecemos na escola, e que passará a ser denotado por $\mathbb{Q}$. Em particular, defina as operações de adição e multiplicação e verifique suas propriedades usuais. ■

Com $\mathbb{Z}$ e $\mathbb{Q}$ *formalmente* apresentados, já é possível determinar suas *cardinalidades* por meio dos resultados vistos na subseção anterior (entre outros exercícios elementares propostos no final do capítulo):

(i) como existem funções injetoras $\mathbb{N}\to\mathbb{Z}$ e $\mathbb{Z}\to\mathbb{Q}$, resulta que $\mathbb{N}\precsim\mathbb{Z}\precsim\mathbb{Q}$;

(ii) por construção, $\mathbb{Q}$ vem de fábrica com uma sobrejeção da forma $\mathbb{Z}\times(\mathbb{Z}\setminus\{0\})\to\mathbb{Q}$, mostrando que $\mathbb{Q}\precsim\mathbb{Z}\times(\mathbb{Z}\setminus\{0\})$;

(iii) em geral, se $A\approx A'$ e $B\approx B'$, então $A\times B\approx A'\times B'$, de modo que por valer $\mathbb{Z}\approx\mathbb{Z}\setminus\{0\}$, resulta $\mathbb{Z}\times(\mathbb{Z}\setminus\{0\})\approx\mathbb{Z}\times\mathbb{Z}$;

(iv) finalmente, $\mathbb{Z}=\bigcup_{n\in\mathbb{N}}\{-n,n\}$, com $\{-n,n\}\precsim\mathbb{N}$ para todo $n\in\mathbb{N}$, resultando em $\mathbb{Z}\precsim\mathbb{N}$ (logo, $\mathbb{Z}\approx\mathbb{N}$) e, como acima, $\mathbb{Z}\times\mathbb{Z}\approx\mathbb{N}\times\mathbb{N}\approx\mathbb{N}$.

Em suma:

$$\mathbb{N}\precsim\mathbb{Z}\precsim\mathbb{Q}\precsim\mathbb{Z}\times(\mathbb{Z}\setminus\{0\})\approx\mathbb{Z}\times\mathbb{Z}\approx\mathbb{N},$$

donde o Teorema de Cantor-Bernstein assegura as clássicas *identidades* $\mathbb{N}\approx\mathbb{Z}\approx\mathbb{Q}$, i.e., tanto $\mathbb{Z}$ quanto $\mathbb{Q}$ são enumeráveis.

§2 O problema da escolha e o sentido da existência

Talvez você tenha notado o símbolo "Ⓒ" perdido em alguns pontos do texto. Caso tenha passado batido, sua primeira ocorrência foi no Teorema 0.1.26. "*Será que tal teorema está protegido por direitos autorais?*".

Resposta: não. A ideia foi marcar no texto os pontos em que *fomos* displicentes com os *problemas de escolha* — e, em inglês, "escolha" se escreve "*choice*" (Ⓒhoice).

Por mais que tenhamos firmado o *compromisso* de *aceitar* conjuntos infinitos, as ferramentas que temos para compreendê-los são *finitárias*: a vida, os símbolos, o papel, a paciência... são recursos indiscutivelmente finitos, o que leva à exigência (bastante) razoável de que *demonstrações* também devem ser *finitas* em algum sentido. Certamente, dado que textos empregam, invariavelmente, apenas finitos caracteres, a noção de finitude que se espera de uma demonstração não deve ser pensada apenas textualmente, mas *computacionalmente*.

Exemplo 0.4.17 (Fundamental)**.** Como ilustração, considere a demonstração apresentada para o Teorema 0.4.12. Secretamente, o argumento *sugere* como definir, recursivamente, uma injeção da forma $\mathbb{N} \to X$ sempre que X for infinito:

- ✓ por X ser infinito, tem-se $X \neq \emptyset$, o que permite *escolher* $x_0 \in X$;
- ✓ por X ser infinito, tem-se $X \neq \{x_0\}$, o que permite *escolher* $x_1 \in X \setminus \{x_0\}$;
- ✓ por X ser infinito...
- ✓ por fim, a função $\mathbb{N} \to X$ que faz $n \mapsto x_n$ é injetora.

Implicitamente, o raciocínio acima descreve um processo de construção *infinitamente arbitrário*: cada novo passo *exige* um `input`, uma escolha *indeterminada* que deve ser feita por quem executa a demonstração. Para ilustrar uma situação que não sofre deste problema, suponha que X seja bem ordenado: neste caso, bastaria definir $x_0 := \min X$, $x_1 := \min X \setminus \{x_0\}$, *e assim sucessivamente*. Pode parecer a mesma coisa, mas não é: no último caso, o argumento explicita que a escolha deve ser feita por meio da boa ordem nativa de X, tomando-se *o* menor dentre os elementos ainda não escolhidos, o que independe de quem executa a demonstração[†]. ▲

O que fazer diante disso? Infelizmente, não há resposta *filosoficamente* fácil, uma vez que o questionamento atinge diretamente a *ontologia matemática*: afinal, quando usamos o verbo "existir" numa sentença matemática, como em "existe solução para a equação $x^2 + 1 = 0$", qual o sentido da palavra "existe"?

Exemplo 0.4.18. Futuramente, veremos que *o tipo de objeto que se convenciona chamar de reta real*, $\mathbb{R}$, é não enumerável e, além disso, pode ser escrito como $\mathbb{R} = \mathbb{A} \cup \mathbb{T}$, onde $\mathbb{A}$ é a coleção dos números reais que são raízes de polinômios cujos coeficientes são racionais[‡], e $\mathbb{T} := \mathbb{R} \setminus \mathbb{A}$. Com alguma paciência, mostra-se que $\mathbb{A}$ é enumerável, ao que você provavelmente se pergunta: e daí? Ora, como $\mathbb{R}$ é não enumerável, conclui-se que $\mathbb{T}$ deve ser não enumerável, pois o contrário acarretaria a enumerabilidade de $\mathbb{R}$ (cf. Exercício 0.47). Portanto, sem exibir qualquer elemento de $\mathbb{T}$, provou-se que $\mathbb{T} \neq \emptyset$. ▲

A aceitação desse tipo de argumento não é unanimidade entre a comunidade matemática: há quem defenda que por mais abstratas que sejam, as entidades matemáticas devem ser *exibidas*. Para essas pessoas, não basta provar que a inexistência de algo leva a uma contradição, é preciso argumentar diretamente a fim de justificar a *existência*.

[†] Trata-se de uma versão menos divertida da *anedota das meias*, criada por Russell para ilustrar o uso do Axioma da Escolha: numa coleção infinita de pares de meias, a fim de tomar uma meia de cada par, precisa-se escolher arbitrariamente uma meia de cada par (já que meias de um mesmo par costumam ser indistinguíveis); já para uma coleção infinita de pares de sapato, basta escolher o pé esquerdo de cada par.

[‡] Como $\sqrt{2}$, raiz do polinômio $x^2 - 2$.

Porém, tal postura construtivista não é predominante: o mais comum, na verdade, é encarar *existência* como *ausência de contradições* perante algum sistema axiomático fixado, o que não resolve o problema ontológico[†], mas pelo menos possibilita um sono tranquilo para quem pratica Matemática.

De volta ao problema da escolha: como aprendi com Hugo C. Botós,

"para quem só *sabe* martelo, todo parafuso é prego".

Aqui, martelo é o modo finitário por meio do qual argumentamos, enquanto os parafusos são os conjuntos infinitos. É para tratá-los (minimamente) como ~~pregos~~ conjuntos finitos que se postula o

Axioma da Escolha. *Para uma família não vazia $\mathcal{A}$ de conjuntos não vazios, existe uma função $f\colon \mathcal{A} \to \bigcup \mathcal{A}$ tal que $f(A) \in A$ para cada $A \in \mathcal{A}$.*

Observação 0.4.19 (Um pouco de contexto, mas não muito)**.** Na tentativa de resolver o primeiro problema da *lista de exercícios* proposta por David Hilbert em 1900, um jovem chamado Ernst Zermelo *demonstrou*, em 1904, o que ficou conhecido como

Teorema 0.4.20 (da Boa Ordenação (©))**.** *Todo conjunto admite uma boa ordem.*

Mais precisamente, qualquer *conjunto* X admite pelo menos uma relação binária $\preceq$ que faz de $(X, \preceq)$ uma boa ordem[‡]. O problema de Hilbert em questão era a *Hipótese do Contínuo*, que consiste em saber se existe (ou não) um conjunto infinito X com $\mathbb{N} \prec X \prec \wp(\mathbb{N})$: no caso, a relação com a noção de boa ordem se dá por um resultado de Cantor que estabelece o conjunto das boas ordens de $\mathbb{N}$ (a menos de *isomorfismo*) como um representante da *primeira* cardinalidade *maior* do que a cardinalidade de $\mathbb{N}$. No entanto, o que interessa para a presente discussão é o fato de Zermelo ter explicitado, pela primeira vez até então, a suposição de que escolhas arbitrárias poderiam ser feitas, i.e., o Axioma da Escolha[††]. △

Embora exista quem aponte tal axioma como incompatível com posturas finitárias, pode-se argumentar que foi justamente o comprometimento com tais posturas que levou à percepção de que escolhas infinitas não podem ser realizadas de maneira *construtiva*. Nesse sentido, em vez de banir os resultados que dependem de tais processos, Zermelo propôs aceitar a *existência* de *escolhas completas* sem a limitação de descrevê-las ou construí-las, algo bem mais *honesto*.

Enfim, vejamos esboços de como justificar as passagens marcadas com © (exceto pelo Teorema da Boa Ordenação, cuja demonstração foge do escopo deste trabalho)[‡‡].

✓ No Teorema 0.1.26, a classe de representantes se obtém por meio do Axioma da Escolha ao considerar $\mathcal{A}$ como a família das classes de equivalência da relação (ou como a própria partição, a depender do caso), i.e., basta tomar $\mathcal{R} := \{f(A) : A \in \mathcal{A}\}$.

[†]Por exemplo: o sistema axiomático fixado descreve entidades que existem de forma independente (num *universo platônico*) ou tudo não passa de um jogo artificial sofisticado? Eu não tenho uma resposta definitiva. Se quiser se aprofundar nesse tipo de discussão para tentar formular a sua própria resposta, você pode conferir o tratado [39] ou a envolvente discussão de Penelope Maddy [27] — se preferir algo mais modesto, o texto de Joel D. Hamkins [16] também é uma boa pedida.

[‡]Tal ordem não tem qualquer tipo de comprometimento com *outras ordens* previamente existentes em X: no caso de $\mathbb{Z}$, por exemplo, poderia ocorrer $3 = \min_{\preceq} \mathbb{Z}$, $-1 = \mathrm{suc}_{\preceq}(3)$, $19 = \mathrm{suc}_{\preceq}(-1)$, etc.

[††]A rigor, em 1904, Zermelo tratou o Axioma da Escolha como mera suposição; foi apenas em 1908 que ele propôs uma axiomatização para a teoria de conjuntos que englobasse sua suposição como um dos axiomas fundamentais.

[‡‡]Para mais detalhes, confira [10, 30].

✓ O Teorema 0.4.9 e seu Corolário 0.4.10 apenas utilizaram o teorema anterior.

✓ Para o Corolário 0.4.11, ao fazer $\mathcal{B} := \{\text{Sob}(\mathbb{N}, X_n) : n \in \mathbb{N}\}$, o Axioma da Escolha *assegura* uma função $f : \mathbb{N} \to \bigcup \mathcal{B}$ com $f(n) \in \text{Sob}(\mathbb{N}, X_n)$ para cada $n \in \mathbb{N}$ (por quê?)$^{\overset{\star}{\star\star}}$.

✓ Finalmente, para o Teorema 0.4.12, com $\mathcal{A} := \wp(X) \setminus \{\emptyset\}$, o Axioma da Escolha dá uma função $f : \mathcal{A} \to X$ com $f(A) \in A$ para todo $A \in \mathcal{A}$, de modo que ao chamar por $\text{seq}\,(X)$ a família das funções da forma $\mathbb{N}_{<n} \to X$ para cada $n \in \mathbb{N}$, pode-se definir a função $\mathcal{O} : \text{seq}\,(X) \to X$ que faz $\mathcal{O}(s) := f(X \setminus \text{im}\,(s))$, que explicitamente *escolhe* um elemento em X não pertencente à imagem da *sequência finita* $s \in \text{seq}\,(X)$. Adaptando-se a ideia da demonstração do Teorema 0.3.9, mostra-se que existe uma (única) função $\psi : \mathbb{N} \to X$ tal que $\psi(n) = \mathcal{O}((\psi(m) : m < n))$ para cada $n \in \mathbb{N}$, que pelo modo com que a função $\mathcal{O}$ foi tomada, deve ser injetora.

Exercício 0.54 ($\overset{\star}{\star\star}$)**.** Surpreendentemente, a demonstração do Teorema de Cantor-Bernstein não usou o Axioma da Escolha. Por quê? ■

Ocasionalmente mencionaremos o uso do Axioma da Escolha ao longo do texto, mas você não precisa se preocupar com isso nesta etapa da sua vida.

§3 Como representar os infinitos?

Outro símbolo foi empregado despretensiosamente no texto: "Ⓡ". Sua ocorrência foi única, na Definição 0.1.27, mas não sem importância: ela indica que a definição está *condenada* pelo Paradoxo de Ⓡussell (Exercício 0.11). De fato, como se discutiu na subseção em questão, um dos modos de eliminar o Paradoxo de Russell consiste em substituir o Axioma da Abstração pelo Axioma da Separação: dada uma propriedade $\mathcal{P}$ e um conjunto A, existe $\{x \in A : x \text{ tem a propriedade } \mathcal{P}\}$. Uma consequência disso foi propositalmente omitida:

Teorema 0.4.21. *Não existe conjunto V para o qual se tenha $x \in V$ para todo x.*

Demonstração. Se existisse, então o Axioma da Separação permitiria definir o conjunto $R := \{x \in V : x \notin x\}$, que traria novamente o Paradoxo de Russell: note que $R \in R$ se, e somente se, $R \notin R$ (certo?)*. □

Observação 0.4.22. A inexistência de um conjunto universo é menos problemática do que parece. Ao estudar Teoria dos Números, por exemplo, $\mathbb{Z}$ é o universo dos objetos estudados (os números inteiros!), e ninguém reclama de se ter "$\mathbb{Z} \notin \mathbb{Z}$", i.e., $\mathbb{Z}$ não é um número inteiro. Da mesma forma, aqui, $\mathbb{V} \notin \mathbb{V}$, i.e., o universo dos conjuntos não é um conjunto. △

Apesar da observação acima, a impossibilidade de considerar $\mathbb{V}$ como conjunto atrapalha o tratamento adotado para as noções de cardinalidade, principalmente dos conjuntos infinitos: no caso de conjuntos finitos, os números naturais servem precisamente como representantes, mas nada foi apresentado aqui para cumprir tal papel no caso de conjuntos infinitos. O modo usual para resolver o problema faz uso dos chamados *ordinais de von Neumann*.

A ideia é quase simples: define-se $0 := \emptyset$, $1 := \{0\}$, $2 := \{0,1\}$, $3 := \{0,1,2\}$, etc. Mais geralmente, um **ordinal** é um conjunto α com duas propriedades peculiares: α é *transitivo*, no sentido de que $x \subseteq \alpha$ sempre que $x \in \alpha$, e $(\alpha, \in_\alpha)$ é uma boa ordem (estrita), onde $\in_\alpha$ é a relação de ordem (estrita) em α dada por $x \in_\alpha y$ se, e somente se, $x \in y$. Com paciência, mostra-se que $0, 1, 2, 3, \ldots$, definidos como acima, são ordinais, mas não só isso: ao considerar a coleção de todos os números naturais implementados dessa forma, ganha-se um conjunto ω que também é um ordinal. Ao repetir o processo de sucessão acima e definir $\omega + 1 := \omega \cup \{\omega\}$, ganha-se novamente um ordinal, e ao definir $\omega + 2 := \omega + 1 \cup \{\omega + 1\}$... Os *números cardinais* desejados *surgem* então como ordinais especiais: um ordinal α é **cardinal** se não existe $\beta \in \alpha$ que admita bijeção com α.

Ao desenvolver as ideias acima com tempo (num curso de Teoria dos Conjuntos, por exemplo), prova-se que todo número natural é um cardinal, assim como ω, que passa a ser chamado de $\aleph_0$, o primeiro cardinal infinito. Depois de argumentar um pouco mais, chega-se à conclusão de que deve existir $\aleph_1$, o primeiro cardinal maior do que $\aleph_0$, bem como $\aleph_2$, o primeiro cardinal maior do que $\aleph_1$, bem como $\aleph_3$, o primeiro... Após chegar a $\aleph_\omega$ (o primeiro cardinal maior do que $\aleph_n$ para todo n natural), chega-se à conclusão de que deve existir $\aleph_{\omega+1}$, o primeiro... Você pegou a ideia, certo? Para mais detalhes, confira [10, 30].

Observação 0.4.23. No caso em que X é um conjunto finito, a expressão "$|X|$" indica o seu número cardinal. Agora, pelo que se discutiu acima, é lícito adotar a mesma notação para X infinito, mesmo que não tenhamos discutido cardinais em detalhes. Se isso te incomodar, ao encontrar expressões como "$|\mathbb{R}| \leq |\wp(\mathbb{N})|$", por exemplo, pense que se trata de uma abreviação estilosa para a afirmação "existe função injetora do tipo $\mathbb{R} \to \wp(\mathbb{N})$". Analogamente, "$|A| = |B|$" abrevia "existe bijeção entre A e B" e assim por diante. $\triangle$

0.5 Exercícios adicionais

Importante: nos exercícios desta seção e pelo restante do texto, as notações para cardinalidades seguem o que se discutiu na Observação 0.4.23.

Essenciais

Exercício 0.55 (Imagens, pré-imagens, restrições, etc. ($\star$))**.** Para uma função $f\colon X \to Y$ e um subconjunto $W \subseteq X$, a **restrição** da função f ao subconjunto W é a função $f|_W\colon W \to Y$ definida por $f|_W(w) := f(w)$ para todo $w \in W$. Dizemos que $g\colon Z \to Y$ **estende** f se $X \subseteq Z$ e $g|_X = f$. Mais ainda, para subconjuntos $A \subseteq X$ e $B \subseteq Y$ fixados, consideram-se:

(i) a **imagem direta** de A por f, definida como $f[A] := \{f(a) : a \in A\}$;

(ii) a **pré-imagem** de B por f, definida como $f^{-1}[B] := \{x \in X : f(x) \in B\}$.

Com base nas definições acima, faça o que se pede a seguir.

a) Mostre que $\operatorname{im}(f|_A) = f[A]$ para qualquer $A \subseteq X$.

b) Mostre que $f[A] \subseteq f[A']$ e $f^{-1}[B] \subseteq f^{-1}[B']$ sempre que $A \subseteq A' \subseteq X$ e $B \subseteq B' \subseteq Y$, respectivamente.

c) Mostre que $f[\emptyset] = \emptyset$ e $f^{-1}[\emptyset] = \emptyset$.

d) Mostre que $f^{-1}[Y] = X$.

e) Para $A \subseteq X$ e $B \subseteq Y$ quaisquer, mostre que valem as inclusões $A \subseteq f^{-1}[f[A]]$ e $f[f^{-1}[B]] \subseteq B$, com as igualdades garantidas se f for injetora (para o primeiro caso) ou sobrejetora (para o segundo caso). ■

Importante. Explicitamente, $y \in f[A]$ se, e somente se, existe algum $a \in A$ com $f(a) = y$, enquanto $x \in f^{-1}[B]$ se, e somente se, $f(x) \in B$. Embora a pressa possa fazer você pensar o contrário, ao escrever "$f^{-1}[B]$", **não se supõe** que a função f seja *invertível*/bijetora: "f^{-1}" se refere à *relação inversa* de f, que pode não ser uma função (confira a discussão na Subseção 0.1.1 §0). Na dúvida, *sempre* ~~que possível~~ use a definição dada para uma notação — e não a sua intuição para o que a notação deveria ser.

Exercício 0.56 ((!))**.** Leia o parágrafo anterior novamente. ■

Definição 0.5.0. Para conjuntos X e Y, denotaremos por Y^X a família de todas as funções da forma $X \to Y$.[†] ¶

Exercício 0.57 ($^{\star}_{\star}$)**.** Para um conjunto X qualquer (possivelmente infinito!), exiba uma bijeção entre $\wp(X)$ e $\{0,1\}^X$. Dica: para um subconjunto $A \subseteq X$, associe uma função "óbvia" da forma $X \to \{0,1\}$; para facilitar, comece com X finito para daí generalizar. ■

Exercício 0.58 ($^{\star}_{\star}$)**.** Mostre que se $|A| = |B|$, então $|\wp(A)| = |\wp(B)|$. ■

Exercício 0.59 ($^{\star}_{\star}$)**.** Sejam A, B, C e D conjuntos. Mostre que se $|A| \leq |C|$ e $|B| \leq |D|$, então $|A^B| \leq |C^D|$. ■

Exercício 0.60 ($^{\star}_{\star}$)**.** Mostre que se $|A| = |C|$ e $|B| = |D|$, então $|A^B| = |C^D|$. ■

Observação 0.5.1. Note que o exercício anterior justifica definir potenciação entre cardinalidades/números cardinais: $|X|^{|Y|} := |X^Y|$. △

Observação 0.5.2. Independentemente do que será discutido nos próximos capítulos acerca das chamadas *indeterminações*, a definição acima obriga que se tenha $0^0 = 1$, o que está de acordo com o fato de que existe uma única função da forma $\emptyset \to \emptyset$. △

Exercício 0.61 ($^{\star}_{\star\star}$)**.** Para conjuntos X, Y e Z *quaisquer*, exiba uma bijeção entre $(X^Y)^Z$ e $X^{Y \times Z}$. Qual a interpretação cardinal disso? ■

Exercício 0.62 ((!))**.** Se você fez o exercício acima apenas para os casos em que X, Y e Z são finitos, volte e faça de novo, para **conjuntos quaisquer**: a hipótese de finitude é irrelevante para a solução do exercício[‡]. ■

Exercício 0.63 ($^{\star}_{\star\star}$)**.** Demonstre as seguintes afirmações.

a) Para todo $n \in \mathbb{N}$ ocorre $|\mathbb{N}_{<n}| = n$.

b) Se X é finito e $x \notin X$, então $|X \cup \{x\}| = |X| + 1$.

c) Se $X \subseteq Y$ e Y é finito, então X é finito e $|X| \leq |Y|$. Dica: indução em $|Y|$ + item anterior.

[†] O motivo da notação exponencial é sugerido, implicitamente, pelo último item do Exercício 0.63.
[‡] De modo geral, só assuma que os conjuntos são finitos quando isso for explicitamente indicado.

d) Se X e Y são finitos, então $X \cup Y$ é finito e $|X \cup Y| \leq |X| + |Y|$. Dica: indução em $|X|$ + item (b) para o subcaso do passo indutivo em que $X \nsubseteq Y$.

e) Se X e Y são finitos e disjuntos, então $|X \cup Y| = |X| + |Y|$.

f) Se $\mathcal{F}$ é finito e todo $F \in \mathcal{F}$ é finito, então $\bigcup \mathcal{F}$ é finito. Dica: indução em $|\mathcal{F}|$.

g) Se X e Y são finitos, então $X \times Y$ é finito e $|X \times Y| = |X| \cdot |Y|$. Dica: note que $X \times Y = \bigcup_{y \in Y} X \times \{y\}$.

h) Se X é finito e $f\colon X \to Y$ é uma função, então $\operatorname{im}(f)$ é finito e $|\operatorname{im}(f)| \leq |X|$.

i) Se $X \times Y$ é finito, então X e Y são finitos.

j) Se X é finito, então $\wp(X)$ é finito e $|\wp(X)| = 2^{|X|}$. Dica: combine o próximo item com o Exercício 0.57.

k) Se X e Y são finitos, então Y^X é finito e $|Y^X| = |Y|^{|X|}$. ∎

Exercício 0.64 $(^{\star}_{\star\star})$**.** Mostre que se X é finito, então existem $|X|!$ bijeções da forma $X \to X$. Dica: indução em $|X|$. ∎

Exercício 0.65 $(^{\star}_{\star})$**.** Mostre que para $n \in \mathbb{N}$, uma função $f\colon \mathbb{N}_{<n} \to \mathbb{N}_{<n}$ é injetora se, e somente se, é sobrejetora. ∎

Exercício 0.66 $(^{\star}_{\star\star})$**.** Mostre que para X finito, uma função $f\colon X \to X$ é injetora se, e somente se, é sobrejetora. ∎

Exercício 0.67 $(^{\star}_{\star})$**.** O resultado acima permanece válido para X infinito? ∎

Exercício 0.68 $(^{\star}_{\star})$**.** Pense rápido: se $f\colon \mathbb{N} \to \mathbb{N}$ é uma função sobrejetora, então existe $n \in \mathbb{N}$ tal que $f^{-1}[\{n\}]$ é finito? ∎

Exercício 0.69 $(^{\star}_{\star})$**.** Seja $f\colon X \to Y$ uma função, com Y enumerável e X não enumerável. Mostre que existe $y \in Y$ tal que $\{x \in X : f(x) = y\}$ é não enumerável. Dica: confira o Exercício 0.74. ∎

Exercício 0.70 $(^{\star}_{\star})$**.** Sejam $(\mathbb{S}, \leq)$ e $(\mathbb{T}, \leq)$ ordens parciais. Uma função $f\colon \mathbb{S} \to \mathbb{T}$ é:

(i) **crescente** se para quaisquer $s, s' \in \mathbb{S}$ valer que $s \leq s' \Rightarrow f(s) \leq f(s')$;

(ii) **decrescente** se para quaisquer $s, s' \in \mathbb{S}$ valer que $s \leq s' \Rightarrow f(s) \geq f(s')$;

(iii) **monótona** se f for crescente ou decrescente.

Sabendo disso, suponha que a ordem de $\mathbb{S}$ seja total.

a) Mostre que se f é crescente e $f(s) < f(s')$, então $s < s'$.

b) Mostre que se f é decrescente e $f(s) < f(s')$, então $s > s'$.

c) Conclua que se f for monótona e bijetora, então a inversa $f^{-1}\colon \mathbb{T} \to \mathbb{S}$ também será monótona[†]. ∎

[†]Na verdade, f é crescente/decrescente se, e somente se, f^{-1} é crescente/decrescente.

Observação 0.5.3. Uma função crescente e injetora satisfaz a implicação *estrita*

$$s < s' \Rightarrow f(s) < f(s'),$$

já que deve ocorrer $f(s) \leq f(s')$, enquanto a injetividade proíbe $f(s) = f(s')$. Uma função em tal condição será chamada de **estritamente crescente**. Funções **estritamente decrescentes** são análogas[†]. $\triangle$

Exercício 0.71 $(\star)$**.** Convença-se de que funções estritamente monótonas são injetoras. $\blacksquare$

Exercício 0.72 $(\substack{\star\\\star\star})$**.** Suponha que $X \subseteq \mathbb{N}$ seja um subconjunto com as seguintes propriedades: $k \in X$ e $n+1 \in X$ sempre que $n \in X$. Em tais condições, mostre que $X = \{n \in \mathbb{N} : k \geq n\}$. Dica: observe que $\{n \in \mathbb{N} : n \geq k\}$ é uma boa ordem natural com a ordem herdada de $\mathbb{N}$ (cf. Teorema 0.3.11 ou a demonstração da Proposição 1.3.1). $\blacksquare$

Preciosismo Como bem observado por Pryscilla Silva, o exercício acima justifica por que podemos fazer indução *a partir de um número natural fixado*. Com alguma paciência, você não deve ter dificuldades para perceber que o resultado vale, mais geralmente, para $X \subseteq \mathbb{Z}$ e $k \in \mathbb{Z}$ quaisquer.

Extras

Exercício 0.73 $(\substack{\star\\\star\star})$**.** Para um conjunto X e uma função $\Phi\colon \operatorname{seq}(X) \to X$, mostre que existe uma única função $\varphi\colon \mathbb{N} \to X$ tal que $\varphi(0) = \Phi(\emptyset)$ e $\varphi(n+1) = \Phi((\varphi(0), \ldots, \varphi(n)))$ para todo $n \in \mathbb{N}$. $\blacksquare$

Exercício 0.74 $(\substack{\star\\\star})$**.** Seja $f\colon X \to Y$ uma função. Mostre que $\mathcal{P} := \{f^{-1}[\{y\}] : y \in \operatorname{im}(f)\}$ é uma partição de X (cf. Definição 0.1.22). $\blacksquare$

Exercício 0.75 $(\substack{\star\\\star})$**.** Seja $\psi\colon X \to Y$ uma função bijetora. Mostre que se $\mathcal{P}$ é uma partição de X, então $\psi(\mathcal{P}) := \{\psi[P] : P \in \mathcal{P}\}$ é uma partição de Y. Em particular, tem-se $|\mathcal{P}| = |\psi(\mathcal{P})|$. $\blacksquare$

Exercício 0.76 $(\substack{\star\\\star\star})$**.** Mostre que existe uma partição infinita $\mathcal{P} := \{P_n : n \in \mathbb{N}\}$ de $\mathbb{N}$ com cada P_n infinito. Dica: arrume uma função sobrejetora $\mathbb{N} \times \mathbb{N} \to \mathbb{N}$ esperta, ou apele para números *primos*. $\blacksquare$

Exercício 0.77 $(\substack{\star\star\\\star\star})$**.** Seja $(\mathcal{N}, i, s)$ um sistema natural. Diremos que $I \subseteq \mathcal{N}$ é **indutivo** se $s(n) \in I$ sempre que $n \in I$. Agora, para $m, n \in \mathcal{N}$, vamos escrever "$m \preceq n$" a fim de indicar que "n pertence a todo conjunto indutivo que contém m". O propósito deste exercício é mostrar que $(\mathcal{N}, \preceq)$ é uma boa ordem natural cujo sistema natural induzido é $(\mathcal{N}, i, s)$.

a) Convença-se de que $\preceq$ é reflexiva e transitiva.

b) Mostre que se $I \subseteq \mathcal{N}$ é indutivo, então $s[I]$ também é.

c) Mostre que se $m, n \in \mathcal{N}$ e $m \npreceq n$, então $s(m) \npreceq s(n)$ e $s(m) \npreceq n$. Dica: pela definição de $\preceq$, deve existir um subconjunto indutivo I satisfazendo $m \in I$ com $n \notin I$; use o item anterior para concluir.

[†] A literatura também costuma xingar de *não decrescente* as coisas que aqui foram chamadas de *crescentes*. Em tais textos, o adjetivo *crescente* se reserva para as situações de desigualdade estrita. Um comentário análogo é válido para funções *não crescentes*.

d) Mostre que para todo $n \in \mathcal{N}$ ocorre $n \leq s(n)$ e $s(n) \not\leq n$. Dica: para a segunda parte, com $X := \{n \in \mathcal{N} : s(n) \not\leq n\}$, note que $i \in X$ e $s(n) \in X$ sempre que $n \in X$.

e) Mostre que $(\mathcal{N}, \leq)$ é uma ordem parcial. Sugestão: verifique a antissimetria pela contrapositiva, i.e., mostre que se $m \neq n$, então $m \not\leq n$ ou $n \not\leq m$, o que pode ser feito por *indução* como no item anterior.

f) Mostre que se $m \leq n$, então $s(m) \leq s(n)$.

g) Mostre que se $I \subseteq \mathcal{N}$ é indutivo, então $s^{-1}[I]$ é indutivo.

h) Definindo "$m < n$" como abreviação para "$m \leq n$ e $m \neq n$", mostre que se $s(m) < s(n)$, então $m < n$. Dica: use o item anterior.

i) Mostre que $m < s(n)$ se, e somente se, $m \leq n$. Dica: primeiro, observe que se $m < s(i)$, então $m = i$; depois, note que se a tese valer para $n \in \mathcal{N}$ e $s(k) < s(s(n))$ para algum k, então $k < s(n)$ em virtude do item anterior; conclua por *indução*.

j) Suponha $X \subseteq \mathcal{N}$ com a seguinte propriedade: para todo $n \in \mathcal{N}$, a ocorrência de $m \in X$ para todo $m < n$ acarreta $n \in X$. Mostre que $X = \mathcal{N}$. Dica: considere $Y := \{n \in \mathcal{N} : \forall m \in \mathcal{N}\ m < n \Rightarrow m \in X\}$ e mostre por indução que $Y = \mathcal{N}$, ponto em que o item anterior será essencial; para concluir, lembre-se de que $n < s(n)$.

k) Mostre que $(\mathcal{N}, \leq)$ é uma boa ordem. Dica: tome $S \subseteq \mathcal{N}$ sem menor elemento e use o item anterior para concluir que $\mathcal{N} \setminus S = \mathcal{N}$, i.e., $S = \emptyset$.

l) Mostre que $i = \min \mathcal{N}$ e $\mathrm{suc}_{\mathcal{N}}(n) = s(n)$ para todo $n \in \mathcal{N}$. ■

AVISO: quatro estrelas se reservam a exercícios que devem ser feitos ao menos uma vez na vida.

Exercício 0.78 $(^{\star}_{\star\star})$**.** Para $m, n \in \mathbb{N}$, mostre que $m < n$ se, e somente se, existe $c \in \mathbb{N}$ tal que $n = m + (c + 1)$. Dica: faça por indução em n (e suponha conhecidas as propriedades operatórias da adição em $\mathbb{N}$). ■

Exercício 0.79 $(^{\star\star}_{\star\star})$**.** Suponha que você ministre um curso em que $0 \notin \mathbb{N}$. Como definir, recursivamente (cf. Exercício 0.36), as operações de soma e multiplicação em $\mathbb{N}$? ■

0.6 Linguagem: revisão de estruturas algébricas

0.6.0 Essencial

§0 Operações binárias e elementos especiais

Definição 0.6.0. Uma função $*: X \times X \to X$ é chamada de **operação binária** em X. Porém, como não trataremos explicitamente de outras *aridades*, não há risco em chamar $*$ simplesmente de operação. Para $x, y \in X$, costuma-se escrever $x * y$ em vez de $*(x, y)$, em alusão às notações já bem estabelecidas para *adição* e *multiplicação*. ¶

A operação $*$ determina o que é $x * y$ em X. Independentemente disso, como $x * y \in X$, é legítimo operar tal elemento com algum $z \in X$, situação em que se escreve $(x * y) * z$. Os parênteses são necessários pois, a princípio, poderia ocorrer $(x * y) * z \neq x * (y * z)$, entre outros comportamentos *indesejados*. Porém, como o tempo é curto, vamos nos focar nos casos que interessam a este curso.

Definição 0.6.1. Seja $*\colon X \times X \to X$ uma operação num conjunto X.

(i) Diremos que a operação $*$ é **associativa** se para quaisquer $x, y, z \in X$ ocorrer $(x * y) * z = x * (y * z)$, o que na prática significa que pode-se escrever $x * y * z$ em vez de $(x * y) * z$ ou $x * (y * z)$.

(ii) Diremos que a operação $*$ é **comutativa** se para quaisquer $x, y \in X$ valer a identidade $x * y = y * x$.

(iii) Diremos que $e \in X$ é **elemento neutro**[†] da operação $*$ se para qualquer $x \in X$ ocorrer $e * x = x = x * e$. ¶

Como você já deve saber, um elemento neutro de uma operação $*$, se existir, é único, o que permite o uso da expressão "o elemento neutro" da operação. Para aquecer os motores, convém destacar isso.

Lema 0.6.2. *Uma operação admite, no máximo, um único elemento neutro.*

Demonstração. Se $e, e' \in X$ são ambos elementos neutros da operação, então $e = e * e'$ por e' ser elemento neutro, enquanto $e * e' = e'$ por e ser elemento neutro. □

Definição 0.6.3. Diremos que $y \in X$ é um **$*$-inverso à direita** de x se valer $x * y = e$. Analogamente, y será dito um **$*$-inverso à esquerda** de x se ocorrer $y * x = e$. Se y for simultaneamente $*$-inverso à direita e à esquerda de x, diremos simplesmente que y é um **$*$-inverso**[‡] de x. Naturalmente, o *prefixo* "$*$-" será abandonado sempre que o contexto permitir. ¶

Exercício 0.80 ($\star$)**.** Sejam X um conjunto e $*$ uma operação em X. Mostre que se $*$ é associativa e tem elemento neutro, então cada $x \in X$ admite, <u>no máximo</u>, um $*$-inverso. ■

Para quem gosta de colecionar terminologias, as próximas são um prato cheio:

- um **semigrupo** é um conjunto não vazio munido de uma operação associativa;
- um **monóide** é um semigrupo com elemento neutro;
- um **grupo** é um monóide em que todo elemento tem um inverso;
- finalmente, um **grupo abeliano** é um grupo cuja operação é comutativa.

Observação 0.6.4. Tal qual se faz com ordens, quando se busca algum tipo de precisão exagerada, escreve-se algo como $(G, *, e)$ para indicar que $*$ é uma operação em G que tem e como elemento neutro, de modo que afirmações do tipo "$(G, *, e)$ *é um grupo abeliano*" abreviam a tediosa frase "$*\colon G \times G \to G$ é uma operação que faz de G um grupo comutativo cujo elemento neutro é e". △

Exemplo 0.6.5. As operações de adição e multiplicação em $\mathbb{N}$, $\mathbb{Z}$ e $\mathbb{Q}$ constituem exemplos elementares das estruturas definidas acima, embora isso não tenha sido sempre óbvio[††].

[†] Também chamado de **zero** ou **unidade** a depender do contexto.

[‡] Ou oposto, simétrico, etc. Tudo depende do contexto.

[††] Historicamente, grupos de *permutações* foram os primeiros a dar as caras, entre as últimas décadas do século XVIII e as primeiras décadas do século XIX. Depois do aparecimento dessas estruturas, *percebeu-se* que entidades cotidianas compostas por *números* também constituíam *modelos* de grupos [19].

- ✓ $(\mathbb{N}, +, 0)$ é um monóide comutativo, mas não é um grupo: não há, por exemplo, $n \in \mathbb{N}$ satisfazendo $n + 1 = 0$.
- ✓ $(\mathbb{Z}, +, 0)$ e $(\mathbb{Q}, +, 0)$ são grupos abelianos.
- ✓ $(\mathbb{Z} \setminus \{0\}, \cdot, 1)$ é um monóide comutativo, mas não é um grupo.
- ✓ $(\mathbb{Q} \setminus \{0\}, \cdot, 1)$ é um grupo abeliano.

Naturalmente, tais proposições só podem ser justificadas rigorosamente num cenário em que se tenha uma *construção* ou, pelo menos, uma definição explícita dos conjuntos e operações envolvidos. Se for de seu interesse verificar essas coisas, confira as construções de $\mathbb{Z}$ e $\mathbb{Q}$ na Subseção 0.4.1 §1. ▲

Praticamente todas as operações elementares consideradas ao longo do texto serão comutativas — a exceção mais marcante é a composição de funções, que em geral não será considerada do ponto de vista algébrico-estrutural. Apesar disso, no contexto que se aproxima, teremos que lidar com duas operações simultaneamente, que serão denotadas pelos símbolos $+$ e $\cdot$.

Em ambos os casos, as operações serão associativas, comutativas e dotadas de elemento neutro. No caso da operação $+$, o elemento neutro será denotado por 0, e o inverso de um elemento x, único em virtude do Lema 0.80, será denotado por $-x$ e chamado de **inverso aditivo** ou **simétrico**. Para a operação $\cdot$, o elemento neutro será denotado por 1, e o inverso de um elemento x, se existir, será denotado por x^{-1} ou $\frac{1}{x}$, caso em que x será dito **invertível**[†].

Proposição 0.6.6. *Sejam $(G, \cdot, 1)$ um monóide e $x \in G$. Para $n \in \mathbb{N}$, defina $x^n \in G$ recursivamente, fazendo $x^0 := 1$ e $x^{n+1} := x^n \cdot x$ para todo $n \in \mathbb{N}$. Se x admitir inverso, para cada $n \in \mathbb{N}$ com $n > 0$ defina ainda $x^{-n} := (x^{-1})^n$. Então, vale o seguinte:*

(i) $x^{m+n} = x^m \cdot x^n$ *para quaisquer* $m, n \in \mathbb{N}$*;*

(ii) $x^{mn} = (x^m)^n$ *para quaisquer* $m, n \in \mathbb{N}$*;*

(iii) se x admite inverso, então as identidades anteriores valem para $m, n \in \mathbb{Z}$.

Demonstração. O argumento é indutivo e depende das propriedades operatórias de $\mathbb{N}$ e $\mathbb{Z}$. Por exemplo: para $m \in \mathbb{N}$ fixado, tem-se:

- ✓ $x^{m+0} = x^m = x^m \cdot 1 = x^m \cdot x^0$;
- ✓ se $x^{m+n} = x^m \cdot x^n$ para algum $n \in \mathbb{N}$, então
$$x^{m+(n+1)} = x^{(m+n)+1} := x^{(m+n)} \cdot x = (x^m \cdot x^n) \cdot x = x^m \cdot (x^n \cdot x) := x^m \cdot x^{n+1},$$
donde a validade do primeiro item segue por indução.

Para o segundo item, o pulo do gato é notar que
$$x^{m(n+1)} = x^{mn+m} = x^{mn} \cdot x^m = (x^m)^n \cdot x^m := (x^m)^{n+1}.$$

A parte final segue, dentre outras coisas, de se observar que $x^{-n} = (x^n)^{-1}$. Os detalhes ficam por sua conta. □

[†]Como sempre, lembre-se: o verbo é "inverter" e não "inverser".

Exercício 0.81 ($^{\star}_{\star}$)**.** Complete a demonstração da proposição anterior. Dica: além do que já foi mencionado, pode ser útil observar que $x^{n+1} = x \cdot x^n$. ■

Será cada vez mais comum utilizar o mesmo símbolo com significados que variam de acordo com o contexto. Na proposição anterior, por exemplo, o símbolo "1" em "$(G, \cdot, 1)$" denota o elemento neutro da operação "$\cdot$" de G. Já em "$x^{n+1} := x^n \cdot x$", o expoente "$n+1$" indica o sucessor de n em $\mathbb{N}$. Esta é a chamada *notação multiplicativa*, em que o *histórico* das iterações é registrado no expoente.

Alternativamente, quando a operação de G é denotada com o símbolo "+", seu elemento neutro é indicado por "0" (caso exista), enquanto o *histórico das iterações* é registrado à esquerda, como explicitado no próximo corolário — esta é a chamada *notação aditiva*. Note que, a seguir, o símbolo "0" em "$0x := 0$" também assume *dupla função*.

Corolário 0.6.7. *Sejam $(G, +, 0)$ um grupo (abeliano) e $x \in G$. Para $n \in \mathbb{N}$, defina $nx \in G$ recursivamente, fazendo $0x := 0$ e $(n+1)x := nx + x$ para todo $n \in \mathbb{N}$. Além disso, defina $(-n)x := n(-x)$ para cada $n \in \mathbb{N}$. Então para quaisquer $m, n \in \mathbb{Z}$ valem as identidades $(m+n)x = mx + nx$ e $m(nx) = (mn)x$.*

§1 Anéis e corpos

Definição 0.6.8. Um **anel** consiste de um conjunto $A \neq \emptyset$ munido de duas operações, $+$ e $\cdot$, e elementos $0, 1 \in A$, onde

(i) $(A, +, 0)$ é um grupo abeliano, cuja operação $+$ é chamada de *adição*,

(ii) $(A, \cdot, 1)$ é um monóide comutativo, cuja operação $\cdot$ é chamada de *multiplicação*, e

(iii) as operações $+$ e $\cdot$ *comutam* entre si, i.e., para quaisquer $a, b, c \in A$ valem as identidades $a \cdot (b + c) = (a \cdot b) + (a \cdot c)$ e $(a + b) \cdot c = (a \cdot c) + (b \cdot c)$. ¶

A fim de não tornar o texto carregado demais, quase sempre iremos nos referir simplesmente *ao anel A*, deixando as operações subentendidas no contexto.

Observação 0.6.9. A rigor, os anéis definidos acima deveriam ser chamados de *anéis comutativos com unidade*, dado que existem contextos que não exigem a comutatividade da multiplicação e tampouco a existência de elemento neutro multiplicativo. △

Exemplo 0.6.10. Os conjuntos numéricos $\mathbb{Z}$ e $\mathbb{Q}$ são anéis quando munidos das operações usuais, enquanto $\mathbb{N}$ não é anel, posto que a adição não define uma estrutura de grupo. ▲

Muitas propriedades operatórias básicas do cenário aritmético ainda são válidas no contexto de anéis[†].

Proposição 0.6.11. *Sejam A um anel e $a \in A$ um elemento qualquer. Então valem as seguintes identidades:*

(i) $0 \cdot a = 0$; (ii) $-a = (-1) \cdot a$.

[†]E aqui cabe uma ressalva ~~que deveria ser óbvia~~. O comportamento que se observa nos números do "dia a dia" não decorre dos axiomas utilizados para *formalizá-los*, pelo contrário: os axiomas postulados para formalizar os números são "bons" justamente por reproduzirem os *comportamentos* observados.

Demonstração. Tais identidades relacionando as operações $+$ e $\cdot$ decorrem diretamente da distributividade exigida na definição de anel. De fato,

$$a\cdot 0 = a\cdot(0+0) = (a\cdot 0)+(a\cdot 0) \Rightarrow 0 = (a\cdot 0)+(-(a\cdot 0)) = (a\cdot 0)+(a\cdot 0)+(-(a\cdot 0)) = a\cdot 0,$$

ao passo que

$$a + ((-1)\cdot a) = (1+(-1))\cdot a = 0\cdot a = 0,$$

donde a igualdade $(-1)\cdot a = -a$ segue da unicidade do *oposto aditivo* de a. □

Observação 0.6.12. Em particular, como o inverso de $-a$ é a, segue que

$$(-1)\cdot(-a) = -(-a) = a,$$

identidade que será utilizada daqui em diante sem menções especiais. △

Ao longo do texto, usaremos simultaneamente a Proposição 0.6.6 e o Corolário 0.6.7: a primeira com respeito à multiplicação do anel, e o segundo com respeito à adição do anel. Assim, se A é um anel, $a \in A$ e $m, n \in \mathbb{N}$, então ma^n indica o que, intuitivamente, seria escrito como

$$\underbrace{\underbrace{a\cdot a\cdot\ldots\cdot a}_{n\text{ vezes}} + \underbrace{a\cdot a\cdot\ldots\cdot a}_{n\text{ vezes}} + \ldots + \underbrace{a\cdot a\cdot\ldots\cdot a}_{n\text{ vezes}}}_{m\text{ vezes}},$$

com uma interpretação análoga para $m \in \mathbb{Z}$ e, se a for invertível, para $n \in \mathbb{Z}$.

Observação 0.6.13. Daqui em diante, como de costume, "ab" também será usado para denotar "$a\cdot b$", o produto entre a e b num anel A. Como o tempo é curto, "$a-b$" abreviará a expressão "$a+(-b)$". △

Definição 0.6.14. Um anel A é chamado de **corpo** se $1 \neq 0$ e todo $a \in A\setminus\{0\}$ é invertível, i.e, se existir $b \in A\setminus\{0\}$ com $ab = 1$. ¶

Exemplo 0.6.15. O primeiro corpo com o qual nos deparamos *explicitamente* é $\mathbb{Q}$, o conjunto dos números racionais com suas operações usuais. A reta real $\mathbb{R}$, protagonista deste curso, também é um corpo. Estes não são os únicos corpos que existem: a Álgebra Comutativa é repleta de tais entidades. Porém, em contextos *introdutórios* de Análise, não costuma ser importante saber disso[†]. ▲

Proposição 0.6.16 (Cancelamento para multiplicação[‡])**.** *Se A é corpo e $\alpha, x, y \in A$ são tais que $\alpha x = \alpha y$ com $\alpha \neq 0_A$, então $x = y$.*

Demonstração. Basta multiplicar α^{-1} dos dois lados da igualdade $\alpha x = \alpha y$. □

Antes de encerrar, cabe uma ressalva acerca da divisibilidade por zero, cuja *polêmica* só se justifica pela falta de discussão adequada: não há lei universal ou *mandamento* vindo das colinas que proíba dividir por zero em absolutamente qualquer contexto — o ponto é que fazer isso em anéis é, na prática, inútil.

De fato, uma identidade do tipo $\frac{\alpha}{0} = \beta$ num anel A depende, implicitamente, de se admitir a existência de $z := \frac{1}{0} \in A$, i.e., um elemento $z \in A$ que satisfaz $0\cdot z = 1$. Ora, como também deve ocorrer $0\cdot z = 0$, resulta que $0 = 1$ e, consequentemente, $0 = x$ para todo $x \in A$, ou seja, $A = \{0\}$. Algebricamente isto não é crime: um conjunto dotado de um único elemento admite tanto uma adição quanto uma multiplicação que o tornam um anel (verifique?)[⋆]. O *problema* é que ao se modelar axiomaticamente os números naturais, inteiros, etc., *espera-se* que ocorra $0 \neq 1$ (vide o Lema 0.7.2).

[†] Não se engane: a Análise é intrinsecamente dependente da Álgebra, embora a recíproca seja falsa — exceto pelo *Teorema* (não tão) *Fundamental da Álgebra*, que *ainda* é importante em *algumas áreas*.

[‡] Ou "Lema da Academia": corpo é *domínio*. Piadoca que aprendi com o Prof. Eduardo Tengan.

0.6.1 Extras

§0 (Adiável) morfismos entre anéis

Definição 0.6.17. Dados anéis A e B, uma função $f\colon A \to B$ é chamada de **morfismo** *de anéis* (ou *de corpos* quando ambos forem corpos) se

(i) $f(1_A) = 1_B$,

(ii) $f(a + a') = f(a) + f(a')$, e

(iii) $f(aa') = f(a)f(a')$. ¶

Acima, por preciosismo, os elementos neutros multiplicativos de A e B foram distinguidos como 1_A e 1_B, respectivamente. Porém, ainda mais precisão poderia ser dada às notações: note que em "$f(a + a') = f(a) + f(a')$", por exemplo o símbolo "$+$" em "$a + a$" indica a adição do anel A, enquanto o mesmo símbolo em "$f(a) + f(a')$" indica a adição em B. Analogamente, a ausência de símbolos operacionais na última cláusula indica as multiplicações em A e B, respectivamente.

Em certo sentido, um morfismo de anéis $f\colon A \to B$ permite que A *manifeste* informações de natureza algébrica em B por meio de f. Por exemplo, se $a \in A$ é tal que $a^2 = 1$ em A, então o mesmo deve ocorrer com $f(a)$ em B, i.e., $f(a)^2 = 1_B$: de fato, deve-se ter

$$1_B = f(1_A) = f(a^2) = f(a)f(a) = f(a)^2.$$

No presente contexto, um tipo muito particular de morfismo de anéis será fundamental:

Proposição 0.6.18. *Para todo anel A existe um único morfismo de anéis $\zeta_A\colon \mathbb{Z} \to A$.*

Demonstração. Fazendo $(G, +, 0) := (A, +, 0_A)$ e $x := 1_A$ no Corolário 0.6.7, resulta que a função $\zeta_A\colon \mathbb{Z} \to A$, dada por $\zeta_A(n) := n \cdot 1_A$ e $\zeta(-n) := -(n \cdot 1_A)$ para qualquer $n \in \mathbb{N}$, é um morfismo de anéis. Fica por sua conta verificar, por indução, que qualquer *outro* morfismo $\psi\colon \mathbb{Z} \to A$ deve ser tal que $\psi = \zeta_A$. □

Exercício 0.82 ($^\star_\star$)**.** Complete a demonstração acima. Dica: deve-se ter $\psi(1) = 1_A = \zeta_A(1)$, $\psi(1 + 1) = 1_A + 1_A = \zeta_A(2)$, ... ■

Ao longo deste capítulo, fixados um anel A e um número inteiro $z \in \mathbb{Z}$, a notação z_A indicará a imagem de z pelo morfismo ζ_A da última proposição, a **interpretação** de z em A. Na prática,

$$z_A = \underbrace{1_A + \ldots + 1_A}_{z \text{ vezes}} \text{ se } z > 0, \text{ enquanto } z_A = \underbrace{-1_A + \ldots + (-1_A)}_{z \text{ vezes}} \text{ se } z < 0.$$

Exercício 0.83 ($^\star_\star$)**.** Mostre que para $z \in \mathbb{Z}$ e $a \in A$, o elemento $za \in A$ (como definido no Corolário 0.6.7) é tal que $za = z_A a$. Conclua que para quaisquer $a, b \in A$ e $m \in \mathbb{Z}$ deve-se ter $m(a + b) = ma + mb$. ■

Exemplo 0.6.19. Para $A := \mathbb{Q}$, o morfismo $\zeta_A\colon \mathbb{Z} \to A$ é, meramente, a inclusão. ▲

Exemplo 0.6.20 (Matrizes)**.** Se você tiver familiaridade com *matrizes*, note que se A denota o *anel das matrizes de ordem* $n \in \mathbb{N} \setminus \{0\}$ (e coeficientes *reais*, por exemplo), então $\zeta_A\colon \mathbb{Z} \to A$ associa cada inteiro $z \in \mathbb{Z}$ à matriz diagonal $(z\delta_{ij})$. ▲

Exercício 0.84 ($\star$)**.** Seja $f\colon A \to B$ um morfismo de anéis.

a) Mostre que $f(0_A) = 0_B$.

b) Mostre que $f(ma) = mf(a)$ para quaisquer $m \in \mathbb{Z}$ e $a \in A$.

c) Mostre que $f(-a) = -f(a)$ para qualquer $a \in A$.

d) Mostre que $f(a^m) = f(a)^m$ para quaisquer $m \in \mathbb{N}$ e $a \in A$. ■

Os últimos comentários, frequentemente úteis, seguem destacados nos próximos exercícios.

Exercício 0.85 ($\star$)**.** Dado um anel A, mostre que $\mathrm{Id}_A\colon A \to A$ é um morfismo de anéis.■

Exercício 0.86 ($\star$)**.** Mostre que a composição de morfismos de anéis é um morfismo de anéis. ■

Exercício 0.87 (Núcleo e injetividade, $\binom{\star}{\star}$)**.** Para um morfismo de anéis $f\colon A \to B$, chama-se de **núcleo** de f ao conjunto $\ker f \coloneqq \{a \in A : f(a) = 0_B\}$.

a) Mostre que f é injetora se, e somente se, seu núcleo é *trivial*, i.e., se $\ker f = \{0_A\}$.

b) Mostre que se A e B são corpos, então f é injetora. ■

§1 (Adiável) espaços vetoriais e transformações lineares

Se você já estudou Álgebra Linear ou, pelo menos, Geometria Analítica, já deve ter visto *espaços vetoriais*. Grosso modo, trata-se de um conjunto cujos elementos interpretamos como se fossem pontos/vetores *num espaço* em que algum sistema de coordenadas permite que façamos contas com esses pontos.

O exemplo clássico é $\mathbb{R}^n$, para $n \in \mathbb{N} \setminus \{0\}$, o conjunto das n-uplas de *números reais* — e Geometria Analítica consiste, essencialmente, em estudar os casos em que $n \leq 3$. Como tais n-uplas são da forma $(a_1, \ldots, a_n)$, com $a_1, \ldots, a_n \in \mathbb{R}$, é possível tanto somá-las umas com as outras quanto multiplicá-las entre si. Formalmente, definem-se operações

$$\begin{array}{rcl} (+)\colon \mathbb{R}^n \times \mathbb{R}^n & \longrightarrow & \mathbb{R}^n \\ \left((a_i)_{1\leq i\leq n}, (b_i)_{1\leq i\leq n}\right) & \longmapsto & (a_1 + b_1, \ldots, a_n + b_n) \end{array} \quad \text{e} \quad \begin{array}{rcl} (\cdot)\colon \mathbb{R}^n \times \mathbb{R}^n & \longrightarrow & \mathbb{R}^n \\ \left((a_i)_{1\leq i\leq n}, (b_i)_{1\leq i\leq n}\right) & \longmapsto & (a_1 \cdot b_1, \ldots, a_n \cdot b_n) \end{array}.$$

No entanto, embora as duas operações sejam razoáveis de um ponto de vista puramente algébrico, apenas a soma tem uma interpretação geométrica *óbvia*: a saber, translação. Ao fazer a multiplicação com a regra acima, verificam-se comportamentos inesperados incompatíveis com a multiplicação usual[†]. Apesar disso, recupera-se alguma intuição geométrica se, em vez de multiplicar vetores com a segunda regra, multiplicarmos apenas *escalares* por vetores, fazendo $\lambda \cdot (a_1, \ldots, a_n) \coloneqq (\lambda a_1, \ldots, \lambda a_n)$. Ao considerar $\mathbb{R}^n$ com a soma $(+)$ e com essa multiplicação por escalares, tem-se um dos principais exemplos de *espaço vetorial*.

[†]Como $u \cdot v = (0, \ldots, 0)$ para $u, v \neq (0, \ldots, 0)$. Há outros problemas de natureza geométrica que não convêm tratar aqui, mas você pode conferir em https://math.stackexchange.com/questions/185888.

Definição 0.6.21. Fixado um corpo K, dizemos que um grupo abeliano $(V, +, 0)$ é um K-**espaço vetorial** se existir uma função

$$\begin{aligned} K \times V &\mapsto V \\ (k, v) &\mapsto kv \end{aligned}$$

chamada de *multiplicação*, satisfazendo as seguintes condições:

(i) $1_K v = v$ para todo $v \in V$;

(ii) $(\alpha + \beta)v = (\alpha v) + (\beta v)$ para quaisquer $\alpha, \beta \in K$ e $v \in V$;

(iii) $k(u + v) = (ku) + (kv)$ para quaisquer $u, v \in V$ e $k \in K$. ¶

Em tal contexto, os elementos de V são chamados de *vetores*, enquanto os membros de K são xingados de *escalares*.

Após introduzirmos $\mathbb{R}$ como *um* corpo ordenado completo, as discussões acima servirão para entender $\mathbb{R}^n$ como um $\mathbb{R}$-espaço vetorial. Naturalmente, como $\mathbb{Q}$ é corpo, os mesmos argumentos já podem ser usados para justificar que $\mathbb{Q}^n$ é $\mathbb{Q}$-espaço vetorial. De modo geral, K^n é K-espaço vetorial, para qualquer $n \in \mathbb{N}$ e corpo K. Na verdade, pode-se extrapolar um pouco mais.

Exemplo 0.6.22. Se X é um conjunto e K é um corpo, então o conjunto K^X das funções da forma $X \to K$ tem uma estrutura natural de K-espaço vetorial com as operações herdadas de K em cada "coordenada". Mais precisamente, para $f, g \in K^X$ e $\lambda \in K$, definem-se $f + g, \lambda \cdot g \in K^X$ por

$$(f + g)(x) := f(x) + g(x) \quad \text{e} \quad (\lambda \cdot g)(x) := \lambda \cdot g(x)$$

para cada $x \in X$. Caso nunca tenha verificado que se trata de um espaço vetorial, faça isso agora $(\star)$. ▲

Espaços vetoriais aparecem com bastante frequência em cursos de Análise Real, embora geralmente sejam mantidos sob disfarce para turmas iniciantes. Aqui, os principais espaços vetoriais considerados serão o *plano cartesiano* $\mathbb{R}^2 := \mathbb{R} \times \mathbb{R}$ e o *espaço das funções contínuas de* $X \subseteq \mathbb{R}$ *em* $\mathbb{R}$, denotado por $\mathcal{C}(X)$ (ou $\mathcal{C}X$, a depender da grafia de X)[†], ambos intrinsecamente ligados a questões sobre a reta real. Para finalizar esta breve e bem intencionada introdução aos espaços vetoriais, convém apresentar as funções que fazem o papel de *morfismos*:

Definição 0.6.23. Uma função $f : X \to Y$ entre K-espaços vetoriais X e Y é dita K-**linear** (ou **transformação** K-**linear**[‡]) se for compatível com as operações de X e Y, i.e., se $f(\alpha u + v) = \alpha f(u) + f(v)$ para quaisquer $\alpha \in K$ e $u, v \in X$. ¶

Exercício 0.88 $(\substack{\star \\ \star})$**.** Mostre que os seguintes resultados para morfismos de anéis permanecem válidos para transformações lineares:

a) itens (a) e (c) do Exercício 0.84;

b) item (b) do Exercício 0.84, para $m \in K$;

c) os Exercícios 0.85 e 0.86;

d) item (a) do Exercício 0.87. ■

[†]Se X for o intervalo $(0, 1)$, por exemplo, é melhor escrever $\mathcal{C}(0, 1)$ do que $\mathcal{C}((0, 1))$, não acha?
[‡]E como de costume, o sufixo "K" é abandonado nas situações em que o corpo é claro pelo contexto.

0.7 Botando ordem na Álgebra

0.7.0 Essencial

§0 Corpos ordenados

Definição 0.7.0. Um corpo $\mathbb{K}$ munido de uma relação de ordem (estrita) total $<$ é chamado de **corpo ordenado** se $<$ for compatível com sua estrutura algébrica, i.e.,

(CO$_{\text{i}}$) $\forall a, b, c \in \mathbb{K} \quad a < b \Rightarrow a + c < b + c$,

(CO$_{\text{ii}}$) $\forall a, b \in \mathbb{K} \quad a > 0_{\mathbb{K}}$ e $b > 0_{\mathbb{K}} \Rightarrow ab > 0_{\mathbb{K}}$. ¶

Exercício 0.89 (Caracterização alternativa via cones ($\substack{\star\\\star}$))**.** Dado um corpo $\mathbb{K}$, mostre que $\mathbb{K}$ admite uma ordem $<$ que torna $(\mathbb{K}, <)$ um corpo ordenado se, e somente se, existir um subconjunto $P \subseteq \mathbb{K}$ com $x + y, xy \in P$ sempre que $x, y \in P$ e tal que, para qualquer $x \in \mathbb{K}$, ocorra um, e somente um, dos seguintes casos: $x = 0_{\mathbb{K}}$, $x \in P$ ou $-x \in P$. Dica: com a ordem em mãos, faça $P := \{x \in \mathbb{K} : x > 0_{\mathbb{K}}\}$ e note que $x < y$ se, e somente se, $y - x \in P$; com P em mãos, use o passo anterior para definir a ordem em $\mathbb{K}$ de modo adequado. ■

Exemplo 0.7.1. O corpo dos números racionais é um corpo ordenado com sua ordem usual. Caso não se lembre: basta definir $P := \left\{\frac{a}{b} \in \mathbb{Q} : ab > 0\right\}$, que satisfaz as condições do exercício anterior (verifique!)$^{\star}_{\star}$, de modo que

$$\frac{\alpha}{\beta} < \frac{\gamma}{\delta} \Leftrightarrow \frac{\gamma}{\delta} - \frac{\alpha}{\beta} = \frac{\beta\gamma - \alpha\delta}{\beta\delta} \in P \Leftrightarrow \beta\delta(\beta\gamma - \alpha\delta) > 0$$

para quaisquer $\alpha, \beta, \gamma, \delta \in \mathbb{Z}$ com $\beta, \delta \neq 0$. Em particular, assumindo $\beta, \delta > 0$ como de costume, resulta que

$$\frac{\alpha}{\beta} < \frac{\gamma}{\delta} \Leftrightarrow \alpha\delta < \beta\gamma,$$

e fica por sua conta testar alguns exemplos com números “de verdade”. ▲

Lema 0.7.2 (Fundamental)**.** *Se $\mathbb{K}$ é um corpo ordenado, então $0_{\mathbb{K}} < 1_{\mathbb{K}}$.*

Demonstração. O contrário daria $0_{\mathbb{K}} > 1_{\mathbb{K}}$, posto que a ordem é total e $\mathbb{K}$ é corpo. Logo, a condição (CO$_{\text{i}}$), com $c := -1_{\mathbb{K}}$, acarretaria $-1_{\mathbb{K}} > 0_{\mathbb{K}}$ e, consequentemente, teria-se $0_{\mathbb{K}} > 1_{\mathbb{K}} = (-1_{\mathbb{K}})(-1_{\mathbb{K}}) > 0_{\mathbb{K}}$ em virtude da condição (CO$_{\text{ii}}$), uma contradição. □

Para o que se discutirá a seguir, é recomendável saber o que é um morfismo de anéis[†] (cf. Subseção 0.6.1 §0).

Teorema 0.7.3. *Se $\mathbb{K}$ é corpo ordenado, então existe um único morfismo injetor de anéis $\rho\colon \mathbb{Q} \to \mathbb{K}$.*

[†]Se preferir ignorar o conselho: pense que $\rho\colon \mathbb{Q} \to \mathbb{K}$ é uma função que permite interpretar cada número racional q como um elemento de $\rho(q) \in \mathbb{K}$, de forma a respeitar as operações de $\mathbb{Q}$ em $\mathbb{K}$: assim, $\rho(q + q') = \rho(q) + \rho(q')$, etc.

Demonstração. Observe que o lema anterior garante que $n_{\mathbb{K}} > 0_{\mathbb{K}}$ para todo $n \in \mathbb{N} \setminus \{0\}$: como $\mathbb{K}$ é corpo ordenado, tem-se $1_{\mathbb{K}} > 0_{\mathbb{K}}$ (pelo lema anterior!); supondo $n_{\mathbb{K}} > 0_{\mathbb{K}}$ para algum $n \geq 1$, tem-se $(n+1)_{\mathbb{K}} := n_{\mathbb{K}} + 1_{\mathbb{K}} > 0_{\mathbb{K}}$, onde a última desigualdade decorre da condição $(\mathrm{CO_i})$.

Logo, $z_{\mathbb{K}} \neq 0_{\mathbb{K}}$ para todo $z \in \mathbb{Z} \setminus \{0\}$. Agora, se $\sigma \colon \mathbb{Q} \to \mathbb{K}$ for *um* morfismo de anéis, então $\sigma|_{\mathbb{Z}} \colon \mathbb{Z} \to \mathbb{K}$ também é um morfismo de anéis, donde a Proposição 0.6.18 obriga que se tenha $\sigma(z) = z_{\mathbb{K}}$ para todo $z \in \mathbb{Z}$. Por outro lado, da identidade

$$\sigma\left(\frac{a}{1} \cdot \frac{1}{a}\right) = \sigma(a) \cdot \sigma\left(\frac{1}{a}\right),$$

não é difícil concluir que $\sigma\left(\frac{1}{a}\right) = \frac{1}{\sigma(a)}$ para todo $a \in \mathbb{Z} \setminus \{0\}$ e, por conseguinte,

$$\sigma\left(\frac{m}{n}\right) = \frac{m_{\mathbb{K}}}{n_{\mathbb{K}}} := \left(\frac{m}{n}\right)_{\mathbb{K}}$$

deve valer para quaisquer $m, n \in \mathbb{Z}$ com $n \neq 0$. Portanto, tudo se resume a observar que a regra acima define, de fato, um morfismo de anéis da forma $\mathbb{Q} \to \mathbb{K}$, que será injetor por ter *núcleo trivial* (cf. Exercício 0.87). □

Exercício 0.90 ($^{\star}_{\star\star}$)**.** Complete os detalhes. ■

Moralmente, o elemento $q_{\mathbb{K}} \in \mathbb{K}$ descrito recursivamente na demonstração anterior *representa* ou *interpreta* o número racional $q \in \mathbb{Q}$ no corpo ordenado $\mathbb{K}$. Para quem tem aversão a Álgebra ~~melhore!~~, pode-se pensar em $\mathbb{K}$ como um *ambiente virtual* no qual é possível *implementar* os números racionais: nesse sentido, a demonstração apenas descreve e justifica o algoritmo de implementação. Se ainda parecer abstrato demais: todo corpo ordenado contém uma *cópia* de $\mathbb{Q}$!

Definição 0.7.4. $\mathbb{Q}_{\mathbb{K}} := \{q_{\mathbb{K}} : q \in \mathbb{Q}\}$. ¶

Muitas propriedades operatórias corriqueiras dos números *reais* que introduziremos em breve são, na verdade, comuns a qualquer corpo ordenado. As mais *úteis* seguem listadas na próxima

Proposição 0.7.5. *Sejam $\mathbb{K}$ um corpo ordenado e $x, y, z \in \mathbb{K}$ elementos quaisquer. Então:*

(i) *$x > 0_{\mathbb{K}}$ se, e somente se, $-x < 0_{\mathbb{K}}$;*

(ii) *se $x > 0_{\mathbb{K}}$ e $y < z$, então $xy < xz$;*

(iii) *se $x < 0_{\mathbb{K}}$ e $y < z$, então $xy > xz$;*

(iv) *$x > 0_{\mathbb{K}}$ se, e somente se, $x^{-1} > 0_{\mathbb{K}}$;*

(v) *se $x \neq 0_{\mathbb{K}}$, então $x^2 > 0_{\mathbb{K}}$;*

(vi) *se $0_{\mathbb{K}} < x < y$, então $0_{\mathbb{K}} < y^{-1} < x^{-1}$.*

Demonstração. Se $x > 0_{\mathbb{K}}$, então $0_{\mathbb{K}} = x - x > 0_{\mathbb{K}} - x = -x$ por conta da condição $(\mathrm{CO_i})$, i.e., $-x < 0_{\mathbb{K}}$. Analogamente mostra-se a recíproca. Os itens (ii) e (iii) seguem de $(\mathrm{CO_{ii}})$ (e do item anterior!) ao se observar que $y < z$ equivale a $z - y > 0_{\mathbb{K}}$. Para o item (iv), note que se $x > 0_{\mathbb{K}}$ e $x^{-1} < 0_{\mathbb{K}}$, então pelo item anterior resultaria $-1_{\mathbb{K}} = (-x)x^{-1} > -x0_{\mathbb{K}} = 0_{\mathbb{K}}$, contrariando o fato de que $0_{\mathbb{K}} < 1_{\mathbb{K}}$; a recíproca segue da identidade $(x^{-1})^{-1} = x$. O quinto item decorre da condição $(\mathrm{CO_{ii}})$ para $x > 0_{\mathbb{K}}$; para $x < 0_{\mathbb{K}}$, o mesmo raciocínio dá $(-x)^2 > 0_{\mathbb{K}}$, enquanto $(-x)^2 = (-1_{\mathbb{K}})^2 x^2 = x^2$. O último é o mais divertido e, por tal razão, ficará por sua conta. □

Exercício 0.91 ($^{\star}_{\star}$)**.** Complete a demonstração acima. Dica: para o item (vi), note que $x^{-1}y^{-1} > 0_{\mathbb{K}}$, daí use algum dos itens anteriores para concluir. ■

Exercício 0.92 ($\star$)**.** Nas condições anteriores, mostre que se $a < c$ e $b < d$ para certos $a, b, c, d \in \mathbb{K}$, então $a + b < c + d$. ■

§1 Valor absoluto e a desigualdade triangular

Definição 0.7.6. Seja $\mathbb{K}$ um corpo ordenado. O **valor absoluto** em $\mathbb{K}$ é a função $|\cdot|_{\mathbb{K}} \colon \mathbb{K} \to \mathbb{K}$ que associa cada $x \in \mathbb{K}$ ao elemento $|x|_{\mathbb{K}} := \max\{x, -x\}$. ¶

O valor absoluto constitui uma *maneira uniforme* de "medir" elementos de $\mathbb{K}$ por meio da *comparação* com os habitantes de seu **cone positivo**, i.e., do subconjunto $\mathbb{K}_{\geq 0} := \{x \in \mathbb{K} : x \geq 0_{\mathbb{K}}\}$, posto que $|x|_{\mathbb{K}} \in \mathbb{K}_{\geq 0}$ para todo $x \in \mathbb{K}$. Essa "*maneira uniforme*" se refere, entre outras coisas, ao fato de que o valor absoluto é compatível tanto com a estrutura algébrica quanto com a ordem de $\mathbb{K}$, no seguinte sentido.

Proposição 0.7.7. *Sejam $\mathbb{K}$ um corpo ordenado e $x, y \in \mathbb{K}$. Então:*

(i) $|x|_{\mathbb{K}} \geq 0_{\mathbb{K}}$;

(ii) $|x|_{\mathbb{K}} = 0_{\mathbb{K}}$ *se, e somente se,* $x = 0_{\mathbb{K}}$;

(iii) $|xy|_{\mathbb{K}} = |x|_{\mathbb{K}}|y|_{\mathbb{K}}$;

(iv) $|x + y|_{\mathbb{K}} \leq |x|_{\mathbb{K}} + |y|_{\mathbb{K}}$.

Demonstração. Os três primeiros itens ficam por sua conta $\binom{\star}{\star}$. Como a desigualdade (iv) acima, chamada de **desigualdade triangular**, será extremamente recorrente no texto, convém demonstrá-la aqui: como $-x \leq |x|_{\mathbb{K}}$ e $-y \leq |y|_{\mathbb{K}}$, tem-se $-(x+y) \leq |x|_{\mathbb{K}} + |y|_{\mathbb{K}}$; como também ocorre $x + y \leq |x|_{\mathbb{K}} + |y|_{\mathbb{K}}$, conclui-se que

$$|x + y|_{\mathbb{K}} = \max\{x + y, -(x+y)\} \leq |x|_{\mathbb{K}} + |y|_{\mathbb{K}}.$$ □

Exercício 0.93 $\binom{\star}{\star}$**.** Para $x, y \in \mathbb{K}$, com $\mathbb{K}$ corpo ordenado, mostre que são equivalentes:

(i) $-y \leq x \leq y$;

(ii) $x \leq y$ e $-x \leq y$;

(iii) $|x|_{\mathbb{K}} \leq y$.

Conclua que $|x - y|_{\mathbb{K}} \leq z$ se, e somente se, $y - z \leq x \leq y + z$. ■

Observação 0.7.8. O exercício acima permanece válido ao trocarmos "$\leq$" por "$<$" (verifique!)*. △

Teorema 0.7.9 (Truque fundamental da Análise)**.** *Para α e β elementos de um corpo ordenado $\mathbb{K}$, são equivalentes:*

(i) $\alpha = \beta$;

(ii) $|\alpha - \beta|_{\mathbb{K}} < \varepsilon$ *para todo* $\varepsilon \in \mathbb{K}$ *com* $\varepsilon > 0_{\mathbb{K}}$.

Demonstração. A direção (i) $\Rightarrow$ (ii) é clara. Para a recíproca, se ocorresse $\alpha \neq \beta$, teríamos $\alpha < \beta$ ou $\beta < \alpha$ e, consequentemente, $\beta - \alpha > 0_{\mathbb{K}}$ ou $\alpha - \beta > 0_{\mathbb{K}}$ (respectivamente), de modo que

$$|\alpha - \beta|_{\mathbb{K}} = \max\{\alpha - \beta, \beta - \alpha\} := \varepsilon > 0_{\mathbb{K}},$$

provando o resultado pela contrapositiva. □

Exercício 0.94 $\binom{\star}{\star}$**.** Considere $\mathbb{K}$ um corpo ordenado e $x, y \in \mathbb{K}$ quaisquer.

a) Mostre que se $x > y > 0_{\mathbb{K}}$, então $x^2 > y^2$.

b) Mostre que se $x < y < 0_{\mathbb{K}}$, então $x^2 > y^2$.

c) Mostre que se $x > 1_{\mathbb{K}}$, então $x^2 > x$ e, se $0_{\mathbb{K}} < x < 1_{\mathbb{K}}$, então $x^2 < x$.

d) Mostre que se $x^2 + y^2 = 0_{\mathbb{K}}$, então $x = y = 0_{\mathbb{K}}$. ■

0.7.1 Extras

§0 Espaços vetoriais ordenados

Corpos não são as únicas estruturas algébricas que podem aparecer acompanhadas de uma ordem compatível. Um exemplo que será importante é o seguinte: fixados um corpo ordenado $\mathbb{K}$ e um conjunto X, diremos que uma função $f: X \to \mathbb{K}$ é **limitada** se existir $M \in \mathbb{K}$ com $M > 0_{\mathbb{K}}$ tal que $|f(x)|_{\mathbb{K}} \leq M$ para todo $x \in X$. Ao considerar $\mathcal{B}(X, \mathbb{K})$, a coleção de todas as funções limitadas de X em $\mathbb{K}$, obtém-se um legítimo $\mathbb{K}$-espaço vetorial[†] com as operações usuais para espaços de funções (cf. Exemplo 0.6.22):

✓ se $f, g \in \mathcal{B}(X, \mathbb{K})$, então $f + g \in \mathcal{B}(X, \mathbb{K})$ pois, para escalares $M, N \in \mathbb{K}_{>0}$ tais que $|f(x)|_{\mathbb{K}} \leq M$ e $|g(x)|_{\mathbb{K}} \leq N$ para todo $x \in X$, tem-se

$$|f(x) + g(x)|_{\mathbb{K}} \leq |f(x)|_{\mathbb{K}} + |g(x)|_{\mathbb{K}} \leq M + N;$$

✓ a função constante nula $\underline{0}: X \to \mathbb{K}$, que faz $\underline{0}(x) := 0_{\mathbb{K}}$ para todo x, é *claramente* limitada;

✓ se $f \in \mathcal{B}(X, \mathbb{K})$ e $\lambda \in \mathbb{K}$, então para $M \in \mathbb{K}_{>0}$ satisfazendo $|f(x)|_{\mathbb{K}} \leq M$ para todo x se verifica

$$|\lambda f(x)|_{\mathbb{K}} = |\lambda|_{\mathbb{K}} |f(x)|_{\mathbb{K}} \leq |\lambda|_{\mathbb{K}} M.$$

A rigor, os pontos acima apenas mostram que obtemos operações legítimas em $\mathcal{B}(X, \mathbb{K})$ ao restringir as operações usuais de $\mathbb{K}^X$ a funções limitadas. No entanto, como as condições operatórias necessárias para elevar $\mathcal{B}(X, \mathbb{K})$ ao patamar de espaço vetorial já são satisfeitas em $\mathbb{K}^X$, segue que $\mathcal{B}(X, \mathbb{K})$ é, de fato, um espaço vetorial[‡]. Mas não só isso: para $f, g \in \mathcal{B}(X, \mathbb{K})$, podemos declarar

$$f \leq g \Leftrightarrow f(x) \leq g(x) \quad \text{para todo } x \in X,$$

que se revela uma ordem (parcial) em $\mathcal{B}(X, \mathbb{K})$ com as seguintes propriedades:

(i) $f \leq g \Rightarrow f + h \leq g + h$ para quaisquer $f, g, h \in \mathcal{B}(X, \mathbb{K})$;

(ii) $f \leq g \Rightarrow rf \leq rg$ para quaisquer $f, g \in \mathcal{B}(X, \mathbb{K})$ e $r \in \mathbb{K}_{\geq 0}$.

Um $\mathbb{K}$-espaço vetorial V com uma ordem parcial $\leq$ satisfazendo condições análogas costuma ser chamado de **espaço vetorial ordenado**. É claro que assim como $\mathcal{B}(X, \mathbb{K})$ é espaço vetorial ordenado, o próprio $\mathbb{K}^X$ também é: o uso de $\mathcal{B}(X, \mathbb{K})$ como exemplo inicial foi apenas uma desculpa para mostrar valores absolutos em ação.

Exercício 0.95 ($^{\star}_{\star\star}$)**.** Por que a ordem em $\mathcal{B}(X, \mathbb{K})$ não é total? ■

Observação 0.7.10 (Cuidado com a desigualdade estrita)**.** Por definição, escrever "$f < g$" para funções f e g em $\mathcal{B}(X, \mathbb{K})$ (ou em $\mathbb{K}^X$) abrevia "$f \leq g$ e $f \neq g$". Portanto, trata-se uma afirmação que não é equivalente a "$f(x) < g(x)$ para todo $x \in X$". Reflita. △

[†] No futuro, trataremos apenas do caso $\mathbb{K} := \mathbb{R}$. Assim, você pode assumir que $\mathbb{K}$ é $\mathbb{R}$ se preferir.
[‡] A terminologia correta é "$\mathcal{B}(X, \mathbb{K})$ é subespaço vetorial de $\mathbb{K}^X$".

§1 Corpos não ordenáveis

Exercício 0.96 $\binom{\star}{\star}$**.** Mostre que se K é um corpo finito, então não existe ordem total $<$ sobre K segundo a qual $(K,<)$ seja um corpo ordenado. ■

O exercício acima mostra que existem corpos nos quais é *impossível* definir uma relação de ordem compatível com suas operações. Outro caso típico é $\mathbb{C}$, o **corpo dos números complexos**. Trata-se essencialmente do plano $\mathbb{R}^2$ com uma *multiplicação entre vetores* que faz de $\mathbb{R}^2$ um corpo. Tipicamente, escreve-se $a+bi \in \mathbb{C}$ em vez de $(a,b) \in \mathbb{R}^2$, de modo que o produto entre dois elementos de $\mathbb{C}$ é feito de tal forma a valer $i^2 = -1$. Explicitamente:

$$(a+bi)\cdot(c+di) := (ac-bd)+(ad+bc)i.$$

O problema é que se $\mathbb{C}$ admitisse uma ordem compatível com suas operações, então deveria ocorrer $i^2 > 0$. No entanto, $i^2 = -1$, e a desigualdade $-1 < 0$ vale em qualquer corpo ordenado.

0.8 Supremos e ínfimos

0.8.0 Essencial

§0 Definição e exemplos

Já temos quase todo o ferramental necessário para *entender* definição da reta real que será apresentada: $\mathbb{R}$ será definido como um corpo ordenado *completo*. É para discutir completude que precisamos voltar alguns passos e introduzir *supremos e ínfimos*.

Definição 0.8.0. Fixada uma *ordem* $(\mathbb{P},\leq)$, um subconjunto A de $\mathbb{P}$ e um elemento $p \in \mathbb{P}$, diremos que p é um **limitante superior** (ou **majorante**)[†] de A se $x \leq p$ ocorrer para todo $x \in A$. Adicionalmente, dizemos que p é o **supremo** de A se p é *o menor limitante superior* de A. Notação: $\sup A$ ou $\sup_{a\in A} a$. ¶

Dualizando a definição de supremo (cf. Subseção 0.2.1 §1), chega-se à noção de *ínfimo*.

Definição 0.8.1. Fixada uma *ordem* $(\mathbb{P},\leq)$, um subconjunto A de $\mathbb{P}$ e um elemento $p \in \mathbb{P}$, diremos que p é um **limitante inferior** (ou **minorante**)[‡] de A se $p \leq x$ ocorrer para todo $x \in A$. Adicionalmente, dizemos que p é o **ínfimo** de A se p é *o maior limitante inferior* de A. Notação: $\inf A$ ou $\inf_{a\in A} a$. ¶

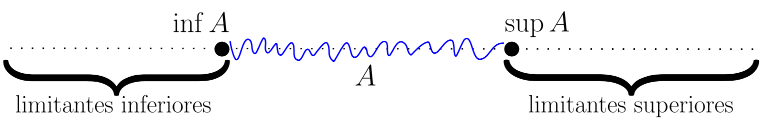


Figura 0.8: O supremo de um conjunto, quando existe, é o *melhor* limitante superior do conjunto. Analogamente, o ínfimo, se existir, é o *melhor* limitante inferior do conjunto.

Exemplo 0.8.2. Máximos são supremos, enquanto mínimos são ínfimos. De fato, se $A \subseteq \mathbb{P}$ tem máximo α, então:

[†]Ou ainda **cota superior**.
[‡]Ou ainda **cota inferior**.

✓ α é limitante superior de A por definição;

✓ se $\beta \in \mathbb{P}$ é algum outro limitante superior de A, então $a \leq \beta$ vale para todo $a \in A$ e, em particular, vale para α já que $\alpha \in A$.

O caso de mínimos é análogo. ▲

Exercício 0.97 ($\star$)**.** Para uma ordem $(\mathbb{P}, \leq)$ e um subconjunto $A \subseteq \mathbb{P}$, suponha que exista $\alpha := \min A$. Mostre que $\alpha = \inf A$. Dica: você pode imitar o argumento anterior ou apelar diretamente para "dualidade" (cf. Subseção 0.2.1 §1). ■

Exemplo 0.8.3 (A recíproca é falsa!)**.** Em $\mathbb{Q}$, o subconjunto $A := \{q \in \mathbb{Q} : q > 0\}$ não tem menor elemento: dado qualquer $q \in A$, tem-se $\frac{q}{2} \in A$ com $\frac{q}{2} < q$ (por quê?!)*, mostrando que nenhum dos elementos de A pode tomar para si o papel de ser *o menor*. O "problema", como você já deve ter percebido, é a ausência de 0 em A: se ocorresse $0 \in A$, então 0 seria, trivialmente, o menor elemento de A. Apesar disso, 0 é o ínfimo de A. ▲

Formalmente, a afirmação final do exemplo anterior pode se justificar com as seguintes observações:

✓ 0 é limitante inferior de A pela definição de A;

✓ se $s \in \mathbb{Q}$ e $s \leq q$ para todo $q \in A$, i.e., se s é limitante inferior de A, então $s \leq 0$, posto que *o contrário levaria a concluir que* $s > 0$, donde seguiria $\frac{s}{2} \in A$ com $\frac{s}{2} < s$, contrariando a suposição de s limitar A inferiormente.

Note que só foi possível concluir "$s > 0$" pois a totalidade da ordem impõe a ocorrência de "$s < 0$", "$s = 0$" ou "$s > 0$", de modo que a negação das duas primeiras forçou a validade da última. Noutras palavras, mostrou-se que nenhum $s > 0$ pode limitar A inferiormente e, portanto, os limitantes inferiores de A devem estar *abaixo* de 0. Mais geralmente, vale o seguinte:

Teorema 0.8.4. *Fixadas uma ordem* <u>*total*</u> $(\mathbb{T}, \leq)$, *um subconjunto* $A \subseteq \mathbb{T}$ *e* $\alpha \in \mathbb{T}$ *um limitante inferior de* A, *são equivalentes:*

(i) $\alpha = \inf A$;

(ii) *para todo* $\beta \in \mathbb{T}$, *se ocorrer* $\beta > \alpha$, *então existe* $a \in A$ *com* $a < \beta$.

Demonstração. Se vale (i), então $\alpha = \max\{l \in \mathbb{T} : l \text{ é limitante inferior de } A\}$, donde segue que se $\beta > \alpha$, então β não pode ser limitante inferior de A, i.e., *precisa* existir $a \in A$ com $\beta \not\leq a$, donde a tricotomia acarreta $a < \beta$, como desejado[†]. Reciprocamente, se vale (ii), então nenhum $\beta > \alpha$ limita A inferiormente ou, equivalentemente (graças à tricotomia), todo β limitante inferior de A satisfaz $\beta \leq \alpha$, donde o restante segue por α ser limitante inferior de A (por hipótese). □

Exercício 0.98 ($^{\star}_{\star}$)**.** Dualize o teorema anterior, i.e., enuncie (e demonstre) a versão para supremos. Dica: explicitamente, "se α é limitante superior de A, então $\alpha = \sup A$ se, e somente se, para todo $\beta < \alpha$ existir $a \in A$ com $\beta < a$". ■

[†] Lembre-se: $\beta \leq a$ se, e somente se, $\beta < a$ ou $\beta = a$.

§1 Supremos e ínfimos em corpos ordenados

Feita esta primeira apresentação, vamos nos ater ao caso em que os supremos e ínfimos são tomados em corpos ordenados. Para aquecer os motores:

Exercício 0.99 ($\star$)**.** Num corpo ordenado $\mathbb{K}$, mostre que $0_{\mathbb{K}} = \inf\{x \in \mathbb{K} : x > 0_{\mathbb{K}}\}$. ■

É bem provável que sua solução para o exercício prove, na verdade, a identidade

$$k = \inf\{x \in \mathbb{K} : x > k\}$$

para qualquer $k \in \mathbb{K}$, o que está absolutamente correto. Isto não foi coincidência.

Definição 0.8.5. Para subconjuntos $A, B \subseteq \mathbb{K}$ e $x \in \mathbb{K}$, definimos:

(i) $A + B \coloneqq \{a + b : a \in A \text{ e } b \in B\}$ e $A + x \coloneqq \{a + x : a \in A\}$;

(ii) $AB \coloneqq \{ab : a \in A \text{ e } b \in B\}$ e $xA \coloneqq \{xa : a \in A\}$;

(iii) $-A \coloneqq \{-a : a \in A\}$. ¶

Exercício 0.100 ($\star$)**.** Pratique!

a) Mostre que $A + B = B + A$, $A \cdot B = B \cdot A$, $A + \emptyset = \emptyset$ e $A \cdot \emptyset = \emptyset$.

b) Mostre que $\{x \in \mathbb{K} : x > 0_{\mathbb{K}}\} + k = \{x \in \mathbb{K} : x > k\}$ para qualquer $k \in \mathbb{K}$.

c) Mostre que $0_{\mathbb{K}}A = \{0_{\mathbb{K}}\}$ e $xyA = x(yA)$ para quaisquer $x, y \in \mathbb{K}$.

d) Para $A \coloneqq \{x \in \mathbb{K} : |x|_{\mathbb{K}} < 1_{\mathbb{K}}\}$, mostre que $-A = A$.

e) Mostre que se $A, B \subseteq \mathbb{K}_{>0}$ (i.e., se $x \in A$ ou $x \in B$, então $x > 0_{\mathbb{K}}$), então vale a inclusão $AB \subseteq \mathbb{K}_{>0}$. ■

Teorema 0.8.6. *Sejam $A, B \subseteq \mathbb{K}$ subconjuntos não vazios e $r \in \mathbb{K}$. Supondo que todos os ínfimos e supremos abaixo existam, valem as seguintes afirmações:*

(i) *se $A \subseteq B$, então $\inf A \geq \inf B$ e $\sup A \leq \sup B$;*

(ii) *se $r \geq 0$, então $\inf(rA) = r \inf A$ e $\sup(rA) = r \sup A$;*

(iii) *se $r \leq 0$, então $\inf(rA) = r \sup A$ e $\sup(rA) = r \inf A$;*

(iv) *se $x \geq 0$ para todo $x \in A \cup B$, então $\inf(AB) = \inf A \inf B$ e $\sup(AB) = \sup A \sup B$;*

(v) *$\inf(A + B) = \inf A + \inf B$ e $\sup(A + B) = \sup A + \sup B$.*

Demonstração. Como diria Jack...

(i) Note que se l é limitante inferior de B, então l também é limitante inferior de A (por quê?)*. Logo, se $\beta \coloneqq \inf B$, então $\beta \in \{L' : L' \text{ é limitante inferior de } A\} \coloneqq C$ e, por valer $\inf A \coloneqq \max C$, segue que $\beta \leq \inf A$.

(ii) Note que o resultado é automático para $r \coloneqq 0_{\mathbb{K}}$. Com $r > 0_{\mathbb{K}}$, e chamando $\alpha \coloneqq \sup A$, temos $x \leq \alpha$ para todo $x \in A$ (pois α limita A superiormente). Logo, se $y \in rA$, então $y = rx$ para algum $x \in A$, acarretando em $y = rx \leq r\alpha$, donde a arbitrariedade de y mostra que $r\alpha$ limita rA superiormente. Logo, $\sup(rA) \leq r\alpha$. Por outro lado, com $\beta \coloneqq \sup(rA)$ e tomando $x \in A$ qualquer, obtemos $rx \leq \beta$ e, por conseguinte, $x \leq \frac{\beta}{r}$, donde segue que $\sup A \leq \frac{\beta}{r}$. Logo, $r \sup A \leq \beta$.

(iii) Nada precisa ser feito para $r := 0_{\mathbb{K}}$. Supondo $r < 0_{\mathbb{K}}$ e fazendo $\delta := \sup A$, mostra-se (como no item anterior) que $r\delta$ é limitante inferior de rA e, por isso, $r\delta \leq \inf(rA)$. Analogamente, com $\varepsilon := \inf(rA)$, mostra-se que $\frac{\varepsilon}{r}$ é limitante superior de A, acarretando $\sup A \leq \frac{\varepsilon}{r}$ e, consequentemente, $r \sup A \geq \varepsilon$.

(iv) Primeiro, observe que sob as condições dadas, $AB = \{0_{\mathbb{K}}\}$ se, e somente se, $A = \{0_{\mathbb{K}}\}$ ou $\mathbb{B} = \{0_{\mathbb{K}}\}$, casos em que as identidades são triviais. Assim, podemos assumir que ambos A e B contêm elementos maiores do que $0_{\mathbb{K}}$. Agora, chamando $\alpha := \sup A$, $\beta := \sup B$ e $\gamma := \sup AB$, mostraremos que $\gamma \leq \alpha\beta$ e $\alpha\beta \leq \gamma$. A primeira desigualdade segue pois $x \leq \alpha$ e $y \leq \beta$ para quaisquer $x \in A$ e $y \in B$, e daí $xy \leq \alpha\beta$ em virtude da hipótese sobre os *sinais*. Para a segunda desigualdade:

- ✓ fixado $y \in B$ com $y > 0_{\mathbb{K}}$, temos $xy \leq \gamma$ para todo $x \in A$, o que assegura $x \leq \frac{\gamma}{y}$ e, consequentemente, $\alpha \leq \frac{\gamma}{y}$;
- ✓ dado que $0_{\mathbb{K}} < \alpha$ (por quê?!)*, resulta $y \leq \frac{\gamma}{\alpha}$ e, como isto vale para qualquer $y' \in B$ com $y' > 0_{\mathbb{K}}$, conclui-se que $\beta \leq \frac{\gamma}{\alpha}$ e, portanto, $\alpha\beta \leq \gamma$.

(v) Novamente, façamos $\alpha := \inf A$, $\beta := \inf B$ e $\gamma := \inf(A + B)$. Como $\alpha \leq x$ e $\beta \leq y$ para quaisquer $x \in A$ e $y \in B$, resulta $\alpha + \beta \leq x + y$ e, consequentemente, $\alpha + \beta \leq \gamma$. Para a desigualdade restante, observe que $\gamma - y \leq x$ para quaisquer $x \in A$ e $y \in B$, o que resulta em $\gamma - y \leq \alpha$, acarretando $\gamma - \alpha \leq y$ e, consequentemente, $\gamma - \alpha \leq \beta$. □

Exercício 0.101 $(^{\star}_{\star\star})$**.** Complete a demonstração do teorema anterior. ■

A demonstração do teorema anterior foi simplificada por uma hipótese preguiçosa: a suposição de que os supremos e ínfimos considerados *sempre* existem. Sem ela, os enunciados ficariam um pouco mais complicados. Por exemplo, no caso de (ii), para o supremo, seria preferível escrever "se $\sup A$ existe, então rA tem supremo e $\sup(rA) = r \sup A$ para qualquer $r \geq 0$": neste caso, seria necessário observar que rA é limitado superiormente para daí provar que $r \sup A$ é o menor limitante superior de rA. Contudo, após reler com atenção a demonstração, você perceberá que, no fundo, foi isso o que provamos — então, na prática, nada se perdeu[†]. Tais ressalvas se aplicam também a (iii), (iv) e (v). O item (i), como veremos, é mais delicado, e só se resolve com a hipótese de *completude* (cf. Exercício 0.139).

0.8.1 Extras

§0 Supremos e ínfimos em ordens

Supremos e ínfimos não são exclusividade de ordens totais. Porém, a vida sem tricotomia é um pouco menos óbvia: sem ela, o Teorema 0.8.4 não se aplica, por exemplo. Ainda assim, tais animais estão em todo lugar.

Exemplo 0.8.7. Dado um conjunto X e subconjuntos $A, B \subseteq X$, existem $\sup\{A, B\}$ e $\inf\{A, B\}$ em $(\wp(X), \subseteq)$? Se sim, quem são? Vejamos:

[†]Explicitamente: se $\sup A$ existe, então $rx \leq r \sup A$ para todo $x \in A$, mostrando que $r \sup A$ limita rA superiormente; agora, se $rx \leq \mu$ para todo $x \in A$, então $x \leq \frac{\mu}{r}$, donde segue que $\sup A \leq \frac{\mu}{r}$ e, por conseguinte, $r \sup A \leq \mu$, mostrando que $r \sup A$ é o menor limitante superior de rA. Faça adaptações análogas para os itens (iii), (iv) e (v)! $(^{\star}_{\star\star})$.

(i) se existir, $\sup\{A, B\}$ deve limitar superiormente o conjunto $\{A, B\}$, acarretando $A, B \subseteq \sup\{A, B\}$ e, além disso, se C for um subconjunto de X com $A, B \subseteq C$, também deverá ocorrer $\sup\{A, B\} \subseteq C$ (o supremo de um conjunto é o seu *menor* limitante superior!);

(ii) analogamente, se existir $\inf\{A, B\}$, este deverá não apenas limitar inferiormente $\{A, B\}$ (i.e., $\inf\{A, B\} \subseteq A, B$), como também satisfazer $D \subseteq \inf\{A, B\}$ para qualquer subconjunto D de X com $D \subseteq A, B$ (o ínfimo de um conjunto é o seu *maior* limitante inferior!).

Parece familiar, não? Explicitamente, $\sup\{A, B\}$ deve ser o menor subconjunto de X a conter tanto A quanto B, e já conhecemos um subconjunto que faz isso: $A \cup B$! E, de fato, tem-se $\sup\{A, B\} = A \cup B$:

✓ $A \cup B$ limita $\{A, B\}$ superiormente, pois ocorre $A, B \subseteq A \cup B$;

✓ $A \cup B$ é o menor limitante superior de $\{A, B\}$, já que $A \cup B \subseteq C$ sempre que $A, B \subseteq C$.

Talvez você possa estar se perguntando: como é possível que as linhas de argumentação acima tenham provado a identidade "$\sup\{A, B\} = A \cup B$", dado que a expressão "$\sup\{A, B\}$" nem sequer apareceu? Resposta: as condições verificadas acima são a *definição de supremo* que, quando existe, é único; assim, ao mostrar que $A \cup B$ tem as propriedades que $\sup\{A, B\}$ deveria ter, conclui-se que $A \cup B$ *é* o supremo procurado. ▲

Exercício 0.102 $(^{\star}_{\star\star})$**.** Mostre que $A \cap B = \inf\{A, B\}$ em $\wp(X)$. ■

Após a introdução de corpos completos na próxima seção, será prática comum tomar supremos e ínfimos apenas de subconjuntos não vazios e limitados. No entanto, a definição não proíbe tais casos, o que traz a pergunta: o que seriam $\sup_{\mathbb{P}} \emptyset$ e $\inf_{\mathbb{P}} \emptyset$ numa ordem $(\mathbb{P}, \leq)$?

Explicitamente, $\sup_{\mathbb{P}} \emptyset$ é o menor dos limitantes inferiores de $\emptyset$. Agora, leia com calma: como todo elemento de $\mathbb{P}$ é limitante superior de $\emptyset$ (por vacuidade!), segue que $\emptyset$ tem supremo em $\mathbb{P}$ se, e somente se, $\mathbb{P}$ tem mínimo. Ou seja, $\sup_{\mathbb{P}} \emptyset = \min \mathbb{P}$. Analogamente, $\emptyset$ tem ínfimo em $\mathbb{P}$ se, e somente se, $\mathbb{P}$ tem máximo, e assim $\inf_{\mathbb{P}} \emptyset = \max \mathbb{P}$. É por isso que em corpos ordenados, não existem $\sup \emptyset$ nem $\inf \emptyset$: corpos ordenados são ilimitados inferior e superiormente (verifique)$^{\star}_{\star}$ e, em particular, não têm máximo e nem mínimo.

§1 (Importante) corpos estendidos e intervalos

Em contraponto ao que se observou acima, é formalmente lícito *acrescentar* pontos num corpo ordenado $\mathbb{K}$ que sirvam como *extremos*. De modo geral, dada uma ordem parcial $(\mathbb{P}, \leq)$, é possível tomar elementos *artificiais*[†] distintos $p, q \notin \mathbb{P}$ e considerar sobre $\mathbb{P}' := \mathbb{P} \cup \{p, q\}$ uma ordem parcial $\preceq$ que estende $\leq$, declarando-se $p \preceq x$ e $x \preceq q$ para qualquer $x \in \mathbb{P}'$, e para $x, y \in \mathbb{P}$, $x \preceq y$ se, e somente se, $x \leq y$; em particular, obtém-se $p = \min \mathbb{P}'$ e $q = \max \mathbb{P}'$. Um *corpo estendido* é a ordem oriunda deste processo aplicado a um corpo ordenado $\mathbb{K}$.

[†]Ou virtuais, fictícios, etc. Não faz diferença, dado que tudo aqui é algum tipo de ficção.

Definição 0.8.8. Denotaremos por $[-\infty, +\infty]_{\mathbb{K}}$ o conjunto $\mathbb{K} \cup \{-\infty, +\infty\}$, onde $-\infty$ e $+\infty$ são elementos não pertencentes a $\mathbb{K}$, com a ordem acima (com $p := -\infty$ e $q := +\infty$), que passa a ser chamado de **corpo estendido**[†]. ¶

Observação 0.8.9. No futuro, $\mathbb{R}$ e seus subcorpos serão os únicos corpos considerados, caso em que $[-\infty, +\infty] := \mathbb{R} \cup \{-\infty, +\infty\}$ será chamada de *reta estendida*. Então, se preferir, pode fingir que $\mathbb{K}$ é $\mathbb{R}$. △

A escolha do símbolo "∞" para indicar os pontos artificiais acrescentados ao corpo $\mathbb{K}$ é arbitrária e segue apenas a prática comum. Dito isso, é importante ressaltar que, embora seja frequente se referir a tais pontos como "infinitos", seria mais correto xingá-los de *ilimitados*, posto que "infinito" costuma se referir à cardinalidade de conjuntos, enquanto "$-\infty$" e "$+\infty$" apenas denotam *extremos* artificiais numa ordem, algo bem mais específico.

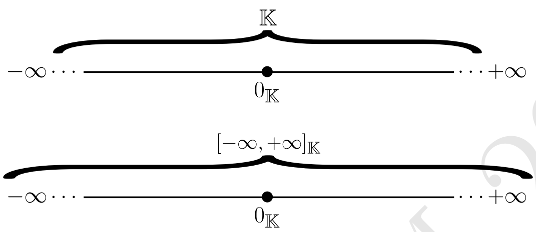


Figura 0.9: Se você já aceita o desenho de cima, por que não aceitar o desenho de baixo?

Definição 0.8.10 (Intervalos abertos)**.** Para $\alpha, \beta \in [-\infty, +\infty]_{\mathbb{K}}$, os conjuntos

$$[-\infty, \beta)_{\mathbb{K}} := \{x \in [-\infty, +\infty]_{\mathbb{K}} : x < \beta\}, \tag{0.1}$$
$$(\alpha, +\infty]_{\mathbb{K}} := \{x \in [-\infty, +\infty]_{\mathbb{K}} : \alpha < x\} \tag{0.2}$$

serão chamados de *intervalos abertos fundamentais* de $[-\infty, +\infty]_{\mathbb{K}}$. Diremos que um subconjunto $I \subseteq [-\infty, +\infty]_{\mathbb{K}}$ é um **intervalo aberto** se I for interseção finita de intervalos fundamentais. Quando $\mathbb{K}$ for claro pelo contexto, os subíndices serão abandonados[‡]. ¶

Sim: a notação acima não está errada, e os intervalos foram "fechados" nos pontos *infinitos*. Você não podia fazer isso em Cálculo I pois ainda não tinha idade para entender certas coisas, como a liberdade poética proporcionada pela Teoria dos Conjuntos. Mas esta fase da sua vida passou. Agora, por exemplo, é completamente lícito considerar o intervalo $[-\infty, 5)_{\mathbb{Q}}$ em $[-\infty, +\infty]_{\mathbb{Q}}$, explicitamente composto pelo ponto $-\infty$ e todos os números racionais menores do que 5, ou seja:

$$[-\infty, 5)_{\mathbb{Q}} = \{-\infty\} \cup \{x \in \mathbb{Q} : x < 5\}.$$

O mesmo se aplica à definição dada para intervalo aberto: secretamente, ela generaliza os intervalos abertos que você já conhecia.

Proposição 0.8.11. *Para $a, b \in [-\infty, +\infty]_{\mathbb{K}}$, o subconjunto*

$$(a, b)_{\mathbb{K}} := \{x \in [-\infty, +\infty]_{\mathbb{K}} : a < x < b\}$$

é um intervalo aberto inteiramente contido em $\mathbb{K}$, i.e., $(a, b)_{\mathbb{K}} \subseteq \mathbb{K}$. Em particular, $\emptyset$, $\mathbb{K}$, $(k, l)_{\mathbb{K}}$, $(-\infty, k)_{\mathbb{K}}$ e $(k, +\infty)_{\mathbb{K}}$ são intervalos abertos contidos em $\mathbb{K}$, para quaisquer $k, l \in \mathbb{K}$.

[†] O mais comum é escrever $\overline{\mathbb{K}}$ em vez de $[-\infty, +\infty]_{\mathbb{K}}$. No entanto, tal notação entraria em conflito com o *fecho topológico*, que será introduzido no próximo capítulo. Em tempo: isto nada tem a ver com as *extensões de corpos* da Teoria de Anéis. Por favor, não insista.

[‡] Em particular, após a introdução de $\mathbb{R}$, sempre consideraremos $\mathbb{K} = \mathbb{R}$.

Demonstração. A segunda parte segue da primeira ao se observar que $\emptyset = (a,b)_{\mathbb{K}}$ sempre que $b \leq a$ e $\mathbb{K} = (-\infty, +\infty)_{\mathbb{K}}$. Para a primeira parte, note que

$$(a,b)_{\mathbb{K}} = [-\infty, b)_{\mathbb{K}} \cap (a, +\infty]_{\mathbb{K}},$$

ou seja, é uma interseção finita de intervalos abertos fundamentais e, por isso, é um intervalo aberto. A inclusão segue pois $-\infty \leq a, b$ e $a, b \leq +\infty$, de modo que se $a < x < b$, então $x \notin \{-\infty, +\infty\}$ e, portanto, $x \in \mathbb{K}$. □

Observação 0.8.12. Embora a proposta seja elevar $-\infty$ e $+\infty$ ao patamar de *pontos*, eles não são elementos do corpo e, por isso, não podem ser operados livremente com os outros *elementos* do corpo. Em particular, não escreva atrocidades do tipo "$-\infty + \infty = 0$". Lembre-se: o mundo não vai acabar se você pensar um pouco antes de escrever uma abobrinha. △

O motivo para tais *intervalos* serem chamados de *abertos* só será abordado no próximo capítulo. Com isso dito, há outra palavra que deveria ter chamado sua atenção: por que tais animais são chamados de intervalos? Afinal, o que *significa* ser um intervalo?

Definição 0.8.13. Seja $(\mathbb{P}, \leq)$ uma ordem. Dizemos que um subconjunto $I \subseteq \mathbb{P}$ é um **intervalo** se para quaisquer $a, b, c \in \mathbb{P}$ ocorrer $c \in I$ sempre que $a \leq c \leq b$ com $a, b \in I$.¶

Figura 0.10: Na ilustração à esquerda, quaisquer dois pontos entre a e b estão na região destacada. Já na figura à direita, há pontos entre a e b que não estão na região destacada.

Exercício 0.103 (⁎⁎). Seja $\mathbb{K}$ um corpo ordenado.

a) Mostre que os intervalos abertos fundamentais de $[-\infty, +\infty]_{\mathbb{K}}$ são intervalos.

b) Mostre que a interseção de intervalos (numa ordem qualquer) é um intervalo.

c) Conclua que os intervalos abertos contidos em $\mathbb{K}$ são intervalos de $\mathbb{K}$. ■

Em particular, após introduzirmos a reta real $\mathbb{R}$, poderemos classificar todos os subconjuntos de $\mathbb{R}$ e de $[-\infty, +\infty]$ que são intervalos, o que condiz com os exemplos usuais que você já conhece do Cálculo I. Um pouco mais adiante, em posse da noção de *conexidade*, veremos que os intervalos são, precisamente, os subconjuntos *conexos* de $\mathbb{R}$, como mandam o *bom senso* e a *intuição*. Isto sugere a pergunta: se, na prática, os intervalos serão exatamente os subconjuntos que já sabíamos ser intervalos, para que serve ter uma definição abstrata de intervalo? Resposta: para simplificar a vida, sempre[†]!

Exercício 0.104 (⁎⁎). Supondo que $(\mathbb{T}, \leq)$ é ordem total, sejam $\alpha, \beta \in \mathbb{T}$, com $\alpha < \beta$. Mostre que se $I, J \subseteq \mathbb{T}$ são intervalos tais que $\alpha \in I$, $\beta \in J$ e $I \cap J = \emptyset$, então para quaisquer $x \in I$ e $y \in J$ deve ocorrer $x < y$. Dica: supondo $x \geq y$, investigue os cenários $\alpha < y$ e $y \leq \alpha$. ■

Futuramente, o exercício acima será usado na demonstração da clássica "lei da conservação do sinal", uma importante propriedade dos limites de ~~redes~~ funções e sequências reais.

[†]Como dizia o Prof. Alexandre "Sasha" Ananin, "a preguiça é a locomotiva do progresso".

Observação 0.8.14 (IMPORTANTE). Assim que nos convencermos da boa definição de $\mathbb{R}$, o subíndice $\mathbb{K}$ será abandonado em todos os lugares. Em particular, um intervalo $(a, b)_{\mathbb{R}}$ será denotado simplesmente por (a, b), o que traz um problema: como saber quando (a, b) indica um intervalo aberto e não um par ordenado $(a, b) \in \mathbb{R}^2$? RESPOSTA: contexto e prática. Por exemplo: no Exercício 0.128 (provar que $|(a, b)| = |\mathbb{R}|$), não faria sentido pensar que (a, b) indica um par ordenado (cf. Exercício 0.10). Evidentemente, você pode adotar alternativas notacionais, mas é importante se acostumar com tais abusos de notação, pois a literatura profissional está repleta disso. $\triangle$

0.9 Completude (no sentido de Dedekind)

0.9.0 Essencial

§0 Cortes e corpos completos

É chegada a hora de apresentar a reta real: a ideia é definir $\mathbb{R}$ como um corpo ordenado *sem buracos*. Evidentemente, isto pressupõe que saibamos o que é um *buraco*. No entanto, por questões estéticas e morais, buracos serão chamados de *cortes*.

Definição 0.9.0. Seja $\mathbb{K}$ um corpo ordenado. Um **corte** em $\mathbb{K}$ é um par (A, B) de subconjuntos não vazios de $\mathbb{K}$ tais que

(i) $A \cap B = \emptyset$ e $A \cup B = \mathbb{K}$,

(ii) para quaisquer $a \in A$ e $b \in B$ ocorre $a < b$. ¶

Exemplo 0.9.1. Ampliando o leque de *intervalos* apresentados na Definição 0.8.10 e na Proposição 0.8.11 (cf. Subseção 0.8.1 §1), fixado um corpo ordenado $\mathbb{K}$ e elementos $\alpha, \beta \in [-\infty, +\infty]_{\mathbb{K}}$, definimos

(i) $[\alpha, \beta]_{\mathbb{K}} := \{x \in [-\infty, +\infty]_{\mathbb{K}} : \alpha \leq x \leq \beta\}$,

(ii) $[\alpha, \beta)_{\mathbb{K}} := \{x \in [-\infty, +\infty]_{\mathbb{K}} : \alpha \leq x < \beta\}$ e

(iii) $(\alpha, \beta]_{\mathbb{K}} := \{x \in [-\infty, +\infty]_{\mathbb{K}} : \alpha < x \leq \beta\}$,

todos intervalos no sentido da Definição 0.8.13. Em particular, para $p \in \mathbb{K}$, ambos os intervalos $(-\infty, p]_{\mathbb{K}}$ e $[p, +\infty)_{\mathbb{K}}$ são subconjuntos de $\mathbb{K}$ (verifique?)* que induzem os chamados *cortes triviais*.

✓ (A, B) com $A := (-\infty, p]_{\mathbb{K}}$ e $B := (p, +\infty)_{\mathbb{K}}$, e

✓ (A, B) com $A := (-\infty, p)_{\mathbb{K}}$ e $B := [p, +\infty)_{\mathbb{K}}$.

Em geral, diremos que um corte (A, B) é **trivial** se existir $p \in \mathbb{K}$ tal que $p = \max A$ ou $p = \min B$. Desse modo, não é difícil perceber que (A, B) é trivial se, e somente se, (A, B) se enquadra em um dos casos acima para algum $p \in \mathbb{K}$ (verifique?)*. ▲

Como o exemplo acima sugere, os subconjuntos A e B na definição do corte (A, B) correspondem aos dois pedaços que se obteriam de $\mathbb{K}$ se este fosse *cortado* num determinado *ponto* de $\mathbb{K}$. É por essa razão que os cortes do exemplo anterior são triviais: para qualquer ponto $p \in \mathbb{K}$ fixado, é *trivial* cortar a reta em p, basta “inclinar a lâmina” para que p fique num dos lados do corte. A grande sacada vem agora: a depender do corpo considerado, podem haver “buracos” que permitam cortes sem extremos, i.e., não triviais.

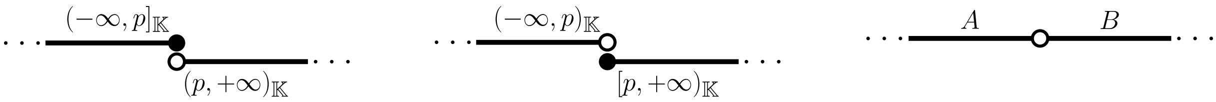


Figura 0.11: Um corte não trivial é, na prática, um "buraco".

Exemplo 0.9.2 (Fundamental: "$\sqrt{2}$")**.** Fixado um corpo ordenado $\mathbb{K}$, os subconjuntos

$$A := \left\{x \in \mathbb{K} : x < 0_{\mathbb{K}} \text{ ou } 0_{\mathbb{K}} \leq x^2 < 2_{\mathbb{K}}\right\} \text{ e } B := \left\{x \in \mathbb{K} : x > 0_{\mathbb{K}} \text{ e } x^2 \geq 2_{\mathbb{K}}\right\}$$

determinam um corte (A, B) em $\mathbb{K}$.

✓ $A \neq \emptyset$ pois $0_{\mathbb{K}}, 1_{\mathbb{K}} \in A$ e $B \neq \emptyset$ pois $2_{\mathbb{K}} \in B$.

✓ A tricotomia da ordem de $\mathbb{K}$ acarreta tanto $A \cap B = \emptyset$ quanto $A \cup B = \mathbb{K}$ (certo?)*.

✓ Finalmente, se $a \in A$ e $b \in B$, então $a < b$: isto é evidente para $a < 0_{\mathbb{K}}$; se ocorresse $a \geq 0_{\mathbb{K}}$ com $a \geq b$, teria-se $a^2 \geq ab \geq b^2 \geq 2_{\mathbb{K}}$, acarretando $a \notin A$.

Pergunta-se: tal corte é, necessariamente, induzido por algum $\alpha \in \mathbb{K}$? Como a discussão no Exemplo 0.9.1 sugere, isto equivale a perguntar sobre a existência de máximo para A ou mínimo para B.

Lema 0.9.3. *Seja $\alpha \in \mathbb{K}$. Se $\alpha = \sup A$ ou $\alpha = \min B$, então $\alpha^2 = 2_{\mathbb{K}}$.*

Demonstração. A prova a seguir é adaptada de Rudin [36]. Essencialmente, a ideia é mostrar que se $\alpha^2 \neq 2_{\mathbb{K}}$, então α não pode ser supremo de A e tampouco pode ser mínimo de B. Algumas considerações iniciais:

(i) como $1_{\mathbb{K}} \in A$ e $b > 0_{\mathbb{K}}$ para todo $b \in B$, podemos supor $\alpha > 0_{\mathbb{K}}$;

(ii) como a ordem de $\mathbb{K}$ é total, de $\alpha^2 \neq 2_{\mathbb{K}}$ restam apenas as alternativas $\alpha^2 < 2_{\mathbb{K}}$ ou $\alpha^2 > 2_{\mathbb{K}}$;

(iii) se ocorrer o primeiro caso, então $\alpha \notin B$ e, portanto, α não pode ser o mínimo de B;

(iv) se ocorrer o segundo caso, então $\alpha \in B$;

(v) se $\alpha \in B$ mas $\alpha \neq \min B$, então existe $b \in B$ com $b < \alpha$, donde segue que α não pode ser o supremo de A (certo?)*.

Desta argumentação preliminar, resulta que basta mostrar duas afirmações.

Afirmação 0. *Se $\alpha^2 < 2_{\mathbb{K}}$, então α não é supremo de A.*

Afirmação 1. *Se $\alpha^2 > 2_{\mathbb{K}}$, então α não é mínimo de B.*

⌜*Demonstração.* Para mostrar que α não é supremo de A, basta obter $\beta \in A$ tal que $\alpha < \beta$ (pois assim α não limitará A superiormente). Para mostrar que α não é mínimo de B, basta obter $\beta \in B$ com $\beta < \alpha$. Ora, escrevendo $\beta := \alpha + \gamma$, precisamos determinar γ de modo que se $\alpha \in A$, então $\gamma > 0_{\mathbb{K}}$ com $\alpha + \gamma \in A$, e se $\alpha \in B$, então $\gamma < 0_{\mathbb{K}}$ com $\alpha + \gamma \in B$. Como Rudin [36], faremos

$$\gamma := -\frac{\alpha^2 - 2_{\mathbb{K}}}{\alpha + 2_{\mathbb{K}}}, \tag{!!}$$

o que assegura as identidades (verifique!)$^{\star}_{\star}$

$$\underbrace{\beta = \alpha - \frac{\alpha^2 - 2_{\mathbb{K}}}{\alpha + 2_{\mathbb{K}}}}_{\text{(I)}} = \frac{2_{\mathbb{K}}\alpha + 2_{\mathbb{K}}}{\alpha + 2_{\mathbb{K}}} \quad \text{e} \quad \underbrace{\beta^2 - 2_{\mathbb{K}} = \frac{2_{\mathbb{K}}(\alpha^2 - 2_{\mathbb{K}})}{(\alpha + 2_{\mathbb{K}})^2}}_{\text{(II)}}.$$

Antes de prosseguir, observe que $\beta > 0_{\mathbb{K}}$ (o contrário daria $\alpha \leq -1_{\mathbb{K}}$). Agora, se $\alpha^2 < 2_{\mathbb{K}}$, então $\alpha^2 - 2_{\mathbb{K}} < 0_{\mathbb{K}}$, donde (I) acarreta $\beta > \alpha > 0_{\mathbb{K}}$, enquanto (II) garante $\beta \in A$. Por outro lado, se $\alpha^2 > 2_{\mathbb{K}}$, então $\alpha^2 - 2_{\mathbb{K}} > 0_{\mathbb{K}}$, donde (I) acarreta $0_{\mathbb{K}} < \beta < \alpha$, enquanto (II) implica $\beta^2 > 2_{\mathbb{K}}$, i.e., $\beta \in B$. ⌟

Enfim, se $\alpha^2 \neq 2_{\mathbb{K}}$ e:

- ✗ $\alpha^2 < 2_{\mathbb{K}}$, então α não pode ser mínimo de B (por (iii)), enquanto a Afirmação 0 prova que α também não pode ser supremo de A;

- ✗ $\alpha^2 > 2_{\mathbb{K}}$, então α não pode ser mínimo de B pela Afirmação 1, enquanto (v) prova que α também não pode ser supremo de A. □

Em posse do lema acima, observe que se $\mathbb{K}$ for um corpo em que *não existe* $\alpha \in \mathbb{K}$ satisfazendo $\alpha^2 = 2_{\mathbb{K}}$, então o corte (A, B) não poderá ser trivial, posto que o lema garante que, se existir, um máximo de A (ou um mínimo de B) deve ter *quadrado* igual a $2_{\mathbb{K}}$ (certo?)*. Em particular, $\mathbb{Q}$ tem buracos! ▲

Exercício 0.105 $\left(\substack{\star\\\star}\right)$**.** Mostre que não existe $q \in \mathbb{Q}$ com $q^2 = 2$. ■

Observação 0.9.4 (De onde Rudin tirou aquele γ?!)**.** Note que o *sinal* do γ procurado é dado, em ambos os casos, por $\alpha^2 - 2_{\mathbb{K}}$: no primeiro caso, temos $\alpha^2 - 2_{\mathbb{K}} < 0_{\mathbb{K}}$ e buscamos $\gamma > 0_{\mathbb{K}}$; no segundo caso, temos $\alpha^2 - 2_{\mathbb{K}} > 0_{\mathbb{K}}$ e buscamos $\gamma < 0_{\mathbb{K}}$. Assim, podemos fazer $\gamma := -(\alpha^2 - 2_{\mathbb{K}})x$ a fim de encontrar valores de x que determinem o *sinal* de $\beta^2 - 2_{\mathbb{K}}$ como desejado. Em outras palavras: caímos na resolução de uma inequação de segundo grau!

[Fingindo que estamos em $\mathbb{R}$]. Com γ dado como acima, podemos reescrever $\beta^2 - 2$ fazendo

$$\beta^2 - 2 = (\alpha + \gamma)^2 - 2 = (\alpha^2 - 2)(1 - 2\alpha x + (\alpha^2 - 2)x^2),$$

e para se ter $\beta^2 - 2 = 0$, há dois valores possíveis para x, a saber $\frac{1}{\alpha + \sqrt{2}}$ e $\frac{1}{\alpha - \sqrt{2}}$ (verifique!)[†]. Logo, basta tomar x no intervalo *real* $\left(0, \frac{1}{\alpha + \sqrt{2}}\right)$ para obter tanto γ quanto β^2 nos *lugares corretos*, em qualquer um dos casos (certo?)$^{\star}_{\star}$. Ora, como espera-se implementar tal solução num corpo ordenado qualquer, precisa-se escolher um x racional neste intervalo. Enfim, como $\sqrt{2} < 2$, basta tomar $x := \frac{1}{\alpha + 2}$. Brilhante[‡]. △

Reconhecido o problema, precisa-se encontrar uma solução: o que exigir sobre um corpo ordenado a fim de não ter buracos?

Exercício 0.106 $\left(\substack{\star\\\star\star}\right)$**.** Sejam $\mathbb{K}$ um corpo ordenado e (A, B) um corte em $\mathbb{K}$. Para $\alpha \in \mathbb{K}$ qualquer, mostre que $\alpha = \sup A$ se, e somente se, $\alpha = \inf B$. Dica: primeiro, faça um desenho. ■

[†]Por "Bhaskara" mesmo! $\left(\substack{\star\\\star}\right)$

[‡]Rudin é conhecido por apresentar argumentos fantásticos sem enfatizar suas possíveis motivações. ~~Nesse aspecto, ele não é melhor que o Elon.~~ Em todo caso, para aprofundar a discussão sobre o que pode ter motivado a escolha do "γ de Rudin", confira https://math.stackexchange.com/questions/141774.

Observe então que se $\alpha = \sup A$ (ou $\alpha = \inf B$), então de duas uma: ou $\alpha = \max A$ ou $\alpha = \min B$. Com efeito, por valer $\mathbb{K} = A \cup B$, tem-se $\alpha \in A$ ou $\alpha \in B$, donde segue que $\alpha = \max A$ ou $\alpha = \min B$. Descobrimos assim como ~~tampar os buracos de um corpo ordenado~~ *descrever* um corpo ordenado *sem* buracos!

Definição 0.9.5. Um corpo ordenado $\mathbb{K}$ é chamado de **completo**[†] se todo subconjunto não vazio $A \subseteq \mathbb{K}$ limitado superiormente tem supremo. ¶

Observação 0.9.6. É equivalente pedir que todo subconjunto não vazio e limitado inferiormente tenha ínfimo (cf. Exercício 0.138). △

Exercício 0.107 ($\star$). Mostre que se $\mathbb{K}$ é um corpo ordenado completo e (A, B) é um corte em $\mathbb{K}$, então (A, B) é trivial. ■

§1 A condição arquimediana

A completude também traz outra consequência fundamental que você provavelmente pensou que era "de graça".

Teorema 0.9.7. *Se $\mathbb{K}$ é um corpo ordenado e completo, então o subconjunto*

$$\mathbb{N}_{\mathbb{K}} := \{n_{\mathbb{K}} : n \in \mathbb{N}\}$$

não é limitado superiormente, i.e., para qualquer $x \in \mathbb{K}$ existe $n \in \mathbb{N}$ tal que $x < n_{\mathbb{K}}$.

Demonstração. Se $\mathbb{N}_{\mathbb{K}}$ fosse limitado superiormente, existiria $\alpha := \sup \mathbb{N}_{\mathbb{K}} \in \mathbb{K}$. Como $\alpha - 1_{\mathbb{K}} < \alpha$, a minimalidade de α como limitante superior de $\mathbb{N}_{\mathbb{K}}$ acarreta a existência de $n \in \mathbb{N}$ tal que $\alpha - 1_{\mathbb{K}} < n_{\mathbb{K}}$. Mas daí $\alpha < n_{\mathbb{K}} + 1_{\mathbb{K}} := (n+1)_{\mathbb{K}} \in \mathbb{N}_{\mathbb{K}}$ (cf. Teorema 0.7.3), uma contradição. □

Definição 0.9.8. Um corpo ordenado $\mathbb{K}$ satisfazendo a tese do teorema acima é chamado de (corpo) **arquimediano**. ¶

Que grande porcaria a propriedade arquimediana, não é mesmo? O conjunto dos naturais é ilimitado?! O que de útil poderia decorrer de uma afirmação tão *trivial*? RESPOSTA:

Proposição 0.9.9. *Dado um corpo ordenado $\mathbb{K}$, são equivalentes:*

(i) ($\mathbb{K}$ é arquimediano) $\mathbb{N}_{\mathbb{K}}$ *não é limitado superiormente em* $\mathbb{K}$*;*

(ii) (ausência de "ilimitados") *não existe $x \in \mathbb{K}$ com $n_{\mathbb{K}} < x$ para todo $n \in \mathbb{N}$;*

(iii) (ausência de "infinitésimos") *não existe $x \in \mathbb{K}$ com $x \neq 0_{\mathbb{K}}$ satisfazendo*

$$|x|_{\mathbb{K}} < \frac{1_{\mathbb{K}}}{n_{\mathbb{K}}}$$

para todo $n \in \mathbb{N} \setminus \{0\}$;

(iv) ($\mathbb{Q}$ é "denso" em $\mathbb{K}$) *se $x, y \in \mathbb{K}$ e $x < y$, então existe $q \in \mathbb{Q}$ tal que $x < q_{\mathbb{K}} < y$.*

[†] Ou *Dedekind-completo*.

Demonstração. Os três primeiros itens são *claramente* equivalentes entre si (já sabé, né?)$\star$. Agora, assumindo (iii), provaremos (iv). Como $x < y$, temos $y - x = |y - x|_{\mathbb{K}} > 0_{\mathbb{K}}$ e, por (iii), existe $n \in \mathbb{N} \setminus \{0\}$ com

$$\frac{1_{\mathbb{K}}}{n_{\mathbb{K}}} < y - x$$

e, por conseguinte, $1_{\mathbb{K}} + nx < ny$ (por quê?)*. Se encontrarmos um número inteiro $m \in \mathbb{Z}$ satisfazendo $nx < m_{\mathbb{K}} \leq 1_{\mathbb{K}} + nx$ ou $nx \leq m_{\mathbb{K}} < 1_{\mathbb{K}} + nx$, então a desigualdade desejada seguirá com $q := \frac{m}{n}$ se $nx < m_{\mathbb{K}}$, ou com $q := \frac{m+1}{n}$ se $nx = m_{\mathbb{K}}$ (certo?)$\star$. Supondo $x \neq 0_{\mathbb{K}}$, a condição (iii) novamente assegura $s \in \mathbb{N} \setminus \{0\}$ com

$$\frac{1_{\mathbb{K}}}{s_{\mathbb{K}}} < \frac{1_{\mathbb{K}}}{|nx|_{\mathbb{K}}},$$

de modo que ao tomar $m := \min\left\{s \in \mathbb{N} \setminus \{0\} : \frac{1_{\mathbb{K}}}{s_{\mathbb{K}}} < \frac{1_{\mathbb{K}}}{|nx|_{\mathbb{K}}}\right\}$ resulta $m_{\mathbb{K}} - 1_{\mathbb{K}} \leq |nx|_{\mathbb{K}} < m_{\mathbb{K}}$ e, consequentemente, $nx < m_{\mathbb{K}} \leq nx + 1_{\mathbb{K}}$ se $x > 0_{\mathbb{K}}$, enquanto $nx \leq -m_{\mathbb{K}} + 1_{\mathbb{K}} < nx + 1_{\mathbb{K}}$ se $x < 0_{\mathbb{K}}$ (percebeu?!)*. O caso $x = 0_{\mathbb{K}}$ fica por sua conta ($\star$).

Finalmente, supondo (iv), mostraremos que $\mathbb{N}$ é ilimitado superiormente: para $r \in \mathbb{K}$ com $r > 0_{\mathbb{K}}$, existem $m, n \in \mathbb{N} \setminus \{0\}$ tais que

$$0_{\mathbb{K}} < \frac{1_{\mathbb{K}}}{n_{\mathbb{K}}} \leq \frac{m_{\mathbb{K}}}{n_{\mathbb{K}}} < \frac{1_{\mathbb{K}}}{r},$$

donde segue que $r < n_{\mathbb{K}}$, como desejado. □

Exercício 0.108 ($\star$)**.** Sejam $\mathbb{K}$ um corpo arquimediano e $x, y \in \mathbb{K}$. Mostre que se $x, y > 0_{\mathbb{K}}$, então existe $N \in \mathbb{N}$ com $Nx > y$. ■

A proposição acima estabelece que para um corpo ordenado $\mathbb{K}$ fixado, são as cópias de $\mathbb{N}$ e $\mathbb{Q}$ em $\mathbb{K}$ que codificam a informação necessária para decidir se $\mathbb{K}$ é arquimediano ou não. Em particular, é de se esperar que o próprio corpo ordenado $\mathbb{Q}$ seja arquimediano, o que de fato ocorre: dados $p, q \in \mathbb{Q}$ distintos, não é difícil perceber que $s := p + r$ é tal que $p < s < q$, onde $r := \frac{|p-q|}{2}$. Em particular, por $\mathbb{Q}$ ter subconjuntos não vazios, limitados superiormente e sem supremo, resulta que a condição arquimediana não garante completude.

0.9.1 Extras

§0 Corpos não arquimedianos

Definição 0.9.10. Dado um corpo ordenado $\mathbb{K}$, diremos que $x \in \mathbb{K}$ é **infinitesimal** em $\mathbb{K}$, ou é um **infinitésimo**, se para todo $n \in \mathbb{N}$ valer $|nx|_{\mathbb{K}} < 1_{\mathbb{K}}$. Analogamente, $x \in \mathbb{K}$ é **ilimitado**[†] em $\mathbb{K}$ se para todo $n \in \mathbb{N}$ valer $n_{\mathbb{K}} < |x|_{\mathbb{K}}$. ¶

É claro que $0_{\mathbb{K}} \in \mathbb{K}$ é infinitesimal em $\mathbb{K}$. Por outro lado, $x \neq 0_{\mathbb{K}}$ é infinitesimal em $\mathbb{K}$ se, e somente se, $\frac{1}{x}$ é ilimitado em $\mathbb{K}$. Logo, $\mathbb{K}$ tem infinitésimos não nulos se, e somente se, $\mathbb{K}$ tem elementos ilimitados (verifique)*. Portanto, corpos arquimedianos são precisamente aqueles nos quais *não* existem infinitésimos não nulos (certo?)*, o que sugere a pergunta: há algum corpo ordenado não arquimediano? Resposta: sim.

[†]Na Wikipedia, você encontrará tais números xingados como "*infinitos*", mas tal terminologia não é adequada, por confundir noções de ordem e cardinalidade.

Proposição 0.9.11. *Se* $\mathbb{D}$ *é um* domínio ordenado[†], *então seu* corpo de frações $\mathbb{K} \coloneqq \mathrm{Frac}(\mathbb{D})$ *admite uma relação de ordem total* $\sqsubseteq$, *compatível com a ordem de* $\mathbb{D}$ *e que faz de* $\mathbb{K}$ *um corpo ordenado.*

Demonstração. Basta encarar o Exemplo 0.7.1 até que ele te encare de volta. □

Observação 0.9.12. O *corpo de frações* K de um domínio D é um corpo cuja construção remete aos racionais quando comparados aos inteiros. Para construir um, basta imitar os passos utilizados na construção de $\mathbb{Q}$ (cf. Exercício 0.53), trocando $\mathbb{Z}$ por D. △

Portanto, a fim de obter um corpo ordenado dotado de infinitésimos, basta encontrar um domínio ordenado $\mathbb{D}$ dotado de um elemento ilimitado x, pois daí seu inverso multiplicativo x^{-1} (no corpo de frações $\mathbb{K}$) será um infinitésimo não trivial.

Exercício 0.109 $(\substack{\star\star\\\star\star})$**.** Para um domínio ordenado $(\mathbb{D}, \preceq)$, considere o *anel de polinômios* na indeterminada x e coeficientes em $\mathbb{D}$, denotado por $\mathbb{D}[x]$.

a) Mostre que $\mathbb{D}[x]$ é um domínio legítimo. Dica: avalie o grau de um produto de polinômios.

b) Dados $p(x), q(x) \in \mathbb{D}[x]$, declare $p(x) \sqsubseteq q(x)$ se, e somente se, $p(x) = q(x)$ ou o coeficiente líder de $q(x) - p(x)$ é (estritamente!) maior do que $0 \in \mathbb{D}$. Mostre que tal relação faz de $\mathbb{D}[x]$ um domínio ordenado em que x é ilimitado.

c) Conclua que $\mathbb{D}(x)$, o corpo de frações do domínio $\mathbb{D}[x]$, é um corpo ordenado que contém infinitésimos. ■

§1 Análise "não *standard*"

O Cálculo não nasceu em sua formulação atual. Originalmente, para definir a *derivada* de uma função "$f(x)$", por exemplo, considerava-se a expressão

$$\frac{f(x+\varepsilon) - f(x)}{\varepsilon},$$

com ε um infinitesimal não nulo, a fim de obter daí sua *parte não infinitesimal*. Por exemplo: com $f(x) \coloneqq x^2$, tem-se

$$\frac{(x+\varepsilon)^2 - x^2}{\varepsilon} = \frac{x^2 + 2x\varepsilon + \varepsilon^2 - x^2}{\varepsilon} = 2x + \varepsilon,$$

donde segue que a derivada de $f(x)$ é $f'(x) = 2x$. A questão é: como justificar isso? Afinal de contas, num certo momento do cálculo, trata-se $\varepsilon \neq 0$ para, posteriormente, agir como se $\varepsilon = 0$. Na época, as definições não se embasavam em entidades abstratas como conjuntos, mas sim em noções geométrico-físicas, de modo que tal tratamento "artificial" causava certo incômodo. O tempo passou e, aos poucos, infinitésimos foram substituídos por limites e a reta geométrica tornou-se um corpo arquimediano completo.

No entanto, nos anos 60 do século passado, técnicas avançadas de Lógica-Matemática foram utilizadas na elaboração do que passou a ser conhecido como *Nonstandard Analysis*, ou *Análise não padrão*, que permite não apenas formalizar e, em certa medida, justificar a metodologia original, como também sistematizar métodos para transferir resultados da Análise não padrão para a Análise usual — e vice-versa. Para saber mais, o texto de Keisler [20] é um excelente ponto de partida.

[†]Cuja definição é a mesma dos corpos ordenados, trocando-se o corpo $\mathbb{K}$ por um *domínio* $\mathbb{D}$. Em tempo, lembre-se de que um anel $\mathbb{D}$ é chamado de **domínio** se $0_{\mathbb{D}} \neq 1_{\mathbb{D}}$ e $xy \neq 0_{\mathbb{D}}$ sempre que $x, y \in \mathbb{D} \setminus \{0_{\mathbb{D}}\}$.

0.10 A reta real: definição, unicidade e cardinalidade

0.10.0 Essencial

§0 Unicidade de corpos completos (a menos de isomorfismo)

Assim como o Axioma de Dedekind-Peano postulou a existência de $\mathbb{N}$, vamos postular a existência ~~da reta real~~ de um corpo ordenado completo.

Axioma da Preguiça Infinita. Existe um corpo ordenado completo.

O nome dado ao axioma acima explicita a sua motivação: preguiça. Enquanto, no caso de $\mathbb{N}$, precisa-se realmente postular sua existência[†], aqui, um corpo ordenado completo poderia ser efetivamente construído, mas o custo seria demasiado alto[‡] em comparação aos benefícios: na prática, apenas transformaríamos o axioma acima num teorema e nunca mais voltaríamos a aproveitar a demonstração (neste texto).

Com isso dito, parece mais razoável dar atenção ao problema da *unicidade*: como já se mencionou anteriormente, a ideia é definir a reta real $\mathbb{R}$ como *um* corpo ordenado completo. Ora, admitindo-se que existe pelo menos um objeto dessa natureza, surge a pergunta: e se existir outro? Se duas construções distintas para corpos ordenados completos forem apresentadas, pode-se garantir que os corpos em questão são similares em algum sentido? Resposta: sim.

Observação 0.10.0. Para o que segue, convém revisar a noção de morfismo de anel na Subseção 0.6.1 §0. △

Definição 0.10.1. Dados corpos ordenados $\mathbb{K}$ e $\mathbb{K}'$, diremos que uma função $f\colon \mathbb{K} \to \mathbb{K}'$ é um **morfismo de corpos ordenados** se f for simultaneamente um morfismo de corpos e uma função crescente[††]. Diz-se que f é um **isomorfismo** de corpos ordenados se existir um morfismo de corpos ordenados $g\colon \mathbb{K}' \to \mathbb{K}$ com $g \circ f = \mathrm{Id}_{\mathbb{K}}$ e $f \circ g = \mathrm{Id}_{\mathbb{K}'}$. Em tais condições, $\mathbb{K}$ e $\mathbb{K}'$ são ditos **isomorfos**. ¶

Exercício 0.110 ($^{\star}_{\star}$)**.** Mostre que se $\mathbb{K}$ é corpo ordenado, então existe um único morfismo de corpos ordenados $f\colon \mathbb{Q} \to \mathbb{K}$. ■

A definição de isomorfismo dada acima é bem mais geral e se aplica, naturalmente, a qualquer contexto no qual uma noção *apropriada* de morfismo estiver disponível. Em certo sentido, enquanto *morfismos* são meios pelos quais *objetos* de um mesmo *tipo* ou *categoria* se *comunicam*, *isomorfismos* são meios que permitem não apenas a troca de informação, mas também a fidelidade nas traduções de um lado para outro. Há, porém, um modo bem mais prático de verificar isomorfismos *no presente contexto*.

Exercício 0.111 ($^{\star}_{\star}$)**.** Sejam A e B anéis e $f\colon A \to B$ um morfismo de anéis.

a) Mostre que se f é bijetora, então f é um isomorfismo de anéis. Dica: a inversa (que existe!) deve satisfazer as condições para ser morfismo.

b) Mostre que se A e B são corpos ordenados e f é bijeção crescente, então f é um isomorfismo de corpos ordenados. Dica: pelo item anterior, f já é um isomorfismo de corpos, enquanto o Exercício 0.70 permite concluir que f^{-1} também é crescente. ■

[†] Grosso modo, postular a existência de $\mathbb{N}$ equivale a assumir que existe ao menos um conjunto infinito. Detalhes mais precisos fogem do escopo do texto.

[‡] Mesmo assim, uma breve discussão é apresentada na Subseção 0.10.1 §1.

[††] Como definido no Exercício 0.70.

Observação 0.10.2. Apesar do que se estabeleceu acima, há outros contextos (ou *categorias*) nos quais a mera bijetividade não é suficiente para atestar o isomorfismo entre os objetos considerados. Por exemplo: na *categoria* dos *espaços topológicos*, que conheceremos superficialmente em breve, *funções contínuas* fazem o papel de morfismos, e nem toda função contínua bijetiva tem inversa contínua. △

Moralmente, dizer que A e B são anéis ou corpos (ordenados) isomorfos significa afirmar que embora A e B possam ter definições distintas, os *comportamentos* que suas estruturas modelam são os mesmos. O próximo exercício pode dar uma ideia mais clara sobre tudo isso no contexto específico dos corpos ordenados.

Exercício 0.112 ($\star$)**.** Sejam $\mathbb{K}$ e $\mathbb{K}'$ corpos ordenados e $f\colon \mathbb{K} \to \mathbb{K}'$ um isomorfismo de corpos ordenados.

a) Mostre que $S \subseteq \mathbb{K}$ é limitado superiormente se, e somente se, $f[S] \subseteq \mathbb{K}'$ é limitado superiormente.

b) Mostre que $S \subseteq \mathbb{K}$ admite um supremo $\alpha \in \mathbb{K}$ se, e somente se, $f[S]$ admite supremo $\beta \in \mathbb{K}'$. Além disso, tem-se $\beta = f(\alpha)$.

c) Mostre que a equação $x^2 - 2_{\mathbb{K}} = 0_{\mathbb{K}}$ tem solução em $\mathbb{K}$ se, e somente se, a equação $x^2 - 2_{\mathbb{K}'} = 0_{\mathbb{K}'}$ tem solução em $\mathbb{K}'$. ■

Pelo que se expôs acima, um modo legítimo de resolver o problema da "unicidade" seria mostrar que quaisquer dois corpos ordenados completos são isomorfos. É precisamente isso o que será feito a seguir.

Lema 0.10.3. *Sejam $\mathbb{A}$ e $\mathbb{K}$ corpos ordenados e, para cada $a \in \mathbb{A}$, considere o subconjunto $\mathbb{Q}_{\mathbb{K},a} := \{q_{\mathbb{K}} \in \mathbb{Q}_{\mathbb{K}} : q_{\mathbb{A}} < a\}$. Se $\mathbb{A}$ é arquimediano e $\mathbb{K}$ é completo, então a correspondência*

$$\begin{aligned} \rho\colon \mathbb{A} &\to \mathbb{K} \\ a &\mapsto \sup \mathbb{Q}_{\mathbb{K},a} \end{aligned} \tag{0.3}$$

é um morfismo de corpos ordenados.

É mais fácil entender a prova do que escrevê-la. A coisa toda é bastante visual, como ilustrado a seguir.

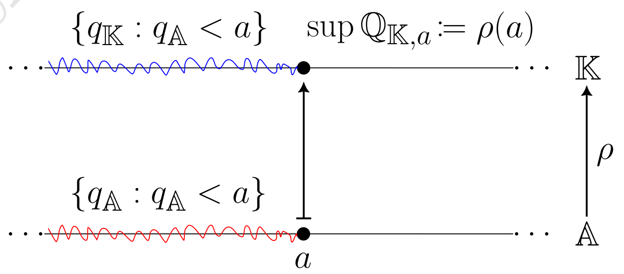


Figura 0.12: Sincronização de corpos.

Para cada $a \in \mathbb{A}$ considera-se, num primeiro momento, a coleção dos racionais (interpretados em $\mathbb{A}$) menores do que a. Ao *interpretar* tais números racionais em $\mathbb{K}$, obtém-se um conjunto limitado superiormente, que por sua vez admite um supremo em virtude da completude de $\mathbb{K}$. Finalmente, ρ apenas associa a ao supremo obtido no passo anterior. Mesmo que $\mathbb{A}$ e $\mathbb{K}$ sejam *construídos* de maneiras distintas, o fato de ambos *interpretarem* cópias *densas* de $\mathbb{Q}$ permite sincronizá-los entre si.

Demonstração. É edificante observar, antes de qualquer outra coisa, que a *relação* ρ é, na verdade, uma função:

- ✓ a propriedade arquimediana de $\mathbb{A}$ permite mostrar tanto que $\mathbb{Q}_{\mathbb{K},a} \neq \emptyset$ quanto a existência de um limitante superior em $\mathbb{K}$ (certo?)*;

- ✓ logo, a completude de $\mathbb{K}$ assegura a existência de um único $\rho(a) \in \mathbb{K}$ digno de ser xingado como $\sup \mathbb{Q}_{\mathbb{K},a}$.

Em outras palavras: ρ associa a cada $a \in \mathbb{A}$ um único $\rho(a) \in \mathbb{K}$, como esperado. Você pode cuidar dos detalhes omitidos acima[†]. Agora, mostraremos que ρ é um morfismo de corpos. Para isso, para quaisquer $a, b \in \mathbb{A}$, precisa-se verificar que

$$\sup \mathbb{Q}_{\mathbb{K},1_{\mathbb{A}}} = 1_{\mathbb{K}}, \tag{0.4}$$
$$\sup \mathbb{Q}_{\mathbb{K},a+b} = \sup \mathbb{Q}_{\mathbb{K},a} + \sup \mathbb{Q}_{\mathbb{K},b} \tag{0.5}$$
$$\sup \mathbb{Q}_{\mathbb{K},ab} = \sup \mathbb{Q}_{\mathbb{K},a} \cdot \sup \mathbb{Q}_{\mathbb{K},b}. \tag{0.6}$$

Identidade (0.4)**.** Ela vale mais geralmente, pois $\sup \mathbb{Q}_{\mathbb{K},q_{\mathbb{A}}} = q_{\mathbb{K}}$ para todo $q \in \mathbb{Q}$. Com efeito, $q_{\mathbb{K}}$ limita $\mathbb{Q}_{\mathbb{K},q_{\mathbb{A}}}$ superiormente e, se $\beta < q_{\mathbb{K}}$, então existe $r \in \mathbb{Q}$ com $r < q$ e $\beta < r_{\mathbb{K}}$, mostrando que $q_{\mathbb{K}}$ é, legitimamente, o menor limitante superior de $\mathbb{Q}_{\mathbb{K},q_{\mathbb{A}}}$.

Identidade (0.5)**.** Pelo Teorema 0.8.6, basta mostrar que $\mathbb{Q}_{\mathbb{K},a+b} = \mathbb{Q}_{\mathbb{K},a} + \mathbb{Q}_{\mathbb{K},b}$, o que será verificado a seguir.

Por um lado, se $q := r + s$ com $r_{\mathbb{A}} < a$ e $s_{\mathbb{A}} < b$, então $r_{\mathbb{A}} + s_{\mathbb{A}} < a + b$, acarretando $q_{\mathbb{K}} = r_{\mathbb{K}} + s_{\mathbb{K}} \in \mathbb{Q}_{\mathbb{K},a+b}$, donde a arbitrariedade de q implica em $\mathbb{Q}_{\mathbb{K},a} + \mathbb{Q}_{\mathbb{K},b} \subseteq \mathbb{Q}_{\mathbb{K},a+b}$. Por outro lado, se $q_{\mathbb{A}} < a+b$, então $0_{\mathbb{A}} < a+b-q_{\mathbb{A}}$ e, pela condição arquimediana satisfeita por $\mathbb{A}$, existem $r, s \in \mathbb{Q}$ com $0_{\mathbb{A}} < r_{\mathbb{A}} < a+b-q_{\mathbb{A}}$ e $a-r_{\mathbb{A}} < s_{\mathbb{A}} < a$. Logo, $q_{\mathbb{A}}-s_{\mathbb{A}} < q_{\mathbb{A}}+r_{\mathbb{A}}-a < b$, com $q_{\mathbb{A}} - s_{\mathbb{A}} \in \mathbb{Q}_{\mathbb{K},b}$ e $s_{\mathbb{A}} \in \mathbb{Q}_{\mathbb{K},a}$, mostrando $q_{\mathbb{A}} = s_{\mathbb{A}} + q_{\mathbb{A}} - s_{\mathbb{A}} \in \mathbb{Q}_{\mathbb{K},a} + \mathbb{Q}_{\mathbb{K},b}$.

Identidade (0.6)**.** A verificação das possíveis variações de *sinal* se reduz ao caso em que $a, b > 0_{\mathbb{A}}$, desde que se saiba da identidade auxiliar $\rho(-a) = -\rho(a)$. De fato, em posse disso, para $a < 0_{\mathbb{A}}$ e $b > 0_{\mathbb{A}}$, por exemplo, resulta

$$-\rho(ab) = \rho(-ab) = \rho((-a)b) = \rho(-a)\rho(b) = -\rho(a)\rho(b),$$

com um raciocínio análogo para o caso em que $a < 0_{\mathbb{A}}$ e $b < 0_{\mathbb{A}}$ (o caso em que $a = 0_{\mathbb{A}}$ ou $b = 0_{\mathbb{A}}$ é, em vista de (0.4), imediato). Tratemos então das identidades *suficientes*.

- ✓ Observe que se $x > 0_{\mathbb{A}}$, então o conjunto $D_x := \{q_{\mathbb{K}} \in \mathbb{Q}_{\mathbb{K},x} : q > 0_{\mathbb{A}}\}$ é tal que $\sup \mathbb{Q}_{\mathbb{K},x} = \sup D_x$ (certo?)*. Daí, observando que $D_{ab} = D_a \cdot D_b$ (verifique!)$^{\star}_{\star}$, a identidade $\rho(ab) = \rho(a)\rho(b)$ para $a, b > 0_{\mathbb{A}}$ segue do Teorema 0.8.6.

- ✓ Para a identidade auxiliar $\rho(-a) = -\rho(a)$, note que não há perda de generalidade em supor $a \notin \mathbb{Q}_{\mathbb{A}}$ (por (0.4)). Daí, note que se $q_{\mathbb{A}} < a$, então $-a < -q_{\mathbb{A}}$, donde a definição de supremo acarreta $\sup \mathbb{Q}_{\mathbb{K},-a} \leq -q_{\mathbb{K}}$. Logo, $q_{\mathbb{K}} \leq -\sup \mathbb{Q}_{\mathbb{K},-a}$ e, novamente pela definição, $\sup \mathbb{Q}_{\mathbb{K},a} \leq -\sup \mathbb{Q}_{\mathbb{K},-a}$. Se a desigualdade fosse estrita, a condição arquimediana garantiria um $p \in \mathbb{Q}$ com $\sup \mathbb{Q}_{\mathbb{K},a} < p_{\mathbb{K}} < -\sup \mathbb{Q}_{\mathbb{K},-a}$, com $a > p_{\mathbb{A}}$ e, consequentemente, $-p_{\mathbb{A}} < -a$, o que implicaria em $-p_{\mathbb{K}} \leq \sup \mathbb{Q}_{\mathbb{K},-a}$, i.e., $-\sup \mathbb{Q}_{\mathbb{K},-a} \leq p_{\mathbb{K}}$, uma contradição. Portanto, $\sup \mathbb{Q}_{\mathbb{K},a} = -\sup \mathbb{Q}_{\mathbb{K},-a}$, como desejado.

[†]Será essencial lembrar que as correspondências $q \mapsto q_{\mathbb{A}}$ e $q \mapsto q_{\mathbb{K}}$ definem (únicos) morfismos de corpos ordenados da forma $\mathbb{Q} \to \mathbb{A}$ e $\mathbb{Q} \to \mathbb{K}$, respectivamente.

A enfadonha discussão acima mostra que ρ é um morfismo de corpos. Resta a ordem: se $a \leq b$ em $\mathbb{A}$, então $\mathbb{Q}_{\mathbb{K},a} \subseteq \mathbb{Q}_{\mathbb{K},b}$ e, novamente pelo Teorema 0.8.6, segue que $\rho(a) \leq \rho(b)$. □

Exercício 0.113 ($^{\star}_{\star}$)**.** Complete os detalhes da demonstração acima. ■

Observação 0.10.4. Convém destacar que o morfismo de corpos ρ é *estritamente crescente*, no sentido da Observação 0.5.3. Há dois modos simples de se convencer disso:

(i) por $\mathbb{A}$ ser arquimediano, existe $q \in \mathbb{Q}$ com $a < q_{\mathbb{A}} < b$ e

$$\rho(a) := \sup \mathbb{Q}_{\mathbb{K},a} < q_{\mathbb{K}} < \sup \mathbb{Q}_{\mathbb{K},b} := \rho(b);$$

(ii) alternativamente, como ρ é um morfismo de corpos, segue que ρ é injetor (item b) do Exercício 0.87). Logo, ρ deve ser estritamente crescente.

Embora o segundo argumento mostre que *qualquer* morfismo de corpos ordenados é estritamente crescente, o primeiro argumento será importante em breve, quando surgir o problema de estimar a cardinalidade de corpos arquimedianos. △

Teorema 0.10.5. *Se $\mathbb{A}$ e $\mathbb{K}$ são corpos ordenados e completos, então o mapa $\rho\colon \mathbb{A} \to \mathbb{K}$ definido em (0.3) é um isomorfismo de corpos ordenados.*

Demonstração. Desta vez o corpo $\mathbb{A}$ também é completo. Logo, o lema anterior permite conjurar *dois* morfismos de corpos ordenados, simultaneamente:

$$\begin{array}{rl} \rho\colon \mathbb{A} \to \mathbb{K} \\ a \mapsto \sup \mathbb{Q}_{\mathbb{K},a} \end{array} \quad \text{e} \quad \begin{array}{rl} \sigma\colon \mathbb{K} \to \mathbb{A} \\ k \mapsto \sup \mathbb{Q}_{\mathbb{A},k} \end{array}$$

Portanto, basta mostrar que um é o inverso do outro. Fixado $a \in \mathbb{A}$, tem-se

$$\sigma(\rho(a)) := \sup \mathbb{Q}_{\mathbb{A},\rho(a)},$$

e busca-se verificar $\sigma(\rho(a)) = a$. Ora, dado $q_{\mathbb{A}} \in \mathbb{Q}_{\mathbb{A},\rho(a)}$, tem-se $q_{\mathbb{K}} < \rho(a)$, e isso proíbe a ocorrência de $a \leq q_{\mathbb{A}}$: caso contrário, teria-se $\rho(a) \leq \rho(q_{\mathbb{A}}) = q_{\mathbb{K}}$. Agora, se $\beta \in \mathbb{A}$ é tal que $\beta < a$, então $\rho(\beta) < \rho(a)$, e existe $q \in \mathbb{Q}$ com $\rho(\beta) < q_{\mathbb{K}} < \rho(a)$, donde segue que $q_{\mathbb{A}} \in \mathbb{Q}_{\mathbb{A},\rho(a)}$ com $\beta < q_{\mathbb{A}}$. Portanto, a é o menor limitante superior de $\mathbb{Q}_{\mathbb{A},\rho(a)}$, i.e., $a = \sigma(\rho(a))$, como queríamos. Analogamente, mostra-se que $\rho(\sigma(k)) = k$ para todo $k \in \mathbb{K}$. □

Exercício 0.114 ($^{\star}_{\star\star}$)**.** Sejam $\mathbb{A}$ e $\mathbb{K}$ corpos ordenados, com $\mathbb{A}$ arquimediano e $\mathbb{K}$ completo.

a) Mostre que se $\varphi\colon \mathbb{A} \to \mathbb{K}$ é um morfismo de corpos ordenados, então $\varphi(q_{\mathbb{A}}) = q_{\mathbb{K}}$ para todo $q \in \mathbb{Q}$. Dica: as correspondências $q \mapsto q_{\mathbb{A}}$ e $q \mapsto q_{\mathbb{K}}$ determinam os únicos morfismos de corpos ordenados da forma $\mathbb{Q} \to \mathbb{A}$ e $\mathbb{Q} \to \mathbb{K}$, respectivamente; por outro lado, a composição entre $q \mapsto q_{\mathbb{A}}$ e φ também determina um morfismo da forma $\mathbb{Q} \to \mathbb{K}$.

b) Conclua que existe um único morfismo de corpos ordenados da forma $\mathbb{A} \to \mathbb{K}$. Em particular, os isomorfismos do teorema anterior são únicos. ■

As discussões acima resolvem o problema da "unicidade" mencionado anteriormente, e tornam quase honesta a próxima

Definição 0.10.6. Denota-se por $\mathbb{R}$ *qualquer* corpo ordenado e completo, que passa a ser chamado de **conjunto dos números reais**, ou apenas de **reta real**. ¶

Devido a tal escolha de notação, perde o sentido carregar "$\mathbb{R}$" como subíndice para indicar em qual corpo ordenado e completo um determinado procedimento ocorre, postura que será aplicada também para o valor absoluto de um número real x, que será denotado por $|x|$ de agora em diante[†].

Além disso, em contextos algébricos ou *analíticos*, será inofensivo considerar como verdadeiras as inclusões próprias $\mathbb{N} \subsetneq \mathbb{Z} \subsetneq \mathbb{Q} \subsetneq \mathbb{R}$ embora, a rigor, existam apenas morfismos injetores que preservam as *estruturas algébricas e de ordem* subjacentes. Se, por um lado, isso soa demasiado arbitrário, por outro, a unicidade assegura que não haveria outra forma de *enxergar* um dentro do outro. É por isso que, na prática, tanto $\mathbb{N}$, quanto $\mathbb{Z}$ e $\mathbb{Q}$ são *substituídos* por suas cópias em $\mathbb{R}$.

§1 A cardinalidade da reta real

Um dos fatos mais marcantes nos desenvolvimentos iniciais da Teoria dos Conjuntos e no estudo de coleções infinitas foi a *constatação* de que o *tipo de infinito* de $\mathbb{R}$ é *estritamente maior* do que o *tipo de infinito* de $\mathbb{N}$, i.e., $\mathbb{N} \prec \mathbb{R}$ ou $|\mathbb{N}| < |\mathbb{R}|$ (lembre-se da Observação 0.4.23!). Explicitamente, isto significa que existe função injetora $\mathbb{N} \to \mathbb{R}$ mas não existe função bijetora entre $\mathbb{N}$ e $\mathbb{R}$. Com o jargão típico dos textos básicos de Análise, isto se reduz a dizer que "$\mathbb{R}$ é não enumerável".

A primeira parte é fácil: como $\mathbb{N}$ *é subconjunto* de $\mathbb{R}$, a inclusão $i\colon \mathbb{N} \to \mathbb{R}$ determina uma função automaticamente injetora. A parte não trivial é mostrar que não pode existir função injetora $\mathbb{R} \to \mathbb{N}$ ou, equivalentemente (cf. Teorema 0.4.9), não existe função sobrejetora $\mathbb{N} \to \mathbb{R}$.

Há várias formas de demonstrar a não enumerabilidade de $\mathbb{R}$. Um argumento bastante comum (*a.k.a.* "diagonalização de Cantor") consiste em tomar *qualquer* função $\varphi\colon \mathbb{N} \to \mathbb{R}$ e definir um *número* $r := r_0, r_1r_2r_3\ldots$ exigindo-se que $r_n \in \{0, \ldots, 9\} \setminus \{a_{n,n}, 1, 9\}$ para todo $n \in \mathbb{N}$, onde

$$\varphi(n) := a_{n,0}, a_{n,1}a_{n,2}\ldots a_{n,n}\ldots$$

indica a expansão de $\varphi(n)$ em *base* 10. Como $r \notin \operatorname{im}(\varphi)$, segue que φ não pode ser sobrejetora (percebeu?)[‡].

Evidentemente, tal argumento depende de um estudo um pouco mais cuidadoso de *séries* e *representações* decimais, o que atrasaria a *formalização* do resultado no texto. Uma alternativa popular no Brasil [24, 25], por exemplo, é apelar para a *compacidade* dos *intervalos fechados e limitados* de $\mathbb{R}$, disfarçada como a *propriedade dos intervalos encaixantes*. Porém, não vejo sentido em apelar prematuramente para propriedades *topológicas* da reta apenas para apresentar um argumento com gosto geométrico.

Aqui, a abordagem adotada será outra: mostraremos que $\mathbb{R}$ está em bijeção com $\wp(\mathbb{N})$, o conjunto das partes de $\mathbb{N}$. Daí, a não enumerabilidade de $\mathbb{R}$ seguirá do Teorema de Cantor (cf. Teorema 0.4.5): como não existe sobrejeção $\mathbb{N} \to \wp(\mathbb{N})$, tampouco pode existir sobrejeção $\mathbb{N} \to \mathbb{R}$, já que $\wp(\mathbb{N}) \approx \mathbb{R}$. Como *punição* pela transgressão imperdoável de não seguir o Elon, *ganharemos* de brinde a existência de bijeção entre $\mathbb{R}$ e $\mathbb{R}^n$ para qualquer $n \in \mathbb{N} \setminus \{0\}$, entre outras coisas (cf. Subseção 0.10.1 §2).

[†] O contexto deixará claro quando "$|x|$" representa o valor absoluto do *número real* x ou a cardinalidade do *conjunto* x. Pelo menos neste texto, conjuntos não são denotados por letras minúsculas, o que pode ajudar a acalmar os ânimos

[‡] Note que $r \neq \varphi(0)$ pois $r_0 \neq a_{n,0}$, $r \neq \varphi(1)$ pois $r_1 \neq a_{n,1}$, etc.

Exercício 0.115 $(^{\star}_{\star})$**.** Mostre que $\wp(\mathbb{Q}) \approx \wp(\mathbb{N})$. Dica: Exercício 0.58 + $\mathbb{N} \approx \mathbb{Q}$. ■

Lema 0.10.7. *Se $\mathbb{A}$ é corpo arquimediano, então $|\mathbb{A}| \leq |\wp(\mathbb{N})|$, i.e., existe função injetora $\mathbb{A} \to \wp(\mathbb{N})$.*

Demonstração. A correspondência $\partial \colon \mathbb{A} \to \wp(\mathbb{Q})$, dada por $\partial(a) := \{q \in \mathbb{Q} : q_{\mathbb{A}} < a\}$, é injetora: se $a, b \in \mathbb{A}$ são distintos, então ocorre $a < b$ ou $b < a$, donde a condição arquimediana assegura a existência de $q \in \mathbb{Q}$ *entre* a e b, acarretando em $\partial(a) \neq \partial(b)$. Portanto, existe função injetora $\mathbb{A} \to \wp(\mathbb{Q})$ e, pelo exercício anterior, existe injeção $\mathbb{A} \to \wp(\mathbb{N})$, como desejado. □

Como $\mathbb{R}$ é arquimediano (por ser completo), segue que $|\mathbb{R}| \leq |\wp(\mathbb{N})|$. O próximo passo é mostrar a desigualdade oposta, i.e., $|\wp(\mathbb{N})| \leq |\mathbb{R}|$, pois daí o Teorema 0.1.10 (Cantor-Bernstein) garantirá $|\mathbb{R}| = |\wp(\mathbb{N})|$. Embora este lado da desigualdade possa ser demonstrado de modo mais rápido por meio de *séries*, é possível maquiar os argumentos com supremos de ~~séries~~ *somas* finitas. A seguir, assume-se que você tenha familiaridade com a notação de somatório ($\sum$). Se não for o caso, confira a Subseção 0.10.1 §0.

Lema 0.10.8. *Se $0 < a < 1$, então* $\sup\left\{\sum_{n \leq m} a^n : m \in \mathbb{N}\right\} = \frac{1}{1-a}$.

Demonstração. Por indução, verifica-se que $a^n > 0$ para todo $n \in \mathbb{N}$: $a^0 := 1$, $a^1 = a > 0$ e, se $a^n > 0$, então $a^{n+1} := a \cdot a^n > 0$ pela Proposição 0.7.5. Por sua vez,

$$\sum_{n \leq m} a^n := \sum_{n=0}^{m} a^n = \frac{1 - a^{m+1}}{1 - a}$$

para qualquer $m \in \mathbb{N}$ (verifique!)$^{\star}_{\star}$.

Pela primeira parte, $S := \{\sum_{n \leq m} a^n : m \in \mathbb{N}\}$ é (evidentemente não vazio e) limitado superiormente (por quê?!)$^{\star}_{\star}$, donde a completude de $\mathbb{R}$ garante a existência de $s \in \mathbb{R}$ com $s = \sup S$. Para finalizar, o Teorema 0.8.6 se aplica e permite fazer

$$as = a \sup S = \sup aS = \underbrace{\sup\{ay : y \in S\} = \sup\{y - 1 : y \in S\}}_{\text{(por quê?!)}^{\star}_{\star}} = \sup(S - 1) = s - 1,$$

acarretando $s = \frac{1}{1-a}$. □

Exercício 0.116 $(\star)$**.** Mostre que $\sup\left\{\sum_{n \leq m} \frac{1}{10^n} : m \in \mathbb{N}\right\} = \frac{10}{9}$. ■

Lema 0.10.9. *Para cada $f \colon \mathbb{N} \to \{0, 1\}$, existe o número real $\psi(f)$ dado por*

$$\psi(f) := \sup\left\{\sum_{n \leq m} \frac{f(n)}{10^n} : m \in \mathbb{N}\right\}.$$

Além disso, ao denotar por $\{0, 1\}^{\mathbb{N}}$ o conjunto das funções da forma $\mathbb{N} \to \{0, 1\}$, a correspondência $\psi \colon \{0, 1\}^{\mathbb{N}} \to \mathbb{R}$ é injetora.

Demonstração. Fixada $f \in \{0,1\}^{\mathbb{N}}$, a ideia é usar a completude de $\mathbb{R}$ para garantir a existência de $\psi(f)$. Para tanto, é suficiente mostrar que o conjunto

$$S(f) := \left\{ \sum_{n \leq m} \frac{f(n)}{10^n} : m \in \mathbb{N} \right\}$$

é não vazio e limitado superiormente: obviamente, tem-se $S(f) \neq \emptyset$; para verificar a limitação, observe que $0 \leq f(n) \leq 1$ para todo $n \in \mathbb{N}$, donde segue que

$$\frac{1}{10^n} f(n) \leq \frac{1}{10^n} \Rightarrow \sum_{n \leq m} \frac{f(n)}{10^n} \leq \sum_{n \leq m} \frac{1}{10^n} \leq \frac{10}{9},$$

para qualquer $m \in \mathbb{N}$.

Isso mostrou que $S(f) \neq \emptyset$ é limitado superiormente para cada $f \in \{0,1\}^{\mathbb{N}}$. Consequentemente, a correspondência $\psi \colon \{0,1\}^{\mathbb{N}} \to \mathbb{R}$ está bem definida, pois o supremo de $S(f)$ é único para cada $f \in \{0,1\}^{\mathbb{N}}$. Resta apenas verificar a injetividade de ψ.

Para $p \in \mathbb{N}$ e $f \in \{0,1\}^{\mathbb{N}}$ fixados, mostraremos que

$$\psi_p(f) := \sup \left\{ \sum_{n \leq m} \frac{f(n+p)}{10^{n+p}} : m \in \mathbb{N} \right\} \leq \frac{1}{9 \cdot 10^{p-1}}. \tag{0.7}$$

De fato, já que $f(j) \leq 1$ para todo $j \in \mathbb{N}$, pode-se fazer

$$\sum_{n \leq m} \frac{f(n+p)}{10^{n+p}} = \frac{1}{10^p} \sum_{n \leq m} \frac{f(n+p)}{10^n} \leq \frac{1}{10^p} \sum_{n \leq m} \frac{1}{10^n} \leq \frac{1}{10^p} \cdot \frac{10}{9} = \frac{1}{9 \cdot 10^{p-1}},$$

donde a desigualdade (0.7) segue.

O último ingrediente da prova consiste em observar que

$$\psi(f) = \sum_{n \leq m} \frac{f(n)}{10^n} + \psi_{m+1}(f) \tag{0.8}$$

para qualquer $m \in \mathbb{N}$, o que segue *principalmente* do Teorema 0.8.6 (por quê?)*.

Enfim, para $g \in \{0,1\}^{\mathbb{N}}$ com $f \neq g$, existe $m := \min\{j \in \mathbb{N} : f(j) \neq g(j)\}$ e, por falta de opções, não há perda de generalidade em supor $f(m) = 0$ e $g(m) = 1$. Segue então de (0.8), bem como da minimalidade de m, que existe $r \in \mathbb{R}$ tal que

$$\psi(f) = r + \psi_{m+1}(f) \text{ e } \psi(g) = r + \frac{1}{10^m} + \psi_{m+1}(g).$$

Por (0.7), finalmente, obtém-se

$$\begin{aligned}
\psi(g) - \psi(f) &= \frac{1}{10^m} + \psi_{m+1}(g) - \psi_{m+1}(f) \geq \frac{1}{10^m} + 0 - \psi_{m+1}(f) \\
&\geq \frac{1}{10^m} - \frac{1}{9 \cdot 10^m} = \frac{8}{9 \cdot 10^m} > 0,
\end{aligned}$$

mostrando que $\psi(f) \neq \psi(g)$. $\square$

Secretamente, a prova acima consiste apenas em tomar sequências infinitas de 0's e 1's e interpretá-las como números reais por meio da expansão decimal. Assim, a sequência constante $(1, \ldots, 1, \ldots)$, por exemplo, se torna o número real que, na rua†, xingaríamos de $1,1111\ldots$

Corolário 0.10.10. *$\mathbb{R}$ é não enumerável.*

Demonstração. Pelo Exercício 0.57 temos $|\wp(\mathbb{N})| = |\{0,1\}^{\mathbb{N}}|$. Logo, mostrou-se que $|\wp(\mathbb{N})| \leq |\mathbb{R}|$. Como já tínhamos $|\mathbb{R}| \leq |\wp(\mathbb{N})|$, a igualdade $|\mathbb{R}| = |\wp(\mathbb{N})|$ segue em virtude do Teorema de Cantor-Bernstein. Por fim, como $\wp(\mathbb{N})$ é não enumerável pelo Teorema 0.4.5 (de Cantor), o resultado segue. □

Exercício 0.117 ($^{\star}_{\star\star}$). Diz-se que $r \in \mathbb{R}$ é **transcendente** se não existe polinômio <u>não nulo</u> $p(x) \in \mathbb{Q}[x]$ com $p(r) = 0$. Mostre que o conjunto dos números transcendentes é não enumerável. Em particular, conclua que $\mathbb{R} \setminus \mathbb{Q}$, o conjunto dos **números irracionais**, é não enumerável. Dica: para a parte final, use o Exercício 0.44. ■

0.10.1 Extras

§0 Somatórios e produtórios

A seguir, $\operatorname{seq}(\mathbb{R})$ denota a coleção das **sequências finitas** de números reais, i.e., $s \in \operatorname{seq}(\mathbb{R})$ se, e somente se, s é uma função da forma $\mathbb{N}_{<n} \to \mathbb{R}$ para algum $n \in \mathbb{N}$. Em particular, note que $\emptyset \in \operatorname{seq}(\mathbb{R})$. Para facilitar as notações, vamos escrever $(f_i : i \leq n)$ para indicar a sequência finita $f \colon \mathbb{N}_{n+1} \to \mathbb{R}$ que a cada $i \leq n$ associa f_i, enquanto $(g_i : i < n)$ indica uma função real g cujo domínio é $\mathbb{N}_{<n}$.

Definição 0.10.11 (Operadores Σ e Π)**.** Definem-se $\sum \colon \operatorname{seq}(\mathbb{R}) \to \mathbb{R}$ e $\prod \colon \operatorname{seq}(\mathbb{R}) \to \mathbb{R}$ da seguinte forma:

(i) $\sum \emptyset := 0$ e $\prod \emptyset := 1$;

(ii) $\sum (f_i : i \leq n) := \left(\sum (f_i : i < n)\right) + f_n$ e $\prod (f_i : i \leq n) := \left(\prod (f_i : i < n)\right) \cdot f_n$ para cada $n \in \mathbb{N}$ e $f := (f_i : i \leq n) \in \mathbb{R}^{n+1}$. ¶

É comum que o primeiro contato com as *definições* acima cause desconforto. Porém, a coisa é bastante simples, e consiste tão somente de um algoritmo de repetição. No caso de Σ, por exemplo, com $x_n \in \mathbb{R}$ para todo $n \in \mathbb{N}$, tem-se

$$
\begin{aligned}
\sum \emptyset &:= 0; \\
\sum (x_0) &:= \left(\sum \emptyset\right) + x_0 = 0 + x_0 = x_0; \\
\sum (x_0, x_1) &:= \left(\sum (x_0)\right) + x_1 = x_0 + x_1; \\
\sum (x_0, x_1, x_2) &:= \left(\sum (x_0, x_1)\right) + x_2 = (x_0 + x_1) + x_2; \\
&\vdots
\end{aligned}
$$

Intuitivamente, $\sum (x_i : i \leq n)$ expressa aquilo que se escreveria como $x_0 + x_1 + \ldots + x_n$. Isto sugere notações bem mais práticas e *maleáveis* do que as anteriores.

†Talvez você se surpreenda ao efetuar o cálculo “$10 \div 9$”, em sua calculadora, por exemplo.

Definição 0.10.12. Sejam $n \in \mathbb{N}$ e $f := (f_i : i \leq n) \in \mathbb{R}^{n+1}$.

(i) Tanto $\displaystyle\sum_{i \leq n} f_i$ quanto $\displaystyle\sum_{i=0}^{n} f_i$ serão usados para denotar $\sum (f_i : i \leq n)$.

(ii) Tanto $\displaystyle\prod_{i \leq n} f_i$ quanto $\displaystyle\prod_{i=0}^{n} f_i$ serão usados para denotar $\prod (f_i : i \leq n)$. ¶

A partir dessas definições, é relativamente simples adaptá-las a fim de dar sentido formal a variações típicas de *somatórios* e *produtórios*, como os listados a seguir:

(i) $\displaystyle\sum_{i=j}^{m} f_i$; (ii) $\displaystyle\prod_{i<m} a_i \sum_{j=0}^{n} b_j$; (iii) $\displaystyle\sum_{i=0}^{m} \sum_{j+k=i} a_j b_k c_i$

(iv) $\displaystyle\prod_{x \in X} h(x)$ para um conjunto finito X e uma função $h\colon X \to \mathbb{R}$;

(v) $\sum F$ para um subconjunto finito $F \subseteq \mathbb{R}$;

(vi) ...

Você provavelmente já tem familiaridade com esse tipo de notação e sabe como operá-las no dia a dia, inclusive em outros ambientes como anéis e espaços vetoriais. Ainda assim, quando alguma propriedade for usada sem maiores explicações, fica o convite para que você a demonstre — são bons exercícios de indução ~~para se fazer pelo menos uma vez na vida~~.

§1 Construções da reta real

Tipicamente, ao encontrar algum tipo de objeto matemático *incompleto* num sentido específico, a coleção das testemunhas da incompleteza/incompletude esconde um modo para completar o que faltava. Foi assim com $\mathbb{Z}$ e com $\mathbb{Q}$. Também é assim com $\mathbb{R}$. Há três modos clássicos para construir um corpo ordenado completo.

✓ *Cortes de Dedekind.* Considera-se a família

$$\mathcal{C} := \{(A, B) \in \wp(\mathbb{Q}) \times \wp(\mathbb{Q}) : (A, B) \text{ é corte de } \mathbb{Q} \text{ e } A \text{ não tem máximo}\},$$

conjunto que é munido de uma estrutura de corpo ordenado completo. A vantagem desta construção está na completude: como a ordem de $\mathcal{C}$ é, essencialmente, a inclusão, supremos se manifestam como reuniões de maneira quase automática. No entanto, a parte algébrica é terrível.

✓ *Sequências de Cauchy racionais.* Considera-se a família

$$\mathcal{S} := \{(q_n)_{n \in \mathbb{N}} \in \mathbb{Q}^{\mathbb{N}} : (q_n)_{n \in \mathbb{N}} \text{ é de Cauchy}\},$$

onde "ser de Cauchy" significa, *grosso modo*, que os termos da sequência se tornam arbitrariamente próximos uns dos outros. Daí, com uma relação de equivalência $\sim$ apropriada sobre $\mathcal{S}$, definem-se sobre o quociente $\mathcal{S}/\!\sim$ operações de adição e multiplicação, bem como uma ordem, que fazem de $\mathcal{S}/\!\sim$ um corpo ordenado completo.

✓ *Completamento uniforme.* Apela-se para a *teoria de espaços uniformes* e seus teoremas de *completamento*, que englobam tipos de estruturas chamadas de *grupos topológicos*, reino em que habita o grupo aditivo $(\mathbb{Q};+;0)$. Em tal cenário, mostra-se que o completamento uniforme de $\mathbb{Q}$ pode ser promovido ao patamar de corpo ordenado (completo).

Tecnicamente, o primeiro método é o menos exigente do ponto de vista terminológico: já teríamos bagagem suficiente para realizar a construção, se não tivéssemos mais o que fazer. Todavia, os meandros envolvidos nessa *implementação* não costumam contribuir para a prática da Análise no dia a dia.

Nesse sentido, o segundo método tem a vantagem de utilizar *ecos* de definições que serão importantes *após* a construção, como as *sequências de Cauchy*. Mesmo assim, as enfadonhas verificações de que as *estruturas* definidas satisfazem as condições desejadas só servem para o contexto da construção da reta: no futuro, quando você *precisar* de *espaços métricos completos*, tudo deverá ser refeito.

O terceiro método não tem essa desvantagem: um único teorema de *completamento* de *espaços uniformes* dá conta de *completar* tanto $\mathbb{Q}$ quanto *espaços métricos* e outras *estruturas uniformes* encontradas na *natureza*. Porém, dado o escopo do texto, desenvolver esse ferramental *apenas* para construir um corpo ordenado completo e, posteriormente, *fazer* Análise, seria descabido[†].

§2 Outras bijeções curiosas

Observação 0.10.13. Para um conjunto X e um número $n \in \mathbb{N}$ fixado, pode-se definir $X^n := X^{\mathbb{N}_{<n}}$, i.e., a coleção de todas as funções da forma $(x_i : i < n)$. Alternativamente, poderíamos definir $X^0 := \{\emptyset\}$, $X^1 := X$, $X^2 := X \times X$ e, mais geralmente, $X^{n+1} := X^n \times X$, o que essencialmente consiste em formalizar n-uplas como se fossem pares ordenados tomados iteradamente: assim, uma tripla (x_0, x_1, x_2) seria, formalmente, $((x_0, x_1), x_2)$. $\triangle$

Entre outras coisas, o Exercício 0.59, que você provavelmente já fez, assegura que se existem funções injetoras $A \to C$ e $B \to D$, então existe uma função injetora $A^B \to C^D$, i.e., há uma correspondência injetiva que associa a cada função do tipo $B \to A$ outra função do tipo $C \to D$. Consequentemente:

$$|\mathbb{R}| = \left|\{0,1\}^{\mathbb{N}}\right| \leq \left|\mathbb{N}^{\mathbb{N}}\right| \leq \left|\left(\{0,1\}^{\mathbb{N}}\right)^{\mathbb{N}}\right|,$$

onde a primeira desigualdade segue pois existe função injetora $\{0,1\} \to \mathbb{N}$, enquanto a segunda decorre da existência de função injetora $\mathbb{N} \to \{0,1\}^{\mathbb{N}}$.

Mas você também fez o Exercício 0.61 e, por isso, sabe que existe bijeção entre as funções da forma $Y \times Z \to X$ e as funções da forma $Z \to X^Y$. Em particular, temos daí

$$|\mathbb{R}| \leq |\mathbb{R}^n| = \left|\left(\{0,1\}^{\mathbb{N}}\right)^n\right| = \left|\left(\{0,1\}^{\mathbb{N}}\right)^{\mathbb{N}_{<n}}\right| = \left|\{0,1\}^{\mathbb{N}\times\mathbb{N}_{<n}}\right| \text{ e } \left|\left(\{0,1\}^{\mathbb{N}}\right)^{\mathbb{N}}\right| = \left|\{0,1\}^{\mathbb{N}\times\mathbb{N}}\right|$$

para qualquer $n \in \mathbb{N}$ com $n > 0$. Por fim, segue do Exercício 0.60 (que é consequência direta do Exercício 0.59) que X^Y e X^Z estão em bijeção sempre que Y e Z estão em bijeção.

[†]É a postura tomada por Bourbaki [4, 5], por exemplo. Se quiser uma releitura em português, confira [28].

Portanto,

$$|\mathbb{R}| \leq |\mathbb{R}^n| \leq \left|\{0,1\}^{\mathbb{N}\times\mathbb{N}_{<n}}\right| = \left|\{0,1\}^{\mathbb{N}}\right| \leq \left|\left(\{0,1\}^{\mathbb{N}}\right)^{\mathbb{N}}\right| = \left|\{0,1\}^{\mathbb{N}\times\mathbb{N}}\right| = \left|\{0,1\}^{\mathbb{N}}\right| = |\mathbb{R}|$$

pois $\mathbb{N}$, $\mathbb{N}\times\mathbb{N}$ e $\mathbb{N}\times\mathbb{N}_{<n}$ têm todos a mesma cardinalidade para qualquer $n \in \mathbb{N}$ com $n > 0$ (por quê?)$^{\star}_{\star}$.

Corolário 0.10.14. *$\mathbb{R}$ e $\mathbb{R}^n$ têm a mesma cardinalidade para todo $n \in \mathbb{N}$ com $n > 0$.*

Corolário 0.10.15. *$\mathbb{R}$ e $\mathbb{R}^{\mathbb{N}}$ têm a mesma cardinalidade. Explicitamente: existem tantas funções da forma $\mathbb{N} \to \mathbb{R}$ quanto números reais.*

Exercício 0.118 $({}^{\star}_{\star\star})$**.** Mostre que $|\mathbb{R}| < \left|\mathbb{R}^{\mathbb{R}}\right|$. ■

§3 Opcional: indução na prática

Se você não tem familiaridade com indução, é bem provável que a *estrutura indutiva* das provas por indução pareça mais complicada do que realmente é. Para sanar isso, o ideal é cursar alguma disciplina introdutória em que se discutam argumentos desse tipo "na prática". Porém, como você pode ter caído de paraquedas em Análise[†], esta breve subsubseção traz alguns exemplos mais detalhados a fim de diminuir seus prejuízos adiante. Mas lembre-se: aprender os pré-requisitos é um problema seu.

A primeira coisa a salientar se refere à notação usualmente empregada para estruturar argumentos por indução. Por exemplo: *se $\mathcal{P}(n)$ é uma afirmação acerca de um número natural n qualquer tal que se verifiquem*

- *a validade de $\mathcal{P}(0)$, e*
- *a validade de $\mathcal{P}(n+1)$ sempre que $\mathcal{P}(n)$ vale,*

então $\mathcal{P}(n)$ vale para todo $n \in \mathbb{N}$.

O que raios é esse $\mathcal{P}(n)$? É uma função? É um polinômio? É um pesadelo? A resposta está no próprio enunciado: é uma *afirmação*! Na prática, $\mathcal{P}(n)$ é uma *sentença matemática* acerca de n, sobre a qual é possível atribuir um valor lógico (verdadeiro ou falso).

Exemplo 0.10.16. Vamos provar que a identidade

$$0^2 + 1^2 + \ldots + n^2 = \frac{n(n+1)(2n+1)}{6}$$

vale para todo $n \in \mathbb{N}$.

Para empregar o esquema de prova indutiva, precisamos identificar o que é "$\mathcal{P}(n)$" aqui. Neste caso, $\mathcal{P}(n)$ é a identidade

$$0^2 + 1^2 + \ldots + n^2 = \frac{n(n+1)(2n+1)}{6}.$$

Veja bem:

[†]Por ser uma pessoa autoditada que decidiu aprender Análise sem pré-requisitos, por exemplo. Ou por ser estudante de um curso de Matemática que avançou mais do que deveria sem ter aprendido o que precisava. Ou por vários outros motivos que não mudam o fato de que não se constrói um prédio a partir do teto.

- $\mathcal{P}(n)$ **não** é a expressão $0^2 + 1^2 + \ldots + n^2$;
- $\mathcal{P}(n)$ **não** é a expressão $\dfrac{n(n+1)(2n+1)}{6}$;
- $\mathcal{P}(n)$ **é** a afirmação de que as duas expressões acima são iguais.

Isto traz um segundo problema: o que é $\mathcal{P}(0)$? Uma resposta correta, porém confusa para quem ainda não se acostumou com os abusos de notação matemática, poderia ser "$\mathcal{P}(0)$ é $\mathcal{P}(n)$ com 0 no lugar de n". No entanto, ao seguir tal resposta displicente ao pé da letra, é possível que quem não tenha familiaridade com recursão entenda que $\mathcal{P}(0)$ é a afirmação

$$0^2 + 1^2 + \ldots + 0^2 = \frac{0(0+1)(2\cdot 0+1)}{6} = 0.$$

O problema na identidade acima não está do lado direito, mas sim do lado esquerdo: o que significa "$0^2 + 1^2 + \ldots + 0^2$"? Resposta: nada. Eis o que está acontecendo: a expressão "$0^2 + 1^2 + \ldots + n^2$" é uma abreviação, um abuso de notação, um modo mais amigável de escrever

$$\sum_{i=0}^{n} i^2,$$

que por sua vez expressa um algoritmo recursivo de soma que depende do número n.

- Para $n \coloneqq 0$, $\displaystyle\sum_{i=0}^{0} i^2 = 0^2 = 0$.
- Para $n \coloneqq 1$, $\displaystyle\sum_{i=0}^{1} i^2 = 0^2 + 1^2 = 1$.
- Para $n \coloneqq 2$, $\displaystyle\sum_{i=0}^{2} i^2 = 0^2 + 1^2 + 2^2 = 1 + 4 = 5$.
- Para $n \coloneqq 3$, $\displaystyle\sum_{i=0}^{3} i^2 = 0^2 + 1^2 + 2^2 + 3^2 = \left(\sum_{i=0}^{2} i^2\right) + 3^2$.
- Para $n \coloneqq k+1$, $\displaystyle\sum_{i=0}^{k+1} i^2 = \left(\sum_{i=0}^{k} i^2\right) + (k+1)^2$.

Então sim, para decidir o que é $\mathcal{P}(0)$, basta substituir 0 no lugar de n em $\mathcal{P}(n)$, mas tal substituição não se faz ao pé da letra. Uma vez entendido isso, o resto se torna álgebra elementar:

- verifica-se $\mathcal{P}(0)$ pois $\dfrac{0(0+1)(2\cdot 0+1)}{6} = 0$;
- assumindo $\mathcal{P}(n)$ válida, temos

$$\underbrace{0^2+1^2+\ldots+n^2}_{(A)}+(n+1)^2 = \underbrace{\frac{n(n+1)(2n+1)}{6}}_{(B)}+\frac{6(n+1)^2}{6}$$
$$= \frac{(n+1)(n(2n+1)+6(n+1))}{6}$$
$$= \frac{(n+1)(2n^2+7n+6)}{6}$$
$$= \frac{(n+1)(n+2)(2n+3)}{6}$$
$$= \frac{(n+1)((n+1)+1)(2(n+1)+1)}{6},$$

onde a igualdade entre (A) e (B) se deve, precisamente, por termos assumido a validade de $\mathcal{P}(n)$; ocorre que esta longa cadeia de identidades assegura a validade da afirmação $\mathcal{P}(n+1)$. Acabou[†]. ▲

Exercício 0.119 $(^{\star}_{\star})$**.** Prove que

$$2+5+\ldots+(2+3n) = \frac{(n+1)(4+3n)}{2}$$

para todo $n \in \mathbb{N}$. Dica: perceba, antes de qualquer outra coisa, que $2 = 2+3\cdot 0$, $5 = 2+3\cdot 1$, $8 = 2+3\cdot 2$, etc. ■

Observação 0.10.17. Na prática, também é comum que o primeiro caso não seja $n \coloneqq 0$, mas sim algum outro número natural (1, ou 2, ou 3, etc.). Em tais situações, o raciocínio é análogo: prova-se $\mathcal{P}(k)$, onde k é o primeiro número para o qual a afirmação vale e, em seguida, prova-se que $\mathcal{P}(n+1)$ vale sempre que $\mathcal{P}(n)$ vale para algum $n \geq k$. △

Exercício 0.120 $(^{\star}_{\star})$**.** Prove que $2^0+2^1+\ldots+2^{n-1} = 2^n-1$ para todo $n \in \mathbb{N}\setminus\{0\}$. Dica: neste caso, o primeiro passo é $n \coloneqq 1$. ■

Exercício 0.121 (?!)**.** Qual o problema da afirmação acima com $n \coloneqq 0$? ■

Exemplo 0.10.18. Vamos provar que 2 divide n^2+n para qualquer $n \in \mathbb{N}$.

Aqui, $\mathcal{P}(n)$ é a afirmação "2 divide n^2+n". Vejamos:

- $\mathcal{P}(0)$ vale pois 2 divide 0^2+0, haja vista que $0^2+0 = 0 = 0\cdot 2$;
- se $\mathcal{P}(n)$ vale, então 2 divide n^2+n, ou seja, existe $k \in \mathbb{N}$ com $2k = n^2+n$; consequentemente,

  $$(n+1)^2+(n+1) = n^2+2n+1+n+1 = (n^2+n)+2n+2 = 2k+2n+2 = 2(k+n+1),$$

  mostrando que 2 também divide $(n+1)^2+(n+1)$, i.e., vale $\mathcal{P}(n+1)$.

Acabou. ▲

Exercício 0.122 $(^{\star}_{\star})$**.** Prove que para todo $n \in \mathbb{N}$, n é múltiplo de 2 ou $n+1$ é múltiplo de 2. Use isso para obter uma demonstração alternativa para a afirmação do exemplo anterior. Dica: tome cuidado com o uso adequado do "ou", já que aqui $\mathcal{P}(n)$ é "2 divide n ou 2 divide $n+1$". ■

[†]A não ser que você tenha encontrado problemas na manipulação algébrica, mas daí este é outro pré-requisito que você precisa dar conta de aprender, independentemente do que você escutar nas suas seções de terapia, rodas de conversa, etc.

Exemplo 0.10.19 (Influência do passo indutivo)**.** Considere o problema de provar que $2^n > n$ para todo $n \in \mathbb{N}$. É claro que $2^0 = 1 > 0$, mostrando que o caso inicial vale. No entanto, ao assumir $2^n > n$ para algum $n \in \mathbb{N}$, é comum que se tente provar que $2^{n+1} > n + 1$ observando que

$$2^{n+1} := 2 \cdot 2^n > 2 \cdot n = n + n.$$

No entanto, pelo modo como n foi tomado, a desigualdade acima nos dá apenas $n + n \geq n$ posto que $n \geq 0$ e, portanto, $2^{n+1} > n$, o que *ainda* não é a desigualdade desejada. Há um jeito maroto de corrigir o problema: se em vez de $n \geq 0$ tivéssemos $n \geq 1$, a mesma linha de raciocínio levaria a $2^{n+1} > n + n \geq n + 1$, como desejado. E para acrescentar tal suposição sobre n, basta verificar que $2^1 > 1$, o que precisa ser feito pois tal desigualdade não decorre de $2^0 > 0$ com o método apresentado acima.

Como de costume, há alternativas. Primeiro, pode-se mostrar por indução que a *sequência* $(2^n)_{n\in\mathbb{N}}$ é estritamente crescente, i.e., $2^m < 2^n$ sempre que $m < n$. Uma vez em posse disso, observe que ao assumir $2^n > n$ como hipótese de indução para algum $n \in \mathbb{N}$, resulta que $2^{n+1} > 2^n$ pela primeira parte (pois $n + 1 > n$), enquanto $2^n \geq n + 1$ (percebeu?)*, acarretando $2^{n+1} > n + 1$, como desejado. ▲

Exercício 0.123 $(^{\star}_{\star})$**.** Mostre que $(2^n)_{n\in\mathbb{N}}$ é estritamente crescente. Dica: considere como $\mathcal{P}(n)$ a afirmação "$2^m < 2^n$ para todo $m \in \mathbb{N}$ tal que $m < n$"; note que $\mathcal{P}(0)$ vale por vacuidade, já que não há $m \in \mathbb{N}$ com $m < 0$. ■

Exemplo 0.10.20 (Importância do caso inicial)**.** É comum negligenciar o passo inicial, mas ele é importante. Por exemplo, ao considerar a afirmação $\mathcal{P}(n)$ como sendo "n^2+5n+1 é múltiplo de 2", verifica-se sem grandes esforços que $\mathcal{P}(n) \Rightarrow \mathcal{P}(n + 1)$ para todo $n \in \mathbb{N}$: **se** $n^2 + 5n + 1 = 2k$, **então** $(n + 1)^2 + 5(n + 1) + 1 = 2(k + n + 3)$. No entanto, a afirmação $\mathcal{P}(n)$ é falsa para todo $n \in \mathbb{N}$ (verifique!)$^{\star}_{\star}$. Moral da história: enquanto não souber cozinhar, siga a receita! ▲

0.11 Exercícios adicionais

Adiante, $\mathbb{K}$ denota um corpo ordenado qualquer, enquanto $\mathbb{R}$ denota *a* reta real.

Essenciais

Exercício 0.124 $(^{\star}_{\star})$**.** Seja $A \subseteq \{x \in \mathbb{K} : x > 0_{\mathbb{K}}\}$ com $A \neq \emptyset$. Mostre que A é ilimitado superiormente se, somente se, $\inf \left\{a^{-1} : a \in A\right\} = 0_{\mathbb{K}}$. ■

Exercício 0.125 $(^{\star}_{\star})$**.** Mostre que se $\mathbb{K}$ é arquimediano, então $\inf\limits_{n\in\mathbb{N}} (2_{\mathbb{K}})^{-n} = 0_{\mathbb{K}}$. Dica: primeiro, mostre que $\{(2_{\mathbb{K}})^n : n \in \mathbb{N}\}$ é ilimitado superiormente em $\mathbb{K}$. ■

Exercício 0.126 $(^{\star}_{\star})$**.** Mostre que

$$\frac{|x+y|_{\mathbb{K}}}{1_{\mathbb{K}} + |x+y|_{\mathbb{K}}} \leq \frac{|x|_{\mathbb{K}}}{1_{\mathbb{K}} + |x|_{\mathbb{K}}} + \frac{|y|_{\mathbb{K}}}{1_{\mathbb{K}} + |y|_{\mathbb{K}}}$$

para quaisquer $x, y \in \mathbb{K}$. Dica: $1_{\mathbb{K}} - 1_{\mathbb{K}} = 0_{\mathbb{K}}$:p ■

Exercício 0.127 $(^{\star}_{\star})$**.** Sejam $\delta \in \mathbb{K}$ e $n \in \mathbb{N}$. Mostre que se $\delta > -1_{\mathbb{K}}$ e $n > 0$, então vale a *desigualdade de Bernoulli*: $(1_{\mathbb{K}} + \delta)^n \geq 1_{\mathbb{K}} + n_{\mathbb{K}}\delta$. Dica: $1_{\mathbb{K}} + \delta > 0_{\mathbb{K}}$ e $n\delta^2 \geq 0_{\mathbb{K}}$. ■

Exercício 0.128 $(^{\star}_{\star})$**.** Dados $a, b \in \mathbb{R}$ com $a < b$, mostre que $|(a, b)| = |\mathbb{R}|$. ■

Exercício 0.129 $(\star)$**.** Dados $a, b \in \mathbb{R}$ com $a < b$, qual a cardinalidade de $\mathbb{Q} \cap (a, b)$? ■

Exercício 0.130 $(\star)$**.** Dados $x, y \in \mathbb{R}$ com $x < y$, mostre que existe um número real $z \in \mathbb{R} \setminus \mathbb{Q}$ com $x < z < y$. ■

Exercício 0.131 $(^{\star}_{\star})$**.** Mostre que $\varphi\colon (-1, 1) \to \mathbb{R}$ dada por $\varphi(x) := \dfrac{x}{1 - |x|}$ é bijetora. ■

Exercício 0.132 $(^{\star}_{\star})$**.** Mostre que $[a, b]$ e (a, b) têm a mesma cardinalidade para quaisquer $a, b \in \mathbb{R}$ com $a < b$. Conclua que $[-\infty, +\infty] := \mathbb{R} \cup \{-\infty, +\infty\}$ e $\mathbb{R}$ têm a mesma cardinalidade. ■

Exercício 0.133 $(^{\star}_{\star})$**.** Mostre que se $x^2 \leq y^2$, para $x, y \in \mathbb{R}$, então $|x| \leq |y|$. ■

Exercício 0.134 $(^{\star}_{\star})$**.** Assuma que para todo $r \in \mathbb{R}$ com $r \geq 0$ exista um único $\alpha \geq 0$ tal que $\alpha^2 = r$. O número α costuma ser indicado por $\sqrt{r}$, a **raiz quadrada de** r

a) Mostre que se $x, y \geq 0$, então deve ocorrer $\sqrt{xy} \leq \dfrac{x + y}{2}$. Dica: $\gamma^2 \geq 0$ para todo $\gamma \in \mathbb{R}$.

b) Com respeito ao item anterior: em que situações pode-se garantir a igualdade? Dica: observe que $\sqrt{x}^2 = x$ para todo $x \geq 0$.

c) Mostre que $\sqrt{x^2} = |x|$ para todo $x \in \mathbb{R}$.

d) Mostre que se $\alpha > \beta \geq 0$, então $\sqrt{\alpha} > \sqrt{\beta}$ ■

Exercício 0.135 $(^{\star}_{\star\star})$**.** Prove que $\sqrt{r}$ existe para todo $r \geq 0$. Dica: para a unicidade, lembre-se de que $\alpha^2 - \beta^2 = (\alpha - \beta)(\alpha + \beta)$; para a existência, imite o que se fez para $r := 2$ (cf. Exemplo 0.9.2), onde pode ser conveniente tratar separadamente os casos em que $0 < r < 1$ e $r \geq 1$. ■

Exercício 0.136 $(^{\star}_{\star\star})$**.** Mostre que vale a recíproca do Exercício 0.107. ■

Extras

Exercício 0.137 $(^{\star}_{\star\star})$**.** Para uma ordem parcial $(\mathbb{P}, \leq)$, mostre que são equivalentes:

(i) todo subconjunto não vazio de $\mathbb{P}$ e limitado superiormente admite supremo;

(ii) todo subconjunto não vazio de $\mathbb{P}$ e limitado inferiormente admite ínfimo.

Dica: note que se $A \subseteq \mathbb{P}$ é subconjunto não vazio e limitado inferiormente, então o conjunto $B := \{p \in \mathbb{P} : p$ é limitante inferior de $A\}$ é não vazio e limitado superiormente (pelos membros de A!). ■

Exercício 0.138 $(^{\star}_{\star})$**.** Mostre que $\mathbb{K}$ é completo se, e somente se, todo subconjunto não vazio e limitado inferiormente tem ínfimo. Dica: encare o exercício anterior até que ele te encare de volta. ■

Exercício 0.139 $(^{\star}_{\star\star})$**.** Sejam A e B subconjuntos de uma ordem parcial $(\mathbb{P}, \leq)$.

a) Mostre que se $\sup A$ e $\sup B$ existem e, para todo $a \in A$ existe $b \in B$ tal que $a \leq b$, então $\sup A \leq \sup B$. Enuncie e demonstre a versão dual para ínfimos.

b) Fazendo $\mathbb{P} := \mathbb{R}$, conclua que se $\emptyset \neq A \subseteq B$ e B tem supremo, então A tem supremo e vale $\sup A \leq \sup B$. Analogamente, se B tem ínfimo, então A tem ínfimo e $\inf A \geq \inf B$. ■

Exercício 0.140 ($^{\star}_{\star\star}$)**.** Uma **norma** num $\mathbb{R}$-espaço vetorial E é uma função $\|\cdot\|\colon E \to \mathbb{R}$ satisfazendo as condições a seguir para quaisquer $x, y \in E$ e $\lambda \in \mathbb{R}$:

(i) $\|x\| \geq 0$;

(ii) $\|x\| = 0$ se, e somente se, $x = 0$;

(iii) $\|\lambda x\| = |\lambda|\|x\|$;

(iv) $\|x + y\| \leq \|x\| + \|y\|$.

Para um conjunto não vazio X qualquer, mostre que a função $\|\cdot\|_\infty\colon \mathcal{B}(X, \mathbb{R}) \to \mathbb{R}$ dada por $\|f\|_\infty := \sup\{|f(x)| : x \in X\}$ define uma norma no $\mathbb{R}$-espaço vetorial $\mathcal{B}(X, \mathbb{R})$ das funções limitadas da forma $X \to \mathbb{R}$ (cf. Subseção 0.7.1 §0). Costuma-se chamar $\|\cdot\|_\infty$ de **norma do supremo**. ■

Capítulo 1

Limites e continuidade

O capítulo anterior apresentou e discutiu a reta real enquanto objeto matemático formal, enfatizando seus aspectos algébricos e de ordem. O próximo passo é entender como tal objeto pode ser usado na "prática": neste capítulo, vamos dar sentido às noções de *limite* na reta (e noutros ambientes similares), e introduziremos as funções que são compatíveis com tais limites — as funções *contínuas*.

1.0 Noções de convergência

1.0.0 Essencial

§0 Limites de sequências e funções (motivação)

Em Cálculo I você provavelmente aprendeu a noção de *limite* de uma função real. Por exemplo, para uma função $f\colon \mathbb{R} \to \mathbb{R}$ e números reais $p, L \in \mathbb{R}$, escreve-se

$$\lim_{x \to p} f(x) = L, \tag{1.0}$$

para dizer que L é *um* **limite real da função** f **quando** x **tende a** p, o que abrevia o seguinte: *para todo* $\varepsilon > 0$ *existe* $\delta > 0$ *tal que* $|f(x) - L| < \varepsilon$ *sempre que* $0 < |x - p| < \delta$.

Por sua vez, se você fez Cálculo II no Bacharelado em Matemática da UESC, então já deve ter se deparado com a noção de *sequência convergente*[‡]. Primeiro, lembre-se de que uma **sequência real** $(x_n)_{n\in\mathbb{N}}$ é, meramente, uma função $s\colon \mathbb{N} \to \mathbb{R}$ onde $s(n) \coloneqq x_n$ para todo n. Daí, para uma sequência $(x_n)_{n\in\mathbb{N}}$ e um número real L, escreve-se

$$\lim_{n \to \infty} x_n = L, \tag{1.1}$$

para dizer que L é *um* **limite real da sequência** $(x_n)_{n\in\mathbb{N}}$ (e alguns ainda insistem em dizer "quando n tende a infinito"). Como no caso de funções, a notação acima também abrevia uma condição um pouco mais técnica: *para todo* $\varepsilon > 0$ *existe* $N \in \mathbb{N}$ *tal que* $|x_n - L| < \varepsilon$ *sempre que* $n \geq N$.

Intuitivamente, tanto em "$\lim\limits_{x\to p} f(x) = L$" quanto em "$\lim\limits_{n\to\infty} x_n = L$", expressa-se a ideia de que conforme elementos num conjunto de índices *avançam* numa *direção* específica ("$x \to p$" no primeiro caso e "$n \to \infty$" no segundo), as imagens indexadas ($f(x)$ no primeiro caso e x_n no segundo) se *aproximam* do limite L. No fundo, ambos são subcasos de um fenômeno mais geral. Chegaremos lá...

[‡]Se fez e não viu, a culpa não é minha.

Exemplo 1.0.0. *Existe no máximo* um *número real* $L \in \mathbb{R}$ *que satisfaz a condição abreviada por* (1.0). Mais precisamente, se $L, L' \in \mathbb{R}$ satisfazem a mesma condição, então $L = L'$. Ora, a fim de mostrar isso, basta garantir que $|L - L'| < \varepsilon$ para todo $\varepsilon > 0$ (cf. Teorema 0.7.9). Tendo em vista que a desigualdade triangular acarreta $|L - L'| \leq |L - f(x)| + |f(x) - L'|$ para qualquer $x \in \mathbb{R}$, convém utilizar a condição satisfeita por L e L', que permite tomar $\delta, \delta' > 0$ tais que

$$0 < |x - p| < \delta \Rightarrow |f(x) - L| < \frac{\varepsilon}{2} \quad \text{e} \quad 0 < |x - p| < \delta' \Rightarrow |f(x) - L'| < \frac{\varepsilon}{2}.$$

Certamente, existe algum $x' \in \mathbb{R}$ satisfazendo as duas condições com respeito a δ e δ': basta tomar x' com $0 < |x' - p| < \min\{\delta, \delta'\}$, por exemplo. Logo,

$$|L - L'| \leq |L - f(x')| + |f(x') - L'| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

como desejado. ▲

Exemplo 1.0.1. *Existe no máximo* um *número real* $L \in \mathbb{R}$ *que satisfaz a condição abreviada por* (1.1). Mais precisamente, se $L, L' \in \mathbb{R}$ satisfazem a mesma condição, então $L = L'$. Novamente, basta assegurar que $|L - L'| < \varepsilon$ para todo $\varepsilon > 0$. Por um lado, para $\gamma > 0$ fixado, existe $N_L \in \mathbb{N}$ tal que $|x_n - L| < \gamma$ sempre que $n \geq N_L$. Por outro lado, também existe $N_{L'} \in \mathbb{N}$ tal que $|x_n - L'| < \gamma$ sempre que $n \geq N_{L'}$, pois $x_n \to L'$. Daí, *se* $N \geq \max\{N_L, N_{L'}\}$, segue que

$$|L - L'| \leq |L - x_N| + |x_N - L'| < \gamma + \gamma = 2\gamma$$

donde a afirmação original segue com $\gamma := \frac{\varepsilon}{2}$. ▲

No primeiro exemplo, o fato de podermos tomar um mesmo ponto x' cumprindo as duas exigências ($|x' - p| < \delta$ e $|x' - p| < \delta'$) permitiu concluir que tanto $|f(x') - L|$ quanto $|f(x') - L'|$ são *suficientemente pequenos*. Já no segundo exemplo, o número N escolhido satisfaz as duas exigências ($N \geq N_L$ e $N \geq N_{L'}$), o que permitiu concluir que tanto $|x_N - L|$ quanto $|x_N - L'|$ são *suficientemente pequenos*. Chega de rodeios.

§1 Limites de redes reais

Definição 1.0.2. Sejam $\mathbb{D}$ um conjunto e $\preceq$ uma relação binária sobre $\mathbb{D}$. Diremos que $(\mathbb{D}, \preceq)$ é uma **pré-ordem** se $\preceq$ for reflexiva e transitiva. Diremos que $\mathbb{D} \neq \emptyset$ é um **conjunto dirigido** pela pré-ordem $\preceq$ se, adicionalmente, valer a seguinte condição de *compatibilidade/refinamento*: para quaisquer $x, y \in \mathbb{D}$ existe $z \in \mathbb{D}$ com $x, y \preceq z$. ¶

Exemplo 1.0.3. *Ordens totais* não vazias são conjuntos dirigidos, já que quaisquer dois elementos x e y da ordem são majorados por $\max\{x, y\}$. Note que $(\mathbb{N}, \leq)$ é dirigido. ▲

Exemplo 1.0.4. Para um ponto $p \in \mathbb{R}$ fixado, sejam $\mathbb{R}_p := \mathbb{R} \setminus \{p\}$ e a relação binária $\preceq$ em $\mathbb{R}_p$ definida da seguinte forma: para $x, y \in \mathbb{R}_p$, vamos escrever $x \preceq y$ para indicar $|y - p| \leq |x - p|$. Não é difícil se convencer de que $\preceq$ é uma pré-ordem sobre $\mathbb{R}_p$ (verifique?)*. Para ver que $(\mathbb{R}_p, \preceq)$ é dirigido, apenas note que se $x, y \in \mathbb{R}_p$, então ocorre $x \preceq y$ ou $y \preceq x$, já que $|x - y|$ e $|y - x|$ são elementos de um conjunto totalmente ordenado. ▲

Exercício 1.0 ($\star$). Mostre que $(\mathbb{R}_p, \preceq)$ não é uma ordem parcial. ■

Uma **rede real**[†] é uma função da forma $\rho\colon \mathbb{D} \to \mathbb{R}$, onde $\mathbb{D}$ é um conjunto dirigido por alguma pré-ordem $\preceq$. Redes serão denotadas por $(x_d)_{d\in\mathbb{D}}$, $(x_d : d \in \mathbb{D})$ ou mesmo $(x_d)_d$ quando o conjunto dirigido $\mathbb{D}$ for claro pelo contexto[‡].

Definição 1.0.5. Um ponto $L \in \mathbb{R}$ será xingado de (*um*) **limite real da rede** $(x_d)_{d\in\mathbb{D}}$ se a seguinte condição for satisfeita: para todo $\varepsilon > 0$ existe $D \in \mathbb{D}$ tal que $|x_d - L| < \varepsilon$ para todo $d \succeq D$. Em tal situação, também se diz que $(x_d)_{d\in\mathbb{D}}$ **converge para** L em $\mathbb{R}$, o que se abrevia com a notação $x_d \to L$. Dizer que uma rede é **convergente em** $\mathbb{R}$ apenas indica que existe um limite real para o qual a rede converge. ¶

Observação 1.0.6 (Interpretação geométrica)**.** Lembre-se de que $x \in (L-\varepsilon, L+\varepsilon)$ se, e somente se, $|L - x| < \varepsilon$ (cf. Exercício 0.93). Assim, uma rede $(x_d)_{d\in\mathbb{D}}$ converge para $L \in \mathbb{R}$ se, e somente se, para todo intervalo da forma $(L-\varepsilon, L+\varepsilon)$ existe um índice $D \in \mathbb{D}$ *a partir do qual* $x_d \in (L-\varepsilon, L+\varepsilon)$, i.e., tal que $x_d \in (L-\varepsilon, L+\varepsilon)$ para todo $d \succeq D$. △

Note que ao ler "$a \preceq b$" como "b é melhor do que a", dizer que L é limite de $(x_d)_{d\in\mathbb{D}}$ passa a significar: para todo $\varepsilon > 0$, $|x_d - L| < \varepsilon$ para $d \in \mathbb{D}$ *suficientemente bom.*

Exercício 1.1 ($\star$)**.** Mostre que uma sequência real $(x_n)_{n\in\mathbb{N}}$ converge para $L \in \mathbb{R}$ (no sentido de sequência convergente) se, e somente se, $x_n \to L$ (no sentido de rede convergente). ■

Exemplo 1.0.7. De volta ao Exemplo 1.0.4, note que uma função $f\colon \mathbb{R} \to \mathbb{R}$ induz uma rede $(f(x))_{x\in\mathbb{R}_p}$ por meio da restrição a $\mathbb{R} \setminus \{p\}$: cada $x \in \mathbb{R}_p$ é associado a $f(x)$. Reciprocamente, para $a \in \mathbb{R}$ fixado, uma rede $(\eta(x))_{x\in\mathbb{R}_p}$ induz uma função $\eta_a\colon \mathbb{R} \to \mathbb{R}$ por meio da regra

$$\eta_a(x) := \begin{cases} a, & \text{se } x = p \\ \eta(x), & \text{caso contrário} \end{cases}.$$

Agora, *se* $L \in \mathbb{R}$ *for tal que* $\lim_{x\to p} f(x) = L$ *no sentido clássico do Cálculo I, então a rede* $(f(x))_{x\in\mathbb{R}_p}$ *converge para* L. Com efeito, para $\varepsilon > 0$ dado, a condição satisfeita por L assegura $\delta > 0$ tal que $|f(x) - L| < \varepsilon$ sempre que $0 < |x - p| < \delta$. Tomando $D \in \mathbb{R}_p$ tal que $0 < |D - p| < \delta$, segue que $|f(x) - L| < \varepsilon$ para qualquer $x \in \mathbb{R}_p$ *melhor* do que D, uma vez que $x \succeq D$ significa $|x - p| \leq |D - p|$, acarretando $0 < |x - p| < \delta$ (por transitividade).

Reciprocamente, *se uma rede* $(\eta(x))_{x\in\mathbb{R}_p}$ *converge para* $L \in \mathbb{R}$*, então qualquer função induzida* $\eta_a\colon \mathbb{R} \to \mathbb{R}$ *é tal que* $\lim_{x\to p} \eta_a(x) = L$: para $\varepsilon > 0$, existe um *momento* $D \in \mathbb{R}_p$ tal que $|\eta_a(x) - L| < \varepsilon$ sempre que $x \succeq D$; tomando $\delta := |D - p|$, segue que se $0 < |x - p| < \delta$, então $x \succeq D$, donde segue o resultado. ▲

Moral da história: limites de redes generalizam, simultaneamente, limites de sequências e limites de funções da forma $\mathbb{R} \to \mathbb{R}$. Veremos adiante que, na verdade, o escopo de redes é bem mais abrangente do que esta breve introdução sugere.

Exercício 1.2 ($\substack{\star\\\star}$)**.** Mostre que se uma rede real $(x_d)_{d\in\mathbb{D}}$ converge para números reais $L, L' \in \mathbb{R}$, então $L = L'$. ■

Definição 1.0.8. Também escreveremos $\lim\limits_{d\in\mathbb{D}} x_d$ para indicar o único número real $L \in \mathbb{R}$, se existir, tal que $x_d \to L$. ¶

Exercício 1.3 ($\substack{\star\\\star}$)**.** Mostre que uma rede real $(x_d)_{d\in\mathbb{D}}$ converge para $L \in \mathbb{R}$ se, e somente se, para todo intervalo aberto $I \subseteq \mathbb{R}$ com $L \in I$ existir $D \in \mathbb{D}$ tal que $x_d \in I$ sempre que $d \succeq D$. ■

[†] A terminologia usual, em inglês, é *net*.

[‡] Como de costume, entende-se que x_d indica a imagem de $d \in \mathbb{D}$ pela função ρ.

1.0.1 Extras

§0 Normas e produtos internos

O Exercício 0.140 introduziu a noção de *espaço vetorial normado*: trata-se de um $\mathbb{R}$-espaço vetorial em que se tem a habilidade de *calcular* ~~o tamanho~~ a *norma* dos vetores. Embora o exemplo apresentado tenha sido relativamente abstrato, há representantes bem mais corriqueiros dessa classe de animal:

(i) a própria reta real, enquanto $\mathbb{R}$-espaço vetorial, com a norma dada pelo seu valor absoluto $|\cdot|\colon \mathbb{R} \to [0, +\infty)$;

(ii) o *plano bidimensional* $\mathbb{R}^2 := \mathbb{R} \times \mathbb{R}$, seu velho conhecido dos tempos da Geometria Analítica, com a norma $\|\cdot\|_2\colon \mathbb{R}^2 \to [0, +\infty)$ que a cada vetor $(a, b) \in \mathbb{R}^2$ associa a *norma euclidiana* do vetor (a, b), definida como $\|(a, b)\|_2 := \sqrt{a^2 + b^2}$;

(iii) o *espaço tridimensional* $\mathbb{R}^3$, outro amigo dos tempos da Geometria Analítica, com a *norma euclidiana*[†] $\|\cdot\|_2\colon \mathbb{R}^3 \to [0, +\infty)$ que a cada vetor (a, b, c) associa a norma euclidiana do vetor (a, b, c), definida como $\|(a, b, c)\|_2 := \sqrt{a^2 + b^2 + c^2}$.

Em geral, para qualquer $n \in \mathbb{N}$ com $n > 0$, pode-se considerar sobre $\mathbb{R}^n$ a **norma euclidiana** $\|\cdot\|_2\colon \mathbb{R}^n \to [0, +\infty)$ que a cada n-upla $(a_1, \ldots, a_n) \in \mathbb{R}^n$ associa o número

$$\|(a_1, \ldots, a_n)\|_2 := \sqrt{a_1^2 + \ldots + a_n^2}.$$

Embora a intuição geométrica torne bastante razoável assumir que a regra acima defina realmente uma *norma*, a terminologia tem atrelada a ela certas condições que precisam ser verificadas para que possamos chamar a coisa de "norma" sem peso na consciência. As condições (i), (ii) e (iii) são corriqueiras:

(i) como $a^2 \geq 0$ para todo $a \in \mathbb{R}$, tem-se $a_1^2 + \ldots + a_n^2 \geq 0$ (certo?)*, o que assegura a boa definição do número $\sqrt{a_1^2 + \ldots + a_n^2}$, que é maior do que ou igual a zero, por definição (cf. Exercício 0.135);

(ii) se $\|(a_1, \ldots, a_n)\|_2 = 0$, então $a_1^2 + \ldots + a_n^2 = 0$ e, consequentemente $a_i^2 = 0$ para todo $i \leq n$ (por quê?!)*, donde segue que $a_i = 0$ para todo $i \leq n$; por outro lado, é claro que $\|(0, \ldots, 0)\|_2 = 0$;

(iii) para $\lambda \in \mathbb{R}$ e $(a_1, \ldots, a_n) \in \mathbb{R}^n$ quaisquer,

$$\begin{aligned} \|\lambda(a_1, \ldots, a_n)\|_2 &= \|(\lambda a_1, \ldots, \lambda a_n)\|_2 = \sqrt{\lambda^2 \cdot (a_1^2 + \ldots + a_n^2)} \\ &\underset{(!)}{=} \sqrt{\lambda^2}\sqrt{a_1^2 + \ldots + a_n^2} \underset{(A)}{=} |\lambda| \cdot \|(a_1, \ldots, a_n)\|_2, \end{aligned}$$

onde a igualdade (!) se deve à identidade $\sqrt{\alpha\beta} = \sqrt{\alpha}\sqrt{\beta}$, válida para quaisquer $\alpha, \beta \geq 0$ (por quê?!)*, enquanto (A) decorre do Exercício 0.134.

A desigualdade triangular (iv), porém, é mais delicada. Para verificá-la, vamos assumir, momentaneamente, a validade da desigualdade

$$|(a_1b_1 + \ldots + a_nb_n)| \leq \|(a_1, \ldots, a_n)\|_2 \cdot \|(b_1, \ldots, b_n)\|_2 \tag{1.2}$$

[†] *Spoiler alert*: o subíndice "$_2$" em "$\|\cdot\|_2$" faz referência aos expoentes envolvidos na expressão da norma, e não à *dimensão* do espaço.

para quaisquer $(a_1, \ldots, a_n), (b_1, \ldots, b_n) \in \mathbb{R}^n$. Para quê? Uma vez que para $\alpha, \beta \geq 0$ se tem $\alpha \leq \beta \Leftrightarrow \alpha^2 \leq \beta^2$ (certo?)*, a fim de mostrar

$$\underbrace{\|(a_1+b_1, \ldots, a_n+b_n)\|_2}_{B} \leq \underbrace{\|(a_1, \ldots, a_n)\|_2}_{C} + \underbrace{\|(b_1, \ldots, b_n)\|_2}_{D},$$

basta obter $B^2 \leq (C+D)^2$. E realmente:

$$\begin{aligned} B^2 &= \sum_{j=1}^n (a_j+b_j)^2 = \sum_{j=1}^n a_j^2 + 2a_jb_j + b_j^2 = C^2 + 2\sum_{j=1}^n a_jb_j + D^2 \\ &\underset{(a)}{\leq} C^2 + 2\left|\sum_{j=1}^n a_jb_j\right| + D^2 \underset{(b)}{\leq} C^2 + 2CD + D^2 = (C+D)^2, \end{aligned}$$

onde (a) vale em geral[†] e (b) decorre, justamente, da suposição (1.2).

Então... por que (1.2) vale? Secretamente, a norma euclidiana foi induzida por um tipo de função chamada de *produto interno*. Um **produto interno** num $\mathbb{R}$-espaço vetorial E é uma função $E \times E \to \mathbb{R}$ que a cada par de vetores $(x, y) \in E \times E$ associa um *escalar* $\langle x, y\rangle$, sujeito às seguintes condições, para quaisquer $x, y, z \in E$ e $\lambda \in \mathbb{R}$:

(I) $\langle x, y\rangle = \langle y, x\rangle$,

(II) $\langle \lambda x, y\rangle = \lambda \cdot \langle x, y\rangle$,

(III) $\langle x+y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$, e

(IV) $\langle x, x\rangle > 0$ se $x \neq 0_E$.

É um exercício simples ($\star$) verificar que a função $\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ que faz

$$((a_1, \ldots, a_n), (b_1, \ldots, b_n)) \mapsto a_1b_1 + \ldots + a_nb_n$$

é um produto interno em $\mathbb{R}^n$. O que traz a pergunta: e daí?

Lema 1.0.9 (Desigualdade de Cauchy-Schwarz)**.** *Se $\langle \cdot, \cdot \rangle \colon E \times E \to \mathbb{R}$ é um produto interno, então ao definir $\|u\| := \sqrt{\langle u, u\rangle}$ para todo $u \in E$, a desigualdade a seguir se verifica para quaisquer $x, y \in E$:*

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

Demonstração, atribuída a Paul Halmos.

$$0 \leq \left\| \frac{x}{\|x\|} \pm \frac{y}{\|y\|} \right\|^2 = 1 + 1 \pm 2\frac{\langle x, y\rangle}{\|x\|\|y\|} \Leftrightarrow \pm\langle x, y\rangle \leq \|x\|\|y\|,$$

pois $\|x\|\|y\| \geq 0$ e $\frac{\langle z, z\rangle}{\|z\|^2} = 1$ para qualquer $z \neq 0$. Os detalhes ficam por sua conta ($\star$). □

Exercício 1.4 ($\star$)**.** Prove a desigualdade (1.2). ■

Exercício 1.5 (Outras normas em $\mathbb{R}^n$ — ($^{\star}_{\star\star}$))**.** Para $(a_1, \ldots, a_n) \in \mathbb{R}^n$, defina os números

$$\|(a_1, \ldots, a_n)\|_1 := |a_1| + \ldots + |a_n| \quad \text{e} \quad \|(a_1, \ldots, a_n)\|_\infty := \max\{|a_1|, \ldots, |a_n|\}.$$

Mostre que as funções $\|\cdot\|_1 \colon \mathbb{R}^n \to \mathbb{R}$ e $\|\cdot\|_\infty \colon \mathbb{R}^n \to \mathbb{R}$ são normas em $\mathbb{R}^n$. ■

[†]Pois $\alpha \leq |\alpha|$ para todo $\alpha \in \mathbb{R}$.

Definição 1.0.10. Costuma-se chamar $\|\cdot\|_1$ de **norma da soma**, enquanto $\|\cdot\|_\infty$, neste contexto, é a **norma do máximo**. ¶

Toda essa ladainha deve ter te convencido de que espaços vetoriais normados são bastante comuns e, justamente por isso, não há razão para temê-los. Com isso dito, ao rever a definição de limite de rede, cabe perguntar: por que considerar apenas redes *reais*? Por exemplo, em $\mathbb{R}^2$, poderíamos dizer que uma sequência $((x_n, y_n))_{n\in\mathbb{N}}$ de *pontos* em $\mathbb{R}^2$ *converge* para um *ponto* $(x, y) \in \mathbb{R}^2$ se para todo $\varepsilon > 0$ existir $N \in \mathbb{N}$ tal que $\|(x_n, y_n) - (x, y)\|_2 < \varepsilon$ sempre que $n \geq N$. Reescrevendo com notação de *gente grande*[†]: uma sequência $(x_n)_{n\in\mathbb{N}}$ em $\mathbb{R}^2$ *converge* para $x \in \mathbb{R}^2$ se para todo $\varepsilon > 0$ existir $N \in \mathbb{N}$ tal que $\|x_n - x\|_2 < \varepsilon$ sempre que $n \geq N$. Ou seja: é o <u>mesmo</u> conceito de sequência convergente em $\mathbb{R}$, e com uma grafia essencialmente idêntica. Isto motiva (que se tente) adaptar tudo o que se fez anteriormente em $\mathbb{R}$ para o contexto de espaços normados! Mas faremos algo um pouco *melhor*.

§1 Espaços métricos

Enquanto uma norma $\|\cdot\|$ num $\mathbb{R}$-espaço vetorial E permite determinar o *tamanho* dos vetores de E, a função

$$\begin{aligned} d_{\|\cdot\|} \colon E \times E &\to \mathbb{R} \\ (x, y) &\mapsto \|x - y\| \end{aligned}$$

define uma *noção de distância*, já que ela calcula o "tamanho" do vetor que "liga" o *ponto* x ao ponto y (Geometria Analítica, lembra?). Note que para $x, y, z \in E$ quaisquer:

(i) $d_{\|\cdot\|}(x, y) \geq 0$ (não há distâncias negativas!);

(ii) $d_{\|\cdot\|}(x, y) = 0$ se, e somente se, $x = y$ (pontos distintos têm distância estritamente positiva);

(iii) $d_{\|\cdot\|}(x, y) = d_{\|\cdot\|}(y, x)$ (a distância entre Maomé e a montanha...);

(iv) $d_{\|\cdot\|}(x, y) \leq d_{\|\cdot\|}(x, z) + d_{\|\cdot\|}(z, y)$ (desigualdade triangular).

Uma das vantagens da formulação acima, que não é óbvia num primeiro momento, é a seguinte: não se mencionam quaisquer operações algébricas em E, mas apenas em $\mathbb{R}$. Isto sugere a possibilidade de discutir problemas que envolvem distâncias mesmo em contextos desprovidos de uma estrutura vetorial.

Definição 1.0.11. Dizemos que uma função $d \colon X \times X \to \mathbb{R}$ é uma **métrica** em X se as condições (i), (ii), (iii) e (iv) acima forem satisfeitas para quaisquer $x, y, z \in X$. O par (X, d) será chamado de **espaço métrico** — e diremos apenas que "X é espaço métrico" quando a métrica for clara pelo contexto. ¶

Percebe-se assim que o *aparato algébrico* empregado na definição dos limites de redes reais era apenas uma distração: no número "$|x_d - L|$" que se busca *minimizar* conforme o índice d *melhora*, a ocorrência da subtração é apenas um artifício para expressar a *distância* entre x_d e L. Dessa forma, a seguinte definição é inevitável:

[†]Ao escrever "$x \in \mathbb{R}^2$", a pessoa adulta percebe que x indica um par ordenado e não um número real.

Definição 1.0.12. Para um espaço métrico (X, d), uma *rede em* X é uma função da forma $\eta\colon \mathbb{A} \to X$ onde $\mathbb{A}$ é algum conjunto dirigido[†] por uma pré-ordem $\preceq$. Diremos que $L \in X$ é (*um*) **limite** (em X) **da rede** $(x_a)_{a\in\mathbb{A}}$ se a seguinte condição for satisfeita: para todo $\varepsilon > 0$ existe $A \in \mathbb{A}$ tal que $d(x_a, L) < \varepsilon$ para todo $a \succeq A$. Em tais situações, diremos que $(x_a)_{a\in\mathbb{A}}$ **converge para** $L \in X$, o que será abreviado com a notação $x_a \to L$. Dizer que uma rede em X é *convergente em* X apenas indica que existe um limite em X para a rede. ¶

Embora pareça mais complicado, na verdade, trata-se de uma simplificação: com os mesmos argumentos que já utilizávamos, obtemos resultados que abrangem mais casos! Por exemplo:

Proposição 1.0.13. *Se $(x_a)_a$ é uma rede num espaço métrico X tal que $x_a \to L$ e $x_a \to L'$ para certos $L, L' \in X$, então $L = L'$.*

Demonstração. A fim de provar que $L = L'$, basta mostrar que $d(L, L') = 0$ (pela condição (ii) na definição de espaço métrico) o que, por sua vez, pode ser alcançado mostrando que $d(L, L') < \varepsilon$ para qualquer $\varepsilon > 0$ tomado. Ora, fixando-se algum $\varepsilon > 0$, a desigualdade triangular (para métricas) assegura

$$d(L, L') \leq d(L, x_a) + d(x_a, L')$$

para qualquer a no domínio da rede $(x_a)_a$. Em particular, por $x_a \to L$ e $x_a \to L'$, existem α e α' no conjunto de índices tais que $d(x_a, L) < \frac{\varepsilon}{2}$ sempre que $a \succeq \alpha$, enquanto $d(x_a, L') < \frac{\varepsilon}{2}$ sempre que $a \succeq \alpha'$. Assim, tomando-se $\alpha'' \succeq \alpha, \alpha'$, que existe pois o domínio da rede é dirigido, obtém-se

$$d(L, L') \leq d(L, x_{\alpha''}) + d(x_{\alpha''}, L') = d(x_{\alpha''}, L) + d(x_{\alpha''}, L') < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

como queríamos. □

Como este curso é voltado para Análise na Reta, você não precisa se preocupar muito com espaços métricos... ~~pelo menos por enquanto.~~ No entanto, conforme as discussões na reta avançarem, naturalmente nosso entendimento sobre espaços métricos e normados avançará, mesmo que subconscientemente.

§2 Exercícios extras

Exercício 1.6 $(^{\star}_{\star})$**.** Sejam (X, d) um espaço métrico e $Y \subseteq X$ um subconjunto. Para $a, b \in Y$, defina $\rho(a, b) \coloneqq d(a, b)$.

a) Mostre que ρ é uma métrica em Y.

b) Para $a \in Y$ e $r > 0$, mostre que $B_\rho(a, r) = B_d(a, r) \cap Y$.

c) Seja $X \coloneqq \mathbb{R}$ e considere d como sendo a métrica usual de $\mathbb{R}$. Determine $B_d(0, \frac{1}{2})$.

d) Para $Y \coloneqq [0, 1)$, determine $B_\rho(0, \frac{1}{2})$. ■

[†]Embora a letra "$\mathbb{D}$" faça alusão a "$\mathbb{D}$irigido", usá-la agora nos *obrigaria* a utilizar a letra "d" para denotar seus elementos, o que entraria em conflito com o uso da mesma letra para nos referirmos à ~~distância~~ *métrica* do espaço. Indicar a métrica pela letra "m" seria ainda pior, já que m é um número natural por *aclamação popular*.

Exercício 1.7 $(\star)$**.** Mostre que se $\langle \cdot, \cdot \rangle$ é um produto interno sobre um espaço vetorial E (cf. Subseção 1.0.1 §0) e $\| \cdot \|$ é a norma induzida pelo produto interno, então se verifica a *identidade do paralelogramo*:

$$\|x+y\|^2 + \|x-y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right)$$

para quaisquer $x, y \in E$. ■

Exercício 1.8 $({}^{\star}_{\star})$**.** Mostre que se $\| \cdot \|$ é uma norma sobre E que satisfaz a identidade do paralelogramo, então existe um único produto interno $\langle \cdot, \cdot \rangle$ cuja norma induzida é $\| \cdot \|$. Dica: faça $\langle x, y \rangle := \dfrac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2\right)$ e prove que isto funciona. ■

Exercício 1.9 $(\star)$**.** Num espaço normado $(E, \| \cdot \|)$, mostre que $\|x-y\| \geq \|x\| - \|y\|$ para quaisquer $x, y \in E$. ■

Exercício 1.10 (?!)**.** Num espaço vetorial E de sua preferência, fixe dois vetores $u, v \in E$ e *descreva* a imagem da função $\gamma \colon [0,1] \to E$ que faz $\gamma(t) := tu + (1-t)v$ para cada $t \in [0,1]$. Dica: faça $E := \mathbb{R}^2$ se não tiver costume com espaços vetoriais abstratos (e, neste caso, aproveite para *esboçar* uma ilustração da imagem de γ). ■

Observação 1.0.14. Não confunda o gráfico de uma função com sua sua imagem. Para $f \colon X \to Y$, o gráfico de f é um subconjunto de $X \times Y$, a saber $\{(x, f(x)) : x \in X\}$, enquanto sua imagem é $\{f(x) : x \in X\}$, um subconjunto de Y. Se isto não lhe parecer óbvio, tire algum tempo para revisar esse assunto, procurar exemplos, etc. △

Exercício 1.11 $({}^{\star}_{\star})$**.** Sejam $(x_n)_n$ uma sequência em X e $x \in X$ um ponto fixado. Assumindo que X é métrico, mostre que $x_n \to x$ se, e somente se, $d(x_n, x) \to 0$ em $\mathbb{R}$. Tal resultado vale para redes? ■

Exercício 1.12 (?!)**.** Você percebeu alguma semelhança entre as normas do máximo em $\mathbb{R}^n$ e do supremo em $\mathcal{B}(X)$? ■

1.1 Estudos de caso

1.1.0 Essencial

§0 Alguns exemplos importantes

Funções constantes constituem os exemplos mais preguiçosos de convergência: se $r \in \mathbb{R}$ e $\mathbb{D}$ é um conjunto dirigido, então a rede $C_r \colon \mathbb{D} \to \mathbb{R}$, que faz $C_r(d) := r$ para todo $d \in \mathbb{D}$ e é usualmente denotada por $(r)_{d \in \mathbb{D}}$, converge precisamente para r. De fato, para $\varepsilon > 0$, qualquer $D \in \mathbb{D}$ é tal que $|C_r(d) - r| < \varepsilon$ sempre que $d \succeq D$, pois

$$|C_r(d) - r| = |r - r| = 0$$

para qualquer $d \in \mathbb{D}$. Em particular:

(i) fazendo $\mathbb{D} := \mathbb{N}$, segue que sequências constantes convergem;

(ii) fazendo $\mathbb{D} := \mathbb{R}_p$ para algum $p \in \mathbb{R}$ (cf. Exemplo 1.0.7), segue que qualquer função $f \colon \mathbb{R} \to \mathbb{R}$ que assume valor constante r em $\mathbb{R} \setminus \{p\}$ é tal que $\lim_{x \to p} f(x) = r$, independentemente do valor de f em p.[†]

[†]Não se preocupe, futuramente, vamos tratar de funções da forma $X \to \mathbb{R}$ com $X \subseteq \mathbb{R}$.

Exemplo 1.1.0 (Convergência não implica constância)**.** A recíproca do que se observou acima é falsa, i.e., existem redes convergentes que não são constantes. Um exemplo simples é a sequência $\left(\frac{1}{2^n}\right)_{n\in\mathbb{N}}$, cujo limite é 0. Vejamos:

✓ para $\varepsilon > 0$, busca-se $N \in \mathbb{N}$ tal que $\left|\frac{1}{2^n} - 0\right| < \varepsilon$ para todo $n \geq N$;

✓ para obter tal N, note que $\frac{1}{\varepsilon} > 0$ e, por $\mathbb{R}$ ser arquimediano, existe $N' \in \mathbb{N}$ tal que
$$\frac{1}{\varepsilon} < N';$$

✓ como $n < 2^n$ para todo $n \in \mathbb{N}$ com $n \geq 2$ (certo?)*, se tomarmos $N \geq \max\{N', 2\}$ teremos
$$\frac{1}{\varepsilon} < N < 2^N \text{(certo?)}^\star,$$
e, consequentemente,
$$\frac{1}{2^N} < \varepsilon;$$

✓ por fim, como $0 < \frac{1}{2} < 1$, resulta que $\left(\frac{1}{2^n}\right)_{n\in\mathbb{N}}$ é estritamente decrescente (cf. Exercício 1.78), de modo que se $n \geq N$, então $\frac{1}{2^n} < \frac{1}{2^N}$.

Por tudo o que se observou acima, conclui-se que $\lim_{n\to\infty} \frac{1}{2^n} = 0$ (certo?)*. ▲

A argumentação acima foi bastante cuidadosa por dois motivos: i) foi o primeiro exemplo explícito de uma rede convergente não trivial; ii) ainda não vimos resultados que permitem "adivinhar" alguns limites de modo rápido e indolor. As próximas seções vão lidar com esse tipo de problema. Mas tudo em seu tempo.

Exemplo 1.1.1 (Uma sequência *divergente*)**.** A definição de convergência não teria graça se tudo convergisse. Para um exemplo simples de rede que não converge, considere a sequência $((-1)^n)_{n\in\mathbb{N}}$: explicitamente, trata-se da sequência $(1, -1, 1, -1, 1, \dots)$ que leva $n \in \mathbb{N}$ ao número $x_n := (-1)^n$, daí $x_0 = 1$, $x_1 = -1$, etc. Intuitivamente, é óbvio que se trata de uma *sequência* que não converge em $\mathbb{R}$.

Para argumentar formalmente, o Exercício 1.3 pode ser útil: se $L \in \mathbb{R}$ fosse limite da sequência $((-1)^n)_{n\in\mathbb{N}}$, então para qualquer intervalo aberto $I \subseteq \mathbb{R}$ com $L \in I$ deveria ser possível encontrar $N \in \mathbb{N}$ tal que $(-1)^n \in I$ sempre que $n \geq N$. Existe algum L com tal propriedade?

✗ para $L := 1$, basta tomar I tal que $-1 \notin I$, pois daí para qualquer $N \in \mathbb{N}$, sempre se tem $(-1)^N \notin I$ ou $(-1)^{N+1} \notin I$ (por quê?)*;

✗ para $L := -1$, basta tomar I tal que $1 \notin I$, pois daí para qualquer $N \in \mathbb{N}$, sempre se tem $(-1)^N \notin I$ ou $(-1)^{N+1} \notin I$ (por quê?)*;

✗ para $L \neq -1, 1$, basta tomar I tal que $-1, 1 \notin I$, pois daí $(-1)^N \notin I$ para todo $N \in \mathbb{N}$.

Em breve teremos ferramentas bem mais práticas para detectar quando uma sequência não converge em $\mathbb{R}$. ▲

Definição 1.1.2. Diremos que uma rede real $(x_d)_{d\in\mathbb{D}}$ é **divergente em** $\mathbb{R}$ se não existe $L \in \mathbb{R}$ tal que $x_d \to L$. Sequências divergentes são definidas analogamente. ¶

Observação 1.1.3 (Cuidado com o conjunto dirigido!)**.** No último exemplo, a divergência de $((-1)^n)_{n\in\mathbb{N}}$ se deve ao fato de que $n+1 > n$ para todo n, de modo que os dois valores possíveis da sequência se alternam à medida que os índices avançam segundo a ordenação usual de $\mathbb{N}$. Isto pode deixar de acontecer se, por exemplo, trocarmos a ordem de $\mathbb{N}$, como no próximo

Exercício 1.13 $({}^{\star}_{\star\star})$**.** Para $\mathcal{N} := \mathbb{N}\setminus\{0\}$, considere a relação $\preceq$ em que se declara $m \preceq n$ se m *divide* n, i.e., se existe $k\in\mathbb{N}$ tal que $n = mk$. Mostre que $(\mathcal{N},\preceq)$ é um conjunto dirigido tal que a rede $((-1)^n)_{n\in\mathcal{N}}$ converge para 1. Dica: i) observe o que acontece com $(-1)^n$ para $n \succeq 2$; ii) releia a dica (i) até perceber que não está escrito "$n \geq 2$", mas sim "$n \succeq 2$"; iii) se você não entendeu o que significa "$n \succeq 2$" neste exercício, releia o enunciado desde o princípio[†]. ■

Com isso dito, não se preocupe: em geral, sempre que $\mathcal{M}$ for um subconjunto infinito de $\mathbb{N}$ e $(x_m)_{m\in\mathcal{M}}$ for uma rede, iremos considerar $\mathcal{M}$ com a ordem usual *herdada* de $\mathbb{N}$, o que na prática significa que podemos tratá-las como sequências[‡]. △

Como veremos, o problema de determinar se uma rede (sequência, função, etc.) converge ou não é bastante delicado. Assim, conhecer critérios rápidos, mesmo que aplicáveis apenas a um número *limitado* de casos, pode ser bastante útil. Por falar em *limitação*, já vimos que uma função $f\colon X\to\mathbb{R}$ é **limitada** se existe $M>0$ tal que $|f(x)|\leq M$ para todo $x\in X$ (cf. Subseção 0.7.1 §0). Em particular, como redes são funções, a mesma definição se aplica a elas — e a sequências, etc.

Teorema 1.1.4. *Toda rede real, limitada e monótona*[††] *converge em* $\mathbb{R}$.

Demonstração. Suponha $(x_d)_{d\in\mathbb{D}}$ crescente e considere $\alpha\in\mathbb{R}$ tal que $\alpha = \sup\{x_d : d\in\mathbb{D}\}$, que existe pois $T := \{x_d : d\in\mathbb{D}\}$ é não vazio e limitado enquanto a reta real é completa (confira as Definições 0.9.5 e 0.10.6 caso já tenha se esquecido). Mostraremos que $x_d\to\alpha$.

Para isso, dado $\varepsilon>0$, temos $\alpha-\varepsilon<\alpha$. Por α ser o supremo de T, $\alpha-\varepsilon$ não pode limitar T superiormente. Logo, existe pelo menos um $D\in\mathbb{D}$ com $x_D > \alpha-\varepsilon$. Enfim, por $(x_d)_{d\in\mathbb{D}}$ ser crescente, temos

$$\alpha-\varepsilon < x_D \leq x_d \leq \alpha < \alpha+\varepsilon$$

para todo $d\succeq D$. O caso decrescente é problema seu. □

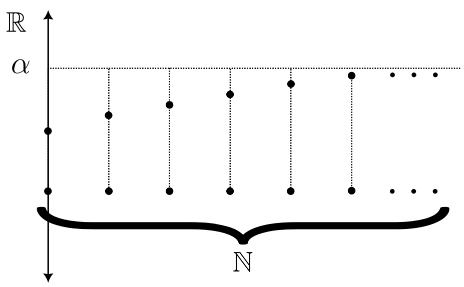


Figura 1.0: O supremo α impede que os pontos escapem "para cima", enquanto o comportamento crescente da sequência "empurra" os pontos até α.

[†]Se nada disso ajudou... (suspiro): "$n\succeq 2$" significa "$2\preceq n$", que por sua vez significa "2 divide n". Por que eu sei disso? Porque li a definição da relação $\preceq$ no começo do enunciado do exercício.

[‡]Isto ficará mais claro após a discussão de *subsequências*.

[††]Isto é, crescente ($d\preceq d' \Rightarrow x_d\leq x_{d'}$) ou decrescente ($d\preceq d'\Rightarrow x_{d'}\leq x_d$). Trata-se da mesma definição apresentada no Exercício 0.70.

Exercício 1.14 ($^{\star}_{\star}$)**.** Mostre que se $(x_d)_{d\in\mathbb{D}}$ é uma rede real, limitada e decrescente, então $\inf\{x_d : d \in \mathbb{D}\}$ é o limite da rede. ■

Observação 1.1.5. A Subseção 1.1.1 §0 traz uma aplicação surpreendente para o resultado acima: a existência de *limites laterais reais* para funções reais limitadas e monótonas. △

Exemplo 1.1.6. Como prometido, agora é bem mais fácil provar que $\lim_{n\to\infty}\frac{1}{2^n} = 0$: dado que $\inf_{n\in\mathbb{N}}\frac{1}{2^n} = 0$ (cf. Exercício 0.125) e $\left(\frac{1}{2^n}\right)_{n\in\mathbb{N}}$ é decrescente, resulta

$$\lim_{n\to\infty}\frac{1}{2^n} = \inf_{n\in\mathbb{N}}\frac{1}{2^n} = 0.$$

Para outro exemplo, observe que $\lim_{n\to\infty}\left(1-\frac{1}{2^n}\right) = 1$, uma vez que $\left(1-\frac{1}{2^n}\right)_{n\in\mathbb{N}}$ é limitada, crescente (certo?)* e valem as identidades

$$\sup_{n\in\mathbb{N}}\left\{1-\frac{1}{2^n} : n\in\mathbb{N}\right\} = 1 - \inf_{n\in\mathbb{N}}\left\{\frac{1}{2^n} : n\in\mathbb{N}\right\} = 1$$

por conta do Teorema 0.8.6 (entendeu?)$^{\star}_{\star}$. Futuramente, isto será ainda mais simples. ▲

Exemplo 1.1.7 (Séries de números reais)**.** Um tipo particular de sequência costuma ser bastante útil em diversos contextos: as *séries*. Fixada uma sequência $(x_n)_{n\in\mathbb{N}}$ em $\mathbb{R}$, a **série** (de números reais) determinada por $(x_n)_{n\in\mathbb{N}}$, geralmente denotada por

$$\sum_{n=0}^{\infty} x_n, \quad \sum_{n\geq 0} x_n \quad \text{ou apenas} \quad \sum x_n,$$

é a *sequência* $(s_n)_{n\in\mathbb{N}}$ formada pelos termos $s_n := \sum_{j=0}^{n} x_j$, que por sua vez costumam ser chamados de **somas parciais da série**.

Intuitivamente, a série $\sum x_n$ *representa* o que seria a *soma* de <u>todos</u> os termos da sequência $(x_n)_{n\in\mathbb{N}}$, o que não faz sentido puramente algébrico já que as operações de $\mathbb{R}$ são *finitárias*, mas que pode fazer sentido do ponto de vista ~~topológico~~ ordenado por meio da noção de limite.

Definição 1.1.8. Diremos que a série de números reais $\sum x_n$ **converge** em $\mathbb{R}$ se existir $S \in \mathbb{R}$ tal que

$$S = \lim_{n\to\infty}\sum_{j=0}^{n} x_j,$$

i.e., se as somas parciais da série convergem para S. Por abuso de notação, em tais situações, o limite da série também será denotado por $\sum x_n$, que xingaremos de **soma da série** $\sum x_n$. ¶

É importante não confundir a sequência $(x_n)_{n\in\mathbb{N}}$ com a série $\sum_{n\geq 0} x_n$. Além das notações serem completamente diferentes[†], cada uma tem um significado preciso: $(x_n)_{n\in\mathbb{N}}$ é a função que associa $x_n \in \mathbb{R}$ a cada $n \in \mathbb{N}$, enquanto $\sum_{n\geq 0} x_n$ associa $x_0 + \ldots + x_n$ a cada $n \in \mathbb{N}$. Mastigando ainda mais, note que a sequência constante $(1)_{n\in\mathbb{N}}$ induz a série $\sum_{n\geq 0} 1$, que por sua vez é a sequência $(1,2,3,4,5,6,\ldots)$. Ao longo do curso, séries serão animais bastante recorrentes. Por ora, vamos apenas reescrever um resultado que já foi utilizado no capítulo anterior (cf. Lema 0.10.8) para ilustrar um pouco mais a aplicabilidade do Teorema 1.1.4. ▲

Exercício 1.15 ($^{\star}_{\star}$)**.** Mostre que se $0 < a < 1$, então $\sum_{n\geq 0} a^n = \frac{1}{1-a}$. ■

[†]Em caso de discordância, procure especialistas em oftalmologia.

§1 Limites na reta estendida

Na discussão anterior, é bem provável que você tenha percebido um segundo exemplo de sequência divergente: a saber, $\sum_{n\geq 0} 1$, a sequência formada por todos os números naturais maiores do que zero. Ela não pode convergir em $\mathbb{R}$ pois, se convergisse, seu limite seria o supremo[†] de $\mathbb{N}$ em $\mathbb{R}$, mas $\mathbb{N}$ é ilimitado em $\mathbb{R}$ (cf. Teorema 0.9.7). Ainda assim, a *divergência* de $\sum_{n\geq 0} 1$ parece ser diferente da *divergência* de $((-1)^n)_{n\in\mathbb{N}}$: enquanto os termos da primeira assumem valores *arbitrariamente altos*, a segunda oscila indefinidamente.

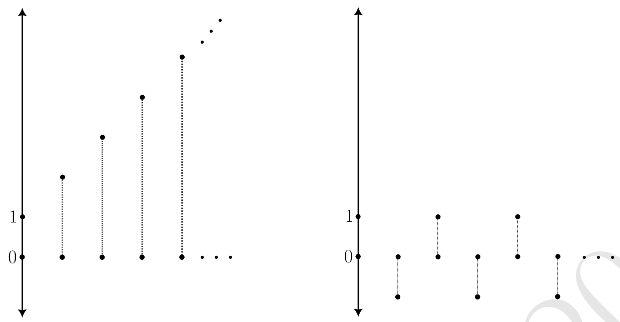


Figura 1.1: A série $\sum_{n\geq 0} 1$ vs. a sequência $((-1)^n)_{n\in\mathbb{N}}$.

Mais precisamente, ao dizer que os termos de uma sequência $(x_n)_{n\in\mathbb{N}}$ assumem *valores arbitrariamente altos*, busca-se abreviar a seguinte condição: para todo $M \in \mathbb{R}$ (na verdade, basta fazer $M > 0$), existe $N \in \mathbb{N}$ tal que $M < x_n$ sempre que $n \geq N$.

Reescrevendo a condição acima por meio de intervalos, os termos da sequência $(x_n)_{n\in\mathbb{N}}$ *assumem valores arbitrariamente altos* se para todo $M \in \mathbb{R}$ existe $N \in \mathbb{N}$ tal que $x_n \in (M, +\infty)$ sempre que $n \geq N$. Esta última formulação sugere que assumir valores arbitrariamente altos possa ter semelhanças com a convergência em $\mathbb{R}$.

Considere, por exemplo, a sequência $\left(1 - \frac{1}{2^n}\right)_{n\in\mathbb{N}}$. Já sabemos que $1 - \frac{1}{2^n} \to 1$, o que por sua vez significa que para todo $\varepsilon > 0$ existe $N \in \mathbb{N}$ tal que

$$\left|1 - \frac{1}{2^n} - 1\right| < \varepsilon$$

sempre que $n \geq N$. Porém, ao analisar este caso por um ponto de vista (mais) geométrico, podemos afirmar que para qualquer $M < 1$ existe $N \in \mathbb{N}$ tal que $1 - \frac{1}{2^n} \in (M, 1)$ para todo $n \geq N$: como nenhum termo da sequência em questão assume valores maiores do que ou iguais a 1, o lado direito dos intervalos pode ser ignorado.

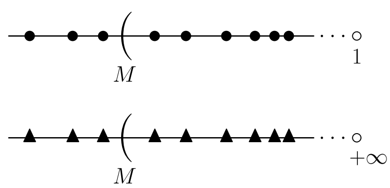


Figura 1.2: As bolinhas convergem para 1, enquanto os triângulos "convergem para $+\infty$".

Assim, ao pensar em convergência menos algebricamente e mais *geometricamente* (por meio de intervalos), percebemos que seria inofensivo "acrescentar" um ponto artificial à

[†] Atenção: isto não é consequência do Teorema 1.1.4: o teorema diz que se uma rede real é monótona e limitada, então ela converge (para o seu supremo ou o seu ínfimo, a depender do caso). Aqui, o que está sendo dito é: se uma rede monótona converge em $\mathbb{R}$, então seu limite é supremo ou ínfimo da imagem da rede (a depender do caso). Confira o Exercício 1.80 para mais detalhes.

direita da reta real a fim de servir como *limite* das sequências (... ou redes) que assumem valores arbitrariamente altos: este ponto é geralmente denotado por $+\infty$, e costuma-se escrever

$$\lim_{n\to\infty} x_n = +\infty$$

para indicar que "$(x_n)_{n\in\mathbb{N}}$ assume valores arbitrariamente altos", por exemplo.

Após refazer toda essa argumentação num espelho, chega-se à definição de sequências que assumem *valores arbitrariamente baixos* e, mais importante, percebe-se que seria inofensivo "acrescentar" um ponto artificial à esquerda da reta real, comumente denotado por "$-\infty$".

Definição 1.1.9. A **reta estendida** é o conjunto $\mathbb{R}\cup\{-\infty,+\infty\}$, com a ordem que você imagina: para $x, y \in [-\infty, +\infty]$, $x < y$ significa

- $x < y$ no sentido usual se $x, y \in \mathbb{R}$ (ou seja, números reais seguem sua ordem "de sempre"), ou
- $x := -\infty$ e $y \in \mathbb{R}\cup\{+\infty\}$, ou
- $y := +\infty$ e $x \in \mathbb{R}\cup\{-\infty\}$. ¶

Em $[-\infty, +\infty]$, passa a fazer sentido escrever $\sup A = +\infty$ sempre que A for um subconjunto não vazio de números reais ilimitado superiormente, assim como $\inf B = -\infty$ para $B \subseteq \mathbb{R}$ com $B \neq \emptyset$ ilimitado inferiormente: para o primeiro caso, $+\infty$ é automaticamente limitante superior de A, mas nenhum $L < +\infty$ pode limitar A superiormente, já que supomos A ilimitado em $\mathbb{R}$.

Exercício 1.16 ($^\star_\star$)**.** Seja $\emptyset \neq B \subseteq \mathbb{R}$. Mostre que se B é ilimitado inferiormente, então $\inf B = -\infty$. Vale a volta? ■

Também podemos ser mais flexíveis com intervalos: além dos intervalos usuais de $\mathbb{R}$, são permitidos os intervalos da forma $[-\infty, \alpha]$, $[-\infty, \alpha)$, $[\beta, +\infty]$ e $(\beta, +\infty]$ para $\alpha, \beta \in [-\infty, +\infty]$. Diremos que $I \subseteq [-\infty, +\infty]$ é um **intervalo aberto fundamental** se I for da forma $[-\infty, \alpha)$ ou $(\alpha, +\infty]$ para algum $\alpha \in [-\infty, +\infty]$. Interseções finitas de intervalos abertos fundamentais serão chamadas de **intervalos abertos**.

Definição 1.1.10. Um ponto $L \in [-\infty, +\infty]$ será chamado de (*um*) **limite estendido da rede** $(x_d)_{d\in\mathbb{D}}$ se a seguinte condição for satisfeita: para todo intervalo aberto $I \subseteq [-\infty, +\infty]$ com $L \in I$ existe $D \in \mathbb{D}$ tal que $x_d \in I$ sempre que $d \succeq D$. Em tal situação, também se diz que $(x_d)_{d\in\mathbb{D}}$ **converge para** L **na reta estendida**, o que se abrevia com a notação $x_d \to L$. Dizer que uma rede é **convergente na reta estendida** apenas indica que existe um limite em $[-\infty, +\infty]$ para o qual a rede converge. ¶

Como intervalos abertos de $\mathbb{R}$ são intervalos abertos de $[-\infty, +\infty]$ (cf. Exercício 0.103), segue que limites reais são limites na reta estendida. A falha da recíproca se dá justamente com as redes arbitrariamente altas ou baixas: elas divergem na reta real, mas convergem na reta estendida.

Proposição 1.1.11. *Se $x_d \to L$ com $L \in \{-\infty, +\infty\}$ e $r \in \mathbb{R}$, então $x_d \not\to r$.*

Demonstração. O argumento é essencialmente *geométrico*: tudo segue pois existem intervalos abertos $I, J \subseteq [-\infty, +\infty]$ tais que $L \in I$, $r \in J$ e $I \cap J = \emptyset$ (certo?!)*. De fato, sabendo disso, como $x_d \to L$, existe $D \in \mathbb{D}$ com $x_d \in I$ para todo $d \succeq D$. Por outro lado, se também valesse $x_d \to r$, então existiria $D' \in \mathbb{D}$ tal que $x_d \in J$ sempre que $d \succeq D'$. Logo, para $D'' \succeq D, D'$ ocorreria tanto $x_{D''} \in I$ quanto $x_{D''} \in J$. □

Exercício 1.17 $({}^{\star}_{\star})$**.** Mostre que se $L, L' \in [-\infty, +\infty]$ e $x_d \to L$ e $x_d \to L'$, então $L = L'$. Dica: imite o argumento anterior. ■

Definição 1.1.12. Para uma rede real $(x_d)_{d\in\mathbb{D}}$, vamos escrever $\lim\limits_{d\in\mathbb{D}} x_d$ para indicar o único ponto $L \in [-\infty, +\infty]$, se existir, tal que $x_d \to L$. ¶

Embora pareça diferente, a abordagem aqui adotada apenas enfatiza os aspectos geométricos das diversas definições clássicas de limites reais e infinitos: a mágica, se há alguma, está nos intervalos.

Exercício 1.18 $({}^{\star}_{\star})$**.** A seguir, para uma sequência $(x_n)_{n\in\mathbb{N}}$ de números reais e um ponto $L \in [-\infty, +\infty]$, a notação "$x_n \to L$" está de acordo com a Definição 1.1.10.

a) Para $L \in \mathbb{R}$, mostre que $x_n \to L$ se, e somente se, para todo $\varepsilon > 0$ existe $N \in \mathbb{N}$ tal que $|x_n - L| < \varepsilon$ sempre que $n \geq N$. Dica: note que um ponto $a \in \mathbb{R}$ pertence a um intervalo aberto $I \subseteq [-\infty, +\infty]$ se, e somente se, existe $\varepsilon > 0$ tal que $(a-\varepsilon, a+\varepsilon) \subseteq I$, e lembre-se de que $b \in (a-\varepsilon, a+\varepsilon)$ se, e somente se, $|b - a| < \varepsilon$.

b) Mostre que $x_n \to +\infty$ se, e somente se, para todo $M \in \mathbb{R}$ com $M > 0$ existe $N \in \mathbb{N}$ tal que $M < x_n$ sempre que $n \geq N$.

c) Mostre que $x_n \to -\infty$ se, e somente se, para todo $M \in \mathbb{R}$ com $M < 0$ existe $N \in \mathbb{N}$ tal que $x_n < M$ sempre que $n \geq N$. ■

Exercício 1.19 $({}^{\star}_{\star})$**.** Adapte o exercício acima para caracterizar $\lim_{x\to p} f(x) = L$ nos casos em que $L \in \mathbb{R}$, $L := +\infty$ e $L := -\infty$, onde $f \colon \mathbb{R} \to \mathbb{R}$ é uma função e $p \in \mathbb{R}$. ■

Exercício 1.20 $({}^{\star}_{\star})$**.** Mostre que $\lim_{n\to\infty} 2^n = +\infty$ e $\lim_{n\to\infty} -(2^n) = -\infty$. A sequência $((-2)^n)_{n\in\mathbb{N}}$ converge em $[-\infty, +\infty]$? ■

Observação 1.1.13. Se você não dormiu nas aulas de Cálculo n, para $n \geq 1$, deve se lembrar de que também existem limites da forma $\lim\limits_{x\to p} f(x)$ para funções reais, onde $p \in \{-\infty, +\infty\}$. Isto será feito aqui, na distante Definição 1.8.5, por meio de algumas noções *topológicas* ainda não discutidas. Com isso dito, caso você se lembre da ideia... △

Exercício 1.21 (?!)**.** Use redes para definir $\lim\limits_{x\to p} f(x) = L$, onde $f \colon \mathbb{R} \to \mathbb{R}$ é uma função, com $p \in \{-\infty, +\infty\}$ e $L \in [-\infty, +\infty]$. ■

Exercício 1.22 $({}^{\star}_{\star})$**.** Mostre que se $(x_d)_{d\in\mathbb{D}}$ é uma rede real e monótona, então existe $L \in [-\infty, +\infty]$ tal que $x_d \to L$. Dica: adapte a prova do Teorema 1.1.4 e a solução do Exercício 1.14 ■

1.1.1 Extras

§0 Limites laterais e funções monótonas

Observação 1.1.14. Embora já tenhamos ferramental para discutir limites de funções da forma $X \to \mathbb{R}$ com $X \subseteq \mathbb{R}$, isto só será feito após termos discutido noções *topológicas*. Com isso dito, os resultados aqui se aplicam nesse contexto levemente mais geral, *mutatis mutandis*. △

Para uma função $f\colon \mathbb{R} \to \mathbb{R}$ e um ponto $p \in \mathbb{R}$, a relação $\preceq$ sobre $\mathbb{R}_p \coloneqq \mathbb{R} \setminus \{p\}$ segundo a qual

$$\lim_{x \to p} f(x) = \lim_{r \in \mathbb{R}_p} f(r),$$

i.e., segundo a qual a definição usual de limite coincide com a noção de limite de rede, declara

$$a \preceq b \text{ se, e somente se, } |b-p| \leq |a-p|.$$

Verbalmente: b é melhor do que a se a distância de b até p é menor do que ou igual à distância de a até p (em outras palavras, quando o assunto é estar próximo de p, b é melhor nisso do que a).

Exercício 1.23 ($\star$)**.** Para $p \coloneqq 0$, dê exemplos de números reais $a, b \in \mathbb{R}$ satisfazendo $a < b$ e $a \succ b$. Observação: note que não se pediu "$a < b$ e $a > b$", mas sim "$a < b$ e $a \succ b$" (os símbolos "$>$" e "$\succ$" indicam coisas diferentes!). ■

O exercício sugere que a relação $\preceq$ sobre $\mathbb{R}_p$ ignora o "lado" pelo qual o ponto se aproxima de p. No entanto, há situações bastante naturais em que faz sentido analisar o comportamento da função levando em conta a direção de aproximação.

Exemplo 1.1.15 (*Spoiler alert*: derivadas)**.** Futuramente, para uma função $f\colon \mathbb{R} \to \mathbb{R}$ e um ponto $p \in \mathbb{R}$, iremos considerar o número

$$f'(p) \coloneqq \lim_{x \to p} \frac{f(x) - f(p)}{x - p},$$

que, quando existir em $\mathbb{R}$, será chamado de *derivada da função f no ponto p.* A rigor, trata-se do limite da rede $\eta_p\colon \mathbb{R}_p \to \mathbb{R}$ dada por

$$x \mapsto \frac{f(x) - f(p)}{x - p}.$$

Em particular, para $f(x) \coloneqq |x|$ e $p \coloneqq 0$, a função η_p correspondente é dada pela regra

$$\eta_p(x) = \frac{|x|}{x} = \begin{cases} 1, & \text{se } x > 0 \\ -1, & \text{se } x < 0 \end{cases},$$

mostrando que os valores da rede $(\eta_p(x))_{x \in \mathbb{R}_p}$ dependem do lado pelo qual x se aproxima de 0: para $x > 0$, $\eta_p(x) = 1$, ao passo que para $x < 0$, $\eta_p(x) = -1$. Como no caso da *sequência* $((-1)^n)_{n \in \mathbb{N}}$, não é difícil perceber que $(\eta_p(x))_{x \in \mathbb{R}_p}$ diverge (certo?)$^{\star}_{\star}$. Ainda assim, ao considerarmos apenas um dos lados, os limites correspondentes parecem existir. ▲

O exemplo anterior ilustra uma situação típica em que *limites laterais* dão as caras. Para uma função $f\colon \mathbb{R} \to \mathbb{R}$ e $p \in \mathbb{R}$, diz-se que $L \in [-\infty, +\infty]$ é **limite de $f(x)$ quando x tende a p pela esquerda** se para todo intervalo aberto $I \subseteq [-\infty, +\infty]$ com $L \in I$, existe $\delta > 0$ tal que $f(x) \in I$ sempre que $0 < p - x < \delta$. Analogamente, diz-se que L é **limite de $f(x)$ quando x tende a p pela direita** se para todo intervalo aberto $I \subseteq [-\infty, +\infty]$ com $L \in I$ existe $\delta > 0$ tal que $f(x) \in I$ sempre que $0 < x - p < \delta$.

Exercício 1.24 ($\star$)**.** Mostre que $0 < p - x < \delta$ se, e somente se, $p - \delta < x < p$. Faça o mesmo para o caso à direita. ■

Exemplo 1.1.16. Seja $g\colon \mathbb{R} \to \mathbb{R}$ tal que $g(x) := \frac{1}{x}$ para todo $x \neq 0$, com $g(0) \in \mathbb{R}$ qualquer[†]. Vamos ver que o limite de $g(x)$ quando x tende a 0 pela esquerda é $-\infty$. Um intervalo aberto $I \subseteq [-\infty, +\infty]$ *em torno de* $-\infty$ é da forma $I = [-\infty, M)$ para algum $M \in \mathbb{R}$, que podemos assumir negativo, i.e., tal que $M < 0$. Agora, busca-se $\delta > 0$ tal que se $0 < -x < \delta$, então $g(x) \in I$, i.e., $g(x) < M$. Ora, como $-x > 0$ e $M < 0$, temos

$$\frac{1}{x} < M \Leftrightarrow -\frac{1}{M} > -x,$$

mostrando que basta tomar $\delta := -\frac{1}{M}$ (certo?)$^{\star}_{\star}$. Analogamente, o limite de $g(x)$ quando x tende a 0 pela direita é $+\infty$. Por fim, não existe limite de $g(x)$ quando x tende a 0: $+\infty$ não pode ser limite de $g(x)$ pois sempre podemos tomar pontos arbitrariamente próximos de 0 pela esquerda, e estes se aproximam de $-\infty$; por sua vez, $-\infty$ também não pode ser limite de $g(x)$ (por quê?)$^{\star}_{\star}$. ▲

Parece a definição de limite usual mas, a rigor, não é: no caso *canhoto*, apenas os pontos x satisfazendo $x < p$ interessam (aqueles à esquerda de p!); no caso *destro*, apenas os pontos x satisfazendo $p < x$ interessam (aqueles à direita de p!).

A princípio, isto proíbe a *aplicação* das propriedades de limites usuais para limites laterais, já que os segundos são diferentes dos primeiros. Porém, como as demonstrações são verdadeiramente análogas, o mais comum é deixar essas verificações como exercícios[‡]. O ponto importante aqui é outro: se você interpretar limites laterais como limites de redes, as verificações nem precisam ser deixadas como exercício, pois as propriedades se tornam corolários diretos.

$(-)$ Limite pela esquerda. Ao considerar $\mathbb{R}_{<p} := (-\infty, p)$ com a relação $\preceq$ *herdada* de $\mathbb{R}_p$, tem-se $a \preceq b$ se, e somente se, $a \leq b$, pois como ambos $a, b \in (-\infty, p)$, tem-se $a < p$ e $b < p$, de modo que

$$a \preceq b \Leftrightarrow |b - p| \leq |a - p| \Leftrightarrow p - b \leq p - a \Leftrightarrow -b \leq -a \Leftrightarrow a \leq b.$$

Agora, como $(\mathbb{R}_{<p}, \preceq)$ é dirigido (certo?)*, podemos considerar a rede $(f(x))_{x\in\mathbb{R}_{<p}}$. Enfim, um ponto $L \in [-\infty, +\infty]$ é limite da rede $(f(x))_{x\in\mathbb{R}_{<p}}$ se, e somente se, para todo intervalo aberto $I \subseteq [-\infty, +\infty]$ com $L \in I$ existe $D \in \mathbb{R}_{<p}$ tal que $f(x) \in I$ sempre que $x \succeq D$. Como no Exemplo 1.0.7, não é difícil perceber que

$$\lim_{x\in\mathbb{R}_{<p}} f(x) = L \text{ se, e somente se, } L \text{ é limite de } f(x) \text{ quando } x \text{ tende a } p \text{ pela esquerda.}$$

$(+)$ Limite pela direita. Ao considerar $\mathbb{R}_{>p} := (p, +\infty)$ com a relação $\preceq$ definida sobre $\mathbb{R}_p$, tem-se $a \preceq b$ se, e somente se, $b \leq a$, pois como ambos $a, b \in (p, +\infty)$, tem-se $p < a$ e $p < b$, acarretando

$$a \preceq b \Leftrightarrow |b - p| \leq |a - p| \Leftrightarrow b - p \leq a - p \Leftrightarrow b \leq a.$$

Agora, como $(\mathbb{R}_{>p}, \preceq)$ é dirigido (certo?)*, podemos considerar a rede $(f(x))_{x\in\mathbb{R}_{>p}}$. Enfim, um ponto $L \in [-\infty, +\infty]$ é limite da rede $(f(x))_{x\in\mathbb{R}_{>p}}$ se, e somente se... Você pegou a ideia, e não deve ter dificuldades para verificar que

$$\lim_{x\in\mathbb{R}_{>p}} f(x) = L \text{ se, e somente se, } L \text{ é limite de } f(x) \text{ quando } x \text{ tende a } p \text{ pela direita.}$$

[†]O valor que g assume em 0 é irrelevante para o que faremos. Na prática, isto significa que estamos considerando, na verdade, a restrição da função g ao subconjunto $\mathbb{R} \setminus \{0\}$. Porém, não precisa se preocupar com isso ainda.

[‡]Que nunca são feitos!

Portanto, já não precisamos nos preocupar em provar que limites laterais, quando existem, são únicos, pois já provamos tal resultado para redes reais quaisquer. Em particular, a unicidade justifica a atribuição de notação específica. Tipicamente, tais limites são denotados por

$$\lim_{x\to p^-} f(x) := \lim_{x\in\mathbb{R}_{<p}} f(x) \quad \text{e} \quad \lim_{x\to p^+} f(x) := \lim_{x\in\mathbb{R}_{>p}} f(x). \tag{1.3}$$

Analogamente, após provarmos as propriedades operatórias dos limites reais de redes reais, como

$$\lim_{d\in\mathbb{D}} x_d + \lim_{d\in\mathbb{D}} y_d = \lim_{d\in\mathbb{D}}(x_d + y_d),$$

teremos, <u>automaticamente</u>, os resultados correspondentes para os diversos tipos de limites de funções, como

$$\lim_{x\to p} f(x)+\lim_{x\to p} g(x) = \lim_{x\to p}(f(x)+g(x)), \quad \lim_{x\to p^-} f(x)+\lim_{x\to p^-} g(x) = \lim_{x\to p^-}(f(x)+g(x)), \quad \text{etc.}$$

Observação 1.1.17 (Importante)**.** Talvez você tenha notado que

$$\lim_{x\to p} f(x) = L \Rightarrow \lim_{x\to p^-} f(x) = \lim_{x\to p^+} f(x) = L,$$

ou, verbalmente: se o limite de $f(x)$ quando x tende a p existe e é igual a L, então ambos os limites laterais de $f(x)$ em p existem e são iguais a L. Isto de fato acontece (por quê)$\overset{\star}{\star\star}$, o que dá um critério prático para decidir quando $\lim_{x\to p} f(x)$ não existe: basta que os limites laterais não existam, ou existam mas não coincidam. Secretamente, trata-se do mesmo fenômeno observado na <u>sequência</u> $((-1)^n)_{n\in\mathbb{N}}$, que não converge por existirem *subsequências* que convergem para pontos distintos. Retornaremos a isso em breve, não se preocupe. △

Por fim, uma consequência inesperada de interpretar limites laterais como redes:

Proposição 1.1.18. *Se $f\colon \mathbb{R}\to\mathbb{R}$ é monótona, então*

$$\lim_{x\to p^-} f(x) \quad e \quad \lim_{x\to p^+} f(x)$$

existem em $\mathbb{R}$ para todo $p\in\mathbb{R}$.

Demonstração. Se f é crescente, por exemplo, então $(f(x))_{x\in\mathbb{R}_{<p}}$ é uma rede crescente e sua imagem é limitada superiormente por $f(p)$, enquanto $(f(x))_{x\in\mathbb{R}_{>p}}$ é decrescente e sua imagem é limitada inferiormente por $f(p)$ (por quê?)$\overset{\star}{\star}$. Logo, pelo Teorema 1.1.4, existem $L, L'\in\mathbb{R}$ tais que

$$\lim_{x\to p^-} f(x) = \lim_{x\in\mathbb{R}_{<p}} f(x) = L \quad \text{e} \quad \lim_{x\to p^+} f(x) = \lim_{x\in\mathbb{R}_{>p}} f(x) = L',$$

como desejado[†]. □

Exercício 1.25 $(\substack{\star\\\star})$**.** Prove a proposição anterior para o caso em que f é decrescente. Sugestão: não prove, mas apenas use o que já foi provado, observando que $-f\colon\mathbb{R}\to\mathbb{R}$ é crescente se f for decrescente[‡]. ■

[†]Cuidado: o Teorema 1.1.4 originalmente exige que a imagem da rede seja limitada (superiormente e inferiormente). Porém, isto foi feito apenas para tornar o enunciado mais legível: na prática, se $(x_d)_{d\in\mathbb{D}}$ é rede crescente e sua imagem é limitada superiormente, o mesmo argumento se aplica para mostrar que $\lim_{d\in\mathbb{D}} x_d = \sup\{x_d : d\in\mathbb{D}\}$; analogamente, se $(x_d)_{d\in\mathbb{D}}$ é decrescente e sua imagem é limitada inferiormente, então $\lim_{d\in\mathbb{D}} x_d = \inf\{x_d : d\in\mathbb{D}\}$.

[‡]Se você for uma pessoa muito honesta, talvez prefira esperar ter em mãos as propriedades operatórias de limites, a fim de usar a identidade $\lim_{d\in\mathbb{D}}(-x_d) = -\lim_{d\in\mathbb{D}} x_d$. É uma atitude muito honrosa, parabéns!

§1 Supremos e ínfimos revistos (parte I)

É possível que, em seu primeiro contato com supremos e ínfimos, você tenha *sentido* que as definições escondiam alguma *noção de* limite — como os estudados em Cálculo I ou II. Agora, após conhecer limites de redes, talvez a sensação esteja mais forte. Por sorte, a impressão está certa, e não é difícil se convencer disso.

De fato, para $S \subseteq \mathbb{R}$ um subconjunto não vazio, podemos considerar S como um conjunto dirigido tanto com a ordenação usual de $\mathbb{R}$ quanto com a ordem inversa. Para evitar confusões, vamos escrever $\overrightarrow{S}$ para indicar o conjunto dirigido $(S, \leq)$ (aqui, $\leq$ é realmente a ordem usual), e $\overleftarrow{S}$ para indicar o conjunto dirigido $(S, \preceq)$, onde $a \preceq b$ se, e somente se, $b \leq a$ (se você prometer não se confundir, pode escrever $\geq$ em vez de $\preceq$).

Proposição 1.1.19. *Nas notações acima,* $\sup S = \lim_{s \in \overrightarrow{S}} s$ *e* $\inf S = \lim_{s \in \overleftarrow{S}} s$.

Demonstração. Note que as redes $(s)_{s \in \overrightarrow{S}}$ e $(s)_{s \in \overleftarrow{S}}$ são, respectivamente, crescente e decrescente, e ambas têm o conjunto S como imagem. Logo, para os casos limitados, tudo segue do Teorema 1.1.4. Os casos ilimitados decorrem do que se observou acerca de supremos e ínfimos de conjuntos ilimitados (cf. Exercício 1.16). Completar os detalhes fica por sua conta $({}^{\star}_{\star\star})$. □

Futuramente, esta abordagem permitirá reobter as propriedades operatórias de supremos e ínfimos como casos particulares das propriedades operatórias dos limites de redes (mas a que custo, não é mesmo?).

§2 Bolas abertas e limites em espaços métricos

Os intervalos na reta estendida permitiram não apenas dar sentido formal aos limites infinitos, como também (permitiram) interpretar de maneira mais intuitiva/geométrica as noções de limites reais. Algo similar pode ser feito em espaços normados e métricos — mas em vez de intervalos, usam-se *bolas*. A fim de motivar a definição, observe que para $a, b \in \mathbb{R}$ com $a < b$, ao fazer $c := \frac{a+b}{2}$ e $r := \frac{b-a}{2}$, resulta

$$(a, b) = (c - r, c + r) = \{x \in \mathbb{R} : |x - c| < r\}.$$

Exercício 1.26 $(\star)$**.** Mostre que as identidades acima estão certas. Dica/observação: note que c é o “centro” do intervalo (a, b). ■

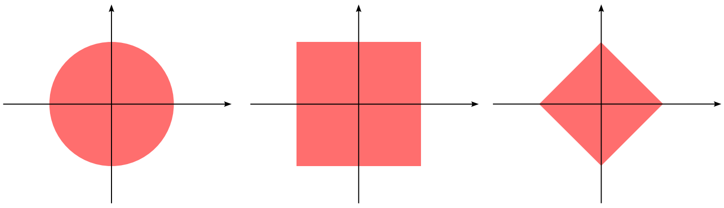
A última *descrição* de intervalo tem a vantagem de não depender da ordem de $\mathbb{R}$ ou de sua estrutura algébrica: ao escrever $(a, b) = \{x \in \mathbb{R} : |x - c| < r\}$, expressa-se que (a, b) é a coleção de todos os pontos da reta cuja distância até c é menor do que r. Ou seja: trata-se de algo *realizável* em *espaços métricos* (cf. Definição 1.0.11).

Definição 1.1.20. Dado um espaço métrico (X, d), para $x \in X$ e $r \in \mathbb{R}$ com $r > 0$, define-se a **d-bola aberta de centro x e raio r** como sendo o conjunto

$$B_d(x, r) := \{y \in X : d(x, y) < r\}.$$

Quando não houver risco de confusão, o sufixo “d-” será suprimido. ¶

Exemplo 1.1.21 (Bolas em $\mathbb{R}^2$)**.** O uso da palavra “bola” é motivado pela interpretação geométrica das *bolas* definidas por meio das normas euclidianas em $\mathbb{R}^2$ (no plano) e $\mathbb{R}^3$ (no espaço), como sugerido pela figura a seguir, à esquerda.



No entanto, você deve abandonar o quanto antes a ideia de que *bolas* precisam ser “redondas”. Chamando por d_∞ a métrica em $\mathbb{R}^2$ induzida pela norma do máximo $\|\cdot\|_\infty$ (cf. Exercício 1.5), não é difícil perceber que $B_{d_\infty}(0,1)$ corresponde ao quadrado central na figura anterior, onde 0 indica o vetor nulo de $\mathbb{R}^2$, por abuso de notação (e é um bom exercício fazer o esboço por conta própria!)$^{\star}_{\star\star}$.

Da mesma forma, chamando por d_1 a métrica em $\mathbb{R}^2$ induzida pela norma da soma $\|\cdot\|_1$ (cf. Exercício 1.5), a bola $B_{d_1}(0,1)$ corresponde ao último quadrado na figura anterior, à direita. Com efeito, deve-se ter $u := (x,y) \in B_{d_1}(0,1)$ se, e somente se, $\|u\|_1 < 1$, i.e., $|x| + |y| < 1$. Ocorre que tal desigualdade equivale a $|x| - 1 < y < 1 - |x|$, e os pontos (x,y) de $\mathbb{R}^2$ com tal propriedade são, precisamente, aqueles que se encontram na interseção das duas regiões esboçadas na Figura 1.3 (verifique!)$^{\star}_{\star\star}$. ▲

Figura 1.3: Os pontos do plano satisfazendo as inequações $|x| - 1 < y$ e $y < 1 - |x|$, respectivamente.

Bolas abertas permitem dar uma interpretação geométrica bastante intuitiva para convergência: uma rede $(x_a)_{a\in\mathbb{A}}$ num espaço métrico (X,d) converge para um ponto $L \in X$ se, e somente se, para toda bola aberta B com $L \in B$ existe $A \in \mathbb{A}$ tal que $x_a \in B$ para todo $a \succeq A$ (cf. Exercício 1.31). Exatamente o que se tem em $\mathbb{R}$ ou em $[-\infty, +\infty]$, apenas trocando intervalos abertos por bolas abertas.

Exercício 1.27 ($^{\star}_{\star}$)**.** Prove a Proposição 1.0.13 usando bolas abertas. Dica: imite a Proposição 1.1.11. ■

Exemplo 1.1.22 (Métrica discreta)**.** As bolas abertas dependem das métricas, e estas variam com o contexto. Para piorar, um mesmo conjunto admite muitas métricas diferentes: você já viu acima, por exemplo, que o plano $\mathbb{R}^2$ admite três métricas distintas. Mas essas três não esgotam o repertório.

Sobre qualquer conjunto X, podemos definir uma métrica $d_X : X \times X \to \{0,1\}$ chamada de (métrica) **discreta**, fazendo

$$d_X(x,y) := \begin{cases} 0, & \text{se } x = y \\ 1, & \text{se } x \neq y \end{cases}.$$

Exercício 1.28 ($\star$)**.** Convença-se de que d_X é, de fato, uma métrica em X. ■

Quando tal métrica é considerada em X, há apenas dois tipos de bolas abertas: se $0 < r < 1$, então $B_{d_X}(x, r) = \{x\}$ pois não há $y \neq x$ satisfazendo $d_X(x, y) < 1$; se $r \geq 1$, então $B_{d_X}(x, r) = X$ pois $d(x, y) \leq 1 \leq r$ para quaisquer $x, y \in X$. Na prática, d_X é capaz de detectar apenas quando dois pontos são distintos, sem qualquer outra nuance de distância — diferente do que acontece nos espaços vetoriais normados, por exemplo. ▲

Há um problema aqui, certo? Se um conjunto admite tantas métricas diferentes, como escolher a correta para se trabalhar? Resposta: não há métrica certa. O que pode existir, a depender do caso, é a métrica mais conveniente para um dado contexto. Para espaços vetoriais do tipo $\mathbb{R}^n$, por exemplo, o mais comum é imitar a experiência empírica que temos no *espaço físico*, o que faz da métrica induzida pela norma euclidiana a escolha *canônica*. Mas mesmo esta escolha pode mudar a depender do que você pretende fazer com a métrica.

Se você precisa construir uma casa, é melhor manter a métrica euclidiana em $\mathbb{R}^3$. Porém, se você trabalha no setor de transportes terrestres, pode preferir a métrica da soma: afinal de contas, para quem não pode voar do ponto (a, b) até o ponto (c, d) em linha reta, é mais conveniente conhecer as distâncias de a até c e de b até d, já que este será o trajeto percorrido[†]. Esta é justamente a distância da soma, $d_1((a, b), (c, d)) := \| (a, b) - (c, d) \|_1 = |a - c| + |b - d|$.

Por outro lado, se você pretende usar métricas apenas para analisar *limites no próprio espaço* e coisas do tipo, a situação fica um pouco melhor a depender do conjunto. Um fato surpreendente sobre $\mathbb{R}^n$, por exemplo, é que qualquer métrica induzida por *alguma* norma produz os mesmos resultados nesse aspecto! Embora seja cedo para provar isso, é relativamente fácil se convencer de que a afirmação vale em $\mathbb{R}^2$ para as três métricas discutidas anteriormente.

(i) Supondo que $(x_a)_{a \in \mathbb{A}}$ é uma rede em $\mathbb{R}^2$ que converge para um ponto $L \in \mathbb{R}^2$ com respeito à métrica euclidiana, vamos mostrar que $x_a \to L$ com respeito à métrica do máximo: para isso, precisamos mostrar que para qualquer d_∞-bola aberta B *em torno* de L existe um índice $A \in \mathbb{A}$ tal que $x_a \in B$ para todo $a \succeq A$; ora, uma vez fixada uma dessas bolas B, certamente existe uma bola aberta euclidiana B' com $L \in B'$ e $B' \subseteq B$, o que permite usar a hipótese de que $x_a \to L$ na métrica euclidiana para obter $A \in \mathbb{A}$ tal que $x_a \in B'$ para todo $a \succeq A$.

(ii) Supondo que $(x_a)_{a \in \mathbb{A}}$ é uma rede em $\mathbb{R}^2$ que converge para um ponto $L \in \mathbb{R}^2$ com respeito à métrica do máximo, vamos mostrar que $x_a \to L$ com respeito à métrica da soma: como no caso anterior, o *Katzensprung*[‡] é mostrar que qualquer d_1-bola aberta centrada em L contém uma d_∞-bola aberta centrada em L (dentro de um quadrado há um "losango").

(iii) Supondo que $(x_a)_{a \in \mathbb{A}}$ é uma rede em $\mathbb{R}^2$ que converge para um ponto $L \in \mathbb{R}^2$ com respeito à métrica da soma, vamos mostrar que $x_a \to L$ com respeito à métrica euclidiana... você entendeu a ideia, certo? $(^{\star}_{\star\star})$.

[†] Pelo menos num cenário com condições mínimas de urbanização.

[‡] Pulo do gato, em alemão.

Exemplo 1.1.23 (Convergência não é bagunça)**.** É importante ter em mente que alterar a métrica pode sim alterar os limites. Por exemplo, com a métrica discreta em $\mathbb{R}$, uma sequência $(x_n)_{n\in\mathbb{N}}$ converge para $L \in \mathbb{R}$ se, e somente se, existe $N \in \mathbb{N}$ tal que $x_n = L$ para todo $n \geq N$. Com efeito, $B := B_{d_{\mathbb{R}}}(L, 1)$ é uma bola aberta legítima em torno de L com a métrica discreta, de modo que se $x_n \to L$, então existe $N \in \mathbb{N}$ tal que $x_n \in B$ para todo $n \geq N$. Uma vez que $B = \{L\}$, o resultado segue. ▲

A causa dos fenômenos observados acima tem nome: *topologia*. Em breve, veremos que todo espaço métrico (X, d) vem de fábrica com uma família de subconjuntos $\mathcal{T}_d$ chamada de *topologia*, que por sua vez é a verdadeira responsável por todas as coisas quando o assunto é *convergência* (pelo menos em contextos *elementares*). Em particular, quando duas métricas induzem a mesma *topologia* num espaço, uma rede convergirá com respeito a uma das métricas se, e somente se, convergir com respeito a outra. Assim, o fenômeno observado em $\mathbb{R}^n$ é apenas sintoma de que as métricas oriundas de normas (em $\mathbb{R}^n$!) geram a mesma topologia, que por sua vez é diferente da topologia induzida pela métrica discreta.

Com isso dito, você pode se perguntar: por que alguém trocaria a norma de $\mathbb{R}^n$, uma vez que a norma euclidiana parece ser a mais intuitiva? RESPOSTA: evitar contas! Como ilustração, veja a próxima proposição, com a qual encerramos esta seção.

Proposição 1.1.24. *Seja $\mathbb{R}^2$ com a métrica d induzida por uma das normas usuais $\|\cdot\|_1$, $\|\cdot\|_2$ ou $\|\cdot\|_\infty$. Para uma rede $((x_a, y_a))_{a\in\mathbb{A}}$ em $\mathbb{R}^2$ e um par $(L, L') \in \mathbb{R}^2$, são equivalentes:*

(i) $\lim_{a\in\mathbb{A}} (x_a, y_a) = (L, L')$ *com respeito à métrica d;*

(ii) $\lim_{a\in\mathbb{A}} x_a = L$ *e* $\lim_{a\in\mathbb{A}} y_a = L'$ *em* $\mathbb{R}$.

Demonstração. As discussões anteriores mostram que basta trabalhar com a norma do máximo $\|\cdot\|_\infty$. E com ela a coisa toda fica bem simples.

($\Rightarrow$) Para mostrar que $x_a \to L$, podemos fixar $\varepsilon > 0$ a fim de encontrar $A \in \mathbb{A}$ tal que $x_a \in (L - \varepsilon, L + \varepsilon)$ sempre que $a \succeq A$. Ora, a bola aberta $B_d((L, L'), \varepsilon)$ é formada por todos os pares $(x, y) \in \mathbb{R}^2$ satisfazendo
$$\| (x, y) - (L, L') \|_\infty = \max\{|x - L|, |y - L'|\} < \varepsilon$$
e, por hipótese, existe $A \in \mathbb{A}$ tal que $\| (x_a, y_a) - (L, L') \|_\infty < \varepsilon$ sempre que $a \succeq A$. Em particular, por valer (trivialmente!) $|x_a - L| \leq \| (x_a, y_a) - (L, L') \|_\infty$, obtemos a condição desejada. Analogamente, mostra-se que $y_a \to L'$.

($\Leftarrow$) Para mostrar que $(x_a, y_a) \to (L, L')$ em $\mathbb{R}^2$, fixa-se uma bola aberta $B_d((L, L'), r)$ em torno de (L, L') a fim de encontrar $A' \in \mathbb{A}$ tal que $\| (x_a, y_a) - (L, L') \|_\infty < r$ sempre que $a \succeq A'$. Desta vez, a hipótese assegura $A_0, A_1 \in \mathbb{A}$ tais que $|x_a - L| < r$ se $a \succeq A_0$ e $|y_a - L'| < r$ se $a \succeq A_1$ (pois $x_a \to L$ e $y_a \to L'$, respectivamente). Daí, basta tomar $A' \succeq A_0, A_1$, que existe por $\mathbb{A}$ ser dirigido, acarretando
$$\| (x_a, y_a) - (L, L') \|_\infty = \max\{|x_a - L|, |y_a - L|\} < r$$
sempre que $a \succeq A'$ (certo?)$^{\star}_{\star}$. □

Poderíamos provar a proposição acima sem apelar para a norma do máximo? Certamente. Porém, algumas alterações seriam necessárias: com a norma euclidiana, por exemplo, também ocorre a desigualdade (um pouco *menos óbvia*†)
$$|x - L|, |y - L'| \leq \| (x, y) - (L, L') \|_2 = \sqrt{(x - L)^2 + (y - L')^2},$$

†Mas é *quase* óbvia, certo? $(^{\star}_{\star})$.

necessária para provar a "ida" ($\Rightarrow$), enquanto

$$\| (x, y) - (L, L') \|_2 = \| (x - L, 0) + (0, y - L') \|_2 \leq |x - L| + |y - L'|$$

(por quê?)$^{\star}_{\star}$, que pode ser usada para demonstrar a "volta" ($\Leftarrow$). Embora nada disso seja terrivelmente mais difícil do que a demonstração apresentada, o apego aos cálculos pode ofuscar o verdadeiro conceito: um problema de convergência em $\mathbb{R}^2$ nada mais é do que dois problemas simultâneos de convergência em $\mathbb{R}$! Nesse sentido, trocar as métricas de vez em quando pode ajudar a não perder tempo com o que não interessa!

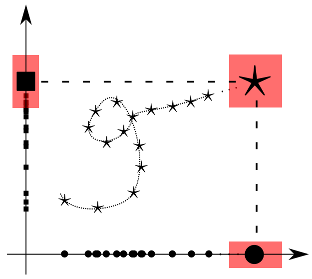


Figura 1.4: As estrelhinhas convergem para a estrelona se, e somente se, os pontinhos convergem para o pontão e os quadradinhos convergem para o quadradão.

§3 Exercícios extras

Exercício 1.29 ($\star$)**.** Mostre que redes constantes em espaços métricos convergem. ■

Exercício 1.30 ($^{\star}_{\star}$)**.** Sejam (X, d) um espaço métrico, $y \in X$ um ponto e $r > 0$. Mostre que se $x \in B_d(y, r)$, então existe $s > 0$ tal que $B_d(x, s) \subseteq B_d(y, r)$. Dica: faça um desenho antes de provar. ■

Exercício 1.31 ($^{\star}_{\star}$)**.** Mostre que uma rede $(x_a)_{a \in \mathbb{A}}$ num espaço métrico X converge para $L \in X$ se, e somente se, para toda bola aberta $B \subseteq X$ com $L \in B$ existir $A \in \mathbb{A}$ tal que $x_a \in B$ para todo $a \succeq A$. Dica: use o exercício anterior para imitar o argumento que você faria se tivesse $X = \mathbb{R}$. ■

Exercício 1.32 (!?)**.** Como definir séries num espaço normado? ■

1.2 A "álgebra ordenada" dos limites

1.2.0 Essencial

Utilizar a definição de convergência para verificar se uma rede (sequência, função, etc.) converge para um ponto escolhido é um processo delicado e artesanal: a depender da rede $(x_d)_d$, pode ser que o método para assegurar o índice D a partir do qual as *distâncias* entre x_d e o candidato a limite se tornem menores do que um $\varepsilon > 0$ dado seja diferente de todos os artifícios conhecidos pela pessoa até o momento — se o método existir, afinal, o ponto escolhido pode não ser limite da rede.

Um modo de (tentar) simplificar isso consiste em expressar uma rede em termos de outras redes cujos limites já são conhecidos (por sorte ou azar). Os dois modos mais comuns são: escrever o termo da rede como uma expressão algébrica envolvendo redes cujos limites são conhecidos; comparar (no sentido de ordem) os termos da rede com os termos de outras redes cujos limites são conhecidos.

§0 Operações com limites reais

Vamos verificar as *propriedades operatórias dos limites*, o que em linguagem humana significa que analisaremos como os limites de redes reais se comportam diante das operações algébricas usuais. Por ser o primeiro contato formal com tais propriedades no texto, o foco inicial será nas situações em que as redes convergem para pontos na reta real.

Lema 1.2.0. *Se $(x_d)_{d\in\mathbb{D}}$ é uma rede real que converge em $\mathbb{R}$, então $|\lim_{d\in\mathbb{D}} x_d| = \lim_{d\in\mathbb{D}} |x_d|$.*

Demonstração. Fazendo $L := \lim_{d\in\mathbb{D}} x_d$, mostraremos que $\lim_{d\in\mathbb{D}} |x_d| = |L|$. Para isso, fixado $\varepsilon > 0$, busca-se $D \in \mathbb{D}$ tal que $||x_d| - |L|| < \varepsilon$ sempre que $d \succeq D$ (certo?)*. Mas isto é muito simples: basta notar que $||\alpha| - |\beta|| \leq |\alpha - \beta|$ para quaisquer $\alpha, \beta \in \mathbb{R}$ (verifique!)$^\star_\star$ e daí apelar para o fato de que existe $D \in \mathbb{D}$ tal que $|x_d - L| < \varepsilon$ sempre que $d \succeq D$. Os detalhes ficam por sua conta. □

Teorema 1.2.1. *Sejam $(x_d)_{d\in\mathbb{D}}$ e $(y_d)_{d\in\mathbb{D}}$ redes reais sobre o mesmo conjunto dirigido $\mathbb{D}$, tais que $x_d \to x$ e $y_d \to y$ para certos $x, y \in \mathbb{R}$.*

(i) *A rede $(z_d)_{d\in\mathbb{D}}$ dada por $z_d := x_d + y_d$ para todo d é tal que $z_d \to x + y$, i.e.,*

$$\lim_{d\in\mathbb{D}}(x_d + y_d) = \lim_{d\in\mathbb{D}} x + \lim_{d\in\mathbb{D}} y_d.$$

(ii) *A rede $(z_d)_{z\in\mathbb{D}}$ dada por $z_d := x_d \cdot y_d$ para todo d é tal que $z_d \to x \cdot y$, i.e.,*

$$\lim_{d\in\mathbb{D}}(x_d \cdot y_d) = \lim_{d\in\mathbb{D}} x_d \cdot \lim_{d\in\mathbb{D}} y_d.$$

(iii) *Se $y \neq 0$ e $y_d \neq 0$ para todo $d \in \mathbb{D}$, então a rede $(z_d)_{d\in\mathbb{D}}$ dada por $z_d := \frac{x_d}{y_d}$ é tal que $z_d \to \frac{x}{y}$, i.e.,*

$$\lim_{d\in\mathbb{D}} \frac{x_d}{y_d} = \frac{\lim\limits_{d\in\mathbb{D}} x_d}{\lim\limits_{d\in\mathbb{D}} y_d}.$$

Demonstração. Ao longo desta prova, ε indica um número real estritamente maior do que 0. Para o primeiro caso, devemos encontrar $D \in \mathbb{D}$ tal que $|z_d - (x + y)| < \varepsilon$ sempre que $d \succeq D$. Ora, temos

$$|z_d - (x + y)| = |x_d + y_d - x - y| = |(x_d - x) + (y_d - y)| \leq \underbrace{|x_d - x|}_{(A)} + \underbrace{|y_d - y|}_{(B)}.$$

Como $x_d \to x$, podemos deixar (A) tão pequeno quanto desejado, bastando para isso ajustar o índice d. Da mesma forma, (B) pode ser feito tão pequeno quanto quisermos pois $y_d \to y$. Queremos $(A) + (B) < \varepsilon$, certo? Então, se fizermos $(A), (B) < \frac{\varepsilon}{2}$, teremos a desigualdade desejada![†] Mais precisamente: como $x_d \to x$ e $y_d \to y$, existem $D_0, D_1 \in \mathbb{D}$ tais que $|x_d - x| < \frac{\varepsilon}{2}$ sempre que $d \succeq D_0$ e $|y_d - y| < \frac{\varepsilon}{2}$ sempre que $d \succeq D_1$; por $\mathbb{D}$ ser

[†] Felizmente, a vida é curta demais para explicar esse tipo de raciocínio em todas as situações em que precisarmos escolher valores "espertos" para épsilons intermediários. "*Mas como eu poderia adivinhar que era para escolher $\frac{\varepsilon}{2}$? A vida é muito cruel!*". Resposta: não adivinhe. Por exemplo, pedindo apenas que $|x_d - x| < \varepsilon'$ e $|y_d - y| < \varepsilon'$, chega-se a $(A) + (B) < 2\varepsilon'$; como o que se busca é $(A) + (B) < \varepsilon$, o problema fica resolvido se $2\varepsilon' < \varepsilon$, ou seja, $\varepsilon' < \frac{\varepsilon}{2}$. De modo geral, é dessa maneira que se "descobrem" valores adequados para os épsilons intermediários. Por que não escrever sempre assim? Simples: papel é caro.

dirigido, existe $D \succeq D_0, D_1$ e daí, por transitividade, se $d \succeq D$, então $d \succeq D_0$ e $d \succeq D_1$, acarretando
$$|z_d - (x+y)| \leq |x_d - x| + |y_d - y| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

O segundo caso é análogo: busca-se $D \in \mathbb{D}$ tal que $|x_d y_d - xy| < \varepsilon$ sempre que $d \succeq D$. A diferença, desta vez, é que precisamos “somar zero” de um jeito *maroto* duas vezes:

✓ primeiro fazemos
$$|x_d y_d - xy| \leq |x_d y_d + (x_d y - x_d y) - xy| \leq |x_d| \cdot |y_d - y| + |y| \cdot |x_d - x|;$$

✓ depois[†], note que $|x_d| = |x_d + (-x + x)| \leq |x_d - x| + |x|$, de modo que em virtude da desigualdade anterior, resulta
$$|x_d y_d - xy| \leq (|x_d - x| + |x|) \cdot |y_d - y| + |y| \cdot |x_d - x| = |x_d - x| \cdot (|y_d - y| + |y|) + |x| \cdot |y_d - y|.$$

Acabou: como $y_d \to y$, existe $D_0 \in \mathbb{D}$ tal que $|y_d - y| < \frac{\varepsilon}{2(|x|+1)} := \alpha$ sempre que $d \succeq D_0$; chamando $\beta := \alpha + |y|$, existe $D_1 \in \mathbb{D}$ tal que $|x_d - x| < \frac{\varepsilon}{2\beta}$ sempre que $d \succeq D_1$ (pois $x_d \to x$); enfim, para $D \succeq D_0, D_1$, verifica-se
$$|x_d y_d - xy| < \frac{\varepsilon}{2\beta} \cdot \beta + |x| \cdot \frac{\varepsilon}{2(|x|+1)} < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$
para todo $d \succeq D$ (certo?)[‡].

Para o terceiro caso, basta mostrar que $\lim_{d \in \mathbb{D}} \frac{1}{y_d} = \frac{1}{y}$, pois daí o item (ii) assegura
$$\lim_{d \in \mathbb{D}} \frac{x_d}{y_d} = \lim_{d \in \mathbb{D}} \left(x_d \cdot \frac{1}{y_d} \right) = x \cdot \frac{1}{y} = \frac{x}{y}.$$

Uma prova simples de que $\lim_{d \in \mathbb{D}} \frac{1}{y_d} = \frac{1}{y}$ requer certo grau de malícia: busca-se $D \in \mathbb{D}$ tal que $\left|\frac{1}{y_d} - \frac{1}{y}\right| < \varepsilon$, e temos
$$\left|\frac{1}{y_d} - \frac{1}{y}\right| = \left|\frac{y - y_d}{y_d \cdot y}\right| = \frac{|y_d - y|}{|y \cdot y_d|}.$$

Enquanto a condição $y_d \to y$ nos permite “controlar” o número $|y_d - y|$ diretamente, não é tão claro como fazer a mesma coisa com $\frac{1}{|y \cdot y_d|}$. A ideia aqui é apelar para a *conservação de sinal*: como $y \neq 0$, existe $\alpha \in \mathbb{R}$ com $0 < \alpha < |y|$ e por valer $|y_d| \to |y|$ (pelo lema anterior), existe $D_0 \in \mathbb{D}$ tal que $\alpha < |y_d|$ para todo $d \succeq D_0$ (por quê?)*. Consequentemente, $\frac{1}{|y| \cdot |y_d|} < \frac{1}{|y|\alpha}$. Agora, por existir $D_1 \in \mathbb{D}$ tal que $|y_d - y| < \varepsilon\alpha|y|$ para todo $d \succeq D_1$, basta tomar $D \succeq D_0, D_1$, pois assim $d \succeq D$ acarreta
$$\left|\frac{1}{y_d} - \frac{1}{y}\right| \leq \frac{|y_d - y|}{|y \cdot y_d|} < \frac{1}{|y|\alpha} \cdot \varepsilon\alpha|y| = \varepsilon. \qquad \square$$

Exercício 1.33 ($\star$). Adapte alguns dos argumentos acima para os casos de sequências e funções reais. ■

[†]Não podemos *pedir* $D \in \mathbb{D}$ tal que $|y_d - y| < \frac{\varepsilon}{2|x_d|}$ sempre que $d \succeq D$ pois x_d também está *atrelado* ao índice d, ou seja: não é uma constante como $|x|$ e $|y|$, que estão *fixados*.

[‡]Pois $\frac{|x|}{|x|+1} < 1$ para todo $x \in \mathbb{R}$ (verifique!)*. Precisa-se fazer isso pois poderia ocorrer $|x| = 0$, o que nos obrigaria a considerar tal caso separadamente.

Observação 1.2.2. A formulação do teorema anterior deve ser levada à sério, no seguinte sentido: se $\lim_{d\in\mathbb{D}} x_d$ e $\lim_{d\in\mathbb{D}} y_d$ são números reais, então o limite de $(z_d)_{d\in\mathbb{D}}$ existe e é dado pela expressão correspondente. **Não se pode assegurar** que os limites de $(x_d)_{d\in\mathbb{D}}$ ou $(y_d)_{d\in\mathbb{D}}$ existem apenas com a garantia de que $(z_d)_{d\in\mathbb{D}}$ converge em $\mathbb{R}$.

Por exemplo: fazendo $x_n := (-1)^n$ e $y_n := (-1)^{n+1}$ para todo $n \in \mathbb{N}$, verifica-se

$$\lim_{n\to\infty}(x_n + y_n) = 0 \quad \text{e} \quad \lim_{n\to\infty}(x_n y_n) = -1,$$

mas tanto $(x_n)_{n\in\mathbb{N}}$ quanto $(y_n)_{n\in\mathbb{N}}$ divergem em $\mathbb{R}$. Apesar disso, esse detalhe é tipicamente omitido, o que pode confundir pessoas incautas. Por exemplo, para $\alpha \in \mathbb{R}$, é comum encontrar por aí

$$\lim_{n\to\infty}\left(\alpha - \frac{1}{2^n}\right)^2 = \left(\lim_{n\to\infty}\left(\alpha - \frac{1}{2^n}\right)\right)^2 = \left(\lim_{n\to\infty}\alpha - \lim_{n\to\infty}\frac{1}{2^n}\right)^2 = (\alpha - 0)^2 = \alpha^2,$$

o que, a rigor, deveria ser interpretado da direita para a esquerda: como $\lim_{n\to\infty}\alpha = \alpha$ e $\lim_{n\to\infty}\frac{1}{2^n} = 0$, resulta $\lim_{n\to\infty}\left(\alpha - \frac{1}{2^n}\right) = \alpha$; logo, $\lim_{n\to\infty}\left(\alpha - \frac{1}{2^n}\right)^2 = \alpha^2$. Todavia, a vida é curta demais para não vivê-la perigosamente. $\triangle$

Se $\lim_{d\in\mathbb{D}} y_d \neq 0$, então não é preciso *exigir* $y_d \neq 0$ para todo d a fim de *estimar* $\lim_{d\in\mathbb{D}}\frac{1}{y_d}$. Na verdade, por existir $D \in \mathbb{D}$ tal que $y_d \neq 0$ sempre que $d \succeq D$ (*déjà vu*?), pode-se considerar o subconjunto $D^\uparrow := \{d \in \mathbb{D} : d \succeq D\}$, que ainda é dirigido pela pré-ordem $\preceq$ de $\mathbb{D}$. Isso permite tratar $(y_d)_{d\in D^\uparrow}$ como uma rede legítima[†]. Daí, não é difícil perceber que para $L \in [-\infty, +\infty]$, tem-se $\lim_{d\in\mathbb{D}} y_d = L$ se, e somente se, $\lim_{d\in D^\uparrow} y_d = L$. Assim, podemos *interpretar* $\lim_{d\in\mathbb{D}}\frac{1}{y_d}$ como sendo $\lim_{d\in D^\uparrow}\frac{1}{y_d}$, mesmo que, a rigor, o número $\frac{1}{y_d}$ não esteja definido para certos índices $d \in \mathbb{D}$.

Exercício 1.34 ($^{\star}_{\star}$)**.** Verifique as afirmações acima. $\blacksquare$

No caso de sequências, por exemplo, se $(y_n)_{n\in\mathbb{N}}$ é tal que $y_n \to y \neq 0$, então existe $N \in \mathbb{N}$ tal que $y_n \neq 0$ sempre que $n \geq N$: em tal situação, para qualquer $\varepsilon > 0$, existe $N' \geq N$ tal que $\left|\frac{1}{y_n} - \frac{1}{y}\right| < \varepsilon$ para qualquer $n \geq N'$. Já no caso de uma função real $f\colon \mathbb{R} \to \mathbb{R}$ e um ponto $p \in \mathbb{R}$ tal que $\lim_{x\to p} f(x) = L \neq 0$, existe $r > 0$ tal que $f(x) \neq 0$ sempre que $0 < |x - p| < r$. Logo, para qualquer $\varepsilon > 0$, exite δ com $0 < \delta < r$ tal que $\left|\frac{1}{f(x)} - \frac{1}{L}\right| < \varepsilon$ para todo x com $0 < |x - p| < \delta$.

Definição 1.2.3. Para simplificar futuras notações, escreveremos $\lim\limits_{d\succeq D} x_d$ para indicar o limite de uma rede $(x_d)_{d\in\mathbb{D}}$ restrita ao subconjunto de índices $D^\uparrow$ sempre que $(x_d)_{d\in\mathbb{D}}$ convergir. ¶

Exemplo 1.2.4. No caso de sequências, podemos reinterpretar limites do tipo $\lim_{n\geq N} x_n$ como limites usuais de sequências, afinal de contas, $\{n \in \mathbb{N} : n \geq N\} = \{N + n : n \in \mathbb{N}\}$, e daí não é difícil se convencer de que

$$\lim_{n\geq N} x_n = \lim_{n\to\infty} x_{N+n}.$$

Pense um pouco antes de discordar ou concordar[‡]. $\blacktriangle$

[†] *Spoiler*: explicitamente, trata-se de uma *sub-rede*, que por sua vez generaliza a ideia de *subsequência*.

[‡] Lembre-se: em Matemática, "pensar" é abreviação para "faça as contas até se convencer". Outra coisa: um panda morre toda vez que alguém confunde $\lim\limits_{n\to\infty} x_{N+n}$ com $\lim\limits_{n\to\infty} x_N + n$.

Exemplo 1.2.5 (Teste da divergência para séries de números reais)**.** As propriedades operatórias de limites dão um critério muito simples para dizer quando uma série $\sum a_n$ de números reais não converge em $\mathbb{R}$.

Proposição 1.2.6. *Se* $\sum a_n$ *converge em* $\mathbb{R}$*, então* $a_n \to 0$.

Demonstração. Chamando $\sum a_n = a$ e $s_n \coloneqq a_0 + \ldots + a_n$ para cada $n \in \mathbb{N}$, temos $s_n \to a$ por definição e, pelo que se discutiu acima,

$$\lim_{n\to\infty} s_{n+1} = a.$$

Agora, por um lado, temos

$$\lim_{n\to\infty} (s_{n+1} - s_n) = \lim_{n\to\infty} (a_0 + \ldots + a_n + a_{n+1} - (a_0 + \ldots + a_n)) = \lim_{n\to\infty} a_{n+1}$$

e, por outro lado,

$$\lim_{n\to\infty} (s_{n+1} - s_n) = \lim_{n\to\infty} s_{n+1} - \lim_{n\to\infty} s_n = a - a = 0,$$

mostrando que $\lim_{n\to\infty} a_{n+1} = \lim_{n\to\infty} a_n = 0$. □

Infelizmente, a recíproca não é verdadeira: oportunamente veremos que $\sum \frac{1}{n+1} = +\infty$, embora se tenha $\lim_{n\to\infty} \frac{1}{n+1} = 0$. ▲

Exemplo 1.2.7. *Para* $a \in \mathbb{R}$ *com* $|a| < 1$*, verifica-se* $\lim_{n\to\infty} a^n = 0$. De fato, se $0 \leq a < 1$, então existe $L \in \mathbb{R}$ com $\lim_{n\to\infty} a_n = L$ pois $(a^n)_{n\in\mathbb{N}}$ é limitada e decrescente (certo?)*. Como $(a^{n+1})_{n\in\mathbb{N}}$ também deve convergir para L, obtemos

$$L = \lim_{n\to\infty} a^{n+1} = \lim_{n\to\infty} a \cdot a^n = \lim_{n\to\infty} a \cdot \lim_{n\to\infty} a^n = aL,$$

acarretando $L(1-a) = 0$ e, consequentemente, $L = 0$. Isto também resolve o caso em que $-1 < a \leq 0$: pela parte anterior, $\lim_{n\to\infty} |a^n| = 0$, mas isto acarreta $\lim_{n\to\infty} a^n = 0$ (por quê?)[†]. Em particular, reobtemos a identidade $\sum_{n=0}^{\infty} a^n = \frac{1}{1-a}$, desta vez para qualquer a com $|a| < 1$: para verificar (faça isso!)*, pode ser útil rever a prova do Lema 0.10.8. ▲

Exercício 1.35 ($^\star_\star$)**.** Mostre que se $f : \mathbb{R} \to \mathbb{R}$ é função polinomial, então

$$\lim_{x\to p} f(x) = f(p)$$

para qualquer $p \in \mathbb{R}$. Dica: indução. ■

§1 Operações com limites na reta estendida

O Teorema 1.2.1 foi enunciado e demonstrado para limites reais. As situações em que as redes convergem para $-\infty$ ou $+\infty$, porém, são mais delicadas: precisa-se de cuidado para estender as propriedades *operatórias* uma vez que $-\infty$ e $+\infty$ não são números reais, o que impede que eles sejam *algebricamente operados* com os membros de $\mathbb{R}$. A reta estendida não é bagunça!

[†]Faça por sua conta agora ($^{\star}_{\star\star}$) ou confira o Exercício 1.76.

Teorema 1.2.8. *Sejam $(x_d)_{d\in\mathbb{D}}$ e $(y_d)_{d\in\mathbb{D}}$ redes de números reais sobre o mesmo conjunto dirigido $\mathbb{D}$, com $x_d \to x$ e $y_d \to y$, tais que $x \in [-\infty, +\infty]$ e $y \in \{-\infty, +\infty\}$.*[†]

(i) *Se $x \neq +\infty$ e $y = -\infty$, então $x_d + y_d \to -\infty$.*

(ii) *Se $x \neq -\infty$ e $y = +\infty$, então $x_d + y_d \to +\infty$.*

(iii) *Se $x > 0$, então $x_d \cdot y_d \to -\infty$ caso $y = -\infty$, e $x_d \cdot y_d \to +\infty$ se $y = +\infty$.*

(iv) *Se $x < 0$, então $x_d \cdot y_d \to +\infty$ caso $y = -\infty$, e $x_d \cdot y_d \to -\infty$ se $y = +\infty$.*

Demonstração. A argumentação fica bastante simplificada se apelarmos para a interpretação "geométrica", i.e., intervalos.

(i) Para um intervalo aberto $I \subseteq [-\infty, +\infty]$ em torno de $-\infty$, digamos $I := [-\infty, -M)$ para algum $M > 0$, busca-se $D \in \mathbb{D}$ tal que $x_d + y_d \in I$ sempre que $d \succeq D$. Há dois casos: se $x = -\infty$, então existem $D_0, D_1 \in \mathbb{D}$ tais que $x_d < -M$ (se $d \succeq D_0$) e $y_d < -M$ (se $d \succeq D_1$), de modo que ao escolher qualquer $D \succeq D_0, D_1$, resulta

$$x_d + y_d < -2M < -M$$

sempre que $d \succeq D$; se $-\infty < x$, então para qualquer $r > x$ com $r > 0$ existe $D_0 \in \mathbb{D}$ tal que $x_d < r$ para todo $d \succeq D_0$ (por quê?)$^{\star}_{\star}$, e também existe $D_1 \in \mathbb{D}$ tal que $y_d < -(M + r)$, de modo que para qualquer $D \in \mathbb{D}$ com $D \succeq D_0, D_1$ se verifica

$$x_d + y_d < r - (M + r) = -M$$

sempre que $d \succeq D$, como queríamos.

(ii) O argumento é análogo e ficará por sua conta $\binom{\star}{\star}$[‡].

(iii) Se $y = +\infty$ e $M > 0$, precisamos encontrar $D \in \mathbb{D}$ tal que $x_d y_d \in (M, +\infty]$ para todo $d \succeq D$. Ora, como $0 < x$, existe $\alpha \in \mathbb{R}$ com $0 < \alpha < x$ e, daí, por $x_d \to x$ e $y_d \to y$, obtemos $D_0, D_1 \in \mathbb{D}$ tais que

$$\begin{aligned} d \succeq D_0 &\Rightarrow \alpha < x_d, \text{ e} \\ d \succeq D_1 &\Rightarrow \frac{M}{\alpha} < y_d, \end{aligned}$$

de modo que para $D \succeq D_0, D_1$, vale $M < x_d y_d$ sempre que $d \succeq D$. O caso $y = -\infty$ é análogo e, por isso, ficará por sua conta.

(iv) Adivinha? Exercício! □

Exercício 1.36 $\binom{\star}{\star}$. Complete a demonstração anterior, e observe que as posições de x e y podem ser trocadas em todos os casos[††]. Se preferir, pode argumentar com sequências ou limites de funções. ■

[†]Leia com atenção: se escrevêssemos $y \in \{A, B\}$, significaria que $y = A$ ou $y = B$. Analogamente, ao escrever $y \in \{-\infty, +\infty\}$, deve-se entender que $y = -\infty$ ou $y = +\infty$.

[‡]Sugestão: se não quiser sujar as mãos, note que os itens (iii) e (iv) garantem $-x_d \to -x \in [-\infty, +\infty)$ e $-y_d \to -\infty$, donde (i) acarreta $-x_d - y_d \to -\infty$ e, novamente por (iv), resulta $x_d + y_d \to +\infty$.

[††]Isto segue da comutatividade das operações em $\mathbb{R}$ e da unicidade dos limites em $[-\infty, +\infty]$.

Exemplo 1.2.9 (As temidas *indeterminações*)**.** Na prática, o que as propriedades operatórias anteriores *fazem*? RESPOSTA: elas permitem determinar, *a priori*, os limites de certas redes a partir dos limites de outras redes possivelmente mais simples, desde que algumas condições algébricas sejam satisfeitas. Assim, uma situação de *indeterminação* é aquela em que não se tem uma regra algébrica que permita determinar o limite *a priori*:

✗ não significa que o limite automaticamente não existe (em $\mathbb{R}$ ou em $[-\infty, +\infty]$);

✗ não significa que o limite é $+\infty$;

✗ não significa que o limite é $-\infty$;

✗ não significa que é proibido se perguntar se o limite existe.

Exercício 1.37 (!!!)**.** Leia os itens acima dez vezes. ■

Por exemplo: se $x_d \to +\infty$ e $y_d \to -\infty$, gostaríamos de atribuir um significado para "$+\infty + (-\infty)$" que permitisse determinar $\lim_{d\in\mathbb{D}}(x_d + y_d)$. Isto é possível? RESPOSTA: não, pois depende do caso!

(i) Chamando $x_n \coloneqq 2n$ e $y_n \coloneqq -n$ para cada $n \in \mathbb{N}$, temos
$$\lim_{n\to\infty}(x_n + y_n) = \lim_{n\to\infty}(2n - n) = \lim_{n\to\infty} n = +\infty.$$

(ii) Chamando $x_n \coloneqq n$ e $y_n \coloneqq -2n$ para cada $n \in \mathbb{N}$, temos
$$\lim_{n\to\infty}(x_n + y_n) = \lim_{n\to\infty}(n - 2n) = \lim_{n\to\infty} -n = -\infty.$$

(iii) Para $r \in \mathbb{R}$ fixado e chamando $x_n \coloneqq n + r$ e $y_n \coloneqq -n$ para cada $n \in \mathbb{N}$, temos
$$\lim_{n\to\infty}(x_n + y_n) = \lim_{n\to\infty}(n + r - n) = \lim_{n\to\infty} r = r.$$

(iv) Chamando $x_n \coloneqq n$ e $y_n \coloneqq -n + (-1)^n$ para cada $n \in \mathbb{N}$, temos $x_n + y_n = (-1)^n$ para todo n, e $((-1)^n)_{n\in\mathbb{N}}$ é uma sequência divergente em $[-\infty, +\infty]$.

Logo, para qualquer valor L que você pensava que deveria ser atribuído a "$+\infty+(-\infty)$", existem sequências $(x_n)_n$ e $(y_n)_n$ tais que $x_n \to +\infty$, $y_n \to -\infty$ e $\lim_{n\to\infty}(x_n + y_n) \neq L$. O mesmo tipo de fenômeno ocorre com outras expressões, como "$0 \cdot (\pm\infty)$". Lembre-se de alguns exemplos por conta própria ($^\star_\star$). ▲

Limites de redes do tipo $\left(\frac{x_d}{y_d}\right)_{d\in\mathbb{D}}$, com $(x_d)_{d\in\mathbb{D}}$ e $(y_d)_{d\in\mathbb{D}}$ convergentes na reta estendida, também podem ser (pré) determinados em algumas situações específicas.

Proposição 1.2.10. *Sejam $(x_d)_{d\in\mathbb{D}}$ e $(y_d)_{d\in\mathbb{D}}$ redes de números reais sobre o mesmo conjunto dirigido $\mathbb{D}$, com $x_d \to x$ e $y_d \to y$. Se $x \in \mathbb{R}$ e $y \in \{-\infty, +\infty\}$, então $\lim_{d\in\mathbb{D}} \dfrac{x_d}{y_d} = 0$.*

Demonstração. Primeiro, observe que apesar de se ter $y_d \to y \neq 0$, poderia ocorrer $y_d = 0$ para *alguns* índices. No entanto, existe $D \in \mathbb{D}$ tal que $y_d \neq 0$ para todo $d \succeq D$ (certo?)[†], de modo que mostraremos, na verdade, $\lim_{d\succeq D} \frac{x_d}{y_d} = 0$.

[†]Este argumento foi usado algumas vezes no texto. Se ainda não parecer claro porque isto pode ser feito, espere até a discussão de monotonicidade, que virá a seguir.

Segundo, note que é suficiente assegurar $\lim_{d\succeq D}\frac{1}{y_d}=0$: se isso estiver provado, então $(x_d)_{d\succeq D}$ e $\left(\frac{1}{y_d}\right)_{d\succeq D}$ serão redes reais cujos limites são números reais e, portanto,

$$\lim_{d\succeq D}\frac{x_d}{y_d}=\lim_{d\succeq D}x_d\cdot\lim_{d\succeq D}\frac{1}{y_d}=x\cdot 0=0.$$

Finalmente, provemos $\frac{1}{y_d}\to 0$: para $\varepsilon>0$, $I:=\left[-\infty,-\frac{1}{\varepsilon}\right)$ é um intervalo aberto em torno de $-\infty$, de modo que se $y_d\to-\infty$, então existe $D'\succeq D$ tal que $d\succeq D'$ acarreta $y_d<-\frac{1}{\varepsilon}$ ou, equivalentemente[†],

$$-\varepsilon<\frac{1}{y_d}<0<\varepsilon,$$

i.e., $\left|\frac{1}{y_d}\right|<\varepsilon$. Para o caso em que $y_d\to+\infty$, basta repetir o argumento com $I:=\left(\frac{1}{\varepsilon},+\infty\right]$. Os detalhes ficam por sua conta ($^\star_\star$). □

O caso em que $y_d\to 0$ é um pouco mais delicado do que o anterior pois a multiplicação *sente* o *sinal* dos termos y_d, i.e., a direção por meio da qual os pontos se aproximam de 0. Por exemplo, após estudarmos o *Teorema do Confronto* (o que ocorrerá muito em breve), poderemos concluir de maneira indolor que

$$\lim_{n\to\infty}\left(-\frac{1}{2}\right)^n=0,$$

ao passo que $((-1)^n2^n)_{n\in\mathbb{N}}$ diverge em $[-\infty,+\infty]$: para n par ocorre $(-1)^n2^n>0$ e, para n ímpar, $(-1)^n2^n<0$. Porém, nos casos em que os sinais de $(y_d)_{d\in\mathbb{D}}$ se estabilizam num dos lados de 0, é possível determinar (*a priori*) o limite de $\left(\frac{1}{y_d}\right)_{d\in\mathbb{D}}$, o que por sua vez permite determinar o limite de redes do tipo $\left(\frac{x_d}{y_d}\right)_{d\in\mathbb{D}}$ num leque um pouco maior de situações (cf. Exercício 1.41).

Proposição 1.2.11. *Seja $(y_d)_{d\in\mathbb{D}}$ uma rede de números reais tal que $y_d\to 0$.*

(i) *Se existe $D\in\mathbb{D}$ tal que $y_d>0$ para todo $d\succeq D$, então $\lim_{d\in\mathbb{D}}\frac{1}{y_d}=+\infty$.*

(ii) *Se existe $D\in\mathbb{D}$ tal que $y_d<0$ para todo $d\succeq D$, então $\lim_{d\in\mathbb{D}}\frac{1}{y_d}=-\infty$.*

Demonstração. Pelo item (iv) do Teorema 1.2.8, basta mostrar o primeiro caso: para $M>0$, a hipótese assegura $D'\succeq D$ tal que $0<y_d<\frac{1}{M}$ sempre que $d\succeq D'$ (por quê?)*, donde segue que $M<\frac{1}{y_d}$. Em outras palavras, mostramos que para todo intervalo aberto I em torno de $+\infty$ existe D' tal que $\frac{1}{y_d}\in I$ sempre que $d\succeq D'$. □

Exercício 1.38 ((?!))**.** Esboce o gráfico da função $g\colon\mathbb{R}\setminus\{0\}\to\mathbb{R}\setminus\{0\}$ dada por $g(x):=\frac{1}{x}$ e reflita sobre os resultados anteriores. ■

[†]Pois $y_d<0$ em tais casos.

§2 Monotonicidade, sanduíche e confronto

Um modo alternativo de mostrar que $2^n \to +\infty$ em $[-\infty, +\infty]$ (cf. Exercício 1.20) consiste em apelar para a desigualdade $n < 2^n$, válida para qualquer $n \in \mathbb{N}$: dado um intervalo I em torno de $+\infty$, digamos $I := (M, +\infty]$ para algum $M > 0$, resulta que $2^n \in I$ sempre que $n \geq M$, pois $2^n > n \geq M$, i.e., $2^n \in I$.

Em certo sentido, os limites das sequências $(n)_{n\in\mathbb{N}}$ e $(2^n)_{n\in\mathbb{N}}$ *concordam* com o comportamento dos seus termos: verifica-se $n \leq 2^n$ para n *suficientemente bom*† e

$$+\infty = \lim_{n\to\infty} n \leq \lim_{n\to\infty} 2^n = +\infty.$$

Outros casos similares já passaram por aqui:

(i) $0 \leq \frac{1}{2^n}$ para todo n e $0 = \lim_{n\to\infty} 0 \leq \lim_{n\to\infty} \frac{1}{2^n}$;

(ii) $1 - \frac{1}{2^n} \leq 1$ para todo n e $\lim_{n\to\infty}\left(1 - \frac{1}{2^n}\right) \leq \lim_{n\to\infty} 1 = 1$;

(iii) para $a \in (0,1)$, $1 + a + \ldots + a^n \leq \frac{1}{1-a}$ para todo n e $\sum a^n \leq \frac{1}{1-a}$.

Os exemplos acima *sugerem* que algo mais geral ocorre: se $x_d \to x$, $y_d \to y$ e vale $x_d \leq y_d$ para d suficientemente bom, então também vale $x \leq y$. Na prática:

$$x_d \leq y_d \text{ para } d \text{ suficientemente bom} \xRightarrow[\text{“passando o limite”}^{(\heartsuit)}]{} \lim_{d\in\mathbb{D}} x_d \leq \lim_{d\in\mathbb{D}} y_d$$

onde $(\heartsuit)$ também costuma ser substituído por expressões como "fazendo n tender a infinito" (no caso de sequências) ou "tomando x suficientemente próximo de p" (no caso de funções), etc. A sugestão está certa, e se deve ao Exercício 0.104 (cf. Figura 1.5, a seguir).

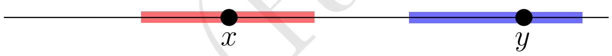


Figura 1.5: Pontos no intervalo vermelho são estritamente menores do que os pontos no intervalo azul.

Teorema 1.2.12. *Sejam $(x_d)_{d\in\mathbb{D}}$ e $(y_d)_{d\in\mathbb{D}}$ redes reais sobre o mesmo conjunto dirigido $\mathbb{D}$, tais que $x_d \to x$ e $y_d \to y$ para certos $x, y \in [-\infty, +\infty]$.*

(i) *(Monotonicidade): se existe $D \in \mathbb{D}$ tal que $x_d \leq y_d$ para todo $d \succeq D$, então $x \leq y$.*

(ii) *(Conservação): se $x < y$, então existe $D \in \mathbb{D}$ tal que $x_d < y_d$ para todo $d \succeq D$.*

Demonstração. Tratemos de (i) pela contrapositiva. Supondo $y < x$ (negou-se a tese...), existem intervalos abertos e disjuntos $I, J \subseteq [-\infty, +\infty]$ com $y \in I$ e $x \in J$. Agora,

- como $y_d \to y$, existe $D_0 \in \mathbb{D}$ tal que $y_d \in I$ sempre que $d \succeq D_0$ e,
- como $x_d \to x$, também existe $D_1 \in \mathbb{D}$ tal que $x_d \in J$ sempre que $d \succeq D_1$.

Note que ao escolher $D_2 \in \mathbb{D}$ melhor do que D_0 e D_1, o Exercício 0.104 garante que $y_d < x_d$ para todo $d \succeq D_2$ (por quê?)*. Em particular, se existisse $D \in \mathbb{D}$ como no enunciado do item (i), poderíamos tomar $d \succeq D_2, D$, mas isto acarretaria $y_d < x_d$ e $x_d \leq y_d$ simultaneamente, o que é absurdo. Portanto, não existe tal D (negou-se a hipótese...).

O item (ii) é um pouco mais fácil. Tomando novamente intervalos abertos e disjuntos $I, J \subseteq [-\infty, +\infty]$ tais que $x \in I$ e $y \in J$, o Exercício 0.104 assegura que $a < b$ sempre que $a \in I$ e $b \in J$. Logo, basta escolher $D \in \mathbb{D}$ de modo que se garanta $x_d \in I$ e $y_d \in J$ para todo $d \succeq D$, o que pode ser feito pois $x_d \to x$ e $y_d \to y$ (convença-se disso!)*. □

†Lembrando que, neste caso, melhor = maior no sentido usual.

Corolário 1.2.13 (Conservação de sinal)**.** *Seja $(z_d)_{d\in\mathbb{D}}$ uma rede real que converge para um ponto $L \in [-\infty, +\infty]$.*

(i) *Se $L > 0$, então existe $D \in \mathbb{D}$ tal que $z_d > 0$ para todo $d \succeq D$.*

(ii) *Se $L < 0$, então existe $D \in \mathbb{D}$ tal que $z_d < 0$ para todo $d \succeq D$.*

Demonstração. Segue do item (ii) do teorema anterior: para $L > 0$, basta fazer $x_d := z_d$ e $y_d := 0$ para todo $d \in \mathbb{D}$; para $L < 0$, faça $x_d := 0$ e $y_d := z_d$ para todo $d \in \mathbb{D}$. □

Corolário 1.2.14 (Conservação de sinal, para sequências)**.** *Seja $(x_n)_{n\in\mathbb{N}}$ uma sequência de números reais que converge para $L \in [-\infty, +\infty]$.*

(i) *Se $L > 0$, então existe $N \in \mathbb{N}$ tal que $x_n > 0$ para todo $n \geq N$.*

(ii) *Se $L < 0$, então existe $N \in \mathbb{N}$ tal que $x_n < 0$ para todo $n \geq N$.*

Corolário 1.2.15 (Conservação de sinal, para funções)**.** *Sejam $f : \mathbb{R} \to \mathbb{R}$ uma função, $p \in \mathbb{R}$ um ponto e $L \in [-\infty, +\infty]$ tal que $\lim_{x\to p} f(x) = L$.*

(i) *Se $L > 0$, então existe $\delta > 0$ tal que $f(x) > 0$ para todo $x \in \mathbb{R} \setminus \{p\}$ tal que $|x - p| < \delta$.*

(ii) *Se $L < 0$, então existe $\delta > 0$ tal que $f(x) < 0$ para todo $x \in \mathbb{R} \setminus \{p\}$ tal que $|x - p| < \delta$.*

E assim por diante...

Exercício 1.39 ($\star$)**.** Convença-se de que os dois últimos corolários são instâncias particulares do Corolário 1.2.13. Interprete graficamente. ■

Exemplo 1.2.16 (Como se livrar do truque de Rudin (cf. Observação 0.9.4))**.** Vamos mostrar que o conjunto $A := \{x \in \mathbb{Q} : x < 0 \text{ ou } x^2 < 2\}$ satisfaz $(\sup A)^2 = 2$ (em $\mathbb{R}$) sem apelar para truques ou cartolas. Primeiro, existe $\alpha \in \mathbb{R}$ com $\sup A = \alpha$ pois $\mathbb{R}$ é corpo completo (cf. Exemplo 0.9.2). Agora, para qualquer $n \in \mathbb{N}$, tem-se

$$\alpha - \frac{1}{2^n} < \alpha < \alpha + \frac{1}{2^n}.$$

Da definição de supremo, a primeira desigualdade assegura uma testemunha $a_n \in A$ de que $\alpha - \frac{1}{2^n}$ não é supremo de A, i.e., $\alpha - \frac{1}{2^n} < a_n$, enquanto a segunda desigualdade implica que $\alpha + \frac{1}{2^n} \notin A$ (o contrário daria $\frac{1}{2^n} \leq 0$). Logo,

$$\left(\alpha - \frac{1}{2^n}\right)^2 < a_n^2 < 2 \leq \left(\alpha + \frac{1}{2^n}\right)^2$$

para qualquer $n \in \mathbb{N}$ (pois $a_n \in A$ com $a_n > 0$, uma vez que $\alpha \geq 1$). Ao "fazer n tender a infinito", resulta (cf. Observação 1.2.2)

$$\alpha^2 = \lim_{n\to\infty} \left(\alpha - \frac{1}{2^n}\right)^2 \leq 2 \leq \lim_{n\to\infty} \left(\alpha + \frac{1}{2^n}\right)^2 = \alpha^2$$

como desejado. ▲

Exercício 1.40 ($\overset{\star}{\star}$)**.** Use o raciocínio anterior para resolver o Exercício 0.135. ■

Observação 1.2.17. Note que não é possível concluir $\lim_{d\in\mathbb{D}} x_d < \lim_{d\in\mathbb{D}} y_d$ sabendo-se apenas que $x_d < y_d$ para todo $d \in \mathbb{D}$. Já vimos alguns exemplos ao longo do texto. Encontre alguns deles para ter certeza. $\triangle$

Intuitivamente, a convergência de uma rede "por baixo" de outra empurra o limite da segunda para cima (se este existir). Em particular, se duas redes que convergem para o mesmo ponto "sanduicharem" uma terceira rede, então...

Proposição 1.2.18 (Sanduíche). *Sejam $(x_d)_{d\in\mathbb{D}}$, $(y_d)_{d\in\mathbb{D}}$ e $(z_d)_{d\in\mathbb{D}}$ redes reais sobre o mesmo conjunto dirigido $\mathbb{D}$ e considere $L \in [-\infty, +\infty]$. Se $\lim\limits_{d\in\mathbb{D}} x_d = \lim\limits_{d\in\mathbb{D}} z_d = L$ e existe $D \in \mathbb{D}$ com $x_d \leq y_d \leq z_d$ para todo $d \succeq D$, então $y_d \to L$.*

Demonstração. Dado um intervalo aberto $I \subseteq [-\infty, +\infty]$ com $L \in I$, busca-se $D \in \mathbb{D}$ tal que $y_d \in I$ para todo $d \succeq D$. Ora, como $x_d \to L$ e $z_d \to L$, existem $D', D'' \in \mathbb{D}$ satisfazendo

$$d \succeq D' \Rightarrow x_d \in I \quad \text{e} \quad d \succeq D'' \Rightarrow z_d \in I.$$

Por $\mathbb{D}$ ser dirigido, existe $\widetilde{D} \succeq D, D', D''$. Logo, se $d \succeq \widetilde{D}$, então $x_d \leq y_d \leq z_d$ com $x_d, z_d \in I$, donde o fato de I ser intervalo (cf. Definição 0.8.13) acarreta $y_d \in I$, como desejado. $\square$

Argumentos de "sanduíche" são muito comuns e, a depender da fome, podem ser usados até mesmo em situações em que não parece haver pães.

Proposição 1.2.19 (Confronto). *Sejam $(x_d)_{d\in\mathbb{D}}$ e $(y_d)_{d\in\mathbb{D}}$ redes reais sobre o mesmo conjunto dirigido $\mathbb{D}$, com $y_d \to 0$. Se existe $D \in \mathbb{D}$ tal que $\{x_d : d \succeq D\}$ é um subconjunto limitado de $\mathbb{R}$, então $x_d y_d \to 0$.*

Demonstração. Pelo Lema 1.2.0, $|y_d| \to 0$. Agora, para $M > 0$ com $|x_d| < M$ para todo $d \succeq D$, note que $-M|y_d| \leq |x_d y_d| \leq M|y_d|$, com $\lim\limits_{d\in\mathbb{D}} -M|y_d| = \lim\limits_{d\in\mathbb{D}} M|y_d| = 0$. Logo, $|x_d y_d| \to 0$ pela proposição anterior e, novamente pela proposição anterior, $x_d y_d \to 0$, haja vista valer $-|x_d y_d| \leq x_d y_d \leq |x_d y_d|$. $\square$

Exercício 1.41 (Confronto no infinito — $(^{\star}_{\star\star})$). Prove as afirmações a seguir.

a) Sejam $(x_n)_{n\in\mathbb{N}}$ e $(y_n)_{n\in\mathbb{N}}$ sequências reais. Se existe $C > 0$ tal que $C < |x_n|$ para todo n, existe $N \in \mathbb{N}$ tal que $x_n \cdot y_n > 0$ para todo $n \in \mathbb{N}$ com $n \geq N$ e $\lim_{n\to\infty} y_n = 0$, então $\lim_{n\to\infty} \frac{x_n}{y_n} = +\infty$.

b) Sejam $f, g \colon \mathbb{R} \to \mathbb{R}$ funções reais e $p \in \mathbb{R}$. Se existe $C > 0$ tal que $C < |f(x)|$ para todo x, existe $r > 0$ tal que $f(x) \cdot g(x) < 0$ para todo $x \in \mathbb{R}$ tal que $0 < |x - p| < r$ e $\lim_{x\to p} g(x) = 0$, então $\lim_{x\to p} \frac{f(x)}{g(x)} = -\infty$.

c) Sejam $(x_d)_{d\in\mathbb{D}}$ e $(y_d)_{d\in\mathbb{D}}$ redes reais. Se existe $C > 0$ tal que $C < |x_d|$ para todo d, existe $D \in \mathbb{D}$ tal que $x_d \cdot y_d > 0$ para todo $d \succeq D$ e $y_d \to 0$, então $\lim_{d\in\mathbb{D}} \frac{x_d}{y_d} = +\infty$.

d) Sejam $(x_d)_{d\in\mathbb{D}}$ e $(y_d)_{d\in\mathbb{D}}$ redes reais. Se existe $C > 0$ tal que $C < |x_d|$ para todo d, existe $D \in \mathbb{D}$ tal que $x_d \cdot y_d < 0$ para todo $d \succeq D$ e $y_d \to 0$, então $\lim_{d\in\mathbb{D}} \frac{x_d}{y_d} = -\infty$

Dicas: i) se você se sentir confortável com redes, basta provar os dois últimos itens, pois os outros são corolários; ii) em vez de investigar $\frac{x_d}{y_d}$, atente-se para o que a proposição anterior tem a dizer sobre o comportamento de $\frac{y_d}{x_d}$, e daí retorne para o caso desejado por meio da Proposição 1.2.11. $\blacksquare$

DIVAGAÇÃO. Há uma pergunta bastante inocente que talvez você não tenha feito até agora por manter algum tipo de respeito implícito pelas in(sti)tuições físicas: se permitimos que as redes reais convirjam para pontos na reta estendida, por que não considerar redes *na reta estendida*? Mais precisamente: por que não considerar coisas do tipo $(x_d)_{d\in\mathbb{D}}$ com $x_d \in [-\infty, +\infty]$ e, possivelmente, $x_d \in \{-\infty, +\infty\}$? RESPOSTA: se você quiser, pode!

Embora essa divagação não traga vantagens do ponto de vista algébrico, já que $[-\infty, +\infty]$ não é um corpo, há pelo menos uma aplicação útil do ponto de vista de ordem: uma vez que a demonstração apresentada para o *Teorema do Sanduíche* não fez uso de qualquer coisa algébrica de $\mathbb{R}$, as redes $(x_d)_d$, $(y_d)_d$ e $(z_d)_d$ no enunciado poderiam ser tomadas em $[-\infty, +\infty]$. Em particular, fazendo $z_d := +\infty$ para todo d temos $z_d \to +\infty$, de modo que se $\lim_{d\in\mathbb{D}} x_d = +\infty$ e $x_d \leq y_d$ para d suficientemente bom, conclui-se que $y_d \to +\infty$! Ao repetir as considerações para $z_d := -\infty$, chega-se ao seguinte resultado útil, que pode ser provado sem a gambiarra acima (faça isso, se quiser)$^{\star}_{\star}$ — mas a que custo, não é mesmo?

Corolário 1.2.20 (Explosão/implosão). *Sejam $(x_d)_{d\in\mathbb{D}}$ e $(y_d)_{d\in\mathbb{D}}$ redes reais sobre um mesmo conjunto dirigido $\mathbb{D}$. Se existe $D \in \mathbb{D}$ tal que $x_d \leq y_d$ para todo $d \succeq D$, então*

(i) *$y_d \to +\infty$ se ocorrer $x_d \to +\infty$, e*

(ii) *$x_d \to -\infty$ se ocorrer $y_d \to -\infty$.*

Exercício 1.42 $\left(^{\star}_{\star}\right)$**.** Sejam $(x_n)_{n\in\mathbb{N}}$ e $(y_n)_{n\in\mathbb{N}}$ sequências de números reais tais que $x_n \leq y_n$ para n suficientemente grande[†].

a) Mostre que se $\sum x_n = +\infty$, então $\sum y_n = +\infty$.

b) Mostre que se $\sum y_n = -\infty$, então $\sum x_n = -\infty$. ■

1.2.1 Extras

§0 Integrais de Riemann como limites de redes (parte I)

Para $a, b \in \mathbb{R}$ com $a < b$, xingaremos de **partição do intervalo** $[a, b]$ qualquer sequência finita de números reais $\mathcal{P} := (a_0, \ldots, a_n)$ com $n > 0$ e $a := a_0 < a_1 < \ldots < a_{n-1} < a_n := b$. Embora, a rigor, não sejam as mesmas *partições* introduzidas na Definição 0.1.22, há uma relação clara entre ambas: se $\mathcal{P} := (a_0, \ldots, a_n)$ é uma partição do *intervalo* $[a, b]$, então $\mathcal{P}' := \{[a_i, a_{i+1}) : i < n\}$ é uma partição do *conjunto* $[a, b)$ (certo?!)*.

Agora, fixada uma partição $\mathcal{P} := (a_0, \ldots, a_n)$ do intervalo $[a, b]$, uma ***tag***[‡] de $\mathcal{P}$ é uma sequência $T := (t_1, \ldots, t_n)$ de números reais satisfazendo $a_{i-1} \leq t_i \leq a_i$ para cada $i \in \{1, \ldots, n\}$. A coleção dos pares $(\mathcal{P}, T)$ em que $\mathcal{P}$ é uma partição do intervalo $[a, b]$ e T é uma *tag* de $\mathcal{P}$ será denotada por $\mathrm{Par}_{\mathcal{R}}[a, b]$, a família das **partições de Riemann** do intervalo $[a, b]$. Você já deve imaginar onde isso tudo vai dar.

Definição 1.2.21. Seja $f : [a, b] \to \mathbb{R}$ uma função e $(\mathcal{P}, T)$ uma partição de Riemann do intervalo. A **soma de Riemann** da função f com respeito à partição $(\mathcal{P}, T)$ é o número real

$$\sum_{(\mathcal{P},T)} f := \sum_{i=1}^{n} f(t_i)(a_i - a_{i-1}),$$

[†]Isto é: existe $N \in \mathbb{N}$ tal que $x_n \leq y_n$ para todo $n \geq N$.

[‡]Nacionalistas podem preferir expressões (mais longas) em português, como *marcação*, *etiqueta*, *pontilhamento*, etc. Particularmente, tenho mais apego ao tempo do que ~~à bandeira~~ ao léxico.

onde $\mathcal{P} := (a_0, \ldots, a_n)$ e $T := (t_1, \ldots, t_n)$. ¶

Quando a função f é não negativa, pode-se pensar em $\sum_{(\mathcal{P},T)} f$ como uma aproximação *tosca* do que seria a área da região plana determinada pelo gráfico de f com o eixo *horizontal*. Intuitivamente, a fim de tornar a aproximação menos *tosca*, i.e., torná-la uma aproximação *melhor* do que seria a área *real* da figura, basta *refinar* as partições, no sentido de acrescentar mais pontos a ela, processo em que se diminuem os *tamanhos* dos subintervalos da forma $[a_i, a_{i+1})$.

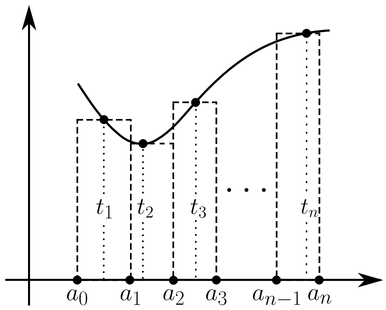


Figura 1.6: Ilustração típica de uma soma de Riemann.

Para tornar mais precisas as considerações acima, vamos associar a cada partição $\mathcal{P} := (a_0, \ldots, a_n)$ do intervalo $[a, b]$ o número $\|\mathcal{P}\| := \max\{a_i - a_{i-1} : 1 \leq i \leq n\}$, que xingaremos de **norma da partição**.

Definição 1.2.22. Diremos que $f\colon [a, b] \to \mathbb{R}$ é **Riemann-integrável** se existir $L \in \mathbb{R}$ com a seguinte propriedade: para todo $\varepsilon > 0$, existe $\delta > 0$ tal que para qualquer partição de Riemann $(\mathcal{P}, T)$ de $[a, b]$ se tenha $\left|L - \sum_{(\mathcal{P},T)} f\right| < \varepsilon$ sempre que $\|\mathcal{P}\| < \delta$. ¶

Em contextos que tratam unicamente do limite de sequências e funções com domínio real, a definição acima não costuma ser muito prática, posto que ela consiste numa *terceira forma de limite*†. Porém, as ferramentas apresentadas até agora permitem interpretar a Definição 1.2.22 como a mera exigência de que uma rede apropriada converge: ao declarar $(\mathcal{P}, T) \preceq (\mathcal{P}', T')$ se, e somente se, $\|\mathcal{P}'\| \leq \|\mathcal{P}\|$, resulta que $(\mathrm{Par}_{\mathcal{R}}\,[a, b], \preceq)$ é um conjunto dirigido, o que segue essencialmente da totalidade da ordem de $\mathbb{R}$. Com efeito, para $(\mathcal{P}, T), (\mathcal{P}', T') \in \mathrm{Par}_{\mathcal{R}}\,[a, b]$, basta ver que $\|\mathcal{P}\| \leq \|\mathcal{P}'\|$ ou $\|\mathcal{P}'\| \leq \|\mathcal{P}\|$.

Observação 1.2.23. Existe outra forma bastante natural de *dirigir* $\mathrm{Par}_{\mathcal{R}}\,[a, b]$ que é frequentemente útil: para partições $\mathcal{P} := (a_0, \ldots, a_n)$ e $\mathcal{P}' := (a'_0, \ldots, a'_m)$ de $[a, b]$ com *tags* T e T', respectivamente, vamos escrever tanto $\mathcal{P} \sqsubseteq \mathcal{P}'$ quanto $(\mathcal{P}, T) \sqsubseteq (\mathcal{P}', T')$ para indicar que

$$\mathrm{im}\,(\mathcal{P}) = \{a_0, \ldots, a_n\} \subseteq \{a'_0, \ldots, a'_m\} = \mathrm{im}\,(\mathcal{P}'),$$

†Talvez por isso seja comum substituí-la pela noção (equivalente) de *integrabilidade de Darboux*, como feito em [24, 25], que se baseia nas noções de supremo e ínfimo.

já que escrever apenas $\mathcal{P} \subseteq \mathcal{P}'$ seria imoral. Agora, não é difícil perceber que $(\mathrm{Par}_{\mathcal{R}}[a,b], \sqsubseteq)$ é um conjunto dirigido: para a condição de refinamento, dadas partições $\mathcal{P}$ e $\mathcal{P}'$ como acima, basta definir $\mathcal{P} \sqcup \mathcal{P}' := (p_0, \ldots, p_{k-1})$, onde $k := |\operatorname{im}(\mathcal{P}) \cup \operatorname{im}(\mathcal{P}')|$, $p_0 := a$ e, para cada $i < k$, $p_{i+1} := \min((\operatorname{im}(\mathcal{P}) \cup \operatorname{im}(\mathcal{P}')) \setminus \{p_0, \ldots, p_i\})$ pois, com isso, verifica-se $\mathcal{P}, \mathcal{P}' \sqsubseteq \mathcal{P} \sqcup \mathcal{P}'$ (certo?!)[†]. As relações entre $\preceq$ e $\sqsubseteq$ são discutidas a seguir. △

Exercício 1.43 ($^\star_\star$). Sejam $(\mathcal{Q}, R)$ e $(\mathcal{S}, T)$ partições de Riemann do intervalo $[a,b]$.

a) Mostre que se $(\mathcal{Q}, R) \sqsubseteq (\mathcal{S}, T)$, então $\|\mathcal{S}\| \leq \|\mathcal{Q}\|$. Conclua que $(\mathcal{Q}, R) \preceq (\mathcal{S}, T)$.

b) Para $a := 0$ e $b := 10$, considere $\mathcal{P} := (0, 5, 10)$ e $\mathcal{Q} := (0, 3, 6, 10)$, com *tags* T e T' quaisquer. Mostre que $\mathcal{P} \not\sqsubseteq \mathcal{Q}$, mas $\|\mathcal{P}\| \geq \|\mathcal{Q}\|$. Conclua que $(\sqsubseteq) \subsetneq (\preceq)$. ■

Moralmente, a pré-ordem $\preceq$ determina que conforme as normas das partições diminuem, elas se tornam *melhores* ou mais próximas do que seria uma partição *ideal*, justamente o que se espera de um conjunto dirigido. Para facilitar futuras referências, diremos que uma partição $\mathcal{P}$ é **mais fina** do que outra partição $\mathcal{P}'$ se ocorrer $\|\mathcal{P}'\| \leq \|\mathcal{P}\|$, nomenclatura que também será utilizada para partições de Riemann — já que as *tags* não influenciam a ordenação. Em particular, ganha-se o direito de considerar redes indexadas por $\mathrm{Par}_{\mathcal{R}}[a,b]$, o que permite enunciar o próximo

Teorema 1.2.24. *Para uma função $f\colon [a,b] \to \mathbb{R}$, as seguintes afirmações são equivalentes:*

(i) *f é Riemann-integrável;*

(ii) *a rede $\left(\sum_{(\mathcal{P},T)} f\right)_{(\mathcal{P},T) \in \mathrm{Par}_{\mathcal{R}}[a,b]}$ converge em $\mathbb{R}$.*

Demonstração. Com f Riemann-integrável e $\varepsilon > 0$, para *um* número real L como na Definição 1.2.22, existe $\delta > 0$ tal que $\left|L - \sum_{(\mathcal{P},T)} f\right| < \varepsilon$ sempre que $\|\mathcal{P}\| < \delta$. Logo, para mostrar que a rede converge, basta exibir uma partição de Riemann $(\mathcal{P}', T')$ para $[a,b]$ com $\|\mathcal{P}'\| < \delta$, pois daí sempre que $(\mathcal{P}, T) \succeq (\mathcal{P}', T')$ teremos $\left|L - \sum_{(\mathcal{P},T)} f\right| < \varepsilon$. Ora, tomando $N \in \mathbb{N}$ com $N > \frac{b-a}{\delta}$, que existe por $\mathbb{R}$ ser arquimediano (certo?)*, considere $\mathcal{P}' := (a_0, \ldots, a_N)$ onde $a_j := a + j \cdot \frac{b-a}{N}$ para cada $j \leq N$, e tome T' uma *tag* qualquer em $\mathcal{P}'$, como por exemplo $T' := (a_1, \ldots, a_N)$: por construção, $\|\mathcal{P}'\| = \frac{b-a}{N} < \delta$. Para a recíproca, com a rede convergindo para $L \in \mathbb{R}$ e fixado $\varepsilon > 0$, basta tomar $\delta < \|\mathcal{P}'\|$, onde $(\mathcal{P}', T')$ é tal que $\left|L - \sum_{(\mathcal{P},T)} f\right| < \varepsilon$ sempre que $(\mathcal{P}, T) \succeq (\mathcal{P}', T')$. □

Corolário 1.2.25. *Se $f\colon [a,b] \to \mathbb{R}$ é Riemann-integrável e $L, L' \in \mathbb{R}$ satisfazem a Definição 1.2.22, então $L = L'$*

Demonstração. Pelo teorema anterior, L e L' são limites da mesma rede real. Logo, $L = L'$ pela unicidade dos limites de redes reais. □

[†]Por exemplo: com $a := 0$ e $b := 15$, ao considerar $\mathcal{P} := (0, 3, 7, 9, 15)$ e $\mathcal{P}' := (0, 1, 5, 7, 11, 15)$, o procedimento descrito para $\mathcal{P} \sqcup \mathcal{P}'$ resulta em $k = 8$ e $p_0 := 0$, com os demais pontos da partição definidos recursivamente, $p_1 = \min\{1, 3, 5, 7, 9, 11, 15\} = 1$, $p_2 = \min\{3, 5, 7, 11, 15\} = 3$, $p_3 = 5\ldots$ de modo que $\mathcal{P} \sqcup \mathcal{P}' = (0, 1, 3, 5, 7, 9, 11, 15)$. Viu? Não é tão difícil inventar um exemplo para entender uma definição. Faça isso você também!

Definição 1.2.26. Dada uma função Riemann-integrável $f\colon [a,b] \to \mathbb{R}$, a **integral de Riemann de f em** $[a,b]$, denotada por

$$\int_a^b f(t)\,\mathrm{d}t,$$

é o único $L \in \mathbb{R}$ satisfazendo as condições da Definição 1.2.22. ¶

Exemplo 1.2.27. Funções constantes são Riemann-integráveis. Com efeito, se $f(x) = r$ para todo $x \in [a,b]$, então

$$\sum_{(\mathcal{P},T)} f = \sum_{i=1}^{n} f(t_i)(a_i - a_{i-1}) = \sum_{i=1}^{n} r(a_i - a_{i-1}) = r\sum_{i=1}^{n}(a_i - a_{i-1}) = r(b-a)$$

para qualquer partição de Riemann $(\mathcal{P},T)$ do intervalo $[a,b]$, com $\mathcal{P} = (a_0,\dots,a_n)$ e $T = (t_1,\dots,t_n)$, mostrando que a rede é constante e, portanto, convergente. Em particular, $\displaystyle\int_a^b r\,\mathrm{d}t = r(b-a)$. ▲

Exemplo 1.2.28 (Monotonicidade da integral de Riemann)**.** Se $f,g\colon [a,b] \to \mathbb{R}$ são funções Riemann-integráveis com $f \leq g$ e $(\mathcal{P},\mathcal{T})$ é uma partição de Riemann de $[a,b]$, digamos que com $\mathcal{P} := (a_0,\dots,a_n)$ e $T := (t_1,\dots,t_n)$, então

$$\sum_{(\mathcal{P},T)} f := \sum_{i=1}^{n} f(t_i)(a_i - a_{i-1}) \leq \sum_{i=1}^{n} g(t_i)(a_i - a_{i-1}) =: \sum_{(\mathcal{P},T)} g.$$

Dada a arbitrariedade da partição tomada, o Teorema 1.2.12 acarreta a desigualdade $\displaystyle\int_a^b f(t)\,\mathrm{d}t \leq \int_a^b g(t)\,\mathrm{d}t$, como desejado. ▲

Exemplo 1.2.29 (Linearidade da integral de Riemann)**.** Para funções $f,g\colon [a,b] \to \mathbb{R}$ Riemann-integráveis e constantes $\alpha,\beta \in \mathbb{R}$ fixadas, a função $\alpha f + \beta g$ ainda é Riemann-integrável. Com efeito, para uma partição de Riemann $(\mathcal{P},T)$ de $[a,b]$, digamos que $\mathcal{P} := (a_0,\dots,a_n)$ e $T := (t_1,\dots,t_n)$, tem-se

$$\begin{aligned}\sum_{(\mathcal{P},T)} (\alpha f + \beta g) &:= \sum_{i=1}^{n}(\alpha f + \beta g)(t_i)(a_i - a_{i-1}) = \sum_{i=1}^{n}(\alpha f(t_i) + \beta g(t_i))(a_i - a_{i-1}) \\ &= \alpha\sum_{i=1}^{n} f(t_i)(a_i - a_{i-1}) + \beta\sum_{i=1}^{n} g(t_i)(a_i - a_{i-1}) =: \alpha\sum_{(\mathcal{P},T)} f + \beta\sum_{(\mathcal{P},T)} g.\end{aligned}$$

Logo, existe a integral de Riemann de $\alpha f + \beta g$ em $[a,b]$, já que

$$\begin{aligned}\int_a^b \alpha f(t)\,\mathrm{d}t + \int_a^b \beta g(t)\,\mathrm{d}t &= \lim_{\mathrm{Par}_{\mathcal{R}}[a,b]} \sum_{(\mathcal{P},T)} \alpha f + \lim_{\mathrm{Par}_{\mathcal{R}}[a,b]} \sum_{(\mathcal{P},T)} \beta g \\ &= \lim_{\mathrm{Par}_{\mathcal{R}}[a,b]} \left(\alpha\sum_{(\mathcal{P},T)} f + \beta\sum_{(\mathcal{P},T)} g\right) \\ &= \lim_{\mathrm{Par}_{\mathcal{R}}[a,b]} \sum_{(\mathcal{P},T)} (\alpha f + \beta g) =: \int_a^b (\alpha f + \beta g)(t)\,\mathrm{d}t,\end{aligned}$$

como desejado. ▲

Os procedimentos aqui já permitem calcular integrais de forma indolor? Certamente não! Porém, a mera formulação de integrais em termos de redes revela, pelo menos, que as propriedades acima fazem parte do mesmo leque de propriedades de sequências convergentes, limites de funções, etc. É claro que, no momento de aprofundar as discussões sobre integrais, suas particularidades virão à tona — mas o mesmo ocorre com sequências, séries, limites de funções, etc. Para encerrar este primeiro contato, vamos ver que a condição de Riemann-integrabilidade *limita*, literalmente, o escopo das funções que podem ser integradas.

Proposição 1.2.30. *Se* $f\colon [a,b] \to \mathbb{R}$ *é Riemann-integrável, então* f *é limitada.*

Demonstração. Suponha que não. Por f ser Riemann-integrável, existe $L \in \mathbb{R}$ tal que

$$\int_a^b f(t)\,\mathrm{d}t = L,$$

o que assegura uma partição $\mathcal{P}$ do intervalo $[a,b]$ tal que

$$\left| L - \sum_{(\mathcal{Q},T)} f \right| < 1. \tag{1.4}$$

para toda partição de Riemann $(\mathcal{Q},T)$ satisfazendo $\|\mathcal{Q}\| \leq \|\mathcal{P}\|$.

Agora, chamando $\mathcal{P} := (a_0, \ldots, a_n)$, existe $i \in \mathbb{N}$ com $0 < i \leq n$ tal que a restrição de f ao intervalo $[a_{i-1}, a_i]$ é ilimitada: caso contrário, f seria limitada em $[a_0, a_1]$, $[a_1, a_2], \ldots$ e $[a_{n-1}, a_n]$ e, por conseguinte, f seria limitada em $[a,b]$ (lembre-se da frase que iniciou esta demonstração!). Isto permite escolher $t_i, t_i' \in [a_{i-1}, a_i]$ com $|f(t_i) - f(t_i')| \cdot (a_i - a_{i-1}) \geq 2$, uma vez que $\sup\{|f(t)| : t \in [a_{i-1}, a_i]\} = +\infty$ (verifique! $(^{\star}_{\star})$)†. E daí? Ora, tomando $t_j = t_j' \in [a_{j-1}, a_j]$ arbitrariamente para $j \neq i$, obtemos duas tags $T := (t_1, \ldots, t_n)$ e $T' := (t_1', \ldots, t_n')$ tais que

$$\left| \sum_{(\mathcal{P},T)} f - \sum_{(\mathcal{P},T')} f \right| = |f(t_i) - f(t_i')|\,(a_i - a_{i-1}) \geq 2.$$

Porém, como a desigualdade (1.4) vale para $(\mathcal{P},T)$ e $(\mathcal{P},T')$, também vale

$$\left| \sum_{(\mathcal{P},T)} f - \sum_{(\mathcal{P},T')} f \right| \leq \left| L - \sum_{(\mathcal{P},T)} f \right| + \left| L - \sum_{(\mathcal{P},T')} f \right| < 1 + 1 = 2,$$

contrariando a conclusão anterior. □

Exercício 1.44 $(\star)$**.** Mostre que se a rede $\left(\sum_{(\mathcal{P},T)} f\right)_{(\mathcal{P},T)\in \mathrm{Par}_{\mathcal{R}}[a,b]}$ converge em $\mathbb{R}$ com respeito à ordem $\sqsubseteq$, então f é limitada. Dica: imite a demonstração anterior. ■

Exercício 1.45 $(^{\star}_{\star})$**.** Mostre que se $f\colon [a,b] \to \mathbb{R}$ é Riemann-integrável, então vale a desigualdade $\left|\int_a^b f(t)\,\mathrm{d}t\right| \leq \|f\|_\infty (b-a)$, onde $\|f\|_\infty := \sup\{|f(x)| : x \in [a,b]\}$. ■

†Dica/sugestão: note que se $\sup\{|a| : a \in A\} = +\infty$, então $\sup\{|a-b| : a, b \in A\} = +\infty$, afinal de contas, para $M > 0$ e $a \in A$ qualquer, existe $b \in A$ com $M + |a| < |b|$, donde segue que $M < |b| - |a| \leq |a-b|$.

§1 Operações com limites em espaços normados

Muitas das propriedades operatórias vistas para limites reais se estendem de forma muito natural para espaços normados, feitas as devidas ressalvas oriundas das restrições algébricas. A grande sacada é perceber como as desigualdades utilizadas anteriormente se generalizam. Por sorte, a maior parte do trabalho consiste em trocar adequadamente as ocorrências de "$|\cdot|$" por "$\|\cdot\|$".

Lema 1.2.31. *Seja $(E, \|\cdot\|)$ um espaço normado. Para vetores $u, v, x, y \in E$ e escalares $\alpha, \beta \in \mathbb{R}$, valem as seguintes desigualdades:*

(i) $|\|u\| - \|v\|| \leq \|u - v\|$;

(ii) $\|u + v - (x + y)\| \leq \|u - x\| + \|v - y\|$;

(iii) $\|\alpha x - \beta y\| \leq |\alpha - \beta| \cdot (\|x - y\| + \|y\|) + |\beta| \cdot \|x - y\|$.

Exercício 1.46 $({}^{\star}_{\star\star})$**.** Prove as desigualdades anteriores. Dica: observe que para verificar as desigualdades análogas em $\mathbb{R}$, tudo o que você fez foi usar as propriedades fundamentais do valor absoluto (cf. Proposição 0.7.7) que, por sua vez, são justamente as condições que definem normas (cf. Exercício 0.140 e Subseção 1.0.1 §0); evidentemente, esta dica é inútil se você ignorou as desigualdades em $\mathbb{R}$. ■

Corolário 1.2.32. *Sejam $(x_d)_{d\in\mathbb{D}}$ e $(y_d)_{d\in\mathbb{D}}$ redes num espaço normado $(E, \|\cdot\|)$, e $(\lambda_d)_{d\in\mathbb{D}}$ uma rede real, todas sobre um mesmo conjunto dirigido $\mathbb{D}$, tais que $x_d \to x$, $y_d \to y$ e $\lambda_d \to \lambda$ para certos $x, y \in E$ e $\lambda \in \mathbb{R}$.*

(i) *A rede $(z_d)_{d\in\mathbb{D}}$ dada por $z_d := \|x_d\|$ para todo d é tal que $z_d \to \|x\|$, i.e.,*

$$\left\|\lim_{d\in\mathbb{D}} x_d\right\| = \lim_{d\in\mathbb{D}} \|x_d\|.$$

(ii) *A rede $(z_d)_{d\in\mathbb{D}}$ dada $z_d := x_d + y_d$ para todo d é tal que $z_d \to x + y$, i.e.,*

$$\lim_{d\in\mathbb{D}}(x_d + y_d) = \lim_{d\in\mathbb{D}} x_d + \lim_{d\in\mathbb{D}} y_d.$$

(iii) *A rede $(z_d)_{d\in\mathbb{D}}$ dada por $z_d := \lambda_d \cdot x_d$ para todo d é tal que $z_d \to \lambda \cdot x$, i.e.,*

$$\lim_{d\in\mathbb{D}}(\lambda_d \cdot x_d) = \lim_{d\in\mathbb{D}} \lambda_d \cdot \lim_{d\in\mathbb{D}} x_d.$$

Demonstração. Os mesmos argumentos utilizados nas provas do Lema 1.2.0 e do Teorema 1.2.1 se aplicam. Os detalhes ficam por sua conta $({}^{\star}_{\star})$. □

É preciso ter em mente que neste contexto, a multiplicação é feita entre escalares e vetores, e não entre vetores: "multiplicar" vetores em dimensão maior do que 1 costuma ser algo *complexo*. Com isso dito, na situação particular em que se pensa no *produto interno*, a propriedade esperada se mantém.

Exercício 1.47 $({}^{\star}_{\star\star})$**.** Sejam E um $\mathbb{R}$-espaço vetorial, $\langle\cdot,\cdot\rangle\colon E \to \mathbb{R}$ um produto interno (cf. Subseção 1.0.1 §0) e considere sobre E a norma $\|\cdot\|$ induzida pelo produto interno[†].

[†] Aquela que declara $\|x\| := \sqrt{\langle x, x\rangle}$ para todo $x \in E$.

Mostre que se $(x_d)_{d\in\mathbb{D}}$ e $(y_d)_{d\in\mathbb{D}}$ são redes em E tais que $x_d \to x$ e $y_d \to y$, então $\langle x_d, y_d\rangle \to \langle x, y\rangle$, i.e.,

$$\lim_{d\in\mathbb{D}} \langle x_d, y_d\rangle = \left\langle \lim_{d\in\mathbb{D}} x_d, \lim_{d\in\mathbb{D}} y_d \right\rangle.$$

Dica: você já parou para pensar que se escrevêssemos "$u \bullet v$" em vez de "$\langle u, v\rangle$", as propriedades do produto interno seriam parecidíssimas com as propriedades da multiplicação usual? Depois que perceber, não se esqueça da desigualdade de Cauchy-Schwarz. ■

Se a multiplicação entre vetores já é inexistente, com ainda mais razão a *divisão* entre vetores nem chega a fazer sentido. Apesar disso, escrevendo $\frac{x}{\lambda}$ para indicar $\frac{1}{\lambda}x$, resultado da multiplicação entre o <u>escalar</u> $\frac{1}{\lambda}$ e o vetor x, segue que

$$\frac{x_d}{\lambda_d} \to \frac{x}{\lambda}$$

sempre que $\lim_{d\in\mathbb{D}} \lambda_d \neq 0$ em $\mathbb{R}$ e $\lim_{d\in\mathbb{D}} x_d \in E$.

As situações *indeterminadas* são igualmente delicadas, com o agravante dimensional de que os *pontos* no infinito ficam mais complicados e, por isso, não costumam ser considerados em contextos elementares[†]. Porém, alguns resultados são similares. Encerraremos esta seção com um deles.

Proposição 1.2.33. *Sejam $(x_d)_{d\in\mathbb{D}}$ uma rede num espaço normado E e $(\lambda_d)_{d\in\mathbb{D}}$ uma rede em $\mathbb{R}$, ambas definidas sobre o mesmo conjunto dirigido $\mathbb{D}$, tais que $\lim_{d\in\mathbb{D}} x_d := x \in E$ e $\lim_{d\in\mathbb{D}} \lambda_d := \lambda \in [-\infty, +\infty]$.*

(i) Se $\lambda \in \{-\infty, +\infty\}$, então $\lim\limits_{d\in\mathbb{D}} \frac{x_d}{\lambda_d} = 0$.

(ii) Se $x \neq 0$, $\lambda = 0$ e $\lambda_d \neq 0$ para todo d, então $\lim_{d\in\mathbb{D}'} \frac{\|x_d\|}{|\lambda_d|} = +\infty$.

Demonstração. O primeiro item é bem mais simples do que parece: já sabemos que em tais condições, $\lim_{d\in\mathbb{D}} \frac{1}{\lambda_d} = 0$ em $\mathbb{R}$, de modo que por $\left(\frac{1}{\lambda_d}\right)_{d\in\mathbb{D}}$ ser uma rede em $\mathbb{R}$, o resultado segue do terceiro item no último corolário[‡]! O segundo item é um pouco malicioso, mas não tanto. Neste caso, fazendo $y_d := \frac{1}{|\lambda_d|}$ no item (i) da Proposição 1.2.11, resulta que

$$\lim_{d\in\mathbb{D}} \frac{1}{|\lambda_d|} = +\infty,$$

donde o restante é consequência do Teorema 1.2.8. □

§2 Exercícios extras

Exercício 1.48 (Cf. Exercício 1.32 — ($\substack{\star\\\star}$))**.** Seja $(x_n)_{n\in\mathbb{N}}$ uma sequência de vetores num espaço normado $(E, \|\cdot\|)$. Mostre que se a série $\sum x_n$ converge em E, então $x_n \to 0$. ■

Exercício 1.49 ($\substack{\star\\\star}$)**.** Sejam $(x_n)_n$ uma sequência em X e $x \in X$ um ponto fixado. Assumindo que X é normado, mostre que $x_n \to x$ se, e somente se, $\|x_n - x\| \to 0$ em $\mathbb{R}$. Mostre que o resultado permanece válido para redes. ■

[†]Em $\mathbb{R}$ existem apenas duas *direções*, "registradas" por $-\infty$ e $+\infty$. Quantas direções existem em $\mathbb{R}^2$?

[‡]Não custa lembrar que existe $D \in \mathbb{D}$ tal que $\frac{1}{\lambda_d} \neq 0$ para todo $d \succeq D$, e é apenas a partir de tal índice que garantimos a boa definição de $\frac{1}{\lambda_d}$.

Exercício 1.50 ($\star$). Sejam $(x_d)_{d\in\mathbb{D}}$ e $(y_d)_{d\in\mathbb{D}}$ redes num espaço normado $(E, \|\cdot\|)$. Mostre que se $x_d \to x$ e $x_d + y_d \to z$, então $y_d \to z - x$. Dica: finja que são sequências em $\mathbb{R}$ e depois adapte o argumento para o contexto geral. ■

Exercício 1.51 ($\star\star$). Mostre que os vindouros Exercícios 1.91, 1.92 e 1.95 permanecem válidos em espaços normados. ■

Exercício 1.52 ($\star\star$). Sejam $(E, \|\cdot\|)$ um espaço normado e $f\colon [a,b] \to E$ uma função. Defina o que significa dizer que f é Riemann-integrável neste contexto, e prove que funções Riemann-integráveis são limitadas. Dica: note que a definição de uma soma de Riemann $\sum_{(\mathcal{P},T)} f$ *ainda* faz sentido, *mutatis mutandis*. ■

Exercício 1.53 ($\star\star$). Com $E := \mathbb{R}^2$ e $f\colon [a,b] \to E$, sejam $f_0, f_1\colon [a,b] \to \mathbb{R}$ tais que $f(x) = (f_0(x), f_1(x))$ para todo $x \in [a,b]$. Em tais condições, mostre que ao munir E com uma das três normas usuais, tem-se f Riemann-integrável se, e somente se, f_0 e f_1 são Riemann-integráveis. Dica: verifique a identidade

$$\int_a^b f(w)\,\mathrm{d}w = \left(\int_a^b f_0(t)\,\mathrm{d}t, \int_a^b f_1(t)\,\mathrm{d}t\right)$$

por meio da Proposição 1.1.24. ■

Exercício 1.54 (!?). Como você definiria a integral de uma função limitada da forma $[0,1] \times [0,1] \to \mathbb{R}$? ■

1.3 O critério de Cauchy

1.3.0 Essencial

§0 Subsequências (e o Teorema de Bolzano-Weierstrass)

Ao longo das seções anteriores, deparamo-nos diversas vezes com situações em que a partir de uma rede dada (ou sequência, função, etc.), consideramos uma restrição dela a um subconjunto particular a fim de analisar limites. Finalmente, faremos isso de modo preciso, com foco especial no caso das sequências.

Definição 1.3.0. Uma *sequência* $(y_n)_{n\in\mathbb{N}}$ é xingada de **subsequência** de outra sequência $(x_n)_{n\in\mathbb{N}}$ se existe uma função estritamente crescente $h\colon \mathbb{N} \to \mathbb{N}$ tal que $y_n = x_{h(n)}$ para todo $n \in \mathbb{N}$. Geralmente, escreve-se $h(k) := n_k$ para cada $k \in \mathbb{N}$ a fim de denotar a subsequência $(y_n)_{n\in\mathbb{N}}$ como $(x_{n_k})_{k\in\mathbb{N}}$. ¶

$$\begin{tikzcd} \mathbb{N} \arrow[r, "x_\bullet"] & \mathbb{R} \\ \mathbb{N} \arrow[u, "h"] \arrow[ur, "y_\bullet"'] & \end{tikzcd} \qquad \begin{tikzcd} h(n) \arrow[r, mapsto] & x_{h(n)} = y_n \\ n \arrow[u] \arrow[ur] & \end{tikzcd}$$

Figura 1.7: Um resumo diagramático da definição de subsequência. Lembre-se de que h precisa ser estritamente crescente.

Para a sequência $((-1)^n)_{n\in\mathbb{N}}$, por exemplo, ao fazer $n_k := 2k$ para cada $k \in \mathbb{N}$, ganha-se a *subsequência* $\left((-1)^{2k}\right)_{k\in\mathbb{N}}$, que por sua vez é a sequência constante $(1)_{k\in\mathbb{N}}$, já que $(-1)^{2k} = 1$ para todo $k \in \mathbb{N}$. Analogamente, ao fazer $m_k := 2k+1$ para cada $k \in \mathbb{N}$, obtém-se a subsequência $\left((-1)^{2k+1}\right)_{k\in\mathbb{N}}$, que desta vez é a sequência constante $(-1)_{k\in\mathbb{N}}$, já que $(-1)^{2k+1} = -1$ para todo $k \in \mathbb{N}$.

Há um detalhe sutil escondido aqui: a rigor, chamando $\mathcal{N} := \{2k : k \in \mathbb{N}\}$, os objetos $\left((-1)^{2k}\right)_{k\in\mathbb{N}}$ e $((-1)^n)_{n\in\mathcal{N}}$ são distintos! De fato, enquanto o primeiro é uma função da forma $\mathbb{N} \to \mathbb{R}$, o segundo é uma função da forma $\mathcal{N} \to \mathbb{R}$. A inquietante semelhança entre ambas é culpa da bijeção $h\colon \mathbb{N} \to \mathcal{N}$ que faz, justamente, $h(n) := 2n$ para cada n: com a ordem herdada de $\mathbb{N}$, o k-ésimo elemento de $\mathcal{N}$ é, precisamente, $2k$, de modo que o k-ésimo termo da sequência $\left((-1)^{2k}\right)_{k\in\mathbb{N}}$ coincide com o k-ésimo termo da rede $((-1)^n)_{n\in\mathcal{N}}$.

Exercício 1.55 ((!))**.** Pare e reflita sobre o que foi escrito acima. Faça um desenho se for preciso, associando alguns números naturais particulares aos seus dobros. Dica: faça esse tipo de coisa sem que alguém precise pedir para você fazer isso! ■

Parece ser uma distinção irrelevante, certo? Desta vez sim.

Proposição 1.3.1 (Opcional: para quem prefere ter um sono tranquilo)**.** *Sejam $(x_n)_{n\in\mathbb{N}}$ uma sequência real e $\mathcal{N} \subseteq \mathbb{N}$ um subconjunto infinito.*

(i) *Existe uma única função bijetora $h\colon \mathbb{N} \to \mathcal{N}$ estritamente crescente.*

(ii) *A subsequência $\left(x_{h(k)}\right)_{k\in\mathbb{N}}$ converge para $L \in [-\infty, +\infty]$ se, e somente se, a rede $(x_n)_{n\in\mathcal{N}}$ converge para L ao considerarmos $\mathcal{N}$ com a ordem induzida de $\mathbb{N}$.*

Demonstração. Embora isto possa ser feito no braço, já temos tecnologia para proceder mais rapidamente, já que $(\mathcal{N}, \leq)$ é uma boa ordem natural com a ordem $\leq$ herdada de $\mathbb{N}$: para $n \in \mathcal{N}$ qualquer, o menor elemento de $\mathcal{N} \setminus \{m \in \mathbb{N} : m \leq n\} \neq \emptyset$ é o sucessor de n em $\mathcal{N}$ (certo?)*; se $n \in \mathcal{N} \setminus \{\min \mathcal{N}\}$, então existe $k := \max\{m \in \mathcal{N} : m < n\}$ (por quê?)$^\star_\star$, cujo sucessor em $\mathcal{N}$ é n. Logo, pelo Teorema de Dedekind (cf. Teorema 0.3.11), existe um único isomorfismo $h\colon \mathbb{N} \to \mathcal{N}$. Alternativamente, você pode definir $h(0) := \min \mathcal{N}$ e, supondo $h(j)$ definido para todo $j \leq n$, para algum $n \in \mathbb{N}$, fazer

$$h(n+1) := \min\left(\mathcal{N} \setminus \{h(j) : j \leq n\}\right),$$

mas daí será problema seu mostrar que h tem as propriedades desejadas[†].

Vamos ao que interessa, o segundo item. Se a subsequência $\left(x_{h(k)}\right)_{k\in\mathbb{N}}$ converge para L e $I \subseteq [-\infty, +\infty]$ é um intervalo aberto em torno de L, então existe $K \in \mathbb{N}$ tal que $x_{h(k)} \in I$ sempre que $k \geq K$. Para mostrar que $(x_n)_{n\in\mathcal{N}}$ converge para L enquanto rede, basta tomar $N := h(K)$: se $n \in \mathcal{N}$, então existe $k \in \mathbb{N}$ tal que $h(k) = n$, de modo que se $n \geq N$, então $k \geq K$ (pois h é estritamente crescente!) e, por conseguinte, $x_{h(k)} = x_n \in I$, como desejado. A recíproca é análoga (e um bom exercício para praticar ($^\star_\star$)). □

Moral da História. Em algumas situações, fixada uma sequência real $(x_n)_{n\in\mathbb{N}}$, pode ser mais fácil definir uma subsequência de $(x_n)_{n\in\mathbb{N}}$ por meio de alguma função $k \mapsto n_k$ que descreve o subconjunto infinito de $\mathbb{N}$ cujos índices serão considerados. Porém, em outras situações, fazer isso é apenas chato: como descrever, por exemplo, a subsequência $(x_{n_k})_{k\in\mathbb{N}}$ de $(x_n)_{n\in\mathbb{N}}$ em que n_k é o k-ésimo número natural primo? É muito mais razoável pensar em $(x_p)_{p\in\mathcal{P}}$ onde $\mathcal{P} := \{p \in \mathbb{N} : p \text{ é primo}\}$.

[†]Se seguir por este caminho, no final do dia descobrirá que h é o isomorfismo entre $\mathbb{N}$ e $\mathcal{N}$.

Subsequências são dispositivos úteis tanto para detectar divergência quanto para determinar convergência.

Proposição 1.3.2. *Sejam* $(x_n)_{n\in\mathbb{N}}$ *uma sequência real e* $(x_{n_k})_{k\in\mathbb{N}}$ *uma subsequência. Se* $\lim_{n\to\infty} x_n = L$ *para algum* $L \in [-\infty, +\infty]$, *então* $\lim_{k\to\infty} x_{n_k} = L$.

Demonstração. A grande sacada está na "cofinalidade" dos índices n_k em $\mathbb{N}$: o fato de que para todo $M \in \mathbb{N}$ existe $K \in \mathbb{N}$ tal que $n_K \geq M$. Isto ocorre pois a função $h\colon \mathbb{N} \to \mathbb{N}$ que faz $h(k) := n_k$ é estritamente crescente por hipótese, o que assegura $n_0 \geq 0$, $n_1 \geq 1$ (pois $n_1 > n_0 \geq 0$ e...)[†] e, mais geralmente, $n_k \geq k$ para todo $k \in \mathbb{N}$ (certo?)*. Por que isto é útil? Simples: fixado um intervalo aberto I em torno de L, existe $N \in \mathbb{N}$ tal que $x_n \in I$ sempre que $n \geq N$, mas agora basta tomar $K := N$, pois daí $n_K \geq N$, acarretando $n_k \geq N$ para todo $k \geq K$ e, por conseguinte, $x_{n_k} \in I$. □

Corolário 1.3.3 (Importante)**.** *Se uma sequência real* $(x_n)_{n\in\mathbb{N}}$ *tem (pelo menos) duas subsequências que convergem para limites distintos em* $[-\infty, +\infty]$, *então* $(x_n)_{n\in\mathbb{N}}$ *diverge em* $[-\infty, +\infty]$, *i.e., não existe* $L \in [-\infty, +\infty]$ *tal que* $x_n \to L$.

Demonstração. Sejam $(x_{n_k})_{k\in\mathbb{N}}$ e $(x_{m_k})_{k\in\mathbb{N}}$ subsequências de $(x_n)_{n\in\mathbb{N}}$ tais que $x_{n_k} \to L'$ e $x_{m_k} \to L''$, com $L', L'' \in [-\infty, +\infty]$. Se existe $L \in [-\infty, +\infty]$ tal que $x_n \to L$, então a proposição anterior acarreta

$$\lim_{k\to\infty} x_{n_k} = \lim_{k\to\infty} x_{m_k} = L,$$

donde o Exercício 1.17 assegura $L = L'$ e $L = L''$. □

Exemplo 1.3.4. Agora é muito fácil verificar que $((-1)^n)_{n\in\mathbb{N}}$ não converge: a subsequência $((-1)^{2n+1})_{n\in\mathbb{N}}$ converge para -1, enquanto a subsequência $((-1)^{2n})_{n\in\mathbb{N}}$ converge para 1. Acabou. ▲

Exemplo 1.3.5. A sequência $((-2)^n)_{n\in\mathbb{N}}$ não converge em $[-\infty, +\infty]$: com efeito, a subsequência $\left((-2)^{2k}\right)_{k\in\mathbb{N}}$ converge para $+\infty$, já que $2^k \leq 2^{2k} = (-2)^{2k}$ para todo $k \in \mathbb{N}$ e $2^k \to +\infty$; por outro lado, $\left((-2)^{2k+1}\right)_{k\in\mathbb{N}}$ converge para $-\infty$ uma vez que $(-2)^{2k+1} = -(2^{2k+1}) \leq -2^k$ e $-2^k \to -\infty$. Note que em ambos os casos, o Corolário 1.2.20 foi usado sem menção explícita. Acostume-se. ▲

Exemplo 1.3.6 (Série harmônica)**.** Já vimos que se $\sum a_n$ é uma série de números reais que converge em $\mathbb{R}$, então $a_n \to 0$. A vida seria fácil demais se a recíproca fosse verdadeira. A **série harmônica** $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n}$ é um exemplo clássico de que a vida não é fácil[‡].

Suponha conhecida uma sequência $(x_n)_{n\geq 1}$ tal que

(i) $0 \leq x_n \leq \frac{1}{n}$ para todo $n \geq 1$, com $x_n < \frac{1}{n}$ quando n é ímpar, e

(ii) $\frac{1}{n} = x_{2n-1} + x_{2n}$ para cada $n \geq 1$.

[†] Complete o raciocínio ($\star$).

[‡] Na verdade, se você pesquisar por *Análise Não-Arquimediana*, verá que a vida também não é fácil quando vale a recíproca.

Observe então que

$$\sum_{n\geq 1} x_n := \lim_{n\to\infty}\sum_{j=1}^{n} x_j = \lim_{k\to\infty}\sum_{j=1}^{2k} x_j = \lim_{k\to\infty}\sum_{j=1}^{k} x_{2j-1}+x_{2j} = \lim_{k\to\infty}\sum_{j=1}^{k}\frac{1}{j} = \sum_{n\geq 1}\frac{1}{n},$$

onde a segunda igualdade se deve ao fato de $(x_1+\ldots+x_{2k})_{k\in\mathbb{N}}$ ser subsequência da sequência $(x_1+\ldots+x_n)_{n\in\mathbb{N}}$, que converge na reta estendida (*possivelmente* para $+\infty$)[†].

Se ocorresse $\displaystyle\sum_{n\geq 1}\frac{1}{n}\in\mathbb{R}$, então valeria que $\displaystyle\sum_{n\geq 1}\frac{1}{n}-\sum_{n\geq 1}x_n = 0$. Porém, pelas propriedades operatórias dos limites,

$$\sum_{n\geq 1}\frac{1}{n}-\sum_{n\geq 1}x_n = \sum_{n\geq 1}\frac{1}{n}-x_n > 0,$$

já que $\frac{1}{n}-x_n\geq 0$ para todo n, com desigualdade estrita sempre que n é ímpar. Logo, deve-se ter $\displaystyle\sum_{n\geq 1}\frac{1}{n}\notin\mathbb{R}$ e, portanto, $\displaystyle\sum_{n=1}^{\infty}\frac{1}{n}=+\infty$. ▲

Exercício 1.56 $({}^{\star}_{\star})$**.** Defina uma sequência $(x_n)_{n\geq 1}$ com as propriedades requiridas no exemplo anterior. Dica: para $n\geq 1$ par, faça $x_n := \frac{1}{n}$. ■

Um dos modos de *determinar* convergência por meio de subsequências se dá na próxima

Proposição 1.3.7. *Sejam $\mathcal{A}$ e $\mathcal{B}$ subconjuntos infinitos de $\mathbb{N}$ tais que $\mathbb{N}=\mathcal{A}\cup\mathcal{B}$. Se $(x_n)_{n\in\mathbb{N}}$ é uma sequência real tal que $(x_a)_{a\in\mathcal{A}}$ e $(x_b)_{b\in\mathcal{B}}$ convergem para um mesmo limite $L\in[-\infty,+\infty]$, então $\lim_{n\to\infty}x_n = L$.*

Demonstração. Para um intervalo aberto I em torno de L, precisamos encontrar $N\in\mathbb{N}$ tal que $x_n\in I$ sempre que $n\geq N$. Como existem $A\in\mathcal{A}$ e $B\in\mathcal{B}$ tais que $x_a\in I$ sempre que $a\in\mathcal{A}$ satisfaz $a\geq A$ e $x_b\in I$ sempre que $b\in\mathcal{B}$ com $b\geq B$, basta escolher $N\geq A, B$: neste caso, se $n\in\mathbb{N}$ for tal que $n\geq N$, então $n\in\mathcal{A}$ com $n\geq A$ ou $n\in\mathcal{B}$ com $n\geq B$ (pois $\mathbb{N}=\mathcal{A}\cup\mathcal{B}$ por hipótese!), donde segue que $x_n\in I$, como desejado. □

Exemplo 1.3.8. Na hora do aperto, é bastante comum esquecer de usar o Teorema do Confronto para simplificar o cálculo de alguns limites. Por exemplo, um jeito relativamente *blasé* de verificar

$$\lim_{n\to\infty}\left(-\frac{1}{2}\right)^n = 0,$$

consiste em observar que $\left(-\frac{1}{2}\right)^n = (-1)^n\cdot\frac{1}{2^n}$, com $((-1)^n)_{n\in\mathbb{N}}$ limitada e $\frac{1}{2^n}\to 0$, donde a igualdade desejada segue por confronto. No entanto, também seria possível apelar para a proposição anterior, que fornece uma solução mais humilde, mas igualmente satisfatória:

$$\lim_{k\to\infty}\left(-\frac{1}{2}\right)^{2k+1} = \lim_{k\to\infty}\left(-\frac{1}{2}\cdot\frac{1}{2^k}\cdot\frac{1}{2^k}\right) = -\frac{1}{2}\cdot\lim_{k\to\infty}\frac{1}{2^k}\cdot\lim_{k\to\infty}\frac{1}{2^k} = 0,$$
$$\lim_{k\to\infty}\left(-\frac{1}{2}\right)^{2k} = \lim_{k\to\infty}\left(\frac{1}{2^k}\cdot\frac{1}{2^k}\right) = \lim_{k\to\infty}\frac{1}{2^k}\cdot\lim_{k\to\infty}\frac{1}{2^k} = 0.$$

“Ah, mas eu não poderia calcular apenas $\lim_{n\to\infty}\frac{1}{2^n}=0$ *e apelar para o Exercício 1.76?? Por que o mundo é assim???”*... Calma, você acabou de perceber que existem jeitos diferentes de resolver um mesmo problema. Parabéns. ▲

[†] Certo? $({}^{\star}_{\star})$.

Spoiler alert: a semelhança com limites laterais, do Cálculo I, não é mera coincidência (cf. Subseção 1.3.1 §0).

Subsequências ainda são úteis num terceiro aspecto: *consolação*. Numa grande gama de casos, mesmo quando uma sequência *não* converge, é possível obter uma subsequência convergente. Esta ideia aparentemente boba terá consequências avassaladoras ao longo do curso, a começar com o teorema que dá nome a esta subsubseção.

Lema 1.3.9. *Toda sequência em $\mathbb{R}$ admite subsequência monótona.*

Demonstração (adaptada de [30]). Dada uma sequência real $(x_n)_n$, pode-se considerar o conjunto $[\mathbb{N}]^2$ de todos os subconjuntos de $\mathbb{N}$ com precisamente dois elementos, e daí definir a função $c\colon [\mathbb{N}]^2 \to \{\mathsf{A}, \mathsf{V}\}$ que faz $c(\{m,n\}) := \mathsf{A}$ se $m < n$ e $x_m < x_n$, e $c(\{m,n\}) := \mathsf{V}$ se $m < n$ com $x_m \geq x_n$. Vamos com calma...

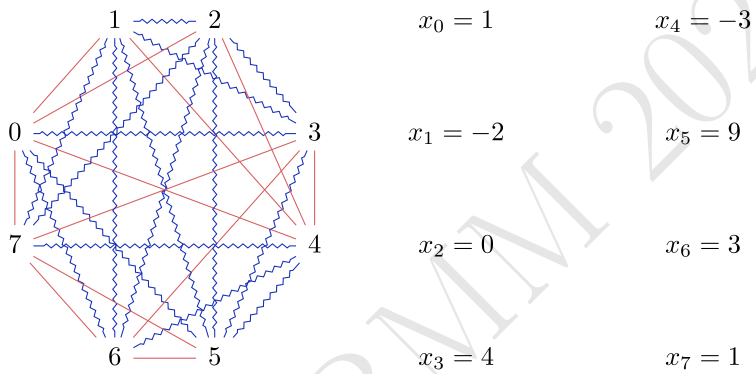


Figura 1.8: Exemplo de um *subgrafo* finito do grafo descrito acima (ondulado = azul).

Intuitivamente, a construção proposta consiste em considerar o *grafo infinito* cujos vértices são <u>todos</u> os números naturais e cujas arestas são todas as possíveis ligações entre eles. Nesse sentido, a função c *pinta* uma "aresta" $m \bullet\!\!-\!\!\!-\!\!\bullet\, n$ em que $m < n$: de Azul se ocorrer $x_m < x_n$; de Vermelho caso contrário[†].

Por que alguém faria isso? *Muito simples*! Um subconjunto infinito $M \subseteq \mathbb{N}$ no qual qualquer aresta ligando seus vértices tenha a mesma cor induz uma subsequência monótona $(x_m)_{m \in M}$: (estritamente) crescente se a cor for Azul; decrescente se a cor for Vermelha. Por exemplo, se $\mathsf{c} = \mathsf{A}$, então $x_m < x_n$ sempre que $m, n \in M$ com $m < n$, ou seja: a *subsequência* $(x_m)_{m \in M}$ é estritamente crescente. O raciocínio é análogo se $\mathsf{c} = \mathsf{V}$.

O passo fundamental na demonstração de que existe um subconjunto $M \subseteq \mathbb{N}$ com a propriedade desejada faz uso (de uma variação) do Princípio da Casa dos Pombos, como expresso no Exercício 0.47: *se X é infinito, $A, B \subseteq X$ são tais que $A \cap B = \emptyset$ e $A \cup B = X$, então A é infinito ou B é infinito.* Em particular, se $P \subseteq [\mathbb{N}]^2$ é infinito, então P é união disjunta dos subconjuntos $\{p \in P : c(p) = \mathsf{A}\}$ e $\{p \in P : c(p) = \mathsf{V}\}$, donde segue que pelo menos um deles deve ser infinito.

$\ulcorner$**Afirmação.** *Fixados um subconjunto infinito $S \subseteq \mathbb{N}$ e um elemento $s \in S$, existem um subconjunto infinito $G_{S,s} \subsetneq S$ e uma cor $\mathsf{c}_{S,s} \in \{\mathsf{A}, \mathsf{V}\}$ tais que $s < \min G_{S,s}$ e $c(\{s,n\}) = \mathsf{c}_{S,s}$ para qualquer $n \in G_{S,s}$.*

[†] Evidentemente, A e V são apenas modos psicologicamente agradáveis de denotar 0 e 1, x e $\{x\}$ ou, mais geralmente, quaisquer dois conjuntos A e V com $A \neq V$.

Demonstração. Em outras palavras, existe um subconjunto infinito de S cujas arestas que ligam seus elementos ao número s têm todas a mesma cor. Para se dar conta disso, note que o subconjunto $P := \{\{s,n\} : n > s \text{ e } n \in S\} \subseteq [\mathbb{N}]^2$ é infinito e, pelo argumento do parágrafo anterior, existe uma cor $\mathsf{C} \in \{\mathsf{A}, \mathsf{V}\}$ tal que $Q := \{p \in P : c(p) = \mathsf{C}\}$ é infinito. Daí, basta tomar $G_{S,s} := (\bigcup Q) \setminus \{s\}$ e $\mathsf{c}_{S,s} := \mathsf{C}$. ⌟

Dito isso, mostraremos que existe uma sequência estritamente crescente $(k_n)_{n\in\mathbb{N}}$ de números naturais tais que, para todo $n \in \mathbb{N}$, existe $\mathsf{c}_n \in \{\mathsf{A}, \mathsf{V}\}$ com $c(\{k_n, k_m\}) = \mathsf{c}_n$ para todo $m > n$. De fato, em vista da argumentação anterior, basta proceder recursivamente:

✓ $S_0 := \mathbb{N}$, $k_0 := \min S_0$ e $\mathsf{c}_0 := \mathsf{c}_{S_0,k_0}$;

✓ $S_1 := G_{S_0,k_0}$, $k_1 := \min S_1$ e $\mathsf{c}_1 := \mathsf{c}_{S_1,k_1}$;

✓ para $n \geq 1$, e supondo definidos $S_0, \ldots, S_n \subseteq \mathbb{N}$, $k_i \in S_i$ e $\mathsf{c}_i \in \{\mathsf{A}, \mathsf{V}\}$ para cada $i < n$, tais que $S_n \subsetneq S_{n-1} \subsetneq \ldots \subsetneq S_0$ com todos infinitos e $c(\{k_i, m\}) = \mathsf{c}_i$ para todo $m \in S_{i+1}$, faz-se $k_n := \min S_n$, $\mathsf{c}_n := \mathsf{c}_{S_n,k_n}$ e $S_{n+1} := G_{S_n,k_n}$, pois assim ainda se verifica S_{n+1} infinito com $S_{n+1} \subsetneq S_n$, $k_n \in S_n$ e $c(\{k_n, m\}) = \mathsf{c}_n$ para todo $m \in S_{n+1}$.

Note que ao proceder desta forma, $c(\{k_n, k_m\}) = \mathsf{c}_n$ sempre que $m > n$ (certo?)$^{\star}_{\star}$. Finalmente, a função $\varphi\colon \mathbb{N} \to \{\mathsf{A}, \mathsf{V}\}$, que faz $\varphi(n) := \mathsf{c}_n$ para cada $n \in \mathbb{N}$, tem imagem finita, donde o Princípio da Casa dos Pombos garante um subconjunto infinito $T \subseteq \mathbb{N}$ e $\mathsf{c} \in \{\mathsf{A}, \mathsf{V}\}$ com $\varphi(t) = \mathsf{c}$ para todo $t \in T$. Isto acarreta $c(\{k_s, k_t\}) = \mathsf{c}$ para quaisquer $s, t \in T$ distintos (certifique-se!)$^{\star}_{\star}$. Logo, basta fazer $M := \{k_t : t \in T\}$. □

Observação 1.3.10 (Opcional: heurística do argumento). Essencialmente, a sequência de subconjuntos $(S_n)_n$ seleciona, a cada passo n, um subconjunto infinito de índices $S_{n+1} \subsetneq S_n$ cujos termos correspondentes são "monótonos" com respeito ao primeiro termo de S_n, mas sem afetar o tipo de monotonicidade (crescente ou decrescente) que seus termos mantêm com o primeiro termo de S_{n-1}. Com base nas ilustrações da Figura 1.9:

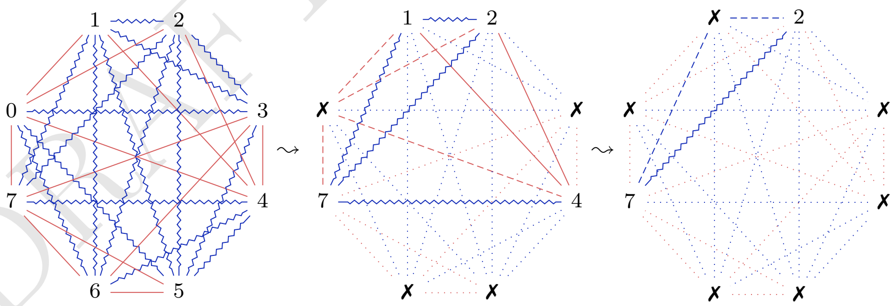


Figura 1.9: Ilustração do procedimento anterior, com $\mathsf{c}_0 := \mathsf{V}$ e $\mathsf{c}_1 := \mathsf{A}$.

(i) com $\mathsf{c}_0 := \mathsf{V}$ e $S_1 := \{1, 2, 4, 7, \ldots\}$, indica-se que o conjunto de vértices S_1 é infinito e tal que todas as arestas ligando os seus vértices a 0 são vermelhas — isto é, $x_0 \geq x_s$ para todo $s \in S_1$; note que no segundo grafo restam apenas os vértices de S_1;

(ii) com $k_1 = 1$ o menor elemento de S_1, a ocorrência de $\mathsf{c}_1 := \mathsf{A}$ com $S_2 := \{2, 7, \ldots\} \subsetneq S_1$ indica que S_2 é um subconjunto infinito de vértices (de S_1!) cujas arestas que ligam seus vértices a 1 são todas azuis — ou seja, $x_1 < x_s$ para todo $s \in S_2$; note que no terceiro grafo restam apenas os vértices de S_2;

(iii) observe que por valer $S_2 \subsetneq S_1$, *ainda* se tem $x_0 \geq x_s$ para todo $s \in S_2$!

Portanto, ao encontrar um subconjunto infinito de $\mathbb{N}$ cujos c_n's coincidem, assegura-se que o tipo de monotonicidade que termos correspondentes na sequência mantêm uns com os outros é o mesmo. $\triangle$

Exercício 1.57 (Opcional — $(\substack{\star\\ \star\star})$)**.** Use o Exercício 0.73 para formalizar o processo recursivo para a definição da sequência de subconjuntos $(S_n)_{n\in\mathbb{N}}$. Dica: aplique o exercício para uma função $\Phi\colon \operatorname{seq}(X) \to X$ adequada, com $X \coloneqq \{S \subseteq \mathbb{N} : S \text{ é infinito}\}$. $\blacksquare$

Teorema 1.3.11 (Bolzano-Weierstrass). *Toda sequência limitada de números reais admite subsequência que converge em $\mathbb{R}$.*

Demonstração. Pelo lema anterior, $(x_n)_n$ tem uma subsequência monótona $(y_n)_n$. Agora, por $(x_n)_n$ ser limitada, segue que $(y_n)_n$ é limitada (certo?!)*. Logo, $(y_n)_n$ tem um limite em $\mathbb{R}$ (pelo Teorema 1.1.4 ou pelo Exercício 1.14). $\square$

Exercício 1.58 $(\substack{\star\\ \star})$**.** Pense rápido: nenhuma sequência ilimitada tem subsequência que converge em $\mathbb{R}$? Dica: não pense rápido. $\blacksquare$

§1 O critério (de completude) de Cauchy

A primeira aplicação do Teorema de Bolzano-Weierstrass será mostrar que em $\mathbb{R}$, a noção de convergência não é afetada por um defeito prático grave presente na definição de limite: o próprio limite. Como assim?

Veja bem: de um ponto de vista meramente linguístico, ao dizer algo do tipo "a sequência $(x_n)_n$ é convergente", é natural entender isso como uma propriedade da sequência. Na prática, isto exigiria que fôsssemos capazes de *decidir* se uma sequência é convergente ou não apenas estudando o comportamento de seus termos. Porém, não é o caso: por definição, a sequência é convergente se ela admite um limite, e até agora isto é tudo o que sabemos[†]... até agora!

Definição 1.3.12. Uma rede $(x_d)_{d\in\mathbb{D}}$ em $\mathbb{R}$ é **de Cauchy**[‡] se para todo $\varepsilon > 0$ existir $D \in \mathbb{D}$ tal que $|x_d - x_{d'}| < \varepsilon$ sempre que $d, d' \succeq D$. Em particular, uma sequência $(x_n)_{n\in\mathbb{N}}$ de números reais é **de Cauchy** se para todo $\varepsilon > 0$ existe $N \in \mathbb{N}$ tal que $|x_m - x_n| < \varepsilon$ sempre que $m, n \geq N$. ¶

Intuitivamente, os termos de uma sequência (ou rede) de Cauchy se tornam *arbitrariamente* próximos uns dos outros[††]. É claro que se os termos de uma sequência (ou rede) se tornam arbitrariamente próximos de um limite, então eles se tornam arbitrariamente próximos entre si. Em outras palavras:

Proposição 1.3.13. *Toda rede real que converge em $\mathbb{R}$ é de Cauchy.*

Demonstração. Seja $(x_d)_{d\in\mathbb{D}}$ uma rede real tal que $x_d \to L$, com $L \in \mathbb{R}$. Para mostrar que $(x_d)_{d\in\mathbb{D}}$ é de Cauchy, devemos fixar $\varepsilon > 0$ e mostrar que existe $D \in \mathbb{D}$ tal que $|x_d - x_{d'}| < \varepsilon$

[†]Em outras palavras: se você encontrar uma sequência real que nunca viu na vida, o único modo de garantir que ela converge é encontrar o seu limite... mas como cada ponto de $[-\infty, +\infty]$ é um candidato em potencial, poderíamos precisar de infinitos testes!

[‡]Lê-se "coxí".

[††]Essa expressão significa que a distância entre os termos fica menor do que qualquer $\varepsilon > 0$ escolhido, desde que se tomem termos suficientemente bons. Escolhido como? *Arbitrariamente*, i.e., conforme o *arbítrio* de quem tem o poder de escolha/decisão: você.

sempre que $d, d' \succeq D$. Como de costume, o truque usual se aplica: para quaisquer $d, d' \in \mathbb{D}$ vale

$$|x_d - x_{d'}| \leq |x_d - L| + |x_{d'} - L|,$$

o que permite escolher $D \in \mathbb{D}$ tal que $|x_d - L| < \frac{\varepsilon}{2}$ sempre que $d \succeq D$ pois, assim, $|x_d - x_{d'}| < \varepsilon$ sempre que $d, d' \succeq D$ (verifique?!)*. □

Em certo sentido, a condição de Cauchy é um *sintoma* de convergência ou, em linguajar clássico, a condição de Cauchy é *necessária* para convergência: se uma sequência (ou rede) converge, então necessariamente ela será de Cauchy. Num mundo ideal, tal condição deveria ser suficiente: toda sequência (ou rede) de Cauchy converge para algum limite! Para variar as frustrações da vida, desta vez veremos que a reta real é um desses mundos ideais.

Lema 1.3.14. *Seja $(x_n)_{n \in \mathbb{N}}$ uma sequência de números reais.*

(i) *Se $(x_n)_n$ é de Cauchy, então $(x_n)_n$ é limitada.*

(ii) *Se $(x_n)_n$ é de Cauchy e admite pelo menos uma subsequência convergente, então $(x_n)_n$ converge para o mesmo limite da subsequência.*

Demonstração. Como a condição de Cauchy pode ser usada para *qualquer* $\varepsilon > 0$, podemos fixar um específico, como $\varepsilon := 1$, e daí obtemos $N \in \mathbb{N}$ tal que $|x_m - x_n| < 1$ sempre que $m, n \geq N$. Como mostrar que $(x_n)_n$ é limitada daí? Ora, note que para $n \geq N$ qualquer, deve-se ter

$$|x_n| = |x_n - x_N + x_N| \leq |x_n - x_N| + |x_N| < 1 + |x_N| := M',$$

mostrando que os termos da sequência ficam limitados por M' para todo $n \geq N$. Restam apenas os termos $x_0, \ldots, x_{N-1}$ mas, para estes, certamente ocorre

$$|x_i| \leq \max\{|x_0|, \ldots, |x_{N-1}|\} := M''$$

para todo $i < N$. Dessa forma, basta tomar $M := \max\{M', M''\}$ para provar o item (i).

Para a segunda afirmação, seja $(x_{n_k})_k$ subsequência de $(x_n)_n$ convergente em $\mathbb{R}$, digamos $x_{n_k} \to L$ com $L \in \mathbb{R}$. Note que para $\varepsilon > 0$ existem

✓ $N \in \mathbb{N}$ tal que $|x_m - x_n| < \varepsilon$ para quaisquer $m, n \geq N$ (pela condição de Cauchy), e

✓ $K \in \mathbb{N}$ tal que $|x_{n_k} - L| < \varepsilon$ para qualquer $k \geq K$ (pois $\lim\limits_{k \to \infty} x_{n_k} = L$).

Como $k \mapsto n_k$ é estritamente crescente, existe $j \in \mathbb{N}$ com $j \geq K$ e $n_j \geq N$, donde segue que

$$|x_n - L| \leq |x_n - x_{n_j}| + |x_{n_j} - L| < 2\varepsilon$$

para todo $n \geq N$, mostrando que $x_n \to L$ (por quê?)*. □

Teorema 1.3.15 (A reta é Cauchy-completa)**.** *Toda sequência de Cauchy converge em $\mathbb{R}$.*

Demonstração. Pelo lema anterior, uma sequência $(x_n)_n$ de Cauchy em $\mathbb{R}$ é limitada. Por Bolzano-Weierstrass, $(x_n)_n$ admite subsequência convergente em $\mathbb{R}$. Novamente pelo lema anterior, resulta que $(x_n)_n$ converge. □

Corolário 1.3.16 (Opcional — mas bem útil)**.** *Toda rede real de Cauchy converge em $\mathbb{R}$.*

Demonstração. Seja $(x_d)_{d\in\mathbb{D}}$ uma rede de Cauchy em $\mathbb{R}$. Para cada $n \in \mathbb{N}$, a condição de Cauchy assegura um $D_n \in \mathbb{D}$ satisfazendo $|x_d - x_{d'}| < \frac{1}{2^n}$ para quaisquer $d, d' \succeq D_n$, e a condição de compatibilidade dos conjuntos dirigidos permite supor $D_{n+1} \succeq D_n$ para todo $n \in \mathbb{N}$ (verifique?)$^{\star}_{\star\star}$. Com isso, resulta que $(x_{D_n})_{n\in\mathbb{N}}$ é uma <u>sequência</u> de Cauchy em $\mathbb{R}$: para $\varepsilon > 0$, basta tomar $N \in \mathbb{N}$ tal que $\frac{1}{2^N} < \varepsilon$ pois daí, para $m, n \geq N$, teremos $D_m, D_n \succeq D_N$ (entendeu?)[†]. Logo, existe $L \in \mathbb{R}$ tal que $x_{D_n} \to L$ (pelo teorema anterior!). Em posse disso, mostraremos que $\lim_d x_d = L$.

Fixado $\varepsilon > 0$, buscamos $D \in \mathbb{D}$ tal que $d \succeq D$ assegure $|x_d - L| < \varepsilon$. Para encontrá-lo, note que

$$|x_d - L| \leq \underbrace{|x_d - x_{D_n}|}_{(A)} + \underbrace{|x_{D_n} - L|}_{(B)},$$

para quaisquer $d \in \mathbb{D}$ e $n \in \mathbb{N}$, em que a parcela (A) pode ser controlada pela condição de Cauchy e a parcela (B) pode ser controlada pois $\lim_{n\to\infty} x_{D_n} = L$. Para controlar (A), observe que já sabemos que $|x_d - x_{D_n}| < \frac{1}{2^n}$ para qualquer $d \succeq D_n$, o que sugere escolher $N \in \mathbb{N}$ tal que $\frac{1}{2^N} < \frac{\varepsilon}{2}$. Para controlar (B), pode-se tomar $N' \in \mathbb{N}$ tal que $|x_{D_n} - L| < \frac{\varepsilon}{2}$ para todo $n \geq N'$. Enfim, com $M := \max\{N, N'\}$, basta fazer $D := D_M$:

- ✓ como $(D_n)_{n\in\mathbb{N}}$ é crescente e $M \geq N$, a ocorrência de $d \succeq D_M$ acarreta $d \succeq D_N$ e, dessa forma, $(A) < \frac{\varepsilon}{2}$;
- ✓ como $M \geq N'$, resulta $|x_{D_M} - L| < \frac{\varepsilon}{2}$;
- ✓ logo, se $d \succeq D_M$, então $|x_d - L| < \varepsilon$, como queríamos. □

E o Kiko? O critério de Cauchy permite assegurar a existência de limites sem conhecê-los. Daí, aliando isto à Proposição 1.3.2, pode-se reduzir o problema de calcular um limite complicado ao processo de determinar o limite de uma *subsequência* mais simples. O Exemplo 1.3.28 traz uma ilustração com integrais.

1.3.1 Extras

§0 Cofinalidade, sub-redes e limites laterais revisitados

Se você curtiu a ideia de redes, deve ter se perguntado: qual seria o análogo de subsequências para redes? Em outras palavras: o que seria uma *sub-rede*? Embora existam respostas não equivalentes para esta pergunta, o presente contexto permite ignorar o panorama geral das sub-redes e focar apenas no caso simples — que felizmente será útil.

Definição 1.3.17. Sejam $(\mathbb{P}, \leq)$ uma ordem parcial e $C \subseteq \mathbb{P}$ um subconjunto. Diz-se que C é **cofinal** em $\mathbb{P}$ se para todo $p \in \mathbb{P}$ existe $c \in C$ com $p \leq c$. ¶

Exemplo 1.3.18 (Subconjuntos cofinais de $\mathbb{N}$)**.** *Um subconjunto $\mathcal{C} \subseteq \mathbb{N}$ é cofinal se, e somente se, $\mathcal{C}$ é infinito.* De fato, se $\mathcal{C}$ é cofinal e $F \subseteq \mathcal{C}$ é um subconjunto finito de $\mathcal{C}$, então existe $n \in \mathcal{C}$ tal que $\max F < 1 + \max F \leq n$, mostrando que $\mathcal{C} \setminus F \neq \emptyset$ para qualquer subconjunto finito $F \subseteq \mathcal{C}$. Em particular, $\mathcal{C}$ deve ser infinito. Reciprocamente, se $\mathcal{C}$ é infinito e $m \in \mathbb{N}$, então $\mathcal{C} \setminus \{n \in \mathcal{C} : n \leq m\}$ é não vazio, ou seja, existe $c \in \mathcal{C}$ com $m \leq c$. Pela arbitrariedade do $m \in \mathbb{N}$ tomado, resulta que $\mathcal{C}$ é cofinal. ▲

[†]Conclua a demonstração de que $(x_{D_n})_{n\in\mathbb{N}}$ é sequência de Cauchy ($^{\star}_{\star}$).

Exemplo 1.3.19 (Subconjuntos cofinais de $\mathbb{R}_p$)**.** Lembre-se de que para $p \in \mathbb{R}$, denota-se por $\mathbb{R}_p$ o conjunto $\mathbb{R}\setminus\{p\}$ com a relação $\preceq$ que declara $a \preceq b$ se, e somente se, $|b-p| \leq |a-p|$ (cf. Exemplo 1.0.4). Quais são os subconjuntos cofinais de $\mathbb{R}_p$?

Proposição 1.3.20. *Nas condições anteriores, $X \subseteq \mathbb{R}_p$ é cofinal se, e somente se, todo intervalo aberto $I \subseteq \mathbb{R}$ em torno de p intercepta X.*

Demonstração. Supondo X cofinal e tomando $I := (\alpha, \beta)$ com $p \in I$, podemos escolher $\alpha', \beta' \in \mathbb{R}$ tais que $\alpha < \alpha' < p < \beta' < \beta$ (certo?)*. Tomando $\gamma \in \mathbb{R}_p$ tal que $\gamma \succeq \alpha', \beta'$, a cofinalidade de X permite escolher $x \in X$ com $\gamma \preceq x$, e o modo como γ foi obtido garante que $x \in I$, pois temos (verifique!)*

$$|x-p| \leq |\gamma - p| \leq \begin{cases} p-\alpha' < p-\alpha \Rightarrow \alpha - p < x-p < p-\alpha \Rightarrow \alpha < x < 2p-\alpha \\ \text{e} \\ \beta' - p < \beta - p \Rightarrow p-\beta < x-p < \beta - p \Rightarrow 2p-\beta < x < \beta \end{cases} \Rightarrow$$
$$\Rightarrow \alpha < x < \beta.$$

Reciprocamente, se X satisfaz a condição e $y \in \mathbb{R}_p$, então para $r := |y-p| > 0$ podemos considerar $J := (p-r, p+r)$, um legítimo intervalo aberto de $\mathbb{R}$ em torno de p. Pela hipótese, existe $x \in X \cap J$, que satisfaz $|x-p| < r$, como desejado. □

Assim, para $p := 0$, digamos, são exemplos de subconjuntos cofinais em $\mathbb{R}_0$:

(i) $(0, b)$ para qualquer $b \in [-\infty, +\infty]$ com $b > 0$;

(ii) $\left\{\frac{1}{2^n} : n \in \mathbb{N}\right\}$;

(iii) $(a, 0)$ para qualquer $a \in [-\infty, +\infty]$ com $a < 0$;

(iv) $\left\{-\frac{1}{2^n} : n \in \mathbb{N}\right\}$;

(v) $(a, b) \setminus \{0\}$ para quaisquer $a, b \in [-\infty, +\infty]$ com $a < 0$ e $b > 0$.

Observe que ao fazer $a := -\infty$ e $b := +\infty$ nos exemplos acima, obtemos $\mathbb{R}_p = (-\infty, 0) \cup (0, +\infty)$, i.e., expressamos o conjunto dirigido $\mathbb{R}_p$ como reunião de dois subconjuntos cofinais. Agora compare com a situação da Proposição 1.3.7. ▲

A definição de sub-rede que adotaremos consiste em restringir redes a subconjuntos cofinais.

Exercício 1.59 ($\star\star$)**.** Sejam $\mathbb{D}$ um conjunto dirigido. Mostre que se $C \subseteq \mathbb{D}$ é cofinal, então a restrição da relação $\preceq$ ao subconjunto C faz de $(C, \preceq)$ um conjunto dirigido. Dica: antes de qualquer outra coisa, perceba que você deve mostrar que para $a, b \in C$, existe $c \in C$ tal que $a, b \preceq c$. ■

Definição 1.3.21. Sejam $\mathbb{D}$ um conjunto dirigido e $(x_d)_{d\in\mathbb{D}}$ uma rede real. Uma **sub-rede** de $(x_d)_{d\in\mathbb{D}}$ é uma rede $(x_d)_{d\in\mathcal{C}}$ onde $\mathcal{C} \subseteq \mathbb{D}$ é um subconjunto cofinal. Explicitamente, trata-se da restrição da função $\rho\colon \mathbb{D} \to \mathbb{R}$ ao subconjunto cofinal $\mathcal{C} \subseteq \mathbb{D}$. ¶

Exemplo 1.3.22. Subsequências *são* sub-redes (cf. Proposição 1.3.1). ▲

Exemplo 1.3.23. Para um ponto $p \in \mathbb{R}$ qualquer, os subconjuntos $\mathbb{R}_{>p} := (p, +\infty)$ e $\mathbb{R}_{<p} := (-\infty, p)$ são ambos cofinais em $\mathbb{R}_p$ (cf. Proposição 1.3.20). Logo, para uma função $f\colon \mathbb{R} \to \mathbb{R}$, as redes $(f(x))_{x\in\mathbb{R}_{<p}}$ e $(f(x))_{x\in\mathbb{R}_{>p}}$ são sub-redes de $(f(x))_{x\in\mathbb{R}_p}$ (cf. Subseção 1.1.1 §0). ▲

Exemplo 1.3.24 (Cf. Subseção 1.2.1 §0)**.** Fixado um intervalo $[a, b]$, considere $\operatorname{Par}_{\mathcal{R}}[a, b]$ o conjunto das partições de Riemann de $[a, b]$ dirigido pela comparação entre normas, i.e., onde $(\mathcal{P}, T) \preceq (\mathcal{P}', T')$ se, e somente se, $\|\mathcal{P}'\| \leq \|\mathcal{P}\|$. Se para cada $n \in \mathbb{N}$ fixarmos uma partição de Riemann $(\mathcal{P}_n, T_n)$ tal que $\|\mathcal{P}_n\| \to 0$, então $\mathcal{C} := \{(\mathcal{P}_n, T_n) : n \in \mathbb{N}\}$ é cofinal em $\operatorname{Par}_{\mathcal{R}}[a, b]$. Com efeito, para qualquer partição $(\mathcal{P}, T) \in \operatorname{Par}_{\mathcal{R}}[a, b]$, existe $N \in \mathbb{N}$ tal que $\|\mathcal{P}_n\| < \|\mathcal{P}\|$ para todo $n \geq N$, donde resulta $(\mathcal{P}, T) \preceq (\mathcal{P}_N, T_N)$.

Analogamente, ao *escolher* uma *tag* $T_{\mathcal{P}}$ para cada partição $\mathcal{P}$ do intervalo $[a, b]$, resulta que a família $\mathcal{D} := \{(\mathcal{P}, T_{\mathcal{P}}) : \mathcal{P} \text{ é partição de } [a, b]\}$ também é cofinal em $\operatorname{Par}_{\mathcal{R}}[a, b]$: para $(\mathcal{P}, T) \in \operatorname{Par}_{\mathcal{R}}[a, b]$, temos $(\mathcal{P}, T) \preceq (\mathcal{P}, T_{\mathcal{P}})$ (por favor!)$^{\star/2}$. ▲

Exercício 1.60 $({}^{\star}_{\star\star})$**.** Com as notações acima, mostre que $\mathcal{D}$ também é cofinal em $(\operatorname{Par}_{\mathcal{R}}[a, b], \sqsubseteq)$. É possível assegurar que $\mathcal{C}$ é cofinal em $(\operatorname{Par}_{\mathcal{R}}[a, b], \sqsubseteq)$? Dica: para a parte difícil, observe que a reunião dos pontos que ocorrem nas partições é enumerável, i.e., $\bigcup_{n\in\mathbb{N}} \operatorname{im}(\mathcal{P}_n)$ é um conjunto enumerável. ■

A propriedade fundamental das subsequências, a saber, a preservação de convergência, permanece válida para sub-redes:

Proposição 1.3.25. *Sejam $(x_d)_{d\in\mathbb{D}}$ uma rede real e $(x_d)_{d\in\mathcal{C}}$ uma sub-rede. Se para algum $L \in [-\infty, +\infty]$ ocorrer $\lim_{d\in\mathbb{D}} x_d = L$, então $\lim_{d\in\mathcal{C}} x_d = L$.*

Demonstração. Dado um intervalo aberto $I \subseteq [-\infty, +\infty]$ em torno de L, existe $D \in \mathbb{D}$ tal que $x_d \in I$ sempre que $d \succeq D$. Como $\mathcal{C}$ é cofinal em $\mathbb{D}$, existe $C \in \mathcal{C}$ com $D \preceq C$. Logo, para todo $c \in \mathcal{C}$ com $c \succeq C$ deve valer $x_c \in I$ já que $c \succeq D$ por transitividade. □

Corolário 1.3.26. *Se $f\colon [a, b] \to \mathbb{R}$ é Riemann-integrável e $((\mathcal{P}_n, T_n))_{n\in\mathbb{N}}$ é uma sequência de partições de Riemann de $[a, b]$ tal que $\|\mathcal{P}_n\| \to 0$, então*

$$\lim_{n\to\infty} \sum_{(\mathcal{P}_n, T_n)} f = \int_a^b f(t)\, \mathrm{d}t.$$

Demonstração. Basta tomar $\mathcal{C}$ como no exemplo anterior e aplicar a proposição acima. □

Exercício 1.61 $(\star)$**.** Seja $f\colon [a, b] \to \mathbb{R}$ uma função Riemann-integrável. Mostre que

$$\int_a^b f(t)\, \mathrm{d}t = \lim_{(\mathcal{P}, T_{\mathcal{P}})\in\mathcal{D}} \sum_{(\mathcal{P}, T_{\mathcal{P}})} f$$

para qualquer escolha de *tags* $T_{\mathcal{P}}$ para as partições $\mathcal{P}$ de $[a, b]$. ■

Exemplo 1.3.27. Como no Corolário 1.3.3, se uma rede real $(x_d)_{d\in\mathbb{D}}$ admite duas sub-redes que convergem para limites distintos, então a rede $(x_d)_{d\in\mathbb{D}}$ não pode convergir. Para um exemplo simples de aplicação, considere a função $f\colon [0, 1] \to \mathbb{R}$ dada por

$$f(x) := \begin{cases} 1, & \text{se } x \in \mathbb{Q} \cap [0, 1], \\ 0, & \text{se } x \in [0, 1] \setminus \mathbb{Q} \end{cases}.$$

Note então que para qualquer partição $\mathcal{P} := (p_0, \ldots, p_n)$ de $[0,1]$, podemos definir *tags* $A_{\mathcal{P}} := (a_1, \ldots, a_n)$ e $B_{\mathcal{P}} := (b_1, \ldots, b_n)$ para $\mathcal{P}$ tais que $a_i \in \mathbb{Q} \cap [0,1]$ e $b_i \in [0,1] \setminus \mathbb{Q}$ para cada $i \in \{1, \ldots, n\}$, de modo que $\sum_{(\mathcal{P}, A_{\mathcal{P}})} f = 1$ e $\sum_{(\mathcal{P}, B_{\mathcal{P}})} f = 0$ (certo?!)$^{\star}_{\star}$. Daí, ao fazer $\mathcal{D} := \{(\mathcal{P}, A_{\mathcal{P}}) : \mathcal{P}$ é partição de $[0,1]\}$ e $\mathcal{D}'' := \{(\mathcal{P}, B_{\mathcal{P}}) : \mathcal{P}$ é partição de $[0,1]\}$, resulta que

$$\lim_{(\mathcal{P}, A_{\mathcal{P}}) \in \mathcal{D}} \sum_{(\mathcal{P}, A_{\mathcal{P}})} f = 1 \quad \text{e} \quad \lim_{(\mathcal{P}, B_{\mathcal{P}}) \in \mathcal{D}'} \sum_{(\mathcal{P}, B_{\mathcal{P}})} f = 0,$$

donde segue que f não pode ser Riemann-integrável (percebeu?)*. ▲

Exemplo 1.3.28. Em breve, provaremos que toda função *contínua* da forma $[a,b] \to \mathbb{R}$ é Riemann-integrável. Assumindo tal resultado conhecido, e supondo que a função $f \colon [0,1] \to \mathbb{R}$ dada por $f(x) := x^2$ é *contínua*, como calcular $\int_0^1 t^2 \,\mathrm{d}t$? Antes de pensar que se trata de um problema bobo, lembre-se de que ainda não temos o *Teorema Fundamental do Cálculo*!

Para contornar a ausência desse poderosíssimo canhão, vamos cozinhar uma sequência $((\mathcal{P}_n, T_n))_n$ de partições apropriadas de $[0,1]$ e calcular o limite da sequência $\left(\sum_{(\mathcal{P}_n, T_n)} f\right)_n$ correspondente. Para $n \in \mathbb{N}$ com $n > 0$, sejam $\mathcal{P}_n := (a_0, \ldots, a_n)$ e $T_n := (t_1, \ldots, t_n)$ definidos por $a_i := t_i := \frac{i}{n}$ para cada $i \leq n$. Como $\|\mathcal{P}_n\| = \frac{1}{n} \to 0$, é lícito prosseguir:

$$\begin{aligned}
\lim_{n \to \infty} \sum_{(\mathcal{P}_n, T_n)} f &= \lim_{n \to \infty} \sum_{i=1}^{n} \left(\frac{i}{n}\right)^2 \frac{1}{n} = \lim_{n \to \infty} \frac{1}{n^3} \sum_{i=1}^{n} i^2 = \lim_{n \to \infty} \frac{1}{n^3} \frac{n(n+1)(2n+1)}{6} \\
&= \lim_{n \to \infty} \frac{2n^2 + 3n + 1}{6n^2} = \lim_{n \to \infty} \frac{2 + \frac{3}{n} + \frac{1}{n^2}}{6} = \frac{2 + \lim_{n \to \infty} \frac{3}{n} + \lim_{n \to \infty} \frac{1}{n^2}}{6} = \frac{2}{6} = \frac{1}{3},
\end{aligned}$$

pois $1^2 + 2^2 + 3^2 + \ldots + n^2 = \frac{n(n+1)(2n+1)}{6}$ para qualquer $n \in \mathbb{N}$ (verifique?!)*. Talvez agora você entenda o "Fundamental" no *Teorema Fundamental do Cálculo*. ▲

Corolário 1.3.29 (da Proposição 1.3.25)**.** *Se $f \colon \mathbb{R} \to \mathbb{R}$ e $p \in \mathbb{R}$ são tais que $\lim_{x \to p} f(x) = L$ para $L \in [-\infty, +\infty]$, então $\lim_{x \to p^-} f(x) = \lim_{x \to p^+} f(x) = L$.*

É nesse sentido que subsequências e limites laterais são semelhantes: todos são exemplos de sub-redes em contextos apropriados. Assim, a Proposição 1.3.7 pode ser encarada como um resultado análogo ao que você aprendeu em Cálculo I acerca de limites laterais: se ambos existem e coincidem, então o limite da função no ponto existe. Mais geralmente, o que se tem é:

Proposição 1.3.30. *Sejam $(x_d)_{d \in \mathbb{D}}$ uma rede real, $L \in [-\infty, +\infty]$ um ponto e $\mathcal{C}, \mathcal{C}' \subseteq \mathbb{D}$ subconjuntos cofinais em $\mathbb{D}$. Se $\mathbb{D} = \mathcal{C} \cup \mathcal{C}'$, então são equivalentes:*

(i) $\lim_{d \in \mathbb{D}} x_d = L$;

(ii) $\lim_{d \in \mathcal{C}} x_d = \lim_{d \in \mathcal{C}'} x_d = L$.

Demonstração. A Proposição 1.3.25 dá conta de (i) $\Rightarrow$ (ii). Para a recíproca, fixado um intervalo aberto $I \subseteq [-\infty, +\infty]$ com $L \in I$, existem $C \in \mathcal{C}$ e $C' \in \mathcal{C}'$ tais que $x_d \in I$ sempre que $d \in \mathcal{C}$ e $d \succeq C$ ou $d \in \mathcal{C}'$ e $d \succeq C'$. Como $\mathbb{D}$ é dirigido, existe $D \in \mathbb{D}$ com $D \succeq C, C'$, de modo que se $d \succeq D$, então por $\mathbb{D} = \mathcal{C} \cup \mathcal{C}'$ resulta que $d \in \mathcal{C}$ e $d \succeq C$ ou $d \in C'$ e $d \succeq c'$: em todo caso, $x_d \in I$. □

Corolário 1.3.31. *Sejam $f\colon \mathbb{R} \to \mathbb{R}$ uma função e $p \in \mathbb{R}$ um ponto. Se*

$$\lim_{x\to p^-} f(x) = \lim_{x\to p^+} f(x) = L$$

para algum $L \in [-\infty, +\infty]$, então $\lim_{x\to p} f(x) = L$.

Demonstração. Já vimos que todos os limites acima são limites de rede:

$$\lim_{x\to p} f(x) = \lim_{x\in\mathbb{R}_p} f(x)$$
$$\lim_{x\to p^-} f(x) = \lim_{x\in\mathbb{R}_{<p}} f(x)$$
$$\lim_{x\to p^+} f(x) = \lim_{x\to\mathbb{R}_{>p}} f(x)$$

Dado que $\mathbb{R}_{<p} := (-\infty, p)$ e $\mathbb{R}_{>p} := (p, +\infty)$ são cofinais em $\mathbb{R}_p = \mathbb{R}_{<p} \cup \mathbb{R}_{>p}$, o resultado segue da proposição anterior. □

Exercício 1.62 (Opcional — ($^{\star}_{\star}$)). Prove o corolário anterior diretamente, i.e., sem aplicar a proposição anterior. Dica: *adapte* a prova da proposição anterior. ■

Não é difícil perceber[†] que na última proposição, se ocorrer $\mathbb{D} = \mathcal{C}_0 \cup \mathcal{C}_1 \cup \ldots \cup \mathcal{C}_n$ para algum $n \in \mathbb{N}$, com $\mathcal{C}_i$ cofinal para todo $i \leq n$, então o limite de $(x_d)_{d\in\mathbb{D}}$ existe e é L se, e somente se, todas as sub-redes $(x_d)_{d\in\mathcal{C}_i}$ também convergem para o mesmo limite L. Este resultado não se estende para o caso em que $\mathbb{D}$ é reunião de *infinitos* subconjuntos cofinais (reflita). Secretamente, é por isso que a noção de limite lateral perde parte do apelo em dimensões maiores...

§1 Espaços métricos completos

Muito do que se discutiu nas seções anteriores pode ser reciclado para o contexto mais geral dos espaços métricos (cf. Definição 1.0.11), mas algum cuidado é necessário pois nem todo espaço métrico é *bom* como a reta real. Os fatos a seguir são generalizações simples do que vimos — e as verificações serão problema seu ($^{\star}_{\star}$).

(i) As noções de subsequência (e, mais geralmente, sub-redes) são idênticas.

(ii) A preservação de convergência por subsequências (e sub-redes) também permanece.

(iii) A condição de Cauchy se expressa de maneira análoga para sequências e redes em espaços métricos: onde antes havia "$|x_m - x_n| < \varepsilon$", substitua por "$d(x_m, x_n) < \varepsilon$".

(iv) Ainda é verdade que redes convergentes em espaços métricos são de Cauchy. Além disso, se uma sequência de Cauchy $(x_n)_n$ tem subsequência que converge para y, então $x_n \to y$.

(v) Num espaço normado, sequências de Cauchy ainda são *limitadas*[‡].

Exercício 1.63 ($^{\star}_{\star}$). Verifique as afirmações acima! ■

[†]Se discordar, encare como exercício ($^{\star}_{\star\star}$).

[‡]Não confunda "ser limitada" com "ter limite": uma sequência $(x_n)_n$ num espaço normado $(E, \|\cdot\|)$ é **limitada** se existe $M \in \mathbb{R}$ com $M > 0$ tal que $\|x_n\| < M$ para todo $n \in \mathbb{N}$. Desta vez a culpa não é minha, mas sim de quem popularizou traduzir "bounded" como "limitado" em vez de "cercado", "acotado", etc.

Os problemas ocorrem nos resultados que envolvem a *existência* de limites...

Exercício 1.64 $(\substack{\star\\\star})$**.** Sejam $X \subseteq \mathbb{R}$ um subconjunto não vazio e $\alpha \in [-\infty, +\infty] \setminus X$. Mostre que se $\alpha = \sup X$, então existe uma sequência $(x_n)_{n\in\mathbb{N}}$ tal que $x_n \in X$ para todo n e $x_n \to \alpha$. Mostre que vale a mesma coisa se $\alpha = \inf X$. ■

Exemplo 1.3.32. O conjunto $\mathbb{Q}$ dos números racionais pode ser encarado como um espaço métrico por meio do valor absoluto (cf. Exercício 1.6). Com isso dito, seja $X \subseteq \mathbb{Q}$ um subconjunto não vazio e limitado tal que $\sup X = \sqrt{2}$ (cf. Exemplo 1.2.16). Pelo exercício anterior, existe $(x_n)_{n\in\mathbb{N}}$ com $x_n \in X$ para todo $n \in \mathbb{N}$ tal que $x_n \to \sqrt{2}$ em $\mathbb{R}$. Como $(x_n)_n$ converge em $\mathbb{R}$, segue que $(x_n)_n$ é de Cauchy em $\mathbb{R}$ e, como as distâncias entre os termos da sequência medidos em $\mathbb{R}$ são iguais às distâncias medidas em $\mathbb{Q}$, resulta que $(x_n)_n$ também é de Cauchy em $\mathbb{Q}$. Mas $(x_n)_n$ não converge em $\mathbb{Q}$! De fato, se existisse $q \in \mathbb{Q}$ tal que $x_n \to q$ em $\mathbb{Q}$, o mesmo argumento das distâncias coincidentes permitiria concluir que $x_n \to q$ em $\mathbb{R}$, e daí $q = \sqrt{2}$, absurdo! Note que este exemplo também mostra que o Teorema de Bolzano-Weierstrass falha em $\mathbb{Q}$.[†] ▲

Definição 1.3.33. Um espaço métrico X é **completo** no *sentido de Cauchy*, assim como sua métrica é dita **completa**, se toda *sequência* de Cauchy em X converge em X. Em particular, se X é espaço vetorial e a métrica completa é induzida por uma norma, diz-se que X é um **espaço de Banach**. ¶

Exercício 1.65 $(\substack{\star\\\star\star})$**.** Adapte o Corolário 1.3.16 para espaços métricos completos. ■

A reta real é o exemplo fundamental de espaço métrico completo, mas está longe de ser o único. Para um exemplo bobo, lembre-se de que qualquer conjunto X pode ser elevado ao patamar de espaço métrico com a métrica discreta (cf. Exemplo 1.1.22), que será completa por um motivo muito simples: suas sequências de Cauchy são as *quase-constantes*, i.e., aquelas que a partir de um certo índice assumem o mesmo valor (cf. Exemplo 1.1.23). Para um exemplo menos bobo, enfrente a próxima proposição, em que $\mathcal{B}(X)$ abrevia $\mathcal{B}(X, \mathbb{R})$, o espaço das funções limitadas da forma $X \to \mathbb{R}$ (cf. Subseção 0.7.1 §0 e o Exercício 0.140).

Atenção. Será cada vez mais importante não confundir f com $f(x)$: a primeira indica *a função como uma totalidade*, enquanto $f(x)$ representa apenas a imagem de um $x \in X$ pela função f. Explicitamente, f é uma coleção de pares ordenados enquanto $f(x)$ é apenas *um número real*. "*Ah, mas quando eu estudava Cálculo I os exercícios eram enunciados assim! Por que as notações mudam toda hora??? Que vida cruel!*". Pois é, já me contaram que eu *também* comia terra quando criança. Vamos seguir com a vida?

Proposição 1.3.34. *O espaço* $(\mathcal{B}(X), \|\cdot\|_\infty)$ *é de Banach para qualquer conjunto* $X \neq \emptyset$.

Demonstração. A ideia é tomar uma sequência de Cauchy $(f_n)_n$ em $\mathcal{B}(X)$ e, por meio da completude de $\mathbb{R}$, obter $f \in \mathcal{B}(X)$ com $f_n \to f$ com respeito à norma $\|\cdot\| := \|\cdot\|_\infty$. Ora, como $|f_m(x) - f_n(x)| \leq \|f_m - f_n\|$ para quaisquer $m, n \in \mathbb{N}$ e $x \in X$, segue que $(f_n(x))_n$ é de Cauchy em $\mathbb{R}$ (certo?)* e, portanto, existe $y \in \mathbb{R}$ com $f_n(x) \to y$. Ao denotar "este" y por $f(x)$, podemos repetir o processo para os demais pontos de X, o que nos dá uma função $f\colon X \to \mathbb{R}$ que faz $z \mapsto f(z)$ para todo z em X. Esta é a candidata natural a limite de $(f_n)_n$. Mostrar que $f \in \mathcal{B}(X)$ será problema seu[‡].

[†]Você precisará adaptar a definição de "ser limitado" para subconjuntos de espaços métricos (confira a definição de *diâmetro* na Proposição 2.0.20).

[‡]Mas confira a demonstração da Proposição 2.1.13 se precisar de ajuda.

Para provar que $f_n \to f$ em $\mathcal{B}(X)$, observe que para $\varepsilon > 0$ existe $N \in \mathbb{N}$ tal que $\|f_m - f_n\| < \varepsilon$ sempre que $m, n \geq N$. Mostraremos que $\|f - f_k\| \leq \varepsilon$ para todo $k \geq N$. Com efeito, neste caso temos

$$|f_m(x) - f_k(x)| \leq \|f_m - f_k\| < \varepsilon$$

para todo $m \geq N$. Como $(f_m(x))_m$ converge para $f(x)$, resulta

$$|f(x) - f_k(x)| = \left|\lim_{m\to\infty} f_m(x) - f_k(x)\right| = \lim_{m\to\infty} |f_m(x) - f_k(x)| \leq \varepsilon,$$

donde a arbitrariedade do ponto x tomado garante $\|f - f_k\| \leq \varepsilon$, como desejado. □

Corolário 1.3.35. *Se $n \in \mathbb{N}$ e $n > 0$, então $(\mathbb{R}^n, \|\cdot\|_\infty)$ é espaço de Banach.*

Demonstração. Tomando $X := \{1, \ldots, n\}$, segue que $\mathbb{R}^n = B(X)$, já que uma n-upla $(a_1, \ldots, a_n) \in \mathbb{R}^n$ nada mais é do que uma função $a\colon X \to \mathbb{R}$, automaticamente limitada por X ser finito. Daí,

$$\|\,(a_1, \ldots, a_n)\,\|_\infty := \max\{|a_1|, \ldots, |a_n|\} = \sup\{|a_x| : x \in X\} = \|a\|_\infty. \qquad \square$$

O corolário anterior permanece válido se trocarmos a norma do máximo em $\mathbb{R}^n$ por qualquer outra norma, mas ainda é cedo para verificar isso.

Observação 1.3.36 (Completude de Dedekind vs. completude de Cauchy)**.** A reta real foi definida como um corpo ordenado e *Dedekind-completo*. Poderíamos, alternativamente, ter definido a noção de sequências de Cauchy em corpos ordenados (por meio do valor absoluto) a fim de introduzir corpos ordenados *Cauchy-completos*: aqueles em que toda sequência de Cauchy converge. Assim, pelo que vimos aqui, a reta real é Cauchy-completa. Porém, pode-se mostrar mais: *um corpo ordenado arquimediano $\mathbb{K}$ é Dedekind-completo se, e somente se, é Cauchy-completo.* A prova da direção que falta não é horrível.

Demonstração (opcional). Seja $S \neq \emptyset$ um subconjunto de $\mathbb{K}$ limitado superiormente, digamos que por $M_0 \in \mathbb{K}$ com $M_0 > 0$. Fixado $x_0 \in S$, define-se $p_0 := \dfrac{1}{2}(x_0 + M_0)$. Daí:

(i) se p_0 for limitante superior de S, faz-se $a_0 := x_0$ e $b_0 := p_0$;

(ii) se não, faz-se $a_0 := p_0$ e $b_0 := M_0$.

Supondo definidos $a_0, \ldots, a_n \in \mathbb{K}$ e $b_0, \ldots, b_n \in \mathbb{K}$ tais que

($\star_i$) para todo $j \leq n$, $b_j - a_j = \dfrac{b_0 - a_0}{2^j}$ e $a_j \leq b_j$,

($\star_{ii}$) para todo $j < n$, $[a_{j+1}, b_{j+1}] \subseteq [a_j, b_j]$, e

($\star_{iii}$) para todo $j \leq n$, b_j limita S superiormente e existe $s \in S$ com $a_j \leq s$,

define-se $p_{n+1} := \dfrac{1}{2}(a_n + b_n)$. Daí:

(i) se p_{n+1} for limitante superior de S, faz-se $a_{n+1} := a_n$ e $b_{n+1} := p_{n+1}$;

(ii) se não, faz-se $a_{n+1} := p_{n+1}$ e $b_{n+1} := b_n$, de modo que as sequências finitas $(a_j)_{j\leq n+1}$ e $(b_j)_{j\leq n+1}$ ainda satisfazem $(\star_i)$, $(\star_{ii})$ e $(\star_{iii})$.

Ao se considerarem as sequências $(a_n)_n$ e $(b_n)_n$ obtidas pelo procedimento recursivo acima, a condição $(\star_{ii})$ garante que a primeira é crescente, enquanto a segunda é decrescente. Por sua vez, a condição $(\star_i)$ assegura que ambas são de Cauchy, donde a hipótese implica que ambas convergem em $\mathbb{K}$, digamos que para $a, b \in \mathbb{K}$, respectivamente. Aplicando-se então, novamente, a condição (i), verifica-se que $a = b$. Por fim, da condição $(\star_{iii})$, segue que $b = \sup S$: como $s \leq b_n$ para quaisquer $s \in S$ e $n \in \mathbb{N}$, resulta $s \leq \lim_n b_n = b$, mostrando que b limita S superiormente; se $x \in \mathbb{K}$ é limitante superior de S, então $a_n \leq x$ para todo $n \in \mathbb{N}$, e daí $a = \lim_n a_n \leq x$, mostrando que $a = b$ é o menor limitante superior de S, o que encerra a prova. □

Para mais detalhes, confira o artigo [15]. △

§2 Exercícios extras

Exercício 1.66 $\binom{\star}{\star}$. Mostre que espaços métricos discretos são completos, i.e., qualquer conjunto X com a métrica discreta é um espaço métrico completo. ■

Exercício 1.67 $\binom{\star}{\star\star}$. Seja $(x_n)_{n\in\mathbb{N}}$ uma sequência num espaço métrico X, com $L \in X$. Mostre que se $x_n \not\to L$, então $(x_n)_{n\in\mathbb{N}}$ admite uma subsequência $(x_{n_k})_{k\in\mathbb{N}}$ tal que nenhuma subsequência de $(x_{n_k})_k$ converge para L. ■

Exercício 1.68 $\binom{\star}{\star}$. Seja $(x_n)_{n\in\mathbb{N}}$ uma sequência de números reais.

a) Mostre que ao declarar $(m, n) \preceq (m', n')$ se, e somente se, $m \leq m'$ e $n \leq n'$, a relação $\preceq$ é uma ordem parcial em $\mathbb{N} \times \mathbb{N}$ que torna $(\mathbb{N} \times \mathbb{N}, \preceq)$ um conjunto dirigido.

b) Mostre que $(x_n)_n$ é de Cauchy se, e somente se, a rede $(x_m - x_n)_{(m,n)\in\mathbb{N}\times\mathbb{N}}$ converge para 0, onde $\mathbb{N} \times \mathbb{N}$ é dirigido pela relação $\preceq$ do item anterior.

c) Mostre que o resultado permanece válido em espaços métricos — com as devidas adaptações. Dica: em vez da rede $(x_m - x_n)_{(m,n)\in\mathbb{N}\times\mathbb{N}}$, considere $(d(x_m, x_n))_{(m,n)\in\mathbb{N}\times\mathbb{N}}$.

d) Mostre que o resultado permanece válido para redes de Cauchy. ■

Exercício 1.69 $\binom{\star}{\star\star}$. Sejam $\mathbb{D} := (\mathbb{R} \setminus \{0\}) \times \mathbb{R}$ e $\rho\colon \mathbb{D} \to \mathbb{R}$ dada por $\rho(x, y) := \dfrac{y}{x}$.

a) Mostre que $\mathbb{D}$ é dirigido pela relação $\preceq$ em que $(x, y) \preceq (x', y')$ se, e somente se, $\|(x', y')\| \leq \|(x, y)\|$, onde $\|\cdot\|$ indica sua norma favorita em $\mathbb{R}^2$.

b) Para $r \in \mathbb{R}$, seja $D_r := \{(x, y) \in \mathbb{D} : \rho(x, y) = r\}$. Mostre que D_r é cofinal em $\mathbb{D}$.

c) Para cada $r \in \mathbb{R}$, mostre que $(\rho(d))_{d\in D_r}$ converge. A rede ρ pode convergir em $\mathbb{R}$?

d) Decida: existe $\displaystyle\lim_{(x,y)\to(0,0)} \frac{y}{x}$? ■

1.4 Alguns critérios de convergência para séries

1.4.0 Essencial

A completude (de Dedekind ou de Cauchy, cf. Observação 1.3.36) é, precisamente, a propriedade que garante a existência de toda sorte de limites que *deveriam existir*. Em particular, ela é usada diariamente para determinar a convergência de séries sem a necessidade de explicitar o valor do limite. Uma ilustração típica se dá na seguinte

Proposição 1.4.0 (Critério de Cauchy para séries)**.** *Uma série* $\sum a_n$ *de números reais converge em* $\mathbb{R}$ *se, e somente se, para todo* $\varepsilon > 0$ *existe* N *tal que* $\left|\sum_{j=m}^{n} a_j\right| < \varepsilon$ *sempre que* $n \geq m \geq N$*.*

Demonstração. A série $\sum a_n$ converge em $\mathbb{R}$ se, e somente se, a sequência $(a_0 + \ldots + a_n)_{n\in\mathbb{N}}$ converge em $\mathbb{R}$. Agora, pelo Teorema 1.3.15, esta sequência converge se, e somente se, é de Cauchy. Para finalizar, *basta perceber* que a condição expressa no enunciado é apenas o critério de Cauchy reescrito para a sequência $(a_0 + \ldots + a_n)_{n\in\mathbb{N}}$. □

Exercício 1.70 ($\star$)**.** Perceba! ■

Ao longo desta seção, veremos outras condições que costumam ser mais simples de verificar.

§0 Convergência absoluta

Definição 1.4.1. Uma série $\sum x_n$ de números reais é **absolutamente convergente** (ou **converge absolutamente**) se $\sum |x_n|$ converge em $\mathbb{R}$. ¶

Assim, por exemplo, $\sum \left(-\frac{1}{2}\right)^n$ converge absolutamente, já que a série formada pelos termos $\left|\frac{1}{2^n}\right|$ já teve sua convergência verificada (cf. Exercício 1.15). O ponto a se destacar é que poderia não ser óbvio determinar se $\sum \left(-\frac{1}{2}\right)^n$ converge no sentido usual, devido a alternância dos sinais. Eis aí uma das maravilhas da completude.

Proposição 1.4.2. *Toda série absolutamente convergente converge.*

Demonstração. Para uma série $\sum x_n$ em $\mathbb{R}$, tem-se

$$\left|\sum_{j=0}^{m} x_j - \sum_{j=0}^{n} x_j\right| \leq \sum_{j=m}^{n} |x_j| \leq \sum_{j=m}^{\infty} |x_j| := S_m \leq \sum_{j=0}^{\infty} |x_j| := M$$

sempre que $m \leq n$. Em particular, se $\sum x_n$ é absolutamente convergente, então $M < +\infty$. Ocorre que em tais condições, ao fazer "$m \to \infty$", resulta que "$S_m \to 0$" (confira o Exercício 1.92). Logo, a série satisfaz a condição da última proposição e, portanto, converge em $\mathbb{R}$. □

Exercício 1.71 ($\substack{\star\\\star}$)**.** Use o Critério de Cauchy para séries aplicado à série $\sum |x_n|$ a fim de demonstrar a proposição anterior sem apelar para o Exercício 1.92. ■

Exercício 1.72 ($\star$)**.** Mostre que se $\sum a_n$ é série absolutamente convergente e $(b_n)_{n\in\mathbb{N}}$ é tal que $b_n \in \{-1, 1\}$ para todo $n \in \mathbb{N}$, então $\sum b_n a_n$ converge em $\mathbb{R}$. ■

Observação 1.4.3. Nem toda série convergente é absolutamente convergente. Um modo relativamente rápido de cozinhar um exemplo faz uso da

Proposição 1.4.4 (Opcional). *Se* $(a_n)_n$ *é uma sequência decrescente de números positivos com* $a_n \to 0$*, então* $\sum(-1)^n a_n$ *converge em* $\mathbb{R}$*.*

Demonstração. Chamando $s_n := \sum_{j \leq n} (-1)^j a_j$, verificam-se sem grandes dificuldades as identidades

$$s_{2n} = s_{2n-2} \underbrace{-a_{2n-1} + a_{2n}}_{\leq 0} \quad \text{e} \quad s_{2n+1} = s_{2n-1} \underbrace{+a_{2n} - a_{2n+1}}_{\geq 0}.$$

Logo, $(s_{2n})_{n \in \mathbb{N}}$ é decrescente, enquanto $(s_{2n+1})_{n \in \mathbb{N}}$ é crescente. Ocorre que uma sequência limita a outra de forma apropriada: como $s_{2n+1} = s_{2n} - a_{2n+1}$ com $a_{2n+1} \geq 0$, segue que $s_{2n+1} \leq s_{2n}$ para todo n, acarretando $s_1 \leq s_{2n}$ e $s_{2n+1} \leq s_0$ para todo $n \in \mathbb{N}$. Consequentemente, existem os limites $\lim_{n\to\infty} s_{2n}$ e $\lim_{n\to\infty} s_{2n+1}$, que coincidem pois

$$\lim_{n\to\infty} s_{2n+1} = \lim_{n\to\infty} s_{2n} - a_{2n+1} = \lim_{n\to\infty} s_{2n} - \lim_{n\to\infty} a_{2n+1} = \lim_{n\to\infty} s_{2n}.$$

O restante segue da Proposição 1.3.7. Os detalhes ficam por sua conta ($\star\atop\star$). □

Para o exemplo prometido: a série

$$\sum_{n \geq 1} \frac{(-1)^n}{n}$$

satisfaz as hipóteses da proposição anterior e, portanto, converge em $\mathbb{R}$; porém, ela não converge absolutamente, já que a série das *normas* de seus termos é a série harmônica! △

Séries convergentes em $\mathbb{R}$ que não convergem absolutamente são chamadas de **condicionalmente convergentes**. Aqui, a escolha de advérbios ("absolutamente" e "condicionalmente") pode ter chamado sua atenção: condicionalmente com respeito a quê? RESPOSTA: a ordenação das parcelas na soma! Confira a Subseção 1.4.1 §1 para mais detalhes (mas não muitos).

§1 Alguns testes clássicos

Em teoria, a primeira proposição desta seção estabelece o melhor critério de convergência para séries, no sentido de que se trata de uma equivalência: nenhuma série convergente é capaz de escapar de seu escopo. No entanto, nem sempre é fácil determinar se uma série satisfaz o critério de Cauchy. E é para facilitar esse aspecto da vida de quem pratica Análise que se investigam outros *testes de convergência*: nem toda série convergente satisfaz todos os testes, mas sempre que uma série satisfaz algum deles, ela converge. Aqui provaremos apenas alguns[†]:

Proposição 1.4.5 (Teste da comparação). *Sejam* $\sum x_n$ *e* $\sum y_n$ *séries de números reais, com* $|x_n| \leq y_n$ *para* n *suficientemente grande. Se* $\sum y_n$ *converge, então* $\sum x_n$ *converge absolutamente.*

[†]Testes mais delicados costumam ser úteis para tratar de *séries de potências* e, por isso, podem esperar até lá. Se você estiver com muita pressa, confira [24, 25]. Se inglês não for um problema, confira [18].

Demonstração. Note que para $m, n \in \mathbb{N}$ com $m \leq n$,

$$\left|\sum_{j=m}^{n} x_j\right| \leq \sum_{j=m}^{n} |x_j| \leq \sum_{j=m}^{n} y_j := A_{m,n}.$$

Daí, como $\sum y_n$ converge em $\mathbb{R}$, para $\varepsilon > 0$ fixado existe $N \in \mathbb{N}$ tal que $A_{m,n} < \varepsilon$ sempre que $n \geq m \geq N$. Consequentemente, $\sum |x_n|$ satisfaz o critério de Cauchy para séries e, portanto, converge (certo?)*. $\square$

Observação 1.4.6. A formulação acima está de acordo com Tao [42], Pugh [32] e, possivelmente, muitas outras obras. Porém, o Teste da Comparação enunciado em [24, 25] é "mais geral": com a mesma tese, Lima pede como hipóteses a convergência (absoluta) de $\sum y_n$ e a existência de uma constante $c > 0$ tal que $|x_n| \leq c y_n$ para n suficientemente grande.

A rigor, trata-se de uma generalização, já que o resultado anterior segue desta versão com $c := 1$. Porém, veja que isto também segue automaticamente da versão que provamos: afinal de contas, se $\sum y_n$ converge em $\mathbb{R}$, então $\sum c y_n = c \sum y_n$ converge em $\mathbb{R}$ (cf. Exercício 1.91) e, se $\sum c x_n$ é absolutamente convergente, então $\frac{1}{c}\sum c x_n = \sum x_n$ é absolutamente convergente. $\triangle$

Exemplo 1.4.7. Ainda não definimos aqui, mas você certamente conhece a função $\operatorname{sen}\colon \mathbb{R} \to \mathbb{R}$ que a cada $x \in \mathbb{R}$ associa $\operatorname{sen}(x)$, a *função seno*. Note que por saber apenas que $|\operatorname{sen}(x)| \leq 1$ para todo $x \in \mathbb{R}$, já *somos capazes de perceber*, de forma indolor, que a série

$$\sum_{n=0}^{\infty} \frac{\operatorname{sen}(x_n)}{2^n}$$

converge absolutamente, para qualquer sequência real $(x_n)_n$ (certo?)*. $\blacktriangle$

Corolário 1.4.8 (Teste de d'Alembert[†])**.** *Se $x_n \neq 0$ para todo $n \in \mathbb{N}$ e existe $c \in \mathbb{R}$ tal que*

$$\left|\frac{x_{n+1}}{x_n}\right| \leq c < 1 \tag{1.5}$$

para n suficientemente grande, então $\sum x_n$ converge absolutamente.

Demonstração. Primeiro, note que com $y_n := c^n$, temos $\sum y_n$ absolutamente convergente (pois $0 < c < 1$). Agora, a sequência $\left(\frac{x_n}{y_n}\right)_{n\in\mathbb{N}}$ é limitada: como existe $N \in \mathbb{N}$ tal que (1.5) se verifica para todo $n \geq N$, resulta que

$$\left|\frac{x_{n+1}}{x_n}\right| \leq c = \frac{y_{n+1}}{y_n}$$

para todo $n \geq N$ ou, equivalentemente,

$$\left|\frac{x_{n+1}}{y_{n+1}}\right| \leq \left|\frac{x_n}{y_n}\right|,$$

[†]Como enunciado em [24, 25]. Trata-se de uma variação do Teste da Razão (*Ratio Test*).

mostrando que a sequência $\left(\left|\frac{x_{k+N}}{y_{k+N}}\right|\right)_{k\in\mathbb{N}}$ é decrescente. Daí não é difícil concluir que $\left(\frac{x_n}{y_n}\right)_{n\in\mathbb{N}}$ é limitada (conclua!)[†]. Por que isso importa? Pelo seguinte: agora, sabemos que existe $B > 0$ tal que $\frac{|x_n|}{|y_n|} \leq B$ para todo $n \in \mathbb{N}$, o que pode ser reescrito como $|x_n| \leq By_n$. Portanto, $\sum x_n$ é absolutamente convergente pelo teste da comparação! □

Exercício 1.73 $(^{\star}_{\star})$**.** Mostre que se $\sum y_n$ converge absolutamente, com $y_n \neq 0$ para todo n, e $\left(\frac{x_n}{y_n}\right)$ é limitada, então $\sum x_n$ é absolutamente convergente. Dica: encare a demonstração anterior até que ela te encare de volta. ■

Exemplo 1.4.9. Fazendo $x_n := \frac{1}{n^2-3n+1}$, será que $\sum x_n$ converge? Resposta: sim. Após verificar que $\sum \frac{1}{n^2}$ converge em $\mathbb{R}$ (cf. Exercício 1.94), basta aplicar o exercício anterior. ▲

Exemplo 1.4.10 (Exponencial escondida)**.** Para qualquer $a \in \mathbb{R}$, a série $\displaystyle\sum_{n=0}^{\infty} \frac{a^n}{n!}$ converge absolutamente: basta notar que

$$\left|\frac{a^{n+1}}{(n+1)!} \cdot \frac{n!}{a^n}\right| = \left|\frac{a}{n+1}\right|,$$

que se torna estritamente menor do que 1 para n suficientemente grande (certo?)*. ▲

Atenção: veja que nos testes acima, a convergência das séries é determinada sem que se expresse o limite delas! Em geral, determinar o valor *exato* de uma série convergente é um problema bem mais delicado e artesanal.

Para mais exemplos de testes de convergência e aplicações, volte duas casas e faça um curso de Cálculo n para $n > 1$. ~~Isto é um curso de Análise Real.~~

1.4.1 Extras

§0 O Teste M de Weierstrass

Embora muito do que se faz na reta real possa ser replicado em espaços normados, há dois aspectos que merecem cuidado: i) espaços normados, em geral, não vêm de fábrica com uma multiplicação entre vetores; ii) espaços normados não costumam ter uma ordem *total* compatível com suas operações. Isto inviabiliza a generalização da maioria dos testes anteriores. Mas não de todos.

Definição 1.4.11. Uma série $\sum x_n$ de vetores num espaço normado $(E, \|\cdot\|)$ é **absolutamente convergente** (ou **converge absolutamente**) se a série $\sum \|x_n\|$ converge na reta real. ¶

Teorema 1.4.12 (Teste M de Weierstrass[‡])**.** *Num espaço de Banach X, toda série absolutamente convergente converge.*

[†]Sugestão: lembre-se do truque utilizado para mostrar que toda sequência de Cauchy é limitada $(^{\star}_{\star})$. Tente realmente se lembrar antes de procurar o trecho exato do texto, que se encontra na página 150.

[‡]"M" de ***Majorantenkriterium***.

~~Déjà vu~~. Demonstração. Para uma série $\sum x_n$ em X, tem-se

$$\left\| \sum_{j=0}^{m} x_j - \sum_{j=0}^{n} x_j \right\| \leq \sum_{j=m}^{n} \|x_j\| \leq \sum_{j=m}^{\infty} \|x_j\| := S_m \leq \sum_{j=0}^{\infty} \|x_j\| := M$$

sempre que $m \leq n$. Em particular, se $\sum x_n$ é absolutamente convergente, então $M < +\infty$. Ocorre que em tais condições, ao fazer "$m \to \infty$" resulta que "$S_m \to 0$" (confira os Exercícios 1.92 e 1.51). Logo, a sequência das somas parciais da série satisfaz o critério de Cauchy e, portanto, a série converge em X. □

Na literatura, o que costuma ser chamado de "Teste M de Weierstrass" é, na verdade, um subcaso do que se apresentou acima: se $(f_n)_n$ é uma *sequência de funções* limitadas da forma $S \to \mathbb{R}$, onde S é um conjunto qualquer, e $(M_n)_n$ é uma sequência de números reais positivos tais que $\|f_n\|_\infty \leq M_n$ para todo n e $\sum M_n < +\infty$, então para cada ponto $x \in S$ a série $\sum f_n(x)$ converge absolutamente e, além disso, a convergência é *uniforme* em S (cf. Definição 2.1.10).

Embora a noção de convergência uniforme vá demorar para ser explicitada no texto (cf. Subseção 2.1.1 §1), ela já está implicitamente presente desde que discutimos o espaço das funções limitadas com a norma do supremo, no seguinte sentido: uma sequência de funções limitadas converge em $(\mathcal{B}(S), \|\cdot\|_\infty)$ para uma função (limitada) se, e somente se, a sequência converge *uniformemente* para a função (cf. Proposição 2.1.13). De qualquer forma, sabendo disso, não é difícil perceber que o enunciado clássico do Teste M de Weierstrass segue do fato de $(\mathcal{B}(S), \|\cdot\|_\infty)$ ser um espaço de Banach (percebeu?)$^{\star}_{\star}$. Futuramente, este será nosso método favorito para definir funções *contínuas* por meio de *séries de funções*.

Proposição 1.4.13 (Opcional — vale a volta)**.** *Se, num espaço normado X, toda série absolutamente convergente converge, então X é de Banach.*

Demonstração. Deve-se mostrar que se $(x_n)_n$ é de Cauchy em X, então $(x_n)_n$ converge em X. Todavia, em *qualquer* espaço normado, para assegurar a convergência de uma sequência de Cauchy basta *encontrar* uma subsequência que converge (cf. Lema 1.3.14 + Exercício 1.63), justamente o que faremos a seguir. Primeiro, tomamos $n_0 \geq 0$ tal que $\|x_m - x_n\| < 1$ sempre que $m, n \geq n_0$. Escolhemos então $n_1 > n_0$ tal que $\|x_m - x_n\| < \frac{1}{2}$ sempre que $m, n \geq n_1$. Procedendo desta forma recursivamente, resulta que

$$\|x_{n_{k+1}} - x_{n_k}\| \leq \frac{1}{2^k}$$

para todo k e, portanto,

$$\sum_{k=0}^{\infty} \|x_{n_{k+1}} - x_{n_k}\| \leq 2,$$

mostrando que a série $\displaystyle\sum_{k=0}^{\infty} x_{n_{k+1}} - x_{n_k}$ é absolutamente convergente. Logo, pela hipótese, existe $x \in X$ tal que

$$x = \sum_{k=0}^{\infty} x_{n_{k+1}} - x_{n_k},$$

donde segue que $x_{n_k} \to x_{n_0} + x$ (cf. Exercícios 1.92 e 1.51). □

§1 Convergência relativa e o surpreendente Teorema de Riemann

Vamos discutir melhor, mas não muito, o sentido das expressões "absolutamente" e "condicionalmente" que acompanharam algumas séries ao longo da última seção. Primeiro, vamos mostrar que a *soma* de uma série absolutamente convergente não se altera se a ordem das parcelas for alterada — e daí a convergência ser *absoluta*, i.e., não relativa à ordem escolhida.

Proposição 1.4.14. *Se $(E, \|\cdot\|)$ é um espaço de Banach e $\sum_{n=0}^{\infty} x_n$ é uma série absolutamente convergente, então*

$$\sum_{n=0}^{\infty} x_n = \sum_{n=0}^{\infty} x_{\varphi(n)}$$

para qualquer bijeção $\varphi\colon \mathbb{N} \to \mathbb{N}$.

Demonstração. Pela discussão anterior, já temos a convergência da série $\sum x_n$ em E. Resta mostrar que $\sum x_{\varphi(n)}$ converge para o mesmo ponto. Ora, para $N \in \mathbb{N}$ qualquer, existe $M \in \mathbb{N}$ com $M \geq N$ tal que $\{0, \ldots, N\} \subseteq \{\varphi(0), \ldots, \varphi(M)\}$ (certo?)$^{\star}_{\star}$. Daí, segue que

$$\left\| \sum_{n=0}^{M} x_{\varphi(n)} - \sum_{n=0}^{\infty} x_n \right\| = \left\| \sum_{n=0}^{M} x_{\varphi(n)} - \sum_{n=0}^{M} x_n - \sum_{k>M} x_n \right\| \leq \sum_{k \geq N} \|x_k\|,$$

e a última converge para 0 conforme $N \to \infty$ (cf. Exercícios 1.92 e 1.51). Os detalhes ficam por sua conta $(^{\star}_{\star})$. □

O resultado acima pode não surpreender tanto, já que as somas com as quais nos acostumamos a viver são *comutativas*, de modo que, em certo sentido, a proposição só faz o trabalho de mostrar que a intuição é preservada pelo formalismo matemático adotado. A situação muda drasticamente com as séries condicionalmente convergentes.

Exercício 1.74 $(^{\star}_{\star\star})$**.** Seja $\sum x_n$ uma série de números reais condicionalmente convergente.

a) Para cada $n \in \mathbb{N}$, considere $x_n^+ := \max\{x_n, 0\}$ e $x_n^- := \min\{x_n, 0\}$. Convença-se de que $\sum x_n = \sum x_n^+ - \sum x_n^-$.

b) Mostre que se $\sum x_n^+$ ou $\sum x_n^-$ convergissem em $\mathbb{R}$, então $\sum x_n$ seria absolutamente convergente. Dica: $|a| = a^+ + a^-$. ■

Exercício 1.75 (Teorema do Rearranjo de Riemann $(^{\star}_{\star} \star ^{\star}_{\star})$)**.** Seja $\sum x_n$ uma série de números reais. Mostre que se $\sum x_n$ é condicionalmente convergente e $r \in \mathbb{R}$, então existe uma bijeção $\varphi\colon \mathbb{N} \to \mathbb{N}$ tal que $\sum x_{\varphi(n)} = r$. Dica: note que o Exercício 1.74 assegura $\sum x_n^+ = +\infty$ e $\sum x_n^- = -\infty$, o que permite selecionar alternadamente parcelas à direita e à esquerda de r. ■

A rigor, uma série satisfazendo a tese da última proposição é dita *incondicionalmente convergente*. Pelo Teorema do Rearranjo de Riemann, resulta que, em $\mathbb{R}$, séries incondicionalmente convergentes são precisamente as séries absolutamente convergentes. Pode-se mostrar que a equivalência permanece válida em espaços de Banach de dimensão finita — mas ela desaparece em dimensão infinita. Contudo, discutir isso foge do escopo deste trabalho.

1.5 Exercícios essenciais

Exercício 1.76 $\left(\substack{\star\\\star}\right)$**.** Mostre que uma sequência $(x_n)_{n\in\mathbb{N}}$ em $\mathbb{R}$ converge para $x \in \mathbb{R}$ se, e somente se, $\lim_{n\to\infty} |x_n - x| = 0$. Dica: primeiro prove para o caso em que $x = 0$. ■

Exercício 1.77 $(\star)$**.** Pense rápido: se $\lim_{n\to\infty} |x_n| = |x|$ para $x \in \mathbb{R}$ e $(x_n)_{n\in\mathbb{N}}$ sequência real, então $\lim_{n\to\infty} x_n = x$? ■

Exercício 1.78 $(\star)$**.** Mostre que se $0 < a < 1$, então $(a^n)_{n\in\mathbb{N}}$ é estritamente decrescente. Dica: Proposição 0.7.5 + indução. ■

Exercício 1.79 $(\star)$**.** Mostre que se $a < -1$, então não existe $L \in [-\infty, +\infty]$ tal que $a^n \to L$. ■

Exercício 1.80 $\left(\substack{\star\\\star}\right)$**.** Mostre que se $(x_n)_{n\in\mathbb{N}}$ é uma sequência real, crescente e que converge em $\mathbb{R}$, então $\sup_{n\in\mathbb{N}} x_n = \lim_{n\to\infty} x_n$. Mostre o análogo para sequências decrescentes. Dica: para $m \in \mathbb{N}$ qualquer, note que $x_m \le \sup_{n\in\mathbb{N}} x_n$ e $x_m \le x_j$ para todo $j \ge m$. ■

Exercício 1.81 $\left(\substack{\star\\\star\star}\right)$**.** Mostre que o resultado anterior permance válido para redes. ■

Exercício 1.82 $(\star)$**.** Mostre que se $x_n \ge 0$ para todo $n \in \mathbb{N}$, então $\sum x_n$ converge em $[0, +\infty]$. O que ocorre se $x_n \le 0$ para todo n? ■

Exercício 1.83 $\left(\substack{\star\\\star}\right)$**.** Sejam $(x_d)_{d\in\mathbb{D}}$ e $(y_d)_{d\in\mathbb{D}}$ redes reais. Mostre que se $x_d \to x$ e $x_d + y_d \to z$, com $x, z \in \mathbb{R}$, então $y_d \to z - x$. Dica:escreva y_d *usando* x_d e $x_d + y_d$. ■

Exercício 1.84 $(\star)$**.** Sejam $(x_n)_{n\in\mathbb{N}}$ e $(y_n)_{n\in\mathbb{N}}$ sequências de números reais. Mostre que se $\lim_{n\to\infty} x_n = x$ e $\lim_{n\to\infty}(x_n + y_n) = z$, com $x, z \in \mathbb{R}$, então $\lim_{n\to\infty} y_n = z - x$. ■

Exercício 1.85 $\left(\substack{\star\\\star}\right)$**.** Seja $(x_n)_{n\in\mathbb{N}}$ uma sequência de números reais estritamente positivos. Mostre que se existe $\lambda \in [0, 1)$ tal que $\frac{x_{n+1}}{x_n} \to \lambda$, então $x_n \to 0$. Dica: para $\mu \in (\lambda, 1)$ fixado, existe $N \in \mathbb{N}$ tal que $\frac{x_{n+1}}{x_n} < \mu$ para todo $n \ge N$, o que permite concluir que a (sub) sequência $(x_n)_{n\ge N}$ é decrescente e limitada inferiormente; para encerrar, imite o truque do Exemplo 1.2.7. ■

Exercício 1.86 $\left(\substack{\star\\\star}\right)$**.** Seja $k \in \mathbb{N} \setminus \{0\}$ fixado.

a) Mostre que $\lim\limits_{n\to\infty} n^k = \lim\limits_{n\to\infty} n! = \lim\limits_{n\to\infty} n^n = +\infty$.

b) Para $a > 1$, mostre que $\lim\limits_{n\to\infty} \frac{n^k}{a^n} = \lim\limits_{n\to\infty} \frac{a^n}{n!} = \lim\limits_{n\to\infty} \frac{n!}{n^n} = 0$. Dica: use o exercício anterior. ■

Exercício 1.87 $\left(\substack{\star\\\star\star}\right)$**.** Sejam $x \in \mathbb{R}$ um número real, $(x_d)_d$ uma rede real e $(r_n)_n$ uma sequência de números reais estritamente positivos com $r_n \to 0$. Mostre que se para todo $N \in \mathbb{N}$ existe $D_n \in \mathbb{D}$ com $|x_d - x| < r_n$ para todo $d \ge D_n$, então $\lim_d x_d = x$. ■

Exercício 1.88 $\left(\substack{\star\\\star\star}\right)$**.** Sejam $(x_d)_{d\in\mathbb{D}}$ e $(y_d)_{d\in\mathbb{D}}$ redes reais, com $y_d \to 0$. Mostre que <u>se</u> a rede $\left(\frac{x_d}{y_d}\right)_{d\in\mathbb{D}}$ converge em $\mathbb{R}$, então $x_d \to 0$. Dica/sugestão: faça primeiro para sequências e depois generalize. ■

Exercício 1.89 $\left(\substack{\star\\\star}\right)$**.** Sejam $(x_n)_{n\in\mathbb{N}}$ e $(y_n)_{n\in\mathbb{N}}$ sequências reais e crescentes tais que $x_n \le y_n$ para todo n. Mostre que se para cada $p \in \mathbb{N}$ fixado existir um número natural $m_p > p$ com $x_{m_p} \ge y_p$, então $\lim_n x_n = \lim_n y_n$. Dica: sua primeira vontade deve ser usar subsequências mas, para isso, defina $n_0 := m_0$ e $n_{k+1} := n_k + m_{k+1}$ para todo $k \in \mathbb{N}$; se não quiser fazer isso, confira o próximo exercício. ■

Exercício 1.90 $\binom{\star}{\star}$**.** Sejam $(x_d)_{d\in\mathbb{D}}$ e $(y_d)_{d\in\mathbb{D}}$ redes reais e crescentes. Mostre que se para cada $a \in \mathbb{D}$ existir $b \succeq a$ tal que $x_b \geq y_a$, então $\lim_d x_d \geq \lim_d y_d$. Dica: antes de qualquer outra coisa, note (!) que $\lim_d x_d = \sup_d x_d$ e $\lim_d y_d = \sup_d y_d$. ■

Exercício 1.91 (Propriedades operatórias de séries $\binom{\star}{\star}$)**.** Para séries $\sum x_n$ e $\sum y_n$ convergentes em $\mathbb{R}$, mostre que $\sum x_n + y_n = \sum x_n + \sum y_n$ e $r \sum x_n = \sum r x_n$ para qualquer $r \in \mathbb{R}$. ■

Exercício 1.92 $\binom{\star}{\star}$**.** Seja $(x_n)_{n\in\mathbb{N}}$ uma sequência em $\mathbb{R}$.

a) (**Séries telescópicas**) Mostre que se $x_n \to 0$, então $\displaystyle\sum_{n=0}^{\infty}(x_n - x_{n+1}) = x_0$. Dica: calcule
$$\sum_{j\leq n} x_j - x_{j+1}$$
explicitamente[†].

b) Fixado $m \in \mathbb{N}$, mostre que $\displaystyle\sum_{n=0}^{\infty} x_n$ converge em $\mathbb{R}$ se, e somente se, $\displaystyle\sum_{n\geq m} x_n$ também converge em $\mathbb{R}$, onde $\displaystyle\sum_{n\geq m} x_n$ indica a série induzida pela sequência $(x_{m+n})_n$. Dica: observe que
$$\lim_{k\to\infty} \sum_{n=0}^{m+k} x_n = \sum_{n=0}^{\infty} x_n,$$
daí use o Exercício 1.83 para encerrar.

c) Mostre que se $\sum x_n$ converge em $\mathbb{R}$, então $\displaystyle\lim_{m\to\infty} \sum_{n\geq m} x_n = 0$. Dica: antes de reclamar, confira o Exercício 1.84. ■

Observação 1.5.0. A série $\displaystyle\sum_{n\geq m} x_n$ acima costuma ser xingada de **rabo da série** (a partir de m), também denotada por $\displaystyle\sum_{n=m}^{\infty} x_n$. △

Exercício 1.93 $\binom{\star}{\star}$**.** Mostre que $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n(n+1)} = 1$. Dica: expresse $\frac{1}{n(n+1)}$ como uma diferença marota. ■

Exercício 1.94 $\binom{\star}{\star}$**.** Mostre que $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^2}$ converge em $\mathbb{R}$. Dica: $\frac{1}{n^2} < \frac{1}{n(n-1)}$. ■

Exercício 1.95 ($\star$)**.** Mostre que se $\sum x_n$ converge em $\mathbb{R}$ e $\sum y_n$ diverge, então $\sum x_n + y_n$ diverge. ■

Exercício 1.96 $\binom{\star}{\star}$**.** Estenda o Teorema 0.8.6 para supremos e ínfimos em $[-\infty, +\infty]$. ■

[†] E não há problema se o resultado do somatório não lhe parecer completamente óbvio — realmente não é *completamente*. O problema, neste caso, é você não fazer, por conta própria, exemplos com valores pequenos de n para perceber o que está acontecendo.

Exercício 1.97 $({}^{\star}_{\star\star})$**.** Sejam $A, B \subseteq \mathbb{R}$ subconjuntos não vazios.

a) Mostre que $\sup A$ e $\inf A$ sempre existem em $[-\infty, +\infty]$.

b) Supondo que $\sup A \coloneqq +\infty$, mostre que $\sup(A+B) = +\infty$.

c) Supondo que $\inf A \coloneqq -\infty$, mostre que $\inf(A+B) = -\infty$.

d) Mostre que se $\inf A \coloneqq -\infty$, então $\inf(rA) = -\infty$ para $r > 0$ e $\sup(rA) = +\infty$ para $r < 0$.

e) Mostre que se $\sup A \coloneqq +\infty$, então $\sup(rA) = +\infty$ para $r > 0$ e $\inf(rA) = -\infty$ para $r < 0$.

f) Poderia ocorrer $\sup A = -\infty$ ou $\inf A = +\infty$? Por quê? ■

Exercício 1.98 $({}^{\star}_{\star})$**.** Resolva o Exercício 1.14 usando apenas o caso crescente e as propriedades operatórias de limites e supremos/ínfimos (na reta estendida). ■

Exercício 1.99 $({}^{\star}_{\star})$**.** Usando o Teorema do Sanduíche, prove que se $(x_n)_{n\in\mathbb{N}}$ é sequência em $\mathbb{R}$ tal que $|x_n| \to 0$, então $x_n \to 0$. Dica: $-|a| \leq a \leq |a|$ para todo $a \in \mathbb{R}$. ■

Exercício 1.100 $({}^{\star}_{\star})$**.** Dado $r \in \mathbb{R}$, mostre que existe uma sequência $(q_n)_{n\in\mathbb{N}}$ de números racionais tal que $\lim_{n\to\infty} q_n = r$. ■

Exercício 1.101 $({}^{\star}_{\star})$**.** Seja $(x_n)_{n\in\mathbb{N}}$ uma sequência de números reais.

a) Mostre que se $(x_n)_n$ é limitada e $a \in \mathbb{R}$ não é limite da sequência, então $(x_n)_n$ admite uma subsequência que converge para algum $b \neq a$. Dica: primeiro, escreva com cuidado o que significa dizer que a é limite de $(x_n)_n$; pronto, agora tudo o que você precisa fazer é negar essa condição com cuidado para cozinhar uma subsequência esperta antes de aplicar o Teorema de Bolzano-Weierstrass[†].

b) Mostre que se $\lim_{n\to\infty} x_n = L$ com $L \in \{-\infty, +\infty\}$, então $\lim_{n\to\infty} |x_n| = +\infty$.

c) Mostre que se $\lim_{n\to\infty} |x_n| = +\infty$, então nenhuma subsequência de $(x_n)_n$ converge em $\mathbb{R}$.

d) Vale a recíproca do item anterior? Isto é: se $(x_n)_n$ não admite subsequências que convergem em $\mathbb{R}$, então $\lim_{n\to\infty} |x_n| = +\infty$? Dica: proceda pela contrapositiva, e mostre que se $|x_n| \not\to +\infty$, então $(x_n)_n$ admite uma subsequência limitada. ■

Exercício 1.102 $({}^{\star}_{\star\star})$**.** Seja $(x_n)_{n\in\mathbb{N}}$ uma sequência em $\mathbb{R}$ e $L \in [-\infty, +\infty]$. Mostre que se $x_n \not\to L$, então $(x_n)_{n\in\mathbb{N}}$ admite uma subsequência $(x_{n_k})_{k\in\mathbb{N}}$ tal que nenhuma subsequência de $(x_{n_k})_k$ converge para L. ■

Exercício 1.103 $({}^{\star}_{\star})$**.** Refaça a prova do Lema 0.10.9 com suas novas habilidades sobre séries. ■

[†]**Atenção:** dizer que a não é limite de $(x_n)_n$ NÃO É DIZER QUE $\lim_{n\to\infty} x_n \neq a$, pois a sequência $(x_n)_n$ poderia simplesmente não convergir.

1.6 Continuidade

Intuitivamente, uma função contínua, digamos que da forma $[a, b] \to \mathbb{R}$, é aquela cujo gráfico pode ser desenhado "sem tirar o lápis do papel". É claro que isso não é uma *definição matemática* no sentido atual do termo mas, enquanto motivação, essa ideia irá nos ajudar a elaborar/encontrar definiç*ões* apropriadas de continuidade. Por exemplo: como *detectar onde* o lápis foi tirado do papel?

Mais precisamente, se o lápis for tirado do papel no *ponto* $(p, f(p))$ de $\mathbb{R}^2$, então deve ter ocorrido um *salto* no gráfico[†]: existe um intervalo aberto $I \subseteq \mathbb{R}$ em torno de $f(p)$ tal que para qualquer intervalo aberto $J \subseteq [a, b]$ em torno de p é possível encontrar $x \in J$ tal que $f(x) \notin I$.

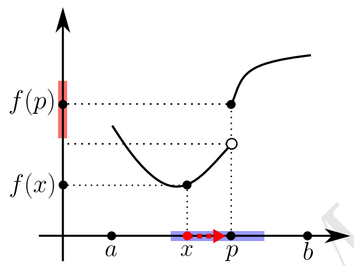


Figura 1.10: Para o intervalo aberto I (em vermelho) em torno de $f(p)$, em qualquer intervalo aberto J (em azul) em torno de p existe x tal que $f(x) \notin I$. Note que se os J's diminuem, os x's convergem para p.

Acima, o intervalo aberto I é *uma* testemunha do "salto", isto é, uma evidência de que o lápis foi tirado do papel no ponto $(p, f(p))$. Note que para servir como uma testemunha efetiva de ~~descontinuidade~~ salto, I realmente precisa encontrar pelo menos um ponto x em qualquer intervalo aberto J em torno de p satisfazendo a condição $f(x) \notin I$.

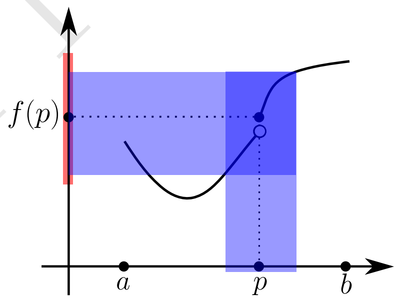


Figura 1.11: Resposta (para a pergunta na Observação 1.6.0): sim.

Observação 1.6.0 (Cuidado com os quantificadores!)**.** Se para *algum* intervalo aberto I em torno de $f(p)$ existir *pelo menos um* intervalo aberto J em torno de p tal que $f(x) \in I$ para todo $x \in J$, ainda seria *possível* que f apresentasse um salto em p?

Na Figura 1.11, qualquer ponto x no intervalo aberto J em torno de p (em azul) tem sua imagem pertencente ao intervalo aberto I em torno de $f(p)$ (em vermelho). No entanto, se diminuíssemos um pouco mais o intervalo vermelho, chegaríamos novamente ao cenário descrito pela Figura 1.10. △

[†] A menos que a pessoa volte a desenhar no mesmo ponto em que parou... Como este caso não é útil para a ilustração, podemos supor que a pessoa esteja trêmula.

Em outras palavras, cada intervalo aberto I em torno de $f(p)$ funciona como uma lupa que utilizamos para checar a existência de saltos. Se a lupa for incapaz de determinar um salto (como na Figura 1.11), melhoramos a lupa, tomando um intervalo aberto I' menor, para enxergar ainda *mais perto*. Se for impossível encontrar uma lupa/intervalo aberto capaz de detectar um salto, a conclusão é uma só: não há salto!

1.6.0 Essencial

§0 Continuidade via convergência

A discussão anterior sugere definir continuidade como ausência de saltos. Porém, em vez de fazer isso apenas para funções da forma $[a,b] \to \mathbb{R}$, vamos considerar funções da forma $X \to \mathbb{R}$, para um subconjunto qualquer $X \subseteq \mathbb{R}$.

Definição 1.6.1 (Continuidade via intervalos, para funções reais)**.** Sejam $X \subseteq \mathbb{R}$ um subconjunto, $p \in X$ um ponto e $f\colon X \to \mathbb{R}$ uma função. Diremos que $f\colon X \to \mathbb{R}$ é **contínua em** p se para todo intervalo aberto $I \subseteq \mathbb{R}$ com $f(p) \in I$ existir um intervalo aberto $J \subseteq \mathbb{R}$ tal que $p \in J$ e $f(x) \in I$ para todo $x \in J \cap X$. Uma função **contínua** é aquela que é contínua em todos os pontos de seu domínio. ¶

Exercício 1.104 (Definição clássica — $(^{\star}_{\star})$)**.** Mostre que $f\colon X \to \mathbb{R}$ é contínua em $p \in X$ se, e somente se, para todo $\varepsilon > 0$ existe $\delta > 0$ tal que $|f(x) - f(p)| < \varepsilon$ sempre que $x \in X$ satisfaz $|x - p| < \delta$. Dica: lembre-se de que para um intervalo aberto $O \subseteq \mathbb{R}$ tal que $z \in O$, existe $r > 0$ tal que $(z - r, z + r) \subseteq O$ (cf. Observação 1.0.6). ■

Num primeiro momento, trata-se da exata negação do que entendemos como a definição de descontinuidade, o que parece fazer desta uma formulação ótima, já que é bastante intuitiva. Porém, ela é pouco prática:

Exemplo 1.6.2 (Ilustrativo, quase nunca será assim)**.** Suponha que se queira provar que a função $f\colon [0,1] \to \mathbb{R}$ dada por $f(t) := t^2$ é contínua em todo ponto do intervalo $[0,1]$. Para isso, fixado $x \in [0,1]$ e um intervalo aberto I em torno de x^2, devemos encontrar um intervalo aberto J em torno de x tal que $y^2 \in I$ sempre que $y \in J \cap [0,1]$. Tomando $I := (x^2 - \varepsilon, x^2 + \varepsilon)$ para algum $\varepsilon > 0$, isto significa encontrar $\delta > 0$ tal que $|x^2 - y^2| < \varepsilon$ sempre que $y \in [0,1] \cap (x - \delta, x + \delta)$. Percebeu que será massante, né? ▲

Exercício 1.105 $(^{\star}_{\star\star})$**.** Encontre um $\delta > 0$ que funciona. Dica: $|x^2 - y^2| = |x - y||x + y|$ e $|x + y| \leq |x| + |y| \leq |x| + |y - x| + |x| = |x - y| + 2|x|$. ■

O exemplo acima, embora chato, pode ajudar as pessoas que não leem as legendas das figuras a perceber que (des) continuidade tem a ver com convergência, o que pode trazer uma definição de continuidade bem mais prática: na Figura 1.10, o intervalo I que testemunha a descontinuidade de f em p é tal que para todo intervalo aberto J em torno de p existe $x \in J$ tal que $f(x) \notin I$. Em particular, ao escolher $x_J \in J$ tal que $f(x_J) \notin I$, podemos fazer J variar na coleção $\mathcal{I}_p$ de *todos os intervalos abertos em torno de* p.

Ora, tal coleção $\mathcal{I}_p$ pode ser encarada como um conjunto dirigido: basta declarar $J \preceq J'$ se, e somente se, $J' \subseteq J$ (verifique que $(\mathcal{I}_p, \preceq)$ é um conjunto dirigido!)[†]. Desse modo, $(x_J)_{J \in \mathcal{I}_p}$ se revela uma rede legítima, tal que $x_J \to p$ mas $f(x_J) \not\to f(p)$. Esta é a contrapositiva da ponta de um *iceberg*.

[†]Dica: a interseção de intervalos abertos em $\mathbb{R}$ é um intervalo aberto de $\mathbb{R}$ $(^{\star}_{\star})$.

Teorema 1.6.3. *Sejam $X \subseteq \mathbb{R}$ um subconjunto e $f\colon X \to \mathbb{R}$ uma função. Para qualquer $p \in X$, são equivalentes:*

(i) *f é contínua em p;*

(ii) *$\lim_{d\in\mathbb{D}} f(x_d) = f(p)$ para toda rede real $(x_d)_{d\in\mathbb{D}}$ em X tal que $x_d \to p$;*

(iii) *$\lim_{n\to\infty} f(x_n) = f(p)$ para toda sequência real $(x_n)_{n\in\mathbb{N}}$ em X tal que $x_n \to p$.*

Demonstração. Para (i) $\Rightarrow$ (ii), dado um intervalo aberto $I \subseteq \mathbb{R}$ em torno de $f(p)$, existe um intervalo aberto $J \subseteq \mathbb{R}$ tal que $p \in J$ e tal que $f(x) \in I$ para todo $x \in X \cap J$. Ora, por hipótese, $x_d \in X$ para todo $d \in \mathbb{D}$ e, como $x_d \to p$, existe $D \in \mathbb{D}$ tal que $d \succeq D$ acarreta $x_d \in X \cap J$ e, portanto, $f(x_d) \in I$. Em outras palavras, mostramos que $f(x_d) \to f(p)$.

A implicação (ii) $\Rightarrow$ (iii) segue pois toda sequência é uma rede. Enfim, (iii) $\Rightarrow$ (i) será ~~um *déjà vu*~~ pela contrapositiva. Se f é descontínua em p e $I \subseteq \mathbb{R}$ é um intervalo aberto em torno de $f(p)$ que testemunha isso, então para todo $n \in \mathbb{N}$ existe $x_n \in \left(p - \frac{1}{2^n}, p + \frac{1}{2^n}\right) \cap X$ tal que $f(x_n) \notin I$, o que resulta numa sequência $(x_n)_{n\in\mathbb{N}}$ tal que $x_n \to p$ mas $f(x_n) \not\to f(p)$ (certo?)$^{\star}_{\star}$. $\square$

USO IMPLÍCITO DO AXIOMA DA ESCOLHA. Como justificar as escolhas dos pontos $x_n \in \left(p - \frac{1}{2^n}, p + \frac{1}{2^n}\right) \cap X$ para cada $n \in \mathbb{N}$? Certamente, como $\left(p - \frac{1}{2^n}, p + \frac{1}{2^n}\right) \cap X \neq \emptyset$ para todo n, algum ponto *pode* ser escolhido. Porém, a rigor, isto pode ser feito apenas individualmente: a fim de *extrapolar* para todo $n \in \mathbb{N}$, precisamos fazer escolhas infinitas vezes, o que não condiz com o ideal de demonstrações como "algoritmos" finitários. Formalmente, o que se faz é assumir que isto pode ser feito: entrou em cena o Axioma da Escolha, discutido na Subseção 0.4.1 §2. "*Ah, eu prefiro o jeito ~~do Elon~~ canônico, ele não usa essas coisas filosóficas de conjuntos!!!*". Você tem o direito de ignorar as justificativas formais — mas aqui, pelo menos, a *escolha* é sua e não minha.

Entre outras coisas, o último teorema também mostra que sequências *bastam* para discutir continuidade na reta real[†]. No entanto, matrizes também bastam para estudar Álgebra Linear em dimensão finita, e nem por isso descartamos as transformações lineares, não é mesmo? Embora sequências sejam *suficientes*, elas não têm a mesma versatilidade que as redes apresentam para descrever as diversas noções de convergência que surgem diariamente na Análise.

§1 Exemplos iniciais

Corolário 1.6.4 (Cf. Exercício 1.35, na Subseção 1.2.0 §0)**.** *Toda função polinomial é contínua.*

Demonstração. Seja $f\colon X \to \mathbb{R}$ dada por $f(t) := a_0 + a_1 t + \ldots + a_m t^m$ para certos $a_0, \ldots, a_m \in \mathbb{R}$. Se $(x_n)_{n\in\mathbb{N}}$ é sequência em X tal que $x_n \to x$ para algum $x \in X$, então

$$\lim_{n\to\infty} f(x_n) = a_0 + a_1 \lim_{n\to\infty} x_n + \ldots + a_m \left(\lim_{n\to\infty} x_n\right)^m = a_0 + a_1 x + \ldots + a_m x^m = f(x). \quad \square$$

Observação 1.6.5 (Continuidade de restrições)**.** Acima, não haveria perda de generalidade em assumir $X = \mathbb{R}$. De fato, em geral, se $X \subseteq Y \subseteq \mathbb{R}$ e $f\colon Y \to \mathbb{R}$ é contínua, então a restrição $f|_X\colon X \to \mathbb{R}$ ainda é contínua (verifique!)*. $\triangle$

[†] E, mais geralmente, em espaços métricos.

Exemplo 1.6.6. A função $\iota\colon \mathbb{R}\setminus\{0\}\to\mathbb{R}$ dada por $\iota(x) := \frac{1}{x}$ é contínua (em todos os pontos do seu domínio $\mathbb{R}\setminus\{0\}$) em virtude do item (iii) no Teorema 1.2.1. Neste caso, não é possível *estender* ι a uma função *contínua* da forma $g\colon\mathbb{R}\to\mathbb{R}$. Com efeito, se $g\colon\mathbb{R}\to\mathbb{R}$ é contínua em todos os pontos, então $\lim_{n\to\infty} g(x_n) = g(0)$ para qualquer sequência $(x_n)_{n\in\mathbb{N}}$ em $\mathbb{R}$ tal que $x_n\to 0$. Agora, se $g(x) = \frac{1}{x}$ para todo $x\neq 0$, então ao fazer $x_n := \frac{1}{2^n}$ para todo $n\in\mathbb{N}$ teríamos $x_n\to 0$ mas $\lim_{n\to\infty} g(x_n) = +\infty$. ▲

Observação 1.6.7 (Opcional: nem todo "erro" está errado)**.** Apesar do que se destacou acima, seria possível estender ι a uma função *contínua* da forma $[-\infty,0)\cup(0,+\infty]\to\mathbb{R}$: neste caso, declarando $\Gamma(+\infty) := \Gamma(-\infty) := 0$ e $\Gamma(x) := \frac{1}{x}$ para todo $x\in\mathbb{R}\setminus\{0\}$, resulta que Γ é *contínua*, no sentido de que se $(x_n)_{n\in\mathbb{N}}$ é sequência em $[-\infty,+\infty]\setminus\{0\}$ tal que $x_n\to x$ para $x\in[-\infty,+\infty]\setminus\{0\}$, então $\Gamma(x_n)\to\Gamma(x)$ (cf. Proposição 1.2.10). Se $\Phi\colon[-\infty,+\infty]\setminus\{0\}\to\mathbb{R}$ fosse outra função com as mesmas propriedades, resultaria que $\Gamma=\Phi$ (certo?)$^{\star}_{\star}$. Logo, Γ fica completamente determinada por ι, o que justifica escrever

$$\Gamma(x) := \frac{1}{x}$$

para todo $x\in[-\infty,+\infty]\setminus\{0\}$. Assim, você não precisa sentir peso na consciência ao escrever "$\frac{1}{+\infty} = 0$" ou "$\frac{1}{-\infty} = 0$", embora seja bom evitar, já que as pessoas não estão prontas para conversar sobre a reta estendida. △

Corolário 1.6.8. *A função <u>valor absoluto</u>[†] $|\cdot|\colon\mathbb{R}\to\mathbb{R}$ é contínua.*

Demonstração. Automático em vista do teorema anterior e do Lema 1.2.0. □

Corolário 1.6.9. *A função $\sqrt{(\cdot)}\colon[0,+\infty)\to\mathbb{R}$ é contínua.*

Demonstração. Vamos considerar primeiro o caso em que $x\neq 0$. Se $(x_n)_{n\in\mathbb{N}}$ é sequência em $[0,+\infty)$ tal que $x_n\to x$, então podemos assumir $\sqrt{x_n}+\sqrt{x}\neq 0$ para todo $n\in\mathbb{N}$ (por quê?)*, e daí

$$\lim_{n\to\infty}\sqrt{x_n}-\sqrt{x} = \lim_{n\to\infty}\frac{x_n-x}{\sqrt{x_n}+\sqrt{x}} = \lim_{n\to\infty}(x_n-x)\cdot\frac{1}{\sqrt{x_n}+\sqrt{x}} = 0,$$

pois $\lim_{n\to\infty}(x_n-x) = 0$ e $\left(\frac{1}{\sqrt{x_n}+\sqrt{x}}\right)_{n\in\mathbb{N}}$ é limitada por $\frac{1}{\sqrt{x}}$ (certo?)*. Disso resulta a identidade $\lim_{n\to\infty}\sqrt{x_n} = \sqrt{x}$ (cf. Exercício 1.84). Se $x=0$, então para $\varepsilon>0$ existe $N\in\mathbb{N}$ tal que $|x_n|<\varepsilon^2$ para todo $n\geq N$. Logo, $|\sqrt{x_n}|<\sqrt{\varepsilon^2}=\varepsilon$ sempre que $n\geq N$, i.e., $\sqrt{x_n}\to 0$ (pode ser útil rever o Exercício 0.134). □

Corolário 1.6.10. *Sejam X e Y subconjuntos de $\mathbb{R}$, bem como funções $f\colon X\to Y$ e $g\colon Y\to\mathbb{R}$. Se f é contínua num ponto $p\in X$ e g é contínua em $f(p)$, então $g\circ f$ é contínua no ponto p.*

Demonstração. Para $(x_n)_{n\in\mathbb{N}}$ sequência real em X, devemos mostrar que se $x_n\to p$, então $(g\circ f)(x_n)\to g(f(p))$. Ora, por f ser contínua em p temos $f(x_n)\to f(p)$. Como $(f(x_n))_{n\in\mathbb{N}}$ é sequência em Y e g é contínua em $f(p)$, resulta $g(f(x_n))\to g(f(p))$, como desejado. □

Exercício 1.106 ($^{\star}_{\star}$)**.** Sejam $f,g\colon X\to\mathbb{R}$ funções reais contínuas em $p\in X$. Mostre que $f+g$ e $f\cdot g$ são contínuas em p. Além disso, se $g(p)\neq 0$, então a função $\frac{f}{g}$ (definida em todo ponto $x\in X$ tal que $g(x)\neq 0$) é contínua em p. ■

[†]... "módulo".

Exemplo 1.6.11. A função $\chi_{\mathbb{Q}} : \mathbb{R} \to \{0,1\}$, que faz $\chi_{\mathbb{Q}}(x) := 1$ se, e somente se, $x \in \mathbb{Q}$, é descontínua em todos os pontos. De fato, para $p \in \mathbb{R} \setminus \mathbb{Q}$, existe uma sequência $(q_n)_{n\in\mathbb{N}}$ em $\mathbb{Q}$ tal que $q_n \to p$ (cf. Exercício 1.100): para cada $n \in \mathbb{N}$ existe q_n no intervalo $\left(p - \frac{1}{2^n}, p + \frac{1}{2^n}\right)$, de modo que $\lim_{n\to\infty} |p - q_n| = 0$ e, consequentemente, $\lim_{n\to\infty} q_n = p$. Todavia, $\lim_{n\to\infty} \chi_{\mathbb{Q}}(q_n) = 1$ e $\chi_{\mathbb{Q}}(p) = 0$. Analogamente (isto é, prove!)*, para $p \in \mathbb{Q}$, existe sequência $(s_n)_{n\in\mathbb{N}}$ em $\mathbb{R} \setminus \mathbb{Q}$ tal que $s_n \to p$, mostrando que $\lim_{n\to\infty} \chi_{\mathbb{Q}}(s_n) = 0$ mas $\chi_{\mathbb{Q}}(p) = 1$. Em ambos os casos, $\chi_{\mathbb{Q}}$ não é contínua em p. ▲

Exemplo 1.6.12 (A função exponencial)**.** O Exemplo 1.4.10 permite definir a função $\exp : \mathbb{R} \to \mathbb{R}$ que a cada número real $x \in \mathbb{R}$ associa $\exp(x) := \sum_{n=0}^{\infty} \frac{x^n}{n!}$, usualmente xingada de **função exponencial**. Vamos ver que tal função é contínua. Primeiro, note que exp é estritamente crescente quando restrita a $[0, +\infty)$, e vale $\exp(a) \geq 1$ sempre que $a \geq 0$ (por quê?)$^{\star}_{\star}$. Agora, observe que para $x, y \in \mathbb{R}$ e $k \in \mathbb{N} \setminus \{0\}$, temos

$$|x^k - y^k| = |x - y| \cdot \left|x^{k-1} + x^{k-2}y + \ldots + xy^{k-2} + y^{k-1}\right| \leq |x - y| \cdot \sum_{j=0}^{k-1} |x|^j |y|^{k-1-j}.$$

Como $\alpha^k \leq \beta^k$ sempre que $0 < \alpha <\leq \beta$ (verifique!)$^{\star}_{\star}$, se $c \geq \max\{|x|, |y|\}$, resulta

$$|x^k - y^k| \leq |x - y| \cdot \sum_{j=0}^{k-1} c^j c^{k-1-j} = |x - y| \cdot \sum_{j=0}^{k-1} c^{k-1} = |x - y| k c^{k-1},$$

de modo que

$$\sum_{k=0}^{m} \frac{|x^k - y^k|}{k!} \leq \sum_{k=1}^{m} |x - y| \frac{k c^{k-1}}{k!} = |x - y| \sum_{j=1}^{m} \frac{c^{k-1}}{(k-1)!} \leq |x - y| \cdot \exp(c).$$

Por fim, resulta

$$|\exp(x) - \exp(y)| = \left| \lim_{n\to\infty} \sum_{k=0}^{n} \frac{x^k - y^k}{k!} \right| \leq \lim_{n\to\infty} \sum_{k=0}^{n} \frac{|x^k - y^k|}{k!} \leq |x - y| \cdot \exp(c)$$

donde segue que para $\varepsilon > 0$ qualquer, $|\exp(x) - \exp(y)| < \varepsilon$ sempre que $|x - y| < \delta$ para um $\delta > 0$ adequado. Ajustar os detalhes será problema seu. ▲

Exercício 1.107 ($^{\star}_{\star\star}$)**.** Você *realmente* encontrou um δ apropriado que depende apenas de x e ε? Dica: note que $|y| \leq |x - y| + |x| < \delta + |x|$ para qualquer $\delta > 0$ que você escolher. ■

Observação 1.6.13. Define-se $e := \exp(1)$, i.e., $e := \sum_{n=0}^{\infty} \frac{1}{n!}$, número real que costuma ser xingado de **constante de Euler**. Futuramente, após provarmos que a identidade $\exp(x + y) = \exp(x) \cdot \exp(y)$ se verifica para quaisquer $x, y \in \mathbb{R}$, seguirá que $\exp(n) = e^n$ para todo $n \in \mathbb{N}$, bem como $\exp(-x) = \frac{1}{\exp(x)}$ para todo $x \in \mathbb{R}$, o que justifica a notação mais comum: $e^x := \exp(x)$. △

Antes de avançar para a Subseção 1.6.1 ou para a Seção 1.7, você pode, alternativamente, conferir as subsubseções §1 e §2 da Subseção 2.1.1 para aprender alguns exemplos mais sofisticados de funções contínuas definidas via séries. A desvantagem de fazer isso agora é que os Teoremas de Weierstrass e do Valor Médio ainda não foram provados. Como não estou com pressa, seguirei por aqui. Boa sorte.

1.6.1 Extras

§0 Continuidade em espaços métricos

Ao trocar os intervalos abertos na definição de continuidade por bolas abertas num espaço métrico (cf. Definição 1.0.11), chega-se à definição de continuidade para funções entre espaços métricos.

Definição 1.6.14. Para espaços métricos (X, d) e (Y, d'), uma função $f: X \to Y$ é **contínua em** $p \in X$ se, e somente se, para toda d'-bola aberta $B \subseteq Y$ com $f(p) \in B$ existe uma d-bola aberta $C \subseteq X$ tal que $p \in C$ e $f[C] \subseteq B$. A função f é **contínua** se for contínua em todos os pontos de seu domínio. ¶

Exercício 1.108 (Definição clássica — $({}^{\star}_{\star\star})$)**.** Mostre que $f: X \to Y$ é contínua em $p \in X$ se, e somente se, para todo $\varepsilon > 0$ existe $\delta > 0$ tal que $d'(f(x), f(p)) < \varepsilon$ sempre que $d'(x, p) < \delta$. Dica: mostre que, em geral, se (Z, ρ) é espaço métrico e $z \in Z$, então z pertence a uma ρ-bola aberta B se, e somente se, existe $r > 0$ tal que $B_\rho(z, r) \subseteq B$. ■

Não deve ser difícil perceber que a definição acima estende a noção de continuidade de funções reais: como sempre, a ideia é trocar "$|x - y|$" por "$d(x, y)$" sem pensar muito no assunto. Talvez a parte menos evidente se refira à interseção presente na Definição 1.6.1: para $X \subseteq \mathbb{R}$, ao definir $d(a, b) := |a - b|$ para quaisquer $a, b \in X$, segue que (X, d) é um espaço métrico. E daí?

Exercício 1.109 $({}^{\star}_{\star})$**.** Com as notações acima, para $x \in X$ e $r > 0$, mostre que vale a identidade $B_d(x, r) = (x - r, x + r) \cap X$. Dica: basta "destrinchar" os significados de $a \in B_d(x, r)$ e $a \in (x - r, x + r) \cap X$; pode ser útil conferir o Exercício 1.6. ■

Com base nisso, você já deve imaginar que o Teorema 1.6.3 permaneça válido para espaços métricos. Imaginou certo! Parabéns... mas terá que provar por conta própria.

Exercício 1.110 $({}^{\star}_{\star\star})$**.** Sejam (X, d) e (Y, d') espaços métricos, e $f: X \to Y$ uma função. Mostre que para qualquer $p \in X$, são equivalentes:

(i) f é contínua em p;

(ii) $\lim\limits_{a \in \mathbb{A}} f(x_a) = f(p)$ para toda rede $(x_a)_{a \in \mathbb{A}}$ em X tal que $x_a \to p$;

(iii) $\lim\limits_{n \to \infty} f(x_n) = f(p)$ para toda sequência $(x_n)_{n \in \mathbb{N}}$ em X tal que $x_n \to p$. ■

Exemplo 1.6.15 (Continuidade das operações em $\mathbb{R}$)**.** As operações usuais de $\mathbb{R}$, a saber $(+): \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ e $(\cdot): \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ são, antes de qualquer outra coisa, funções de $\mathbb{R}^2$ em $\mathbb{R}$. Como $\mathbb{R}^2$ é um espaço métrico, é lícito perguntar se tais operações são contínuas... e a resposta é sim. O mais curioso é que, secretamente, isto já está provado no texto.

Com efeito, em vista do exercício anterior, basta mostrar que $(+)$ e $(\cdot)$ "comutam" com limites de redes ou sequências em $\mathbb{R}^2$. Ora, se $(x_n, y_n)_{n \in \mathbb{N}}$ é uma sequência em $\mathbb{R}^2$ tal que $\lim_{n \to \infty} (x_n, y_n) = (x, y)$, então já sabemos que $\lim_{n \to \infty} x_n = x$ e $\lim_{n \to \infty} y_n = y$ (cf. Proposição 1.1.24). Agora, o Teorema 1.2.1 assegura que

$$\lim_{n \to \infty} (x_n + y_n) = x + y \quad \text{e} \quad \lim_{n \to \infty} (x_n \cdot y_n) = xy,$$

i.e.,

$$\lim_{n \to \infty} (+)(x_n, y_n) = (+)(x, y) \quad \text{e} \quad \lim_{n \to \infty} (\cdot)(x_n, y_n) = (\cdot)(x, y).$$

Moral da história: as propriedades operatórias dos limites reais sempre foram sintomas da continuidade das operações de soma e produto vistas como funções de $\mathbb{R}^2$ em $\mathbb{R}$. ▲

Exercício 1.111 ($\star$)**.** Enuncie e demonstre a versão análoga do Corolário 1.6.10 para funções entre espaços métricos. ■

Exemplo 1.6.16. Ainda por conta do Exercício 1.110, as projeções

$$\begin{aligned}\pi_1\colon \mathbb{R}\times\mathbb{R} &\to \mathbb{R} \\ (x,y) &\mapsto x\end{aligned} \qquad \text{e} \qquad \begin{aligned}\pi_2\colon \mathbb{R}\times\mathbb{R} &\to \mathbb{R} \\ (x,y) &\mapsto y\end{aligned}$$

são contínuas (certo?)$^{\star}_{\star}$. Disto segue que se $F\colon \mathbb{R}\to\mathbb{R}\times\mathbb{R}$ for uma função contínua, então também serão contínuas as funções $\pi_1\circ F$ e $\pi_2\circ F$ (pelo exercício anterior). O que isto significa? Vejamos: por definição, F associa a cada $x\in\mathbb{R}$ um ponto $F(x)\in\mathbb{R}\times\mathbb{R}$ que deve ter duas coordenadas, digamos $F(x) := (f_1(x), f_2(x))$. Assim, o que a afirmação anterior assegura é que se F é contínua, então f_1 e f_2 são contínuas.

A graça é que vale a volta: se $f\colon \mathbb{R}\to\mathbb{R}$ e $g\colon\mathbb{R}\to\mathbb{R}$ são funções contínuas, então também é contínua a função $(f,g)\colon \mathbb{R}\to\mathbb{R}\times\mathbb{R}$ que faz $(f,g)(x) := (f(x), g(x))$. De fato, se $x_n\to x$ em $\mathbb{R}$, então $f(x_n)\to f(x)$ e $g(x_n)\to g(x)$, donde a Proposição 1.1.24 assegura que $(f(x_n), g(x_n))\to(f(x),g(x))$ em $\mathbb{R}^2$, mostrando que (f,g) é contínua (novamente pelo Exercício 1.110). ▲

Exercício 1.112 ($^{\star}_{\star}$)**.** Para $f,g\colon X\to\mathbb{R}$, mostre que $f+g = (+)\circ(f,g)$ e $f\cdot g = (\cdot)\circ(f,g)$. Use isto para provar que $f+g$ e $f\cdot g$ são contínuas sempre que f e g são contínuas. ■

§1 Continuidade de transformações lineares

Quando os espaços métricos são, na verdade, espaços (vetoriais!) normados, temos ainda mais alguns métodos para detectar a continuidade de algumas funções... Infelizmente, isto vale apenas para as *transformações lineares* (cf. Definição 0.6.23), mas não se engane: em dimensões *maiores*[†], transformações lineares costumam ser bem mais interessantes do que em dimensão 1.

Proposição 1.6.17. *Para espaços (vetoriais) normados $(E,\|\cdot\|_E)$ e $(S,\|\cdot\|_S)$ e uma transformação linear $T\colon E\to S$, são equivalentes:*

(i) *T é contínua;*

(ii) *T é contínua em 0_E;*

(iii) *existe $r>0$ tal que $\|T(x)\|_S\leq r\cdot\|x\|_E$ para todo $x\in E$.*

Demonstração. É claro que (i) ⇒ (ii). Para (ii) ⇒ (iii), é mais conveniente usar a definição de continuidade em termos de ε's e δ's. Como T é contínua em 0_E, para $\varepsilon := 1$ existe $\delta>0$ tal que

$$\|T(x)\|_S = \|T(x)-T(0_E)\|_S < 1$$

sempre que $\|x\|_E = \|x-0_E\|_E<\delta$. Logo, para qualquer $x\neq 0_E$ e δ' com $0<\delta'<\delta$ temos $u := \frac{\delta'}{\|x\|_E}\cdot x$ satisfazendo $\|u\|_E<\delta$ (certo?)*, de modo que

$$\|T(u)\|_S = \left\|T\left(\frac{\delta'}{\|x\|_E}x\right)\right\|_S = \frac{\delta'}{\|x\|_E}\|T(x)\|_S < 1 \Rightarrow \|T(x)\|_S < \frac{1}{\delta'}\|x\|_E,$$

[†] Talvez você já tenha ouvido falar que *toda* transformação linear é contínua. Mas isto é falso: em dimensão *infinita* sempre é possível encontrar transformações lineares descontínuas (cf. Observação 2.0.30).

mostrando que basta tomar $r := \frac{1}{\delta'}$. Para (iii) $\Rightarrow$ (i), a linearidade de T dá

$$\|T(x) - T(y)\|_S = \|T(x-y)\|_S \leq r\|x-y\|_E,$$

e assim, $\|T(x) - T(y)\|_S < \varepsilon$ sempre que $\|x-y\|_E < \frac{\varepsilon}{r}$. □

Uma consequência surpreendente desta banalidade?

Teorema 1.6.18. *Seja $(f_n)_{n\in\mathbb{N}}$ uma sequência de funções em $\mathcal{B}[a,b]$, i.e., cada f_n é uma função limitada da forma $[a,b] \to \mathbb{R}$. Se cada f_n é Riemann-integrável e, além disso, $f_n \to f$ em $\mathcal{B}[a,b]$, onde $f\colon [a,b] \to \mathbb{R}$ também é Riemann-integrável, então*

$$\lim_{n\to\infty} \int_a^b f_n(t)\,\mathrm{d}t = \int_a^b f(t)\,\mathrm{d}t.$$

Demonstração. O subconjunto $\mathcal{RI}[a,b]$ das funções Riemann-integráveis é um subespaço vetorial de $\mathcal{B}[a,b]$ (cf. Exemplo 1.2.29), o que permite considerar $\mathcal{RI}[a,b]$ como espaço vetorial normado pela norma do supremo $\|\cdot\|_\infty$. Ainda pelo exemplo supracitado, a função

$$\begin{aligned} \int_a^b \colon \mathcal{RI}[a,b] &\to \mathbb{R} \\ g &\mapsto \int_a^b g(t)\,\mathrm{d}t \end{aligned}$$

é uma transformação linear. Finalmente, pela última proposição, o Exercício 1.45 garante a continuidade da integral enquanto transformação linear, com $r := (b-a)$. Isto encerra a prova (percebeu?)*. □

Na verdade, as hipóteses do último corolário podem ser afrouxadas: uma vez em posse de uma caracterização "melhor" de Riemann-integrabilidade, será possível assegurar que se uma sequência $(f_n)_{n\in\mathbb{N}}$ de funções Riemann-integráveis *converge uniformemente* para uma função $f\colon [a,b] \to \mathbb{R}$, então f será Riemann-integrável (cf. Teorema 2.4.32). Mas ainda é cedo para se preocupar com isso.

§2 Exercícios extras

Exercício 1.113 ($\star$). Lembre-se de que um polinômio real nas indeterminadas x e y é uma expressão polinomial

$$P(x,y) := \sum_{i=0}^{m} \sum_{j=0}^{n} a_{ij} x^i y^j$$

em que $a_{ij} \in \mathbb{R}$ para quaisquer $i \in \{0,\ldots,m\}$ e $j \in \{0,\ldots,n\}$. Uma função $f\colon \mathbb{R}^2 \to \mathbb{R}$ é chamada de (função) *polinomial* se existe um polinômio real $P(x,y)$ nas indeterminadas x e y tais que $f(s,t) = P(s,t)$ para quaisquer $s,t \in \mathbb{R}$. Em tais condições, mostre que f é contínua. Dica: as projeções usuais são polinomiais. ■

Exercício 1.114 ($\star$). Generalize o exercício anterior para funções polinomiais do tipo $\mathbb{R}^n \to \mathbb{R}$. Dica: aplique a mesma dica. ■

Exercício 1.115 ($\star$). Para $m,n \in \mathbb{N}\setminus\{0\}$ com $m < n$, sejam $a_1,\ldots,a_m$ tais que

$$1 \leq a_1 < a_2 < \ldots < a_m \leq n,$$

i.e., uma lista crescente de m números naturais entre 1 e n. Defina a função $\Pi\colon \mathbb{R}^n \to \mathbb{R}^m$ dada por $\Pi(x_1,\ldots,x_n) := (x_{a_1},\ldots,x_{a_m})$. Mostre que Π é contínua. Dica: perceba que $\Pi = (\pi_{a_1},\ldots,\pi_{a_m})$, onde $\pi_{a_i}\colon \mathbb{R}^n \to \mathbb{R}$ é a projeção de $\mathbb{R}^n$ na a_i-ésima coordenada. ■

Sugestão: caso tenha parecido muito simbólico: com $m := 3$ e $n := 7$, considere $a_1 := 2$, $a_2 := 5$ e $a_3 := 6$, por exemplo. Neste caso, $\Pi(x_1, \ldots, x_7) := (x_2, x_5, x_6)$. Não é errado inventar exemplos concretos para entender definições abstratas — o problema é não fazer isso por conta própria.

Exercício 1.116 $\left({}^{\star}_{\star\star}\right)$**.** Mostre que a função determinante é contínua. Sugestão: interprete a função determinante como uma correspondência do tipo $\mathbb{R}^{n^2} \to \mathbb{R}$. Dica: ao seguir a sugestão, mostre por indução que a função determinante se expressa como combinação linear de funções contínuas (lembre-se da fórmula de Laplace para o determinante). ■

Exercício 1.117 $(\star)$**.** Mostre que a norma de um espaço normado é sempre contínua. ■

Exercício 1.118 $\left({}^{\star}_{\star}\right)$**.** Para um espaço normado X, com $v \in X$ fixado, mostre que a função $f\colon \mathbb{R} \to X$, dada por $f(t) := tv$ para todo $t \in \mathbb{R}$, é contínua. ■

Exercício 1.119 (Cf. Exemplo 1.6.16 — $\left({}^{\star}_{\star}\right)$)**.** Seja $F\colon \mathbb{R}^2 \to \mathbb{R}$ uma função.

a) Mostre que se F é contínua, então para quaisquer $a, b \in \mathbb{R}$ fixados, também serão contínuas as funções

$$\begin{aligned} F_a\colon \mathbb{R} &\to \mathbb{R} \\ y &\mapsto F(a,y) \end{aligned} \quad \text{e} \quad \begin{aligned} F^b\colon \mathbb{R} &\to \mathbb{R} \\ x &\mapsto F(x,b) \end{aligned}.$$

Dica: composição de funções contínuas é contínua, e tanto $y \mapsto (a,y)$ quanto $x \mapsto (x,b)$ definem funções contínuas da forma $\mathbb{R} \to \mathbb{R}^2$.

b) Seja $F\colon \mathbb{R}^2 \to \mathbb{R}$ a função que a cada par $(x,y) \neq (0,0)$ associa o número

$$F(x,y) := \frac{xy}{x^2 + y^2},$$

enquanto faz $F(0,0) := 0$. Mostre que F_a e F^b, como definidas acima, são contínuas para quaisquer $a, b \in \mathbb{R}$.

c) A função $F\colon \mathbb{R}^2 \to \mathbb{R}$ no item anterior é contínua? Dica: teste a continuidade em $(0,0)$, e perceba que existem sequências em $\mathbb{R}^2$ que convergem para $(0,0)$, mas cujas imagens por F não convergem para 0 (por exemplo, $(x_n, y_n) := \left(\frac{1}{2^n}, \frac{1}{2^n}\right)$ já faz o serviço, certo?). ■

Exercício 1.120 $\left({}^{\star}_{\star}\right)$**.** Assumindo conhecidos os fatos básicos de trigonometria, mostre que a função $f\colon [0, 2\pi) \to \mathbb{R}^2$ dada por $f(t) := (\cos(t), \operatorname{sen}(t))$ é contínua em todo $t \in \mathbb{R}$. Qual é a imagem de f? E qual é o *gráfico* de f? ■

Observação 1.6.19. Funções contínuas cujos domínios são intervalos da reta real costumam ser chamadas de **caminhos**, *principalmente* quando o intervalo é $[0, 1]$. △

1.7 A topologia da reta

1.7.0 Essencial

§0 Abertos da reta e abertos relativos

Para tratar da última caracterização de continuidade, precisamos revelar outra importante forma de abordar as questões de convergência. Ela se baseia numa observação muito simples: na Figura 1.10, tudo funcionaria de forma idêntica se, em vez de intervalos abertos I e J com $f(p) \in I$ e $p \in J$, tivéssemos apenas subconjuntos I' e J' com $I \subseteq I'$ e $J \subseteq J'$.

Definição 1.7.0. Diremos que um subconjunto $A \subseteq \mathbb{R}$ é **aberto em** $\mathbb{R}$ (ou **de** $\mathbb{R}$) se para todo $x \in A$ existe um intervalo aberto $I \subseteq \mathbb{R}$ tal que $x \in I$ e $I \subseteq A$. ¶

Exercício 1.121 ($\substack{\star \\ \star}$)**.** Mostre que $A \subseteq \mathbb{R}$ é aberto em $\mathbb{R}$ se, e somente se, para todo $x \in A$ existe $r > 0$ tal que $(x - r, x + r) \subseteq A$. ■

Por vacuidade, $\emptyset$ é um subconjunto aberto de $\mathbb{R}$: não há $x \in \emptyset$ que possa testemunhar o contrário. Analogamente, $\mathbb{R}$ é aberto em $\mathbb{R}$ (não há intervalo aberto de $\mathbb{R}$ que não esteja contido em $\mathbb{R}$). Intervalos abertos também são abertos em $\mathbb{R}$, mas não por culpa da gramática: se $I \subseteq \mathbb{R}$ é intervalo aberto e $x \in I$, então o próprio I é um intervalo aberto tal que $x \in I$ e $I \subseteq I$.

Exercício 1.122 ($\star$)**.** Mostre que intervalos abertos são abertos em $\mathbb{R}$ usando a caracterização do exercício anterior[†]. ■

Exemplo 1.7.1 (Nem todo aberto de $\mathbb{R}$ é intervalo aberto)**.** A definição dos abertos de $\mathbb{R}$ se aplica a outros tipos de subconjuntos. Por exemplo, se $A := (0, 2) \cup (3, 5)$, então A é aberto em $\mathbb{R}$, mas não é um intervalo. É aberto em $\mathbb{R}$ pois para todo $x \in A$ existe um intervalo aberto $I \subseteq A$ tal que $x \in I$ e $I \subseteq A$: se $0 < x < 2$, então basta tomar $I := (0, 2)$; se $3 < x < 5$, então basta fazer $I := (3, 5)$. No entanto, A não é intervalo: temos $1, 4 \in A$, com $1 < 3 < 4$, mas $3 \notin A$ (cf. Definição 0.8.13). ▲

Exercício 1.123 (?!)**.** Leia o exemplo anterior dez vezes. Em particular, perceba que um conjunto que se escreve como reunião de intervalos não precisa ser intervalo. ■

Exemplo 1.7.2 (Nem tudo é aberto!)**.** O subconjunto $S := [0, 1)$ não é aberto em $\mathbb{R}$, e a testemunha é o ponto 0: embora $0 \in S$, não há intervalo aberto $I \subseteq \mathbb{R}$ tal que $0 \in I$ e $I \subseteq S$ (certo?)*. Analogamente, $\mathbb{Q}$ não é aberto em $\mathbb{R}$ e, mais geralmente, nenhum subconjunto enumerável de $\mathbb{R}$ pode ser aberto em $\mathbb{R}$ (por quê?)$^{\substack{\star \\ \star}}$. ▲

Lema 1.7.3. *Seja $A \subseteq \mathbb{R}$ um subconjunto. Se para todo $x \in A$ existir um subconjunto $B \subseteq \mathbb{R}$ aberto em $\mathbb{R}$ tal que $x \in B$ e $B \subseteq A$, então A é aberto em $\mathbb{R}$.*

Demonstração. Como B é aberto em $\mathbb{R}$, existe intervalo aberto $I \subseteq \mathbb{R}$ tal que $x \in I$ e $I \subseteq B$. Como $B \subseteq A$, o resultado segue. □

[†] "*Como assim a caracterização do exercício anterior? Por que você usa tantas palavras difíceis??!!*". Veja: no exercício anterior, você mostrou que A é aberto em $\mathbb{R}$ (segundo a Definição 1.7.0) se, e somente se, "para todo $x \in A$ existe $r > 0$ tal que $(x - r, x + r) \subseteq A$". Isto significa que a condição entre aspas *também* serve para determinar o que é um aberto de $\mathbb{R}$. Ou seja, tal condição serve como "caracterização", no sentido de determinar quais características o objeto precisa ter para ser xingado de aberto em $\mathbb{R}$.

Proposição 1.7.4. *Seja Λ um conjunto. Se para cada $\lambda \in \Lambda$ for dado um subconjunto aberto $A_\lambda \subseteq \mathbb{R}$, então $\bigcup_{\lambda\in\Lambda} A_\lambda$ é subconjunto aberto de $\mathbb{R}$.*

Demonstração. Primeiro, lembre-se que $x \in \bigcup_{\lambda\in\Lambda} A_\lambda$ se, e somente se, existe $\lambda' \in \Lambda$ tal que $x \in A_{\lambda'}$ (cf. Definição 0.1.20). Assim, chamando $A := \bigcup_{\lambda\in\Lambda} A_\lambda$, *perceba* que se $x \in A$, então existe um aberto $A_{\lambda'} \subseteq \mathbb{R}$ tal que $x \in A_{\lambda'}$ e $A_{\lambda'} \subseteq A$, donde o resultado segue pelo lema anterior. □

Exercício 1.124 ($\star$)**.** Perceba! ■

Definição 1.7.5 (Abertos relativos)**.** Seja $X \subseteq \mathbb{R}$ um subconjunto. Diremos que $R \subseteq X$ é um subconjunto **aberto de** X (ou **em** X) se existir um subconjunto $A \subseteq \mathbb{R}$ aberto em $\mathbb{R}$ tal que $R = X \cap A$. ¶

Exemplo 1.7.6. Para $X := [0,3]$, o subconjunto $R := [0,1)$ é aberto em X, pois $[0,1) = (-1,1) \cap X$. "*Ah, mas eu havia pensado em fazer* $[0,1) = (-\infty,1) \cap X$*!!! Por que eu nunca acerto uma?!!!*" Calma... A definição de aberto relativo não diz que a testemunha A deva ser única: "existe um" é diferente de "existe um único". ▲

Observação 1.7.7 (Importante)**.** Veja que um subconjunto R pode ser aberto num subconjunto de $\mathbb{R}$ sem ser aberto em $\mathbb{R}$: já vimos que $[0,1)$ não é aberto em $\mathbb{R}$ mas, pelo exemplo anterior, $[0,1)$ é aberto em $[0,3]$. "*Mas por quê??!! Por que um subconjunto pode ser aberto num X mas não ser aberto em $\mathbb{R}$??!! Ahhh eu odeio Análise!!!*"...

Resposta: em certo sentido, ao considerar os abertos relativos de $[0,3]$, ignoram-se todos os pontos de $\mathbb{R}$ fora de $[0,3]$, de modo que os intervalos abertos que causavam problema "deixam de existir". △

Exercício 1.125 ($^\star_\star$)**.** Considere X um subconjunto de $\mathbb{R}$. Mostre que:

a) $\emptyset$ e X são abertos em X;

b) se $A, B \subseteq X$ são abertos em X, então $A \cap B$ é aberto em X;

c) se $A_\lambda \subseteq X$ é aberto em X para cada $\lambda \in \Lambda$, então $\bigcup_{\lambda\in\Lambda} A_\lambda$ é aberto em X. ■

Exercício 1.126 ($^\star_\star$)**.** Mostre que $A \subseteq X$ é aberto em X se, e somente se, para todo $a \in A$ existe um subconjunto $B \subseteq X$, aberto em X, tal que $a \in B$. ■

§1 Continuidade topológica

Para um conjunto Z, que nada precisa ter a ver com $\mathbb{R}$, uma família $\mathcal{T}$ de subconjuntos de Z é chamada de **topologia** em Z se

(i) $\emptyset, Z \in \mathcal{T}$;

(ii) $A \cap B \in \mathcal{T}$ sempre que $A, B \in \mathcal{T}$;

(iii) $\bigcup_{\lambda\in\Lambda} A_\lambda \in \mathcal{T}$ sempre que $\{A_\lambda : \lambda \in \Lambda\} \subseteq \mathcal{T}$.

Dessa forma, ao longo desta primeira parte, mostramos que para $X \subseteq \mathbb{R}$, a *família* $\mathcal{T}_X$, cujos membros são os subconjuntos de X abertos em X, constitui uma *topologia* em X. O que isso tudo tem a ver com continuidade?

Teorema 1.7.8 (Continuidade de funções reais, via pré-imagens de abertos)**.** *Para um subconjunto $X \subseteq \mathbb{R}$, uma função $f\colon X \to \mathbb{R}$ é contínua se, e somente se, $f^{-1}[A]$ for aberto em X sempre que A for aberto em $\mathbb{R}$, onde $f^{-1}[A]$ indica a pré-imagem de A por f.*

Demonstração. Se f é contínua e $A \subseteq \mathbb{R}$ é aberto, pode-se ter $f^{-1}[A] = \emptyset$ ou $f^{-1}[A] \neq \emptyset$, com o primeiro caso automaticamente resolvido pelo exercício anterior. No segundo caso, para $x \in f^{-1}[A]$, temos $f(x) \in A$ pela definição de pré-imagem (cf. Exercício 0.55) e, por A ser aberto em $\mathbb{R}$, existe um intervalo aberto $I \subseteq \mathbb{R}$ tal que $f(x) \in I$ e $I \subseteq A$. Logo, pela definição de continuidade, existe um intervalo aberto $J \subseteq \mathbb{R}$ tal que $x \in J$ e $f(z) \in I$ sempre que $z \in X \cap J$. Perceba o que fizemos: $X \cap J$ é um aberto em X tal que $x \in X \cap J$ e $X \cap J \subseteq f^{-1}[A]$ (percebeu?)*. Ou seja: pelo exercício anterior, mostramos que $f^{-1}[A]$ é aberto em X.

Para a recíproca, basta mostrar que f é contínua em x, para qualquer $x \in X$. Dado um intervalo aberto $I \subseteq \mathbb{R}$ com $f(x) \in I$, precisamos encontrar outro intervalo aberto $J \subseteq \mathbb{R}$ em torno de x tal que $f(z) \in I$ sempre que $z \in X \cap J$. Ora, por I ser aberto em $\mathbb{R}$, sabemos que $f^{-1}[I]$ é aberto em X, digamos $f^{-1}[I] = X \cap O$ para algum aberto $O \subseteq \mathbb{R}$. Como $x \in f^{-1}[I]$ (pois $f(x) \in I$!!), resulta que $x \in O$ e, pela definição dos abertos de $\mathbb{R}$, existe um intervalo aberto $J \subseteq \mathbb{R}$ tal que $x \in J$ e $J \subseteq O$. Finalmente, note que se $z \in X \cap J$, então $z \in X \cap O = f^{-1}[I]$ e, portanto, $f(z) \in I$, como desejado. □

Observação 1.7.9. Pré-imagem não é a imagem pela função inversa.

Exercício 1.127 (!!!)**.** Releia a frase acima cinquenta vezes. ■

Por que a pré-imagem não é a imagem pela função inversa?

Resposta rápida: o significado dos símbolos depende do contexto.

Resposta longa: uma função só admite inversa quando é bijetora, e a pré-imagem de f por A é, *por definição*, o subconjunto $f^{-1}[A] := \{x \in X : f(x) \in A\}$. Ou seja: f **não precisa ser bijetora para considerarmos pré-imagens**. Nas situações em que, por sorte, f é bijetora, aí sim, as duas noções coincidem.

Poderíamos utilizar uma notação menos ambígua para pré-imagens? Sim. Há textos que empregam "$f^{\leftarrow}[A]$" para indicar a pré-imagem de A pela função f, por exemplo. No entanto, a maioria costuma assumir que ninguém ~~em pleno exercício de suas faculdades mentais~~ acrescentaria uma hipótese tão forte como "bijetividade" sem perceber. Portanto, atente-se! △

O salto conceitual que demos foi alto: antes das considerações topológicas, a noção de continuidade (e, por extensão, convergência) parecia mais intuitiva, já que podíamos nos apegar a *números* e *distâncias* para *calcular* e *estimar* a convergência; enquanto isso, uma topologia parece apenas uma sacola cheia de informações abstratas que alguém inventou só para impedir você de se formar. Contudo, não é ~~só~~ para isso que ela serve.

As topologias formalizam a impressão que você provavelmente já tinha de que não precisamos de uma régua ótima para decidir se uma função é contínua ou se uma rede (ou sequência) converge, já que todos os procedimentos envolvem *desigualdades*, estimativas, aproximações, etc. Além disso, em situações em que mais de uma noção de distância é razoável[†], topologias permitem decidir quando tais noções de distância são *equivalentes* — mas este é um problema mais natural em $\mathbb{R}^n$ para $n > 1$. Mesmo assim, na reta real, *argumentos topológicos* costumam ser mais concisos e, simultaneamente, abrangentes, como veremos pelo restante do curso.

[†]Cf. Subseção 1.1.1 §2.

1.7.1 Extras

§0 Espaços topológicos e homeomorfismos

Definição 1.7.10. Um **espaço topológico** é um par $(X, \mathcal{T})$ onde X é um conjunto e $\mathcal{T}$ é uma topologia em X. ¶

Diariamente, os membros da topologia $\mathcal{T}$ de um espaço topológico $(X, \mathcal{T})$ são chamados de **abertos de** X (ou **em** X) **segundo a topologia** $\mathcal{T}$ — ou ainda, se preferir, $\mathcal{T}$**-abertos**. Nesse sentido, os axiomas que definem uma topologia se reescrevem assim: $\emptyset$ e X são abertos, a interseção finita de abertos é aberta e a reunião arbitrária de abertos é aberta. Secretamente, o Teorema 1.7.8 diz como definir *funções contínuas* entre espaços topológicos:

Definição 1.7.11. Para $(X, \mathcal{T})$ e $(Y, \mathcal{S})$ espaços topológicos, uma função $f\colon X \to Y$ é **contínua** se $f^{-1}[S] \in \mathcal{T}$ sempre que $S \in \mathcal{S}$. Verbalmente: "a pré-imagem de abertos de Y é aberta em X". ¶

Exemplo 1.7.12. Ao considerar $\mathbb{R}$ com a topologia $\mathcal{T}_{\mathbb{R}}$ formada pelos abertos de $\mathbb{R}$ e $X \subseteq \mathbb{R}$ com a topologia $\mathcal{T}_X$ formada por seus abertos relativos, segue que $f\colon X \to \mathbb{R}$ é contínua (como na Definição 1.6.1) se, e somente se, f é contínua no sentido da definição anterior (cf. Teorema 1.7.8!). ▲

Exemplo 1.7.13 (Topologia de espaços métricos)**.** Todo espaço métrico (X, d) vem de fábrica com uma topologia $\mathcal{T}_d$ que se diz ser **induzida por** d: um subconjunto $A \subseteq X$ é $\mathcal{T}_d$-aberto se, e somente se, para todo $x \in A$ existe uma bola aberta $B \subseteq A$ tal que $x \in B$ (verifique que isto define uma topologia!)$\stackrel{\star}{\star}$. Classicamente, porém, a definição é diferente: pede-se que para todo $x \in A$ exista $r > 0$ tal que $B_d(x, r) \subseteq A$. A equivalência entre as duas definições se dá pelo seguinte[†]:

Exercício 1.128 $(\stackrel{\star}{\star})$**.** Mostre que bolas abertas são abertas. Dica: você já fez isso (se precisar de ajuda, use a dica do Exercício 1.108). ■

Exercício 1.129 $(\star)$**.** Mostre que $H := \{(x, y) \in \mathbb{R}^2 : x, y > 0\}$ é aberto em $\mathbb{R}^2$ na topologia induzida por qualquer uma das normas $\|\cdot\|_1, \|\cdot\|_2$ ou $\|\cdot\|_\infty$. Dica: faça apenas para uma das normas e daí aplique o próximo exercício. ■

Exercício 1.130 $(\stackrel{\star}{\star})$**.** Mostre que as normas $\|\cdot\|_1$, $\|\cdot\|_2$ e $\|\cdot\|_\infty$ induzem a mesma topologia em $\mathbb{R}^2$. Dica: recicle a discussão feita nos itens (i), (ii) e (iii) na página 123. ■

Exercício 1.131 $(\stackrel{\star}{\star})$**.** Mostre que com a métrica discreta d_X num conjunto X (cf. Exemplo 1.1.22), todo subconjunto de X é d_X-aberto. Conclua que a topologia induzida pela métrica discreta em $\mathbb{R}^2$ não é a topologia induzida pela norma $\|\cdot\|_2$. ■

Agora, para espaços métricos (X, d) e (Y, d'), o que significa dizer que $f\colon X \to Y$ é contínua com respeito às suas topologias?

Resposta: a mesma coisa que a Definição 1.6.14!

[†]Compare com a dica fornecida para o Exercício 1.104.

De fato, se f é "metricamente contínua" e $S \subseteq Y$ é aberto (em Y)[†] segundo a topologia $\mathcal{T}_{d'}$, então para qualquer $x \in f^{-1}[S]$ vale $f(x) \in S$, de modo que a definição de $\mathcal{T}_{d'}$ garante uma d'-bola aberta $B' \subseteq S$ com $f(x) \in B' \subseteq S$. Pela "continuidade métrica", existe uma d-bola aberta $B \subseteq X$ tal que $x \in B$ e $f[B] \subseteq B' \subseteq S$, ou seja, $x \in B$ e $B \subseteq f^{-1}[S]$, mostrando que $f^{-1}[S]$ é $\mathcal{T}_d$-aberto (por quê?)[‡]. Reciprocamente, se f é "topologicamente contínua", então f é "metricamente contínua" por conta do Exercício 1.128. Você pode cuidar dos detalhes $({}^{\star}_{\star\star})$. ▲

A relação com espaços métricos pode levantar algumas perguntas interessantes:

(i) faz sentido falar de funções (topologicamente) contínuas em pontos específicos?

(ii) faz sentido falar de convergência em espaços topológicos?

(iii) as funções contínuas ainda são aquelas que preservam convergência nesse contexto?

A resposta para todas as perguntas é "sim", embora a última tenha ressalvas. Dizer que uma rede $(x_d)_{d\in\mathbb{D}}$ **num espaço topológico** $(X, \mathcal{T})$ **converge** para um ponto $x \in X$ significa que para todo $\mathcal{T}$-aberto $V \subseteq X$ com $x \in V$ existe $D \in \mathbb{D}$ tal que $x_d \in V$ sempre que $d \succeq D$, o que ainda abreviaremos com $x_d \to x$. Note que tal definição estende as noções anteriores de convergência de redes (note!)$^{\star}_{\star}$. Com isso, vale o seguinte.

Teorema 1.7.14. *Para espaços topológicos* $(X, \mathcal{T})$ *e* $(Y, \mathcal{S})$*, uma função* $f\colon X \to Y$ *e um ponto* $p \in X$*, são equivalentes:*

(i) *para todo* $\mathcal{S}$*-aberto* $U \subseteq Y$ *tal que* $f(p) \in U$*, existe um* $\mathcal{T}$*-aberto* $V \subseteq X$ *com* $p \in V$ *e* $f[V] \subseteq U$*;*

(ii) $f(x_d) \to f(p)$ *em* Y *para qualquer rede* $(x_d)_{d\in\mathbb{D}}$ *em* X *tal que* $x_d \to p$*.*

Demonstração. Para (i) ⇒ (ii), basta imitar o argumento da etapa correspondente no Teorema 1.6.3. Para a recíproca, é suficiente trocar $\mathcal{I}_p$ por $\mathcal{T}_p := \{V \in \mathcal{T} : p \in V\}$ no argumento que antecede o mesmo teorema. □

Exercício 1.132 $({}^{\star}_{\star\star})$**.** Complete a demonstração. ■

Definição 1.7.15. Uma função $f\colon X \to Y$ entre espaços topológicos $(X, \mathcal{T})$ e $(Y, \mathcal{S})$ é **contínua em** $p \in X$ se qualquer uma das condições no teorema anterior for satisfeita. ¶

Exercício 1.133 $({}^{\star}_{\star})$**.** Mostre que $f\colon X \to Y$ é contínua se, e somente se, f é contínua em todos os pontos de X. Dica: numa das direções será útil perceber que o Exercício 1.126 permanece válido em espaços topológicos. ■

E as ressalvas? Há duas principais: a primeira se refere à unicidade dos limites, que não é garantida em espaços topológicos quaisquer; a segunda se refere às sequências convergentes, que nem sempre podem substituir as redes no último teorema.

Sim, você pode esbarrar com um espaço topológico $(X, \mathcal{T})$ em que uma rede $(x_d)_{d\in\mathbb{D}}$ converge para dois pontos distintos. Mas não se preocupe: já provamos que em espaços métricos isto não acontece, então não será agora que passará a acontecer (em espaços métricos!). Esse tipo de comportamento *pode* ocorrer em topologias que não são induzidas por métricas.

[†]Com o tempo você perceberá que isto não precisa ser enfatizado a todo momento.

[‡]Um argumento parecido já foi usado anteriormente, junto com um lema esperto; adapte o lema $({}^{\star}_{\star\star})$.

Exercício 1.134 (Opcional — $(^{\star}_{\star})$)**.** Um espaço topológico é chamado de (espaço de) **Hausdorff** se para quaisquer pontos distintos x e y existem abertos disjuntos A e B tais que $x \in A$ e $y \in B$. Mostre que se X é de Hausdorff e $(x_d)_{d\in\mathbb{D}}$ é uma rede em X tal que $x_d \to x$ e $x_d \to y$, então $x = y$. **For fun** $(^{\star}_{\star\star})$**:** mostre que vale a volta![†] ■

A questão das sequências convergentes é mais delicada: existem situações em que $f(x_n) \to f(p)$ sempre que uma sequência $(x_n)_n$ converge para p mas, ainda assim, a função f não é contínua em p no sentido do item (i) do último teorema. Evidentemente, isto nunca ocorre quando as topologias envolvidas são induzidas por métricas: você já provou isso no Exercício 1.110, não precisa se preocupar. Em todo caso, esta é uma das motivações para o estudo de redes: permitir que convergência e continuidade se descrevam mutuamente em cenários *não*-metrizáveis.

Respondidas as perguntas, outro aspecto fundamental da topologia é dar sentido preciso a certas relações de *similaridade* que alguns espaços parecem manter com outros.

Definição 1.7.16. Dois espaços topológicos $(X, \mathcal{T})$ e $(Y, \mathcal{S})$ são **homeomorfos** se existe uma função bijetora contínua $f\colon X \to Y$ cuja inversa $f^{-1}\colon Y \to X$ também é contínua. Em tais situações, a função f é chamada de **homeomorfismo.** ¶

Exemplo 1.7.17. Os intervalos $(-1, 1)$ e $(-\infty, +\infty)$ certamente são distintos. Apesar disso, eles são *parecidos*: é como se $(-\infty, +\infty)$ fosse $(-1, 1)$ *esticado*. Embora as distâncias exatas sejam alteradas nesse processo de "esticamento", *algo se preserva*. Formalmente, o esticamento pode ser dado pela função

$$\begin{aligned} \varphi\colon (-1,1) &\to \mathbb{R} \\ t &\mapsto \frac{t}{1-|t|} \end{aligned}$$

contínua em virtude das propriedades de funções contínuas, bijetora por ter

$$\begin{aligned} \psi\colon \mathbb{R} &\to (-1,1) \\ t &\mapsto \frac{t}{1+|t|} \end{aligned}$$

como inversa, homeomorfismo pois ψ é contínua (certo?)*. O que tal homeomorfismo preserva? Resposta: convergência. Por exemplo: se $(x_n)_{n\in\mathbb{N}}$ é uma sequência em $(-1, 1)$ que converge para $\frac{9}{10}$, então

$$\lim_{n\to\infty} \varphi(x_n) = \varphi\left(\frac{9}{10}\right) = 9,$$

mostrando que φ leva $(x_n)_n$ a pontos arbitrariamente próximos de $\varphi\left(\frac{9}{10}\right)$. Analogamente, ψ leva pontos arbitrariamente próximos de $r \in \mathbb{R}$ a pontos arbitrariamente próximos de $\psi(r)$. ▲

Exemplo 1.7.18 (Informal)**.** Os intervalos $(0, 1)$ e $[0, 1)$ não são homeomorfos. De fato, se $f\colon [0, 1) \to (0, 1)$ fosse um homeomorfismo, então f levaria 0 a algum ponto de $(0, 1)$, digamos $f(0) = \frac{1}{2}$. Porém, se isso ocorresse, então ao retirar o ponto 0 de $[0, 1)$, resultaria que $(0, 1)$ seria homeomorfo ao conjunto $(0, 1) \setminus \left\{\frac{1}{2}\right\} = \left(0, \frac{1}{2}\right) \cup \left(\frac{1}{2}, 1\right)$, mas isto não pode ocorrer, já que o segundo tem "dois pedaços", enquanto o primeiro tem apenas um "pedaço". No próximo capítulo, este raciocínio será formalizado por meio da noção de *conexidade*. ▲

[†]Em particular, todo espaço métrico é de Hausdorff, bem como a reta estendida, mas você já sabia disso (cf. Exercício 1.27).

Exemplo 1.7.19. Para encerrar, note que se d e d' são duas métricas sobre o mesmo conjunto X, então $\mathcal{T}_d = \mathcal{T}_d'$ se, e somente se, a função identidade $\mathrm{Id}_X : X \to X$ é um homeomorfismo entre $(X, \mathcal{T}_d)$ e $(X, \mathcal{T}_{d'})$ (por quê?)$^{\star}_{\star}$. Em particular, se E é um espaço vetorial com duas normas $\|\cdot\|$ e $\|\cdot\|'$, então elas *induzem a mesma topologia*[†] se, e somente se, existem $r, R > 0$ tais que $r\|x\| \leq \|x\|' \leq R\|x\|$ para todo $x \in E$ (cf. Proposição 1.6.17). Futuramente, isto será usado para provar que, quando o assunto é continuidade e convergência em $\mathbb{R}^n$, podemos considerar qualquer[‡] norma em $\mathbb{R}^n$. ▲

§1 Produtos e subespaços

O hábito de lidar com $\mathbb{R}^2$ como um espaço normado costuma camuflar certas propriedades de sua topologia usual que, quando bem analisadas, permitem estender a noção de *produto* para quaisquer espaços topológicos. Nesse sentido, a norma do máximo $\|\cdot\|_\infty$ é providencial.

Proposição 1.7.20. *Um subconjunto $A \subseteq \mathbb{R}^2$ é aberto na topologia induzida por $\|\cdot\|_\infty$ se, e somente se, para qualquer $(x, y) \in A$ existem intervalos abertos $I, J \subseteq \mathbb{R}$ tais que $x \in I$, $y \in J$ e $I \times J \subseteq A$.*

Demonstração. Se $(x, y) \in A$, então existe uma $\|\cdot\|_\infty$-bola aberta $B \subseteq \mathbb{R}^2$ com $(x, y) \in B$ e $B \subseteq A$ (cf. Exemplo 1.7.13), digamos $B := B_{\|\cdot\|_\infty}((c, d), r)$ para certos $c, d, r \in \mathbb{R}$ com $r > 0$. Ocorre que, por definição, um par (s, t) pertence a B se, e somente se, $\|(c, d) - (s, t)\|_\infty < r$, i.e., se $\max\{|c - s|, |d - t|\} < r$. Logo, a "ida" segue ao se fazer $I := (c - r, c + r)$ e $J := (d - r, d + r)$, pois assim

$$(\alpha, \beta) \in I \times J \Leftrightarrow |c - \alpha| < r \text{ e } |d - \beta| < r,$$

acarretando $(x, y) \in I \times J \subseteq B \subseteq A$ (certo?)*. A recíproca será problema seu ($\substack{\star\\\star}$)[††]. □

Secretamente, esta proposição revela que a norma do máximo em $\mathbb{R}^2$ não é *necessária* para *descrever* a topologia de $\mathbb{R}^2$: basta conhecer os abertos de $\mathbb{R}$ e utilizá-los adequadamente para definir os abertos de $\mathbb{R}^2$. A ideia com espaços topológicos quaisquer é idêntica.

Definição 1.7.21. Para espaços topológicos $(X, \mathcal{T})$ e $(Y, \mathcal{S})$, definimos uma *nova* topologia em $X \times Y$, chamada de **topologia produto**, declarando como abertos os subconjuntos $A \subseteq X \times Y$ com a seguinte propriedade: para todo $(x, y) \in A$, existem $U \in \mathcal{T}$ e $V \in \mathcal{S}$ tais que $(x, y) \in U \times V$ e $U \times V \subseteq A$. ¶

Exercício 1.135 ($\substack{\star\\\star}$)**.** Mostre que isto define uma topologia em $X \times Y$. ■

IMPORTANTE. Se você fez o exercício anterior com a devida atenção, deve ter notado que, embora $U \times V$ seja aberto em $X \times Y$ sempre que $U \subseteq X$ e $V \subseteq Y$ são abertos em seus respectivos espaços, nem todo aberto de $X \times Y$ é um produto de abertos! Se duvida, lembre-se que $B := \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 < 1\}$ é aberto em $\mathbb{R}^2$ na topologia induzida por $\|\cdot\|_2$ (certo?)* e, portanto, é aberto de $\mathbb{R} \times \mathbb{R}$ na topologia produto (cf. Exercício 1.130), mas não existem intervalos abertos $I, J \subseteq \mathbb{R}$ tais que $B = I \times J$ (justique!!)$^{\star}_{\star}$. MORAL DA HISTÓRIA: não trate definições como enfeites.

[†]Duas métricas (resp. normas) sobre um mesmo conjunto X (resp. espaço vetorial) são **topologicamente equivalentes** se ambas induzem a mesma topologia.

[‡]E não apenas as normas $\|\cdot\|_1$, $\|\cdot\|_2$ e $\|\cdot\|_\infty$, como no Exercício 1.130.

[††]Dica: não é um bom momento para esquecer da dica sugerida no item a) do Exercício 1.18.

Quando os espaços X e Y são métricos ou (vetoriais) normados, existem métricas ou normas que induzem precisamente a topologia produto, e você não deve ter dificuldades para encontrar algumas delas: para a surpresa de ninguém, basta adaptar as definições de $\|\cdot\|_1$, $\|\cdot\|_2$ e $\|\cdot\|_\infty$. No entanto, lidar diretamente com a topologia produto costuma ser tanto mais fácil quanto menos ambíguo, haja vista que a topologia produto é única (cf. Exercício 1.145).

Exemplo 1.7.22. *A operação de adição é contínua em qualquer espaço normado.* Neste caso, chamando por E o espaço normado, a adição é a função $+\colon E \times E \to E$ que a cada par de vetores $(u, v) \in E \times E$ associa o vetor $u + v \in E$. Se estivéssemos lidando apenas com funções entre espaços métricos ou normados, seríamos obrigados a decidir ou *convencionar* uma norma em $E \times E$ que permitisse discutir a continuidade de funções do tipo $E \times E \to E$. Todavia, topologicamente só há uma coisa natural a fazer: considerar a topologia produto em $E \times E$. Ocorre que ao fazer isso, o resultado desejado se revela uma simples consequência da combinação entre o Corolário 1.2.32, o Teorema 1.7.14 e o vindouro Exercício 1.146 (concorda?)$^{\star}_{\star}$. ▲

Exercício 1.136 $(^{\star}_{\star})$**.** Mostre que a multiplicação por escalar é contínua em qualquer espaço normado. Dica: desta vez você irá considerar $E \times \mathbb{R}$ com a topologia produto. ■

Observação 1.7.23. Você não *precisa* usar redes para ter uma relação amigável com a topologia produto. Por exemplo, para mostrar que a adição é contínua, pode-se fixar uma bola aberta de raio $2\varepsilon > 0$ em torno de $u + v$ e observar que as bolas abertas $U := B(u, \varepsilon)$ e $V := B(v, \varepsilon)$ são tais que $(u, v) \in U \times V$ e $x + y \in B(u + v, 2\varepsilon)$ sempre que $(x, y) \in U \times V$ (certo?)*. Ou seja: mostrou-se que a adição é contínua por meio da definição clássica em termos de abertos (percebeu *né*?!)*. △

Restrições constituem outra forma ainda mais básica para obter espaços topológicos novos a partir de outros previamente conhecidos. Em geral, se $(X, \mathcal{T})$ é espaço topológico e $Y \subseteq X$ é um subconjunto, a família $\mathcal{T}_Y := \{A \cap Y : A \in \mathcal{T}\}$ é uma topologia em Y, chamada de **topologia de subespaço** (compare com a Definição 1.7.5).

Exercício 1.137 $(^{\star}_{\star})$**.** Para X espaço topológico e $Y \subseteq X$ um subconjunto, considere a família $\tau_Y := \{Y \cap A : A \subseteq X \text{ é aberto em } X\}$.

a) Mostre que τ_Y é uma topologia em Y.

b) Supondo X métrico e $Y \subseteq X$, mostre que τ_Y é a topologia induzida pela métrica ρ do Exercício 1.6. ■

Portanto, secretamente, a Definição 1.7.5 apenas descreveu a topologia que um subconjunto $X \subseteq \mathbb{R}$ herda de $\mathbb{R}$ enquanto subespaço topológico. De modo geral, subconjuntos de espaços topológicos sempre serão considerados com suas topologias de subespaços, a menos de menção contrária.

§2 Supremos e ínfimos revistos (parte II)

Nessa altura do campeonato você provavelmente já percebeu que $[-\infty, +\infty]$ é um espaço topológico de maneira bastante natural†. Consequentemente, em virtude do que se discutiu na subsubseção anterior, segue que $[-\infty, +\infty] \times [-\infty, +\infty]$ também é!

†Declare $A \subseteq [-\infty, +\infty]$ aberto se, e somente se, para todo $a \in A$ existir um intervalo aberto $I \subseteq [-\infty, +\infty]$ tal que $a \in I$ e $I \subseteq A$.

Exercício 1.138 ($\substack{\star \\ \star}$)**.** Mostre que a topologia usual de $\mathbb{R}$ coincide com a topologia de $\mathbb{R}$ como subespaço de $[-\infty, +\infty]$. ■

Entre muitas outras coisas, isso nos permite revisitar as indeterminações com bem mais poder de *expressividade*. O Exemplo 1.6.15 já revelou que as propriedades operatórias dos limites reais são, na verdade, manifestação da continuidade das operações em $\mathbb{R}$ vistas como funções da forma $\mathbb{R} \times \mathbb{R} \to \mathbb{R}$. Assim, podemos dizer que as propriedades operatórias dos limites na reta estendida, secretamente, buscavam estender as operações de $\mathbb{R}$ para $[-\infty, +\infty]$ de forma contínua.

$$\begin{tikzcd}
{[-\infty,+\infty] \times [-\infty,+\infty]} \arrow[r, "*_\infty"] & {[-\infty,+\infty]} \\
\mathbb{R} \times \mathbb{R} \arrow[u] \arrow[r, "*"] & \mathbb{R} \arrow[u]
\end{tikzcd}$$

Figura 1.12: Existe uma operação contínua em $[-\infty, +\infty]$ que estende outra operação contínua $*$ em $\mathbb{R}$?

Nesse sentido, as indeterminações apenas provam que isto não é possível com as operações usuais de $\mathbb{R}$. Por exemplo, com $* := (+)$, se existisse uma função contínua $(+)_\infty \colon [-\infty, +\infty] \times [-\infty, +\infty] \to [-\infty, +\infty]$ tal que $a\ (+)_\infty\ b = a + b$ sempre que $a, b \in \mathbb{R}$, então deveria haver um *valor* em $[-\infty, +\infty]$ bem definido para $+\infty\ (+)_\infty\ -\infty$, digamos $L \in [-\infty, +\infty]$. Logo, por continuidade, sempre que $x_n \to +\infty$ e $y_n \to -\infty$, com $x_n, y_n \in \mathbb{R}$ para todo n, deveria ocorrer $\lim\limits_{n \to \infty} x_n + y_n = L$. No entanto, já vimos que para $x_n = n$ e $y_n = (-1)^n - n$, a sequência $(x_n + y_n)_{n \in \mathbb{N}}$ não converge em $[-\infty, +\infty]$! O problema da multiplicação é análogo.

Apesar disso, é possível estender tais operações a um subespaço do espaço topológico $[-\infty, +\infty] \times [-\infty, +\infty]$. Com isso em mente, não é difícil perceber que o Teorema 1.2.8, na verdade, prova que a adição $(+)$ e a multiplicação $(\cdot)$ em $\mathbb{R}$ admitem *extensões contínuas* $(+)_\infty$ e $(\cdot)_\infty$, respectivamente, sobre os subespaços $\mathcal{A}$ e $\mathcal{M}$ de $[-\infty, +\infty] \times [-\infty, +\infty]$ *ilustrados* a seguir. Explicitamente, basta fazer

- ✓ $+\infty\ (+)_\infty\ r := r\ (+)_\infty\ +\infty = +\infty$ para todo $r \in (-\infty, +\infty]$,
- ✓ $-\infty\ (+)_\infty\ r := r\ (+)_\infty\ -\infty = -\infty$ para todo $r \in [-\infty, +\infty)$,
- ✓ $+\infty\ (\cdot)_\infty\ r := r\ (\cdot)_\infty\ +\infty = +\infty$ para todo $r \in (0, +\infty]$,
- ✓ $+\infty\ (\cdot)_\infty\ r := r\ (\cdot)_\infty\ +\infty = -\infty$ para todo $r \in [-\infty, 0)$,
- ✓ $-\infty\ (\cdot)_\infty\ r := r\ (\cdot)_\infty\ -\infty = -\infty$ para todo $r \in (0, +\infty]$, e
- ✓ $-\infty\ (\cdot)_\infty\ r := r\ (\cdot)_\infty\ -\infty = +\infty$ para todo $r \in [-\infty, 0)$.

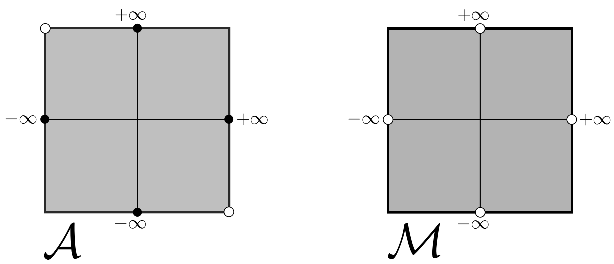


Figura 1.13: Os círculos brancos indicam os pontos de $[-\infty, +\infty] \times [-\infty, +\infty]$ excluídos em $\mathcal{A}$ e $\mathcal{M}$.

Analogamente, a Proposição 1.2.10 apenas garante que a função $\iota\colon \mathbb{R}\setminus\{0\}\to\mathbb{R}\setminus\{0\}$, dada por $\iota(x) := \frac{1}{x}$, se estende a uma função contínua $\Gamma\colon [-\infty,+\infty]\setminus\{0\}\to\mathbb{R}$ em que $\Gamma(\pm\infty)=0$, o que ficará bem menos estranho depois que você se dedicar ao

Exercício 1.139 (?!)**.** Esboce o gráfico da função ι. Sugestão: Geogebra. ■

O que supremos e ínfimos têm a ver com tudo isso? Pouca coisa: o título foi mais uma isca do que algo realmente sério. Ainda assim...

Exercício 1.140 (*For fun* — $\left(\substack{\star\star\\\star\star}\right)$)**.** Sejam $A, B\subseteq\mathbb{R}$ subconjuntos não vazios de $\mathbb{R}$.

a) Mostre que se $(\sup A, \sup B)\in\mathcal{A}$, então $\sup(A+B) = \sup A\ (+)_\infty \sup B$. Sugestão: a correspondência $(a,b)\mapsto a+b$ determina uma função crescente da forma $A\times B\to\mathbb{R}$, onde $A\times B$ pode ser encarado como conjunto dirigido da maneira "óbvia".

b) Repita o item anterior para ínfimos.

c) Use os truques acima para mostrar que $\sup(-A) = -\inf A$.

d) Supondo $A, B\subseteq\{x\in\mathbb{R} : x>0\}$, use o mesmo tipo de raciocínio para mostrar que $\sup(AB) = \sup A\ (\cdot)_\infty \sup B$ e $\inf(AB) = \inf A\cdot\inf B$ sempre que os pares $(\sup A, \sup B)$ e $(\inf A, \inf B)$ pertencem ao conjunto $\mathcal{M}$. ■

§3 Exercícios extras

Exercício 1.141 $\left(\substack{\star\\\star\star}\right)$**.** Para um espaço topológico X, dizemos que $p\in X$ é **ponto de acumulação** de $S\subseteq X$ se $S\cap(V\setminus\{p\})\neq\emptyset$ para todo aberto $V\subseteq X$ tal que $p\in V$.

a) Supondo X métrico, mostre que p é ponto de acumulação de S se, e somente se, existe sequência injetiva $(x_n)_{n\in\mathbb{N}}$ em $S\setminus\{p\}$ tal que $x_n\to p$.

b) Qual das duas implicações acima se verifica num espaço topológico qualquer? ■

Exercício 1.142 $\left(\substack{\star\\\star}\right)$**.** Mostre que todo ponto de um espaço normado E é ponto de acumulação de E. ■

Exercício 1.143 $\left(\substack{\star\\\star}\right)$**.** Mostre que se $X\neq\emptyset$ é espaço topológico, então todo ponto de $X\times\mathbb{Q}$ é ponto de acumulação de $X\times\mathbb{Q}$. ■

Exercício 1.144 $\left(\substack{\star\\\star}\right)$**.** Para espaços topológicos X e Y, mostre que as projeções canônicas $\pi_X\colon X\times Y\to X$ e $\pi_Y\colon X\times Y\to Y$ são contínuas, onde $X\times Y$ tem a topologia produto (cf. Exercício 1.135). Use isso para concluir que $f\colon Z\to X\times Y$ é contínua se, e somente se, $\pi_X\circ f$ e $\pi_Y\circ f$ são contínuas. ■

Exercício 1.145 $\left(\substack{\star\\\star}\right)$**.** Suponha que τ seja uma topologia em $X\times Y$ que torne as projeções π_X e π_Y contínuas. Mostre que todo aberto da topologia produto é também um aberto de τ. Dica: você já percebeu que $A\times B = \pi_X^{-1}[A]\cap\pi_Y^{-1}[B]$ para quaisquer subconjuntos $A\subseteq X$ e $B\subseteq Y$? ■

Exercício 1.146 $\left(\substack{\star\\\star}\right)$**.** Para redes $(x_d)_d$ e $(y_d)_d$ em espaços topológicos X e Y, respectivamente, mostre que $(x_d, y_d)\to(x,y)$ em $X\times Y$ se, e somente se, $x_d\to x$ e $y_d\to y$. Dica: finja que se trata de um problema em $\mathbb{R}^2$ com a norma do máximo. ■

Exercício 1.147 $(\star)$**.** Mostre que o produto interno $\langle\cdot,\cdot\rangle\colon E \times E \to \mathbb{R}$ é contínuo sempre que E é dotado da norma induzida por $\langle\cdot,\cdot\rangle$. Dica: confira o Exercício 1.47. ■

Exercício 1.148 $({}^{\star}_{\star\star})$**.** Mostre que se X é espaço topológico e $\varphi\colon X \to Y$ é bijeção, então existe única topologia em Y tal que φ é homeomorfismo. ■

Exercício 1.149 $({}^{\star}_{\star\star})$**.** Uma **base** (de abertos) para um espaço topológico X é uma família $\mathcal{B}$ de abertos de X tal que para todo $O \subseteq X$ aberto e todo $x \in O$ existe $B \in \mathcal{B}$ tal que $x \in B$ e $B \subseteq O$.

a) Mostre que se $\mathcal{B}$ é base de X, então todo aberto de X é reunião de membros de $\mathcal{B}$.

b) Exiba bases para: a topologia usual de $\mathbb{R}$; a topologia usual de um espaço métrico; a topologia de um produto $X \times Y$, em que X e Y são espaços topológicos quaisquer.

c) Mostre que os intervalos abertos em $[-\infty, +\infty]$ constituem uma base para topologia de $[-\infty, +\infty]$.

d) Generalize o item anterior para ordens totais quaisquer. Compare com as topologias de subespaço para todos os tipos de intervalos de $\mathbb{R}$.

e) Mostre que se X admite uma *base enumerável*[†], então X tem um subconjunto *denso* enumerável. Dica: você precisará saber o que significa dizer que um subconjunto é *denso*; para isso, confira os Exercícios 1.218 e 1.243.

f) Mostre que se X é métrico e tem um denso enumerável, então X tem uma base enumerável. Dica: use raios racionais em torno dos pontos do denso enumerável. ■

1.8 Limites de funções

1.8.0 Essencial

§0 Limites de funções em pontos de acumulação

Vamos finalmente utilizar todo o ferramental desenvolvido ao longo do capítulo para formalizar a noção de limite para funções da forma $X \to \mathbb{R}$, onde X é um subconjunto qualquer de $\mathbb{R}$. Primeiro, a ideia clássica, que está quase certa: para uma função $f\colon X \to \mathbb{R}$ e um ponto $p \in \mathbb{R}$, que pode ou não pertencer a X, dizemos que $L \in \mathbb{R}$ *é um limite de f quando x tende a p* se para todo $\varepsilon > 0$ existe $\delta > 0$ tal que $|f(x) - L| < \varepsilon$ sempre que $0 < |x - p| < \delta$ para $x \in X$. A definição acima *quase* funciona.

Exemplo 1.8.0 (Falha da unicidade de limites)**.** Seja $X \coloneqq (0,1) \cup (3,4)$ e considere a função $f\colon X \to \mathbb{R}$ dada por

$$f(x) \coloneqq \begin{cases} 0, & \text{se } x \in (0,1) \\ 1, & \text{se } x \in (3,4) \end{cases}.$$

PERGUNTA-SE: existe algum $L \in \mathbb{R}$ tal que "$\lim\limits_{x\to 2} f(x) = L$"?

[†]Leia com atenção: dizer que X tem uma base $\mathcal{B}$ enumerável não significa dizer que $\mathcal{B}$ é um aberto enumerável, haja vista que $\mathcal{B}$ é uma **família de abertos**. Se a ideia de coleções de conjuntos ainda lhe parece confusa, volte algumas casas e revise as noções básicas de Teoria dos Conjuntos.

Resposta: sim! Basta tomar $L := 59$. Com efeito, dado qualquer $\varepsilon > 0$, ao fazer $\delta := 1$, a implicação

$$x \in X \quad \text{e} \quad 0 < |x-2| < 1 \Rightarrow |f(x) - 59| < \varepsilon$$

é verdadeira, por um motivo bastante frustrante: não há $x \in X$ tal que $0 < |x-p| < 1$. Ora, mas então por que 59? Pelo mesmo motivo que $L := 200$ funcionaria: para o p escolhido, não é possível garantir a unicidade dos números L's satisfazendo a definição de limite! ▲

A situação do exemplo anterior se deve a uma falha "técnica": o fato de existir um $\delta > 0$ tal que $(2-\delta, 2+\delta) \cap X = \emptyset$ permite que a implicação que caracteriza o limite seja satisfeita por vacuidade. Para impedir isso, portanto, basta pedir que tal δ não exista (cf. Proposição 1.3.20).

Definição 1.8.1. Para $X \subseteq \mathbb{R}$ e $p \in \mathbb{R}$, diremos que p é **ponto de acumulação de** X se para todo $r > 0$ existe $x \in X$ tal que $0 < |x-p| < r$. ¶

Exercício 1.150 ($\star$)**.** A seguir, decida se os pontos $p \in \mathbb{R}$ dados são pontos de acumulação dos subconjuntos S.

a) $S := (0,2)$ e $p := 1$.

b) $S := (0,1]$ e $p := 0$.

c) $S := \mathbb{Z}$ e $p := 200$.

d) $S := \{\sqrt{2}\}$ e $p := \sqrt{2}$.

e) $S := \mathbb{Q}$ e $p := \sqrt{12039810293} - 1$.

f) $S := \left\{\frac{1}{2^n} : n \in \mathbb{N}\right\}$ e $p := 0$. ■

Exercício 1.151 ($\overset{\star}{\star}$)**.** Mostre que se $p \in \mathbb{R}$ é ponto de acumulação de X e $L, L' \in \mathbb{R}$ são limites† de f quando x tende a p, então $L = L'$. Dica: imite o Exemplo 1.0.7. ■

Definição 1.8.2. Como de costume, nas situações em que p é ponto de acumulação de X e $f\colon X \to \mathbb{R}$ é uma função, vamos escrever $\lim\limits_{x\to p} f(x) = L$ para indicar que $L \in \mathbb{R}$ é o único número real L que é limite de f quando x tende a p. ¶

Observação 1.8.3 (É só mais um limite de redes, mas com ressalvas)**.** Para $X \subseteq \mathbb{R}$ fixado, considere $X_p := X \setminus \{p\} \neq \emptyset$. Ao declarar $a \preceq b$ se $a, b \in X_p$ e $|b-p| \leq |a-p|$, elevamos $(X_p, \preceq)$ ao patamar de conjunto dirigido, o que permite (mais uma vez) considerar a rede $(f(x))_{x\in X_p}$. Agora, como no Exemplo 1.0.7, para $L \in \mathbb{R}$,

$$\lim_{x\in X_p} f(x) = L \Leftrightarrow \lim_{x\to p} f(x) = L$$

sempre que p é ponto de acumulação de X.

A sutileza está nas situações em que p *não é ponto de acumulação de* X. Em tais cenários, embora não faça sentido tratar de "$\lim\limits_{x\to p} f(x)$" no sentido clássico, o limite da rede $(f(x))_{x\in X_p}$ ainda *pode* fazer sentido, já que X_p é dirigido independentemente da situação do ponto p (verifique!)$\overset{\star}{\star}$. De qualquer forma, vamos manter a exigência de p ser ponto de acumulação, entre outras coisas, por respeito aos cânones. △

†Com respeito à definição apresentada no começo desta seção!

A observação acima coloca à disposição todos os resultados obtidos anteriormente para redes reais. Daí, como já se estendeu a noção de limite de redes reais, faz sentido considerar os casos em que $\lim\limits_{x\to p} f(x) = L$ com $L \in [-\infty, +\infty]$. Mas faremos mais:

Exercício 1.152 $(^{\star}_{\star})$**.** Para $X \subseteq \mathbb{R}$ e $p \in \mathbb{R}$, mostre que p é ponto de acumulação de X se, e somente se, para todo intervalo aberto $I \subseteq \mathbb{R}$ tal que $p \in I$ existe $x \in X \cap (I \setminus \{x\})$. ■

Definição 1.8.4. Diremos que $p \in [-\infty, +\infty]$ é **ponto de acumulação de** $X \subseteq \mathbb{R}$ se para todo intervalo aberto $I \subseteq [-\infty, +\infty]$ com $p \in I$ existir $x \in X$ com $x \in I$ e $x \neq p$. ¶

Por conta do exercício anterior, a definição acima estende a noção inicial de ponto de acumulação para números reais, ao mesmo tempo em que permite tratar $-\infty$ e $+\infty$ como pontos de acumulação de certos subconjuntos da reta: note que para um subconjunto não vazio $X \subseteq \mathbb{R}$,

✓ X é ilimitado superiormente se, e somente se, $+\infty$ é ponto de acumulação de X, e

✓ X é ilimitado inferiormente se, e somente se, $-\infty$ é ponto de acumulação de X.

Definição 1.8.5. Para uma função $f : X \to \mathbb{R}$, onde $X \subseteq \mathbb{R}$ e $p \in [-\infty, +\infty]$ é ponto de acumulação de X, dizemos que $L \in [-\infty, +\infty]$ é o *limite de f quando x tende a p*, o que se indica com $L = \lim\limits_{x\to p} f(x)$, se para todo intervalo aberto $I \subseteq [-\infty, +\infty]$ com $L \in I$ existe intervalo aberto $J \subseteq [-\infty, +\infty]$ tal que $p \in J$ e $f(x) \in I$ sempre que $x \in J \cap X \setminus \{p\}$. ¶

Chegamos à verdade fundamental sobre a noção usual dos limites de funções: trata-se apenas de um modo de discutir continuidade sob um disfarce não topológico.

Teorema 1.8.6 (Opcional: sempre foi continuidade)**.** *Sejam $X \subseteq \mathbb{R}$ um subconjunto e $p, L \in [-\infty, +\infty]$, com p ponto de acumulação de X. Para uma função $f : X \to \mathbb{R}$, seja $h : X \cup \{p\} \to [-\infty, +\infty]$ a função dada pela seguinte regra:*

$$h(x) := \begin{cases} f(x), & \text{se } x \neq p \\ L, & \text{se } x = p \end{cases}.$$

Sob tais condições, $\lim\limits_{x\to p} f(x) = L$ e, e somente se, a função h é contínua em p.

Evidentemente, o enunciado usa a noção de continuidade topológica discutida no Teorema 1.7.14, juntamente com a descrição da topologia de $[-\infty, +\infty]$ (cf. nota de rodapé †, na página 184). Mas, ao reler todos os pré-requisitos com cuidado, o enunciado se torna um exercício quase simples de tradução (cf. Exercício 1.196).

Exercício 1.153 $(^{\star}_{\star})$**.** Demonstre o teorema anterior supondo $p, L \in \mathbb{R}$. ■

Exercício 1.154 $(^{\star}_{\star\star})$**.** Demonstre o teorema anterior supondo $p, L \in [-\infty, +\infty]$. ■

Corolário 1.8.7 (do Teorema 1.6.3)**.** *Nas condições do teorema anterior, são equivalentes:*

(i) $\lim\limits_{x\to p} f(x) = L$;

(ii) $\lim\limits_{n\to\infty} f(x_n) = L$ *para toda sequência* $(x_n)_{n\in\mathbb{N}}$ *em* $X \setminus \{p\}$ *tal que* $x_n \to p$.

Em particular, f é contínua em $p \in X$ se, e somente se, $\lim\limits_{x\to p} f(x) = f(p)$.

Os **limites laterais** se obtêm então de modo natural via restrição (cf. Subseção 1.1.1 §0). Para $f\colon X \to \mathbb{R}$ e $p \in [-\infty, +\infty]$ ponto de acumulação de X, podemos considerar os subconjuntos $X_{<p} := \{x \in X : x < p\}$ e $X_{>p} := \{x \in X : x > p\}$.

Exercício 1.155 ($\star\atop\star$)**.** Mostre que $p \in [-\infty, +\infty]$ é ponto de acumulação de X se, e somente se, p é ponto de acumulação de $X_{<p}$ ou p é ponto de acumulação de $X_{>p}$. ■

Assim, quando p é ponto de acumulação de $X_{<p}$ (caso em que se diz que p é **ponto de acumulação de X pela esquerda**), faz sentido considerar o limite da função $f|_{X_{<p}}$ no ponto p, denotado por $\lim\limits_{x\to p^-} f(x)$ quando existe. Analogamente, quando p é ponto de acumulação de $X_{>p}$ (caso em que se diz que p é **ponto de acumulação de X pela direita**), faz sentido considerar o limite da função $f|_{X_{>p}}$ no ponto p, denotado por $\lim\limits_{x\to p^+} f(x)$ quando existe. Nas frequentes situações em que p é ponto de acumulação tanto de $X_{<p}$ quanto de $X_{>p}$, ambos os limites $\lim\limits_{x\to p^-} f(x)$ e $\lim\limits_{x\to p^+} f(x)$ podem ser investigados. Neste caso, verifica-se

$$\lim_{x\to p} f(x) = L \text{ se, e somente se, } \lim_{x\to p^-} f(x) = \lim_{x\to p^+} f(x) = L.$$

Exercício 1.156 ($\star\atop\star$)**.** Prove a equivalência anterior. Dica: confira a Proposição 1.3.30.■

Observação 1.8.8. Note que ao fazer $p := -\infty$, por exemplo, p é ponto de acumulação de X se, e somente se, p é ponto de acumulação de X pela direita, mostrando que os limites da forma "$\lim\limits_{x\to -\infty}$" são limites pela direita. Observação análoga vale para $p := +\infty$.

ATENÇÃO. A ordem que dirige $X_{<+\infty}$ é a ordem usual de $\mathbb{R}$ restrita a X (quanto maior, melhor!), enquanto a ordem que dirige $X_{>-\infty}$ é a ordem inversa de $\mathbb{R}$ restrita a X (quanto menor, melhor!). △

§1 Mudança de variáveis

Dadas funções $f\colon X \to \mathbb{R}$ e $g\colon Y \to \mathbb{R}$, com $X, Y \subseteq \mathbb{R}$ e $\operatorname{im}(f) \subseteq Y$, pode-se considerar a composição $g \circ f\colon X \to \mathbb{R}$: lembre-se de que, por definição, $(g \circ f)(x) := g(f(x))$ para cada $x \in X$. Neste caso, se $p \in [-\infty, +\infty]$ é ponto de acumulação de X, faz sentido investigar a existência de $\lim\limits_{x\to p}(g \circ f)(x)$. Até aí, nada de novo...

Porém, *se* existirem $\lim\limits_{x\to p} f(x) := L \in [-\infty, +\infty]$ e $\lim\limits_{y\to L} g(y) := L' \in [-\infty, +\infty]$, a intuição sugere que deveria valer algo como $\lim\limits_{x\to p} g(f(x)) = L'$. Exceto por situações "patologicamente" problemáticas, é exatamente o que acontece.

Exemplo 1.8.9 (Casos patológicos)**.** Com $f\colon \mathbb{R} \to \mathbb{R}$ a função nula e $g\colon \mathbb{R} \to \mathbb{R}$ a função que faz $g(y) := 1$ para $y \neq 0$ e $g(0) := 0$, temos $\lim\limits_{x\to 0} f(x) = 0$, $\lim\limits_{y\to 0} g(y) = 1$, mas $\lim\limits_{x\to 0} g(f(x)) = \lim\limits_{x\to 0} g(0) = 0$. No entanto, note que a função g foi construída de modo *artificial* a fim de gerar a falha: se g fosse contínua em 0, teríamos $g(0) = 1$ e daí $\lim g(f(x)) = g(\lim f(x))$. ▲

Teorema 1.8.10 (Mudança de variáveis — o caso que importa)**.** *Para $L \in [-\infty, +\infty]$, seja $g\colon Y \cup \{L\} \to [-\infty, +\infty]$ função contínua em L. Se $p \in [-\infty, +\infty]$ é ponto de acumulação de $X \subseteq \mathbb{R}$ e $f\colon X \to Y$ é tal que $\lim\limits_{x\to p} f(x) = L$, então $\lim\limits_{x\to p} g(f(x)) = g(L)$.*

Demonstração. As hipóteses asseguram que $(f(x))_{x\in X_p}$ é uma rede em Y que converge para L. Logo, por continuidade, $(g(f(x)))_{x\in X_p}$ converge para $g(L)$, i.e., $\lim\limits_{x\to p} g(f(x)) = g(L)$. $\square$

Observação 1.8.11 (Pedantismo opcional)**.** A rigor, a função g não precisaria estar definida em L, mas daí precisa-se de mais cuidado.

Proposição 1.8.12 (Mudança pedante de variáveis)**.** *Sejam $f\colon X \to \mathbb{R}$ e $g\colon Y \to \mathbb{R}$ funções, com $X, Y \subseteq \mathbb{R}$ e $\operatorname{im}(f) \subseteq Y$. Se $p, L, L' \in [-\infty, +\infty]$ são tais que*

(i) p é ponto de acumulação de X e $\lim\limits_{x\to p} f(x) = L$,

(ii) L é ponto de acumulação de Y e $\lim\limits_{y\to L} g(y) = L'$, e

(iii) existe um intervalo aberto $I \subseteq [-\infty, +\infty]$ com $p \in I$ e $f(x) \neq L$ para todo ponto $x \in (X \cap I) \setminus \{p\}$,

então $\lim\limits_{x\to p} g(f(x)) = L'$.

Demonstração. Basta trocar g pela função $h\colon Y \cup \{L\} \to [-\infty, +\infty]$ que faz $h(y) := g(y)$ para $y \neq L$ e $h(L) := L'$, que será contínua em L, e daí aplicar o teorema anterior, observando que $g(f(x)) = h(f(x))$ para todo $x \in X_p \cap I$ (cf. Exercício 1.195). $\square$

Note que no exemplo anterior, foi justamente a falha da condição (iii) que ocasionou a falha no resultado esperado. $\triangle$

Um modo alternativo de verificar $\lim\limits_{x\to+\infty} \frac{1}{x+1} = 0$ é observar que as *funções* $g(y) := \frac{1}{y}$ e $f(x) := x + 1$ são tais que $\lim\limits_{x\to+\infty} f(x) = +\infty$ e $\lim\limits_{y\to+\infty} g(y) = 0$, com a condição (iii) na Proposição 1.8.12 satisfeita automaticamente[†], donde segue $\lim\limits_{x\to+\infty} \frac{1}{x+1} = \lim\limits_{x\to+\infty} g(f(x)) = 0$.

A terminologia "mudança de variáveis" se deve ao seguinte: a princípio, a função $x \mapsto \frac{1}{x+1}$ associa a *variável* x ao valor $\frac{1}{x+1}$; porém, ao considerar a função que associa a variável $y := x + 1$ ao valor $\frac{1}{y}$, percebe-se que a correspondência original é a composição das funções $x \mapsto y := x + 1$ e $y \mapsto \frac{1}{y}$.

§2 Derivadas

Para encerrar o *essencial* deste capítulo, vamos introduzir um tipo de limite de função fundamental para a Análise.

Definição 1.8.13. Para um subconjunto $X \subseteq \mathbb{R}$, uma função $f\colon X \to \mathbb{R}$ e $a \in X$ um ponto de acumulação de X, a **derivada de** f em a é o limite

$$f'(a) := \lim_{x\to a} \frac{f(x) - f(a)}{x - a} \tag{1.6}$$

se este existir — possivelmente em $[-\infty, +\infty]$. Se, adicionalmente, $f'(a)$ for um número real, diremos que f é **diferenciável em** a. Uma função **diferenciável** é aquela que é diferenciável em todos os pontos de seu domínio. ¶

[†]Uma vez que $f(x) \in \mathbb{R}$ e $L \notin \mathbb{R}$. Alternativamente, você pode observar que a função g utilizada é contínua em $(0, +\infty) \cup \{+\infty\}$, e daí aplicar o teorema anterior.

Alternativamente, pode-se definir

$$f'(a) = \lim_{t \to 0} \frac{f(a+t) - f(a)}{t}, \tag{1.7}$$

essencialmente por mudança de variáveis. Primeiro, observe que $a \in X$ se, e somente se, $0 \in Z \coloneqq \{x - a : x \in X\}$, bem como a é ponto de acumulação de X se, e somente se, 0 é ponto de acumulação de Z.

Exercício 1.157 ($\star\star$)**.** Observe! ■

Agora, note que a função $\varphi \colon Z \to \mathbb{R}$ dada por $\varphi(t) \coloneqq a + t$ é tal que $\lim\limits_{t \to 0} \varphi(t) = a$, com $\operatorname{im}(\varphi) = X$. Por sua vez, a função $\psi \colon X \to [-\infty, +\infty]$ dada por

$$\psi(x) \coloneqq \begin{cases} \dfrac{f(x) - f(a)}{x - a}, & \text{se } x \neq a \\ f'(a), & \text{caso contrário} \end{cases}$$

é tal que $\lim\limits_{x \to a} \psi(x) = f'(a)$. Logo,

$$f'(a) = \lim_{t \to 0} \psi(\varphi(t)) = \lim_{t \to 0} \frac{f(\varphi(t)) - f(a)}{\varphi(t) - a} = \lim_{t \to 0} \frac{f(a+t) - f(a)}{t},$$

como queríamos[†]. A recíproca é análoga (faça!)**.

Observação 1.8.14. O conjunto Z precisou ser explicitado pois a função f só sabe *operar* em X, que é seu domínio por hipótese. Todavia, no dia a dia é relativamente seguro não ter tanto cuidado. △

Intuitivamente, o quociente

$$\frac{f(x) - f(a)}{x - a}$$

expressa o *coeficiente angular* da reta determinada pelos pontos $(x, f(x))$ e $(a, f(a))$, de modo que ao tomar x cada vez mais próximo de a, chega-se cada vez mais perto do coeficiente angular de uma reta *tangente* ao gráfico de f no ponto $(a, f(a))$. Nas situações em que a função f é contínua em a, o limite que define $f'(a)$ é uma *indeterminação* da forma $0/0$, pois $\lim\limits_{x \to a} f(x) - f(a) = 0$ e $\lim\limits_{x \to a} x - a = 0$. Assim, você tem menos um motivo para achar que indeterminações não são importantes.

Proposição 1.8.15. *Se f é diferenciável em $a \in X$, então f é contínua em a.*

Demonstração. Em virtude do Corolário 1.8.7, basta mostrar que $\lim\limits_{x \to a} f(x) = f(a)$ ou, equivalentemente, $\lim\limits_{x \to a} f(x) - f(a) = 0$. Ocorre que para $x \neq a$ no domínio de f, tem-se

$$f(x) - f(a) = (f(x) - f(a)) \cdot \frac{x - a}{x - a} = \frac{f(x) - f(a)}{x - a} \cdot (x - a),$$

com $f'(a) = \lim\limits_{x \to a} \frac{f(x) - f(a)}{x - a}$ e $\lim\limits_{x \to a} x - a = 0$, donde segue que

$$\lim_{x \to a} f(x) - f(a) = \lim_{x \to a} \frac{f(x) - f(a)}{x - a} \cdot (x - a) = f'(a) \cdot 0 = 0. \qquad \square$$

[†]Em resumo: ao fazer $x \coloneqq a + t$ na expressão que define (1.7), somos "forçados" a substituir t por $x - a$ (já que $x = a + t$ se, e somente se, $t = x - a$), e daí se chega (1.6), vice-versa.

Exercício 1.158 ($\star$)**.** Mostre que se $f\colon X \to \mathbb{R}$ é uma função constante, então $f'(a) = 0$ para todo $a \in X$. ■

Exercício 1.159 ($^{\star}_{\star}$)**.** Mostre que se a é ponto de acumulação de $X \subseteq \mathbb{R}$ e $f\colon X \to \mathbb{R}$ é a função inclusão, i.e., $f(x) \coloneqq x$ para todo $x \in X$, então $f'(a) = 1$. ■

Exemplo 1.8.16. Com $f\colon \mathbb{R} \to \mathbb{R}$ dada por $f(x) \coloneqq x^2$ e $a \coloneqq 2$,

$$f'(2) \coloneqq \lim_{t\to 0} \frac{(2+t)^2 - 4}{t} = \lim_{t\to 0} \frac{t^2 + 4t}{t} = \lim_{t\to 0} 4 + t = 4.$$

Em outras palavras, o coeficiente angular da reta tangente ao gráfico de f no ponto $(2, f(2))$ é 4. Portanto, a reta tangente é descrita pela função $r(t) \coloneqq 4(t-2)+4$. ▲

Exemplo 1.8.17 (Clássico)**.** A função $|\cdot|\colon \mathbb{R} \to \mathbb{R}$ não é derivável em 0, pois

$$|0|' \coloneqq \lim_{t\to 0} \frac{|t| + 0}{t} = \lim_{t\to 0} \frac{|t|}{t}$$

não existe (verifique os limites laterais!)$^{\star}_{\star}$. ▲

Exemplo 1.8.18. Para $n \in \mathbb{N}$ com $n \geq 1$, seja $f\colon \mathbb{R} \to \mathbb{R}$ dada por $f(x) \coloneqq x^n g(x)$, em que $g\colon \mathbb{R} \to \mathbb{R}$ é uma função limitada tal que não exista $\lim\limits_{x\to 0} g(x)$; se você já souber o que é, pode fazer $g(x) \coloneqq \operatorname{sen}\left(\frac{1}{x}\right)$ para $x \neq 0$ e $g(0) \coloneqq 0$. Como $\lim\limits_{x\to 0} x^n = 0$, segue por *confronto* que

$$f'(0) \coloneqq \lim_{t\to 0} \frac{f(t) - f(0)}{t} = \lim_{t\to 0} \frac{t^n g(t)}{t} = \lim_{t\to 0} t^{n-1} g(t) = 0$$

desde que se tenha $n > 1$. ▲

Embora sejam definidas como indeterminações, derivadas apresentam comportamentos bastante previsíveis diante de certas operações algébricas. Quem não dormiu nas aulas de Cálculo deve se lembrar da próxima

Proposição 1.8.19 (Propriedades operatórias)**.** *Para $f, g\colon X \to \mathbb{R}$ diferenciáveis em $a \in X$, as funções $f + g$ e $f \cdot g$ são diferenciáveis em a, com*

(i) $(f+g)'(a) = f'(a) + g'(a)$ *e*

(ii) $(f \cdot g)'(a) = f(a)g'(a) + g(a)f'(a)$ (a.k.a. ***regra de Leibniz***)*.*

Além disso, se $g(a) \neq 0$, então $\frac{f}{g}$ é diferenciável em a e

$$\left(\frac{f}{g}\right)'(a) = \frac{f'(a)g(a) - g'(a)f(a)}{(g(a))^2}. \tag{1.8}$$

Demonstração. A primeira identidade segue diretamente da continuidade da soma. Para a segunda, note que com $x \neq a$ tal que $x \in X$,

$$\begin{aligned}\frac{f(x)g(x) - f(a)g(a)}{x-a} &= \frac{f(x)g(x) + (f(x)g(a) - f(x)g(a)) - f(a)g(a)}{x-a} \\ &= f(x) \cdot \frac{g(x) - g(a)}{x-a} + g(a) \cdot \frac{f(x) - f(a)}{x-a},\end{aligned}$$

donde a continuidade de f em a (proposição anterior) acarreta

$$f'(a) = \lim_{x \to a} f(x) \cdot g'(a) + g(a) \cdot f'(a) = f(a)g'(a) + g(a)f'(a).$$

Para o quociente, com $t \neq 0$ tal que $a + t \in X$ (que existe pois...?)$^{\star}_{\star}$, tem-se

$$\frac{\dfrac{1}{g(a+t)} - \dfrac{1}{g(a)}}{t} = \frac{g(a) - g(a+t)}{g(a+t)g(a)} \cdot \frac{1}{t} = -\frac{g(a+t) - g(a)}{t} \cdot \frac{1}{g(a+t)g(a)},$$

donde segue que

$$\left(\frac{1}{g}\right)'(a) := \lim_{t \to 0} \frac{\dfrac{1}{g(a+t)} - \dfrac{1}{g(a)}}{t} = -\frac{g'(a)}{(g(a))^2},$$

uma vez que g é contínua em a (por ser diferenciável). O caso geral segue do item (ii). □

Exercício 1.160 ($^{\star}_{\star}$)**.** Para $f\colon \mathbb{R} \to \mathbb{R}$ dada por $f(x) := x^n$, mostre que $f'(a) = na^{n-1}$ para todo $a \in \mathbb{R}$. Conclua que funções polinomiais são diferenciáveis. Dica: indução + Regra de Leibniz. ■

Exercício 1.161 ($^{\star}_{\star}$)**.** Para f como no exercício anterior, determine $\left(\frac{1}{f}\right)'(a)$ para cada $a \neq 0$. Conclua que para $z \in \mathbb{Z}$ fixado, a função que faz $x \mapsto x^z$ é diferenciável em todo ponto $\neq 0$, e sua derivada é a função que faz $x \mapsto zx^z$. Dica: lembre-se de que para $n > 0$, $x^{-n} := \frac{1}{x^n}$. ■

Exemplo 1.8.20. A função $\sqrt{\ }\colon [0, +\infty) \to [0, +\infty)$ é diferenciável em todo ponto $x > 0$. De fato, observe que para $x, t > 0$,

$$\frac{\sqrt{x+t} - \sqrt{x}}{t} = \frac{1}{t} \cdot \frac{\sqrt{x+t} - \sqrt{x}}{\sqrt{x+t} + \sqrt{x}} \cdot \left(\sqrt{x+t} + \sqrt{x}\right) = \frac{1}{\sqrt{x+t} + \sqrt{x}},$$

e, como $\sqrt{x+t} + \sqrt{x} \to 2\sqrt{x} \neq 0$ quando $t \to 0$, resulta $\sqrt{x}' = \frac{1}{2\sqrt{x}}$.

Note que ao escrever $x^{\frac{1}{2}} := \sqrt{x}$, a expressão acima condiz com a fórmula do exercício anterior: $(x^{\frac{1}{2}})' = \frac{1}{2} \cdot (x^{\frac{1}{2}-1}) = \frac{1}{2\sqrt{x}}$. ▲

Observação 1.8.21 (Derivadas laterais)**.** No exemplo anterior, para $x := 0$, temos

$$\lim_{t \to 0^+} \frac{\sqrt{t}}{t} = \lim_{t \to 0^+} \frac{1}{t}\sqrt{t} \cdot \frac{\sqrt{t}}{\sqrt{t}} = \lim_{t \to 0^+} \frac{1}{\sqrt{t}} = +\infty,$$

posto que $\sqrt{t} \geq 0$ para todo $t \geq 0$ e $\sqrt{\ }$ é contínua em 0, com $\sqrt{0} = 0$ (mudança de variáveis!). Geometricamente, a reta tangente ao gráfico de $\sqrt{\ }$ no ponto 0 é vertical (sugestão: faça um desenho!). Secretamente, esse tipo de situação lida com a noção de *derivada lateral*, que essencialmente consiste em tratar das versões laterais (i.e., à direita e à esquerda) do limite que define a derivada. △

Nas situações em que $f\colon X \to \mathbb{R}$ é *diferenciável*, a derivada $f'(a)$ existe para todo $a \in X$, o que induz uma *nova* função, agora indicada por $f'\colon X \to \mathbb{R}$, que a cada $a \in X$ associa $f'(a)$. Tal função é, muito apropriadamente, xingada de **derivada da função** f. Por ser uma função da forma $X \to \mathbb{R}$, é lícito investigar a derivada de f', denotada por f'' e chamada de **segunda derivada de** f, que se existir permite a procura de sua derivada, f''' (*a.k.a.* **terceira derivada de** f), e assim sucessivamente, *ad nauseam*.

Exemplo 1.8.22. Pelos exercícios anteriores, funções polinomiais são *infinitamente diferenciáveis*, i.e., admitem todas as derivadas iteradas. Usualmente, funções com tal propriedade são xingadas de **suaves**. ▲

Exemplo 1.8.23. De volta ao Exemplo 1.8.18, note que ao assumir g diferenciável em $\mathbb{R} \setminus \{0\}$, a função $f(x) := x^2 g(x)$ é tal que $f'(a) = 2ag(a) + a^2 g'(a)$ se $a \neq 0$, enquanto $f'(a) = 0$. Logo,

$$f''(0) := \lim_{t \to 0} \frac{f'(t) - f'(0)}{t} = \lim_{t \to 0} 2g(t) + t^2 g'(t),$$

que pode não existir a depender da função g escolhida. Com $g(x) := \operatorname{sen}\left(\frac{1}{x}\right)$, por exemplo, veremos que $g'(t) = -\frac{1}{t^2} \cdot \cos\frac{1}{t}$ para $t \neq 0$, situação em que o limite acima não existe: fazendo

$$t_n := \frac{1}{2\pi n + \frac{\pi}{2}} \quad \text{e} \quad s_n := \frac{1}{2\pi(n+1)},$$

tem-se $t_n \to 0$ e $s_n \to 0$, mas $2g(t_n) + t_n^2 g'(t_n) = 2$ e $2g(s_n) + t^2 g'(s_n) = -1$ para todo $n \in \mathbb{N}$ (percebeu?)*. ▲

Para encerrar este primeiro contato com derivadas[†], vamos relembrar

Teorema 1.8.24 (Regra da cadeia)**.** *Sejam X e Y subconjuntos de $\mathbb{R}$, $f\colon X \to \mathbb{R}$ e $g\colon Y \to \mathbb{R}$ funções tais que $f[X] \subseteq Y$. Se f é diferenciável em $a \in X$ e g é diferenciável em $f(a) \in Y$, então $g \circ f$ é diferenciável em a, com $(g \circ f)'(a) = g'(f(a)) \cdot f'(a)$.*

Demonstração. A primeira ideia que vem em mente, possivelmente, consiste em afirmar que para $x \neq a$ seja lícito escrever

$$\frac{g(f(x)) - g(f(a))}{x - a} = \frac{g(f(x)) - g(f(a))}{x - a} \cdot \frac{f(x) - f(a)}{f(x) - f(a)} = \frac{g(f(x)) - g(f(a))}{f(x) - f(a)} \cdot \frac{f(x) - f(a)}{x - a},$$

de modo que o argumento se encerra ao fazer "$x \to a$". Contudo, não há razões para supor $f(x) \neq f(a)$ para x *suficientemente próximo* de a. Isto se remedia com um truque sujo.

Das hipóteses de que $f'(a)$ e $g'(f(a))$ existem, faz sentido definir as funções auxiliares $\varphi\colon X \to \mathbb{R}$ e $\psi\colon Y \to \mathbb{R}$ dadas por

$$\varphi(x) := \frac{f(x) - f(a)}{x - a} - f'(a) \qquad \psi(y) := \frac{g(y) - g(f(a))}{y - f(a)} - g'(f(a))$$

sempre que $x \neq a$ e $y \neq f(a)$, respectivamente, com $\varphi(a) := \psi(f(a)) := 0$. Adiante, será importante notar que φ e ψ são contínuas em a e $f(a)$, respectivamente: note! $\binom{\star}{\star}$. Agora, substituindo y por $f(x)$ na última expressão, segue que

$$\begin{aligned} g(f(x)) - g(f(a)) &= (\psi(f(x)) + g'(f(a))) \cdot (f(x) - f(a)) \\ &= (\psi(f(x)) + g'(f(a))) \cdot (\varphi(x) + f'(a)) \cdot (x - a), \end{aligned} \tag{1.9}$$

acarretando

$$\lim_{x \to a} \frac{g(f(x)) - g(f(a))}{x - a} = \lim_{x \to a} (\psi(f(x)) + g'(f(a))) \cdot \lim_{x \to a} (\varphi(x) + f'(a)) = g'(f(a)) \cdot f'(a),$$

pois ψ e φ são contínuas em $f(a)$ e a, respectivamente[‡]. □

[†] Que continuará pelo restante deste material.

[‡] A continuidade de f em a também é importante, pois ao fazer "$x \to a$", resulta "$f(x) \to f(a)$" e daí "$\psi(f(x)) \to \psi(f(a)) = 0$".

~~Súplica~~ Sugestão. Se você for uma pessoa de mente aberta, considere dar uma chance para a caracterização de Carathéodory para derivadas (cf. Subseção 1.8.1 §0), que permite a elaboração de demonstrações bem mais diretas.

Exemplo 1.8.25 (Cuidado para não se confundir com o Exemplo 1.8.20)**.** Para $x, t > 0$, não é difícil se convencer de que

$$\frac{\frac{1}{\sqrt{x+t}} - \frac{1}{\sqrt{x}}}{t} = -\frac{1}{\sqrt{x+t}\sqrt{x} \cdot (\sqrt{x+t} + \sqrt{x})},$$

o que permite mostrar que

$$\left(\frac{1}{\sqrt{x}}\right)' = -\frac{1}{2x\sqrt{x}}.$$

Ocorre que a função $x \mapsto \frac{1}{\sqrt{x}}$ é a composição das funções $f\colon x \mapsto \sqrt{x}$ e $g\colon y \to \frac{1}{y}$. Logo, pela regra da cadeia,

$$(g \circ f)'(x) = g'(f(x)) \cdot f'(x) = -\frac{1}{\sqrt{x}^2} \cdot \frac{1}{2\sqrt{x}} = -\frac{1}{2x\sqrt{x}},$$

exatamente o resultado esperado. ▲

Exemplo 1.8.26. Para g diferenciável em a com $g(a) \neq 0$, a Regra da Cadeia garante que

$$\left(\frac{1}{g}\right)'(a) = -\frac{g'(a)}{(g(a))^2},$$

justamente a expressão obtida na demonstração da identidade (1.8). ▲

Exemplo 1.8.27 (L'Hôpital)**.** Uma aplicação inusitada das derivadas facilita a <u>estimativa</u> de algumas indeterminações. Supondo $f, g\colon X \to \mathbb{R}$ funções contínuas em a e satisfazendo $f(a) = g(a) = 0$, já vimos que $\lim_{x\to a} \frac{f(x)}{g(x)}$ tem *status* indeterminado. Porém, se f e g forem diferenciáveis em a com $g'(a) \neq 0$, então

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \frac{f'(a)}{g'(a)}.$$

Com efeito, neste caso,

$$f'(a) \cdot \frac{1}{g'(a)} = \lim_{x\to a} \frac{f(x) - f(a)}{x - a} \lim_{x\to a} \frac{x - a}{g(x) - g(a)} = \lim_{x\to a} \frac{f(x)}{g(x)}.$$

Uma vez munidos de ferramentas topológicas mais sofisticadas, poderemos estender o resultado acima para uma gama bem mais ampla de situações: em geral, para pontos $a, L \in [-\infty, +\infty]$, com $I \subseteq [-\infty, +\infty]$ um intervalo aberto tal que $a \in I$ e $f, g\colon I \setminus \{a\} \to \mathbb{R}$ diferenciáveis, com $g'(x) \neq 0$ para todo $x \in I \setminus \{a\}$ e $\lim_{x\to a} \frac{f'(x)}{g'(x)} = L$, tem-se

$$\lim_{x\to a} \frac{f(x)}{g(x)} = L$$

sempre que $\lim_{x\to a} f(x) = \lim_{x\to a} g(x) = 0$ ou $\lim_{x\to a} |g(x)| = +\infty$. A prova, contudo, requer mais ferramentas do que as disponíveis: faltam em nosso arsenal os importantes TVM e TVI, que serão explorados no próximo capítulo. ▲

1.8.1 Extras

§0 Importante: derivadas à moda Carathéodory

Enquanto limites podem ser tratados, de modo geral, em espaços topológicos dos mais abstratos, derivadas requerem algum tipo de aparato algébrico para serem definidas: a fim de dar sentido a algo como

$$\lim_{t\to 0}\frac{f(a+t)-f(a)}{t},$$

precisa-se de uma topologia ("$\lim\limits_{t\to 0}$"), de uma adição e uma multiplicação no domínio da função ("$a+t$" e "$\cdot\frac{1}{t}$"), bem como de uma adição no contradomínio da função ("$f(a+t)-f(a)$"). Além disso, existe o problema de que se o domínio e o contradomínio de f não forem subconjuntos do mesmo espaço ambiente, todas essas operações podem ser *incompatíveis* umas com as outras.

Exemplo 1.8.28. Para uma função $f\colon \mathbb{R}^3\to\mathbb{R}^2$, note que não faz sentido escrever

$$\frac{f(a+t)-f(a)}{t},$$

pois $a, t, a+t\in\mathbb{R}^3$, $f(a+t)-f(a)\in\mathbb{R}^2$ e tampouco há uma multiplicação definida em $\mathbb{R}^3$ que permita dar sentido a algo como "$\frac{1}{t}$" para $t\neq(0,0,0)$. Mesmo que tudo isso fosse resolvido, restaria o problema de multiplicar um elemento de $\mathbb{R}^3$ por outro de $\mathbb{R}^2$. ▲

Um modo bastante esperto de contornar o problema se deve a Carathéodory:

Proposição 1.8.29. *Para um subconjunto $X\subseteq\mathbb{R}$, uma função $f\colon X\to\mathbb{R}$ e $a\in X$ um ponto de acumulação de X, são equivalentes:*

(i) *f é diferenciável em a;*

(ii) *existe $L\colon X\to\mathbb{R}$ contínua em a tal que $f(x)-f(a)=L(x)(x-a)$ para todo $x\in X$.*

Em particular, $f'(a)=L(a)$.

Demonstração. Para (i) $\Rightarrow$ (ii), a ideia é definir a função $L\colon X\to\mathbb{R}$ dada por

$$L(x) := \begin{cases} \dfrac{f(x)-f(a)}{x-a}, & \text{se } x\neq a\\ f'(a), & \text{se } x=a\end{cases}$$

pois daí, por definição, $\lim\limits_{x\to a}L(x)=L(a)$, mostrando que L é continua em a e satisfaz a identidade desejada para todo $x\in X$. Para a recíproca, a identidade assegura

$$\frac{f(x)-f(a)}{x-a}=L(x)$$

sempre que $x\neq a$, de modo que o limite que caracteriza a derivada de f em a é, precisamente, o limite de L quando x se aproxima arbitrariamente do ponto a. □

Parece uma observação artificial, certo? Talvez algo criado apenas para testar a sua habilidade em manipular as definições? Isto está longe de ser verdade.

Exemplo 1.8.30 (Continuidade de funções diferenciáveis revisitada)**.** Note que a continuidade de uma função diferenciável é decorrência automática desta formulação: como $f(x) - f(a) = L(x)(x - a)$, temos

$$\lim_{x\to a} f(x) - f(a) = \lim_{x\to a} L(x) \cdot \lim_{x\to a} x - a = L(a) \cdot 0 = 0,$$

exatamente o que precisou ser feito na demonstração original. ▲

Exemplo 1.8.31 (Regra da cadeia, de novo)**.** Nas condições da Regra da Cadeia, note que a proposição anterior assegura funções $K\colon X \to \mathbb{R}$ e $L\colon Y \to \mathbb{R}$, com K contínua em a e L contínua em $f(a)$, tais que $f(x)-f(a) = K(x)(x-a)$ e $g(y)-g(f(a)) = L(y)(y-f(a))$ para quaisquer $x \in X$ e $y \in Y$. Para mostrar que $g\circ f$ é diferenciável em a, a proposição pede que encontremos uma função $T\colon X \to \mathbb{R}$, contínua em a, tal que $g(f(x))-g(f(a)) = T(x)(x-a)$ para todo $x \in X$. Ao fazer $y \coloneqq f(x)$ na identidade satisfeita por L, obtemos

$$g(f(x)) - g(f(a)) = L(f(x))(f(x) - f(a)) = L(f(x))K(x)(x - a),$$

onde a última igualdade decorre da identidade satisfeita por K. Mas veja só o que apareceu: a expressão $L(f(x))K(x)$ determina uma função do tipo $X \to \mathbb{R}$ que satisfaz a identidade desejada para todo $x \in X$. Logo, basta mostrar que a função $T\colon X \to \mathbb{R}$, dada por $T(x) \coloneqq L(f(x))K(x)$, é contínua em a. Vejamos duas formas de fazer isso:

✓ observe que se $(x_n)_{n\in\mathbb{N}}$ é sequência em X tal que $x_n \to a$, então $f(x_n) \to f(a)$, $L(f(x_n)) \to L(f(a))$ e $K(x_n) \to K(a)$, donde segue que $T(x_n) \to T(a)$;

✓ alternativamente, $L(f(x)) = (L \circ f)(x)$ é contínua em a pois f é contínua em a e L é contínua em $f(a)$, e daí $T(x)$ é o produto de duas funções contínuas em a.

Em todo caso, provamos que $g \circ f$ é diferenciável em a, e *descobrimos* que sua derivada em a é dada por $T(a) = L(f(a))K(a) = g'(f(a))f'(a)$. ▲

Embora os argumentos acima sejam essencialmente idênticos aos que foram apresentados na seção anterior, aqui há uma diferença sutil: não foi preciso adivinhar nada. Como última ilustração, vejamos o importante

Teorema 1.8.32 (da função inversa)**.** *Seja $f\colon X \to Y$ bijeção contínua e diferenciável num ponto de acumulação $p \in X$. Se f^{-1} é contínua em $f(p)$, então f^{-1} é diferenciável em $f(p)$ se, e somente se, $f'(p) \neq 0$. Além disso, $(f^{-1})'(f(p)) = \frac{1}{f'(p)}$.*

Demonstração. Por hipótese, existe $L\colon X \to \mathbb{R}$ contínua em p com $L(x)(x-p) = f(x)-f(p)$ para todo $x \in X$ e $f'(p) = L(p)$. Em particular, por f ser bijetora, para cada $y \in Y$ existe um único $x \in X$ tal que $f(x) = y$ ou, equivalentemente, $f^{-1}(y) = x$. Logo,

$$y - f(p) = f(x) - f(p) = L(x)(x - p) = L(x)(f^{-1}(y) - f^{-1}(f(p))), \tag{1.10}$$

mostrando que se $y \neq f(p)$, então $L(x) = L(f^{-1}(y)) \neq 0$ (percebeu?)*.

Desse modo, se $f'(p) \neq 0$, podemos definir $\widetilde{L}\colon Y \to \mathbb{R}$ por meio da composição

$$\begin{tikzcd} X \arrow[r, "L"] & \mathbb{R} \arrow[d, "\iota"] \\ Y \arrow[u, "f^{-1}"] \arrow[r, "\widetilde{L}"'] & \mathbb{R} \end{tikzcd}$$

$$y \longmapsto \frac{1}{L(f^{-1}(y))}$$

onde $\iota(z) := \frac{1}{z}$ para todo $z \in \mathbb{R} \setminus \{0\}$. Enquanto as hipóteses asseguram a continuidade de $\widetilde{L}$ em $f(p)$ (certo?!)*, (1.10) garante que $\widetilde{L}(y)(y - f(p)) = f^{-1}(y) - f^{-1}(f(p))$ para todo $y \in Y$, i.e., f^{-1} é diferenciável em $f(p)$. A recíproca, assim como a identidade final, segue da Regra da Cadeia (mas não se esqueça do Exercício 1.159), e você pode cuidar dos detalhes $(\substack{\star\\\star})$. $\square$

§1 Derivadas em dimensões maiores

A verdadeira grande vantagem da formulação de Carathéodory é a facilidade para estender a definição de derivada para funções da forma $\mathbb{R}^m \to \mathbb{R}^n$: com $X \subseteq \mathbb{R}^m$ e $f\colon X \to \mathbb{R}^n$, a função L deveria ser capaz de dar sentido a uma equação do tipo $f(x) - f(a) = L(x)(x - a)$ para qualquer $x \in X$. Logo, $L(x)$ deveria ser um objeto que, ao interagir com o vetor $x - a$ em $\mathbb{R}^m$, *retornasse* o vetor $f(x) - f(a)$ em $\mathbb{R}^n$. Após pensar um pouco, talvez a primeira ideia seja pedir que $L(x)$ seja uma matriz — mas, já que a ideia é fazer as coisas com elegância, vamos pedir que $L(x)$ seja uma transformação linear de $\mathbb{R}^m$ em $\mathbb{R}^n$! Parece estranho, mas não se preocupe: nada disso é óbvio.

(i) A ideia não é que L seja uma transformação linear $\mathbb{R}^m \to \mathbb{R}^n$, mas sim que para cada $x \in X$, $L(x)\colon \mathbb{R}^m \to \mathbb{R}^n$ seja uma transformação linear: ou seja, L é uma função que a cada ponto x de X associa uma transformação linear $L(x)$. Sim, é uma função que associa pontos a outras funções. Não, o mundo não te odeia.

(ii) Mesmo com $m = n = 1$, $L(x)$ *sempre foi uma transformação linear*. De fato, uma função $T\colon \mathbb{R} \to \mathbb{R}$ é transformação linear se, e somente se, existe um escalar $a_T \in \mathbb{R}$ tal que $T(r) = a_T r$ para todo $r \in \mathbb{R}$ (certo?)$\overset{\star}{\star}$. Logo, quando fizemos $L(x) \cdot (x - a)$ com $L(x) \in \mathbb{R}$, secretamente tínhamos a transformação linear $T_x\colon \mathbb{R} \to \mathbb{R}$ que a cada $r \in \mathbb{R}$ associa o número $T_x(r) := L(x) \cdot r$, de modo que $L(x) \cdot (x - a)$ é, meramente, $T_x(x - a)$.

Superadas as angústias algébricas, e denotando por $\mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$ o conjunto das transformações lineares do tipo $\mathbb{R}^m \to \mathbb{R}^n$, o próximo passo é entender o que significa dizer que uma função $L\colon X \to \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$ é contínua em $a \in X$. Tipicamente, apela-se para Álgebra Linear e para o *isomorfismo* existente entre $\mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$ e $\mathbb{R}^{mn}$ a fim de tratar transformações lineares como matrizes/vetores com alguma das normas usuais. O caminho aqui será outro: justamente por conta de tal isomorfismo, *sabe-se*[†] que quaisquer duas normas em $\mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$ são equivalentes, o que permite utilizar uma norma mais apropriada a espaços de funções lineares.

Na inofensiva Proposição 1.6.17, provamos que uma transformação linear $T\colon E \to S$ é contínua se, e somente se, existe $r > 0$ tal que $\|T(x)\|_S \leq r\|x\|_E$ para todo $x \in E$. Em particular, isto assegura a boa definição do número real

$$\|T\| := \inf\{r > 0 : \|T(x)\|_S \leq r\|x\|_E \text{ para todo } x \in E\}.$$

Como a notação sugere, ao fazer T percorrer o conjunto de todas as transformações lineares contínuas de E em S, obtemos uma norma $\|\cdot\|\colon \mathcal{L}_c(E, S) \to \mathbb{R}$, onde $\mathcal{L}_c(E, S)$ denota o espaço vetorial das transformações lineares contínuas de E em S.

[†] Cf. Teorema 2.0.27.

Exercício 1.162 ($^{\star}_{\star\star}$)**.** Prove as afirmações acima. Em particular, gaste algum tempo observando que $\mathcal{L}_c(E,S)$ é realmente um espaço vetorial[†]. ■

A grande vantagem da norma $\|\cdot\|$ acima, usualmente chamada de **norma de operador**, é a validade da desigualdade

$$\|T(x)\|_S \leq \|T\| \cdot \|x\|_E$$

para todo $x \in E$: um modo divertido de prová-la consiste em observar que existe uma sequência $(r_n)_{n\in\mathbb{N}}$ de números reais satisfazendo $\|T(x)\|_S \leq r_n\|x\|_E$ e $r_n \to \|T\|$ (por quê?)$^{\star}_{\star}$, e daí $\|T(x)\|_S \leq \|x\|_E \lim_{n\to\infty} r_n = \|x\|_E \cdot \|T\|$

Definição 1.8.33. Para espaços normados X e Y, seja $S \subseteq X$ um subconjunto aberto[‡] e considere $f \colon S \to Y$ uma função. Diremos que f é **diferenciável em** $a \in S$ se existe uma função $\Phi \colon S \to \mathcal{L}_c(X,Y)$, contínua em a, tal que $\Phi(x)(x-a) = f(x) - f(a)$ para todo $x \in S$. ¶

Abusos de notação. Daqui em diante, os subíndices para indicar o espaço ambiente da norma serão abandonados. Assim, ao encontrar algo como "$\|T(x)\| \leq \|T\|\|x\|$", você deve estar ciente de que "$\|T(x)\|$" indica a norma do vetor $T(x)$ num espaço, "$\|x\|$" indica a norma do vetor "x" em outro espaço, enquanto "$\|T\|$" indica a norma da transformação linear T, e cada uma dessas normas é, possivelmente, diferente da outra.

Acima, a função Φ costuma ser chamada de *(função de) inclinação de f* no ponto a, enquanto a transformação linear *contínua* (!) $\Phi(a) \colon X \to Y$ é a **derivada** de f em a. A fim de xingar $\Phi(a)$ por um apelido mais específico, como $f'(a)$, é preciso mostrar que ela independe de Φ. Emergem daí as *derivadas direcionais*.

Proposição 1.8.34. *Nas condições acima,* $\Phi(a)(v) = \lim\limits_{t\to 0} \dfrac{f(a+tv) - f(a)}{t}$ *para todo* $v \in X$*.*

Demonstração. Lembre-se de que $v \in X$ e $a \in S$ estão fixados. Agora, observe que para $t \neq 0$ *suficientemente próximo* de 0, o quociente

$$\frac{f(a+tv) - f(a)}{t} := \frac{1}{t} \cdot (f(a+tv) - f(a))$$

está *bem definido*, já que a função $\gamma \colon \mathbb{R} \to X$ dada por $\gamma(t) := a + tv$ é contínua: como $\gamma(0) = a \in S$ e S é aberto em X, existe $r > 0$ tal que $\gamma(t) \in S$ sempre que $|t| < r$ (por quê?)$^{\star}_{\star}$, mostrando que $a + tv$ pertence ao domínio de f em tais situações.

Para tais valores de t, a definição de Φ assegura

$$f(a+tv) - f(a) - t\Phi(a)(v) = t\Phi(a+tv)(v) - t\Phi(a)(v),$$

pois $\Phi(a+tv)$ é linear e $a + tv - a = tv$ (verifique!)$^{\star}_{\star}$. Logo,

$$\left|\frac{f(a+tv) - f(a)}{t} - \Phi(a)(v)\right| = |\Phi(a+tv)(v) - \Phi(a)(v)| \leq \|\Phi(a+tv) - \Phi(a)\| \cdot \|v\|,$$

que por sua vez pode ser *arbitrariamente controlado*: como γ é contínua em 0 e Φ é contínua em $\gamma(0) = a$, existe $r' > 0$ menor do que r tal que $\|\Phi(a+tv) - \Phi(a)\| < \frac{\varepsilon}{\|v\|}$ para todo $t \neq 0$ com $|t| < r'$. Percebeu que acabou? ($\star$) □

[†] Quando E tem dimensão finita, vale $\mathcal{L}(E,S) = \mathcal{L}_c(E,S)$, mas isto se perde quando a dimensão de E é infinita (cf. Corolário 2.0.29 e Observação 2.0.30).

[‡] É possível *implementar* a definição pedindo que $a \in X$ seja um ponto de acumulação de X, mas a complicação não justifica a generalidade. Aliás, mesmo para funções reais, a maioria das referências lida apenas com funções diferenciáveis definidas em intervalos abertos.

Exercício 1.163 $(^{\star}_{\star\star})$**.** Mostre que se f é diferenciável em $a \in S$, então f é contínua em a. Dica: note que para x no domínio de f vale $0 \leq \|f(x) - f(a)\| \leq \|\Phi(x)\| \cdot \|x - a\|$, e não se esqueça de que existe $\delta > 0$ tal que $\|x - a\| < \delta$ acarreta $\|\Phi(x)\| < 1 + \|\Phi(a)\|$. ■

Exercício 1.164 $(^{\star}_{\star})$**.** Mostre que a norma de um espaço normado não é diferenciável na origem. Dica: se fosse, então existiria uma transformação linear contínua $T\colon E \to \mathbb{R}$ tal que $T(v) = \|v\| \lim_{t\to 0} \dfrac{|t|}{t}$ para todo $v \in E$. Isto parece possível? ■

O limite encontrado na última proposição costuma ser chamado de **derivada de f na direção de v no ponto** a, que se denota, usualmente, por $\frac{\partial f}{\partial v}(a)$. Dessa forma, o que se demonstrou foi que uma função diferenciável num ponto admite todas as *derivadas direcionais* naquele ponto, pois $f'(a)(v) = \frac{\partial f}{\partial v}(a)$.

Com um pouco mais de Álgebra Linear e paciência, desenvolve-se todo o arsenal básico de derivação em $\mathbb{R}^n$, mas isto só será explorado no Capítulo 3. Por ora, há espaço para mais um resultado antes de encerrarmos a discussão.

Lema 1.8.35. *Para espaços normados X, Y e Z, a correspondência $(T, S) \mapsto T \circ S$ determina uma função contínua $\circ\colon \mathcal{L}_c(Y, Z) \times \mathcal{L}_c(X, Y) \to \mathcal{L}_c(X, Z)$.*

Demonstração. Se $T_n \to T$ em $\mathcal{L}_c(Y, Z)$ e $S_n \to S$ em $\mathcal{L}_c(X, Z)$, então $T_n \circ S_n \to T \circ S$, posto que

$$\begin{aligned} \|T_n(S_n(x)) - T(S(x))\| &= \|T_n(S_n(x) - S(x)) + (T_n - T)(S(x))\| \\ &\leq \|T_n\|\|S_n - S\|\|x\| + \|T_n - T\|\|S\|\|x\| \end{aligned}$$

para qualquer $x \in X$, donde segue o resultado (conclua!)$^{\star}_{\star}$. □

Teorema 1.8.36 (Regra da cadeia)**.** *Sejam X, Y e Z espaços normados, $S \subseteq X$ e $T \subseteq Y$ subconjuntos abertos e $f\colon S \to Y$ e $g\colon T \to Z$ funções, com $\operatorname{im}(f) \subseteq Z$. Se f e g são diferenciáveis em $a \in S$ e $f(a) \in T$, respectivamente, então $g \circ f$ é diferenciável em a e $(g \circ f)'(a) = g'(f(a)) \circ f'(a)$.*

Demonstração. Basta imitar a prova apresentada no Exemplo 1.8.31, trocando a multiplicação por composição! Boa sorte! $(^{\star}_{\star\star})$ □

$$\begin{tikzcd}
T \times S \arrow[r, "\Psi\times\Phi"] & \mathcal{L}_c(Y, Z) \times \mathcal{L}_c(X, Y) \arrow[dd, "\circ"] \\
S \times S \arrow[u, "f\times \mathrm{Id}_S"] & \\
S \arrow[u, "\Delta"] \arrow[r, "\Gamma"'] & \mathcal{L}_c(X, Z)
\end{tikzcd}$$

Figura 1.14: Se $\Phi\colon S \to \mathcal{L}_c(X, Y)$ e $\Psi\colon T \to \mathcal{L}_c(Y, Z)$ são as funções inclinação de f e g nos pontos a e $f(a)$, respectivamente, então a composição Γ é uma função inclinação para $g \circ f$ no ponto a: o diagrama acima já te garante a continuidade em a (lembrando que $\Delta(s) := (s, s)$ para todo s).

§2 Exercícios extras

Exercício 1.165 $(^{\star}_{\star})$**.** Para uma transformação linear $T\colon X \to Y$ entre espaços normados, considere os seguintes *valores* (possivelmente em $[-\infty, +\infty]$):

(i) $A \coloneqq \inf\{r > 0 : \|T(x)\| \leq r\|x\| \text{ para todo } x \in X\}$;

(ii) $B \coloneqq \sup\{\|T(x)\| : x \in X \text{ e } \|x\| \leq 1\}$;

(iii) $C \coloneqq \sup\{\|T(x)\| : x \in X \text{ e } \|x\| = 1\}$;

(iv) $D \coloneqq \sup\left\{\frac{\|T(x)\|}{\|x\|} : x \in X \text{ e } x \neq 0\right\}$.

Prove que $A = B = C = D$. ■

Exercício 1.166 $(^{\star}_{\star})$**.** Seja $T\colon \mathbb{R} \to X$ uma transformação linear em que X é um espaço normado. Mostre que $\|T\| = \|T(1)\|$. Em particular, conclua que T é contínua. ■

Exercício 1.167 $(^{\star}_{\star\star})$**.** Para um intervalo aberto $I \subseteq \mathbb{R}$ e uma função $f\colon I \to X$, em que X é um espaço normado, mostre que f é diferenciável em $p \in I$ se, e somente se, existe $v \in X$ tal que

$$\lim_{t\to 0} \frac{f(p+t) - f(p)}{t} = v,$$

i.e., para todo $\varepsilon > 0$ existe $\delta > 0$ tal que se $0 < |t| < \delta$, então $\|f(p+t) - f(p) - tv\| < t\varepsilon$. Em particular, $f'(p)(t) = tv$ para todo $t \in \mathbb{R}$. Dica: a princípio, $f'(p)$ é uma transformação linear (contínua)[†] do tipo $\mathbb{R} \to X$, mas isto permite tratá-la explicitamente como um vetor de X, lembra? ■

Observação 1.8.37. O exercício anterior justifica tratar derivadas de caminhos em X como vetores em X. △

Exercício 1.168 (Derivadas de caminhos — $(\star)$)**.** Para um intervalo aberto $I \subseteq \mathbb{R}$ e funções diferenciáveis $f_1, \ldots, f_n\colon I \to \mathbb{R}$, mostre que a função $f\colon I \to \mathbb{R}^n$ dada por $f(t) \coloneqq (f_1(t), \ldots, f_n(t))$ é diferenciável e, mais ainda,

$$f'(t) = (f_1'(t), \ldots, f_n'(t))$$

para todo $t \in I$. Se $g\colon I \to \mathbb{R}^n$ é diferenciável, o que se pode dizer acerca de $\pi_i \circ g$? Observação: aqui, $\pi_i\colon \mathbb{R}^n \to \mathbb{R}$ é a projeção na i-ésima coordenada ■

Exercício 1.169 (!?)**.** Calcule as derivadas das suas funções favoritas do tipo $I \to \mathbb{R}^n$.■

Exercício 1.170 (Selecionados do "Elão" [26] — $(^{\star}_{\star})$)**.** Sejam I um intervalo aberto de $\mathbb{R}$ e $f\colon I \to X$ diferenciável, onde X é um espaço normado.

a) Mostre que se $a \in I$ é ponto de acumulação de $f^{-1}[\{v\}]$ para algum $v \in X$, então $f'(a) = 0$. Dica: em virtude do Exercício 1.203, existe $(x_n)_n$ em $I \setminus \{a\}$ com $x_n \to a$.

[†]Neste caso, a continuidade é automática (certo?)*. Para uma situação um pouco mais geral, confira o Lema 2.0.26.

b) Suponha $f'(a) \neq 0$ para $a \in I$. Mostre que se existirem $x, v \in X$ e $(t_n)_n$ uma sequência injetiva em I com $t_n \to a$ e $f(t_n) \in \{tv + x : t \in \mathbb{R}\}$ para todo n, então $f'(a)$ é múltiplo de v. Observação: com $X := \mathbb{R}^n$, isto significa dizer que $R := \{tv + x : t \in \mathbb{R}\}$ é *a* reta tangente ao gráfico de f no ponto $(a, f(a))$. Dica: para o caso geral, você pode *assumir*, por enquanto[†], que se $F \subseteq X$ é subespaço vetorial com dimensão finita e $(x_n)_{n\in\mathbb{N}}$ é sequência em F com $x_n \to x$, então $x \in F$.

Exercício 1.171 $(^{\star}_{\star})$**.** Sejam $T\colon X \to Y$ uma transformação linear contínua entre espaços normados, $S \subseteq X$ um subconjunto aberto e $f := T|_S$. Mostre que $f'(x) = T$ para todo $x \in S$. Dica: neste caso, existe uma função constante $L\colon S \to \mathcal{L}_c(X, Y)$ que funciona. ■

Exercício 1.172 $(^{\star}_{\star\star})$**.** Seja $\langle\cdot,\cdot\rangle$ um produto interno num espaço vetorial V e considere sobre V a norma induzida por tal produto interno. Mostre que para $(a, b) \in V \times V$ qualquer,

$$\langle\cdot,\cdot\rangle'(a, b)(h, k) = \langle h, b\rangle + \langle a, k\rangle,$$

i.e., a derivada de $\langle\cdot,\cdot\rangle$ no ponto (a, b) é a transformação linear $V \times V \to \mathbb{R}$ que a cada (h, k) associa o escalar $\langle h, b\rangle + \langle a, k\rangle$. Dica: para $(x, y) \in V \times V$, faça

$$\Phi(x, y)(h, k) := \left\langle h, b + \frac{y - b}{2}\right\rangle + \left\langle a + \frac{x - a}{2}, k\right\rangle$$

para cada par $(h, k) \in V \times V$, o que define uma função inclinação $\Phi\colon V \times V \to \mathcal{L}_c(V \times V, \mathbb{R})$ adequada no ponto (a, b). ■

Exercício 1.173 $(^{\star}_{\star\star})$**.** Sejam $f\colon S \to Y$ e $g\colon S \to Y$ funções diferenciáveis, onde S é subconjunto aberto de um espaço normado X. Mostre que a função $(f, g) : S \times S \to Y \times Y$, que a cada par $(s, t) \in S \times S$ associa o vetor $(f(s), g(t)) \in Y \times Y$, é diferenciável em todos os pontos e, mais ainda,

$$\begin{aligned}(f, g)'(s, t)\colon X \times X &\to Y \times Y\\ (h, k) &\to (f'(s)(h), g'(t)(k)).\end{aligned}$$

Dica: pode ser útil revisar a Subseção 1.7.1 §1; se preferir, assuma que X e Y são espaços euclidianos, mas isto se provará irrelevante. ■

Exercício 1.174 $(^{\star}_{\star})$**.** Para um espaço V com produto interno $\langle\cdot,\cdot\rangle$, considere $f, g\colon S \to V$ funções diferenciáveis e defina $\langle f, g\rangle\colon S \to \mathbb{R}$ fazendo $\langle f, g\rangle(s) := \langle f(s), g(s)\rangle$ para todo $t \in S$. Mostre que

$$\langle f, g\rangle'(s)(x) = \langle f'(s)(x), g(s)\rangle + \langle f(s), g'(s)(x)\rangle$$

para quaisquer $s \in S$ e $x \in X$. Conclua que se $V = \mathbb{R}$, então $f \cdot g$ é diferenciável. Dica: Regra da Cadeia + exercícios anteriores. ■

Exercício 1.175 $(^{\star}_{\star})$**.** Mostre que se $f\colon S \to X$ e $g\colon S \to X$ são diferenciáveis em $a \in S$, então $\alpha f + \beta g$ é diferenciável em a, com $(\alpha f + \beta g)'(a) = \alpha f'(a) + \beta g'(a)$, para quaisquer $\alpha, \beta \in \mathbb{R}$. ■

[†] *Spoiler alert*: subespaços com dimensão finita são de Banach (cf. Corolário 2.0.28) e, por conseguinte, são fechados (cf. Exercício 1.257).

Exercício 1.176 (Opcional: derivadas no sentido de Fréchet ($^{\star}_{\star\star}$))**.** Para espaços normados X e Y, sejam $S \subseteq X$ um subconjunto aberto e considere $f\colon S \to Y$ uma função. Diremos que f é **Fréchet-diferenciável** em $a \in S$ se existe uma transformação linear contínua $T\colon X \to Y$ tal que

$$\lim_{x \to a} \frac{f(x) - f(a) - T(x-a)}{\|x-a\|} = 0.$$

a) Mostre que se f é diferenciável em a, então f é Fréchet-diferenciável em a. Dica: tome $T := f'(a)$ e perceba que $\|\Phi(x)(x-a) - T(x-a)\| \leq \|\Phi(x) - T\|\|x-a\|$.

b) Mostre que se f é Fréchet-diferenciável em a, então f é diferenciável em a e, mais ainda, $f'(a) = T$. *Comentário.* Precisa-se definir uma função inclinação $\Phi\colon S \to \mathcal{L}_c(X,Y)$ no ponto a com $\Phi(a) = T$, porém, desta vez a identidade $\Phi(x)(x-a) = f(x) - f(a)$ não *caracteriza* a transformação linear $\Phi(x)\colon X \to Y$ pois a dimensão de X é, possivelmente, maior do que 1! Para definir $\Phi(x)$ num *complemento* do subespaço vetorial $\{t(x-a) : t \in \mathbb{R}\}$, siga os passos a seguir.

   (i) Suponha que para cada $x \in S \setminus \{a\}$ exista um funcional linear contínuo $\varphi_x\colon X \to \mathbb{R}$ tal que $\varphi_x(x-a) = \|x-a\|$ e $\|\varphi_x\| \leq 1$. Com isso, defina $\Phi(a) := T$ e, para $x \in S \setminus \{a\}$, faça

   $$\Phi(x)(z) := \varphi_x(z) \cdot \frac{f(x) - f(a) - T(x-a)}{\|x-a\|} + T(z)$$

   para cada $z \in X$. Mostre que $\Phi\colon S \to \mathcal{L}_c(X,Y)$ tem as propriedades desejadas.

   (ii) Mostre que se $X = \mathbb{R}^n$, então os funcionais φ_x existem. Dica: há várias formas de proceder, uma delas é fazer $\varphi_x(y) := \left\langle y, \dfrac{x-a}{\|x-a\|} \right\rangle$, onde $\langle \cdot, \cdot \rangle$ denota o produto interno usual de $\mathbb{R}^n$. ■

Observação 1.8.38. A equivalência entre as definições se verifica em qualquer espaço normado. Porém, no caso geral, precisa-se apelar para o Teorema de Hahn-Banach (cf. Brézis [6] ou qualquer outro bom livro de Análise Funcional). Com isso dito, em espaços *euclidianos* (os $\mathbb{R}^n$'s da vida), a formulação de Fréchet costuma ser mais utilizada — possivelmente por favorecer uma abordagem quantitativa para diferenciabilidade, o que facilita questões computacionais. Ela é a mais comum na maioria dos textos que lidam com diferenciabilidade de funções em várias variáveis... *C'est la vie...* Para mais detalhes sobre derivadas à moda Carathéodory, confira os trabalhos [0] e [2]. △

1.9 Exercícios adicionais

Essenciais: cálculos elementares (mas nem tanto)

Exercício 1.177 ($^{\star}_{\star\star}$)**.** Apresente uma definição razoável para "$\lim\limits_{|x| \to +\infty} f(x) = L$", onde entende-se que $f\colon X \to \mathbb{R}$ é uma função e $X \subseteq \mathbb{R}$ é um subconjunto ilimitado tanto superiormente quanto inferiormente, e $L \in [-\infty, +\infty]$. ■

Exercício 1.178 ($^{\star}_{\star}$)**.** Mostre que $\lim\limits_{|x| \to +\infty} \dfrac{1}{x} = 0$. ■

Exercício 1.179 $(\overset{\star}{\star})$. Mostre que não existe $L \in [-\infty, +\infty]$ tal que $\lim_{|x| \to +\infty} x^3 = L$. ∎

Exercício 1.180 $(\overset{\star}{\star})$. Uma função $f\colon \mathbb{R} \to \mathbb{R}$ é dita **par** se $f(x) = f(-x)$ para todo $x \in \mathbb{R}$. Mostre que se f é par, então $\lim_{|x| \to +\infty} f(x)$ existe se, e somente se, $\lim_{x \to +\infty} f(x)$ existe. Em particular, quando existem, $\lim_{x \to +\infty} f(x) = \lim_{|x| \to +\infty} f(x)$. ∎

Exercício 1.181 $(\overset{\star}{\star})$. Uma função $f\colon \mathbb{R} \to \mathbb{R}$ é dita **impar** se $f(x) = -f(x)$ para todo $x \in \mathbb{R}$. Determine em quais situações existe $\lim_{|x| \to +\infty} f(x)$ para f ímpar. ∎

Exercício 1.182 $(\overset{\star}{\star\star})$. Mostre que se $f\colon \mathbb{R} \to \mathbb{R}$ é função polinomial, então deve valer $\lim_{|x| \to +\infty} |f(x)| = +\infty$. Dica: $|\alpha| - |\beta| \leq |\alpha - \beta|$ para quaisquer $\alpha, \beta \in \mathbb{R}$. ∎

Exercício 1.183 $(\star)$. Mostre que $\lim_{x \to +\infty} \sqrt{x} = +\infty$ e $\lim_{x \to +\infty} \sqrt{x} - \sqrt{x+1} = 0$. ∎

Exercício 1.184 $(\overset{\star}{\star})$. Determine $\lim_{x \to +\infty} \sqrt{x + \sqrt{x}} - \sqrt{x}$. Dica: não é 0. ∎

Exercício 1.185 (Parte inteira de um número real — $(\overset{\star}{\star})$). Para cada $x \in \mathbb{R}$ com $x > 0$, considere $N_x := \min\{n \in \mathbb{N} : n > x\} > 0$ e defina o *número natural* $\lfloor x \rfloor := N_x - 1$, usualmente xingado de **parte inteira** de x.

a) Mostre que $\lfloor x \rfloor$ é o único número natural para o qual $x = \lfloor x \rfloor + r$ com $0 \leq r < 1$. Calcule $\lfloor x \rfloor$ para alguns números reais de sua preferência.

b) Para $p \in (0, +\infty) \setminus \mathbb{N}$, mostre que $\lim_{z \to p} \lfloor z \rfloor = \lfloor p \rfloor$.

c) Para $p \in \mathbb{N} \setminus \{0\}$, mostre que $\lim_{z \to p^-} \lfloor z \rfloor = \lfloor p \rfloor - 1$ e $\lim_{z \to p^+} \lfloor z \rfloor = \lfloor p \rfloor$.

d) Em quais pontos a função $\lfloor \cdot \rfloor \colon (0, +\infty) \to \mathbb{R}$ é contínua? ∎

Exercício 1.186 (Limites de funções polinomiais — $(\overset{\star}{\star})$). Para $\lambda \in [-\infty, +\infty] \setminus \{0\}$, defina $\operatorname{sgn}(\lambda) := -1$ para $\lambda < 0$ e $\operatorname{sgn}(\lambda) := 1$ para $\lambda > 0$. Com isso, para $P(t) := \alpha_0 + \ldots + \alpha_n t^n \in \mathbb{R}[t]$ um polinômio de grau $n \geq 1$ e $\lambda \in \{-\infty, +\infty\}$, mostre que

$$\lim_{x \to \lambda} P(x) = \underbrace{\alpha_n \operatorname{sgn}(\lambda)^n \cdot (+\infty)}_{(*)}, \tag{1.11}$$

onde $(*)$ deve ser interpretada com as regras apresentadas na Subseção 1.7.1 §2 (cf. pág. 185). Sugestão: primeiro, observe que $\lim_{x \to \lambda} x = \lambda$; daí, proceda por indução (e use o Teorema 1.2.8 sem dó nem piedade). ∎

Exercício 1.187 $(\star)$. Para $P(t) := -19 + 11t - 5t^7$, calcule $\lim_{x \to -\infty} P(x)$. ∎

Exercício 1.188 $(\overset{\star}{\star})$. Para $X \subseteq \mathbb{R}$, uma função $f\colon X \to \mathbb{R}$ é chamada de **racional** se existem polinômios $P(t), Q(t) \in \mathbb{R}[t]$ tais que $Q(x) \neq 0$ e $f(x) = \dfrac{P(x)}{Q(x)}$ para todo $x \in X$. Suponha que $P(t)$ e $Q(t)$ tenham graus $m, n \in \mathbb{N} \setminus \{0\}$, e considere $\lambda \in \{-\infty, +\infty\}$.

a) Mostre que se $m < n$, então $\lim_{x \to \lambda} \dfrac{P(x)}{Q(x)} = 0$.

b) Mostre que se $m = n$, então $\lim\limits_{x \to \lambda} \dfrac{P(x)}{Q(x)} = \dfrac{p_m}{q_m}$, onde p_m e q_m são os *coeficientes líderes*[†] de $P(t)$ e $Q(t)$, respectivamente.

c) Mostre que se $n > m$, então $\lim\limits_{x \to \lambda} \dfrac{P(x)}{Q(x)} = \dfrac{p_n}{b_m} \cdot \operatorname{sgn}(\lambda)^{n-m} \cdot (+\infty)$. ■

Exercício 1.189 $(\substack{\star \\ \star\star})$**.** Para funções polinomiais $f, g\colon X \to \mathbb{R}$ e $r \in \mathbb{R}$, *investigue*[‡] o comportamento de $\lim\limits_{x \to r} \dfrac{f(x)}{g(x)}$ sabendo que $f(r) = g(r) = 0$. Dica: lembre-se de que você pode escrever $f(x) = (x - r)^N P(x)$ e $g(x) = (x - r)^M Q(x)$, onde P e Q são polinômios satisfazendo $P(r) \neq 0$ e $Q(r) \neq 0$; use as relações entre N e M para te guiar em sua investigação. ■

Exercício 1.190 $(\star)$**.** Mostre que a função $f\colon \mathbb{R} \to \mathbb{R}$ dada por

$$f(x) := \begin{cases} x^2, & \text{se } x \in \mathbb{Q}, \\ 0, & \text{se } x \notin \mathbb{Q} \end{cases}$$

é diferenciável em 0 mas não é contínua em nenhum ponto de $\mathbb{R} \setminus \{0\}$. ■

Exercício 1.191 $(\star)$**.** Para polinômios $p(x)$ e $q(x)$, suponha que as funções polinomiais induzidas em $\mathbb{R}$, digamos $P\colon \mathbb{R} \to \mathbb{R}$ e $Q\colon \mathbb{R} \to \mathbb{R}$ dadas por $P(r) := p(r)$ e $Q(r) := q(r)$, sejam tais que $P = Q$. Mostre que $p = q$. Dica: mostre que um polinômio real não nulo tem no máximo finitas raízes[††]. ■

Exercício 1.192 $(\substack{\star \\ \star})$**.** Para $f, g\colon X \to \mathbb{R}$ funções contínuas, mostre que $\max\{f, g\}\colon X \to \mathbb{R}$ e $\min\{f, g\}\colon X \to \mathbb{R}$ são contínuas. Dica: note que $\max\{\alpha, \beta\} = \dfrac{\alpha + \beta + |\alpha - \beta|}{2}$, e tente encontrar uma expressão parecida para o mínimo. ■

Exercício 1.193 $(\substack{\star \\ \star})$**.** Para $S \subseteq \mathbb{R}$ com $S \neq \emptyset$ e $x \in \mathbb{R}$, definimos a *distância* entre S e x como sendo o número real $d(x, S) := \inf\{|x - s| : s \in S\}$. Mostre que a função $f\colon \mathbb{R} \to \mathbb{R}$, dada por $f(x) := d(x, S)$, é contínua. Dica: mostre que $|f(x) - f(y)| \leq |x - y|$ para quaisquer $x, y \in \mathbb{R}$. ■

Essenciais: resultados importantes sobre limites e derivadas

Exercício 1.194 $(\substack{\star \\ \star})$**.** Seja $p \in [-\infty, +\infty]$ ponto de acumulação de $X \subseteq \mathbb{R}$. Mostre que se $\lim\limits_{x \to p} f(x) = L$ para $L \in [-\infty, +\infty]$, então L é ponto de acumulação do conjunto $\operatorname{im}(f) := \{f(x) : x \in X\}$. Dica: use a definição de ponto de acumulação via intervalos abertos. ■

[†] Coeficiente (escalar, numerozinho de verdade, etc.) que acompanha o monômio de maior grau do polinômio. Por exemplo, em $P(t) := -5 + 90t - 2t^3$, o coeficiente líder é -2.

[‡] Nesse tipo de situação, "investigar" significa analisar por conta própria o que ocorre nas diversas situações possíveis dentro do contexto.

[††] Parece um problema bobo, mas não é, afinal de contas, polinômios e funções polinomiais são animais diferentes. Se duvida, considere o polinômio $p(x) := x^2 + x$, mas com coeficientes no anel $\mathbb{Z}/2\mathbb{Z}$: o polinômio não é nulo, mas $p(r) = \overline{0}$ para todo $r \in \mathbb{Z}/2\mathbb{Z}$.

Exercício 1.195 $\binom{\star}{\star}$**.** Para $X \subseteq \mathbb{R}$ e $p, L \in [-\infty, +\infty]$ tais que p é ponto de acumulação de X, seja $I \subseteq [-\infty, +\infty]$ um intervalo aberto com $p \in I$. Finalmente, para uma função $f\colon X \to \mathbb{R}$, chame de g a restrição de f ao subconjunto $X \cap I$. Com isso, mostre que $\lim_{x \to p} f(x) = L$ se, e somente se, $\lim_{x \to p} g(x) = L$. Dica: antes de reclamar, faça um desenho para interpretar o que o enunciado está dizendo. ■

Exercício 1.196 $\binom{\star}{\star}$**.** Para $X, Y \subseteq [-\infty, +\infty]$, mostre que $f\colon X \to Y$ é contínua em $p \in X$ se, e somente se, para todo intervalo aberto $I \subseteq [-\infty, +\infty]$ tal que $f(p) \in I$ existe um intervalo aberto $J \subseteq [-\infty, +\infty]$ tal que $p \in J$ e $f(x) \in I$ sempre que $x \in J \cap I$. ■

Exercício 1.197 $\binom{\star}{\star}$**.** Sejam $X \subseteq \mathbb{R}$ um subconjunto e $p \in [-\infty, +\infty]$ um ponto de acumulação de X. Para funções $f, g\colon X \to \mathbb{R}$, suponha que

(i) $\lim_{x \to p} f(x) = 0$, e

(ii) existe um intervalo aberto $I \subseteq [-\infty, +\infty]$ e um número real $M > 0$ tal que $p \in I$ e $|g(x)| < M$ para todo $x \in I \cap X$ com $x \neq p$.

Sob tais condições, mostre que $\lim_{x \to p} f(x)g(x) = 0$. Dica: imite a demonstração da Proposição 1.2.19, ou perceba como aplicar a proposição diretamente para obter o resultado pedido. ■

Exercício 1.198 $\binom{\star}{\star}$**.** Dê exemplos de que o resultado anterior é falso sem a condição (i). Dica: você já conhece exemplos com $X := \mathbb{N}$ e $p := +\infty$. ■

Exercício 1.199 $\binom{\star}{\star}$**.** Para $X \subseteq \mathbb{R}$ e $p \in X$ um ponto de acumulação de X, suponha que $f\colon X \to \mathbb{R}$ seja monótona e diferenciável em p.

a) Mostre que se f é crescente, então para todo $x \in X \setminus \{p\}$ se verifica

$$\frac{f(x) - f(p)}{x - p} \geq 0.$$

b) Mostre que se f é crescente, então $f'(p) \geq 0$.

c) Mostre que se f é decrescente, então $f'(p) \leq 0$. Dica: $-f$ é crescente. ■

Exercício 1.200 $\binom{\star}{\star}$**.** Mostre que a função $f\colon \mathbb{R} \to \mathbb{R}$ dada por $f(x) := x^3$ é estritamente crescente e diferenciável. É verdade que $f'(p) > 0$ para todo $p \in \mathbb{R}$? ■

Exercício 1.201 (*For fun* — $(\star)$)**.** Use a continuidade da função $t \mapsto t^2$ para mostrar que o conjunto $A := \{x \in \mathbb{R} : x > 0 \text{ e } x^2 > 2\}$ não tem mínimo. Analogamente, mostre que $B := \{x \in \mathbb{R} : x > 0 \text{ e } x^2 < 2\}$ não tem máximo. Compare com a prova do Lema 0.9.3. ■

Exercício 1.202 (Teste da derivada para máximos locais — $\binom{\star}{\star}$)**.** Um ponto $a \in X$ é dito **máximo local** de $f\colon X \to \mathbb{R}$ se existe um subconjunto $V \subseteq X$ aberto em X com $a \in V$ tal que $f(a) = \max\{f(v) : v \in V\}$. A definição de **mínimo local** é análoga. Diremos que a é **extremo local** se a for ponto de máximo ou de mínimo local[†]. Supondo que a função $f\colon X \to \mathbb{R}$ seja diferenciável em $p \in X$, onde p é **ponto de acumulação bilateral**[‡] de X, mostre que se p for extremo local de f, então $f'(p) = 0$. Dica: supondo $f'(p) \neq 0$, use "conservação de sinal" para investigar o comportamento de $\dfrac{f(x) - f(p)}{x - p}$ para $x \neq p$ tanto à direita quanto à esquerda de p. ■

[†] O adjetivo "local" costuma ser omitido ou trocado por "**absoluto**" na ocorrência de $V = X$.

[‡] Isto é, p é simultaneamente ponto de acumulação de X pela esquerda e pela direita.

Essenciais: mais conceitos topológicos na reta

Exercício 1.203 $\binom{\star}{\star}$**.** Para $X \subseteq \mathbb{R}$, mostre que $p \in [-\infty, +\infty]$ é ponto de acumulação de X se, e somente se, existe sequência $(x_n)_{n\in\mathbb{N}}$ em $X \setminus \{p\}$ tal que $x_n \to p$. Além disso:

a) se p é ponto de acumulação de X pela esquerda, então podemos supor $x_n < p$ para todo n;

b) se p é ponto de acumulação de X pela direita, então podemos supor $p < x_n$ para todo n. ■

Exercício 1.204 $\binom{\star}{\star}$**.** Seja $(x_n)_{n\in\mathbb{N}}$ sequência em $\mathbb{R}$. Mostre que se L é ponto de acumulação do conjunto $\{x_n : n \in \mathbb{N}\}$, então $(x_n)_{n\in\mathbb{N}}$ admite subsequência que converge para L. Vale a recíproca? ■

Exercício 1.205 $\binom{\star}{\star}$**.** Para $X \subseteq \mathbb{R}$, dizemos que $p \in X$ é **ponto isolado em** X se existe um intervalo aberto $I \subseteq \mathbb{R}$ tal que $p \in I$ e $I \cap X = \{p\}$. Um subconjunto $S \subseteq \mathbb{R}$ é chamado de **discreto** se todo ponto de S é isolado em S.

a) Mostre que se $p \in X$ é isolado em X, então qualquer função $f : X \to \mathbb{R}$ é contínua em p. Dica: use a definição de continuidade em termos de intervalos abertos.

b) Mostre que se X é discreto, então qualquer função $f : X \to \mathbb{R}$ é contínua.

c) Mostre que se $S \subseteq \mathbb{R}$ é discreto, então S é, no máximo, enumerável. Dica: para $x \in S$, mostre que existem racionais p_x, q_x tais que $p_x < x < q_x$ e $(p_x, q_x) \cap X = \{x\}$; observe então que se $x, y \in X$ com $x \neq y$, então $p_x \neq p_y$. ■

Exercício 1.206 $\binom{\star}{\star}$**.** Mostre que se $S \subseteq \mathbb{R}$ é discreto, então todo subconjunto $O \subseteq S$ é aberto em S. ■

Exercício 1.207 $\binom{\star}{\star}$**.** Dado um subconjunto discreto $S \subseteq \mathbb{R}$, é possível que exista $x \in \mathbb{R}$ que seja ponto de acumulação de S? É possível que $x \in S$? ■

Exercício 1.208 (IMPORTANTE — $\binom{\star}{\star}$**).** Dizemos que um subconjunto $F \subseteq \mathbb{R}$ é **fechado em** $\mathbb{R}$ se $\mathbb{R} \setminus F$ é um subconjunto aberto de $\mathbb{R}$. Mostre que para $F \subseteq \mathbb{R}$, são equivalentes:

(i) F é fechado em $\mathbb{R}$;

(ii) para qualquer sequência $(x_n)_{n\in\mathbb{N}}$ em F ocorre $x \in F$ sempre que $x_n \to x$. ■

Exercício 1.209 $(\star)$**.** Mostre que:

a) intervalos fechados de $\mathbb{R}$ são fechados em $\mathbb{R}$;

b) $\{x\}$ é fechado em $\mathbb{R}$ para todo $x \in \mathbb{R}$. ■

Exercício 1.210 $\binom{\star}{\star}$**.** Mostre que:

a) $\emptyset$ e $\mathbb{R}$ são fechados em $\mathbb{R}$;

b) $F \cup G$ é fechado em $\mathbb{R}$ sempre que $F, G \subseteq \mathbb{R}$ são fechados em $\mathbb{R}$;

c) $\bigcap_{\lambda\in\Lambda} F_\lambda$ é fechado em $\mathbb{R}$ sempre que $\{F_\lambda : \lambda \in \Lambda\}$ é família não vazia de subconjuntos fechados de $\mathbb{R}$.

Dica: use o Lema 0.4.14 para *dualizar* o que você sabe a respeito dos abertos de $\mathbb{R}$. ■

Exercício 1.211 $({}^{\star}_{\star})$**.** Para $S \subseteq \mathbb{R}$ e $p \in \mathbb{R}$, dizemos que p é **ponto aderente de** S (ou **a** S) se $V \cap S \neq \emptyset$ para todo subconjunto aberto $V \subseteq \mathbb{R}$ tal que $p \in V$. Mostre que p é ponto aderente de S se, e somente se, existe sequência $(x_n)_{n\in\mathbb{N}}$ em S tal que $x_n \to p$. ■

Exercício 1.212 $({}^{\star}_{\star})$**.** O **fecho** de um subconjunto $S \subseteq \mathbb{R}$ em $\mathbb{R}$, denotado por $\overline{S}$, é a coleção de todos os pontos aderentes a S.

a) Mostre que $\overline{S}$ é fechado em $\mathbb{R}$.

b) Mostre que $S \subseteq \overline{S}$.

c) Mostre que se $F \subseteq \mathbb{R}$ é fechado e $S \subseteq F$, então $\overline{S} \subseteq F$.

d) Mostre que S é fechado em $\mathbb{R}$ se, e somente se, $\overline{S} = S$. ■

Exercício 1.213 $(\star)$**.** Para $a, b \in \mathbb{R}$, mostre que $\overline{(a,b)} = [a,b]$. ■

Exercício 1.214 $(\star)$**.** Mostre que $\overline{\mathbb{Q}} = \mathbb{R}$ e $\overline{\mathbb{R} \setminus \mathbb{Q}} = \mathbb{R}$. ■

Exercício 1.215 $({}^{\star}_{\star})$**.** Para um subconjunto $S \subseteq \mathbb{R}$ e um ponto $p \in \mathbb{R}$, dizemos que p é **ponto interior de** S (ou **a** S) se existe um aberto $V \subseteq \mathbb{R}$ tal que $p \in V$ e $V \subseteq S$. O **interior** de um conjunto S, denotado por $\operatorname{int}(S)$, é a coleção de seus pontos interiores.

a) Mostre que $\operatorname{int}(S)$ é aberto em $\mathbb{R}$.

b) Mostre que $\operatorname{int}(S) \subseteq S$.

c) Mostre que se $A \subseteq S$ com A aberto em $\mathbb{R}$, então $A \subseteq \operatorname{int}(S)$.

d) Mostre que S é aberto em $\mathbb{R}$ se, e somente se, $\operatorname{int}(S) = S$. ■

Exercício 1.216 $({}^{\star}_{\star\star})$**.** Para $S \subseteq \mathbb{R}$, mostre que $\operatorname{int}(S) = \mathbb{R} \setminus \overline{\mathbb{R} \setminus S}$. ■

Exercício 1.217 $({}^{\star}_{\star})$**.** Sejam $S, T \subseteq \mathbb{R}$ subconjuntos quaisquer.

a) Mostre que se $S \subseteq T$, então $\operatorname{int}(S) \subseteq \operatorname{int}(T)$ e $\overline{S} \subseteq \overline{T}$. Dica: prove apenas uma das inclusões e use o exercício anterior para obter a outra.

b) Mostre que $\operatorname{int}(S \cap T) = \operatorname{int}(S) \cap \operatorname{int}(T)$ e $\overline{S \cup T} = \overline{S} \cup \overline{T}$. Dica: prove apenas uma das identidades e use o exercício anterior para obter a outra. ■

Exercício 1.218 $({}^{\star}_{\star})$**.** Para $X \subseteq \mathbb{R}$, dizemos que $D \subseteq X$ é **denso em** X se todo aberto não vazio de X contém pelo menos um ponto de D.

a) Mostre que $D \subseteq \mathbb{R}$ é denso em $\mathbb{R}$ se, e somente se, $\overline{D} = \mathbb{R}$.

b) Dê exemplos de subconjuntos densos de $\mathbb{R}$.

c) Mostre que $D \subseteq X$ é denso se, e somente se, para todo $x \in X$ existe uma sequência $(x_n)_{n\in\mathbb{N}}$ de pontos em D tal que $x_n \to x$.

d) Mostre que se $D \subseteq X$ é denso e $f, g \colon X \to \mathbb{R}$ são funções contínuas tais que $f(y) = g(y)$ para todo $y \in D$, então $f = g$. ■

Exercício 1.219 $\left(\substack{\star\\\star}\right)$**.** Mostre que $f\colon \mathbb{R} \to \mathbb{R}$ é contínua se, e somente se, $f^{-1}[G]$ é fechado em $\mathbb{R}$ sempre que $G \subseteq \mathbb{R}$ é fechado em $\mathbb{R}$. ■

Exercício 1.220 $\left(\substack{\star\\\star}\right)$**.** Para $X \subseteq \mathbb{R}$ e $S \subseteq X$, dizemos que S é **fechado em** X se $X \setminus S$ é aberto em X. Mostre que se $F \subseteq \mathbb{R}$ é fechado em $\mathbb{R}$, então $S \subseteq F$ é fechado em F se, e somente se, é fechado em $\mathbb{R}$. Dica: se precisar de cola, confira a demonstração do Corolário 2.0.19. ■

Exercício 1.221 $\left(\substack{\star\\\star}\right)$**.** Mostre que $f\colon X \to \mathbb{R}$ é contínua se, e somente se, $f^{-1}[G]$ é fechado em X sempre que $G \subseteq \mathbb{R}$ é fechado em $\mathbb{R}$. ■

Exercício 1.222 $(\star)$**.** Classifique as afirmações a seguir como VERDADEIRAS ou FALSAS e justifique.

a) $\mathbb{Q}$ é fechado em $\mathbb{R}$.

b) $\mathbb{Q}$ é aberto em $\mathbb{R}$.

c) $\mathbb{R} \setminus \mathbb{Q}$ é aberto em $\mathbb{R}$.

d) $[0, 2)$ é aberto em $[0, 3]$ mas não é aberto em $\mathbb{R}$.

e) $(0, 3]$ é fechado em $(0, 4)$.

f) Todo subconjunto finito de $\mathbb{R}$ é fechado em $\mathbb{R}$.

g) Se um subconjunto de $\mathbb{R}$ não é aberto, então é fechado.

h) Todo ponto de acumulação de um conjunto é aderente ao conjunto.

i) Todo ponto aderente a um conjunto é de acumulação do conjunto.

j) $\mathbb{R} \setminus \mathbb{Q}$ não tem pontos aderentes.

k) Todo $x \in \mathbb{R}$ é ponto de acumulação de $\mathbb{R} \setminus \mathbb{Q}$.

l) Um subconjunto enumerável de $\mathbb{R}$ tem interior vazio.

m) Todo subconjunto de $\mathbb{R}$ não enumerável tem ponto interior. ■

Observação 1.9.0. Eu sei que é difícil, mas resista à tentação de achar que a afirmação "x é ponto interior de S" seja sinônimo de "$x \in S$". Existe uma definição precisa para o que significa a primeira afirmação: confira o Exercício 1.215. Como sempre, fica o alerta: se uma expressão tem uma definição dentro de um contexto matemático, então é ela que você deve respeitar, e não (a sua interpretação d) o dicionário. △

Exercício 1.223 $\left(\substack{\star\\\star\star}\right)$**.** Mostre que $S \subseteq \mathbb{R}$ é fechado em $\mathbb{R}$ e discreto se, e somente se, S não tem pontos de acumulação. Dê exemplos de conjuntos infinitos satisfazendo tal condição. ■

Exercício 1.224 $\left(\substack{\star\\\star}\right)$**.** Para $S \subseteq \mathbb{R}$ um subconjunto qualquer, mostre que S é fechado em $\mathbb{R}$ se, e somente se, toda sequência de Cauchy em S converge para algum ponto de S. ■

Exercício 1.225 $(\star)$**.** Mostre que para todo $x \in \mathbb{R}$, $\mathbb{R} \setminus \{x\}$ é aberto e denso em $\mathbb{R}$. ■

Exercício 1.226 $\left(\substack{\star\\ \star}\right)$**.** Mostre que $S \subseteq \mathbb{R}$ é aberto em $\mathbb{R}$ e denso se, e somente se, $\mathbb{R} \setminus S$ é fechado em $\mathbb{R}$ e tem interior vazio. Dica: pode ser útil conferir o Exercício 1.216. ■

Exercício 1.227 $\left(\substack{\star\\ \star\star}\right)$**.** Mostre que se $A_n \subseteq \mathbb{R}$ é aberto e denso para cada $n \in \mathbb{N}$, então $\bigcap_{n\in\mathbb{N}} A_n$ é denso. Dica: mostre que $V \cap \left(\bigcap_{n\in\mathbb{N}} A_n\right) \neq \emptyset$ para qualquer intervalo aberto $V \neq \emptyset$; para isso, obtenha uma sequência de pontos $(x_n)_n$ em V, bem como uma sequência $(r_n)_n$ de números reais positivos tais que, entre outras coisas, $[x_{n+1} - r_{n+1}, x_{n+1} + r_{n+1}] \subseteq A_{n+1} \cap (x_n - r_n, x_n + r_n)$ e $r_n < \frac{1}{2^n}$ para todo n. ■

Exercício 1.228 $\left(\substack{\star\\ \star\star}\right)$**.** Mostre que se $F_n \subseteq \mathbb{R}$ é fechado em $\mathbb{R}$ e tem interior vazio para cada $n \in \mathbb{N}$, então $\bigcup_{n\in\mathbb{N}} F_n$ tem interior vazio. Dica: você, possivelmente, acabou de resolver este exercício. ■

Exercício 1.229 $(\star)$**.** Mostre que existe família $\{A_n : n \in \mathbb{N}\}$ de abertos densos de $\mathbb{R}$ tal que $\bigcap_{n\in\mathbb{N}} A_n = \mathbb{R} \setminus \mathbb{Q}$. ■

Exercício 1.230 $\left(\substack{\star\\ \star}\right)$**.** Mostre que $\mathbb{Q}$ não é interseção enumerável de abertos em $\mathbb{R}$. Dica: note que todo subconjunto de $\mathbb{R}$ que contém $\mathbb{Q}$ também é denso. ■

Exercício 1.231 $\left(\substack{\star\\ \star\star}\right)$**.** Mostre que não existe família enumerável $\{F_n : n \in \mathbb{N}\}$ de fechados de $\mathbb{R}$ tais que $\bigcup_{n\in\mathbb{N}} F_n = \mathbb{R} \setminus \mathbb{Q}$. Dica: se fosse, o exercício anterior estaria errado. ■

Exercício 1.232 $\left(\substack{\star\\ \star}\right)$**.** Para $S \subseteq \mathbb{R}$ e $x \in \mathbb{R}$, mostre que $x \in \overline{S}$ se, e somente se, $d(x, S) = 0$, onde $d(x, S) := \inf\{|x - y| : y \in S\}$. ■

Exercício 1.233 $\left(\substack{\star\\ \star\star}\right)$**.** Para $F, G \subseteq \mathbb{R}$ fechados não vazios e disjuntos, mostre que existe função contínua $h\colon \mathbb{R} \to [0, 1]$ tal que $h(x) = 0$ para todo $x \in F$ e $h(x) = 1$ para todo $x \in G$. Dica: para f e g tais que $f + g \neq 0$, faz sentido escrever $\frac{f}{f+g}$. ■

Exercício 1.234 $\left(\substack{\star\\ \star\star}\right)$**.** Mostre que o espaço vetorial $\mathcal{C}(\mathbb{R})$, das funções contínuas da forma $\mathbb{R} \to \mathbb{R}$, tem dimensão infinita. ■

Extras em dimensão 1

Exercício 1.235 (Teorema de Merten sobre produtos de Cauchy — $\left(\substack{\star\\ \star\star}\right)$)**.** Sejam $\sum_{n=0}^{\infty} a_n := A$ e $\sum_{n=0}^{\infty} b_n := B$ séries de números reais, ambas convergentes em $\mathbb{R}$, com pelo menos uma delas absolutamente convergente. Para cada $k \in \mathbb{N}$, seja $c_k := \sum_{l=0}^{k} a_l b_{k-l}$. Por fim, seja $\varepsilon > 0$.

a) Chamando $A_n := \sum_{i=0}^{n} a_i$, $B_n := \sum_{i=0}^{n} b_i$ e $C_n := \sum_{i=0}^{n} c_i$ para cada $n \in \mathbb{N}$, mostre que $C_n = \sum_{i=0}^{n} a_{n-i} B_i$ e, por conseguinte, $C_n = \left(\displaystyle\sum_{i=0}^{n} a_{n-i}(B_i - B)\right) + A_n B$. Dica: faça alguns casos particulares antes de provar por indução.

b) Assumindo que $\sum_{n=0}^{\infty} a_n$ converge absolutamente, mostre que existe $N \in \mathbb{N}$ tal que para qualquer $n \geq N$ se verifique

$$|B_n - B| \leq \frac{\varepsilon}{3} \cdot \frac{1}{1 + \displaystyle\sum_{n=0}^{\infty} |a_n|}.$$

c) Mostre que existe $M \in \mathbb{N}$ tal que $|a_n| \leq \dfrac{\varepsilon}{3N \max\{|B_i - B| + 1 : i \leq N-1\}}$ para todo $n \geq M$.

d) Mostre que existe $L \in \mathbb{N}$ tal que $|A_n - A| \leq \dfrac{\varepsilon}{3} \cdot \dfrac{1}{|B|+1}$ sempre que $n \geq L$.

e) Conclua que $C_n \to AB$. Dica: desigualdade triangular + itens anteriores. ■

Observação 1.9.1. O exercício acima lida com o problema do *produto de séries*. Note que o produto de duas expressões polinomiais $a_0 + a_1 t + \ldots + a_n t^n$ e $b_0 + b_1 t + \ldots + b_m t_m$ é dado por

$$a_0 b_0 + (a_0 b_1 + b_0 a_1)t + (a_0 b_2 + a_1 b_1 + b_0 a_2)t^2 + \ldots$$

de modo que o coeficiente que acompanha o termo t^k é dado por

$$c_k := a_0 b_k + a_1 b_{k-1} + a_2 b_{k-2} + \ldots + a_0 b_k = \sum_{l=0}^{k} a_l b_{k-l}.$$

Nesse sentido, o exercício acima nos mostra que quando uma das séries é absolutamente convergente, o produto de ambas se comporta como se fosse o produto de duas expressões polinomiais infinitas. △

Exercício 1.236 (Função exponencial revisitada — $\left(\substack{\star \\ \star\star}\right)$)**.** Considere a função $\exp(x) := e^x$ definida no Exemplo 1.6.12 (cf. Observação 1.6.13).

a) Use o exercício anterior para mostrar que $e^{x+y} = e^x \cdot e^y$ para quaisquer $x, y \in \mathbb{R}$. Dica: note que e^x e e^y são séries absolutamente convergentes, e assim $e^x e^y$ se expressa como a série $\sum_{n=0}^{\infty} c_n$ do exercício anterior; calcule c_n explicitamente[†].

b) Com o item anterior, mostre que $e^x > 0$ para todo $x \in \mathbb{R}$. Além disso, mostre que exp é função estritamente crescente.

c) Mostre que $\lim\limits_{x \to 0} \dfrac{e^x - 1}{x} = 1$ e conclua que $\exp'(x) = \exp(x)$ para todo $x \in \mathbb{R}$. Dica: mostre que $1 + x \leq e^x \leq \dfrac{1}{1-x}$ para todo $x \in (0,1)$, e daí use os itens anteriores para mostrar que tal desigualdade vale para todo x tal que $|x| < 1$.

d) Mostre que $\lim\limits_{x \to +\infty} e^x = +\infty$ e $\lim\limits_{x \to -\infty} e^x = 0$. Dica: para o primeiro limite, observe que $e > 1$, $e^2 = e \cdot e > 2$, $e^3 > 3$, etc.; para o segundo limite, use mudança de variáveis. ■

Exercício 1.237 $\left(\substack{\star \\ \star}\right)$**.** Seja $f \colon I \to \mathbb{R}$ uma função monótona e limitada, onde $I \subseteq \mathbb{R}$ é um intervalo aberto.

a) Mostre que $f(p^-) := \lim\limits_{x \to p^-} f(x)$ e $f(p^+) := \lim\limits_{x \to p^+} f(x)$ existem e são números reais para todo $p \in I$. Dica: isto já foi feito, embora você possa ter ignorado (cf. Proposição 1.1.18).

b) Mostre que $f(p^-) \leq f(p) \leq f(p^+)$ para todo $p \in I$. Dica: monotonicidade.

[†] Você precisará do Binômio de Newton.

c) Assumindo que f é crescente, mostre que se $p, q \in I$ e $p < q$, então $f(p^+) \leq f(q^-)$. Adapte o resultado para o caso decrescente. Dica: use o Exercício 1.203 para escrever $f(p^+)$ e $f(q^-)$ como limites de *sequências*, e daí conclua por monotonicidade.

d) Seja $T \subseteq I$ o conjunto dos pontos de I em que f é descontínua. Mostre que T é, no máximo, enumerável. Dica: se $p \in T$, então $f(p^-) < f(p^+)$, o que permite escolher um número racional $r_p \in (f(p^-), f(p^+))$.

e) Mostre que o resultado do item anterior permanece válido para intervalos fechados. ■

Exercício 1.238 (Colagem de funções contínuas — $(\substack{\star\\\star})$)**.** Sejam A e B subconjuntos de $X \subseteq \mathbb{R}$ tais que $X = A \cup B$, e considere $f \colon X \to \mathbb{R}$ uma função.

a) Mostre que se A e B são abertos em X e $f|_A$ e $f|_B$ são contínuas, então f é contínua.

b) Mostre que se A e B são fechados em X e $f|_A$ e $f|_B$ são contínuas, então f é contínua[†].

c) Mostre que sem as hipóteses anteriores o resultado é falso. ■

Exercício 1.239 (Ponto fixo de Banach, disfarçado — $(\substack{\star\\\star\star})$)**.** Seja $X \subseteq \mathbb{R}$ um subconjunto não vazio e fechado em $\mathbb{R}$ e considere uma função $f \colon X \to X$. Se existir $K \in (0, 1)$ tal que $|f(x) - f(y)| < K|x - y|$ para quaisquer $x, y \in X$, mostre que existe um único $p \in X$ tal que $f(p) = p$. Dica: supondo que p existe, argumente via contradição para provar sua unicidade; para assegurar a existência, escolha $p_0 \in X$ e defina $p_1 := f(p_0)$, $p_2 := f(p_1) \ldots$ note que se $p := \lim_{n\to\infty} p_n$ existir, então $f(p) = p$, de modo que resta apenas mostrar que $(p_n)_{n\in\mathbb{N}}$ converge — o que, por completude, equivale a provar que $(p_n)_n$ é de Cauchy. ■

Extras em espaços topológicos e afins

Exercício 1.240 $(\substack{\star\\\star})$**.** Num espaço topológico $(X, \mathcal{T})$, $F \subseteq X$ é **fechado** em X se $X \setminus F$ é um aberto de X. Mostre que

a) $\emptyset$ e X são fechados em X,

b) $F \cup G$ é fechado sempre que $F, G \subseteq X$ são fechados,

c) $\bigcap_{\lambda\in\Lambda} F_\lambda$ é fechado sempre que $\{F_\lambda : \lambda \in \Lambda\}$ é família não vazia de subconjuntos fechados de X.

Dica: use o Lema 0.4.14 para *dualizar* tudo o que você sabe a respeito dos abertos de um espaço topológico. ■

Exercício 1.241 $(\substack{\star\\\star})$**.** Para X e Y espaços topológicos, mostre que $f \colon X \to Y$ é contínua se, e somente se, $f^{-1}[G]$ é fechado em X sempre que $G \subseteq Y$ é fechado. ■

Exercício 1.242 $(\substack{\star\\\star})$**.** Para X espaço topológico e $F \subseteq X$ fechado, mostre que $S \subseteq F$ é fechado em F (com a topologia de subespaço) se, e somente se, S é fechado em X. Dica: se sua solução para o Exercício 1.220 for "boa", ela servirá para o presente exercício. ■

[†]Um subconjunto $S \subseteq X$ é fechado em X se $X \setminus S$ é aberto em X.

Exercício 1.243 $\binom{\star}{\star\star}$**.** Adapte as definições de ponto aderente, fecho, ponto interior, interior e densidade para espaços métricos ou topológicos. Depois, adapte os resultados provados em $\mathbb{R}$ para o contexto que você escolheu. Dica: os resultados que envolvem sequências permanecem válidos em espaços métricos, mas devem ser substituídos por redes caso você tenha optado por espaços topológicos. ■

Exercício 1.244 $\binom{\star}{\star}$**.** Adapte o Exercício 1.193 para espaços métricos. ■

Exercício 1.245 $\binom{\star}{\star}$**.** Mostre que o Exercício 1.234 permanece válido para qualquer espaço métrico infinito, i.e.: mostre que se X é métrico e infinito, então o espaço vetorial das funções contínuas da forma $X \to \mathbb{R}$ tem dimensão infinita. Dica: note que os exercícios que antecedem o supracitado problema também valem para espaços métricos. ■

Exercício 1.246 $\binom{\star}{\star}$**.** Num espaço métrico (X, d), definimos a d**-bola fechada de centro** $x \in X$ **e raio** $r > 0$ como sendo o conjunto

$$B_d[x, r] \coloneqq \{y \in X : d(x, y) \leq r\}.$$

Mostre que $B_d[x, r]$ é fechado em (X, d). Dica: você já tem tecnologia para resolver isto numa linha (cf. Exercícios 1.244, 1.241 + o fato de $[0, r]$ ser fechado em $\mathbb{R}$). ■

Exercício 1.247 $\binom{\star}{\star}$**.** Pense rápido: num espaço métrico (X, d) qualquer, vale a identidade $\overline{B_d(x, r)} = B_d[x, r]$? E se, em vez de métrico, tivermos um espaço vetorial normado? ■

Exercício 1.248 $\binom{\star}{\star}$**.** Dizemos que um subconjunto C de um espaço vetorial E é **convexo** se para quaisquer $t \in [0, 1]$ e $x, y \in C$ valer $tx + (1 - t)y \in C$. Mostre que se E é normado e C é convexo, então $\overline{C}$ é convexo. ■

Exercício 1.249 $\binom{\star}{\star}$**.** Mostre que $\mathbb{R}^2$ e $C \coloneqq \{(x, y, x^2 + y^2) : x, y \in \mathbb{R}\}$ são homeomorfos. Dica: mostre que a bijeção "óbvia" entre $\mathbb{R}^2$ e C é contínua e tem inversa contínua. ■

Exercício 1.250 $\binom{\star}{\star}$**.** Mostre que se B é uma bola aberta num espaço normado $(E, \|\cdot\|)$, então B e E são homeomorfos. Dica: imite o Exemplo 1.7.17. ■

Exercício 1.251 $\binom{\star}{\star}$**.** Mostre que se B e B' são bolas abertas num espaço normado, então B e B' são homeomorfas. ■

Exercício 1.252 $\binom{\star}{\star}$**.** Os dois exercícios anteriores permanecem válidos em espaços métricos? Dica: investigue a métrica discreta num conjunto com pelo menos dois pontos. ■

Exercício 1.253 $(\star)$**.** Mostre que se duas normas num espaço vetorial induzem a mesma topologia, então as bolas abertas com respeito a uma norma são homeomorfas às bolas abertas com respeito a outra norma. ■

Exercício 1.254 $(\star)$**.** Mostre que a *bola fechada* $B \coloneqq \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$ é homeomorfa ao *quadrado fechado* $Q \coloneqq \{(x, y) \in \mathbb{R}^2 : \max\{|x|, |y|\} \leq 1\}$. ■

Exercício 1.255 $(\star)$**.** Repita o exercício anterior trocando "$\leq$" por "$<$" e "$=$". ■

Exercício 1.256 $\binom{\star}{\star}$**.** Mostre que subespaços vetoriais próprios de espaços normados têm interior vazio. Dica: primeiro, faça um desenho como se isto fosse um problema de transladar vetores em Geometria Analítica. ■

Exercício 1.257 (Cf. Exercício 1.224 — $\binom{\star}{\star}$)**.** Para X espaço métrico e $F \subseteq X$, mostre que:

a) se F é completo com a métrica de subespaço, então F é fechado;

b) se X é completo e F é fechado, então F é completo. ■

Exercício 1.258 $(^{\star}_{\star\star})$**.** Para um espaço métrico (X, d), defina $d' \colon X \times X \to [0, 1]$ fazendo

$$d'(x, y) := \begin{cases} d(x, y), & \text{se } d(x, y) \leq 1 \\ 1, & \text{se } d(x, y) > 1 \end{cases}.$$

Mostre que (X, d') é espaço métrico homeomorfo a (X, d). ■

Exercício 1.259 $(^{\star}_{\star\star})$**.** Adapte o Exercício 1.238 para espaços topológicos quaisquer. ■

Exercício 1.260 $(^{\star}_{\star\star})$**.** Dizemos que uma função $f \colon \mathbb{R} \to \mathbb{R}$ tem **gráfico fechado** se para quaisquer $x, y \in \mathbb{R}$, a ocorrência de $x_n \to x$ e $f(x_n) \to y$ para uma sequência real $(x_n)_n$ for suficiente para garantir que $f(x) = y$.

a) Mostre que se f é contínua, então f tem gráfico fechado.

b) Mostre que se f tem gráfico fechado e f é limitada, então f é contínua. Dica: suponha que não, daí use o Exercício 1.102 e o Teorema de Bolzano-Weierstrass para concluir.

c) Mostre que a hipótese de limitação é indispensável. Dica: $\lim\limits_{x \to 0} \dfrac{1}{|x|} = +\infty$. ■

Exercício 1.261 (Teorema de Baire — $(^{\star}_{\star\star})$)**.** Um espaço topológico X é chamado de (**espaço de**) **Baire** se a tese do Exercício 1.227 é satisfeita em X. Mostre que todo espaço métrico completo é espaço de Baire. Dica: a `label` do Exercício 1.227 é `hiddenbaire`. ■

Exercício 1.262 $(^{\star}_{\star\star})$**.** Mostre que se um espaço normado tem dimensão infinita enumerável, então o espaço não é de Banach. Dica: entre outras coisas, use o Exercício 1.256 juntamente com o Teorema de Baire. ■

Exercício 1.263 (Ponto fixo de Banach sem máscaras — $(^{\star}_{\star})$)**.** Adapte o Exercício 1.239 para espaços métricos completos. ■

Capítulo 2

Os teoremas fundamentais da Análise

Os teoremas que serão discutidos neste capítulo são *fundamentais* para a Análise Real. Com isso dito, não pense que os teoremas anteriores são *menos importantes*. Na verdade, como tudo o que discutiremos se *fundamenta* no que se apresentou previamente, segue que os resultados anteriores podem ser considerados ainda *mais* fundamentais. Feita a ressalva, o título se justifica pelo seguinte: os teoremas que veremos, além de importantes, ajudam a revelar os *aspectos fundamentais* da Análise, no sentido de serem as noções que verdadeiramente permitem desenvolver e generalizar a teoria: *compacidade*, *conexidade* e *completude*.

2.0 Teorema de Heine-Borel-Lebesgue

2.0.0 Essencial

§0 Compacidade e o Teorema de Heine-Borel-Lebesgue

Vamos começar a discussão com um problema inocente: existe uma função contínua e ilimitada $[0,1) \to \mathbb{R}$? Intuitivamente, se existisse, seria como se pudéssemos desenhar uma linha de comprimento infinito e sem tirar o lápis do papel em momento algum. Embora, na vida real, papel e lápis sejam recursos finitos, isto não nos impede de imaginar que enquanto objeto abstrato, existe uma função em tais condições.

Exercício 2.0 ($\star$)**.** Convença-se de que $f\colon [0,1) \to \mathbb{R}$ dada por $f(x) := \frac{x}{1-x}$ serve. ■

Aceitar o exemplo anterior impede que você apele para a finitude dos recursos naturais para tratar do verdadeiro problema: se, no lugar do intervalo $[0,1)$, considerarmos o intervalo $[0,1]$, a resposta para a pergunta original passa a ser negativa. Por quê?

Definição 2.0.0. Dizemos que $\mathcal{U}$ é uma **cobertura aberta** para um subconjunto $X \subseteq \mathbb{R}$ se $\mathcal{U}$ for uma família de subconjuntos abertos de $\mathbb{R}$ tal que para todo $x \in X$ exista $U \in \mathcal{U}$ tal que $x \in U$. O subconjunto $X \subseteq \mathbb{R}$ é chamado de **compacto** se toda cobertura aberta de X admite subcobertura finita, i.e., se existe $\mathcal{V} \subseteq \mathcal{U}$, com $\mathcal{V}$ finito e $X \subseteq \bigcup_{V \in \mathcal{V}} V$. ¶

Exemplo 2.0.1. *A reta não é compacta.* De fato, ao considerar a *coleção*[†] de intervalos abertos $\mathcal{U} := \{(-n, n) : n \in \mathbb{N} \setminus \{0\}\}$, segue que para todo $x \in \mathbb{R}$ existe algum aberto $(-n, n) \in \mathcal{U}$ tal que $x \in (-n, n)$. No entanto, nenhum subconjunto finito de $\mathcal{U}$ é cobertura para $\mathbb{R}$: afinal de contas para $n_0, \ldots, n_m \in \mathbb{N} \setminus \{0\}$ quaisquer, sempre existe um número real $x \in \mathbb{R} \setminus ((-n_0, n_0) \cup (-n_1, n_1) \cup \ldots \cup (-n_m, n_m))$. ▲

Exercício 2.1 $({}^{\star}_{\star})$**.** Justifique a última afirmação. ■

O exemplo anterior sugere que subconjuntos ilimitados de $\mathbb{R}$ não podem ser compactos, o que na contrapositiva se lê assim:

Proposição 2.0.2. *Se $K \subseteq \mathbb{R}$ é compacto, então K é limitado.*

Exercício 2.2 $({}^{\star}_{\star})$**.** Prove a proposição acima. ■

Todavia, limitação não basta para assegurar compacidade.

Exemplo 2.0.3. *O intervalo aberto $(0, 1)$ não é compacto.* Neste caso, note que

$$\mathcal{U} := \left\{ \left(0, 1 - \frac{1}{2^n}\right) : n \in \mathbb{N} \setminus \{0\} \right\}$$

é uma cobertura aberta para $(0, 1)$: para $r \in (0, 1)$, sabemos que existe $N \in \mathbb{N}$ tal que $\left|1 - \frac{1}{2^N} - 1\right| < 1 - r$, e daí $r < 1 - \frac{1}{2^N}$. Ainda assim, não há subconjunto finito de $\mathcal{U}$ capaz de cobrir $(0, 1)$: para $k_0, \ldots, k_n \in \mathbb{N} \setminus \{0\}$ quaisquer, basta tomar $K := \max\{k_0, \ldots, k_n\}$ e notar que $\left(0, 1 - \frac{1}{2^{k_i}}\right) \subseteq \left(0, 1 - \frac{1}{2^K}\right)$, de modo que $1 - \frac{1}{2^{K+1}} \in (0, 1)$ mas não pertence a nenhum dos intervalos $\left(0, 1 - \frac{1}{2^{k_i}}\right)$. Faça um desenho se julgar necessário $(\star)$. ▲

Exercício 2.3 $({}^{\star}_{\star\star})$**.** A argumentação apresentada no exemplo anterior é desnecessária em virtude do que se discutiu no Exemplo 2.0.1. Por quê?! ■

A sugestão do último exemplo é um pouco mais sutil: lembre-se de que $F \subseteq \mathbb{R}$ é *fechado* em $\mathbb{R}$ se $\mathbb{R} \setminus F$ é aberto em $\mathbb{R}$ (cf. Exercício 1.208).

Lema 2.0.4. *Se $K \subseteq \mathbb{R}$ é compacto, então K é fechado em $\mathbb{R}$.*

Demonstração. Mostraremos que $\mathbb{R} \setminus K$ é aberto e, para isso, usaremos o Lema 1.7.3: para $x \in \mathbb{R} \setminus K$, vamos encontrar um subconjunto $B \subseteq \mathbb{R}$ aberto em $\mathbb{R}$ tal que $x \in B$ e $B \subseteq \mathbb{R} \setminus K$. Como nada precisa ser feito se $K = \emptyset$ (certo?)*, podemos supor $K \neq \emptyset$. Agora, com o ponto $x \in \mathbb{R} \setminus K$ fixado, para cada $y \in K$ podemos encontrar subconjuntos abertos em $\mathbb{R}$, digamos $A_y, B_y \subseteq \mathbb{R}$, tais que $x \in B_y$, $y \in A_y$ e $\underline{A_y \cap B_y = \emptyset}$ (certo?)*. Por K ser compacto e $\mathcal{U} := \{A_y : y \in K\}$ ser uma cobertura aberta para K, existem $y_0, \ldots, y_n \in K$ tais que $K \subseteq A_{y_0} \cup \ldots \cup A_{y_n}$. Para encerrar, vamos ver que $B := \bigcap_{i \leq n} B_{y_i}$ cumpre o que se pede:

✓ B é aberto por ser interseção finita de abertos;

✓ $x \in B$ pois $x \in B_{y_i}$ para todo $i \leq n$;

✓ $B \subseteq \mathbb{R} \setminus K$ pois, se $z \in K$, então $z \in A_{y_i}$ para algum $i \leq n$, o que impede que z pertença a B_{y_i} (percebeu?)* e, portanto, $z \notin B$. □

[†] Já passou da hora de você parar de encrencar com o uso da palavra "conjunto" e "coleção" para designar conjuntos cujos elementos são outros conjuntos.

Exercício 2.4 (?!). Por que foi preciso tomar B como uma interseção *finita* de abertos? Por que não fazer $B := B_y$ para *algum* $y \in K$? ■

Num primeiro momento, os resultados acima não nos dizem quem são os compactos de $\mathbb{R}$, mas o contrário: quais subconjuntos não podem ser. Por exemplo, $(0,1)$ não pode ser compacto pois não é fechado (embora seja limitado): existe sequência $(x_n)_{n\in\mathbb{N}}$ em $(0,1)$ com $x_n \to 1$, mas $1 \notin (0,1)$ (cf. Exercício 1.208); $[0,+\infty)$ não pode ser compacto pois não é limitado (embora seja fechado)[†]. Agora, será verdade que se K é fechado em $\mathbb{R}$ e limitado, então K é compacto?

Teorema 2.0.5 (Heine-Borel-Lebesgue...[‡]). *Para $K \subseteq \mathbb{R}$, são equivalentes:*

(i) *K é compacto;*

(ii) *K é fechado e limitado;*

(iii) *todo subconjunto infinito de K tem ponto de acumulação em K;*

(iv) *toda sequência de K tem subsequência que converge em K.*

Demonstração. A implicação (i) ⇒ (ii) já foi feita. Por sua vez, a implicação (ii) ⇒ (iii) segue do Teorema de Bolzano-Weierstrass em cinco passos:

(I) se $A \subseteq K$ é infinito, então existe uma sequência *injetiva* $(x_n)_{n\in\mathbb{N}}$ em A,

(II) como K é limitado, segue que $(x_n)_{n\in\mathbb{N}}$ é limitada;

(III) Bolzano-Weierstrass assegura subsequência $(x_{n_k})_{k\in\mathbb{N}}$ que converge para algum $x \in \mathbb{R}$;

(IV) como K também é fechado, resulta que $x \in K$ (cf. Exercício 1.208);

(V) como a sequência $(x_n)_{n\in\mathbb{N}}$ em A é injetora, segue que x é ponto de acumulação de A (convença-se disso!)$^{\star}_{\star\star}$.

Para (iii) ⇒ (iv), fixada uma sequência $(x_n)_n$ de K, dois casos podem ocorrer:

(A) $\{x_n : n \in \mathbb{N}\}$ pode ser finito, e daí basta tomar uma subsequência constante formada por algum termo que se repita infinitamente;

(B) $\{x_n : n \in \mathbb{N}\}$ pode ser infinito, e daí a hipótese garante um ponto de acumulação x em K, donde ~~é fácil~~ não é difícil obter uma subsequência que converge para o ponto de acumulação (cf. Exercício 1.204).

A última implicação, (iv) ⇒ (i), ficará bem mais simples em posse da seguinte

⌜**Afirmação.** *Para toda coleção $\mathcal{U}$ de abertos de $\mathbb{R}$ existe $\mathcal{V} \subseteq \mathcal{U}$, com $\mathcal{V}$ enumerável, tal que $\bigcup_{U\in\mathcal{U}} U = \bigcup_{V\in\mathcal{V}} V$.*

Demonstração. Seja $S := \bigcup_{U\in\mathcal{U}} U$. Para cada $x \in S$, podemos *escolher* um aberto $U_x \in \mathcal{U}$, bem como números racionais $a_x, b_x \in \mathbb{Q}$ tais que $x \in (a_x, b_x) \subseteq U_x$. Como a família $\mathcal{B} := \{(a,b) : a,b \in \mathbb{Q} \text{ e } a < b\}$ é enumerável (pois $\mathbb{Q} \times \mathbb{Q}$ é enumerável!), resulta que a família $\mathcal{C} := \{(a_x, b_x) : x \in S\}$ também é enumerável, já que $\mathcal{C} \subseteq \mathcal{B}$. Reescrevendo $\mathcal{C} := \{(a_n, b_n) : n \in \mathbb{N}\}$, segue que para cada $n \in \mathbb{N}$ existe $x_n \in S$ com $(a_n, b_n) \subseteq U_{x_n}$, e daí não é difícil perceber $S \subseteq \bigcup_{n\in\mathbb{N}} U_{x_n}$, como desejado. ⌟

[†]Note que $\mathbb{R} \setminus [0,+\infty) = (-\infty, 0)$, um (intervalo) aberto de $\mathbb{R}$. Analogamente, $[a,b]$ é fechado em $\mathbb{R}$ para quaisquer $a, b \in \mathbb{R}$, pois $\mathbb{R} \setminus [a,b] = (-\infty, a) \cup (b, +\infty)$.

[‡]...Bolzano-Weierstrass-Fréchet-Hausdorff... (cf. Observação 2.0.23).

Enfim, supondo (iv), provaremos que K é compacto. Dada uma cobertura aberta $\mathcal{U}$ para K, a afirmação assegura uma subcobertura enumerável $\mathcal{V} := \{V_n : n \in \mathbb{N}\} \subseteq \mathcal{U}$, e mostraremos que $\mathcal{V}$ admite subcobertura finita para K (e, portanto, $\mathcal{U}$ também admite): ora, se não fosse o caso, para cada $n \in \mathbb{N}$ poderíamos escolher $x_n \in K \setminus (V_0 \cup \ldots \cup V_n)$, o que resulta numa sequência $(x_n)_{n\in\mathbb{N}}$ que não tem subsequências que convergem para pontos de K (confira o próximo exercício). $\square$

Exercício 2.5 ($\star\star\star$)**.** Complete os detalhes da demonstração.

a) Mostre que $p \in \mathbb{R}$ é ponto de acumulação de $S \subseteq \mathbb{R}$ se, e somente se, existe sequência injetora $(x_n)_{n\in\mathbb{N}}$ em $S \setminus \{p\}$ tal que $x_n \to p$. Dica: para a direção difícil ($\Rightarrow$), use a condição de acumulação para cozinhar uma sequência adequada, como na demonstração de que (iii) $\Rightarrow$ (i) no Teorema 1.6.3.

b) Mostre que não existe $x \in K$ que seja limite de alguma subsequência da sequência obtida no final da demonstração. Dica: se $x \in K$, então existe algum $V_m \in \mathcal{V}$ tal que $x \in V_m$. $\blacksquare$

Exemplo 2.0.6. $[a, b] \subseteq \mathbb{R}$ é compacto para quaisquer $a, b \in \mathbb{R}$. $\blacktriangle$

Exemplo 2.0.7. Se $(x_n)_{n\in\mathbb{N}}$ é sequência real tal que $x_n \to x$, com $x \in \mathbb{R}$, então o subconjunto $\{x_n : n \in \mathbb{N}\} \cup \{x\}$ é compacto. $\blacktriangle$

Exemplo 2.0.8. Subconjuntos finitos de $\mathbb{R}$ são compactos[†]. $\blacktriangle$

Exercício 2.6 ($\star$)**.** Convença-se de que os exemplos acima são, de fato, exemplos. $\blacksquare$

Antes de prosseguir, convém notar que a compacidade é *intrínseca*, no sentido de que os abertos da cobertura podem ser considerados contidos no próprio subconjunto — e não precisam ser vistos como abertos maiores em $\mathbb{R}$. Mais precisamente:

Exercício 2.7 ($\star\star$)**.** Mostre que para $K \subseteq \mathbb{R}$, são equivalentes:

(i) K é compacto;

(ii) para toda coleção $\mathcal{U}$ de abertos de K tal que $K = \bigcup_{U\in\mathcal{U}} U$, existe $\mathcal{V} \subseteq \mathcal{U}$, com $\mathcal{V}$ finito e $K = \bigcup_{V\in\mathcal{V}} V$. $\blacksquare$

Se esta não é sua primeira vez estudando Análise, você deve ter sentido falta dos ~~superestimados~~ intervalos encaixantes, fundamentais nos cânones [24, 25]. Eles estão na Subseção 2.0.1 §0, não se preocupe.

§1 Os Teoremas de Weierstrass e do Valor Médio

De volta ao problema inicial:

Lema 2.0.9. *Sejam $K \subseteq \mathbb{R}$ um subconjunto e $f : K \to \mathbb{R}$ uma função. Se K é compacto e f é contínua, então $\operatorname{im}(f)$ é compacto.*

Demonstração. Ora, se $\mathcal{V}$ é uma cobertura aberta para $\operatorname{im}(f)$, então $K \subseteq \bigcup_{V\in\mathcal{V}} f^{-1}[V]$, com cada $f^{-1}[V]$ aberto em K em virtude da continuidade de f. Em outras palavras, $\{f^{-1}[V] : V \in \mathcal{V}\}$ é uma cobertura por abertos de K. Daí, por K ser compacto, existem $V_0, \ldots, V_n \in \mathcal{V}$ tais que $K \subseteq \bigcup_{j\leq n} f^{-1}[V_j]$ e, consequentemente, $\operatorname{im}(f) \subseteq V_0 \cup \ldots \cup V_n$. $\square$

[†] Este já era óbvio em virtude da primeira definição de compacidade...

Corolário 2.0.10 (Weierstrass). *Seja* $f\colon [a,b] \to \mathbb{R}$. *Se* f *é contínua, então* f *tem máximo e mínimo.*

Demonstração. O domínio de f é fechado e limitado e, portanto, compacto. Logo, o subconjunto $K := \operatorname{im}(f)$ também é compacto (pelo lema anterior), donde segue que deve ser fechado e limitado. Por ser limitado, existem $m := \inf K$ e $M := \sup K$, enquanto a garantia de K ser fechado permite concluir que $m, M \in K$: basta apelar para os Exercícios 1.64 e 1.208. Acabou. □

Uma das inúmeras consequências do Teorema de Weiertrass recebe um título próprio, tanto por sua interpretação geométrica quanto pela vasta gama de aplicações.

Lema 2.0.11 (Rolle). *Para* $a, b \in \mathbb{R}$ *com* $a < b$ *e uma função contínua* $f\colon [a,b] \to \mathbb{R}$*, diferenciável no intervalo* (a,b) *e tal que* $f(a) = f(b)$*, existe* $c \in (a,b)$ *com* $f'(c) = 0$.

Demonstração. Pelo Teorema de Weierstrass, f tem um ponto de máximo e um ponto de mínimo. Se tais pontos coincidirem com os extremos do intervalo, então f é constante (certo?!)* e, portanto, tem derivada nula (cf. Exercício 1.158). Se, porém, algum desses pontos pertencer a (a,b), então será um ponto de acumulação *bilateral* de $[a,b]$, extremo *local* de f em que a função é diferenciável. Logo, sua derivada será nula (cf. Exercício 1.202). □

Teorema 2.0.12 (do **valor médio**, T.V.M.). *Para* $a, b \in \mathbb{R}$ *com* $a < b$ *e uma função contínua* $f\colon [a,b] \to \mathbb{R}$*, diferenciável no intervalo* (a,b)*, existe* $c \in (a,b)$ *satisfazendo* $f'(c) = \frac{f(b)-f(a)}{b-a}$.

Demonstração. A ideia é manipular f para obter uma função g satisfazendo as hipóteses do Lema de Rolle, de modo que a partir da identidade $g'(c) = 0$ se possa concluir $f'(c) = \frac{f(b)-f(a)}{b-a}$, ou seja: precisa-se que $g'(c) = f'(c) - \frac{f(b)-f(a)}{b-a}$. Um modo razoável de fazer isso consiste em definir $g(x) := f(x) - \gamma x$, onde $\gamma := \frac{f(b)-f(a)}{b-a}$: tem-se g contínua em $[a,b]$, diferenciável em (a,b) e tal que $g(a) = g(b)$ (verifique?!)$^\star_\star$, exatamente as exigências do Lema de Rolle. □

Você se lembra daquela história de somar uma constante nos exercícios de *antiderivação* que te obrigavam a fazer em Cálculo I? Eis o culpado.

Corolário 2.0.13. *Sejam* $I \neq \emptyset$ *um intervalo de* $\mathbb{R}$ *e* $f\colon I \to \mathbb{R}$ *contínua. Se* $f'(p) = 0$ *para todo ponto* p *no* *<u>interior</u>* *de* I*, então* f *é constante em* I.

Demonstração. Primeiro, note que para $a, b \in I$ com $a < b$, as hipóteses acarretam $(a,b) \subseteq I$ com f contínua em $[a,b]$ e diferenciável em (a,b). Logo, pelo T.V.M., existe $c \in (a,b)$ tal que

$$f'(c) = \frac{f(b) - f(a)}{b - a},$$

donde segue que $f(b) = f(a)$ (percebeu?)*. Isto mostra que f é constante no *interior* do intervalo I (cf. Exercício 1.215). Agora, se existir $x \in I$ que não é ponto interior de I, então é possível obter $(x_n)_{n\in\mathbb{N}}$, onde cada x_n é ponto interior de I, tal que $x_n \to x$, e daí $f(x_n) \to f(x)$, donde o resultado segue. □

Exercício 2.8 ($^\star_\star$). Complete os detalhes da demonstração anterior. ■

Corolário 2.0.14. *Sejam* $f, g\colon [a,b] \to \mathbb{R}$ *funções contínuas em* $[a,b]$ *e diferenciáveis em* (a,b)*. Se* $f'(p) = g'(p)$ *para todo ponto* $p \in (a,b)$*, então existe* $C \in \mathbb{R}$ *tal que* $f(x) = g(x) + C$ *para todo* $x \in [a,b]$.

Demonstração. A função $h\colon [a,b] \to \mathbb{R}$ dada por $h(x) := f(x) - g(x)$ é tal que h é contínua em $[a,b]$ e $h'(p) = 0$ para todo $p \in (a,b)$, donde segue que existe $C \in \mathbb{R}$ com $h(x) = C$ para todo $x \in [a,b]$. □

2.0.1 Extras

§0 Compacidade em espaços métricos e topológicos

Secretamente, o Exercício 2.7 dá a dica de como definir *compacidade* em espaços topológicos.

Definição 2.0.15. Dizemos que $\mathcal{U}$ é uma **cobertura aberta** para um espaço topológico $(X, \mathcal{T})$ se $\mathcal{U}$ for uma família de $\mathcal{T}$-abertos de X tal que para todo $x \in X$ exista $U \in \mathcal{U}$ tal que $x \in U$. Um espaço topológico $(X, \mathcal{T})$ é **compacto** se toda cobertura aberta para X admite subcobertura finita. ¶

Nesse sentido, o exercício supracitado apenas revela que quando X é *subespaço topológico* de outro espaço, os abertos das coberturas podem ser tomados no espaço que contém X (pense a respeito)$^{\substack{\star\\\star}}$. Em todo caso, agora podemos discutir compacidade em *ambientes* bem mais gerais do que $\mathbb{R}$. Vamos começar devagar.

Exercício 2.9 ($\star$)**.** Mostre que se X e Y são espaços topológicos homeomorfos, então X é compacto se, e somente se, Y é compacto. ■

Exercício 2.10 ($\substack{\star\\\star}$)**.** Mostre que $[-\infty, +\infty]$ é espaço compacto. ■

Exemplo 2.0.16. *Um espaço discreto é compacto se, e somente se, é finito.* Por um lado, se X é discreto e compacto, então a cobertura aberta $\{\{x\} : x \in X\}$ admite subcobertura finita e, portanto, X é finito (certo?)*. A recíproca é automática. ▲

Exemplo 2.0.17. *Sequências convergentes são compactas.* Mais precisamente, se $(x_n)_{n\in\mathbb{N}}$ é sequência num espaço topológico X e $x_n \to x$, então $K := \{x_n : n \in \mathbb{N}\} \cup \{x\}$ é subespaço compacto. Com efeito, se $\mathcal{U}$ é cobertura aberta para K por abertos de X, então existe $U \in \mathcal{U}$ tal que $x \in U$, e daí $\{x_n : n \geq N\} \subseteq U$ para algum $N \in \mathbb{N}$, de modo que restam apenas finitos pontos fora de U para serem cobertos por abertos de $\mathcal{U}$. ▲

Exercício 2.11 ($\substack{\star\\\star}$)**.** Mostre que se X é compacto e $K \subseteq X$ é fechado (cf. Exercício 1.240), então K é compacto com a topologia de subespaço. Dica: se $\mathcal{U}$ é coleção de abertos de X e $K \subseteq \bigcup_{U\in\mathcal{U}} U$, então $\mathcal{V} := \mathcal{U} \cup \{X \setminus K\}$ é cobertura aberta para X. ■

Exercício 2.12 ($\substack{\star\\\star}$)**.** Mostre que se X é de Hausdorff e $K \subseteq X$ é compacto, então K é fechado. Dica: imite a prova do Lema 2.0.4. ■

Os dois exercícios anteriores sugerem que compacidade conversa bem com fechados. Na verdade, a relação é bem mais profunda, já que fechados são complementares de abertos[†].

Exercício 2.13 ($\substack{\star\\\star}$)**.** Sejam $\mathcal{U}$ e $\mathcal{F}$ famílias não vazias de subconjuntos de X.

a) Mostre que $X = \bigcup_{U\in\mathcal{U}} U$ se, e somente se, $\emptyset = \bigcap_{U\in\mathcal{U}} X \setminus U$. Dica: Lema 0.4.14.

b) Mostre que $\emptyset = \bigcap_{F\in\mathcal{F}} F$ se, e somente se, $X = \bigcup_{F\in\mathcal{F}} X \setminus F$. Dica: Lema 0.4.14. ■

[†]Para saber mais sobre conjuntos fechados, confira os exercícios da Seção 1.9.

Dizemos que um conjunto não vazio $\mathcal{F}$ de subconjuntos de X tem a **propriedade da interseção finita (p.i.f.)** se para quaisquer $n \in \mathbb{N}$ e $F_0, \ldots, F_n \in \mathcal{F}$ valer que $F_0 \cap \ldots \cap F_n \neq \emptyset$.

Teorema 2.0.18. *Para um espaço topológico X, são equivalentes:*

(i) *X é compacto;*

(ii) *toda família $\mathcal{F} \neq \emptyset$ de fechados de X com a p.i.f. é tal que $\bigcap_{F\in\mathcal{F}} F \neq \emptyset$;*

(iii) *para toda família $\mathcal{F} \neq \emptyset$ de fechados de X satisfazendo $\bigcap_{F\in\mathcal{F}} F = \emptyset$, existem $F_0, \ldots, F_n \in \mathcal{F}$, para algum $n \in \mathbb{N}$, tais que $F_0 \cap \ldots \cap F_n = \emptyset$*

Demonstração. Primeiramente, note que (ii) e (iii) são apenas a contrapositiva uma da outra e, portanto, são equivalentes. Agora, se X é compacto e $\mathcal{F} \neq \emptyset$ é uma família de fechados tal que $\bigcap_{F\in\mathcal{F}} F = \emptyset$, então o exercício anterior assegura que $\{X \setminus F : F \in \mathcal{F}\}$ é uma cobertura aberta para X. Por compacidade, existem $F_0, \ldots, F_n \in \mathcal{F}$ tais que $X = \bigcup_{j\leq n} X \setminus F_j$ e daí, novamente pelo exercício anterior,

$$\emptyset = \bigcap_{j\leq n} X \setminus (X \setminus F_j) = \bigcap_{j\leq n} F_j,$$

como desejado. A recíproca será problema seu $({}^{\star}_{\star})$. $\square$

Corolário 2.0.19 (Intervalos encaixantes)**.** *Seja $(I_n)_{n\in\mathbb{N}}$ uma sequência de intervalos não vazios, fechados e limitados de $\mathbb{R}$. Se $I_{n+1} \subseteq I_n$ para todo n, então $\bigcap_{n\in\mathbb{N}} I_n \neq \emptyset$.*

Demonstração. Se mostrarmos que cada I_n é fechado em I_0, seguirá que $\{I_n : n \in \mathbb{N}\}$ é uma família de fechados do compacto I_0 com a p.i.f. e, pelo teorema anterior, deverá valer $\bigcap_{n\in\mathbb{N}} I_n \neq \emptyset$. Vejamos:

(i) por definição, $F \subseteq I_0$ é fechado em I_0 se, e somente se, $I_0 \setminus F$ é aberto em I_0;

(ii) por sua vez, $I_0 \setminus F$ é aberto em I_0 se, e somente se, existe $A \subseteq \mathbb{R}$ aberto tal que $I_0 \setminus F = A \cap I_0$;

(iii) consequentemente, $F = I_0 \setminus (A \cap I_0) = (I_0 \setminus A) \cup (I_0 \setminus I_0) = I_0 \cap (\mathbb{R} \setminus A)$, mostrando que $F = I_0 \cap G$, com $G = \mathbb{R} \setminus A$ fechado em $\mathbb{R}$;

(iv) em particular, por I_0 ser fechado em $\mathbb{R}$, podemos (leia-se: você pode!)* concluir que $F \subseteq I_0$ é fechado em I_0 se, e somente se, F é fechado em $\mathbb{R}$. $\square$

Curiosamente, ao tentar generalizar a propriedade dos intervalos encaixantes para espaços métricos, recaímos não em compacidade, mas sim em completude! Para simplificar as notações, diremos que uma sequência $(F_n)_{n\in\mathbb{N}}$ de subconjuntos de X é *decrescente* se para todo $n \in \mathbb{N}$ valer $F_{n+1} \subseteq F_n$. A seguir, para $S \subseteq X$, $\operatorname{diam}(S) := \sup\{d(x,y) : x, y \in S\}$ indica o **diâmetro** de S.

Proposição 2.0.20 (Fechados encaixantes — versão métrica)**.** *Seja $(F_n)_n$ uma sequência decrescente de subconjuntos fechados de um espaço métrico X cuja métrica é d. Se X é completo e* $\operatorname{diam}(F_n) \to 0$, *então existe um único ponto em $\bigcap_{n\in\mathbb{N}} F_n$.*

Demonstração. Pode-se fixar, para cada $n \in \mathbb{N}$, um elemento $x_n \in F_n$, o que resulta numa sequência $(x_n)_n$ em X que é de Cauchy: para $m, n \in \mathbb{N}$ com $m \geq n$, tem-se $x_n \in F_n \subseteq F_m$ e $x_m \in F_m$, donde segue que $d(x_m, x_n) \leq \operatorname{diam}(F_m)$ e, por valer $\operatorname{diam}(F_m) \to 0$, o restante da afirmação segue. Por conta da completude, existe $x \in X$ com $x_n \to x$, ponto que deve pertencer a F_m para todo m: como $(x_n)_{n \geq m}$ é uma sequência em F_m que é subsequência de $(x_n)_n$, podemos inferir que $\lim\limits_{n \geq m} x_n = x$ (por ser subsequência!) e, por conseguinte, $x \in F_m$ (pois este é fechado!). Finalmente, se $y \in F_m$ para todo m, então $d(x, y) \leq \operatorname{diam}(F_m)$ para todo m e, portanto, $d(x, y) = 0$, i.e., $x = y$. □

Exercício 2.14 (Opcional — $(^{\star}_{\star\star})$). Suponha que para toda sequência decrescente $(F_n)_n$ de fechados não vazios de um espaço métrico X satisfazendo $\operatorname{diam}(F_n) \to 0$ tenha-se um único elemento na interseção $\bigcap_{n \in \mathbb{N}} F_n$. Mostre que X é completo. Dica: para uma sequência de Cauchy $(x_n)_{n \in \mathbb{N}}$ no espaço, considere os fechados $F_n := \overline{\{x_m : m \geq n\}}$ para cada $n \in \mathbb{N}$. ■

Isto sugere que compacidade e completude tenham algum tipo de relação em espaços métricos. Num primeiro momento, é tentador usar a caracterização de compacidade via p.i.f. juntamente com o exercício anterior para concluir que todo espaço métrico compacto é completo. Está certo (certo?)$^{\star}_{\star}$, mas faremos melhor.

Exercício 2.15 (!?). Caso não tenha visto, confira a definição topológica de ponto de acumulação apresentada no Exercício 1.141 e perceba que ela generaliza todas as noções anteriores de ponto de acumulação. ■

Definição 2.0.21. Diremos que um espaço métrico é **totalmente limitado** se toda sequência admitir subsequência de Cauchy. ¶

Teorema 2.0.22. *Para um espaço métrico K, são equivalentes:*

(i) *K é compacto;*

(ii) *todo subconjunto infinito de K tem ponto de acumulação;*

(iii) *toda sequência em K tem subsequência convergente;*

(iv) *K é completo e totalmente limitado.*

Demonstração. A primeira implicação é simples: se $A \subseteq K$ não tem pontos de acumulação em K, então para cada $x \in K$ existe um aberto $V_x \subseteq K$ com $x \in V_x$ que testemunha o fato de A não se acumular em x, ou seja, tal que $V_x \cap (A \setminus \{x\}) = \emptyset$. Pela compacidade, existem $x_0, \ldots, x_n \in K$ com $K = \bigcup_{j \leq n} V_{x_j}$, mas isto obriga que A esteja contido em $\{x_0, \ldots, x_n\}$. A implicação (ii) ⇒ (iii) segue os mesmos moldes da implicação "(iii) ⇒ (iv)" no Teorema 2.0.5 e, por isso, será problema seu $(^{\star}_{\star\star})$[†]. A implicação (iii) ⇒ (iv) é divertida: se toda sequência tem subsequência convergente, então toda sequência de Cauchy converge[‡] bem como toda sequência tem subsequência de Cauchy, i.e., o espaço métrico é completo e totalmente limitado. A carga dramática fica por conta da implicação (iv) ⇒ (i), cuja prova é adaptada de Efe Ok [31].

⌜**Afirmação.** *Se X é espaço métrico totalmente limitado, então para todo $r > 0$ existe um subconjunto finito $F_r \subseteq X$ tal que $X = \bigcup_{x \in F_r} B[x, r]$, onde $B[x, r] := \{y \in X : d(x, y) \leq r\}$ é a d-bola fechada de centro x e raio r* (cf. Exercício 1.246).

[†]Em particular, observe que o item (a) do Exercício 2.5 permanece válido em espaços métricos $(^{\star}_{\star})$.

[‡]Confira a discussão sobre espaços métricos completos, na Subseção 1.3.1 §1.

Demonstração. Se não fosse o caso, existiria $r > 0$ tal que $\{B[x, r] : x \in X\}$ não tem subcobertura finita. Em particular, para $x_0 \in X$ fixado, existe $x_1 \in X \setminus B[x_0, r]$, o que dispara a existência de $x_2 \in X \setminus B[x_0, r] \cup B[x_1, r]$, e assim sucessivamente. A sequência $(x_n)_n$ obtida recursivamente desse processo não admite subsequências de Cauchy, já que para quaisquer $m, n \in \mathbb{N}$ deve-se ter $d(x_m, x_n) > r$. ⌟

Agora, assumindo (iv), suponha que K não seja compacto e fixe uma cobertura aberta $\mathcal{U}$ sem subcobertura finita. Em vista da afirmação anterior, podemos supor que existe um subconjunto finito F_0 tal que $K = \bigcup_{x \in F_0} B[x, 1]$, donde se obtém $x_0 \in F_0$ tal que $B_0 := B[x_0, 1]$ não admite subcobertura finita por abertos de $\mathcal{U}$: caso contrário, $\mathcal{U}$ teria uma subcobertura finita para K. Uma vez que limitação total é uma propriedade hereditária para subespaços (por quê?)$^{\star}_{\star}$, pode-se repetir o argumento a fim de obter outra bola fechada $B_1 := B\left[x_1, \frac{1}{2}\right] \subseteq B[x_0, 1]$ que não admite subcobertura finita por abertos de $\mathcal{U}$. Procedendo recursivamente, obtém-se uma sequência decrescente de *fechados*[†] $(B_n)_n$ cujo diâmetro converge para 0. Logo, a completude permite conjurar $x \in \bigcap_{n \in \mathbb{N}} B_n$, o que leva a uma contradição: como existem $U \in \mathcal{U}$ com $x \in U$ e $r > 0$ com $B(x, r) \subseteq U$, existe $N \in \mathbb{N}$ com $x_N \in B\left(x, \frac{r}{2}\right)$ e $\frac{1}{2^N} < \frac{r}{2}$, acarretando $B_N \subseteq U$ (percebeu?)*, o que contraria a escolha de B_N. Portanto, K é compacto. □

Observação 2.0.23 (Sobre as encarnações de compacidade)**.** A história da *compacidade* se confunde com a própria história da Análise e da Topologia Geral. É evidente que não se começou do geral para o particular, como feito aqui, o que não é um problema, posto que não se reinventa o fogo toda vez que alguém sente frio. Dito isso, convém destacar algumas coisas.

(a) A *constatação* de que (intervalos) fechados e limitados satisfazem as condições da Definição 2.0.0 é usualmente *creditada* a Heine e Borel (por volta de 1894). Apesar disso, em seus argumentos, ambos trataram apenas de coberturas enumeráveis de abertos. Costuma-se atribuir a Lebesgue (1904) a versão para coberturas de cardinalidade arbitrária — por isso o nome "Teorema de Borel-Lebesgue" encontrado em textos que não usam a Definição 2.0.0 como caracterização inicial de compacidade.

(b) As caracterizações de compacidade (ii) e (iii) do último teorema, assim como o próprio termo "compacidade", surgiram no contexto de *espaços métricos*, introduzidas por Fréchet em 1906. Apenas na década seguinte (1914) a equivalência com a *propriedade de Borel-Lebesgue* (das coberturas abertas) foi observada, por Hausdorff.

(c) A caracterização geral de compacidade em termos de fechados com a *propriedade da interseção finita* (Teorema 2.0.18) foi *registrada* pela primeira vez por Riesz (1908).

Todavia, tudo o que se fez acima foi anterior à definição atual de *espaço topológico*, que só ocorreu em 1922, com Kuratowski, mas com a influência de trabalhos anteriores, como o de Hausdorff em 1914. Entre outras coisas, isto serve de alerta para a possibilidade de encontrar textos que definam compacidade por meio de condições diferentes (mas equivalentes), e que enunciem a Definição 2.0.0 como uma caracterização alternativa garantida por algum teorema *com nome de gente*[‡]. Por fim, um alerta: as equivalências no último teorema só podem ser garantidas em espaços métricos. △

[†]Assim como bolas abertas são abertas (cf. Exercício 1.128), bolas fechadas também são fechadas (cf. Exercício 1.246).

[‡]Mais informações sobre a História da Topologia Geral podem ser encontradas nos textos de Engelking [11] e Willard [43], ou em trabalhos voltados explicitamente para a História da Análise, como a recente obra de Thomas Sonar [41]. Para uma revisão sobre a história da compacidade, confira [33].

§1 O Lema de Riesz e a dimensão de espaços normados

Nenhuma das caracterizações anteriores de compacidade parece ser tão simples quanto a que foi apresentada no item (ii) do Teorema 2.0.5: $K \subseteq \mathbb{R}$ é compacto se, e somente se, K é fechado e limitado. Por que ela foi ignorada nas considerações anteriores?

Primeiro: ao lidar com compacidade de espaços topológicos, deve-se ter em mente que todo espaço é fechado em si mesmo, o que inutiliza tal noção para uma definição geral. Analogamente, limitação só faz sentido em espaços métricos ou normados. Assim, a caracterização de compacidade como "fechado + limitado" só tem *chances* de funcionar em (subespaços de) espaços métricos.

Segundo: tal caracterização não tem chances de funcionar em todos os espaços métricos, já que qualquer métrica admite uma métrica limitada e *topologicamente equivalente* (cf. Exercício 1.258), de modo que a própria reta real se revela um contraexemplo (por quê?!)*. Talvez tenhamos mais sorte com espaços normados...

Terceiro: essa caracterização não funciona nem em espaços normados.

Teorema 2.0.24 (*A.k.a.* Lema de Riesz)**.** *Sejam $(E, \|\cdot\|)$ um espaço normado e $S \subseteq E$ um subespaço vetorial com $\overline{S} \neq E$. Para cada $\alpha \in (0,1)$ existe $x \in E$ com $\|x\| = 1$ e $\|s - x\| \geq \alpha$ para todo $s \in S$.*

Demonstração. Para $x_0 \in E \setminus \overline{S}$, seja $R := d(x_0, S)$, que satisfaz $R > 0$ (por quê?)[†]. Como $R < \frac{R}{\alpha}$, existe $x_1 \in S$ com $\|x_0 - x_1\| \leq \frac{R}{\alpha}$, e daí basta fazer $x := \frac{1}{\|x_0 - x_1\|}(x_0 - x_1)$. □

Exercício 2.16 $(^\star_\star)$**.** Verifique os detalhes omitidos na demonstração anterior. ■

Corolário 2.0.25. *Se um espaço normado $(E, \|\cdot\|)$ tem dimensão infinita, então a bola fechada (e limitada) $B[0,1] := \{x \in E : \|x\| \leq 1\}$ não é compacta.*

Demonstração. Fixado $x_0 \in E$ com $\|x_0\| = 1$, pode-se fazer $S := [x_0]$ o subespaço de E gerado por x_0, que satisfaz $\overline{S} = S \neq E$ em virtude da suposição sobre a dimensão de E (cf. Corolário 2.0.28 a seguir, além do Exercício 1.257). Logo, o teorema anterior assegura $x_1 \in E \setminus S$ com $\|x_1\| = 1$ e $\|x_0 - x_1\| \geq \frac{1}{2}$. Como $[x_0, x_1]$ ainda tem dimensão finita, existe $x_2 \in E \setminus [x_0, x_1]$ com $\|x_2\| = 1$ e $\|x_i - x_2\| \geq \frac{1}{2}$ para $i \in \{0, 1\}$. Como $[x_0, x_1, x_2]$ tem dimensão finita... Recursivamente, conclui-se que o subconjunto $B[0,1]$ admite uma sequência sem subsequência de Cauchy e, portanto, não pode ser compacto em vista do Teorema 2.0.22 (certo?)*. □

Moral da história: em espaços métricos quaisquer, não se pode assegurar a compacidade de bolas fechadas, embora elas sejam fechadas (cf. Exercício 1.246) e, obviamente, limitadas. Restam os espaços normados de dimensão finita[‡]. Desta vez a resposta é afirmativa: e é bem fácil provar com a norma do máximo, por exemplo.

Exercício 2.17 $(^{\star}_{\star\star})$**.** Considere $\mathbb{R}^n$ com a norma do máximo (cf. Exercício 1.5). Mostre que $K \subseteq \mathbb{R}^n$ é compacto se, e somente se, K é fechado em $\mathbb{R}^n$ e limitado (na norma do máximo). Dica: para a implicação que não é automática, mostre que toda sequência em K tem subsequência convergente e, para isso, aplique o Teorema de Bolzano-Weierstrass n-vezes. ■

[†]Dica $(^{\star}_{\star\star})$: confira o Exercício 1.244.

[‡]Na verdade, esta é uma conclusão apressada. Pode-se mostrar que um espaço *metrizável* X admite uma métrica d que caracteriza os subconjuntos compactos de X como sendo os fechados e limitados se, e somente se, X tem um denso enumerável e todo ponto de X está contido no interior de um compacto. A prova de tal caracterização, no entanto, foge tremendamente do escopo deste trabalho. Confira o Exercício 4.2.C de Engelking [11] para mais detalhes.

Este é um bom momento para você encrencar com a escolha da norma em $\mathbb{R}^n$: "*Ah! Por que não outra? Eu nunca escolho a norma certa?!!!*". RESPOSTA: tanto faz.

Lema 2.0.26. *Sejam $(E, \|\cdot\|)$ um espaço normado e $T\colon \mathbb{R}^n \to E$ uma transformação linear. Se $\mathbb{R}^n$ é dotado da norma do máximo, então T é contínua.*

Demonstração. Chame por $B := \{e_1, \ldots, e_n\}$ a base canônica de $\mathbb{R}^n$, no sentido usual da Álgebra Linear. Uma vez que para cada $x \in \mathbb{R}^n$ existem únicos escalares $a_1, \ldots, a_n \in \mathbb{R}$ com $x = \sum_{i \le n} a_i e_i$, resulta

$$\|T(x)\| \le \sum_{i \le n} |a_i| \|T(e_i)\| \le \underbrace{\max\{|a_1|, \ldots, |a_n|\}}_{\|x\|_\infty} \cdot \underbrace{n \max\{\|T(e_1)\|, \ldots, \|T(e_n)\|\}}_{R}, \qquad (2.0)$$

mostrando que T é contínua (cf. Proposição 1.6.17). □

Teorema 2.0.27. *Quaisquer duas normas em $\mathbb{R}^n$ são topologicamente equivalentes*[†].

Demonstração. Basta mostrar que qualquer norma em $\mathbb{R}^n$ é equivalente à norma do máximo (certo?)*. Por simplicidade, vamos escrever "E" para indicar o espaço $\mathbb{R}^n$ dotado de uma norma $\|\cdot\|$ qualquer, enquanto "$\mathbb{R}^n$" indicará tal espaço com a norma $\|\cdot\|_\infty$. Note que a função $T\colon \mathbb{R}^n \to E$, que faz $T(x) := x$ para todo $x \in \mathbb{R}^n$, é linear e, assim, o lema anterior assegura sua continuidade.

Para a continuidade da inversa de T, note que $\mathbb{S}_{\mathbb{R}^n} := \{x \in \mathbb{R}^n : \|x\|_\infty = 1\}$ é subespaço compacto de $(\mathbb{R}^n, \|\cdot\|_\infty)$: é fechado por ser pré-imagem do fechado $\{1\}$ pela função contínua $\|\cdot\|_\infty\colon \mathbb{R}^n \to [0, +\infty)$ (cf. Exercício 1.241), e é limitado por definição. Agora, como a função $\varphi := \|\cdot\| \circ T\colon \mathbb{R}^n \to [0, +\infty)$ é contínua, o Teorema de Weierstrass garante que φ deve assumir mínimo em $\mathbb{S}_{\mathbb{R}^n}$, i.e., existe $u \in \mathbb{S}_{\mathbb{R}^n}$ tal que $\|T(u)\| \le \|T(v)\|$ para todo $v \in \mathbb{S}_{\mathbb{R}^n}$, com $\|T(u)\| > 0$ por motivos que você irá descobrir $(^\star_\star)$. Logo, se $y \in E$ e $y \ne 0$, então $v := \frac{y}{\|y\|_\infty} \in \mathbb{S}_{\mathbb{R}^n}$ e

$$\|y\| = \|T(y)\| = \|y\|_\infty \|T(v)\| \ge \|y\|_\infty \|T(u)\|,$$

mostrando que basta tomar $r := \frac{1}{\|T(u)\|}$ para assegurar $\|y\|_\infty = \|T^{-1}(y)\|_\infty \le r\|y\|$ para todo $y \in E$. □

Exercício 2.18 $(^\star_\star)$**.** Complete os detalhes da demonstração anterior. Dica: entre outras coisas, será importante notar que o Teorema de Weierstrass permanece válido para qualquer função contínua da forma $X \to \mathbb{R}$, onde X é espaço *topológico* compacto. ■

Exercício 2.19 $(^\star_\star)$**.** Perceba que o argumento do último teorema prova, na verdade, o seguinte: se $(E, \|\cdot\|)$ é um espaço normado e $T\colon \mathbb{R}^n \to E$ é um isomorfismo linear[‡], então T é um homeomorfismo ao considerarmos $\mathbb{R}^n$ com a topologia induzida por $\|\cdot\|_\infty$. Dica: apague o trecho "$T(x) := x$" e todos os correspondentes da demonstração e verifique a desigualdade "$\|T^{-1}(y)\|_\infty \le r\|y\|$" com um pouco mais de cuidado (lembre-se também de que $T^{-1}(y) = x$ se, e somente se, $T(x) = y$). ■

As consequências do teorema anterior vão longe. O próximo resultado, por exemplo, combina o Corolário 1.3.35 com a desigualdade obtida no Exemplo 1.7.19:

[†] Isto é, induzem a mesma topologia (cf. Exemplo 1.7.19).
[‡] Transformação linear invertível.

Corolário 2.0.28. *Para quaisquer normas* $\|\cdot\|$ *e* $\|\cdot\|'$ *em* $\mathbb{R}^n$*, existem constantes* $r, R > 0$ *tais que*

$$r\|x\| \leq \|x\|' \leq R\|x\|$$

para todo $x \in \mathbb{R}^n$*. Em particular,* $(\mathbb{R}^n, \|\cdot\|)$ *é espaço de Banach com respeito a qualquer norma* $\|\cdot\|$ *e, consequentemente, qualquer espaço normado com dimensão finita é de Banach.*

Demonstração. A existência das constantes decorre diretamente do último teorema. Já para a segunda afirmação, sabemos que $(\mathbb{R}^n, \|\cdot\|_\infty)$ é espaço de Banach, onde $\|\cdot\|_\infty$ é a norma do máximo. Agora, para qualquer outra norma $\|\cdot\|$ em $\mathbb{R}^n$, a existência das constantes $r > 0$ e $R > 0$ como acima, para $\|\cdot\|' := \|\cdot\|_\infty$ permite mostrar que uma sequência é de Cauchy em $(\mathbb{R}^n, \|\cdot\|)$ se, e somente se, é de Cauchy em $(\mathbb{R}^n, \|\cdot\|_\infty)$, donde o resultado segue. □

Exercício 2.20 $(^{\star}_{\star})$**.** Complete os detalhes da demonstração anterior. ■

Por sua vez, ao combinar os resultados anteriores, você descobrirá porque nunca precisou se preocupar com a continuidade das transformações lineares que encontrava na escola (ou em Cálculo n, para $n > 1$).

Corolário 2.0.29. *Sejam* X *e* Y *espaços normados e* $T\colon X \to Y$ *uma transformação linear. Se* X *tem dimensão finita, então* T *é contínua.*

Exercício 2.21 $(^{\star}_{\star})$**.** Prove o corolário anterior. ■

Observação 2.0.30 (Opcional: transformações lineares descontínuas)**.** Se $(E, \|\cdot\|)$ é espaço normado com dimensão infinita, então existe uma transformação linear $T\colon E \to \mathbb{R}$ descontínua. Para simplificar o argumento, suponha que $\mathcal{B} := \{b_n : n \in \mathbb{N}\}$ seja uma base para E e, para cada $n \in \mathbb{N}$, defina

$$\varphi(b_n) := 2^n\|b_n\|.$$

Pelo que você já aprendeu em Álgebra Linear, existe uma única transformação linear $T\colon E \to \mathbb{R}$ tal que $T(b_n) = \varphi(b_n)$ para todo n. Para encerrar, note que se $x_n = \frac{1}{\|b_n\|2^n}b_n$ para todo n, obtemos $x_n \to 0_E$ (por quê?)* mas $T(x_n) = 1$ para todo n.

Para o caso de dimensão infinita qualquer, você só precisa saber que todo subconjunto linearmente independente de E está contido numa base e repetir o argumento anterior, *mutatis mutandis.* △

Independentemente do seu gosto (ou desgosto) pelos espaços normados com dimensão infinita,
não estão lá (cf. Subseção 2.1.1 §3).

Exercício 2.22 (Opcional — $(^{\star}_{\star\star})$)**.** Mostre que $\mathcal{B}(X)$ contém um subconjunto *linearmente independente*[†] em bijeção com X, para qualquer conjunto $X \neq \emptyset$. Conclua que se X é

[†]Um subconjunto S de um espaço vetorial E é **linearmente independente** se para quaisquer $n \in \mathbb{N}$ e $x_0, \ldots, x_n \in S$, uma igualdade do tipo $\alpha_0 x_0 + \ldots + \alpha_n x_n = 0$ em E só for possível com $\alpha_0 = \ldots = \alpha_n = 0$. Em Álgebra Linear, mostra-se que a dimensão de um espaço vetorial coincide com a cardinalidade de qualquer subconjunto linearmente independente maximal.

infinito, então $\mathcal{B}(X)$ tem dimensão infinita. Dica: para $x \in X$ fixado, a função $\delta_x \colon X \to \mathbb{R}$, que a cada $y \in X$ associa o número $\delta_x(y)$ dado por

$$\delta_x(y) := \begin{cases} 1, & \text{se } x = y \\ 0, & \text{se } x \neq y \end{cases},$$

é limitada (faça primeiro com $X := \{0, 1\}$ para entender o que está havendo). ■

Exercício 2.23 (${}^{\star}_{\star\star}$)**.** Para um espaço normado E, mostre que uma transformação linear $\varphi \colon E \to \mathbb{R}$ é contínua se, e somente se, $\ker \varphi := \{x \in E : \varphi(x) = 0\}$ é fechado em E. Dica: para o lado difícil, proceda pela contrapositiva. ■

§2 Exercícios extras

Exercício 2.24 (${}^{\star}_{\star\star}$)**.** Uma função $f \colon X \to Y$ entre espaços métricos X e Y satisfaz a *condição de Lipschitz* se existe $c > 0$ tal que $d_Y(f(a), f(b)) \leq c \cdot d_X(a, b)$ para quaisquer $a, b \in X$.

a) Investigue o seu arsenal de exercícios resolvidos para encontrar exemplos de funções que satisfazem a condição de Lipschitz.

b) Mostre que se X é um subconjunto compacto de um espaço normado E, e Y é normado, então o número

$$\|f\|_L := \sup \left\{ \frac{\|f(a) - f(b)\|}{\|a - b\|} : a, b \in X \text{ e } a \neq b \right\}$$

existe para toda $f \colon X \to Y$ satisfazendo a condição de Lipschitz.

c) Sob as hipóteses acima, mostre que o conjunto das funções da forma $X \to Y$ que satisfazem a condição de Lipschitz é um espaço vetorial com as operações usuais, e que $\|\cdot\|_L$ define uma norma em tal espaço. ■

Exercício 2.25 (${}^{\star}_{\star}$)**.** Um espaço topológico X é um **espaço de Lindelöf** se toda cobertura aberta para X admite subcobertura enumerável.

a) Mostre que $\mathbb{R}$ é um espaço de Lindelöf.

b) Mostre que se $X = \bigcup_{n \in \mathbb{N}} K_n$, com $K_n \subseteq X$ compacto para todo n, então X é de Lindelöf.

c) Mostre que $\mathbb{R}^n$ é espaço de Lindelöf para todo $n \in \mathbb{N} \setminus \{0\}$.

d) Mostre que se X é um espaço métrico de Lindelöf, então X tem um denso enumerável. Dica: use a condição de Lindelöf infinitas vezes.

e) Mostre que se X é métrico e sua topologia tem uma base enumerável, então X é de Lindelöf. Dica: tente provar que $\mathbb{R}$ é de Lindelöf usando os intervalos abertos com extremos racionais até perceber o que importa.

f) Conclua que para espaços métricos, são equivalentes: ter base enumerável; ser de Lindelöf; ~~ser separável~~ ter denso enumerável ■

Exercício 2.26 ($^{\star}_{\star\star}$)**.** Um espaço topológico X é dito **enumeravelmente compacto** se toda cobertura aberta e enumerável para X tem subcobertura finita.

a) Mostre que X é compacto se, e somente se, é enumeravelmente compacto e de Lindelöf.

b) Mostre que X é enumeravelmente compacto se, e somente se, para toda família enumerável $\{F_n : n \in \mathbb{N}\}$ de fechados com a p.i.f. valer $\bigcap_{n\in\mathbb{N}} F_n \neq \emptyset$. Dica: imite o Teorema 2.0.22.

c) Mostre que se X é métrico e enumeravelmente compacto, então todo subconjunto infinito de X tem ponto de acumulação. Dica: se $A \subseteq X$ é infinito e não tem pontos de acumulação, então existe $A' := \{a_n : n \in \mathbb{N}\} \subseteq A$ infinito enumerável e sem pontos de acumulação; note então que $U_n := X \setminus \{a_m : m \geq n\}$ é fechado aberto para todo $n \in \mathbb{N}$ e $X = \bigcup_{n\in\mathbb{N}} U_n$.

d) Conclua que um espaço métrico X é compacto se, e somente se, é enumeravelmente compacto. ■

Exercício 2.27 ($\star$)**.** Mostre que não existe espaço vetorial normado não trivial que seja compacto. Dica: encare a norma até que ela te encare de volta. ■

2.1 Teorema de Heine-Cantor

2.1.0 Essencial

§0 Continuidade uniforme e o Teorema de Heine-Cantor

A definição de continuidade para uma função $f: X \to \mathbb{R}$ pede que f seja contínua em cada ponto de X. Por sua vez, isto significa que para cada $x \in X$ e para cada $\varepsilon > 0$ existe um $\delta > 0$ tal que $|f(x) - f(y)| < \varepsilon$ sempre que $y \in X$ satisfaz $|x - y| < \delta$. A sutileza aqui é a seguinte: uma vez que se "trava" o ponto x e o $\varepsilon > 0$, o $\delta > 0$ encontrado funciona para este x e este ε, razão pela qual alguns textos inclusive escrevem $\delta_{x,\varepsilon}$.

Definição 2.1.0. Uma função $f: X \to \mathbb{R}$ é **uniformemente contínua** se para todo $\varepsilon > 0$ existe $\delta > 0$ tal que $|f(x) - f(y)| < \varepsilon$ para quaisquer $x, y \in X$ satisfazendo $|x - y| < \delta$. ¶

Exercício 2.28 ($\star/2$)**.** Mostre que toda função uniformemente contínua é contínua. ■

Talvez a primeira pergunta que esta definição suscite é a seguinte: "*mas não é a mesma coisa??*". Se fosse a mesma coisa, a definição não teria sobrevivido ao teste do tempo e chegado até aqui, não é mesmo?

Exemplo 2.1.1. *A função $f: \mathbb{R} \to \mathbb{R}$ dada por $f(x) := x^2$ não é uniformemente contínua.* Para isso, precisamos encontrar *pelo menos* um $\varepsilon > 0$ tal que *para qualquer* $\delta > 0$ *existam* $x, y \in \mathbb{R}$ satisfazendo $|x - y| < \delta$ mas $|x^2 - y^2| \geq \varepsilon$ (certo?)[†]. Veremos que $\varepsilon := 1$ funciona para qualquer $\delta > 0$.

[†]Como sempre, saber negar afirmações feitas em *linguagem de primeira ordem* se revela algo essencial. Nesta altura do campeonato, destrinchar esse tipo de coisa *deveria ser* problema seu.

Observe que para $x, z > 0$ quaisquer, o número $y := x + z$ é tal que

$$|x^2 - y^2| = |x^2 - (x^2 + 2xz + z^2)| = |2xz + z^2| = 2xz + z^2 = z(2x + z). \tag{2.1}$$

Assim, ao colocar $z := \frac{\delta}{2}$, garantimos $|x - y| < \delta$ para qualquer x escolhido (desde que y seja tomado como $x + z$), enquanto as identidades em (2.1) mostram que basta encontrar x tal que

$$\frac{\delta}{2} \cdot \left(2x + \frac{\delta}{2}\right) \geq 1,$$

o que ocorre para todo x satisfazendo $x \geq \frac{1}{\delta} - \frac{\delta}{4}$ (certo?)*.

Geometricamente, isto significa que ao fixar um $\delta > 0$, podemos encontrar pontos x e $y \in (x - \delta, x + \delta)$ tais que $|x^2 - y^2| \geq 1$. ▲

Se ainda parecer confuso, considere o seguinte exemplo: seja $P(A, B, C)$ o *predicado* "A e B conhecem C". Neste caso, "$\forall A \forall B \exists C \ P(A, B, C)$" expressaria algo similar a "quaisquer duas pessoas conhecem pelo menos uma pessoa em comum", enquanto "$\forall A \exists C \forall B \ P(A, B, C)$" seria algo como "toda pessoa conhece alguém que todas as pessoas conhecem". Tais afirmações são equivalentes? Reflita, faça desenhos, medite... até se convencer de que a resposta correta é "não".

Agora sua impressão pode ter mudado: em vez de achar que se trata do mesmo conceito, você pode estar se perguntando se funções uniformemente contínuas *existem*.

Exercício 2.29 ($\star$)**.** Caso tenha feito o Exercício 2.24, mostre que funções que satisfazem a condição de Lipschitz são uniformemente contínuas. ■

Exercício 2.30 ($\star$)**.** Mostre que a função identidade $\mathrm{Id}_{\mathbb{R}} \colon \mathbb{R} \to \mathbb{R}$ é uniformemente contínua. Dica: a função identidade é dada por $\mathrm{Id}_{\mathbb{R}}(x) := x$ para todo $x \in \mathbb{R}$, cujo gráfico é a diagonal do plano cartesiano. ■

Exercício 2.31 ($^{\star}_{\star}$)**.** Pense rápido: se $f \colon \mathbb{R} \to \mathbb{R}$ e $g \colon \mathbb{R} \to \mathbb{R}$ são uniformemente contínuas, então $f \cdot g$ é uniformemente contínua? Observação: lembre-se de que $f \cdot g$ e $f \circ g$ representam coisas diferentes — e se você não lembra quais são essas coisas, é um sinal de que você deve revisar alguns assuntos, não acha? ■

Não seria ótimo ter algum teste para decidir sem muito esforço se uma função é uniformemente contínua?

Teorema 2.1.2 (Heine-Cantor)**.** *Se $X \subseteq \mathbb{R}$ é compacto e $f \colon X \to \mathbb{R}$ é contínua, então f é uniformemente contínua.*

Demonstração. O cerne da ideia é simples: para $\varepsilon > 0$, a continuidade de f permite encontrar, em torno de cada $x \in X$, ~~uma bola aberta $B_x := B(x, \delta_x)$~~ um intervalo aberto $I_x := (x - \delta_x, x + \delta_x)$ tal que $|f(x) - f(z)| < \varepsilon$ sempre que $z \in I_x$, o que torna irresistível usar a compacidade de X para tomar uma subcobertura finita de $\{I_x : x \in X\}$, para daí escolher δ como o menor dos δ_x's correspondentes (por que isso pode ser feito?)*.

Há um problema: o próximo passo *seria* mostrar que se z e z' são pontos quaisquer de X com $|z - z'| < \delta$, então $|f(z) - f(z')| < 2\varepsilon$; porém, para assegurar isso, seria preciso ter z e z' num mesmo intervalo da forma I_x, pois daí resultaria a desigualdade

$$|f(z) - f(z')| \leq |f(z) - f(x)| + |f(x) - f(z')| < 2\varepsilon. \tag{2.2}$$

No entanto, sabemos apenas da existência de x e x', possivelmente distintos, com $z \in I_x$ e $z' \in I_{x'}$, de modo que $|z - x'| \leq |z - z'| + |z' - x'| < \delta + \delta_{x'} \leq 2\delta_{x'}$. Ao encarar atentamente a última desigualdade, percebe-se o que fazer para salvar o argumento: dividir por 2 os raios da forma δ_x! Agora sim!

Para cada $x \in X$, seja $\delta_x > 0$ como no começo da demonstração e considere o intervalo aberto $J_x := \left(x - \frac{\delta_x}{2}, x + \frac{\delta_x}{2}\right)$. Note $J_x \subseteq I_x$ para todo x, bem como $X \subseteq \bigcup_{x \in X} J_x$ (note!)$^{\star}_{\star}$. Por X ser compacto, existe um subconjunto finito $F \subseteq X$ com $X \subseteq \bigcup_{x \in F} B\left(x, \frac{\delta_x}{2}\right)$, o que permite definir $\delta := \min\left\{\frac{\delta_x}{2} : x \in F\right\}$. Dessa forma, sempre que $z, z' \in X$ satisfazem $|z - z'| < \delta$, existe um $x \in X$ tal que $z, z' \in I_x$: certamente existe $x \in F$ tal que $z \in J_x$ (certo?!)*, de modo que

$$|z' - x| \leq |z' - z| + |z - x| < \delta + \frac{\delta_x}{2} \leq \frac{\delta_x}{2} + \frac{\delta_x}{2} = \delta_x,$$

i.e., $z' \in I_x$ e $z \in J_x \subseteq I_x$. Para encerrar, basta repetir o argumento em (2.2). □

§1 Extensão de funções uniformemente contínuas

Okay. Continuidade uniforme não é a mesma coisa que continuidade clássica. Mas e daí?

A pergunta é pertinente, mas uma resposta honesta pode ser inútil nesta etapa da vida: continuidade uniforme não preserva apenas a estrutura topológica dos espaços, mas sim a estrutura uniforme induzida pelas métricas. Viu só? Eu avisei. A resposta pragmática talvez soe melhor: dá para fazer mais coisas com funções uniformemente contínuas!

Exercício 2.32 ($^{\star}_{\star}$)**.** Seja $f\colon X \to \mathbb{R}$ função uniformemente contínua.

a) Mostre que se $(x_n)_{n \in \mathbb{N}}$ é sequência de Cauchy em X, então $(f(x_n))_{n \in \mathbb{N}}$ é de Cauchy.

b) Conclua que a função $\iota\colon \mathbb{R} \setminus \{0\} \to \mathbb{R}$ dada por $\iota(x) := \frac{1}{x}$ não é uniformemente contínua. ■

Embora pareça uma banalidade, o exercício anterior esconde uma propriedade fundamental das funções uniformemente contínuas.

Teorema 2.1.3 (Extensão de funções uniformemente contínuas)**.** *Seja $X \subseteq \mathbb{R}$ e $f\colon X \to \mathbb{R}$ uma função. Se f é uniformemente contínua, então existe uma única função contínua $F\colon \overline{X} \to \mathbb{R}$ tal que $F(x) = f(x)$ para todo $x \in X$.*

Demonstração. A unicidade segue com um dos argumentos que você poderia ter usado para resolver o item (d) do Exercício 1.218: se $z \in \overline{X}$, então existe $(z_n)_{n \in \mathbb{N}}$ em X tal que $z_n \to z$, de modo que se $F, G\colon \overline{X} \to \mathbb{R}$ forem funções contínuas satisfazendo $F(x) = G(x) = f(x)$ para todo $x \in X$, então

$$F(z) = \lim_{n \to \infty} F(z_n) = \lim_{n \to \infty} f(z_n) = \lim_{n \to \infty} G(z_n) = \lim_{n \to \infty} G(z).$$

Secretamente, o argumento acima também dá a dica de como *definir* F. Para cada $z \in \overline{X}$ existe $(z_n)_n$ em X com $z_n \to z$, donde a continuidade uniforme de f garante que $(f(z_n))_n$ é de Cauchy em $\mathbb{R}$ e, portanto, convergente para algum ponto, que muito apropriadamente será chamado de $F(z)$. Note que de um ponto de vista formal, isto já define uma função $F\colon \overline{X} \to \mathbb{R}$, de modo que resta apenas mostrar sua continuidade uniforme. Ora, fixado $\varepsilon > 0$, há $\delta > 0$ tal que $|f(x) - f(y)| < \varepsilon' < \varepsilon$ sempre que $x, y \in X$

satisfazem $|x-y|<3\delta$. Logo, se $z,z'\in\overline{X}$ satisfazem $|z-z'|<\delta$, então para $N\in\mathbb{N}$ suficientemente grande deve-se ter

$$|z_N-z'_N|\leq|z_N-z|+|z-z'|+|z'_N-z'|<3\delta,$$

acarretando $|f(z_N)-f(z'_N)|<\varepsilon'$ e, pela continuidade do valor absoluto $|\cdot|\colon\mathbb{R}\to[0,+\infty)$ aliada à monotonicidade dos limites (Proposição 1.2.12),

$$|F(z)-F(z')|=\left|\lim_{n\to\infty}f(z_n)-\lim_{n\to\infty}f(z'_n)\right|=\lim_{n\to\infty}|f(z_n)-f(z'_n)|\leq\varepsilon'<\varepsilon,$$

como desejado (certo?)$^{\star}_{\star}$. □

Excesso de zelo. O argumento acima mostrou que a família $P_z:=\{y\in\mathbb{R}:\exists\,(z_n)_n$ em X tal que $z_n\to z$ e $f(z_n)\to y\}$ é não vazia para cada $z\in\overline{X}$, o que permitiu apelar para o Axioma da Escolha a fim de *conjurar* $F\colon\overline{X}\to\mathbb{R}$ satisfazendo $F(z)\in P_z$ para todo z. Contudo, há quem goste de destacar que o Axioma da Escolha é supérfluo aqui, pois P_z tem apenas um elemento: se $z_n\to z$, $z'_n\to z$ e $y,y'\in\mathbb{R}$ são tais que $f(z_n)\to y$ e $f(z'_n)\to y'$, então $|y-y'|\leq|y-f(z_n)|+|f(z_n)-f(z'_n)|+|f(z'_n)-y'|$, que pode ser controlado por f ser uniformemente contínua. Tal preciosismo, no entanto, é desnecessário em contextos que assumem (ou ignoram) o Axioma da Escolha.

Exemplo 2.1.4 (Adiável: potências reais)**.** Para $a\in\mathbb{R}$ com $a\geq 0$, e $n\in\mathbb{N}\setminus\{0\}$, já não é uma tarefa difícil mostrar que o conjunto

$$P_{a,n}:=\{x\in[0,+\infty):x^n<a\}$$

admite supremo (faça isso!)$^{\star}_{\star\star}$. Chamando por $\alpha:=\sup P_{a,n}$, uma simples adaptação do argumento apresentado no Exemplo 1.2.16 permite mostrar que $\alpha^n=a$ (já sabe né?)$^{\star}_{\star\star}$. Em outras palavras, mostramos que existe um número real $\alpha\geq 0$ tal que $\alpha^n=a$. Como de costume nessas situações, α é único: se $\beta\geq 0$ também satisfaz $\beta^n=a$, então

$$0=\alpha^n-\beta^n=(\alpha-\beta)\cdot\left(\alpha^{n-1}+\alpha^{n-2}\beta+\ldots+\alpha\beta^{n-2}+\beta^{n-1}\right),$$

de modo que se ocorresse $\alpha\neq\beta$, então resultaria $\alpha^n\neq\beta^n$ (por quê?!)$^{\star}_{\star}$.

Definição 2.1.5. O número α acima é denotado por $\sqrt[n]{a}$, a n-**ésima raiz de** a. ¶

Exercício 2.33 ($^{\star}_{\star}$)**.** Mostre que a função $\sqrt[n]{(\cdot)}\colon[0,+\infty)\to\mathbb{R}$ é contínua e injetora. Dica: imite a demonstração do Corolário 1.6.9, ou espere a próxima seção para fazer isto sem esforço. ■

O que tudo isso tem a ver com os problemas desta seção? Veja: a discussão acima dá sentido ao número real $r^{\frac{m}{n}}:=\sqrt[n]{r^m}$ sempre que $m,n\in\mathbb{N}$ com $n\neq 0$ e $x\geq 0$, donde segue que r^q está definido para $q\in\mathbb{Q}$ sempre que $q\geq 0$. De fato, se $\frac{a}{b}=\frac{c}{d}$ são frações positivas, então $\sqrt[b]{r^a}=\sqrt[d]{r^c}$ (verifique!)$^{\star}_{\star}$. Ao definir $r^{-q}:=\frac{1}{r^q}$ para $q<0$, obtemos uma função

$$r^{(\cdot)}\colon\mathbb{Q}\to\mathbb{R}$$

que a cada $q\in\mathbb{Q}$ associa a q-**ésima potência** de $r>0$. Talvez agora você tenha percebido onde iremos chegar: uma boa definição de r^x para $x\in\mathbb{R}$ *qualquer* deve concordar com a definição anterior, ou seja, deve resultar numa função que *estenda* a função $r^{(\cdot)}\colon\mathbb{Q}\to\mathbb{R}$. Faremos isso por meio do Teorema 2.1.3.

(i) Note que a função $r^{(\cdot)}$ é estritamente crescente (note!!)$^{\star}_{\star}$.

(ii) Para $p, q \in \mathbb{Q}$, as propriedades de potenciação racional que você praticou na Escola[†] aliadas ao que você aprendeu aqui justificam as identidades a seguir:

$$|r^p - r^q| = \left|r^{p-q+q} - r^q\right| = \left|r^{p-q}r^q - r^q\right| = |r^q| \cdot \left|r^{p-q} - 1\right| = r^q \cdot \left|r^{p-q} - 1\right|.$$

(iii) Pelo primeiro passo, ao restringir a função $r^{(\cdot)}$ ao subconjunto $X := [-K, K] \cap \mathbb{Q}$, onde $K \in \mathbb{N}$, segue que $r^q \leq M := \max\{r^K, r^{-K}\}$ para qualquer $q \in X$.

(iv) Supondo conhecido o limite $\lim\limits_{q \to 0} r^q = 1$, para $\varepsilon > 0$ dado garante-se que existe $\delta > 0$ tal que $|r^w - 1| < \frac{\varepsilon}{M}$ sempre que $w \in \mathbb{Q}$ satisfaz $|w| < \delta$, donde segue que $|r^p - r^q| < \varepsilon$ para quaisquer $p, q \in [-K, K] \cap \mathbb{Q}$ com $|p - q| < \delta$.

(v) Pelo Teorema 2.1.3, existe uma única função contínua $R_K \colon [-K, K] \to \mathbb{R}$ tal que $R_K(q) = r^q$ sempre que $q \in [-K, K] \cap \mathbb{Q}$.

(vi) Finalmente, a função procurada é a *colagem* das funções R_K, i.e., $R := \bigcup_{K \in \mathbb{N}} R_K$ (cf. Lema 0.3.8), que explicitamente faz $R(x) = R_K(x)$ para qualquer $K \in \mathbb{N}$ tal que $x \in [-K, K]$.

A continuidade da função R se verifica de modo quase automático: se $(x_n)_n$ é sequência em $\mathbb{R}$ com $x_n \to x$ e $x \in \mathbb{R}$, então existem $K \in \mathbb{N}$ e $N \in \mathbb{N}$ tais que $x_n \in (-K, K)$ para todo $n \geq N$ e daí

$$R(x) = R_K(x) = R_K\left(\lim_{n \geq N} x_n\right) = \lim_{n \geq N} R_K(x_n) = \lim_{n \geq N} R(x_n) \Rightarrow R(x) = \lim_{n \to \infty} R(x_n).$$

Em particular, essa construção formaliza o modo intuitivo pelo qual interpretamos algo como $3^{\sqrt{2}}$, por exemplo: trata-se do limite de *qualquer* sequência da forma $(3^{q_n})_{n \in \mathbb{N}}$, onde $(q_n)_{n \in \mathbb{N}}$ é uma sequência de números racionais que converge para $\sqrt{2}$. ▲

Felizmente, ao longo das próximas seções, teremos ferramentas bem mais práticas para assegurar a existência e as propriedades das funções potência. É por essa razão que a discussão anterior pode ser encarada como ~~opcional~~ adiável.

Exercício 2.34 $({}^{\star\star}_{\star\star})$**.** Prove que $\lim\limits_{q \to 0} r^q = 1$.

a) Generalize o item (a) do Exercício 0.134 e mostre que $\sqrt[n]{x_1 \cdot \ldots \cdot x_n} \leq \frac{x_1 + \ldots + x_n}{n}$ para quaisquer $n \in \mathbb{N} \setminus \{0\}$ e $x_1, \ldots, x_n \geq 0$. Dica: aplique o resultado do Exercício 0.134 várias vezes (ou confira o Exercício 2.110).

b) Use o item anterior para mostrar que se $q \in \mathbb{Q} \cap [0, 1]$ e $s \geq 0$, então $(1 + s)^q \leq 1 + sq$. Dica: com $q = \frac{m}{n}$ e $n \geq m$, considere $x_1 = x_2 = \ldots = x_m = (1 + s)$ e $x_{m+1} = \ldots = x_n = 1$ no item anterior.

c) Mostre que se $q \in \mathbb{Q} \cap [0, 1]$ e $r > 1$, então $1 \leq r^q \leq 1 + rq$. Dica: $r^q = (1 + (r - 1))^q$.

d) Conclua que $\lim\limits_{q \to 0} r^q = 1$, conforme $q \to 0$ com $q \in \mathbb{Q}$. Dica: avalie os limites laterais via sanduíche. ■

A Subseção 2.1.1 traz outras aplicações da continuidade uniforme, sendo a primeira delas a mais importante para o contexto deste curso.

[†] E que agora já tem maturidade para demonstrar por conta própria! $({}^{\star}_{\star\star})$.

2.1.1 Extras

§0 Integrais de Riemann como limites de redes (parte II)

Vamos usar a continuidade uniforme das funções contínuas da forma $[a,b] \to \mathbb{R}$ para provar um canhão. Não custa avisar: é bom revisar os conteúdos sobre integrais de Riemann apresentados na Subseção 1.2.1 §0.

Teorema 2.1.6. *Se $f\colon [a,b] \to \mathbb{R}$ é contínua, então f é Riemann-integrável.*

Demonstração. Pela completude de $\mathbb{R}$, basta mostrar que a rede $\left(\sum_{(\mathcal{P},T)} f\right)_{(\mathcal{P},T)}$ é de Cauchy (cf. Corolário 1.3.16). Para isso, a continuidade uniforme de $f\colon [a,b] \to \mathbb{R}$, assegurada pelo Teorema de Heine-Cantor, será fundamental:

✓ fixado $\varepsilon > 0$, existe $\delta > 0$ tal que $|f(x) - f(y)| < \frac{\varepsilon}{b-a}$ sempre que $x, y \in [a,b]$ satisfazem $|x - y| < \delta$;

✓ isso permitirá utilizar tal δ para determinar um valor para as normas das partições de Riemann capaz de garantir a desigualdade de Cauchy correspondente ao ε fixado.

Por exemplo: note que se $T := (t_1, \ldots, t_n)$ e $T' := (t'_1, \ldots, t'_n)$ são duas *tags* de uma partição $\mathcal{P} := (a_0, \ldots, a_n)$ que satisfaz $\|\mathcal{P}\| < \delta$, então

$$\left|\sum_{(\mathcal{P},T)} f - \sum_{(\mathcal{P},T')} f\right| = \left|\sum_{i=1}^{n} (f(t_i) - f(t'_i))(a_i - a_{i-1})\right| \leq \sum_{i=1}^{n} |f(t_i) - f(t'_i)|(a_i - a_{i-1})$$
$$< \frac{\varepsilon}{b-a}(b-a) = \varepsilon,$$

pois $|t_i - t'_i| \leq |a_i - a_{i-1}| < \delta$.

O problema, portanto, consiste em ajustar o argumento para que se atendam as exigências da presente *versão do critério de Cauchy*†. Para fazer isso, pode ser útil recordar a definição da relação "$\sqsubseteq$" entre partições, introduzida na Observação 1.2.23.

⌜**Afirmação.** *Nas presentes condições, existe $\delta > 0$ tal que se $(\mathcal{P}, T)$ é um partição com $\|\mathcal{P}\| < \delta$ e $(\mathcal{Q}, T')$ é outra partição com $\mathcal{P} \sqsubseteq \mathcal{Q}$, então $\left|\sum_{(\mathcal{P},T)} f - \sum_{(\mathcal{Q},T')} f\right| < \varepsilon$.*

Demonstração. Primeiro, para $\mathcal{P} := (p_0, \ldots, p_n)$ e $\mathcal{Q} := (q_0, \ldots, q_m)$, a relação $\mathcal{P} \sqsubseteq \mathcal{Q}$ apenas indica $\{p_0, \ldots, p_n\} \subseteq \{q_0, \ldots, q_m\}$. Agora, pela continuidade uniforme de f em $[a,b]$, existe $\delta > 0$ tal que $|f(x) - f(y)| < \frac{\varepsilon}{b-a}$ sempre que $|x - y| < \delta$, e é justamente este o δ procurado. Com efeito, se $(\mathcal{P}, T)$ satisfaz $\|\mathcal{P}\| < \delta$ e $\mathcal{P} \sqsubseteq \mathcal{Q}$, então cada p_i da partição $\mathcal{P}$ corresponde (de maneira estritamente crescente) a um q_{j_i} da partição $\mathcal{Q}$, de tal forma que

$$\sum_{s=j_i+1}^{j_{i+1}} q_s - q_{s-1} = p_{i+1} - p_i,$$

para $i \in \{0, \ldots, n\}$.‡ Em particular, $q_{j_0} = q_0 = p_0 = a$, $q_{j_n} = q_m = p_n = b$ e, mais geralmente,

$$\mathcal{Q} = (\underbrace{q_{j_0}}_{q_0=p_0=a}, q_{j_0+1}, \ldots, \underbrace{q_{j_1}}_{p_1}, q_{j_1+1}, \ldots, \underbrace{q_{j_{n-1}}}_{p_{n-1}}, \ldots, \underbrace{q_{j_n}}_{q_m=p_n=b}).$$

†Se você estiver com muita pressa, pode ser melhor apelar para o vindouro critério de Riemann-Darboux (cf. Teorema 2.4.9), que torna desnecessário o ajuste feito aqui.

‡Não é um bom momento para confundir $j_i + 1$ com j_{i+1}.

Logo, para $T' := (t'_1, \ldots, t'_m)$ uma *tag* de $\mathcal{Q}$,

$$\begin{aligned}\left|\sum_{(\mathcal{P},T)} f - \sum_{(\mathcal{Q},T')} f\right| &= \left|\sum_{i=1}^{n} f(t_i)(p_i - p_{i-1}) - \sum_{j=1}^{m} f(t'_j)(q_j - q_{j-1})\right| \\ &= \left|\sum_{i=1}^{n}\sum_{s=j_{i-1}+1}^{j_i} f(t_i)(q_s - q_{s-1}) - \sum_{i=1}^{n}\sum_{s=j_{i-1}+1}^{j_i} f(t'_s)(q_s - q_{s-1})\right| \\ &\leq \sum_{i=1}^{n}\sum_{s=j_{i-1}+1}^{j_i} |f(t_i) - f(t'_s)||q_s - q_{s-1}| < \frac{\varepsilon}{b-a}\cdot(b-a) = \varepsilon,\end{aligned}$$

pois $|t_i - t_s| < \delta$ para quaisquer i entre 1 e n e s entre $j_{i-1}+1$ e j_i. ⌟

Para encerrar, com o δ da afirmação referente a $\frac{\varepsilon}{2}$, note que se $(\mathcal{P}, T)$ e $(\mathcal{P}', T')$ são partições com $\|\mathcal{P}\|, \|\mathcal{P}'\| < \delta$, então para a partição $\mathcal{Q}$ obtida por meio da reunião dos pontos de $\mathcal{P}$ e $\mathcal{P}'$, e qualquer *tag* T'' para $\mathcal{Q}$, resulta

$$\left|\sum_{(\mathcal{P},T)} f - \sum_{(\mathcal{P}',T')} f\right| \leq \left|\sum_{(\mathcal{P},T)} f - \sum_{(\mathcal{Q},T'')} f\right| + \left|\sum_{(\mathcal{Q},T'')} f - \sum_{(\mathcal{P}',T')} f\right| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

mostrando que a rede $\left(\sum_{(\mathcal{P},T)} f\right)_{(\mathcal{P},T)}$ é de Cauchy, como desejado. □

Isto cumpre a promessa feita no Exemplo 1.3.28. Resta encontrar um método que trivialize o cálculo de algumas integrais — a saber, o Teorema Fundamental do Cálculo. Embora o ingrediente principal já esteja em nosso poder desde a seção anterior, demonstrá-lo agora faria com que o Teorema do Valor Intermediário fosse o último grande resultado da parte *principal* do texto — algo tremendamente anticlimático.

§1 Convergência uniforme vs. convergência simples

No capítulo anterior, a insistência em utilizar intervalos abertos tinha o objetivo implícito de naturalizar argumentos topológicos. Já os frequentes flertes com espaços de funções buscavam normalizar a ideia de pensar em funções como vetores ou pontos num espaço — e, portanto, passíveis de serem manipulados por meio de sequências e redes convergentes. Você pode até achar estranho, mas já tem feito isso com frequência.

Exemplo 2.1.7. No distante Exemplo 1.2.7, mostramos que

$$\lim_{n\to\infty} a^n = 0$$

sempre que $|a| < 1$. Outro modo de interpretar isso consiste em considerar, para cada $n \in \mathbb{N}$, a função $f_n \colon (-1,1) \to \mathbb{R}$ dada por $f_n(t) := t^n$, juntamente com a função $f \colon (-1,1) \to \mathbb{R}$ constante nula, e observar que

$$\lim_{n\to\infty} f_n(t) = f(t)$$

para todo $t \in (-1,1)$. *Em certo sentido*, os termos da sequência (de funções!) $(f_n)_{n\in\mathbb{N}}$ se *aproximam* arbitrariamente da função f. Nossa próxima tarefa será explicitar o *sentido*. ▲

Na definição a seguir, lembre-se de que Y^X indica a coleção de *todas* as funções da forma $X \to Y$.

Definição 2.1.8. Para um conjunto dirigido $\mathbb{D}$ e um subconjunto $X \subseteq \mathbb{R}$, uma função $\varphi \colon \mathbb{D} \to \mathbb{R}^X$ será chamada de **rede de funções reais** e, como de costume, será indicada por $(f_d)_{d\in\mathbb{D}}$. Para uma função $f \colon X \to \mathbb{R}$, diremos que $(f_d)_{d\in\mathbb{D}}$ **converge pontualmente**[†] para f, o que indicaremos por $f_d \to_{\mathrm{p}} f$, se ocorrer

$$\lim_{d\in\mathbb{D}} f_d(x) = f(x)$$

para todo $x \in X$. Em tais situações, é comum dizer que f é *o* **limite pontual**[‡] da rede $(f_d)_{d\in\mathbb{D}}$, de modo que se pode escrever $\lim_{d\in\mathbb{D}} f_d$ para indicar a própria função f. ¶

Exemplo 2.1.9. Com as notações do exemplo anterior, temos $f_n \to_{\mathrm{p}} f$. Analogamente, fazendo

$$\begin{aligned} g_n \colon (-1,1] &\to \mathbb{R} \\ t &\mapsto t^n \end{aligned} \quad \text{e} \quad \begin{aligned} g \colon (-1,1] &\to \mathbb{R} \\ t &\mapsto \begin{cases} 0, & \text{se } t \in (-1,1) \\ 1, & \text{caso contrário} \end{cases} \end{aligned}$$

resulta $g_n \to_{\mathrm{p}} g$. Todavia, ao definir $h_n \colon [-1,1] \to \mathbb{R}$ por $h_n(t) := t^n$ para todo $t \in [-1,1]$, não existe $h \colon [-1,1] \to \mathbb{R}$ tal que $h_n \to_{\mathrm{p}} h$: uma função h em tais condições deveria ser tal que $h(-1) \in \mathbb{R}$ com $h(-1) = \lim_{n\to\infty}(-1)^n$, mas já sabemos que esse limite não existe. ▲

Exercício 2.35 ((!?))**.** Esboce gráficos de funções da forma $t \mapsto t^n$ (no Geogebra, por exemplo) e observe o que acontece quando n aumenta. ■

A sequência $(g_n)_{n\in\mathbb{N}}$ acima é menos *inútil* do que parece: note que cada termo da sequência é uma função contínua, mas mesmo assim *o* limite da sequência é a função descontínua g. Um fato curioso acerca disso é que o *próprio* Cauchy não se deu conta de que continuidade não é preservada por limites pontuais — mas em defesa dele, algumas décadas se passaram até que Weierstrass percebesse exatamente como resolver o problema[††]. Nesse sentido, é instrutivo esquecer do exemplo anterior e mergulhar com fé na

Conjectura de Cauchy. *Se cada $f_n \colon X \to \mathbb{R}$ é contínua em $x_0 \in X$ e f é limite pontual de $(f_n)_{n\in\mathbb{N}}$, então f é contínua em x_0.*

"Demonstração". Com $\varepsilon > 0$ tomado arbitrariamente, precisa-se encontrar $\delta > 0$ tal que a ocorrência de $|x_0 - x| < \delta$ com $x \in X$ acarrete $|f(x_0) - f(x)| < \varepsilon$. Ora, pela desigualdade triangular, para cada $n \in \mathbb{N}$ e $x \in X$ pode-se fazer

$$|f(x_0) - f(x)| \leq |f(x_0) - f_n(x_0)| + |f_n(x_0) - f_n(x)| + |f_n(x) - f(x)|.$$

Dado que $\lim\limits_{n\to\infty} f_n = f$ e cada f_n é contínua em x_0, todas as parcelas acima podem ser feitas menores do que $\frac{\varepsilon}{3}$ para n suficientemente grande e x suficientemente próximo de x_0. Logo, f é contínua em x_0. □

Exercício 2.36 (?!)**.** Tente encontrar o erro na "demonstração" acima antes de prosseguir. Sugestão: faça isso por cinco minutos e, se não conseguir, avance. ■

[†] Também chamada de *convergência simples*.

[‡] O uso do artigo definido "o" é justificado pelo Exercício 2.111.

[††] A princípio, um "rapaz" chamado Abel percebeu a falha no argumento de Cauchy, bem como alguns exemplos em que o argumento funcionava. Weierstrass conseguiu isolar as propriedades comuns de tais exemplos — e assim se chegou à noção de convergência uniforme.

Eis o que está acontecendo. Realmente, cada parcela na última desigualdade pode ser controlada em virtude das hipóteses. Porém, não há garantias de que o controle seja simultâneo! Mais precisamente, pela continuidade de f_n em x_0, encontra-se $\delta_n > 0$ de modo que $|x_0 - x| < \delta_n$ com $x \in X$ *obriga* $|f_n(x_0) - f_n(x)| < \frac{\varepsilon}{3}$. Analogamente, pela convergência pontual, há $N_0 \in \mathbb{N}$ suficientemente grande para o qual $n \geq N_0$ acarreta $|f(x_0) - f_n(x_0)| < \frac{\varepsilon}{3}$. Contudo, a parcela $|f_n(x) - f(x)|$ *pode* não ser controlada pelo mesmo N_0: sabemos que existe $N_x \in \mathbb{N}$ para o qual $|f_n(x) - f(x)| < \frac{\varepsilon}{3}$ sempre que $n \geq N_x$, mas N_x depende de x! Isto inviabiliza a escolha de um $\delta > 0$ apropriado que dependa apenas do ponto x_0 e da constante $\varepsilon > 0$.

O problema na argumentação acima desaparece se, na definição da convergência pontual, exigirmos que o mesmo $N \in \mathbb{N}$ ateste a convergência $f_n(x) \to f(x)$ para todo $x \in X$. Em outras palavras, trata-se de uma instância da próxima

Definição 2.1.10. Nas condições da Definição 2.1.8, diremos que a rede $(f_d)_{d\in\mathbb{D}}$ **converge uniformemente para** f, o que será abreviado por "$f_d \to_{\mathrm{u}} f$", se para todo $\varepsilon > 0$ existir $D \in \mathbb{D}$ tal que $|f_d(x) - f(x)| < \varepsilon$ para quaisquer $x \in X$ e $d \succeq D$. ¶

Não se atribui uma notação específica para f em tais casos pois a notação anterior já serve:

Exercício 2.37 ($^{\star}_{\star}$)**.** Mostre que se toda pessoa conhece alguém que todas as pessoas conhecem, então quaisquer duas pessoas conhecem pelo menos uma pessoa em comum. Conclua que, nas condições anteriores, se $f_d \to_{\mathrm{u}} f$, então $f_d \to_{\mathrm{p}} f$. ■

Teorema 2.1.11. *Sejam $X \subseteq \mathbb{R}$ um subconjunto, $f\colon X \to \mathbb{R}$ uma função e $x_0 \in X$ um ponto. Se $(f_d)_{d\in\mathbb{D}}$ é uma rede de funções da forma $X \to \mathbb{R}$, todas contínuas em x_0 e $f_d \to_{\mathrm{u}} f$, então f é contínua em x_0.*

Demonstração. Como acima, para quaisquer $x \in X$ e $d \in \mathbb{D}$ se verifica

$$|f(x_0) - f(x)| \leq |f(x_0) - f_d(x_0)| + |f_d(x_0) - f_d(x)| + |f_d(x) - f(x)|.$$

Agora, para $\varepsilon > 0$ fixado, a convergência uniforme assegura um índice $D \in \mathbb{D}$ tal que $|f_d(y) - f(y)| < \frac{\varepsilon}{3}$ para quaisquer $y \in X$ e $d \succeq D$, o que em particular vale para x_0. Por sua vez, f_D é contínua em x_0, o que garante um *intervalo* aberto $I \subseteq \mathbb{R}$ com $x_0 \in X \cap I$ tal que $|f_D(x_0) - f_D(x)| < \frac{\varepsilon}{3}$ sempre que $x \in I \cap X$. Portanto, se $x \in I \cap X$, então

$$|f(x_0) - f(x)| \leq |f(x_0) - f_D(x_0)| + |f_D(x_0) - f_D(x)| + |f_D(x) - f(x)| < \varepsilon,$$

como desejado. □

Continuidade Uniforme $\neq$ Convergência Uniforme. Cuidado! As expressões se parecem, mas têm significados diferentes. Na primeira, o "uniforme" se refere ao $\delta > 0$ que assegura a condição de continuidade para um certo $\varepsilon > 0$, que deve servir *uniformemente* para todos os pontos de X. Na segunda, o "uniforme" se refere ao índice $D \in \mathbb{D}$ que assegura a condição de convergência para um certo $\varepsilon > 0$, que deve servir *uniformemente* para todos os pontos de X.

Exemplo 2.1.12. Evidentemente, a sequência $(g_n)_{n\in\mathbb{N}}$ do Exemplo 2.1.9 converge pontualmente para g, mas não converge uniformemente, caso contrário g seria contínua. No entanto, a mera continuidade do limite pontual não assegura a uniformidade da convergência: com $(f_n)_{n\in\mathbb{N}}$ como no Exemplo 2.1.7, note que por valer $\lim\limits_{x\to 1^-} x^n = 1$ para qualquer $n \in \mathbb{N}$, segue que para $\varepsilon := \frac{1}{2}$ existe $\delta > 0$ tal que $|x^n - 1| < \varepsilon$ sempre que $1 - \delta < x < 1$, mas isto acarreta $|f_n(x) - f(x)| > \frac{1}{2}$ (percebeu?)*, mostrando que nenhum $n \in \mathbb{N}$ pode funcionar uniformemente para todo $x \in (-1, 1)$. ▲

Exercício 2.38 $(\substack{\star\\\star})$**.** Para $x \in (0,1]$ e $n \in \mathbb{N}$, defina $f_n(x) := (1-x)^n$. Mostre que existe uma função $f\colon (0,1] \to \mathbb{R}$ tal que $\displaystyle\sum_{n=0}^{\infty} f_n = f$. A função f é contínua? ■

Embora pareça novidade, a convergência uniforme já deu as caras no texto em outras ocasiões (cf. Proposição 1.3.34).

Proposição 2.1.13. *Sejam $(f_n)_{n\in\mathbb{N}}$ uma sequencia de funções reais, todas limitadas e definidas em $X \subseteq \mathbb{R}$. Se existe $f\colon X \to \mathbb{R}$ tal que $f_n \to_{\mathrm{u}} f$, então $f \in \mathcal{B}(X)$ e $f_n \to f$ como sequência do espaço normado $(\mathcal{B}(X), \|\cdot\|_\infty)$. Reciprocamente, se f é limitada e $f_n \to f$ em $(\mathcal{B}(X), \|\cdot\|_\infty)$, então $f_n \to_{\mathrm{u}} f$.*

Demonstração. Primeiro, se $f_n \to_{\mathrm{u}} f$, então $|f(x)| \leq |f_n(x)| + |f(x) - f_n(x)|$ para quaisquer $x \in X$ e $n \in \mathbb{N}$. Em particular, tomando $N \in \mathbb{N}$ tal que $|f(x) - f_N(x)| < 1$ para todo x, segue que $\|f\|_\infty \leq \|f_N\|_\infty + 1$, mostrando que $f \in \mathcal{B}(X)$. Agora, provaremos que $f_n \to f$ em $\mathcal{B}(X)$. Com efeito, fixados $\varepsilon, \varepsilon' > 0$ com $\varepsilon' < \varepsilon$, existe $N \in \mathbb{N}$ tal que $|f_n(y) - f(y)| < \varepsilon'$ para quaisquer $n \geq N$ e $y \in X$. Logo, para tais valores de n, $\sup_{x\in X} |f_n(x) - f(x)| \leq \varepsilon' < \varepsilon$, i.e., $\|f_n - f\|_\infty < \varepsilon$.

Reciprocamente, se $f_n, f \in \mathcal{B}(X)$ para cada $n \in \mathbb{N}$ e $f_n \to f$ enquanto sequência do espaço normado $(\mathcal{B}(X), \|\cdot\|_\infty)$, então para cada $\varepsilon > 0$ existe $N \in \mathbb{N}$ tal que $n \geq N$ acarreta $\|f_n - f\|_\infty = \sup_{x\in X} |f_n(x) - f(x)| < \varepsilon$. Como $|f_n(y) - f(y)| \leq \sup_{x\in X} |f_n(x) - f(x)|$ para qualquer $y \in X$, mostrou-se em particular que $|f_n(y) - f(y)| < \varepsilon$ para quaisquer $n \geq N$ e $y \in X$. Em outras palavras: $f_n \to_{\mathrm{u}} f$. □

Observação 2.1.14 (Opcional — convergência uniforme de funções ilimitadas)**.** A convergência uniforme por si só é mais geral do que a convergência em $(\mathcal{B}(X), \|\cdot\|_\infty)$ pois apenas as funções *limitadas* moram em $\mathcal{B}(X)$, enquanto a convergência uniforme não impõe restrições de limitação. Por exemplo: com $X := \mathbb{R}$, defina $f_n\colon X \to \mathbb{R}$ fazendo $f_n(x) := x + \frac{1}{2^n}$ para cada $n \in \mathbb{N}$ e $x \in X$, e note que $f_n \to_{\mathrm{u}} \mathrm{Id}_{\mathbb{R}}$, mas $f_n, \mathrm{Id}_{\mathbb{R}} \notin \mathcal{B}(\mathbb{R})$ para todo $n \in \mathbb{N}$. Com isso dito, talvez te interesse saber que existe uma *métrica* em $\mathbb{R}^X$ cuja convergência induzida é, precisamente, a convergência uniforme (cf. Exercício 2.114). △

Para que *serve* tudo isso? ~~Típica pergunta das agências de fomento~~

Corolário 2.1.15. *Para $X \subseteq \mathbb{R}$ qualquer, seja $\mathcal{C}(X)$ a coleção de todas as funções contínuas da forma $X \to \mathbb{R}$. Se X é compacto, então $\mathcal{C}(X)$ é subespaço vetorial de $\mathcal{B}(X)$. Além disso, $(\mathcal{C}(X), \|\cdot\|_\infty)$ é espaço de Banach.*

Demonstração. A inclusão $\mathcal{C}(X) \subseteq \mathcal{B}(X)$ é o Teorema de Weierstrass, enquanto a condição vetorial se deve às propriedades operatórias dos limites (cf. Exercício 1.106). Após observar que a equivalência no Exercício 1.208 permanece válida para fechados de espaços métricos $(\substack{\star\\\star\star})$, a condição de Banach/completude de $(\mathcal{C}(X), \|\cdot\|_\infty)$ decorre do Teorema 2.1.11 aliado ao Exercício 1.257 e à Proposição 1.3.34. Convença-se de que é isso mesmo $(\substack{\star\\\star})$. □

Ainda não imagina para que isso poderia servir?

§2 Séries de potência

Seja $(a_n)_{n\in\mathbb{N}}$ uma sequência de números reais. Ao considerar a função $S_n\colon \mathbb{R} \to \mathbb{R}$ que a cada $x \in \mathbb{R}$ associa o número real $\sum_{j=0}^{n} a_j x^j$, obtemos uma legítima função contínua. Não seria ótimo ter uma função S que fizesse o papel de limite das funções S_n? Mais do que isso, como cada S_n é contínua, não seria ainda melhor se a função S fosse limite uniforme de $(S_n)_{n\in\mathbb{N}}$, já que isso garantiria automaticamente a continuidade de S?

Seja $\mathscr{C} := \{x \in \mathbb{R} : (S_n(x))_{n\in\mathbb{N}}$ converge em $\mathbb{R}\}$. Evidentemente $0 \in \mathscr{C}$, o que não é grande coisa. Porém, o salto interessante vem agora: se $x \in \mathscr{C} \setminus \{0\}$, então $y \in \mathscr{C}$ sempre que $|y| < |x|$. De fato, por $x \in \mathscr{C}$, temos $a_n x^n \to 0$, o que assegura $M > 0$ tal que $|a_n x^n| \leq M$ para todo n. Logo, se y é tal que $|y| < |x|$, então

$$|a_n y^n| = |a_n x^n| \cdot \frac{|y|^n}{|x|^n} \leq M \left|\frac{y}{x}\right|^n \Rightarrow \sum_{n=0}^{\infty} |a_n y^n| \leq M \sum_{n=0}^{\infty} \left|\frac{y}{x}\right|^n < +\infty$$

onde a última desigualdade segue pois $\frac{|y|}{|x|} < 1$.

Exercício 2.39 $\left(\substack{\star\\\star}\right)$**.** Use as considerações anteriores para mostrar que $\mathscr{C}$ é um intervalo. Dica: este é um bom momento para se lembrar de que existe uma definição geral de intervalo (cf. Definição 0.8.13). ■

Definição 2.1.16. Com as notações anteriores, o **raio de convergência**[†] da **série de potências** $\sum_{n=0}^{\infty} a_n x^n$ é o supremo de $\mathscr{C}$ na reta estendida. ¶

Exercício 2.40 $(\star)$**.** Mostre que o raio de convergência R de uma série de potências é sempre maior do que ou igual a 0. Além disso, se $|x| < R$, então a série converge absolutamente no ponto x. ■

Teorema 2.1.17. *Se $R > 0$ é o raio de convergência da série de potências $\sum_{n=0}^{\infty} a_n x^n$, então a função $S\colon (-R, R) \to \mathbb{R}$ dada por $S(x) := \sum_{n=0}^{\infty} a_n x^n$ é contínua.*

Demonstração. A boa definição de S segue da condição de supremo: se $z \in (-R, R)$, então $|z| < R$ e, por conseguinte, existe $x \in \mathscr{C}$ com $|z| < x < R$, donde segue que tanto $S(z)$ quanto $S(-z)$ estão bem definidos (entendeu?)*. O poder da convergência uniforme será usado para assegurar a continuidade de S: vamos mostrar que se $\rho \in (0, R)$, então S é (uniformemente) contínua em $[-\rho, \rho]$, o que por sua vez garante a continuidade de S no intervalo $(-R, R)$ (por que isto basta? $\left(\substack{\star\\\star}\right)$)[‡].

Para tornar as notações menos ambíguas, vamos escrever $\widetilde{S} := S|_{[-\rho,\rho]}$ para indicar a restrição de S ao intervalo $[-\rho, \rho]$, e $p_n\colon [-\rho, \rho] \to \mathbb{R}$ para indicar a função dada por $p_n(x) := a_n x^n$. Uma vez que $\widetilde{S}(x) = \sum_{n=0}^{\infty} p_n(x)$ para todo $x \in [-\rho, \rho]$, basta mostrar que esta série de funções é absolutamente convergente em $(\mathcal{C}[-\rho, \rho], \|\cdot\|_\infty)$, pois este é um espaço de Banach e, portanto, vale o *Teste M de Weierstrass* (cf. Teorema 1.4.12).Vejamos: pelo modo como tomamos ρ, existe $z \in \mathscr{C}$ tal que $\rho < z < R$ e, com os argumentos do começo desta subsubseção, existe $M > 0$ satisfazendo $\|p_j\|_\infty = |a_j \rho^j| \leq M \left(\frac{\rho}{z}\right)^j := M_j$ (por quê?!)$^\star_\star$, com $\sum M_j < +\infty$. Logo, as somas parciais induzidas pela sequência $(p_n)_{n\in\mathbb{N}}$ convergem uniformemente para a função $\widetilde{S}$, como desejado. □

Observação 2.1.18 (Raio de convergência — descrição quantitativa)**.** A definição adotada para o raio de convergência é útil para algumas situações teóricas, mas do ponto de vista prático ela é pouco eficiente. Um dos métodos para resolver isso consiste em considerar a sequência auxiliar $\left(\sqrt[n]{|a_n|}\right)_{n\in\mathbb{N}}$. Note que quando tal sequência é limitada, o conjunto

$$I := \left\{\rho \in \mathbb{R} : \rho > 0 \text{ e } \sqrt[n]{|a_n|} < \frac{1}{\rho} \text{ para } n \text{ suficientemente grande}\right\}$$

[†]É relativamente comum "confundir" o raio de convergência com o "método" para determiná-lo. A definição adotada aqui segue o recente texto de Johar [18].

[‡]Se precisar de inspiração, confira o último parágrafo do Exemplo 2.1.4.

é não vazio (por quê?)* e tal que $\gamma \in I$ sempre que $0 < \gamma < \rho$ para algum $\rho \in I$ (certo?)$^{\star}_{\star}$, donde segue que I é um intervalo de $\mathbb{R}$ que contém $(0, r)$, onde $r := \sup I$.

Ocorre que com as notações anteriores, $r = R$, i.e., r é o supremo de $\mathscr{C}$.

✓ $r \leq R$ pois $(-r, r) \subseteq \mathscr{C}$: se $|x| < r$, então existe $\rho \in I$ com $|x| < \rho \leq r$ e daí

$$\sqrt[n]{|a_n x^n|} = |x| \sqrt[n]{|a_n|} < \frac{|x|}{\rho} < 1$$

para n suficientemente grande, donde a convergência decorre do Testinho da Raiz (cf. Exercício 2.65);

✓ $R \leq r$ pois $(0, R) \subseteq I$: se $0 < x < R$ então $\sum_{n=0}^{\infty} a_n x^n$ converge em $\mathbb{R}$, donde em particular segue que existe $N \in \mathbb{N}$ tal que $|a_n x^n| < 1$ para todo $n \geq N$, acarretando $\sqrt[n]{|a_n|} < \frac{1}{|x|}$.

Em particular, se $\sqrt[n]{|a_n|} \to 0$, então o raio de convergência de $\sum_{n=0}^{\infty} a_n x^n$ é $+\infty$ (certo?)*, o que se revela um critério bastante prático de decisão. △

Exemplo 2.1.19 (Revisitando a exponencial)**.** Como ilustração da discussão anterior, vamos rever a função exponencial $\exp\colon \mathbb{R} \to [0, +\infty)$ dada por

$$\exp(x) := e^x := \sum_{n=0}^{\infty} \frac{x^n}{n!}.$$

De forma bastante artesanal, mostramos que a função acima está bem definida (cf. Exemplo 1.4.10) e é contínua (cf. Exemplo 1.6.12). Mas tudo isso seguiria automaticamente com os argumentos do exemplo anterior. Com efeito, temos

$$\sqrt[n]{\left|\frac{1}{n!}\right|} = \frac{1}{\sqrt[n]{n!}},$$

e a sequência dos termos acima é limitada: na verdade, por valer $(n!)^2 \geq n^n$ para qualquer $n \geq 3$ (verifique!)$^{\star}_{\star}$, não é difícil concluir que $\sqrt[n]{n!} \geq \sqrt{n}$, donde segue que $\sqrt[n]{n!} \to +\infty$ e, consequentemente,

$$\lim_{n\to\infty} \frac{1}{\sqrt[n]{n!}} = 0.$$

Por sua vez, a igualdade acima permite mostrar ~~e você já mostrou, certo?~~ que o raio de convergência da série que define e^x é $+\infty$. Portanto, exp é contínua em todos os pontos de seu domínio! Alternativamente, no nosso caso, como já havíamos provado que $\sum \frac{x^n}{n!}$ converge para todo $x \in \mathbb{R}$, a definição *qualitativa* de raio de convergência também resulta em $R = +\infty$ (pois $\mathscr{C} = \mathbb{R}$)†. ▲

Se a praticidade do argumento acima ainda não te convenceu de que *generalizar é preciso*, considere o problema de derivar a função exponencial. Embora você já tenha feito

†Em particular, confira o Exercício 2.107.

isso[†], um modo muito prático de obter $\exp'(x) = \exp(x)$ para todo x seria derivar a série que define e^x termo a termo. Se fosse lícito fazer isso, teríamos

$$\left(\sum_{n=0}^{\infty} \frac{x^n}{n!}\right)' = \sum_{n=0}^{\infty} \frac{nx^{n-1}}{n!} = \sum_{n=1}^{\infty} \frac{x^{n-1}}{(n-1)!} = \sum_{n=0}^{\infty} \frac{x^n}{n!} = e^x.$$

Seria coincidência demais este método estar errado mas, ainda assim, produzir a resposta correta. A questão é que sua justificativa é um pouco mais delicada do que o exemplo anterior faz parecer.

Teorema 2.1.20. *Seja $(f_n)_{n\in\mathbb{N}}$ uma sequência de funções contínuas da forma $[a,b] \to \mathbb{R}$. Se*

(i) *$f\colon [a,b] \to \mathbb{R}$ é limite uniforme da sequência $(f_n)_n$,*

(ii) *cada f_n é diferenciável em (a,b), e*

(iii) *existe $g\colon (a,b) \to \mathbb{R}$ tal que $f_n' \to_{\mathrm{u}} g$,*

então f é diferenciável em (a,b) e, mais ainda, $f_n'(p) \to f'(p)$ para todo $p \in (a,b)$.

Demonstração. Pela Proposição 1.8.29, a fim de mostrar que f é diferenciável em $p \in (a,b)$, basta exibir $L\colon (a,b) \to \mathbb{R}$ contínua em p tal que $L(x)(x-p) = f(x) - f(p)$ para todo $x \in (a,b)$. Ora, pela mesma proposição, sabemos que para cada $n \in \mathbb{N}$ existe $L_n\colon (a,b) \to \mathbb{R}$ contínua em p tal que $L_n(p) = f_n'(p)$ e $L_n(x)(x-p) = f_n(x) - f_n(p)$. Se conseguirmos assegurar que a sequência de funções $(L_n)_{n\in\mathbb{N}}$ converge (em $\mathcal{B}(a,b)$) para uma função L, todo o restante seguirá, pois

$$L(x)\cdot(x-p) = \lim_{n\to\infty} L_n(x)(x-p) = \lim_{n\to\infty} (f_n(x) - f_n(p)) = f(x) - f(p),$$

com L contínua em p e $f_n'(p) = L_n(p) \to L(p) = f'(p)$.

Primeiro, $L_n \in \mathcal{B}(a,b)$ para cada $n \in \mathbb{N}$: pela continuidade, existe $\delta > 0$ tal que $|L_n(x) - L_n(p)| < 1$ sempre que $x \in (a,b)$ satisfaz $|x-p| < \delta$, o que garante a desigualdade $|L_n(x)| \leq |L_n(p)| + 1$ para tais valores de x; agora, se $|x-p| \geq \delta$, então

$$|L_n(x)| \leq \left|\frac{f_n(x) - f_n(p)}{\delta}\right| \leq \frac{2\|f_n\|_\infty}{\delta} \quad \text{(por quê?!)}^\star.$$

Segundo: $(L_n)_{n\in\mathbb{N}}$ é de Cauchy em $\mathcal{B}(a,b)$. Fixados $\varepsilon, \varepsilon' > 0$ com $\varepsilon' < \varepsilon$, para $m,n \in \mathbb{N}$ e $x \in (a,b) \setminus \{p\}$, temos

$$|L_m(x) - L_n(x)| = \frac{|(f_m(x) - f_n(x)) + (f_m(p) - f_n(p))|}{|x-p|} = |f_m'(\gamma) - f_n'(\gamma)|, \tag{2.3}$$

onde $\gamma \in (a,b) \setminus \{p\}$ é um número entre x e p dado pelo T.V.M com respeito à função diferenciável $f_m - f_n$. Entra em cena a função g do item (iii): como $f_n' \to_{\mathrm{u}} g$, existe $N \in \mathbb{N}$ tal que $|f_k'(y) - g(y)| < \frac{\varepsilon'}{2}$ para quaisquer $k \geq N$ e $y \in (a,b)$, de modo que

$$|L_m(x) - L_n(x)| \leq |f_m'(\gamma) - g(\gamma)| + |g(\gamma) + f_n'(\gamma)| < \varepsilon'$$

para quaisquer $m,n \geq N$ e $x \in (a,b)$, acarretando $\|L_m - L_n\|_\infty \leq \varepsilon' < \varepsilon$. Por fim, como $\mathcal{B}(a,b)$ é de Banach (cf. Proposição 1.3.34), existe $L \in \mathcal{B}(a,b)$ com as propriedades desejadas. $\square$

[†]No Exercício 1.236. Caso não tenha feito: perceba que com a tecnologia do capítulo anterior, tem-se $\exp'(p) = \lim_{t\to 0} \frac{e^{p+t} - e^p}{t} = e^p \lim_{t\to 0} \frac{e^t - 1}{t} = e^p$, pois $e^{p+t} = e^p e^t$ e $\lim_{t\to 0} \frac{e^t - 1}{t} = 1$.

Observação 2.1.21 (Enfraquecimento das hipóteses [25])**.** Você percebeu que a hipótese (i) foi bem pouco usada? Isto se deve ao fato de que ela é mais forte do que precisávamos. Assumindo meramente que $(f_n(c))_{n\in\mathbb{N}}$ é de Cauchy para algum $c \in [a,b]$, o resultado segue das demais hipóteses: de fato, para quaisquer $x \in [a,b]$ e $m,n \in \mathbb{N}$, se verifica

$$|f_m(x) - f_n(x)| \leq |f_m(c) - f_n(c)| + |x-c||f'_m(\gamma) - f'_n(\gamma)|,$$

onde γ é dado pelo T.V.M., o que permite mostrar que $(f_n)_{n\in\mathbb{N}}$ é de Cauchy em $\mathcal{B}[a,b]$ e, portanto, *existe* uma (única) função f como no item (i). Analogamente, a função g também é desnecessária: bastaria pedir que $(f'_n)_{n\in\mathbb{N}}$ fosse *uniformemente de Cauchy*† (para se convencer, encare atentamente as identidades em (2.3)). △

Corolário 2.1.22. *Se $R > 0$ é o raio de convergência de $f(x) := \sum_{n=0}^{\infty} a_n x^n$, então f é diferenciável em $(-R,R)$ e $f'(x) = \sum_{n=1}^{\infty} na_n x^{n-1} = \sum_{n=0}^{\infty}(n+1)a_{n+1}x^n$ para todo $x \in (-R,R)$. Em particular, R é o raio de convergência de $f'(x)$. Consequentemente, séries de potência são **suaves**, i.e., infinitamente diferenciáveis.*

Demonstração. Apesar do enunciado, convém mostrar a segunda afirmação primeiro. Para $x \in \mathbb{R}$, observe que $\sum_{n=1}^{\infty} na_n x^{n-1}$ converge se, e somente se, $x\sum_{n=1}^{\infty} na_n x^{n-1} = \sum_{n=1}^{\infty} na_n x^n$ converge. Logo, o raio de convergência da primeira série de potências é igual ao raio de convergência da segunda série de potências. Por outro lado, como os intervalos

$$I := \left\{\rho \in \mathbb{R} : \rho > 0 \text{ e } \sqrt[n]{|a_n|} < \frac{1}{\rho} \text{ para } n \text{ suficientemente grande}\right\}$$

e

$$I' := \left\{\rho' \in \mathbb{R} : \rho' > 0 \text{ e } \sqrt[n]{|na_n|} < \frac{1}{\rho'} \text{ para } n \text{ suficientemente grande}\right\}$$

têm supremos iguais, segue que $\sum_{n=0}^{\infty} a_n x^n$ e $\sum_{n=1}^{\infty} na_n x^n$ têm o mesmo raio de convergência.

Provemos o restante. Para $x \in (-R,R)$, considere ρ satisfazendo $x < \rho < R$ e $f_n \colon [-\rho,\rho] \to \mathbb{R}$ a função dada por $f_n(x) = \sum_{j=0}^{n} a_j x^j$. Como $f'_n = \sum_{j=1}^{n} ja_j x^{j-1}$ e R é o raio de convergência de $g(x) := \sum_{n=1}^{\infty} na_n x^{n-1}$, segue que $f'_n \to_{\mathrm{u}} g$, o que nos coloca nas hipóteses do teorema anterior, assegurando $f'(x) = g(x)$, como desejado. □

Exercício 2.41 ($^{\star}_{\star\star}$)**.** Prove que $R := \sup I = \sup I' := R'$. Dica: mostre as inclusões $(0,R) \subseteq I'$ e $(0,R') \subseteq I$; para a primeira inclusão será útil notar que $\lim_{n\to\infty} \sqrt[n]{n} = 1$ e, para isso, observe que $\sqrt[n]{n} \geq 1$ para todo $n \geq 1$, o que permite escrever $\sqrt[n]{n} = 1 + c_n$ com $c_n \geq 0$ e daí $n = (1+c_n)^n \geq \frac{n(n-1)}{2}c_n^2$ (pelo Binômio de Newton), o que em particular acarreta $c_n \to 0$. ■

†Isto não é a mesma coisa que dizer "$(f'_n)_{n\in\mathbb{N}}$ é de Cauchy em $\mathcal{B}(a,b)$", já que isto pressupõe a limitação das derivadas f'_ns (cf. Observação 2.1.14).

§3 Uniformidade em espaços métricos

Enquanto a noção de continuidade se adequa perfeitamente a espaços topológicos, a continuidade uniforme é bem mais delicada. Explicitamente, ela não lida diretamente com o comportamento dos pontos que se tornam arbitrariamente próximos de um ponto fixado, mas sim com os pontos que se tornam arbitrariamente próximos uns dos outros[†]. Nesse sentido, pelo menos, é fácil estender a definição para funções entre espaços métricos.

Definição 2.1.23. Para espaços métricos (X, d) e (Y, d'), uma função $f : X \to Y$ é **uniformemente contínua** se para todo $\varepsilon > 0$ existe $\delta > 0$ tal que $d'(f(a), f(b)) < \varepsilon$ sempre que $a, b \in X$ satisfazem $d(a, b) < \delta$. ¶

Exercício 2.42 ($\star$)**.** Mostre que toda função uniformemente contínua é contínua. ■

Com alguma paciência, não é difícil estender para espaços métricos os principais resultados sobre continuidade uniforme provados na seção anterior,

(i) O Teorema de Heine-Cantor permanece válido, e sua demonstração é essencialmente a mesma: basta trocar os intervalos abertos $I_x := (x - \delta_x, x + \delta_x)$ e $J_x := \left(x - \frac{\delta_x}{2}, x + \frac{\delta_x}{2}\right)$ por bolas abertas $B_x := B(x, \delta_x)$ e $C_x := B\left(x, \frac{\delta_x}{2}\right)$, respectivamente, além de substituir expressões do tipo "$|x - y| < \varepsilon$" por "$d(x, y) < \varepsilon$".

(ii) A argumentação para o teorema de extensão de funções uniformemente contínuas (Teorema 2.1.3) também é idêntico, desde que se considere $f : X \to Y$ uniformemente contínua com $X \subseteq Z$ e tanto Y quanto Z métricos completos. Em particular, o argumento para unicidade se recicla para este contexto[‡].

(iii) Para qualquer conjunto X (não necessariamente um subconjunto de $\mathbb{R}$), podemos tratar de redes de funções da forma $X \to Y$, onde Y é espaço métrico[††] e, como nos itens acima, adaptar as Definições 2.1.8 e 2.1.10 para convergência pontual e uniforme, respectivamente, de redes de funções. Com isso dito, para X espaço topológico, verifica-se de forma análoga que o limite uniforme de funções contínuas do tipo $X \to Y$ ainda é uma função contínua! Em particular, $\mathcal{C}(X)$ é subespaço fechado de $(\mathcal{B}(X), \|\cdot\|_\infty)$ sempre que X é espaço topológico compacto, donde segue que $(\mathcal{C}(X), \|\cdot\|_\infty)$ é espaço de Banach para qualquer compacto X.

Exercício 2.43 ($^{\star}_{\star}$)**.** Verifique as afirmações acima. ■

As considerações anteriores, embora inegavelmente abstratas, podem ser usadas em contextos quase corriqueiros. Para aquecer, revise atentamente os Teoremas 1.6.18, 2.1.6 e 2.1.11.

Exercício 2.44 ($^{\star}_{\star}$)**.** Revise os teoremas acima. Em seguida, demonstre o corolário a seguir. ■

[†]Lembrar da diferença entre sequências convergentes e sequências de Cauchy é um sinal de que sua intuição está melhor do que imagina.

[‡]*For fun*: usando redes em vez de sequências, você pode inclusive usar a "mesma" prova para mostrar que se X é espaço topológico, Y é espaço de Hausdorff e $f, g : X \to Y$ são funções contínuas que coincidem num denso de X, então $f = g$ ($^{\star}_{\star}$). É possível provar apenas usando abertos? Sim, mas não seria tão legal.

[††]No caso específico da convergência pontual, bastaria que Y fosse topológico.

Corolário 2.1.24 (Integração é contínua com a norma do supremo)**.** *Para* $a, b \in \mathbb{R}$ *com* $a < b$*, a função*

$$\begin{aligned} \int_a^b : \ (\mathcal{C}[a,b], \|\cdot\|_\infty) &\to \mathbb{R} \\ f &\mapsto \int_a^b f(t)\,\mathrm{d}t \end{aligned}$$

é uma (*transformação linear*) *contínua.*

O corolário acima é interessante, e o fato de conseguirmos prová-lo quase sem esforço evidencia uma das vantagens da generalização mesmo quando o problema considerado é particular. No entanto, há um problema: convergência uniforme é difícil — não (necessariamente) para entender, mas para assegurar.

Exemplo 2.1.25 (Bolas uniformes)**.** Se você ainda não tomou a iniciativa de fazer alguns desenhos por conta própria para entender o quão restritiva é a convergência uniforme, vamos fazer juntos. Fixada uma função contínua $f\colon [a,b] \to \mathbb{R}$, como *descrever* $B := B_{\|\cdot\|_\infty}(f, r)$, a *bola aberta uniforme* de centro f e raio $r > 0$ em $\mathcal{C}[a,b]$? Explicitamente, $g \in B$ se, e somente se,

$$\max\{|g(x) - f(x)| : x \in [a,b]\} = \sup_{x\in[a,b]} |g(x) - f(x)| = \|g - f\|_\infty < r.$$

Exercício 2.45 ($\star$)**.** Por que foi possível trocar "sup" por "max"? ■

Com isso, não é difícil perceber que a bola B pode ser *interpretada* como uma "faixa aberta de raio r" em torno de f. Mais precisamente, ao denotar por

$$F(f, r) := \{(x, y) \in [a,b] \times \mathbb{R} : |y - f(x)| < r\}$$

a *faixa aberta de raio* r *em torno de* f, segue que uma função contínua $g\colon [a,b] \to \mathbb{R}$ pertence à bola B se, e somente se, seu gráfico é subconjunto de $F(f, r)$, i.e.,

$$g \in B \Leftrightarrow \operatorname{graf}(g) := \{(x, y) \in [a,b] \times \mathbb{R} : y = g(x)\} \subseteq F(f, r).$$

Exercício 2.46 ($\substack{\star \\ \star}$)**.** Faça a sua parte: prove a equivalência acima e faça um desenho para interpretar geometricamente. ■

Assim, pode-se pensar que se $f_n \to_{\mathrm{u}} f$ nesse contexto, por exemplo, então para qualquer $\varepsilon > 0$ existe $N \in \mathbb{N}$ tal que $\operatorname{graf}(f_n) \subseteq F(f, \varepsilon)$ para todo $n \geq N$, ou seja: os gráficos das f_n's se tornam arbitrária e uniformemente próximos do gráfico de f. ▲

Se, por um lado, a discussão anterior ajuda a tornar o Corolário 2.1.24 mais intuitivo, por outro lado ela revela o quão restritiva é a convergência uniforme. Isto sugere buscar por outras *noções de convergência* para funções, tal qual fizemos em $\mathbb{R}^n$.

Definição 2.1.26. Para uma função contínua $f\colon [a,b] \to \mathbb{R}$, define-se

$$\|f\|_1 := \int_a^b |f(t)|\,\mathrm{d}t,$$

chamada de **norma** L_1 de f. ¶

Com tudo o que já sabemos, é quase fácil mostrar que $\|\cdot\|_1$ define uma *seminorma* em $\mathcal{C}[a,b]$, i.e.,

(i) $\|f\|_1 \geq 0$ para toda $f \in \mathcal{C}[a,b]$, bem como $\|0\|_1 = 0$;

(ii) $\|\lambda f\|_1 := \int_a^b |\lambda f(t)|\,\mathrm{d}t = |\lambda| \int_a^b |f(t)|\,\mathrm{d}t = |\lambda| \cdot \|f\|_1$ para quaisquer $\lambda \in \mathbb{R}$ e $f \in \mathcal{C}[a,b]$;

(iii) por valer $|f+g| \leq |f| + |g|$, resulta $\|f+g\|_1 \leq \|f\|_1 + \|g\|_1$.

Falta mostrar que se $\|f\|_1 = 0$, então $f = 0$, o que será bem fácil após desenvolvermos algumas técnicas para o cálculo de integrais[†]. Assumindo que $\|\cdot\|_1$ é uma norma em $\mathcal{C}[a,b]$, temos a seguinte:

Proposição 2.1.27 (Integração é contínua com a norma L_1)**.** *Para $a, b \in \mathbb{R}$ com $a < b$, a função*

$$\int_a^b : (\mathcal{C}[a,b], \|\cdot\|_1) \to \mathbb{R}$$
$$f \mapsto \int_a^b f(t)\,\mathrm{d}t$$

é uma (transformação linear) contínua.

Demonstração. Para $f\colon [a,b] \to \mathbb{R}$ contínua, temos $-|f| \leq f \leq |f|$ (cf. Subseção 0.7.1 §0) com $|f|$ contínua e daí, pela monotonicidade da integral,

$$\int_a^b -|f(t)|\,\mathrm{d}t \leq \int_a^b f(t)\,\mathrm{d}t \leq \int_a^b |f(t)|\,\mathrm{d}t,$$

i.e.,

$$\left|\int_a^b f(t)\,\mathrm{d}t\right| \leq \int_a^b |f(t)|\,\mathrm{d}t := \|f\|_1$$

donde o resultado segue da Proposição 1.6.17 (percebeu?)*. □

Na verdade, a proposição anterior *refina* o Corolário 2.1.24, pelo seguinte:

Exercício 2.47 $\binom{\star}{\star}$**.** Mostre que a função identidade $\mathrm{Id}\colon (\mathcal{C}[a,b], \|\cdot\|_\infty) \to (\mathcal{C}[a,b], \|\cdot\|_1)$ é contínua. Conclua que a continuidade da integração no Corolário 2.1.24 decorre da continuidade de Id e da proposição anterior. Dica: $-\|f\|_\infty \leq |f| \leq \|f\|_\infty$. ■

Observação 2.1.28 (Bijeções contínuas sem inversas contínuas)**.** A inversa da função acima não é contínua! Na prática, isto significa que existem sequências de funções contínuas cujas normas L_1 convergem para 0, mas cujos gráficos não convergem uniformemente para a função nula. Mais uma vez, será bem fácil verificar isso após termos ferramentas para calcular integrais de verdade. △

[†]Intuitivamente, porém, a coisa toda é simples: se $f \neq 0$, então $|f(x)| > 0$ para algum $x \in [a,b]$, e daí $|f(y)| > 0$ para todo y num intervalo aberto em torno de x (por conservação de sinal); assim, basta "desenhar" uma função contínua $g\colon [a,b] \to \mathbb{R}$ que fique "entre" 0 e $|f|$ neste intervalo aberto e cuja integral positiva seja fácil de calcular.

A discussão poderia se prolongar por, pelo menos, um livro inteiro! Não é brincadeira: ao investigar $(\mathcal{C}[a,b], \|\cdot\|_1)$ por mais tempo, acabaríamos percebendo que sua métrica não é completa, no sentido de que existem sequências de Cauchy que não convergem; isso motivaria uma breve discussão sobre *completamento de espaços normados*, o que permitiria *enxergar* $\mathcal{C}[a,b]$ como subespaço denso de um espaço de Banach $\mathscr{L}[a,b]$. Em particular, ao notar que transformações lineares contínuas são (sempre!) uniformemente contínuas, poderíamos aplicar o Teorema 2.1.3 (cf. Exercício 2.43) para estender a *integração* $\int_a^b \colon \mathcal{C}[a,b] \to \mathbb{R}$ e obter uma "transformação linear contínua"

$$\int_{[a,b]} \colon \mathscr{L}[a,b] \to \mathbb{R}.$$

Este seria o começo de uma bela jornada pela maravilhosa Teoria da Medida, já que $\mathscr{L}[a,b]$ é uma das manifestações do espaço das funções *Lebesgue-mensuráveis no intervalo* $[a,b]$ e *integráveis* (no sentido de Lebesgue). A função $\int_{[a,b]}$ acima? Ora: é a própria integral de Lebesgue!

Exercício 2.48 ($\star$)**.** Mostre que *toda transformação linear contínua* entre espaços normados é uniformemente contínua. Dica: encare a Proposição 1.6.17 até que ela te encare de volta. ■

Alternativamente, poderíamos retornar ao título desta subsubseção a fim de saber se faz sentido discutir convergência e continuidade uniformes em espaços mais gerais do que os métricos, e a resposta seria "sim": os *espaços uniformes*. Isto abriria margem para discutir *grupos topológicos* e *espaços vetoriais topológicos* como os exemplos mais importantes de espaços uniformes — e este seria o ambiente perfeito para retornar ao problema da *convergência pontual*, que costuma ser deixada de lado em textos introdutórios por conta de sua incompatibilidade com métricas.

Porém, como eu pretendo que este livro seja fechado e limitado, nada disso será tratado aqui... pelo menos não em detalhes.

2.2 Teorema do Valor Intermediário

2.2.0 Essencial

§0 Conexidade e intervalos

Por que os subconjuntos $(0,1)$ e $(2,3) \cup (3,4)$ são distintos? A depender da sua disposição e tempo para escrever, muitas respostas elaboradas podem ser concebidas. Porém, a mais intuitiva delas é fazer um desenho e perceber que $(0,1)$ tem um pedaço, enquanto $(2,3) \cup (3,4)$ tem dois. O problema é que desenhos não podem ser encarados como respostas formais e, mesmo que fossem, faltaria explicar: o que é "pedaço"? A noção de *conexidade* é a maneira profissional de responder isso.

Definição 2.2.0. Um subconjunto $X \subseteq \mathbb{R}$ é **desconexo** se existem abertos $A, B \subseteq \mathbb{R}$ disjuntos tais que $X = (X \cap A) \cup (X \cap B)$ com $X \cap A \neq \emptyset$ e $X \cap B \neq \emptyset$. Neste caso, costuma-se dizer que $X \cap A$ e $X \cap B$ determinam *uma* **cisão** de X. Subconjuntos **conexos** são aqueles que não são desconexos. ¶

Os subconjuntos da forma $X \cap A$, com $A \subseteq \mathbb{R}$ aberto e $X \cap A \neq \emptyset$, são os *pedaços* de X. Assim, dizer que X é desconexo consiste em afirmar que X pode ser descrito como a reunião de dois pedaços que não se interceptam, ou seja: X tem dois pedaços. Consequentemente, dizer que X tem só um pedaço corresponde a dizer que não se pode descrever X por meio de pedaços disjuntos. Com isso dito, no contexto dos subconjuntos de $\mathbb{R}$, a conexidade é sinônimo de algo que já temos visto desde o começo do capítulo anterior.

Exercício 2.49 ($\overset{\star}{\star}$)**.** Mostre que $X \subseteq \mathbb{R}$ é desconexo se, e somente se, existe $A \subseteq X$ aberto em X tal que $A \neq \emptyset$ e $X \setminus A \neq \emptyset$ também é aberto em X. ■

Teorema 2.2.1. *Um subconjunto $C \subseteq \mathbb{R}$ é conexo se, somente se, C é intervalo.*

Demonstração. Por absurdo, suponha que C seja intervalo desconexo. Neste caso, existe um subconjunto $W \subseteq C$, não vazio e aberto em C, tal que $C \setminus W \neq \emptyset$ também é aberto em C (pelo exercício anterior). Tomando $x \in W$ e $y \in C \setminus W$, pode-se assumir, sem perda de generalidade, que $x < y$. Agora, o subconjunto $B \coloneqq W \cap (-\infty, y) = \{w \in W : w < y\}$ é não vazio e limitado superiormente (por quê?)*, o que permite tomar $\beta \in \mathbb{R}$ com $\beta \coloneqq \sup B$, que pertence a C por este ser intervalo (certo?)*. A pergunta a se fazer então é: onde, especificamente, está β?

- ✗ Se $\beta \in W$, então por W ser aberto em C, existe $r > 0$ tal que $z \in W$ sempre que $z \in C$ satisfaz $|z - \beta| < r$ (certo?)*. Por outro lado, como $\beta \in W$ e $y \in C \setminus W$ é um limitante superior de B, resulta $\beta < y$, o que assegura a não vacuidade do intervalo (β, y), que por sua vez é subconjunto do <u>intervalo</u> C. Finalmente, ao tomar $c \in (\beta, y)$ com $|\beta - c| < r$, contraria-se a escolha de β como supremo de B, uma vez que $c \in B$ satisfaz $\beta < c$ (por quê?)$^{\overset{\star}{\star}}$.

- ✗ Resta apenas $\beta \in C \setminus W$. Neste caso, por $C \setminus W$ ser aberto em C, existe $U \subseteq \mathbb{R}$ aberto com $C \cap U = C \setminus W$ e, por conseguinte, existem $\alpha, \gamma \in \mathbb{R}$ com $\beta \in (\alpha, \gamma) \subseteq U$. Ora, como todo $c \in C$ no intervalo (α, γ) deve pertencer a $C \setminus W$, segue que α é um limitante superior de B (verifique!)$^{\overset{\star}{\star\star}}$, estritamente menor do que β, contrariando sua escolha como supremo.

As considerações acima mostram que β não pode existir, o que é absurdo pois B "tem direito" a um supremo. Portanto, se C é intervalo, então C é conexo.

Por sorte, a recíproca é mais simples: se C não é intervalo, então existem $a, b \in C$ e $c \in \mathbb{R} \setminus C$ tais que $a < c < b$, donde segue que $C = (C \cap (-\infty, c)) \cup (C \cap (c, +\infty))$, com $(-\infty, c) \cap (c, +\infty) = \emptyset$, $a \in C \cap (-\infty, c)$ e $b \in C \cap (c, +\infty)$, mostrando que C não é conexo. □

Exercício 2.50 ($\overset{\star}{\star\star}$)**.** Certifique-se de que você verificou os detalhes omitidos! Dica: faça desenhos. ■

§1 O Teorema do Valor Intermediário e aplicações

Por que tanto trabalho para reescrever a definição de intervalo?

Teorema 2.2.2 (do **v**alor **i**ntermediário, T.V.I.)**.** *Se I é um intervalo e $f : I \to \mathbb{R}$ é contínua, então* $\operatorname{im}(f)$ *é um intervalo.*

Demonstração. Pelo teorema anterior, basta mostrar que $\operatorname{im}(f)$ é conexo. Suponha que não: se A e B são abertos em $\operatorname{im}(f)$, não vazios, disjuntos e $\operatorname{im}(f) = A \cup B$, então

$$I = f^{-1}[\operatorname{im}(f)] = f^{-1}[A \cup B] = f^{-1}[A] \cup f^{-1}[B],$$

com $f^{-1}[A]$ e $f^{-1}[B]$ abertos em I, não vazios e disjuntos (verifique!)*, i.e., I seria desconexo. Absurdo! $\square$

E o que significa, na prática, dizer que $\operatorname{im}(f)$ é um intervalo? Ora, $J \subseteq \mathbb{R}$ é intervalo se $y \in J$ sempre que $x \leq y \leq z$ para $x, z \in J$. No caso de $\operatorname{im}(f)$, isto significa que se $a, b \in I$ e $\gamma \in \mathbb{R}$ é um número real *entre* $f(a)$ e $f(b)$, então $\gamma \in \operatorname{im}(f)$, e isto *permite* mostrar que existe x *entre* a e b tal que $f(x) = \gamma$. Verbalmente: um valor intermediário entre $f(a)$ e $f(b)$ deve ser imagem de um ponto entre a e b.

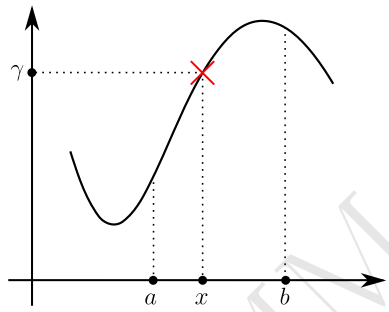


Figura 2.0: Ilustração clássica do T.V.I.

Exercício 2.51 $(^{\star}_{\star})$**.** *Permita-se*: justifique a última afirmação, i.e., prove que se $f\colon I \to \mathbb{R}$ é contínua e $a, b \in I$ com $a < b$, então para todo γ *entre*[†] $f(a)$ e $f(b)$ existe $x \in [a, b]$ tal que $f(x) = \gamma$. Dica: $[a, b]$ também é um intervalo. ■

Vamos aos refrescos!

Corolário 2.2.3 (Existência de raízes n-ésimas)**.** *Sejam $n \in \mathbb{N} \setminus \{0\}$ e $\alpha \in \mathbb{R}$.*

(i) *Se n é par e $\alpha \geq 0$, então existe único $\beta \geq 0$ tal que $\beta^n = \alpha$.*

(ii) *Se n é ímpar, então existe único $\beta \in \mathbb{R}$ tal que $\beta^n = \alpha$.*

Demonstração. Já sabemos que a função $\varphi\colon \mathbb{R} \to \mathbb{R}$ dada por $\varphi(x) := x^n$ é contínua. Pelo T.V.I., se n é par, então $\operatorname{im}(\varphi) = [0, +\infty)$: como $\varphi(0) = 0$ e $\lim\limits_{x \to +\infty} \varphi(x) = +\infty$, segue que para qualquer $\alpha > 0$ existe $x > 0$ com $\alpha < \varphi(x)$, donde o fato de $\operatorname{im}(\varphi)$ ser intervalo assegura a existência de $\beta > 0$ com $\varphi(\beta) = \alpha$. Para n ímpar, note que $\lim\limits_{x \to -\infty} \varphi(x) = -\infty$ e repita o argumento (faça isso!)*. Por fim, a unicidade é reflexo da injetividade de φ no caso de n ímpar, e da injetividade de $\varphi|_{[0,+\infty)}$ para n par, o que você certamente já verificou em algum momento da sua vida (se não, já sabe, né?)$^{\star}_{\star}$. $\square$

Definição 2.2.4. Em ambos os casos, o número β é indicado por $\sqrt[n]{\alpha}$, chamado de raiz n-ésima de α. ¶

O último corolário garante apenas a existência e unicidade das raízes n-ésimas, mas nada é dito acerca da continuidade de tais funções. Isto também seguirá automaticamente dos próximos resultados!

[†]Trata-se de um mero abuso de linguagem para indicar que $f(a) \leq \gamma \leq f(b)$ ou $f(b) \leq \gamma \leq f(a)$.

Corolário 2.2.5 (Em intervalos, injetividade + continuidade = monotonicidade)**.** *Sejam I um intervalo e $f\colon I \to \mathbb{R}$ uma função injetora. Se f é contínua, então f é monótona.*

Demonstração honesta. Se I tem apenas um ponto, acabou. No caso que interessa, existem $a, b \in I$ com $a < b$, e deve-se ter $f(a) \neq f(b)$ por conta da injetividade. Veremos que se ocorrer $f(a) < f(b)$, então f é (estritamente) crescente. Primeiro, para $x \in (a, b)$ qualquer, deve-se ter $f(a) < f(x)$, pois o contrário acarretaria $f(x) < f(a)$ (já que $f(x) = f(a)$ não pode ocorrer), e daí o T.V.I. garantiria $y \in (x, b)$ tal que $f(y) = f(a)$ (por quê?!)[†], contrariando a injetividade.

Com isso, mostraremos que f é estritamente crescente em $[a, b]$. Se não fosse o caso, existiriam $\alpha, \beta \in (a, b)$ com $\alpha < \beta$ e $f(\beta) < f(\alpha)$. Porém, pelo passo anterior, deve-se ter $f(a) < f(\beta) < f(\alpha)$, donde novamente poderíamos apelar para o T.V.I. a fim de conjurar $z \in (a, \alpha)$ tal que $f(z) = f(\beta)$, contrariando a injetividade. Portanto, f é crescente em $[a, b]$. Com argumentação análoga, mostra-se que $f(x) < f(a)$ para todo $x \in I$ com $x < a$ enquanto $f(b) < f(y)$ para todo $y \in I$ com $b < y$, donde segue que f é crescente em I. O caso em que $f(a) > f(b)$ se reduz ao anterior trocando f por $-f$. □

Demonstração desonesta (*Opcional*)*.* Note que a definição de conexidade faz sentido em quaisquer espaços topológicos quando lida em termos de abertos. Em particular, pode-se mostrar que para qualquer intervalo $I \subseteq \mathbb{R}$, o subconjunto $M := \{(x, y) \in I \times I : x < y\}$ é conexo em $\mathbb{R}^2$. E o Kiko? Veja: se f é contínua e injetora, então a função

$$\begin{aligned} g\colon \quad M &\to \mathbb{R} \\ (x, y) &\mapsto f(y) - f(x) \end{aligned}$$

é contínua e, mais importante, nunca se anula (por conta da injetividade!), donde segue que sua imagem é um intervalo que não contém o ponto 0. Ora, neste caso, ou $g(x, y) > 0$ para todo par $(x, y) \in M$ ou $g(x, y) < 0$ para todo par $(x, y) \in M$, o que explicitamente significa dizer que f é estritamente crescente ou estritamente decrescente. Para mais detalhes, confira as discussões da Subseção 2.2.1 §1. □

Corolário 2.2.6. *Sejam I um intervalo e $f\colon I \to \mathbb{R}$ uma função contínua e injetora. Se J denota a imagem de f, então a inversa $f^{-1}\colon J \to I$ é contínua.*

Demonstração. Primeiro, note que o T.V.I. assegura que J é intervalo, enquanto o último corolário garante que f é estritamente crescente ou estritamente decrescente. Agora, o restante da prova consiste em perceber se f é estritamente crescente[‡], então f leva intervalos de um tipo a intervalos do mesmo tipo: por exemplo, para $\alpha, \beta \in I$ com $\alpha < \beta$, tem-se $f[[\alpha, \beta)] = [f(\alpha), f(\beta))$ pois

$$\begin{aligned} \alpha \leq x < \beta \Rightarrow f(\alpha) \leq f(x) < f(\beta) &\qquad \text{por ser estritamente crescente} \\ f(\alpha) \leq \gamma < f(\beta) \Rightarrow \exists x \in [\alpha, \beta) \text{ tal que } f(x) = \gamma &\qquad \text{pelo T.V.I.} \end{aligned}$$

Analogamente, $f[(\alpha, \beta)] = (f(\alpha), f(\beta))$ e $f((\alpha, \beta]] = (f(\alpha), f(\beta)]$. Daí, como vale a identidade $(f^{-1})^{-1}[K] = f[K]$ para qualquer subconjunto $K \subseteq I$, conclui-se que as pré-imagens de abertos de I por f^{-1} são abertos de J (pense a respeito)$^{\star}_{\star}$. □

Corolário 2.2.7. *Para $n \in \mathbb{N} \setminus \{0\}$, a função $\sqrt[n]{(\cdot)}$ é contínua.*

[†] É extremamente recomendável fazer um desenho por conta própria para não se perder no argumento.

[‡] Para lidar com o caso estritamente decrescente, basta notar que em tal situação, $-f$ é estritamente crescente.

Corolário 2.2.8. *A inversa da função* $\exp\colon \mathbb{R} \to (0, +\infty)$ *é contínua.*

Demonstração. A continuidade da exponencial já foi estabelecida desde o Exemplo 1.6.12. Pelo Exercício 1.236, já sabemos que a função exp é injetora (pelo item (b)) e tem por imagem o intervalo $(0, +\infty)$ (pelo T.V.I.). Logo, sua inversa é contínua. □

Definição 2.2.9. A inversa da função exponencial obtida no corolário anterior é chamada de **logarítmo** e indicada por $\ln\colon (0, +\infty) \to \mathbb{R}$. ¶

Observação 2.2.10 (Bijeções contínuas sem inversas contínuas)**.** O Corolário 2.2.6 pode passar a impressão de que *toda* injeção contínua $f\colon I \to \mathbb{R}$ admite uma inversa contínua definida na imagem de f, para qualquer *subconjunto* $I \subseteq \mathbb{R}$. Mas isto é apenas fruto do ~~desligamento involuntário do cérebro~~ costume de esquecer que a reta admite subconjuntos que não são intervalos.

Por exemplo, a função $f\colon [0, 1) \cup [2, 3) \to [0, 2)$, que faz $f(x) \coloneqq x$ para $x \in [0, 1)$ e $f(x) \coloneqq x - 1$ para $x \in [2, 3)$, é contínua em todos os pontos de seu domínio, mas sua inversa não é contínua. Num primeiro momento, você pode resistir a este exemplo, argumentando que o próprio domínio de f tem um salto, o que nos traz novamente ao mantra: não trate intuições ou analogias ("ah! mas é que função contínua é aquela que eu não tiro o lápis do papel para desenhar!!!") como se fossem as definições!

No caso, o domínio de f é o subconjunto $X \coloneqq [0, 1) \cup [2, 3)$ *com a topologia de subespaço herdada de* $\mathbb{R}$, que tem como abertos todos os subconjuntos de X que são da forma $A \cap X$ para algum $A \subseteq \mathbb{R}$ aberto. Nesta altura do campeonato, você não deveria ter problemas para acreditar que f é contínua para todo $x \in [0, 1) \cup (2, 3)$. A confusão pode ocorrer no ponto 2, mas isto se resolve rápido:

(i) via ε's-δ's, note que para $\varepsilon > 0$, qualquer $x \in X$ com $|x - 2| < \min\{\varepsilon, 1\}$ satisfaz $|f(x) - f(2)| < \varepsilon$, pois da desigualdade $|x - 2| < 1$ resulta $x \in [2, 3)$, e daí $|f(x) - f(2)| = |x - 1 - 1| = |x - 2| < \varepsilon$;

(ii) via pré-imagens de intervalos abertos, note que para qualquer $\varepsilon \in (0, 1)$, temos $f^{-1}[(1 - \varepsilon, 1 + \varepsilon)] = (1 - \varepsilon, 1) \cup [2, 2 + \varepsilon)$, que é um aberto $\underline{\text{de } X}$.

Por outro lado, $f^{-1}\colon [0, 2) \to [0, 1) \cup [2, 3)$ não tem *qualquer* chance de ser contínua em todos os pontos, já que $[0, 2)$ é conexo enquanto $[0, 1) \cup [2, 3)$ não é. Se quiser praticar, verifique que f^{-1} é descontínua, precisamente, no ponto 1 $(\star)$. △

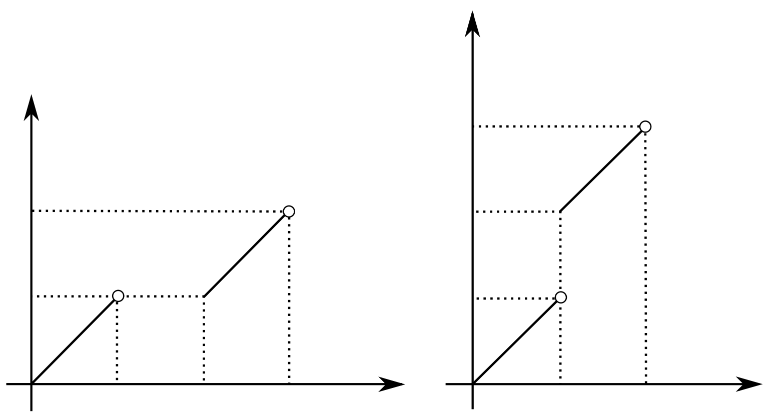


Figura 2.1: Esboço do desenho que, tenho certeza, você já havia feito antes de chegar aqui, pois as indiretas pedindo proatividade foram suficientes.

2.2.1 Extras

§0 Componentes conexas

A ideia de *pedaço* pode ter ativado sua memória muscular para a distante noção de partição 0.1.22. É bem por aí mesmo:

Lema 2.2.11. *Sejam $\Lambda \neq \emptyset$ um conjunto de índices e, para cada $\lambda \in \Lambda$, considere um subconjunto conexo $C_\lambda \subseteq \mathbb{R}$. Se existir $x \in \bigcap_{\lambda\in\Lambda} C_\lambda$, então $\bigcup_{\lambda\in\Lambda} C_\lambda$ é subconjunto conexo de $\mathbb{R}$.*

Demonstração. Com os resultados da seção anterior, basta mostrar que $C := \bigcup_{\lambda\in\Lambda} C_\lambda$ é intervalo. Ora, se $a \in C$ e $a < y \leq x$, então existe $\lambda \in \Lambda$ tal que $a \in C_\lambda$, de modo que $y \in C_\lambda$ pois C_λ é intervalo e $x \in C_\lambda$. Analogamente, se $x \leq y < b$ para algum $b \in C$, então $y \in C$. Enfim, para $a, b \in C$ quaisquer e y com $a < y < b$, deve-se ter $y \leq x$ ou $x \leq y$, donde o resultado desejado segue. □

Corolário 2.2.12. *Para $X \subseteq \mathbb{R}$ e $x \in X$, existe subconjunto conexo $C \subseteq X$ tal que $x \in C$ com a seguinte propriedade: sempre que $D \subseteq X$ é conexo e $x \in D$, verifica-se $D \subseteq C$.*

Demonstração. Note que $\{x\}$ é um subconjunto conexo de X que contém x como elemento[†] (certo?), mostrando que $\mathscr{D} := \{D \subseteq X : x \in D \text{ e } D \text{ é conexo}\}$ é uma família não vazia de subconjuntos conexos de $\mathbb{R}$ com um ponto em comum. Logo, pelo lema anterior, $C := \bigcup_{D\in\mathscr{D}} D$ é conexo. Por construção, C satisfaz as condições impostas no enunciado. □

Exercício 2.52 $\left(^{\star}_{\star}\right)$**.** Complete os detalhes da demonstração. ■

Definição 2.2.13. O subconjunto C no corolário anterior é chamado de (a) **componente conexa** de x em X, que denotaremos por $\mathrm{Con}_X(x)$. ¶

Explicitamente, $\mathrm{Con}_X(x)$ é o maior subconjunto conexo contido em X que contém o ponto x e, justamente por ser o elemento máximo de um subconjunto numa ordem parcial, sua unicidade é garantida — o que justifica a adoção de uma notação específica. Agora, secretamente, $\mathrm{Con}_X(x)$ é a classe de equivalência de x numa relação de equivalência pouco óbvia: para $a, b \in X$, vamos escrever $a \sim_X b$ para indicar a existência de um subconjunto conexo $C \subseteq X$ tal que $a, b \in C$.

Exercício 2.53 $\left(^{\star}_{\star}\right)$**.** Mostre que $\sim_X$ é relação de equivalência em X. ■

Com isso, indicando por $\widetilde{a}^X$ a classe de equivalência de $a \in X$, temos

✓ $\widetilde{a}^X \subseteq \mathrm{Con}_X(a)$ pois, se $b \sim_X a$, então existe $C \subseteq X$ conexo tal que $a, b \in C$, acarretando $C \subseteq \mathrm{Con}_X(a)$ e, consequentemente, $b \in \mathrm{Con}_X(a)$, e

✓ $\mathrm{Con}_X(a) \subseteq \widetilde{a}^X$ pois, se $b \in \mathrm{Con}_X(a)$, então o próprio $\mathrm{Con}_X(a)$ é um conexo que testemunha $b \sim_X a$.

Uma vez que as classes de equivalência de uma relação de equivalência particionam seu domínio, resulta que a família $\{\mathrm{Con}_X(x) : x \in X\}$, das componentes conexas de X, particiona X numa família de subconjuntos conexos dois a dois disjuntos.

Exemplo 2.2.14. Para qualquer $x \in \mathbb{R}$ temos $\mathrm{Con}_{\mathbb{R}}(x) = \mathbb{R}$ pois $\mathbb{R}$ é conexo. ▲

[†]Leia com cuidado, atentando-se para as diferenças entre "$\{x\}$", "x" e "X". Não, elas não são sutis.

Exemplo 2.2.15. Para $X := (0,2) \cup (2,5)$ temos $\mathrm{Con}_X(x) = (0,2)$ sempre que $x \in (0,2)$, enquanto $\mathrm{Con}_X(x) = (2,5)$ se $x \in (2,5)$. ▲

Exemplo 2.2.16. Para $q \in \mathbb{Q}$ qualquer, $\mathrm{Con}_{\mathbb{Q}}(q) = \{q\}$. Isto segue pois nenhum subconjunto de $\mathbb{Q}$ com pelo menos dois pontos é conexo: se $C \subseteq \mathbb{Q}$ e $a, b \in C$ com $a < b$, então existe número irracional y com $a < y < b$ e daí os subconjuntos $A := (-\infty, y) \cap C$ e $B := (y, +\infty) \cap C$ determinam uma cisão de C. ▲

Exercício 2.54 ($\star$)**.** Mostre que se $C \subseteq \mathbb{R}$ é conexo e finito, então $|C| \leq 1$. Use isso para reobter a afirmação feita no exemplo anterior acerca das componentes conexas de $\mathbb{Q}$. ■

Entre outras coisas, componentes conexas permitem *descrever* completa e precisamente os abertos de $\mathbb{R}$ em termos de intervalos abertos.

Teorema 2.2.17. *Seja $A \subseteq \mathbb{R}$ um subconjunto não vazio. Se A é aberto em $\mathbb{R}$, então existe uma <u>única</u> família $\mathcal{C}$ de intervalos abertos, não vazios e dois-a-dois disjuntos, tal que $A = \bigcup \mathcal{C}$. Além disso, $\mathcal{C}$ é enumerável.*

Demonstração. O primeiro passo é mostrar $\mathrm{Con}_A(x)$ é aberto em $\mathbb{R}$ para cada $x \in A$. Ora, se $y \in \mathrm{Con}_A(x)$, então $y \in A$ e, por A ser aberto em $\mathbb{R}$, existe $\varepsilon > 0$ tal que $y \in I := (y - \varepsilon, y + \varepsilon) \subseteq A$. Como $C := \mathrm{Con}_A(x) \cup I$ é subconjunto conexo de A que contém x (certo?)*, resulta que $\mathrm{Con}_A(x) \cup I \subseteq \mathrm{Con}_A(x)$, o que equivale a $I \subseteq \mathrm{Con}_A(x)$ (percebeu?)*. Portanto, $\mathrm{Con}_A(x)$ é intervalo aberto.

Note que embora $\mathcal{C} := \{\mathrm{Con}_A(x) : x \in A\}$ seja uma partição de A, ocorrem infinitas *repetições*, no sentido de que certamente ocorre $\mathrm{Con}_A(x) = \mathrm{Con}_A(y)$ para $x \neq y$ (por quê?)$^{\star}_{\star}$. Para jogar fora os excessos, vamos apelar para o Axioma da Escolha (cf. Teorema 0.1.26)[†]: existe uma *classe de representantes* $\mathcal{R} \subseteq A$ para $\mathcal{C}$, i.e., tal que $\mathcal{C} = \{\mathrm{Con}_A(r) : r \in \mathcal{R}\}$ e $\mathrm{Con}_A(r) \cap \mathrm{Con}_A(r') = \emptyset$ sempre que $r, r' \in \mathcal{R}$ com $r \neq r'$.

Isto permite provar a unicidade de $\mathcal{C}$ de maneira indolor. Com efeito, se $\mathcal{D}$ fosse *outra* família de intervalos abertos não vazios e dois-a-dois disjuntos tal que $\bigcup \mathcal{D} = A$, então para cada $D \in \mathcal{D}$ teríamos

$$D = D \cap A = D \cap \left(\bigcup_{r \in \mathcal{R}} \mathrm{Con}_A(r) \right) = \bigcup_{r \in \mathcal{R}} D \cap \mathrm{Con}_A(r),$$

donde a conexidade de D *obrigaria* $D \cap \mathrm{Con}_A(r) \neq \emptyset$ para um único $r \in \mathcal{R}$, acarretando $D \cap \mathrm{Con}_A(r) = D$ e, por conseguinte, $D \subseteq \mathrm{Con}_A(r)$ (por quê?!)$^{\star}_{\star}$. Ao repetir o argumento num espelho, obtém-se $D = \mathrm{Con}_A(r)$ e, portanto, $\mathcal{D} \subseteq \mathcal{C}$ (verifique!)$^{\star}_{\star}$. Analogamente, mostra-se $\mathcal{C} \subseteq \mathcal{D}$ (mostre!)$^{\star}_{\star}$.

Para encerrar, por $\mathbb{Q}$ ser (enumerável e) denso em $\mathbb{R}$, deve-se ter $\mathcal{C}$ enumerável: para cada $C \in \mathcal{C}$ existe $q_C \in \mathbb{Q} \cap C$, o que define uma injeção $\mathcal{C} \to \mathbb{Q}$. Os detalhes ficam por sua conta. □

Exercício 2.55 ($^{\star}_{\star}$)**.** Complete a demonstração anterior. ■

Assim, a descrição de um aberto de $\mathbb{R}$ em termos de suas componentes conexas caracteriza completamente o aberto. Em particular, declarando

$$\mu((a,b)) := \lim_{x \to b} x - \lim_{y \to a} y \tag{2.4}$$

[†]Isto não precisaria ser feito, mas não há razões, no presente contexto, para que se evite o uso do Axioma da Escolha.

para quaisquer $a, b \in [-\infty, +\infty]$ com $a < b$, obtemos um modo bastante razoável para *medir* intervalos abertos não vazios que, pelo teorema anterior, se estende de forma natural para atribuir uma *medida* a qualquer aberto não vazio de $\mathbb{R}$. E para o caso do vazio, faremos $\mu(\emptyset) := 0$.

§1 Conexidade em outros espaços

Você já deve suspeitar que a definição de conexidade faz sentido em contexto topológico. Não se enganou:

Definição 2.2.18. Um espaço topológico X é **desconexo** se existem abertos não vazios de X, digamos U e V, tais que $X = U \cup V$ e $U \cap V = \emptyset$. Diremos que X é **conexo** se X não for desconexo. ¶

Como a noção de intervalo se perde em contextos desprovidos de ordem, não teremos mais como usá-los como muleta. Por sorte, há outros subterfúgios:

Proposição 2.2.19. *Para um espaço topológico X, são equivalentes:*

(i) *X é conexo;*

(ii) *se $F, G \subseteq X$ são fechados não vazios tais que $X = F \cup G$, então $F \cap G \neq \emptyset$;*

(iii) *se $A \subseteq X$ é aberto e fechado, i.e.,* ***faberto****, então $A \in \{\emptyset, X\}$;*

(iv) *toda função contínua $f : X \to \{0, 1\}$ é constante*[†]*.*

Demonstração. Todas as implicações serão provadas pela contrapositiva, a começar com (i) $\Rightarrow$ (ii). Se $F, G \subseteq X$ são fechados disjuntos com $X = F \cup G$, então para $U := X \setminus F$ e $V := X \setminus G$ se verifica $X = F \cup G = (X \setminus G) \cup (X \setminus F) = V \cup U \subseteq X$, com U e V abertos disjuntos não vazios, i.e., X não é conexo. Agora, se $A \subseteq X$ é faberto, com $A \notin \{\emptyset, X\}$, então X se expressa como a reunião disjunta $A \cup (X \setminus A)$, com A e $X \setminus A$ fechados disjuntos não vazios. Portanto, vale (ii) $\Rightarrow$ (iii). Por sua vez, se $f : X \to \{0, 1\}$ é contínua e não constante, então $f^{-1}[\{0\}] \neq \emptyset$ e $f^{-1}[\{0\}] \neq X$. Mas $\{0, 1\}$ é discreto e $\{0\}$ é *faberto*, donde a continuidade de f garante que $f^{-1}[\{0\}]$ também é faberto, resultando em (iii) $\Rightarrow$ (iv). Finalmente, se X não é conexo, então existem abertos não vazios e disjuntos $U, V \subseteq X$ tais que $X = U \cup V$. Daí, tomando-se $f : X \to \{0, 1\}$ como a função característica de U, tem-se f contínua e não constante. □

Exercício 2.56 ($^{\star}_{\star}$). Reescreva a demonstração anterior no seu ritmo e abra os detalhes com calma. ■

Exemplo 2.2.20. Espaços métricos enumeráveis são desconexos[‡]. De fato, para um espaço métrico enumerável X com pelo menos dois pontos, onde um deles é x, a família $A := \{d(x, y) : y \neq x\}$ é um subconjunto enumerável de $(0, \sup A)$, donde segue que existe r com $0 < r < \sup A \leq +\infty$ tal que $r \notin A$. Disto segue que $B_d(x, r) = B_d[x, r]$, i.e., a bola aberta de centro x e raio r é um faberto não trivial de X. ▲

[†] *Naturalmente*, $\{0, 1\}$ é assumido com a topologia discreta.

[‡] Exceto no caso em que o espaço tem apenas um ponto. Lembre-se de que pela definição adotada aqui (cf. Definição 0.4.7), conjuntos finitos também podem ser enumeráveis.

Olhando com atenção, o argumento acima adapta uma das ideias para provar que $\mathbb{Q}$ é desconexo, o que a princípio parecia depender mais de intervalos do que de conexos. Com isso em mente, não deve ser surpresa que vale o

Teorema 2.2.21 (do valor intermediário, caso geral)**.** *Se $f\colon X \to Y$ é função contínua sobrejetora e X é conexo, então Y é conexo.*

Demonstração. Sob tais condições, se $Y = U \cup V$, então $X = f^{-1}[U] \cup f^{-1}[V]$ e, além disso, vale a identidade $f^{-1}[U] \cap f^{-1}[V] = f^{-1}[U \cap V]$. Logo, se Y for desconexo, então X é desconexo (percebeu?)*. $\square$

Evidentemente, um subespaço Y de X é conexo se Y for conexo com a topologia herdada de X. Com isso em mente, não é complicado perceber que a hipótese de sobrejetividade anterior pode ser abandonada se trocar a tese por "$\operatorname{im}(f)$ é conexa".

Exercício 2.57 ($^{\star}_{\star}$)**.** Cuide dos detalhes. ■

Exemplo 2.2.22 (Conexidade do círculo)**.** Se você já conhece as funções trigonométricas básicas, deve saber que $\mathbb{S}^1 := \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$ é imagem da função $\varphi\colon \mathbb{R} \to \mathbb{R}^2$ dada por $\varphi(t) = (\operatorname{sen}(t), \cos(t))$. Como tal função é contínua (cf. Exemplo 1.6.16), resulta que $\mathbb{S}^1$ é conexo. Sem apelar para trigonometria, observe que a conexidade se preserva por homeomorfismos, e daí descreva um *quadrado* como imagem contínua de um intervalo — e conclua por meio do Exercício 1.255. ▲

Para provar que $\mathbb{R}^2$ e, mais geralmente, $\mathbb{R}^n$ é conexo para qualquer $n \in \mathbb{N} \setminus \{0\}$, é conveniente recordar da topologia produto, discutida na Subseção 1.7.1 §1: seus abertos são, por definição, reuniões de produtos cartesianos de abertos.

Lema 2.2.23 (cf. Lema 2.2.11)**.** *Sejam X um espaço topológico e $\mathcal{X}$ uma família de subespaços conexos de X tal que $X = \bigcup \mathcal{X}$. Se existir $x \in \bigcap \mathcal{X}$, então X é conexo.*

Demonstração. Mostraremos que X satisfaz a condição (iv) na última caracterização de conexidade. Ora, dada uma função contínua $f\colon X \to \{0, 1\}$, para concluir que f é constante basta tomar $y \in X \setminus \{x\}$ e verificar que $f(y) = f(x)$. Por hipótese, existe $Y \in \mathcal{X}$ tal que $y \in Y$. Daí, como a restrição $f|_Y := f_Y\colon Y \to \{0, 1\}$ é contínua e Y é conexo, f_Y deve ser constante, donde se obtém $f(y) = f(x)$, como queríamos. $\square$

Corolário 2.2.24. *Se X e Y são conexos, então $X \times Y$ é conexo.*

Demonstração. Para $y \in Y$ fixado, o subconjunto $U_x := (X \times \{y\}) \cup (\{x\} \times Y)$ é (subespaço de $X \times Y$) conexo para cada $x \in X$, já que $X \times \{y\}$ e $\{x\} \times Y$ são conexos (certo?)$^{\star}_{\star}$ e têm interseção não vazia. Daí, por ocorrer $X \times Y = \bigcup_{x \in X} U_x$ com $X \times \{y\} \subseteq \bigcap_{x \in X} U_x$, resulta que $X \times Y$ é conexo[†]. $\square$

Para não prolongar demais o assunto, observe que o lema anterior também garante que podemos definir componentes conexas em espaços topológicos por meio da mesma condição de maximalidade (observe!)$^{\star}_{\star\star}$. Isto traz para o jogo uma ferramenta muito útil:

Lema 2.2.25. *Se para quaisquer $x, y \in X$ existir uma função contínua $\gamma\colon [0, 1] \to X$ tal que $\gamma(0) = x$ e $\gamma(1) = y$, então X é conexo.*

[†]Faça um desenho! ($^{\star}_{\star}$)

Demonstração. Fixado $x \in X$, a hipótese do enunciado permite mostrar que todo ponto $y \in X$ pertence à componente conexa de x, já que $\operatorname{im}(\gamma)$ é um conexo que contém x e y. Portanto, X é conexo. $\square$

Espaços satisfazendo a hipótese do lema anterior são chamados de *conexos por caminhos*. Eles entraram em cena nesta breve introdução apenas para cumprir uma promessa, mas não se engane: conexidade por caminhos é bem mais útil do que esta mera aplicação faz parecer.

Exercício 2.58 ($^\star_\star$)**.** Um subconjunto C de um espaço vetorial é chamado de **convexo** se $tx + (1-t)y \in C$ para quaisquer $x, y \in C$ e $t \in [0,1]$. Mostre que se C é convexo, então C é conexo. Dica: mostre que C é conexo por caminhos. $\blacksquare$

Exemplo 2.2.26 (A promessa[†])**.** *Para qualquer intervalo $I \subseteq \mathbb{R}$, o conjunto*

$$M := \{(x,y) \in I \times I : x < y\}$$

é subconjunto conexo de $\mathbb{R}^2$. Por conta do que já se discutiu, basta provar que M é convexo. Para isso, precisa-se mostrar que $t(x,y) + (1-t)(a,b) \in M$ para quaisquer pontos $(x,y), (a,b) \in M$ e $t \in [0,1]$. O resultado é óbvio[‡] para $t \in \{0,1\}$. Para $0 < t < 1$, note que por valer $x < y$ e $a < b$, deve-se ter $tx < ty$ e $(1-t)a < (1-t)b$, acarretando $tx + (1-t)a < ty + (1-t)b$. Resta verificar que $tx + (1-t)a, ty + (1-t)b \in I$. Ora: se $x \leq a$, então $tx \leq ta$ e $(1-t)x \leq (1-t)a$, desigualdades que resultam em $x \leq tx + (1-t)a \leq a$ (verifique!)*, donde o fato de I ser intervalo assegura a pertinência desejada. Fica por sua conta mostrar que $ty + (1-t)y \in I$. $\blacktriangle$

§2 O conjunto de Cantor

Para encerrar as discussões sobre a topologia da reta e seus subconjuntos, vamos ver muito brevemente um conjunto que costuma desafiar a intuição de iniciantes.

Definição 2.2.27. O subconjunto

$$\mathfrak{C} := [0,1] \setminus \left(\bigcup_{n \in \mathbb{N}} \left(\bigcup_{k=0}^{3n-3} \left(\frac{1+3k}{3^n}, \frac{2+3k}{3^n} \right) \right) \right)$$

é chamado de **conjunto de Cantor**. ¶

Legal né? E daí?

Para uma interpretação psicologicamente menos desgastante, vejamos a descrição apresentada por Dan Ma em seu excelente blog[††]. Por simplicidade, seja $S_n := \{0,1\}^n$ a coleção das sequências de 0's e 1's da forma $\mathbb{N}_{<n} \to \{0,1\}$: assim, $S_0 = \{\emptyset\}$, $S_1 = \{(0),(1)\}$, $S_2 = \{(0,0),(0,1),(1,0),(1,1)\}$ e por aí vai. A construção de $\mathfrak{C}$ consiste em utilizar a *estrutura* subjacente de $S := \bigcup_{n \in \mathbb{N}} S_n$ para definir uma sequência decrescente $(F_n)_{n \in \mathbb{N}}$ de fechados (encaixantes!) de modo a garantir $\mathfrak{C} = \bigcap_{n \in \mathbb{N}} F_n$.

[†] Isto encerra a demonstração alternativa do Corolário 2.2.5.
[‡] Dessa vez é mesmo: se discorda, faça as contas. Aposto que concordará depois que fizer isso.
[††] https://dantopology.wordpress.com/2010/05/21/the-cantor-set-i/.

Para $n := 0$, fazemos $B_\emptyset := F_0 := [0,1]$. Para $n := 1$, fazemos $B_{(0)} := \left[0, \frac{1}{3}\right]$, $B_{(1)} := \left[\frac{2}{3}, 1\right]$ e $F_1 := B_{(0)} \cup B_{(1)}$; note que, até *agora*, i.e., para $n \leq 1$, verifica-se $F_n = \bigcup_{s \in S_n} B_s$; a ideia é repetir o processo nos passos seguintes. Para $n := 2$, por exemplo, perceba que os elementos de S_2 se obtêm a partir dos elementos de S_1 por meio da concatenação de uma nova entrada: $(0,0)$ e $(1,0)$ obtidos com a concatenação de 0, $(0,1)$ e $(1,1)$ obtidos com a concatenação de 1. Com isso em mente, $B_{(0,0)}$ e $B_{(0,1)}$ se obtêm a partir de $B_{(0)}$ por meio da exclusão do *terço médio* $\frac{1}{3} \cdot \left(\frac{1}{3}, \frac{2}{3}\right) = \left(\frac{1}{9}, \frac{2}{9}\right)$, onde $a \cdot I := \{a \cdot x : x \in I\}$, de modo que $B_{(0,0)} := \left[0, \frac{1}{9}\right]$ e $B_{(0,1)} := \left[\frac{2}{9}, \frac{1}{3}\right]$. Analogamente, exclui-se o terço médio de $B_{(1)}$, a saber $\frac{2}{3} + \frac{1}{3} \cdot \left(\frac{1}{3}, \frac{2}{3}\right) = \left(\frac{7}{9}, \frac{8}{9}\right)$, onde $a + J := \{a + y : y \in J\}$, a fim de definir $B_{(1,0)} := \left[\frac{2}{3}, \frac{7}{9}\right]$ e $B_{(1,1)} := \left[\frac{8}{9}, 1\right]$.

Exercício 2.59 $({}^{\star}_{\star})$**.** Desenhe o que foi feito acima. ■

O procedimento geral consiste em considerar o intervalo fechado $B_f \subseteq [0,1]$, para $f \in S_n$, definido num passo prévio, e daí definir $B_{(f,0)}$ e $B_{(f,1)}$ por meio da exclusão do terço médio $2\left(\sum_{i=0}^{n-1} \frac{f(i)}{3^{i+1}}\right) + \left(\frac{1}{3^n}, \frac{2}{3^n}\right)$, indicando que deve-se excluir de B_f o intervalo aberto obtido pela *translação* do intervalo aberto $\left(\frac{1}{3^n}, \frac{2}{3^n}\right)$ pelo número $2\left(\sum_{i=0}^{n-1} \frac{f(i)}{3^{i+1}}\right)$. Ao fazer isso para cada $f \in S_n$, obtêm-se B_g para cada $g \in S_{n+1}$, o que permite definir $F_{n+1} := \bigcup_{g \in S_{n+1}} B_g$.

Exercício 2.60 $({}^{\star}_{\star\star})$**.** Descreva B_g para cada $g \in S_3$. ■

Agora, não é a coisa mais difícil do mundo perceber que o conjunto de cantor $\mathfrak{C}$ definido anteriormente é, precisamente, $\bigcap_{n \in \mathbb{N}} F_n$. Além disso:

(i) $\mathfrak{C}$ é compacto por ser um fechado do compacto $[0,1]$;

(ii) para cada função $f \colon \mathbb{N} \to \{0,1\}$, $\left(B_{f|_{\mathbb{N}_{<n}}}\right)_{n \in \mathbb{N}}$ é uma sequência decrescente de intervalos fechados, limitados e não vazios, donde a compacidade de $[0,1]$ assegura $x_f \in \bigcap_{n \in \mathbb{N}} B_{f|_{\mathbb{N}_{<n}}} \subseteq \mathfrak{C}$;

(iii) dado que para funções $f, g \colon \mathbb{N} \to \{0,1\}$ distintas existe n tal que $B_{f|_n} \cap B_{g|_n} = \emptyset$, resulta que a correspondência $f \mapsto x_f$ define uma injeção $2^{\mathbb{N}} \to \mathfrak{C}$, mostrando que $\mathfrak{C}$ é, de fato, *infinito*;

(iv) por fim, não é difícil se convencer de que qualquer aberto $A \subseteq \mathbb{R}$ é tal que $A \cap \mathfrak{C}$ é um faberto em $\mathfrak{C}$, de modo que, com um pouco mais de reflexão, conclui-se que para todo $x \in \mathfrak{C}$, a componente conexa de x em $\mathfrak{C}$ é $\{x\}$.

Exercício 2.61 $({}^{\star}_{\star\star})$**.** Complete os detalhes. ■

Em certo sentido, cada função $f \colon \mathbb{N} \to \{0,1\}$ funciona como o endereço de um ponto em $\mathfrak{C}$: quanto maior o índice n, menor será o fechado $B_{f|_{\mathbb{N}_{<n}}}$ correspondente, e maior a precisão de sua localização. Sob essa ótica, a coisa não é tão diferente da ideia de representação decimal de um número real, que será abordada na seção de exercícios.

§3 Exercícios extras

Exercício 2.62 $({}^{\star}_{\star\star})$**.** Sejam X um espaço topológico, bem como subconjuntos A e B de X. Mostre que se $A \subseteq B \subseteq \overline{A}$ com A conexo, então B é conexo. Dica: use redes e funções contínuas (cf. Proposição 2.2.19)! ■

Exercício 2.63 $({}^{\star}_{\star})$**.** Suponha que $*$ seja uma operação na reta estendida que faça de tal conjunto um grupo, e seja $\mu\colon [-\infty, +\infty] \to [-\infty, +\infty]$ a função em que $\mu(x)$ é o $*$-inverso de x, para cada $x \in [-\infty, +\infty]$. Mostre que $*$ e μ não podem ser simultaneamente contínuas. Dica: de modo geral, se G é um grupo com uma topologia cujas *operações* (a binária e a inversão) são contínuas, então para quaisquer $x, y \in G$ existe homeomorfismo $\varphi\colon G \to G$ tal que $\varphi(x) = y$. ■

Exercício 2.64 $({}^{\star}_{\star\star})$**.** O conjunto de Cantor é conexo? ■

2.3 Exercícios adicionais

Essenciais

Exercício 2.65 (Testinho da Raiz — $({}^{\star}_{\star})$)**.** Seja $(a_n)_{n\in\mathbb{N}}$ sequência de números reais. Mostre que se existir $c \in (0,1)$ tal $\sqrt[n]{|a_n|} \leq c$ para n suficientemente grande, então $\sum a_n$ é absolutamente convergente. Dica: $\sum c^n$ converge. ■

Exercício 2.66 $({}^{\star}_{\star})$**.** Mostre que se $X \subseteq \mathbb{R}$ é limitado e $f\colon X \to \mathbb{R}$ é uniformemente contínua, então f é limitada. Dica: lembre-se de que funções uniformemente contínuas admitem extensões contínuas (cf. Teorema 2.1.3). ■

Exercício 2.67 $(\star)$**.** O resultado anterior permanece válido se f for apenas contínua? ■

Exercício 2.68 $({}^{\star}_{\star})$**.** Sejam $a, b \in \mathbb{R}$ com $a < b$. Mostre que $f\colon (a,b) \to \mathbb{R}$ é uniformemente contínua se, e somente se, f é contínua e os limites laterais $\lim\limits_{x\to a^+} f(x)$ e $\lim\limits_{x\to b^-} f(x)$ existem na reta real. Dica: cf. Exercício 2.66 e o Teorema de Heine-Cantor. ■

Exercício 2.69 $({}^{\star}_{\star})$**.** Suponha que toda sequência em $X \subseteq \mathbb{R}$ admita subsequência de Cauchy. Mostre que $\overline{X}$ é compacto. ■

Exercício 2.70 $(\star)$**.** Mostre que se $X, Y \subseteq \mathbb{R}$ são conexos e $X \cap Y \neq \emptyset$, então $X \cup Y$ é conexo. Dica: não complique sua vida e lembre-se de que, na reta real, conexo = intervalo. Observação: faça este exercício mesmo se você tiver visto a versão geral deste enunciado em alguma subseção anterior, a ideia é fortalecer sua "memória muscular". ■

Exercício 2.71 $({}^{\star}_{\star})$**.** Sejam $C, D \subseteq \mathbb{R}$ subconjuntos tais que C é conexo e $C \subseteq D \subseteq \overline{C}$. Mostre que D é conexo. Dica: não invente moda, você já conhece os conexos da reta. ■

Exercício 2.72 $(\star)$**.** Mostre que se $\left\{\frac{1}{2^n} : n \in \mathbb{N}\right\} \cup \{0\}$ é compacto. ■

Exercício 2.73 $({}^{\star}_{\star})$**.** Pense rápido: se $(x_d)_d$ é rede real tal que $x_d \to x$, com $x \in \mathbb{R}$, então $\{x_d : d \in \mathbb{D}\} \cup \{x\}$ é compacto? ■

Exercício 2.74 $({}^{\star}_{\star})$**.** Mostre que o conjunto de Cantor $\mathfrak{C}$ tem a cardinalidade de $\mathbb{R}$. ■

Exercício 2.75 (Representação decimal — $({}^{\star}_{\star\star})$)**.** Vamos formalizar a ideia de representação decimal neste exercício.

a) Para aquecer, mostre que se $(r_n)_{n\in\mathbb{N}}$ é uma sequência de números naturais entre 0 e 9, então $\displaystyle\sum_{n=0}^{\infty}\frac{r_n}{10^n}$ é um número real. Dica: use intervalos encaixantes ou o critério de Cauchy.

b) Mostre que se $r \in \mathbb{R}$, então existem $r_0 \in \mathbb{Z}$ e uma sequência $(r_n)_{n>0}$ de números naturais entre 0 e 9 tais que $r = \displaystyle\sum_{n=0}^{\infty}\frac{r_n}{10^n}$. Dica: supondo $r \geq 0$ para começar, a propriedade arquimediana assegura $r_0 \in \mathbb{N}$ tal que $r_0 \leq r < r_0 + 1$, bem como $r_1 \in \{0, \ldots, 9\}$ com $r_0 + \frac{r_1}{10} \leq r < r_0 + \frac{r_1}{10} + \frac{1}{10}$, bem como $r_2 \in \{0, \ldots, 9\}$ tal que...

c) Mostre que se $x \in \mathbb{R}$ admite representações decimais distintas, então x é da forma $\dfrac{n}{10^k}$ para certos $n, k \in \mathbb{Z}$. Dica: se $x = \displaystyle\sum_{n=0}^{\infty}\frac{x_n}{10^n} = \sum_{n=0}^{\infty}\frac{y_n}{10^n}$, então $\displaystyle\sum_{n=0}^{\infty}\frac{x_n - y_n}{10^n} = 0$. ■

Exercício 2.76 $({}^{\star}_{\star\star})$**.** Generalize o exercício anterior para representações em qualquer base $b \in \mathbb{N} \setminus \{0\}$. ■

Exercício 2.77 $(\star)$**.** Mostre que se $f : X \to \mathbb{R}$ tem *gráfico fechado* (cf. Exercício 1.260) e sua imagem está contida num compacto, então f é contínua. O resultado permanece válido se X for compacto e f tiver imagem ilimitada? ■

Exercício 2.78 $({}^{\star}_{\star})$**.** Dizemos que uma função $f : \mathbb{R} \to \mathbb{R}$ é **fechada** se $f[G]$ é fechado em $\mathbb{R}$ para todo $G \subseteq \mathbb{R}$ fechado. Mostre que funções polinomiais são fechadas. Dica: mostre que se existe $(x_n)_{n\in\mathbb{N}}$ em G tal que $f(x_n) \to y$, então $(x_n)_{n\in\mathbb{N}}$ *tem que ser limitada* (por quê?!). ■

Exercício 2.79 $({}^{\star}_{\star})$**.** Dê exemplos de funções $\mathbb{R} \to \mathbb{R}$ que não são fechadas. Sugestão: considere a função $g(x) := \dfrac{1}{x^2+1}$. ■

Exercício 2.80 $({}^{\star}_{\star})$**.** Mostre que se $f : [0,1] \to [0,1]$ é contínua, então existe $x \in [0,1]$ tal que $f(x) = x$. Dica: encare a função contínua $g(x) = x - f(x)$ até que ela te encare de volta. ■

Exercício 2.81 (Darboux - $({}^{\star}_{\star})$)**.** Seja $f : [a,b] \to \mathbb{R}$ diferenciável.

a) Mostre que se $f'(a) < 0$, então existe $x \in (a,b)$ tal que $f(x) < f(a)$. Dica: conservação de sinal + definição da derivada.

b) Faça o mesmo assumindo $f'(b) > 0$.

c) Mostre que se $f'(a) < 0 < f'(b)$, então existe $c \in (a,b)$ tal que $f'(c) = 0$. Dica: pelo Teorema de Weierstrass, existe $c \in [a,b]$ em que f assume valor mínimo; note que pelos itens acima, deve-se ter $a < c < b$, e daí aplique o Exercício 1.202.

d) Mostre que se $f'(a) < d < f'(b)$, então existe $c \in (a,b)$ com $f'(c) = d$. Dica: considere $g(x) = f(x) - dx$ no lugar de f no item anterior. ■

Exercício 2.82 $({}^{\star}_{\star})$**.** Se $f : I \to \mathbb{R}$ é diferenciável no intervalo I e existe $M > 0$ tal que $|f'(x)| \leq M$ para todo $x \in I$, então $|f(x) - f(y)| \leq M \cdot |x - y|$ para quaisquer $x, y \in I$. Em particular, mostre que f é uniformemente contínua. Dica: T.V.M. ■

Exercício 2.83 $({}^{\star}_{\star})$**.** Mostre que $f(x) \coloneqq \sqrt{x}$ é uniformemente contínua em $[0, +\infty)$ e conclua que a recíproca da parte final do exercício anterior é falsa. Dica: observe que $|\sqrt{x} - \sqrt{y}|^2 \leq |x - y|$ pois, mais geralmente, $|a - b| \leq a + b$ sempre que $a, b \geq 0$. ■

Exercício 2.84 (Teste da derivada — $({}^{\star}_{\star})$)**.** Seja $f \colon I \to \mathbb{R}$ diferenciável no intervalo I.

a) Mostre que f é crescente[†] se, e somente se, $f'(x) \geq 0$ para todo $x \in I$. Dica: a "ida" é o Exercício 1.199; para a volta, use o T.V.M.

b) Mostre ainda que se $f'(x) > 0$ para todo $x \in I$, então f é bijetora, sua imagem é um intervalo J e a inversa g de f é diferenciável em todo ponto $y \in J$, com $g'(f(x)) = \frac{1}{f'(x)}$ para todo $x \in I$. Dica: perceba que o item anterior permite mostrar que f será estritamente crescente e, portanto, injetora, fazendo com que f caia nas hipóteses do Corolário 2.2.6; para encerrar, basta aplicar o Teorema 1.8.32). ■

Exercício 2.85 $({}^{\star}_{\star})$**.** Adapte o exercício anterior para o caso decrescente. ■

Exercício 2.86 (Pontos críticos — $({}^{\star}_{\star\star})$)**.** Dizemos que $p \in X$ é **ponto crítico** de uma função diferenciável $f \colon X \to \mathbb{R}$ se $f'(p) = 0$.

a) Mostre que extremos locais de uma função diferenciável são pontos críticos.

b) Mostre que a recíproca do item anterior é falsa em geral. Dica: $f(x) \coloneqq x^3$.

c) Mostre que se $f \colon \mathbb{R} \to \mathbb{R}$ é diferenciável e $f' \colon \mathbb{R} \to \mathbb{R}$ é contínua, então o conjunto dos pontos críticos de f é fechado. Dica: $\{0\}$ é fechado em $\mathbb{R}$.

d) Mostre que se I é intervalo aberto, $f \colon I \to \mathbb{R}$ é diferenciável em I e $c \in I$ é ponto crítico tal que $f''(c) \neq 0$, então c é extremo local. Dica: conservação de sinal + exercício anterior. ■

Exercício 2.87 $({}^{\star}_{\star})$**.** Seja $f \colon (a, b) \to \mathbb{R}$ limitada e diferenciável. Mostre que se $\lim\limits_{x \to a^+} f(x)$ ou $\lim\limits_{x \to b^-} f(x)$ não existem, então $f' \colon (a, b) \to \mathbb{R}$ é ilimitada. Dica: confira o Exercício 2.82, especialmente a última parte. ■

Exercício 2.88 $({}^{\star}_{\star})$**.** Prove que toda função polinomial real não constante e de grau ímpar admite pelo menos uma raiz real. Dica: T.V.I. ■

Exercício 2.89 $({}^{\star}_{\star\star})$**.** Mostre que se $h \colon \mathbb{R} \to \mathbb{R}$ é um homeomorfismo (estritamente) decrescente, então existe $x \in \mathbb{R}$ tal que $h(x) = x$. Dica: suponha $h(0) \neq 0$ e investigue a função $g \coloneqq \mathrm{Id}_{\mathbb{R}} - h$ em dois pontos espertos, sendo que um deles você já tem. ■

Exercício 2.90 $({}^{\star}_{\star})$**.** Mostre que não existe homeomorfismo $h \colon [0, 1] \to [0, 1]$ tal que $h(0) = \frac{1}{2}$. Dica: suponha que existe. ■

Exercício 2.91 $({}^{\star}_{\star})$**.** Mostre que a soma de funções uniformemente contínuas é contínua, e que o produto de funções uniformemente contínuas limitadas também é uniformemente contínuo. ■

Exercício 2.92 $(\star)$**.** Mostre que a composição de funções uniformemente contínuas é uniformemente contínua. ■

[†]Lembre-se: "não decrescente" para quem seguir o Elon [24, 25].

Exercício 2.93 $(^{\star}_{\star})$**.** Mostre que se $f, g\colon X \to \mathbb{R}$ são uniformemente contínuas, então $\max\{f, g\}$ e $\min\{f, g\}$ são uniformemente contínuas. ■

Exercício 2.94 $(^{\star}_{\star})$**.** Investigue as derivadas das funções raízes n-ésimas nos pontos em que estão definidas. Em particular, mostre que nenhuma delas é diferenciável em 0. Por fim, prove que todas são estritamente crescentes. Dica: use o Teorema da Função Inversa (cf. Teorema 1.8.32) para as duas primeiras partes. ■

Exercício 2.95 $(^{\star}_{\star\star})$**.** Para um intervalo aberto $J \subseteq \mathbb{R}$ tal que $0 \in J$ e uma função $\rho\colon I \to \mathbb{R}$ n-vezes diferenciável, mostre que $\rho^{(i)}(0) = 0$ para todo $i \leq n$ se, e somente se, $\lim\limits_{h\to 0} \dfrac{\rho(h)}{h^n} = 0$, onde $\rho^{(i)}(0)$ indica a i-ésima derivada de ρ em 0. Sugestão: siga as etapas a seguir.

a) Note que para $i \leq 1$ a equivalência segue quase automaticamente da definição de derivada/continuidade.

b) Supondo a implicação válida para $n \geq 1$, note que se ρ é $n+1$ vezes diferenciável e $\rho^{(i)}(0) = 0$ para todo $i \leq n+1$, então $r := \rho'$ satisfaz a hipótese de indução e, por isso $\lim\limits_{h\to 0} \dfrac{r(h)}{h^n} = 0$. Use isso, juntamente com o T.V.M., para mostrar que para $\varepsilon > 0$ existe $\delta > 0$ tal que $|h| < \delta$ acarreta $\left|\dfrac{\rho(h)}{h^{n+1}}\right| < \varepsilon$. Dica: $\left|\dfrac{\rho(h)}{h^{n+1}}\right| = \dfrac{1}{|h|^n} \cdot \left|\dfrac{\rho(h) - \rho(0)}{h}\right|$ para todo $h \neq 0$, cujo segundo fator pode ser estimado por $\rho'(c) = r(c)$ para algum c satisfazendo $0 < |c| < |h|$, pelo T.V.M.

c) Supondo a recíproca válida para $n \geq 1$, para ρ satisfazendo $n+1$ vezes diferenciável satisfazendo $\lim\limits_{h\to 0} \dfrac{\rho(h)}{h^{n+1}} = 0$, investigue a função $\varphi(x) := \rho(x) - \dfrac{\rho^{(n+1)}(0)}{(n+1)!}x^{n+1}$. Dica: note que φ satisfaz a hipótese de indução e, por conta disso, $\varphi(0) = \varphi'(0) = ... = \varphi^{(n)} = 0$; como também vale $\varphi^{(n+1)}(0) = 0$, você pode usar o item anterior para concluir que $\lim\limits_{h\to 0} \dfrac{\varphi(h)}{h^{n+1}} = 0$, e daí ter uma surpresa. ■

Exercício 2.96 (Polinômios e séries de Taylor — $(^{\star}_{\star})$)**.** Sejam $I \subseteq \mathbb{R}$ um intervalo aberto e $f\colon I \to \mathbb{R}$ uma função n vezes diferenciável. Para $a \in I$, o **polinômio de Taylor de ordem n da função f em a** é

$$p(t) := f(a) + f'(a)t + \frac{f''(a)}{2!}t^2 + \ldots + \frac{f^{(n)}(a)}{n!}t^n.$$

a) Mostre que $J := \{h : a + h \in I\}$ é aberto e $0 \in J$. Dica: a função $\varphi\colon \mathbb{R} \to \mathbb{R}$ dada por $\varphi(x) := a + x$ é contínua.

b) Mostre que a função $r\colon J \to \mathbb{R}$ dada por $r(h) := f(a+h) - p(h)$ é n-vezes diferenciável em J. Dica: escreva $r(h) = f(\varphi(h)) - p(h)$, onde $\varphi(h) := a + h$, e aplique a Regra da Cadeia.

c) Mostre que a função r do item anterior satisfaz $\lim\limits_{h\to 0} \dfrac{r(h)}{h^n} = 0$. Dica: você já fez isso.

d) Mostre que se $q(t)$ é um polinômio de grau $\leq n$ tal que $\lim\limits_{h\to 0} \dfrac{f(a+h) - q(h)}{h^n} = 0$, então $p(t) = q(t)$. Dica: use o exercício anterior e lembre-se de como derivar polinômios iteradamente. ■

Observação 2.3.0. A ideia do polinômio de Taylor de f em a é servir como uma aproximação polinomial de f para pontos próximos de a. Lembre-se: conforme o *acrésimo* h se torna pequeno ("$h \to 0$"), $a + h$ se aproxima de a. Desse modo, em vez de calcular $f(a + h)$, podemos meramente calcular $p(h)$, algo possivelmente mais simples em termos computacionais. E embora possa ocorrer $p(h) \neq f(a + h)$, a diferença é *pequena*: como $|p(h) - f(a + h)| = |r(h)|$, o fato de termos $\lim_{h\to 0} \frac{r(h)}{h^n} = 0$ garante que o "erro" $r(h)$ se torna pequeno a uma *velocidade* muito maior do que o acréscimo h diminui. O que será que ocorreria com $n \to \infty$? △

Exercício 2.97 $({}^{\star}_{\star})$**.** Para um intervalo aberto $I \subseteq \mathbb{R}$ e $f\colon I \to \mathbb{R}$ suave, a **série de Taylor de f em torno de** $a \in I$ é a série

$$P(x) := \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!} x^n$$

a) Para $I := (-1, 1)$, determine a série de Taylor $P(x)$ de $f(x) = \frac{1}{1-x}$ em 0.

b) Com $P(x)$ e $f(x)$ como no item anterior, mostre que $P(x) = f(x)$ para todo $x \in I$.

c) Defina $g(x) := e^{-\frac{1}{x^2}}$ para $x \neq 0$ e $g(0) = 0$. Mostre que g é suave. Dica: perceba que $g^{(n)}(x) = g(x)p_n\left(\frac{1}{x}\right)$ para todo $x \neq 0$, onde p_n é um polinômio.

d) Mostre que a série de Taylor de g em torno de 0 se anula em todos os pontos, mas $g(x) \neq 0$ para todo $x \neq 0$. ■

Observação 2.3.1. Uma função suave $f\colon A \to \mathbb{R}$ definida num aberto $A \subseteq \mathbb{R}$ é **analítica** se para todo $a \in A$ existe um intervalo aberto $I \subseteq A$ com $a \in I$ tal que $f(x) = P(x - a)$ para todo $x \in I$, onde P é a série de Taylor de f em torno de a. Em outras palavras, tais funções são, localmente, séries de potências e, como tais, podem ser manipuladas *quase* como se fossem polinômios. △

Exercício 2.98 $({}^{\star}_{\star\star})$**.** Sejam $f, g\colon I \to \mathbb{R}$ funções analíticas definidas num intervalo I e considere $X \subseteq I$ um subconjunto que tem um ponto de acumulação $x_0 \in I$. Mostre que se $f(x) = g(x)$ para todo $x \in X$, então $f = g$. Dica: com $h := f - g$, note que $h^{(n)}(x_0) = 0$ para todo n, mostrando que $A := \{x \in I : h^{(n)}(x) = 0 \text{ para todo } n \in \mathbb{N}\}$ é não vazio; mostre então que A é aberto em I e fechado em I (cf. Exercício 2.49). ■

Exercício 2.99 $({}^{\star}_{\star})$**.** Mostre que se duas séries de potências (*centradas* no mesmo ponto) coincidem na interseção dos intervalos definidos por seus raios de convergência, então seus coeficientes coincidem. Dica: tais funções são analíticas. ■

Exercício 2.100 $({}^{\star}_{\star})$**.** Seja $f\colon \mathbb{R} \to \mathbb{R}$ uma função diferenciável em todos os pontos satisfazendo $f'(x) = f(x)$ para todo $x \in \mathbb{R}$ e $f(0) = 1$.

a) Mostre que $\lim_{x\to 0} \frac{f(x) - 1}{x} = 1$.

b) Mostre que $f(-x) = \frac{1}{f(x)}$ para todo $x \in \mathbb{R}$. Em particular, $f(x) > 0$ para todo x. Dica: defina $g(x) = f(x)f(-x)$, derive e encare o Corolário 2.0.14 até que ele te encare de volta.

c) Mostre que $f(x+y) = f(x)+f(y)$ para quaisquer $x, y \in \mathbb{R}$. Dica: para $y \in \mathbb{R}$ fixado, defina $g(x) = \dfrac{f(x+y)}{f(x)}$, derive e imite o item anterior.

d) Mostre que $\lim\limits_{x\to+\infty} f(x) = +\infty$ e $\lim\limits_{x\to-\infty} f(x) = 0$. Dica: para o primeiro, note que $f(x) > x$ para todo $x \in \mathbb{R}$ e, para o segundo, observe que $\lim\limits_{x\to-\infty} f(x) = \lim\limits_{x\to+\infty} f(-x) = \lim\limits_{x\to+\infty} \dfrac{1}{f(x)}$.

e) Enfim, mostre que $f(x) = e^x$ para todo x. Dica: $g(x) := \dfrac{f(x)}{e^x}$ e pipipi popopó. ■

Observação 2.3.2. Num primeiro momento, você pode se perguntar qual a vantagem do exercício anterior, que na prática apenas reobteve os resultados do Exercício 1.236: bem... se você realmente fez o Exercício 1.236, então certamente deve ter percebido que o exercício anterior é trivial em comparação. Nesse sentido, isto ilustra as vantagens técnicas de usar as ferramentas de convergência uniforme (e séries de potências) na definição de funções: com algumas estimativas simples (que não dependem do Exercício 1.236), o Exemplo 2.1.19 mostrou que exp é uma função contínua que satisfaz $f(0) = 1$ e, com os recursos do Corolário 2.1.22, $\exp'(x) = \exp(x)$. Num segundo momento... perceba que você (provavelmente) acabou de resolver sua primeira *equação diferencial*. △

Exercício 2.101 $(^{\star}_{\star})$**.** Mostre que a função $\ln\colon (0, +\infty) \to \mathbb{R}$ (cf. Definição 2.2.9) é suave. Além disso, $\ln'(x) = \frac{1}{x}$ para todo $x > 0$. Dica: cf. Exercício 2.84, item (b). ■

Exercício 2.102 $(^{\star}_{\star\star})$**.** Para um intervalo aberto $I \subseteq \mathbb{R}$, dizemos que $f\colon I \to \mathbb{R}$ é **convexa** se, para quaisquer $a, b \in I$ e $t \in [0, 1]$ ocorrer

$$f(ta + (1-t)b) \leq tf(a) + (1-t)f(b).$$

a) Assumindo f convexa em $I := (a, b)$, mostre que para $c \in I$, $F\colon (c, b) \to \mathbb{R}$ é crescente, onde
$$F(x) := \frac{f(x) - f(c)}{x - c}$$
para todo $x \in (c, b)$. Dica: álgebra básica?

b) Como no item anterior, mostre que $G(x) := \frac{f(c)-f(x)}{c-x}$ é crescente em (a, c).

c) Com f diferenciável, mostre que f é convexa se, e somente se, f' é crescente. Em particular, mostre que se f é duas vezes diferenciável, então f é convexa se, e somente se, $f''(x) \geq 0$ para todo $x \in I$.

d) Nas mesmas condições, $g\colon I \to \mathbb{R}$ é chamada de **côncava** se $-g$ é convexa. Adapte os resultados anteriores para funções côncavas. ■

Exercício 2.103 $(^{\star}_{\star})$**.** Para $r > 0$ e $x \in \mathbb{R}$ defina $f(x) := e^{x\ln(r)}$.

a) Mostre que $f(q) = r^q$ para todo $q \in \mathbb{Q}$. Conclua que $f(x) = r^x$ para todo $x \in \mathbb{R}$.

b) Para $z \in \mathbb{R}$, defina $g(x) := x^z$ para todo $x > 0$. Mostre $g'(x) = zx^{z-1}$ para todo $x > 0$. ■

Exercício 2.104 $({}^{\star}_{\star})$**.** Seja $(I_n)_{n\in\mathbb{N}}$ uma sequência de intervalos não vazios de $\mathbb{R}$, com $I_{n+1} \subseteq I_n$ para todo $n \in \mathbb{N}$.

a) Se todos os intervalos forem ilimitados, é possível garantir que $\bigcap_{n\in\mathbb{N}} I_n \neq \emptyset$? Por quê? Dica: mais do que filosofar sobre o imposível, é importante pensar em contraexemplos.

b) Se os intervalos não forem fechados, é possível garantir que $\bigcap_{n\in\mathbb{N}} I_n \neq \emptyset$? Dica: mais do que filosofar sobre o impossível, é importante pensar em contraexemplos.

c) Redemonstre o Corolário 2.0.19 sem usar o método apresentado no texto. Dica: você pode usar as outras caracterizações de compacidade ou, ainda, o Teorema de Bolzano-Weierstrass.

d) Definindo $\operatorname{diam}(I) := \sup I - \inf I$, mostre que se todos os intervalos forem fechados e $\operatorname{diam}(I_n) \to 0$, então existe um único ponto em $\bigcap_{n\in\mathbb{N}} I_n$. Dica: se estiver sem ideias, imite a demonstração da Proposição 2.0.20. ■

Exercício 2.105 (L'Hôpital — $({}^{\star}_{\star\star})$)**.** Sejam $f, g\colon X \to \mathbb{R}$ funções tais que $X \subseteq \mathbb{R}$, com $p \in X$ um ponto de acumulação de X. Suponha que exista $\delta > 0$ tal que

✓ f e g são diferenciáveis em $(p-\delta, p+\delta)$,

✓ $g'(x) \neq 0$ para todo $x \in (p-\delta, p+\delta) \setminus \{p\}$,

✓ $\lim\limits_{x\to p} f(x) = \lim\limits_{x\to p} g(x) = 0$ ou $\lim\limits_{x\to p} f(x) = \lim\limits_{x\to p} g(x) = \pm\infty$, e $\lim\limits_{x\to p} \dfrac{f'(x)}{g'(x)} = L \in \mathbb{R}$.

a) Sob as condições acima, mostre que $\lim\limits_{x\to p} \dfrac{f(x)}{g(x)} = L$.

b) Mostre que para $X := (a, +\infty)$ e $p := +\infty$, o resultado permanece válido se as duas primeiras condições forem substituídas por

✓ f e g serem diferenciáveis em $(a, +\infty)$, e

✓ existe $K > a$ tal que $g'(x) \neq 0$ para todo $x > K$. ■

Exercício 2.106 $({}^{\star}_{\star})$**.** Sejam $(f_n)_{n\in\mathbb{N}}$ sequência de funções contínuas da forma $X \to \mathbb{R}$, onde $X \subseteq \mathbb{R}$. Mostre que se $f_n \to_{\mathrm{u}} f$ para alguma função $f\colon X \to \mathbb{R}$, então

$$\lim_{n\to\infty} \lim_{m\to\infty} f_n(x_m) = \lim_{m\to\infty} \lim_{n\to\infty} f_n(x_m)$$

para qualquer sequência $(x_n)_{n\in\mathbb{N}}$ convergente em X. ■

Exercício 2.107 $({}^{\star}_{\star})$**.** Mostre que se para algum $x \in \mathbb{R} \setminus \{c\}$, a série $f(x) := \sum\limits_{n=0}^{\infty} a_n (x-c)^n$ converge, então existe $r > 0$ tal que $f(z)$ existe para todo $z \in (c-r, c+r)$, e a função $f\colon (c-r, c+r) \to \mathbb{R}$ é suave. ■

Exercício 2.108 (Funções trigonométricas — $({}^{\star}_{\star})$)**.** Para $x \in \mathbb{R}$, considere os números

$$\operatorname{sen}(x) := \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1} \quad \text{e} \quad \cos(x) := \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n}.$$

a) Mostre que os números acima estão bem definidos para qualquer $x \in \mathbb{R}$. Dica: use o Teste de D'Alembert.

b) Mostre que as funções $\operatorname{sen}(x)$ e $\cos(x)$ são suaves. Dica: automaticamente pelo Corolário 2.1.22.

c) Mostre que $\operatorname{sen}'(x) = \cos(x)$ e $\cos'(x) = -\operatorname{sen}(x)$.

d) Mostre que $\lim\limits_{x \to 0} \dfrac{\operatorname{sen}(x)}{x} = 1$. Dica: $\operatorname{sen}'(0) = \cos(0)$.

e) Mostre que $\operatorname{sen}^2(x) + \cos^2(x) = 1$ para todo $x \in \mathbb{R}$. Dica: primeiro, observe que a função $\operatorname{sen}^2(x) + \cos^2(x)$ é uma função constante; depois, descubra a constante†.

f) Mostre que $|\operatorname{sen}(x)| \leq 1$ e $|\cos(x)| \leq 1$ para todo $x \in \mathbb{R}$. Dica: note que valem as desigualdades $(1 - \operatorname{sen}(x))(1 + \operatorname{sen}(x)) \geq 0$ e $(1 - \cos(x))(1 + \cos(x)) \geq 0$.

g) Prove outras fórmulas típicas envolvendo funções trigonométricas, aproveite para definir tangente, cotangente, etc. ■

Exercício 2.109 (Teorema de Dini — $(\substack{\star \\ \star\star})$)**.** Sejam $X \subseteq \mathbb{R}$ um subconjunto compacto, $(f_n)_{n \in \mathbb{N}}$ uma sequência de funções contínuas da forma $X \to \mathbb{R}$ e $f \colon X \to \mathbb{R}$ uma função contínua tal que $f_n \to_{\mathrm{p}} f$.

a) Para $\varepsilon > 0$, mostre que $F_n := \{x \in X : f(x) - f_n(x) \geq \varepsilon\}$ é fechado. Dica: continuidade.

b) Com F_n como no item anterior, mostre que se $(f_n(x))_{n \in \mathbb{N}}$ é crescente para todo $x \in X$, então $F_n \supseteq F_{n+1}$ para todo n.

c) Mostre que $\bigcap_{n \in \mathbb{N}} F_n = \emptyset$. Dica: convergência pontual.

d) Com a hipótese do item (b), mostre que existe $N \in \mathbb{N}$ tal que $F_N = \emptyset$ e, com isso, conclua que $f_n \to_{\mathrm{u}} f$. Dica: compacidade.

e) Mostre que a conclusão anterior permanece válida se $(f_n(x))_{n \in \mathbb{N}}$ é decrescente para todo $x \in X$. ■

Exercícios extras

Exercício 2.110 (Desigualdade AM-GM‡ — $(\substack{\star \\ \star})$)**.** Mostre que

$$\sqrt[n]{x_1 \cdot \ldots \cdot x_n} \leq \frac{x_1 + \ldots + x_n}{n}$$

para quaisquer $n \in \mathbb{N} \setminus \{0\}$ e $x_1, \ldots, x_n \geq 0$. Dica: o caso $n := 2$ é o Exercício 0.134; com base neste caso, prove a desigualdade para todo n da forma 2^k (mas livre-se da raíz antes!, cf. Exercício 2.94); para obter o caso geral, escolha k com $2^k > n$ e, para algum $z > 0$ *adequado*, perceba que

$$x_1 \cdot \ldots \cdot x_n \cdot z^{2^k - n} \leq \left(\frac{x_1 + \ldots + x_n + (2^k - n)z}{2^k} \right)^{2^k};$$

se seu z for realmente adequado, a expressão desejada aparecerá daí. ■

†Em particular, não é um bom momento para esquecer que $0^0 := 1$.

‡*Arithmetic mean — geometric mean.*

Exercício 2.111 $\left(\substack{\star\\ \star}\right)$**.** Nas condições da Definição 2.1.8, mostre que se $f_d \to_{\mathrm{p}} f$ e $f_d \to_{\mathrm{p}} g$, então $f = g$. ■

Exercício 2.112 $\left(\substack{\star\\ \star\star}\right)$**.** Mostre que não existe injeção contínua da forma $\mathbb{R}^2 \to \mathbb{R}$. ■

Exercício 2.113 $\left(\substack{\star\\ \star\star}\right)$**.** Uma $\mathbb{R}$-álgebra é um $\mathbb{R}$-espaço vetorial V com uma multiplicação $\bullet\colon V \times V \to V$ que comuta com o produto por escalares de V e de tal forma que $(V, +, \bullet)$ é um anel, onde $+$ é a adição nativa de V. Dizemos que V é $\mathbb{R}$-álgebra de divisão se, adicionalmente, $(V, +, \bullet)$ for um corpo[†]. Mostre que se V é uma $\mathbb{R}$-álgebra que tem dimensão ímpar como espaço vetorial, então $\dim V = 1$. Dica: se $n > 0$, então $\mathbb{R}^n \setminus \{0\}$ é conexo; além disso, para $a \in V \setminus \{0\}$, a função $v \mapsto av$ é um isomorfismo linear. ■

Exercício 2.114 $\left(\substack{\star\\ \star}\right)$**.** Para $f, g\colon X \to \mathbb{R}$, seja $d_\infty(f, g) := \min\left\{1, \sup\limits_{x \in X} |f(x) - g(x)|\right\}$.

a) Mostre que d_∞ é uma métrica em $\mathbb{R}^X$.

b) Para uma sequência $(f_n)_{n \in \mathbb{N}}$ de funções da forma $X \to \mathbb{R}$, mostre que $f_n \to f$ no espaço métrico $\left(\mathbb{R}^X, d_\infty\right)$ se, e somente se, $f_n \to_{\mathrm{u}} f$.

c) Dizemos que uma sequência de funções $(f_n)_{n \in \mathbb{N}}$ da forma $X \to \mathbb{R}$ é **uniformemente de Cauchy** se para todo $\varepsilon > 0$ existe $N \in \mathbb{N}$ tal que $|f_m(x) - f_n(x)| < \varepsilon$ para quaisquer $x \in X$ e $m, n \geq N$. Mostre que $(f_n)_{n \in \mathbb{N}}$ é uniformemente de Cauchy se, e somente se, $(f_n)_{n \in \mathbb{N}}$ é de Cauchy em $\left(\mathbb{R}^X, d_\infty\right)$.

d) Mostre que $\left(\mathbb{R}^X, d_\infty\right)$ é completo. ■

Exercício 2.115 $\left(\substack{\star\\ \star}\right)$**.** Mostre que o Teorema de Dini permanece válido desde que X seja espaço topológico *enumeravelmente compacto* (cf. Exercício 2.26). ■

2.4 Teorema Fundamental do Cálculo

2.4.0 Essencial

§0 Uma abordagem axiomática para o T.F.C.

Uma das ideias básicas da integração é a seguinte: dada uma função $f\colon [a, b] \to \mathbb{R}$, onde $a, b \in \mathbb{R}$ são tais que $a \leq b$, desejamos associar um número real $\mu(f)$ que costuma ser chamado de *integral definida de* f em $[a, b]$. Quando ocorre $f \geq 0$, a intuição que temos do número $\mu(f)$ diz que ele deveria representar a área da região limitada pelas curvas $x = a$, $x = b$, $y = f(x)$ e $y = 0$.

Um pouco mais rigorosamente, para quaisquer $a, b \in \mathbb{R}$ com $a \leq b$, buscamos definir uma função

$$\int_a^b : \mathcal{I}[a, b] \to \mathbb{R},$$

que a cada $f \in \mathcal{I}[a, b]$ associa o número $\displaystyle\int_a^b f \in \mathbb{R}$, onde $\mathcal{I}[a, b]$ é algum conjunto *razoável* de funções da forma $[a, b] \to \mathbb{R}$. Dadas as intuições geométricas, bem como as exigências da vida, esperamos que a correspondência $\displaystyle\int_a^b$ satisfaça as seguintes condições:

[†]O produto $\bullet$ não precisa ser comutativo, mas não se preocupe com isso.

(i) se $f\colon [a,b] \to \mathbb{R}$ é contínua, então $f \in \mathcal{I}[a,b]$;

(ii) se $f\colon [a,b] \to \mathbb{R}$ é tal que $f \in \mathcal{I}[a,b]$, então f é limitada;

(iii) se $f\colon [a,b] \to \mathbb{R}$ é uma função constante, digamos $f = c$, então $f \in \mathcal{I}[a,b]$ e

$$\int_a^b c = c(b-a); \tag{2.5}$$

(iv) se $f, g \in \mathcal{I}[a,b]$ com $f \leq g$ em $[a,b]$, então

$$\int_a^b f \leq \int_a^b g; \tag{2.6}$$

(v) para $f\colon [a,b] \to \mathbb{R}$ e $c \in [a,b]$, deve-se ter $f \in \mathcal{I}[a,b]$ se, e somente se, $f \in \mathcal{I}[a,c]$ e $f \in \mathcal{I}[c,b]$;

(vi) se $f \in \mathcal{I}[a,b]$ e $c \in [a,b]$, então

$$\int_a^b f = \int_a^c f + \int_c^b f. \tag{2.7}$$

A lista acima apenas elenca as propriedades que gostaríamos que uma definição de integral possuísse. Uma vez dada uma definição *efetiva* de integral, digamos $\mathcal{D}$, o conjunto $\mathcal{I}[a,b]$ passa a ser xingado de *conjunto das funções $\mathcal{D}$-integráveis*. Como a abordagem aqui adotada sugere, existem, *possivelmente*, diversos tipos de integrais que satisfazem os axiomas acima. A depender do contexto, um tipo de integração pode se adequar melhor do que outro.

Desse modo, as exigências (i) e (ii) acima se traduzem em dizer que "toda função contínua é integrável" e "toda função integrável é limitada", respectivamente. A condição (iii) pede que "funções constantes sejam integráveis", enquanto a condição (v) pede que "para qualquer $c \in [a,b]$, f é integrável em $[a,b]$ se, e somente se, f é integrável[†] em $[a,c]$ e $[c,b]$".

Se você não ignorou as subseções extras sobre *integrais de Riemann* (Subseções 1.2.1 §0 e 2.1.1 §0), já deve ter percebido que ela satisfaz quase todas as condições impostas acima: a condição (i) decorre do Teorema 2.1.6, (ii) segue da Proposição 1.2.30, enquanto (iii) e (iv) foram abordadas nos Exemplos 1.2.27 e 1.2.28, respectivamente — resta apenas verificar (v) e (vi). No entanto, antes de fazer isso, será dramaticamente melhor provar o

Teorema 2.4.0 (Fundamental do Cálculo)**.** *Suponha que para cada $a, b \in \mathbb{R}$ com $a \leq b$ existam $\mathcal{I}[a,b]$ e $\int_a^b \colon \mathcal{I}[a,b] \to \mathbb{R}$ satisfazendo as condições anteriores. Para $f \in \mathcal{I}[a,b]$ e $x \in [a,b]$, defina*

$$F(x) := \int_a^x f.$$

(I) *A função $F\colon [a,b] \to \mathbb{R}$ definida pela correspondência acima é contínua em cada $c \in [a,b]$.*

(II) *Se f é contínua em $c \in [a,b]$, então F é diferenciável em c, e $F'(c) = f(c)$.*

[†]Embora, a rigor, devêssemos escrever "... $f|_{[a,c]}$ e $f|_{[c,b]}$ são integráveis". Mas a vida é curta.

(III) *Se f é contínua e $G\colon [a,b] \to \mathbb{R}$ é função contínua satisfazendo $G'(x) = f(x)$ para cada $x \in (a,b)$, então*
$$\int_a^b f = G(b) - G(a).$$

Demonstração. Provaremos os itens em ordem. Primeiro, como o axioma (ii) diz que toda função em $\mathcal{I}[a,b]$ é limitada, existe $M \geq 0$ tal que $|f| \leq M$. Se ocorrer $M = 0$, acabou (certo?)*. Por isso, vamos supor $M > 0$. Para $c \in [a,b]$, queremos mostrar que F é contínua em c. Para tanto, fixado $\varepsilon > 0$, devemos encontrar $\delta > 0$ tal que para $x \in [a,b]$, $|x-c| < \delta$ acarrete $|F(x) - F(c)| < \varepsilon$. Para isso, vamos manipular a *expressão* $F(x) - F(c)$ marotamente.

Graças ao axioma (vi), para $x \geq c$ podemos afirmar que $F(x) - F(c) = \displaystyle\int_c^x f$, pois
$$F(x) - F(c) = \int_a^x f - \int_a^c f = \int_a^c f + \int_c^x f - \int_a^c f = \int_c^x f.$$

Analogamente, se $x \leq c$, então $F(x) - F(c) = -\displaystyle\int_x^c f$ (percebeu?)*.

Agora, se $a \leq \alpha \leq \beta \leq b$, então $-M \leq f \leq M$ em $[\alpha, \beta]$. Logo, os axiomas (iii) e (iv) garantem
$$-M(\beta - \alpha) = \int_\alpha^\beta -M \leq \int_\alpha^\beta f \leq \int_\alpha^\beta M = M(\beta - \alpha),$$
donde segue que $\left|\displaystyle\int_\alpha^\beta f\right| \leq M(\beta - \alpha)$.

Consequentemente, deve ocorrer
$$|F(x) - F(c)| \leq M|x - c|,$$
(por quê?!)$^\star_\star$, donde é fácil concluir que F é (uniformemente) contínua.

Provemos o item (II). Supondo a continuidade de f em $c \in [a,b]$, desejamos mostrar que o limite
$$F'(c) := \lim_{x \to c} \frac{F(x) - F(c)}{x - c}$$
existe e, mais ainda, ocorre $F'(c) = f(c)$. Ora, pela continuidade de f em c, dado $\varepsilon > 0$, existe $\delta > 0$ para o qual $|x - c| < \delta$ com $x \in [a,b]$ implica em
$$f(c) - \varepsilon < f(x) < f(c) + \varepsilon.$$

Logo, para $x > c$ com $|x - c| < \delta$, os axiomas (iii) e (iv) nos dão
$$(f(c) - \varepsilon)(x - c) = \int_c^x (f(c) - \varepsilon) \leq \int_c^x f \leq \int_c^x (f(c) + \varepsilon) = (f(c) + \varepsilon)(x - c),$$
donde segue que
$$f(c) - \varepsilon \leq \frac{1}{x - c} \int_c^x f \leq f(c) + \varepsilon.$$

Como, neste caso, $F(x) - F(c) = \displaystyle\int_c^x f$, a desigualdade acima se traduz em
$$\left|\frac{F(x) - F(c)}{x - c} - f(c)\right| < \varepsilon,$$

como queríamos. O caso “$x < c$” será problema seu ($\substack{\star\\\star}$).

Finalmente, provemos (III). Pela parte (II), sabemos que F é *uma antiderivada* de f. Logo, pelo T.V.M. (cf. Corolário 2.0.14), existe uma constante $C \in \mathbb{R}$ tal que $G(x) = F(x) + C$ para todo $x \in [a, b]$. Consequentemente,

$$G(b) - G(a) = (F(b) + C) - (F(a) + C) = F(b) - F(a) = \int_a^b f - \int_a^a f = \int_a^b f,$$

pois $\int_a^a f = 0$ (certo?)*, como desejado. □

Observação 2.4.1. A combinação do Teorema 2.1.6 com os itens (I) e (II) costuma ser chamada de *Primeiro* ***T****eorema* ***F****undamental do* ***C****álculo* (de agora em diante abreviado T.F.C.): explicitamente, se $f : [a, b] \to \mathbb{R}$ é contínua, então f é Riemann-integrável e sua *integral indefinida* F é uma antiderivada de f, i.e., $F' = f$. O *Segundo T.F.C.*, por sua vez, costuma ser enunciado como o item (III), ou como uma generalização que combina os itens anteriores. Discutiremos mais algumas sutilezas na Subseção 2.4.1 §2. △

Nossa argumentação[†] mostra que, uma vez definida uma *noção* de integração satisfazendo os axiomas (i), …, (vi), teremos o T.F.C. para a noção de integral em questão. Em particular, independentemente da definição adotada, o item (III) do T.F.C. nos diz que

$$\int_a^b f = G(b) - G(a),$$

qualquer que seja a antiderivada G de f, enquanto o item (II) garante que pelo menos uma antiderivada existe. Assim, a menos de mostrar a *existência* de uma *integração*, já sabemos qual o único valor possível para a integral de uma função contínua $f : [a, b] \to \mathbb{R}$.

Exercício 2.116 ($\star$)**.** Recalcule a integral do Exemplo 1.3.28. Dica: derive $f(x) := \frac{x^3}{3}$.■

§1 A integral de Riemann revisitada

Tudo o que resta fazer é mostrar que a integral de Riemann *funciona*. Mais precisamente, chamando por $\mathcal{R}[a, b]$ o conjunto das funções da forma $[a, b] \to \mathbb{R}$ para as quais o limite

$$\int_a^b f(t)\,\mathrm{d}t := \lim_{(\mathcal{P},T)\in \mathrm{Par}_{\mathcal{R}}\,[a,b]} \sum_{(\mathcal{P},T)} f$$

existe em $\mathbb{R}$ (cf. Definição 1.2.26), mostraremos que

$$\int_a^b (\cdot)\,\mathrm{d}t : \mathcal{R}[a, b] \to \mathbb{R}$$

satisfaz os axiomas (i), …, (vi).

Preciosismos. A princípio, a Definição 1.2.26 trata apenas do caso $a < b$, o que pode incomodar pessoas preciosistas, já que os axiomas de integral desta seção pedem $a \le b$. No entanto, se permitirmos partições $(p_0, \ldots, p_n)$ tais que $p_0 \le p_1 \le \ldots \le p_n$, resulta

$$\sum_{(\mathcal{P},T)} f = 0$$

[†]Adaptada do texto de Pete L. Clark [8], que por sua vez é inspirado pela magnífica obra de Lang [22].

para qualquer partição de Riemann $(\mathcal{P}, T)$ de $[a, a]$, acarretando

$$\int_a^a f(t)\,\mathrm{d}t = 0,$$

justamente o que se espera de uma noção abstrata de integral com as propriedades estipuladas. Por sua vez, permitir tais partições nos casos em que $a < b$ em nada altera os resultados já obtidos, uma vez que os pontos consecutivos iguais numa partição teriam suas parcelas correspondentes anuladas na soma de Riemann. *Moral da história*: você não precisa se preocupar com os casos em que $a = b$. Para $a \geq b$, confira o Exercício 2.141.

Observação 2.4.2 (Por que raios esse "dt"?)**.** Antes de prosseguir, é interessante discutir a presença misteriosa do termo "dt" para denotar a integral de uma função. A princípio, as razões são históricas. Intuitivamente, a integral foi pensada como uma soma infinita das áreas de retângulos (cf. Figura 1.6) de altura $f(t)$ e cuja base em torno de t é *muito pequena* (um *infinitesimal* dt)[†]. No entanto, nada nas condições anteriores ou na própria definição da integral de Riemann faz alusão explícita a infinitesimais. Então, por que manter? Resposta: praticidade. Como assim?

Embora seja inútil nas discussões teóricas, na prática, pode ser útil que algo na notação indique o que é variável e o que é constante. Por exemplo, ao considerar a função $f\colon [0,1] \times [0,1] \to \mathbb{R}$ dada por $f(x,y) := x^2 + y$, a notação torna quase automático interpretar os significados de

$$\int_0^1 f(x,y)\,\mathrm{d}x \quad \text{e} \quad \int_0^1 f(x,y)\,\mathrm{d}y.$$

Há também motivos mnemônicos. Considere o caso em que $g\colon [a,b] \to \mathbb{R}$ é diferenciável com $g'\colon [a,b] \to \mathbb{R}$ contínua e $f\colon I \to \mathbb{R}$ *contínua*, onde $\operatorname{im}(g) \subseteq I$. Sob tais condições, vale que

$$\int_a^b f'(g(t)) \cdot g'(t)\,\mathrm{d}t = \int_{g(a)}^{g(b)} f(u)\,\mathrm{d}u.$$

De fato, sob tais condições a função $h(t) := f(g(t))g'(t)$ é contínua em $[a,b]$ e, portanto, Riemann-integrável, donde segue que

$$\int_a^b h(t)\,\mathrm{d}t = H(b) - H(a)$$

para qualquer $H\colon [a,b] \to \mathbb{R}$ satisfazendo $H'(x) = h(x)$ para todo $x \in (a,b)$. Ora, pela Regra da Cadeia, $H := F \circ g$ serve, onde F é alguma antiderivada de f (certo?!)*. Porém, para se lembrar disso, costuma-se pensar na "substituição" $u = g(t)$ que *resulta* em

$$\frac{\mathrm{d}u}{\mathrm{d}t} = g'(t) \text{``}\Rightarrow\text{''}\ \mathrm{d}u = g'(t)\,\mathrm{d}t.$$

Para formalizar esta última trapaça, é preciso apelar para *formas diferenciais* ou *Análise não standard*, dois temas que não serão abordados aqui. Em todo caso, fica justificada a obsessão por carregar o "dt" por todos os lados. △

[†]Inclusive, o símbolo da integral, $\int$, é um "S", de ~~esperança~~ soma. Já havia notado isso?

Exercício 2.117 (?!)**.** Perceba que (i), …, (iv) já foram. Dica: revise as seções que tratam desses assuntos ao longo do texto. ■

Proposição 2.4.3. *A "ida" na condição* (v) *é satisfeita.*

Demonstração. Supondo que $f\colon [a,b] \to \mathbb{R}$ é Riemann-integrável, mostraremos que a restrição de f ao intervalo $[a,c]$ é integrável. Por simplicidade, vamos escrever g em vez de $f|_{[a,c]}$. Como no Teorema 2.1.6 (cf. Corolário 1.3.16), mostraremos que a rede

$$\begin{aligned} \sum g\colon \operatorname{Par}_{\mathcal{R}}[a,c] &\to \mathbb{R} \\ (\mathcal{P},T) &\mapsto \sum_{(\mathcal{P},T)} g \end{aligned}$$

é de Cauchy. Como $\sum f\colon \operatorname{Par}_{\mathcal{R}}[a,b] \to \mathbb{R}$ é de Cauchy, para $\varepsilon > 0$ existe $\delta > 0$ tal que

$$\left| \sum_{(\mathcal{Q},S)} f - \sum_{(\mathcal{Q}',S')} f \right| < \varepsilon$$

sempre que $\|\mathcal{Q}\|, \|\mathcal{Q}'\| < \delta$. A afirmação natural é a de que o mesmo δ funciona para $\sum g$.

Com efeito, se $(\mathcal{P},T), (\mathcal{P}',T') \in \operatorname{Par}_{\mathcal{R}}[a,c]$ são partições de Riemann satisfazendo $\|\mathcal{P}\|, \|\mathcal{P}'\| < \delta$, então para $N \in \mathbb{N}$ satisfazendo $N > \frac{b-c}{\delta}$, podemos fazer $\gamma_j := c + j\frac{b-c}{N}$ para cada $j \in \{0, \dots, N\}$ e definir as partições em $\operatorname{Par}_{\mathcal{R}}[a,b]$,

$$\mathcal{Q} := \Bigg(\underbrace{p_0, \dots, p_m = c = \gamma_0}_{\mathcal{P}}, \gamma_1, \dots, \gamma_N \Bigg) \ \text{ e } \ \mathcal{Q}' := \Bigg(\underbrace{p'_0, \dots, p'_n = c = \gamma_0}_{\mathcal{P}'}, \gamma_1, \dots, \gamma_N \Bigg)$$

com as *tags* $S := (t_1, \dots, t_m, \gamma_1, \dots, \gamma_N)$ e $S' := (t'_1, \dots, t'_n, \gamma_1, \dots, \gamma_N)$, respectivamente, que satisfazem $\|\mathcal{Q}\|, \|\mathcal{Q}'\| < \delta$. Daí,

$$\begin{aligned} \left| \sum_{(\mathcal{P},T)} g - \sum_{(\mathcal{P}',T')} g \right| &= \left| \sum_{(\mathcal{P},T)} g + \sum_{i=1}^{N} f(\gamma_i)(\gamma_i - \gamma_{i-1}) - \sum_{i=1}^{N} f(\gamma_i)(\gamma_i - \gamma_{i-1}) - \sum_{(\mathcal{P}',T')} g \right| \\ &= \left| \sum_{(\mathcal{Q},S)} f - \sum_{(\mathcal{Q},S')} f \right| < \varepsilon, \end{aligned}$$

como desejado (certo?!)*. Analogamente (leia-se: faça você! $(\substack{\star\\\star})$), mostra-se que $f|_{[c,b]}$ também é Riemann-integrável. □

A recíproca é um pouco mais delicada: a ideia é mostrar que se $g := f|_{[a,c]}$ e $h := f|_{[c,b]}$ são Riemann-integráveis, então $f\colon [a,b] \to \mathbb{R}$ é Riemann-integrável e vale

$$\int_a^b f(t)\,\mathrm{d}t = \underbrace{\int_a^c f(t)\,\mathrm{d}t}_{(A)} + \underbrace{\int_c^b f(t)\,\mathrm{d}t}_{(B)},$$

o que, em particular, mostra a validade da condição (vi).

Num primeiro momento, parece fácil. Por hipótese, para $\varepsilon > 0$ fixado, existem $\delta_0, \delta_1 > 0$ tais que

$$\left|\sum_{(\mathcal{P},T)} g - A\right| < \varepsilon \quad \text{e} \quad \left|\sum_{(\mathcal{P}',T')} h - B\right| < \varepsilon \tag{2.8}$$

sempre que $(\mathcal{P},T) \in \mathrm{Par}_{\mathcal{R}}[a,c]$ e $(\mathcal{P}',T') \in \mathrm{Par}_{\mathcal{R}}[c,b]$ satisfazem $\|\mathcal{P}\| < \delta_0$ e $\|\mathcal{P}'\| < \delta_1$. Instintivamente, a reação natural a isso é escolher $\delta := \min\{\delta_0, \delta_1\}$ na esperança de que se $(\mathcal{Q},T)$ for partição de Riemann de $[a,b]$ satisfazendo $\|\mathcal{Q}\| < \delta$, então

$$\left|\sum_{(\mathcal{Q},T)} f - (A+B)\right| \leq \left|\sum_{(\mathcal{Q}',T')} f - A\right| + \left|\sum_{(\mathcal{Q}'',T'')} f - B\right|, \tag{2.9}$$

onde $(\mathcal{Q}',T')$ é a *restrição* da partição $(\mathcal{Q},T)$ ao intervalo $[a,c]$, enquanto $(\mathcal{Q}'',T'')$ é a *restrição* da partição $(\mathcal{Q},T)$ ao intervalo $[c,b]$.

Porém, há um problema no raciocínio: com $\mathcal{Q} := (q_0,\ldots,q_n)$, a mera desigualdade $\|\mathcal{Q}\| < \delta$ não garante $c = q_i$ para algum i, e assim as restrições $(\mathcal{Q}',T')$ e $(\mathcal{Q}'',T'')$ não são tão *canônicas* como poderíamos pensar a princípio.

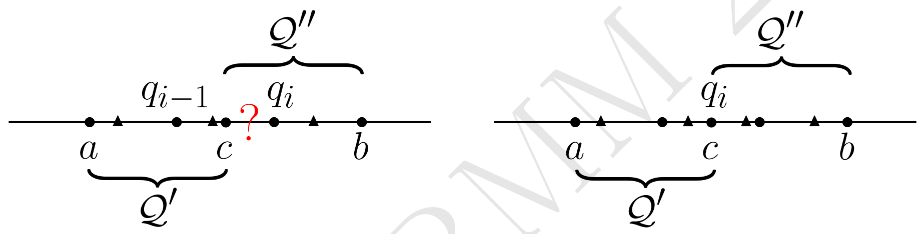


Figura 2.2: Do lado esquerdo, em que c não é um dos pontos da partição de $[a,b]$, representados por bolinhas, precisamos escolher um novo ponto para ser *tag* (os triângulos) do subíntervalo $[c,q_i]$ da partição $\mathcal{Q}''$. Observe que isto não acontece do lado direito.

Curiosamente, este problema desapareceria completamente se *pudéssemos* trocar a pré-ordem $\preceq$ de $\mathrm{Par}_{\mathcal{R}}[a,b]$ pela ordem $\sqsubseteq$ (confira a Observação 1.2.23 para se recordar da definição): com ela, em vez de $\delta_0 > 0$ e $\delta_1 > 0$, garantimos partições $\mathcal{D}_0 := (d_0,\ldots,d_m)$ de $[a,c]$ e $\mathcal{D}_1 := (d'_0,\ldots,d'_n)$ de $[c,b]$ tais que

$$\left|\sum_{(\mathcal{P},T)} f - A\right| < \varepsilon \quad \text{e} \quad \left|\sum_{(\mathcal{P}',T')} f - B\right| < \varepsilon$$

sempre que $\mathcal{D}_0 \sqsubseteq \mathcal{P}$ e $\mathcal{D}_1 \sqsubseteq \mathcal{P}'$, de modo que $\mathcal{D} := (d_0,\ldots,d_m = d'_0 = c,\ldots,d'_n)$ é uma partição de $[a,b]$ em que o truque esboçado em (2.9) funciona.

De fato, se uma partição de Riemann $(\mathcal{Q},T) \in \mathrm{Par}_{\mathcal{R}}[a,b]$ satisfaz $\mathcal{D} \sqsubseteq \mathcal{Q} := (q_0,\ldots,q_k)$, então $c = q_j$ para algum j e daí

$$\begin{aligned}\left|\sum_{(\mathcal{Q},T)} f - (A+B)\right| &= \left|\sum_{i=1}^{j} f(t_i)(q_i - q_{i-1}) - A + \sum_{i=j+1}^{k} f(t_i)(q_i - q_{i-1}) - B\right| \\ &\leq \left|\sum_{(\mathcal{Q}',T')} f - A\right| + \left|\sum_{(\mathcal{Q}'',T'')} f - B\right|,\end{aligned}$$

onde $\mathcal{Q}' := (q_0,\ldots,q_j)$, $T' := (t_1,\ldots,t_j)$, $\mathcal{Q}'' := (q_j,\ldots,q_k)$ e $T'' := (t_{j+1},\ldots,t_k)$, com $\mathcal{D}_0 \sqsubseteq \mathcal{Q}'$ e $\mathcal{D}_1 \sqsubseteq \mathcal{Q}''$. Quer uma boa notícia?

Lema 2.4.4. *Para uma função $f\colon [a,b] \to \mathbb{R}$ e um número $L \in \mathbb{R}$, são equivalentes:*

(i) $\displaystyle\int_a^b f(t)\,\mathrm{d}t = L$; (ii) $\displaystyle\lim_{(\mathcal{P},T)\in(\mathrm{Par}_{\mathcal{R}}\,[a,b],\sqsubseteq)} \sum_{(\mathcal{P},T)} f = L$.

Observação 2.4.5. Como a demonstração que enfrentaremos é bastante árdua, convém perguntar: por que não assumir simplesmente a condição (ii) como "a definição de integral" desde o começo? RESPOSTA: pois as duas são úteis.

Note que somente com base na condição (ii), por exemplo, não poderíamos usar o argumento de cofinalidade do Exemplo 1.3.28 para calcular $\int_0^1 x^2\,\mathrm{d}x$, já que a coleção de partições $\{\mathcal{P}_n : n \in \mathbb{N}\}$ lá definida não é cofinal em $\mathrm{Par}_{\mathcal{R}}\,[0,1]$ com respeito à ordem $\sqsubseteq$ (cf. Exercício 1.60). △

Demonstração. A implicação (i) ⇒ (ii) segue pois $(\mathcal{Q},R) \preceq (S,T)$ sempre que se tem $(\mathcal{Q},R) \sqsubseteq (\mathcal{S},T)$ (pense a respeito!)$^{\star}_{\star}$. O problema é a recíproca. Assumindo (ii), para $\varepsilon,\gamma > 0$ quaisquer, existe uma partição de Riemann $\left(\widetilde{\mathcal{P}},T\right)$ para $[a,b]$ tal que

$$\Delta := \left|\sum_{(\mathcal{Q},S)} f - L\right| < \varepsilon$$

sempre que $\left(\widetilde{\mathcal{P}},T\right) \sqsubseteq (\mathcal{Q},S)$ e, a partir disso, buscamos $\delta > 0$ capaz de assegurar $\Delta < \gamma$ sempre que $\|\mathcal{Q}\| < \delta$ (cf. Teorema 1.2.24). É chegado o momento de apelar para Darboux[†].

⌜**Afirmação (Truque de Darboux).** *Para uma partição $\mathcal{Q} := (q_0,\ldots,q_n)$ de $[a,b]$ e uma função limitada $f\colon [a,b] \to \mathbb{R}$, considere os números reais*[‡]

$$m_i := \inf\{f(x) : x \in [q_{i-1},q_i]\} \quad \text{e} \quad M_i := \sup\{f(x) : x \in [q_{i-1},q_i]\}$$

para cada $i \in \{1,\ldots,n\}$, bem como as somas ~~de Riemann~~

$$L(f,\mathcal{Q}) := \sum_{i=1}^{n} m_i(q_i - q_{i-1}) \quad \text{e} \quad U(f,\mathcal{Q}) := \sum_{i=1}^{n} M_i(q_i - q_{i-1}).$$

Nestas condições, se $\mathcal{Q}'$ for uma partição obtida a partir de $\mathcal{Q}$ por meio do acréscimo de um único ponto, então

$$L(f,\mathcal{Q}') \leq L(f,\mathcal{Q}) + 2\|f\|\|\mathcal{Q}\| \quad \text{e} \quad U(f,\mathcal{Q}') \geq U(f,\mathcal{Q}) - 2\|f\|\|\mathcal{Q}\|.$$

Demonstração. Seja z o ponto adicional. Como $\mathcal{Q}$ é partição, existe um único número natural $k \in \{1,\ldots,n\}$ tal que $q_{k-1} < z < q_k$. Agora, com $m := \inf\{f(x) : x \in [q_{k-1},z]\}$ e $m' := \inf\{f(x) : x \in [z,q_k]\}$, temos

$$\begin{aligned} L(f,\mathcal{Q}') - L(f,\mathcal{Q}) &= m(z - q_{k-1}) + m'(q_k - z) - m_k(q_k - q_{k-1}) \\ &\leq \|f\|_\infty(z - q_{k-1}) + \|f\|_\infty(q_k - z) + \|f\|_\infty(q_k - q_{k-1}) \\ &= 2\|f\|_\infty(q_k - q_{k-1}) \leq 2\|f\|_\infty\|\mathcal{Q}\|, \end{aligned}$$

onde a primeira igualdade se deve à suposição acerca de $\mathcal{Q}'$ (percebeu?)*. A segunda desigualdade se prova de forma análoga (isto é: prove!)$^{\star}_{\star}$. ⌟

[†]Esta é uma escolha mais didática do que técnica, haja vista que muitos textos utilizam *somas de Darboux*, o que torna razoável apresentá-las para você em algum momento. Com isso dito, vale o adendo de que tais somas são completamente evitáveis, mesmo para a demonstração do critério de Riemann-Darboux (cf. Teorema 2.4.9): para mais detalhes, confira o artigo de Gordon [13].

[‡]Que existem por f ser limitada (cf. Exercício 1.44).

Com as notações acima, observe que se $\mathcal{Q}''$ for uma partição de $[a, b]$ obtida a partir de $\mathcal{Q}$ por meio do acréscimo de *dois* pontos, então

$$L(f, \mathcal{Q}'') \leq L(f, \mathcal{Q}') + 2\|f\|_\infty \|\mathcal{Q}'\| \leq L(f, \mathcal{Q}) + 4\|f\|_\infty \|\mathcal{Q}\| \quad \text{e}$$

$$U(f, \mathcal{Q}) \leq U(f, \mathcal{Q}') + 2\|f\|_\infty \|\mathcal{Q}'\| \leq U(f, \mathcal{Q}'') + 4\|f\|_\infty \|\mathcal{Q}\|,$$

onde $\mathcal{Q}'$ é uma partição *intermediária* entre $\mathcal{Q}$ e $\mathcal{Q}''$. Procedendo recursivamente, se $\mathcal{P} := (p_0, \ldots, p_m)$ é partição de $[a, b]$, resulta

$$L(f, \mathcal{Q} \sqcup \mathcal{P}) \leq L(f, \mathcal{Q}) + 2m\|f\|_\infty \|\mathcal{Q}\| \quad \text{e} \quad U(f, \mathcal{Q}) \leq U(f, \mathcal{Q} \sqcup \mathcal{P}) + 2m\|f\|_\infty \|\mathcal{Q}\|, \tag{2.10}$$

pois $\mathcal{Q} \sqcup \mathcal{P}$ se obtém a partir de $\mathcal{Q}$ com o acréscimo de $k < m$ pontos.

Por outro lado, as definições de $L(f, \mathcal{Q})$ e $U(f, \mathcal{Q})$ claramente asseguram

$$L(f, \mathcal{Q}) \leq \sum_{(\mathcal{Q}, S)} f \leq U(f, \mathcal{Q})$$

para qualquer *tag* S de $\mathcal{Q}$ (por quê?!)*, de modo que ao trocar $\mathcal{P}$ por $\widetilde{\mathcal{P}} := (\widetilde{p}_0, \ldots, \widetilde{p}_m)$ em (2.10), obtemos

$$L(f, \mathcal{Q} \sqcup \widetilde{\mathcal{P}}) - 2m\|f\|_\infty \|\mathcal{Q}\| \leq \sum_{(\mathcal{Q}, S)} f \leq U(f, \mathcal{Q} \sqcup \widetilde{\mathcal{P}}) + 2m\|f\|_\infty \|\mathcal{Q}\|. \tag{2.11}$$

Ainda não acabou, mas estamos quase lá: em virtude do pequeno exercício que você fará após esta longa demonstração (cf. Exercício 2.118), a desigualdade anterior acarreta

$$L(f, \widetilde{\mathcal{P}}) - 2m\|f\|_\infty \|\mathcal{Q}\| \leq \sum_{(\mathcal{Q}, S)} f \leq U(f, \widetilde{\mathcal{P}}) + 2m\|f\|_\infty \|\mathcal{Q}\|.$$

O último *Katzensprung* é utilizar as definições de

$$L(f, \widetilde{\mathcal{P}}) := \sum_{i=1}^{m} \alpha_i (\widetilde{p}_i - \widetilde{p}_{i-1}) \quad \text{e} \quad U(f, \widetilde{\mathcal{P}}) := \sum_{i=1}^{m} \beta_i (\widetilde{p}_i - \widetilde{p}_{i-1})$$

para relacioná-las com somas de Riemann legítimas em que a hipótese acerca de $\widetilde{\mathcal{P}}$ possa ser, finalmente, utilizada. Como definimos $\alpha_i := \inf\{f(x) : x \in [\widetilde{p}_{i-1}, \widetilde{p}_i]\}$ e $\beta_i := \sup\{f(x) : x \in [\widetilde{p}_{i-1}, \widetilde{p}_i]\}$ para cada $i \in \{1, \ldots, m\}$, as propriedades de supremo e ínfimo nos dão pontos $l_i, u_i \in [\widetilde{p}_{i-1}, \widetilde{p}_i]$ satisfazendo

$$\alpha_i + \frac{\varepsilon}{b - a} > f(l_i) \quad \text{e} \quad \beta_i - \frac{\varepsilon}{b - a} < f(u_i),$$

donde segue que

$$\alpha_i (\widetilde{p}_i - \widetilde{p}_{i-1}) > f(l_i)(\widetilde{p}_i - \widetilde{p}_{i-1}) - \frac{\varepsilon}{b - a} (\widetilde{p}_i - \widetilde{p}_{i-1})$$

e

$$\beta_i (\widetilde{p}_i - \widetilde{p}_{i-1}) < f(u_i)(\widetilde{p}_i - \widetilde{p}_{i-1}) + \frac{\varepsilon}{b - a} (\widetilde{p}_i - \widetilde{p}_{i-1}).$$

Ao somar todas as parcelas correspondentes conforme i varia de 1 a m, obtemos as desigualdades a seguir[†], onde $l := (l_1, \ldots, l_m)$ e $u := (u_1, \ldots, u_m)$ são *tags* legítimas da partição $\widetilde{\mathcal{P}}$:

$$L(f, \widetilde{\mathcal{P}}) > \sum_{(\widetilde{\mathcal{P}}, l)} f - \frac{\varepsilon}{b-a} \underbrace{\left(\sum_{i=1}^{m} \widetilde{p}_i - \widetilde{p}_{i-1} \right)}_{b-a} = \sum_{(\widetilde{\mathcal{P}}, l)} f - \varepsilon$$

e, analogamente,

$$U(f, \widetilde{\mathcal{P}}) < \sum_{(\widetilde{\mathcal{P}}, u)} f + \varepsilon.$$

Logo, de (2.11) obtemos

$$-\varepsilon - 2m\|f\|_\infty \|\mathcal{Q}\| + \sum_{(\widetilde{\mathcal{P}}, l)} f \leq \sum_{(\mathcal{Q}, \mathcal{S})} f \leq \sum_{(\widetilde{\mathcal{P}}, u)} f + \varepsilon + 2m\|f\|_\infty \|\mathcal{Q}\|. \tag{2.12}$$

Finalmente, pelo modo como a partição $\widetilde{\mathcal{P}}$ foi tomada, temos

$$-\varepsilon - \sum_{(\widetilde{\mathcal{P}}, t)} f < -L < -\sum_{(\widetilde{\mathcal{P}}, t)} f + \varepsilon \tag{2.13}$$

para qualquer *tag* t em $\widetilde{\mathcal{P}}$ (certo?)*. Em particular, ao fazer $t := l$ na primeira desigualdade e $t := u$ na segunda desigualdade, a "soma" de (2.12) e (2.13) resulta em

$$-2\varepsilon - 2m\|f\|_\infty \|\mathcal{Q}\| < \sum_{(\mathcal{Q}, \mathcal{S})} f - L < 2\varepsilon + 2m\|f\|_\infty \|\mathcal{Q}\|,$$

i.e.,

$$\left| \sum_{(\mathcal{Q}, \mathcal{S})} f - L \right| < 2(\varepsilon + m\|f\|_\infty \|\mathcal{Q}\|).$$

Portanto, tomando $\varepsilon := \frac{\gamma}{4}$, garante-se $\Delta < \gamma$ sempre que $\|\mathcal{Q}\| < \frac{\gamma}{4m\|f\|_\infty}$, onde m é o número de pontos na partição $\widetilde{\mathcal{P}}$. Suspiro... □

Exercício 2.118 ($^\star_\star$)**.** Sejam $\mathcal{A}$ e $\mathcal{B}$ partições de $[a, b]$, com $\mathcal{A} \sqsubseteq \mathcal{B}$, bem como uma função limitada $f\colon [a, b] \to \mathbb{R}$.

a) Mostre que $L(f, \mathcal{A}) \leq L(f, \mathcal{B})$. Dica: faça por indução no número de pontos que $\mathcal{B}$ adiciona à partição $\mathcal{A}$.

b) Mostre que $U(f, \mathcal{B}) \leq U(f, \mathcal{A})$. Dica: use a dica anterior. ■

Corolário 2.4.6. *A "volta" da condição* (v)*, bem como a condição* (vi)*, são satisfeitas pela integral de Riemann.*

Demonstração. Pelo que já se discutiu, segue do lema anterior. Se ainda duvida, reflita. □

Se você preferir a abordagem para integrais de Riemann que **não** usa somas de Riemann, mas sim *somas de Darboux*, a Observação 2.4.11 na próxima subseção mostrará a equivalência entre as duas formulações. No entanto, as propriedades da integral de Riemann não serão redemonstradas no texto por meio de supremos e ínfimos: os cânones já fazem isso muito bem.

[†]Que você checará com todo o cuidado, tenho certeza ($^\star_\star$).

2.4.1 Extras (Importante)

§0 Os critérios de Darboux e Lebesgue para integrais de Riemann

Embora o critério de Cauchy tenha se mostrado bastante eficaz para caracterizar a *Riemann-integrabilidade* de funções da forma $[a,b] \to \mathbb{R}$, as particularidades "combinatórias" das partições de Riemann permitem obter critérios ainda mais práticos. O primeiro deles está relacionado com as *somas de Darboux*, implicitamente utilizadas na seção anterior.

Definição 2.4.7. Para uma função limitada $f\colon [a,b] \to \mathbb{R}$ e uma partição $\mathcal{P} := (p_0, \ldots, p_n)$ de $[a,b]$, os números

$$L(f,\mathcal{P}) := \sum_{i=1}^{n} m_i(p_i - p_{i-1}) \quad \text{e} \quad U(f,\mathcal{P}) := \sum_{i=1}^{n} M_i(p_i - p_{i-1})$$

definidos no Truque de Darboux (cf. Teorema 2.4.3), são chamadas de **somas de Darboux** de f com respeito à partição $\mathcal{P}$: $L(f,\mathcal{P})$ é a soma **inferior**[†], enquanto $U(f,\mathcal{P})$ é a soma **superior**[‡]. ¶

Observação 2.4.8. Explicitamente, somas de Darboux são *quase* somas de Riemann com respeito às "*tags*" $L_{\mathcal{P}} := (m_1, \ldots, m_n)$ e $U_{\mathcal{P}} := (M_1, \ldots, M_n)$, onde para cada i se consideram $m_i := \inf\{f(x) : x \in [p_{i-1}, p_i]\}$ e $M_i := \sup\{f(x) : x \in [p_{i-1}, p_i]\}$. Se f fosse contínua, tais ínfimos e supremos pertenceriam às imagens e, portanto, poderiam ser *realizados* efetivamente por pontos da partição. Assim, assumir f limitada permite remediar a possível falta de um representante. △

Teorema 2.4.9 (Critério de Riemann-Darboux)**.** *Uma função limitada $f\colon [a,b] \to \mathbb{R}$ é Riemann-integrável se, e somente se, para todo $\varepsilon > 0$ existe uma partição $\mathcal{Q} := (q_0, \ldots, q_n)$ de $[a,b]$ tal que*

$$\left| \sum_{(\mathcal{Q},T)} f - \sum_{(\mathcal{Q},T')} f \right| < \varepsilon$$

para quaisquer tags T e T' de $\mathcal{Q}$.

Demonstração. A "ida" é evidente. Para a recíproca, pelo que se discutiu na Observação 2.4.4, basta mostrar que a rede $\left(\sum_{(\mathcal{P},T)} f\right)_{(\mathcal{P},T)\in \mathrm{Par}_{\mathcal{R}}[a,b]}$ é de Cauchy, mas com $\mathrm{Par}_{\mathcal{R}}[a,b]$ dirigido pela pré-ordem $\sqsubseteq$ (por quê?!)*. Por sorte, o trabalho pesado já está (quase) todo feito: tomando uma partição $\mathcal{Q}$ como no enunciado com respeito a $\varepsilon' > 0$, veremos que $\left|\sum_{(\mathcal{A}_0,T_0)} f - \sum_{(\mathcal{A}_1,T_1)} f\right|$ é *controlável* desde que $\mathcal{Q} \sqsubseteq \mathcal{A}_0, \mathcal{A}_1$, quaisquer que sejam as *tags* T_j de $\mathcal{A}_j$.

Pelo Exercício 2.118, temos

$$L(f,\mathcal{Q}) \leq L(f,\mathcal{A}_j) \leq \sum_{(\mathcal{A}_j,T_j)} f \leq U(f,\mathcal{A}_j) \leq U(f,\mathcal{Q})$$

para $j \in \{0,1\}$, donde resulta (por quê?)$^\star_\star$

$$L(f,\mathcal{Q}) - U(f,\mathcal{Q}) \leq \sum_{(\mathcal{A}_0,T_0)} f - \sum_{(\mathcal{A}_1,T_1)} f \leq U(f,\mathcal{Q}) - L(f,\mathcal{Q}),$$

[†]Em inglês: *lower*, por isso o "L".
[‡]Em inglês: *upper*, por isso o "U".

que por sua vez equivale a

$$\left|\sum_{(\mathcal{A}_0,T_0)} f - \sum_{(\mathcal{A}_1,T_1)} f\right| \leq U(f,\mathcal{Q}) - L(f,\mathcal{Q}).$$

Adaptando o que se fez no último *Katzensprung* do Lema 2.4.4, mostra-se (já sabe né?)$^{\star}_{\star\star}$ que existem *tags* $l := (l_1,\ldots,l_n)$ e $u := (u_1,\ldots,u_n)$ de $\mathcal{Q}$ satisfazendo

$$L(f,\mathcal{Q}) > \sum_{(\mathcal{Q},l)} f - \varepsilon' \quad \text{e} \quad U(f,\mathcal{Q}) < \sum_{(\mathcal{Q},u)} f + \varepsilon'. \tag{2.14}$$

Logo,

$$U(f,\mathcal{Q}) - L(f,\mathcal{Q}) < \sum_{(\mathcal{Q},u)} f - \sum_{(\mathcal{Q},l)} f + 2\varepsilon' \leq \left|\sum_{(\mathcal{Q},u)} f - \sum_{(\mathcal{Q},l)} f\right| + 2\varepsilon' < \varepsilon' + 2\varepsilon' = 3\varepsilon',$$

onde a última desigualdade se deve à hipótese sobre $\mathcal{Q}$ (complete os demais detalhes!)$^{\star}_{\star}$. □

Exemplo 2.4.10 (Compare com o Exemplo 1.3.27)**.** Para um subconjunto $S \subseteq [a,b]$, considere $\chi_S\colon [a,b] \to \{0,1\}$ a *função característica* de S, i.e., tal que $\chi_S(x) = 1$ se, e somente se, $x \in S$. Agora, com $\mathbb{Q} \cap [0,1] = \{q_n : n \in \mathbb{N}\}$, tome $F_m := \{q_n : n \leq m\}$ bem com sua função característica $\chi_{F_m}\colon [0,1] \to \{0,1\}$ para cada $m \in \mathbb{N}$. Pelo Exercício 2.136, $\int_0^1 F_m(t)\,\mathrm{d}t = 0$ para todo m. Observe então que $\chi_{\mathbb{Q}} = \lim_{m\to\infty} \chi_{F_m}$ (certo?!)*, mas $\chi_{\mathbb{Q}}$ não é Riemann-integrável, já que a rede de suas somas de Riemann não é de Cauchy: para qualquer partição $\mathcal{P}$ de $[0,1]$, podemos tomar *tags* compostas unicamente por números racionais, bem como *tags* compostas unicamente por irracionais (pense a respeito)$^{\star}_{\star}$. ▲

Observação 2.4.11 (Integrais de Darboux vs. integrais de Riemann)**.** Somas de Darboux não são somas de Riemann. Com isso dito, somas de Darboux podem ser usadas para descrever a integral de Riemann. Por ser uma abordagem muito comum na literatura, convém apresentá-la brevemente aqui — e faço isso apenas por este motivo.

Primeiro, *observe* (!)$^{\star}_{\star}$ que o Exercício 2.118 assegura que os conjuntos

$$\mathcal{L}_a^b(f) := \{L(f,\mathcal{P}) : \mathcal{P} \in \mathrm{Par}\,[a,b]\} \quad \text{e} \quad \mathcal{U}_a^b(f) := \{U(f,\mathcal{P}) : \mathcal{P} \in \mathrm{Par}\,[a,b]\}$$

são, respectivamente, limitados superior e inferiormente, onde $\mathrm{Par}\,[a,b]$ indica o conjunto das partições de $[a,b]$ (sem *tags*!). Na verdade, somas de Darboux inferiores limitam inferiormente as somas de Darboux superiores, enquanto as segundas limitam superiormente as primeiras.

Definição 2.4.12. Para uma função limitada $f\colon [a,b] \to \mathbb{R}$, os números reais

$$\underline{\int_a^b} f(t)\,\mathrm{d}t := \sup \mathcal{L}_a^b(f) \quad \text{e} \quad \overline{\int_a^b} f(t)\,\mathrm{d}t := \inf \mathcal{U}_a^b(f)$$

são xingados, respectivamente, de **integral de Darboux inferior** e **integral de Darboux superior** de f. Quando

$$\underline{\int_a^b} f(t)\,\mathrm{d}t = \overline{\int_a^b} f(t)\,\mathrm{d}t,$$

a função f é dita **Darboux-integrável**, caso em que o número acima passa a ser chamado de a **integral de Darboux** de f, que denotaremos simplesmente por $\int_a^b f(t)\,\mathrm{d}t$. ¶

Exercício 2.119 $(^{\star}_{\star})$**.** Mostre que a desigualdade $\underline{\mathcal{D}\!\!\!\int_a^b} f(t)\,\mathrm{d}t \leq \overline{\mathcal{D}\!\!\!\int_a^b} f(t)\,\mathrm{d}t$ se verifica para qualquer função limitada $f\colon [a,b] \to \mathbb{R}$. ■

Tipicamente, após definir a noção de integral acima, é comum investir um tempo considerável na verificação de que as condições elencadas no começo da Subseção 2.4.0 §0 são satisfeitas, tudo por meio das propriedades elementares de supremos e ínfimos. Não é algo particularmente difícil, mas é suficientemente não trivial para criar caráter. Em todo caso, isto não precisará ser feito aqui por conta da próxima... △

Proposição 2.4.13. *Uma função limitada $f\colon [a,b] \to \mathbb{R}$ é Riemann-integrável se, e somente se, é Darboux-integrável. Em particular, se f é Riemann-integrável, então*

$$\mathcal{D}\!\!\!\int_a^b f(t)\,\mathrm{d}t = \int_a^b f(t)\,\mathrm{d}t.$$

Demonstração. Primeiro, observe $(^{\star}_{\star})$ que ao considerar $\mathrm{Par}_{\mathcal{R}}\,[a,b]$ dirigido pela relação $\sqsubseteq$, temos

$$\lim_{(\mathcal{P},T)\in\mathrm{Par}_{\mathcal{R}}\,[a,b]} L(f,\mathcal{P}) = \underline{\mathcal{D}\!\!\!\int_a^b} f(t)\,\mathrm{d}t \quad \text{e} \quad \lim_{(\mathcal{P},T)\in\mathrm{Par}_{\mathcal{R}}\,[a,b]} U(f,\mathcal{P}) = \overline{\mathcal{D}\!\!\!\int_a^b} f(t)\,\mathrm{d}t.$$

Logo, ao assumir f Darboux-integrável, a Riemann-integrabilidade de f segue por “sanduíche” (cf. Proposição 1.2.18), haja vista a desigualdade

$$L(f,\mathcal{P}) \leq \sum_{(\mathcal{P},T)} f \leq U(f,\mathcal{P})$$

ser válida para qualquer partição de Riemann $(\mathcal{P},T)$ de $[a,b]$. Reciprocamente, se f é Riemann-integrável, então ao fixar uma *tag* $T_{\mathcal{P}}$ para cada partição $\mathcal{P}$ de $[a,b]$, temos $\{(\mathcal{P},T_{\mathcal{P}}) : (\mathcal{P},T) \in \mathrm{Par}_{\mathcal{R}}\,[a,b]\}$ cofinal em $(\mathrm{Par}_{\mathcal{R}}\,[a,b],\sqsubseteq)$ (cf. Subseção 1.3.1 §0 + Exercício 1.61), acarretando

$$\lim_{\mathcal{P}} \sum_{(\mathcal{P},T_{\mathcal{P}})} f = \int_a^b f(t)\,\mathrm{d}t.$$

Com isso, para $\varepsilon > 0$ fixado, não é difícil adaptar as desigualdades obtidas em (2.14) (pág. 276) para mostrar que

$$\underline{\mathcal{D}\!\!\!\int_a^b} f(t)\,\mathrm{d}t = \lim_{\mathcal{P}} L(f,P) \geq \int_a^b f(t)\,\mathrm{d}t - \varepsilon \quad \text{e} \quad \overline{\mathcal{D}\!\!\!\int_a^b} f(t)\,\mathrm{d}t = \lim_{\mathcal{P}} U(f,\mathcal{P}) \leq \int_a^b f(t)\,\mathrm{d}t + \varepsilon, \tag{2.15}$$

donde o resultado segue. □

Exercício 2.120 $(^{\star}_{\star\star})$**.** Complete os detalhes da última demonstração. ■

Diferente da proposição anterior, o próximo teorema é *realmente* importante, por estabelecer um critério qualitativo para Riemann-integrabilidade tremendamente eficaz.

Teorema 2.4.14 (Critério de Lebesgue)**.** *Uma função $f\colon [a,b] \to \mathbb{R}$ é Riemann-integrável se, e somente se, f é limitada e contínua a menos de um conjunto de medida nula.*

Acima, falta explicitar o que significa “conjunto de medida nula”. Embora isto seja a ponta de um *iceberg* gigantesco chamado Teoria da Medida, por ora basta saber que se trata de uma generalização da *medida de intervalos* proposta em (2.4) (pág. 252).

Definição 2.4.15. Dizemos que $N \subseteq \mathbb{R}$ **tem medida nula** se para cada $\varepsilon > 0$ existe uma coleção $\{I_n\}_{n\in\mathbb{N}}$ de intervalos abertos (possivelmente vazios) satisfazendo $N \subseteq \bigcup_{n\in\mathbb{N}} I_n$ e $\sum_{n=0}^{\infty} \mu(I_n) \leq \varepsilon$. ¶

Exemplo 2.4.16. *Qualquer subconjunto enumerável de $\mathbb{R}$ tem medida nula.* Com efeito, para $S := \{x_n : n \in \mathbb{N}\}$ e $\varepsilon > 0$, em torno de cada x_n podemos considerar o intervalo $I_n := \left(x_n - \frac{\varepsilon}{2^{n+2}}, x_n + \frac{\varepsilon}{2^{n+2}}\right)$, que satisfaz $\mu(I_n) = \frac{\varepsilon}{2^{n+1}}$ e, por conseguinte,

$$\sum_{n=0}^{\infty} \mu(I_n) = \varepsilon \sum_{n=0}^{\infty} \frac{1}{2^{n+1}} = \varepsilon$$

mostrando que S satisfaz a condição anterior. ▲

Exemplo 2.4.17. *Nem todo conjunto de medida nula é enumerável.* Aqui o exemplo canônico é $\mathfrak{C}$, o conjunto de Cantor (cf. Definição 2.2.27), e o argumento é adaptado de [32]. Primeiro, para $\varepsilon > 0$, fixe $N \in \mathbb{N}$ com $\frac{2^N}{3^N} < \varepsilon$ (isto pode ser feito, certo?)*. Agora, note que $\mathfrak{C} \subseteq F_N$, que por sua vez é uma reunião de 2^N intervalos fechados de comprimento $\frac{1}{3^N}$, digamos $I_1, \ldots, I_{2^N}$ (reflita um pouco)*. Pelo modo como N foi tomado, é possível estender cada intervalo I_j a um intervalo (a_j, b_j) contendo I_j e satisfazendo $\mu((a_j, b_j)) = \frac{\varepsilon}{2^N}$ (por quê?)*. Por fim, note que $\mu((a_1, b_1)) + \ldots + \mu((a_{2^N}, b_{2^N})) = \varepsilon$. ▲

Exemplo 2.4.18. *Nem todo conjunto tem medida nula.* Com efeito, subconjuntos com interior não vazio não podem ter medida nula[†]: confira o Exercício 2.138. ▲

Exercício 2.121 ($\star$)**.** Mostre que se $M \subseteq N \subseteq \mathbb{R}$ e N tem medida nula, então M tem medida nula. ■

Aqui, a ideia de medida não deve ser confundida com a noção de cardinalidade: a última tem a ver com a *quantidade* de elementos do conjunto, enquanto a primeira quantifica o “espaço” ocupado na reta pelo conjunto[‡]. Com isso dito, tais conjuntos podem ser considerados como *desprezáveis*, no sentido de que o espaço ocupado por eles pode ser ajustado de modo a se tornar arbitrariamente pequeno. Aliando tal mantra ao Exemplo 2.4.16, chega-se à seguinte conclusão natural:

Proposição 2.4.19. *Seja $\{X_n : n \in \mathbb{N}\}$ uma coleção enumerável de subconjuntos de $\mathbb{R}$. Se cada X_n tem medida nula, então $\bigcup_{n\in\mathbb{N}} X_n$ tem medida nula.*

Demonstração. Fixado $\varepsilon > 0$, para cada $n \in \mathbb{N}$ existe uma cobertura enumerável de intervalos abertos para X_n, digamos $\{I_{n,k} : k \in \mathbb{N}\}$, satisfazendo

$$\sum_{k=0}^{\infty} \mu(I_{n,k}) \leq \frac{\varepsilon}{2^{n+1}},$$

[†]Cuidado: aqui, “não tem medida nula” não é equivalente a “tem medida positiva”, pois não definimos *uma* noção geral de *medida*.

[‡]Talvez fique mais fácil aceitar em $\mathbb{R}^2$, em que a noção de medida natural corresponde ao que entendemos por *área*: neste caso, é intuitivo aceitar que segmentos de reta têm área zero, embora tenham uma quantidade não enumerável de elementos.

o que sugere definir $\mathcal{I} := \{I_{n,k} : n, k \in \mathbb{N}\}$, pois assim

$$\sum_{n=0}^{\infty}\left(\sum_{k=0}^{\infty}\mu(I_{n,k})\right) \leq \sum_{n=0}^{\infty}\frac{\varepsilon}{2^{n+1}} \leq \varepsilon.$$

O problema é que a definição pede que a família de intervalos seja indexada por $\mathbb{N}$, enquanto a família $\mathcal{I}$ que encontramos está indexada por $\mathbb{N}\times\mathbb{N}$. Ora, como $\mathbb{N}\times\mathbb{N}$ tem a mesma cardinalidade de $\mathbb{N}$, não é difícil perceber que podemos reescrever $\mathcal{I} = \{J_m : m \in \mathbb{N}\}$ (certo?)$^{\star}_{\star}$, mas a princípio isto traz o problema de determinar $\displaystyle\sum_{m=0}^{\infty}\mu(J_m)$, não? □

Exercício 2.122 $\left({}^{\star}_{\star}\right)$**.** Resolva o impasse anterior. Sugestão: dê um jeito de perceber que $\sum_{m=0}^{n}\mu(J_m) \leq \varepsilon$ para todo $n \in \mathbb{N}$. ■

Corolário 2.4.20 (Princípio da casa dos pombos para conjuntos de medida zero)**.** *Seja $\{X_n : n \in \mathbb{N}\}$ uma coleção enumerável de subconjuntos de $\mathbb{R}$. Se $X := \bigcup_{n\in\mathbb{N}} X_n$ não tem medida nula, então existe $n \in \mathbb{N}$ tal que X_n não tem medida nula.*

Exercício 2.123 $(\star)$**.** Prove o corolário acima. ■

O Princípio da Casa dos Pombos é uma das peças fundamentais na demonstração do Critério de Lebesgue: na "ida", a fim de mostrar que o conjunto $D \subseteq [a, b]$ dos pontos de descontinuidade de f tem medida nula, vamos dar um jeito de escrever $\mathcal{D} = \bigcup_{n\in\mathbb{N}}\mathcal{D}_n$ e mostrar que cada $\mathcal{D}_n$ tem medida nula. Falta, no entanto, um ingrediente para a "volta".

Lema 2.4.21 (número de Lebesgue)**.** *Se $K \subseteq \mathbb{R}$ é compacto e $\mathcal{U}$ é uma cobertura aberta para K, então existe $\delta > 0$ satisfazendo a seguinte condição: sempre que $S \subseteq K$ tem diâmetro*[†] *menor do que δ, existe $U \in \mathcal{U}$ tal que $S \subseteq U$. O número δ costuma ser xingado de (um)* ***número de Lebesgue*** *de $\mathcal{U}$.*

Demonstração. Trata-se de uma adaptação da prova para o Teorema de Heine-Cantor. Para cada $x \in K$ existem $U_x \in \mathcal{U}$ com $x \in U_x$ e $\delta_x > 0$ com $I_x := (x - \delta_x, x + \delta_x) \subseteq U_x$. Por K ser compacto, existem $x_0, \ldots, x_n \in K$ com $K \subseteq \bigcup_{i\leq n}\left(x_i - \frac{\delta_{x_i}}{2}, x_i + \frac{\delta_{x_i}}{2}\right)$, o que permite tomar $\delta' := \min\left\{\frac{\delta_{x_i}}{2} : i \leq n\right\}$ e $0 < \delta < \delta'$. Por fim, se $S \subseteq K$ é não vazio, digamos que com $z \in S$, e $\operatorname{diam}(S) \leq \delta$, então $S \subseteq (z - \delta', z + \delta')$ e este intervalo, por sua vez, está contido em I_{x_i}, onde i é tal que $z \in \left(x_i - \frac{\delta_{x_i}}{2}, x_i + \frac{\delta_{x_i}}{2}\right)$. Os detalhes ficam por sua conta $\left({}^{\star}_{\star\star}\right)$. □

Demonstração do Critério de Lebesgue. Primeiro, vamos *investigar* o conjunto $\mathcal{D} := \{x \in [a, b] : f \text{ é descontínua em } x\}$ com mais calma. Pela caracterização de continuidade em termos de convergência, temos $x \in \mathcal{D}$ se, e somente se, existem $\varepsilon > 0$ e uma sequência $(x_n)_{n\in\mathbb{N}}$ em $[a, b]$ tais que $x_n \to x$ e $|f(x_m) - f(x)| \geq \varepsilon$ para qualquer $m \in \mathbb{N}$ (certo?)[‡]. Logo, ao declarar

$$\mathcal{D}_k := \left\{ x \in [a,b] \ : \begin{array}{c} \text{existe } (x_n)_{n\in\mathbb{N}} \subseteq [a,b] \text{ tal que } x_n \to x \text{ e} \\ |f(x_m) - f(x)| \geq \dfrac{1}{2^k} \text{ para todo } m \in \mathbb{N} \end{array} \right\}$$

[†]O diâmetro de S é o número $\operatorname{diam}(S) := \sup\{|x - y| : x, y \in S\}$ (cf. Proposição 2.0.20).

[‡]Dica: a princípio, temos apenas "$x \in \mathcal{D}$ se, e somente se, existe $(x_n)_{n\in\mathbb{N}}$ com $x_n \to x$ e $f(x_n) \not\to f(x)$", mas daí basta usar a "solução" do Exercício 1.102.

para cada $k \in \mathbb{N}$, resulta $\mathcal{D} = \bigcup_{k\in\mathbb{N}} \mathcal{D}_k$.

Antes de prosseguir: verifique a igualdade anterior, e aproveite a oportunidade para observar que $\mathcal{D}_k \subseteq \mathcal{D}_{k+1}$ para todo $k \in \mathbb{N}$ $\binom{\star}{\star}$!

Agora sim, supondo f Riemann-integrável, já sabemos que f é limitada (cf. Proposição 1.2.30). Para mostrar que $\mathcal{D}$ tem medida nula, basta mostrar que $\mathcal{D}_k$ tem medida nula para todo $k \in \mathbb{N}$. Para tanto, fixado $\varepsilon > 0$, existe uma partição $\mathcal{P} := (p_0, \ldots, p_m)$ de $[a, b]$ tal que

$$\left| \sum_{(\mathcal{P},T)} f - \sum_{(\mathcal{P},S)} f \right| < \frac{\varepsilon}{2^k}$$

para quaisquer *tags* T e S associadas à partição $\mathcal{P}$. A sacada agora é perceber que se $\mathcal{D}_k \cap (p_{i-1}, p_i) \neq \emptyset$, então existem $s_i, t_i \in (p_{i-1}, p_i)$ distintos tais que $|f(t_i) - f(s_i)| \geq \frac{1}{2^k}$ e, sem perda de generalidade, podemos assumir $f(s_i) < f(t_i)$. Perceba! $\binom{\star}{\star}$. Enfim, chamando $J := \{j \leq m : (p_{j-1}, p_j) \cap D_k \neq \emptyset\}$, resulta

$$\frac{1}{2^k} \sum_{j\in J} (p_j - p_{j-1}) \leq \sum_{j\in J} (f(t_j) - f(s_j))(p_j - p_{j-1}) = \sum_{(\mathcal{P},T)} f - \sum_{(\mathcal{P},S)} f < \frac{\varepsilon}{2^k}, \tag{2.16}$$

onde $T := (t_1, \ldots, t_n)$ e $S := (s_1, \ldots, s_n)$ são tais que $t_i := s_i$ para todo $i \in \{1, \ldots, m\} \setminus J$.

Desse modo, a desigualdade (2.16) mostra que $\{(p_{j-1}, p_j) : j \in J\}$ é uma coleção finita de intervalos abertos satisfazendo

$$\mathcal{D}_k \setminus \{p_0, \ldots, p_m\} \subseteq \bigcup_{j\in J} (p_{j-1}, p_j) \quad \text{e} \quad \sum_{j\in J} \mu((p_{j-1}, p_j)) \leq \varepsilon,$$

donde o resultado desejado segue. Completar os poucos detalhes omitidos será problema seu $\binom{\star}{\star}$.

Para a recíproca, assumindo f limitada e $\mathcal{D}$ com medida nula, mostraremos que f é Riemann-integrável por meio do Critério de Riemann-Darboux (cf. Teorema 2.4.9). Fixados $\gamma, \gamma', \varepsilon > 0$, a hipótese assegura uma coleção $\{I_n : n \in \mathbb{N}\}$ de intervalos abertos satisfazendo

$$\mathcal{D}' := \mathcal{D} \cup \{a, b\} \subseteq \bigcup_{n\in\mathbb{N}} I_n \quad \text{e} \quad \sum_{n=0}^{\infty} \mu(I_n) < \gamma.$$

Por sua vez, para cada $x \in [a, b] \setminus \mathcal{D}'$, existe um intervalo aberto $J_x \subseteq (a, b)$ tal que $|f(x) - f(y)| < \gamma'$ para todo $y \in J_x$. Agora, o maravilhoso lema anterior garante um número de Lebesgue $\delta > 0$ para a cobertura aberta $\mathcal{U} := \{I_n : n \in \mathbb{N}\} \cup \{J_x : x \in [a, b] \setminus \mathcal{D}'\}$. Consequentemente, fixando uma partição $\mathcal{P} := (p_0, \ldots, p_m)$ para $[a, b]$ com $\|\mathcal{P}\| < \delta$, resulta que para cada $i \leq m$, uma das duas possibilidades a seguir ocorre:

(a) existe $n \in \mathbb{N}$ tal que $[p_{i-1}, p_i] \subseteq I_n$,

(b) ou existe $x \in [a, b] \setminus \mathcal{D}'$ tal que $[p_{i-1}, p_i] \subseteq J_x$.

Tratemos do critério de Riemann-Darboux. Fixadas *tags* T e T' para $\mathcal{P}$, temos

$$\begin{aligned} \left| \sum_{(\mathcal{P},T)} f - \sum_{(\mathcal{P},T')} f \right| &\leq \sum_{i=1}^{m} |f(t_i) - f(t_i')||p_i - p_{i-1}| \\ &= \sum_{i\in A} |f(t_i) - f(t_i')||p_i - p_{i-1}| + \sum_{i\in B} |f(t_i) - f(t_i')||p_i - p_{i-1}|, \end{aligned}$$

onde $A := \{i \leq m : \text{ocorre } (a) \text{ para } [p_{i-1}, p_i]\}$ e $B := \{i \leq m : \text{não ocorre } (a) \text{ para } [p_{i-1}, p_i]\}$. Note que para cada $i \in A$, a hipótese de f ser limitada (finalmente!) acarreta $|f(t_i) - f(t_i')||p_i - p_{i-1}| \leq 2\|f\|_\infty |p_i - p_{i-1}|$ e, consequentemente,

$$\sum_{i\in A} |f(t_i) - f(t_i')||p_i - p_{i-1}| \leq \sum_{i\in A} 2\|f\|_\infty \mu(I_{n_i}) \leq 2\|f\|_\infty \gamma.$$

Por outro lado, para $i \in B$, $|f(t_i) - f(t_i')| \leq 2\gamma'$ (por quê?!)$^\star_\star$, donde segue que

$$\sum_{i\in B} |f(t_i) - f(t_i')||p_i - p_{i-1}| \leq \sum_{i\in B} 2\gamma'(p_i - p_{i-1}) \leq 2\gamma'(b-a).$$

Portanto, escolhendo $\gamma < \frac{\varepsilon}{4\|f\|_\infty}$ e $\gamma' < \frac{\varepsilon}{4(b-a)}$, resulta

$$\left| \sum_{(\mathcal{P},T)} f - \sum_{(\mathcal{P},T')} f \right| < \varepsilon$$

para quaisquer *tags* T e T' de $\mathcal{P}$, como desejado. □

Observação 2.4.22. A prova acima adapta os argumentos de Pugh [32] e Gerald Edgar[†] com o propósito de omitir a noção de *oscilação*[‡]. Se, por qualquer motivo, tais definições forem importantes para você, confira o texto de Pugh [32], os cânones de Lima [24, 25] ou, se tiver apenas uma curiosidade mórbida, o Exercício 2.152. △

Vamos aos refrescos!

Corolário 2.4.23. *Se $f : [a, b] \to \mathbb{R}$ é Riemann-integrável, então $|f|$ é Riemann-integrável.*

Demonstração. Como a função $|\cdot| : \mathbb{R} \to [0, +\infty)$ é contínua, segue que se $|f|$ é descontínua em $p \in [a, b]$, então f é descontínua em p. Logo, o resultado segue do teorema anterior aliado ao Exercício 2.121. □

Exercício 2.124 ($^\star_\star$). Para $f : [a, b] \to \mathbb{R}$, considere $f^+ := \max\{f, 0\}$ e $f^- := \max\{-f, 0\}$.

a) Mostre que $f^+ \geq 0$, $f^- \geq 0$, $f = f^+ - f^-$ e $|f| = f^+ + f^-$.

b) Mostre que f é Riemann-integrável se, e somente se, f^+ e f^- são Riemann-integráveis, e vale $\displaystyle\int_a^b f(t)\,dt = \int_a^b f^+(t)\,dt - \int_a^b f^-(t)\,dt.$ ■

Observação 2.4.24. As funções f^+ e f^- acima são xingadas, respectivamente, de **partes positiva** e **negativa** de f, e costumam ser bem úteis quando precisamos assumir que a função não muda de sinal. △

Corolário 2.4.25. *Se $f, g : [a, b] \to \mathbb{R}$ são Riemann-integráveis, então $fg : [a, b] \to \mathbb{R}$ é Riemann-integrável.*

Demonstração. Se fg é descontínua em $p \in [a, b]$, então f ou g são descontínuas em p. Conclua ($\star$). □

[†] Cf. https://math.stackexchange.com/a/163409/.

[‡] A omissão não se deve apenas ao senso estético, mas também a questões logísticas: oscilações em nada contribuem para os demais tópicos abordados no texto.

Exercício 2.125. Mostre que se $f\colon [a,b] \to \mathbb{R}$ é monótona, então f é Riemann-integrável. Dica: confira o Exercício 1.237 e combine com o Exemplo 2.4.16. ■

Exemplo 2.4.26. A conclusão do Exemplo 2.4.10 se torna automática com o critério de Lebesgue (por favor!)*. Mais interessante, porém, é perceber que as provas de (v) e (vi) nos *axiomas de integração* ficam bem mais simples agora. Para o axioma (v), por exemplo, basta notar que $f\colon [a,b] \to \mathbb{R}$ é limitada se, e somente se, $f|_{[a,c]}$ e $f|_{[c,b]}$ são limitadas, enquanto $D_f = D_{f|_{[a,c]}} \cup D_{f|_{[c,b]}}$ (entendeu?)*†. Já para a identidade (vi), note primeiro que para $a \leq \alpha \leq \beta \leq b$, a função

$$g(x) := \begin{cases} 1, & \text{se } \alpha \leq x \leq \beta \\ 0, & \text{caso contrário} \end{cases}$$

é *obviamente*‡ Riemann-integrável, de modo que para $f\colon [a,b] \to \mathbb{R}$ Riemann-integrável resulta fg Riemann-integrável em $[a,b]$ (pelo corolário anterior). Logo, as integrais

$$\int_\alpha^\beta f(t)\,\mathrm{d}t \quad \text{e} \quad L := \int_a^b f(t)g(t)\,\mathrm{d}t$$

existem, e mostraremos que ambas coincidem.

Primeiro, note que a família

$$\mathscr{C} := \{(\mathcal{P},T) \in \mathrm{Par}_{\mathcal{R}}[a,b] : \alpha, \beta \in \mathrm{im}(\mathcal{P})\}$$

é cofinal em $\mathrm{Par}_{\mathcal{R}}[a,b]$ (com respeito à ordem $\sqsubseteq$), de modo que $\lim\limits_{(\mathcal{P},T)\in\mathscr{C}} \sum\limits_{(\mathcal{P},T)} fg = L$. Em particular, para $\varepsilon > 0$, existe uma partição $\left(\widetilde{\mathcal{P}}, \widetilde{T}\right) \in \mathscr{C}$ tal que $\left|\sum_{(\mathcal{P},T)} fg - L\right| < \varepsilon$ sempre que $(\mathcal{P},T) \in \mathscr{C}$ satisfaz $\mathcal{P} \sqsupseteq \widetilde{\mathcal{P}}$. Ao "restringir" $\left(\widetilde{\mathcal{P}}, \widetilde{T}\right)$ a uma partição $\left(\widetilde{\mathcal{Q}}, \widetilde{S}\right)$ de $[\alpha,\beta]$, não é difícil perceber que se $(\mathcal{Q},S) \in \mathrm{Par}_{\mathcal{R}}[\alpha,\beta]$ for tal que $\widetilde{\mathcal{Q}} \sqsubseteq \mathcal{Q}$, então existe uma "extensão" $(\mathcal{Q}',S') \in \mathscr{C}$ tal que $\widetilde{\mathcal{P}} \sqsubseteq \mathcal{Q}'$ e $\sum_{(\mathcal{Q}',S')} fg = \sum_{(\mathcal{Q},S)} f$, donde a igualdade desejada segue (percebeu?)$^\star_\star$.

Para encerrar, basta definir g como acima com $\alpha := a$ e $\beta := c$, bem como $h(x) := 1$ se $x \in (c,b]$ e $h(x) = 0$ caso contrário, pois daí $f = fg + fh$ e

$$\int_a^b f(t)\,\mathrm{d}t = \int_a^b f(t)g(t) + f(t)h(t)\,\mathrm{d}t = \int_a^c f(t)\,\mathrm{d}t + \int_c^b f(t)\,\mathrm{d}t$$

(cf. Exercício 2.136). ▲

Observação 2.4.27 (Opcional: sub-redes de Willard)**.** Secretamente, o último argumento usa uma noção de sub-rede mais geral do que a empregada na Definição 1.3.21.

Definição 2.4.28. Diremos que uma função $h\colon \mathbb{A} \to \mathbb{B}$ entre conjuntos dirigidos $\mathbb{A}$ e $\mathbb{B}$ é **de Willard** se h for crescente e para todo $B \in \mathbb{B}$ existir $A \in \mathbb{A}$ com $h(A) \succeq B$. Diremos que uma rede $\varphi := (y_a)_{a\in\mathbb{A}}$ é **W-sub-rede** de $\psi := (x_b)_{b\in\mathbb{B}}$ se existir uma função de Willard $h\colon \mathbb{A} \to \mathbb{B}$ tal que $\psi \circ h = \varphi$, i.e., tal que $y_a = x_{h(a)}$ para todo $a \in \mathbb{A}$. ¶

†Evidentemente, D_g denota o conjunto dos pontos no domínio de g em que g é descontínua.
‡Desculpe, mas desta vez é óbvio mesmo.

$$
\begin{tikzcd}
\mathbb{B} \arrow[r, "\psi"] & X \\
\mathbb{A} \arrow[u, "h"] \arrow[ur, "\varphi"'] &
\end{tikzcd}
\qquad\qquad
\begin{tikzcd}
h(a) \arrow[r, mapsto] & x_{h(a)} = y_a \\
a \arrow[u, mapsto] \arrow[ur] &
\end{tikzcd}
$$

Figura 2.3: Compare com o diagrama de subsequências na Figura 1.7, na página 143.

Sub-redes no sentido da Definição 1.3.21 são W-sub-redes, pois se $\mathcal{C}$ é um subconjunto cofinal de $\mathbb{A}$, então a inclusão $i\colon \mathcal{C} \to \mathbb{A}$ é uma função de Willard. No entanto, W-sub-redes permitem flexibilizar o modo como os conjuntos de índices $\mathbb{A}$ e $\mathbb{B}$ se relacionam, mas sem perder a propriedade mais interessante das sub-redes e subsequências:

Proposição 2.4.29. *Sejam $(y_a)_a$ e $(x_b)_b$ redes reais tais que $(y_a)_a$ é W-sub-rede de $(x_b)_b$. Se $x_b \to z$, então $y_a \to z$.*

Demonstração. Dado um aberto V com $z \in V$, busca-se $A \in \mathbb{A}$ tal que $y_a \in V$ sempre que $a \succeq A$. Ora, a existência da função de Willard $h\colon \mathbb{A} \to \mathbb{B}$ como na definição permite tomar A tal que $h(A) \succeq B$, onde B é tal que $x_b \in V$ sempre que $b \succeq B$, o qual existe pois $x_b \to z$ por hipótese: de fato, se $a \succeq A$, então $h(a) \succeq h(A) \succeq B$ e, portanto, $y_a = x_{h(a)} \in V$. $\square$

Exercício 2.126 (Opcional — $(\star)$). Adapte a definição de W-sub-redes para redes em espaços topológicos e perceba que a última proposição permanece válida. $\blacksquare$

No caso do exemplo anterior, secretamente definimos uma função $h\colon \mathscr{C} \to \mathrm{Par}_{\mathcal{R}}\,[\alpha,\beta]$ em que $h((\mathcal{P},T))$ é restrição da partição $(\mathcal{P},T) \in \mathscr{C} \subseteq \mathrm{Par}_{\mathcal{R}}\,[a,b]$ a uma partição de Riemann do intervalo $[\alpha,\beta]$. O fato de valer a identidade

$$\sum_{(\mathcal{P},T)} fg = \sum_{h((\mathcal{P},T))} f$$

significa que o diagrama

$$
\begin{tikzcd}
\mathrm{Par}_{\mathcal{R}}\,[\alpha,\beta] \arrow[r, "\sum_{\bullet} f"] & \mathbb{R} \\
\mathscr{C} \arrow[u, "h"] \arrow[ur, "\sum_{\bullet} fg"'] &
\end{tikzcd}
$$

comuta. Uma vez que a função h é de Willard, a igualdade das integrais (limites das redes) segue da proposição anterior (pense a respeito!)***. $\triangle$

Corolário 2.4.30. *Se $f\colon [a,b] \to \mathbb{R}$ é Riemann-integrável e*

$$\int_a^b |f(t)|\,\mathrm{d}t = 0,$$

então $S(f) := \{x \in [a,b] : f(x) \neq 0\}$ tem medida nula.

Demonstração. Como $|f| = f^+ + f^-$ (cf. Exercício 2.124), temos

$$\int_a^b f^+(t)\,\mathrm{d}t = \int_a^b f^-(t)\,\mathrm{d}t = 0,$$

(certo?)**, de modo que não há perda de generalidade em assumir $f \geq 0$ (entendeu?)*. Agora, supondo que o conjunto $S(f)$ não tem medida nula, deve-se ter

$$S(f) \not\subseteq \{x \in [a,b] : f \text{ é descontínua em } x\} \text{ (por quê?)}^{\star},$$

donde segue que f é contínua em algum ponto $p \in S(f)$. Por continuidade e conservação de sinal, existem $c, d \in [a, b]$ com $c < d$ tais que $f(x) > \frac{f(p)}{2}$ para todo $x \in [c, d]$ (certo?)*. Logo, ao considerar a função g do exemplo anterior, obtemos

$$0 \leq \frac{f(p)}{2} \cdot g(x) \leq f(x)$$

para todo $x \in [a, b]$ e, por conseguinte,

$$0 \leq \frac{f(p)}{2}(d - c) = \int_c^d \frac{f(p)}{2}\, \mathrm{d}t = \int_a^b \frac{f(p)}{2} \cdot g(t)\, \mathrm{d}t \leq \int_a^b f(t)\, \mathrm{d}t = 0,$$

o que contraria a suposição acerca de $f(p)$. $\square$

Na verdade, o corolário anterior mostra que $S(f)$ está contido no conjunto dos pontos de descontinuidade de f sempre que $\int_a^b |f(t)|\, \mathrm{d}t = 0$. Consequentemente, se f é contínua...

Exercício 2.127 $({}^{\star}_{\star})$**.** Mostre que se $f\colon [a, b] \to \mathbb{R}$ é contínua e

$$\int_a^b |f(t)|\, \mathrm{d}t = 0,$$

então $f = 0$. Dica: precisa mesmo? $\blacksquare$

Corolário 2.4.31 (A norma L_1 revisitada (cf. Definição 2.1.26))**.** *A função*

$$\begin{aligned} \|\cdot\|_1 \colon \mathcal{C}[a, b] &\to \mathbb{R} \\ f &\mapsto \int_a^b |f(t)|\, \mathrm{d}t \end{aligned}$$

define uma norma no espaço vetorial $\mathcal{C}[a, b]$*.*

Demonstração. O exercício anterior mostra que se $\|f\|_1 = 0$, então $f = 0$. As demais condições já foram verificadas na página 244. $\square$

Para encerrar esta lista inicial de refrescos[†], vamos apresentar a versão *integral* do Teorema 2.1.20 (compare também com os Corolários 2.1.24 e 2.1.27).

Teorema 2.4.32. *Seja* $(f_n)_{n\in\mathbb{N}}$ *uma sequência de funções Riemann-integráveis da forma* $[a, b] \to \mathbb{R}$*. Se* $f\colon [a, b] \to \mathbb{R}$ *é limite uniforme de* $(f_n)_{n\in\mathbb{N}}$*, então* f *é Riemann-integrável e*

$$\int_a^b f(t)\, \mathrm{d}t = \lim_{n\to\infty} \int_a^b f_n(t)\, \mathrm{d}t.$$

Demonstração. Chamando por $\mathcal{D}$ a coleção dos pontos de descontinuidade de f e $\mathcal{D}_n$ a coleção dos pontos de descontinuidade de f_n, temos

$$\mathcal{D} \subseteq \bigcup_{n\in\mathbb{N}} \mathcal{D}_n$$

em virtude do Teorema 2.1.11 (percebeu, né?)*. Logo, se cada f_n é contínua a menos de um conjunto de medida nula, então a mesma condição se verifica para f. Por sua vez, se cada f_n é limitada, então f é limitada (cf. Proposição 2.1.13). Portanto, a Riemann-integrabilidade de f segue do Critério de Lebesgue.

Para a igualdade, note que

$$\left|\int_a^b f(t)\, \mathrm{d}t - \int_a^b f_n(t)\, \mathrm{d}t\right| = \left|\int_a^b f(t) - f_n(t)\, \mathrm{d}t\right| \leq (b - a)\|f - f_n\|_\infty,$$

donde o resultado desejado segue (por quê?!!)*. $\square$

[†]Mais refrescos são propostos na Seção 2.5.

§1 Integração imprópria

A rigor, a integração de Riemann só pode ser realizada para funções limitadas da forma $[a,b] \to \mathbb{R}$. Porém, intervalos abertos podem ser reescritos como reuniões de intervalos fechados, o que abre uma margem de esperança para estender o escopo de funções Riemann-integráveis.

Seja $f\colon (a,b] \to \mathbb{R}$ uma função real tal que, para cada $\varepsilon \in (0, b-a)$, a restrição $f|_{[a+\varepsilon,b]}$ é Riemann-integrável. Em tal situação, para cada $\varepsilon > 0$ fica bem definido o número

$$\varphi(\varepsilon) := \int_{a+\varepsilon}^{b} f(t)\,\mathrm{d}t, \tag{2.17}$$

de modo que passa a fazer sentido considerar o limite

$$\int_a^b f(t)\,\mathrm{d}t := \lim_{\varepsilon \to 0^+} \varphi(\varepsilon) := \lim_{\varepsilon \to 0^+} \int_{a+\varepsilon}^{b} f(t)\,\mathrm{d}t,$$

xingado de **integral imprópria** da função f. Quando tal limite existe (em $\mathbb{R}$), dizemos que a integral imprópria de f **converge** (e costuma-se dizer que f é uma função **integrável**). Quando tal limite não existe (em $\mathbb{R}$), a tradição manda dizer que a integral imprópria de f **diverge**.

Observação 2.4.33. Note que se $f\colon (a,b] \to \mathbb{R}$ é tal que $f|_{[c,b]}$ é Riemann-integrável para cada $c \in (a,b]$, então vale a inclusão

$$\mathcal{N} := \{x \in (a,b] : f \text{ é descontínua em } x\} \subseteq \bigcup_{n \in \mathbb{N}} \mathcal{D}_n,$$

onde $\mathcal{D}_n := \left\{x \in \left[a + \frac{1}{2^n}, b\right] : f|_{\left[a+\frac{1}{2^n}, b\right]} \text{ é descontínua em } x\right\}$ para cada $n \in \mathbb{N}$, mostrando que $\mathcal{N}$ tem medida nula. Logo, se uma função $f\colon (a,b] \to \mathbb{R}$ pode ter sua integral imprópria definida (embora, possivelmente, inexistente), então o conjunto dos pontos de descontinuidade tem medida nula. Naturalmente, se a função $f\colon (a,b] \to \mathbb{R}$ é descontínua num conjunto de medida nula e limitada em cada intervalo da forma $[c,b]$, então podemos definir sua integral imprópria. Tacitamente, pelo restante da seção, assumiremos que todas as funções satisfazem tais condições. △

Curiosamente, uma função <u>limitada</u> $f\colon (a,b] \to \mathbb{R}$ para a qual faça sentido definir sua integral imprópria não é essencialmente uma novidade.

Proposição 2.4.34. *Seja $f\colon (a,b] \to \mathbb{R}$ uma função nas condições anteriores. Se f é limitada, então qualquer extensão $g\colon [a,b] \to \mathbb{R}$ de f é Riemann-integrável e satisfaz*

$$\int_a^b g(t)\,\mathrm{d}t = \int_a^b f(t)\,\mathrm{d}t.$$

Demonstração. A Riemann-integrabilidade de g segue do Critério de Lebesgue. Agora, pelo Exercício 2.143, temos

$$\int_a^b g(t)\,\mathrm{d}t = \lim_{c \to a^+} \int_c^b g(t)\,\mathrm{d}t.$$

O restante fica por sua conta ($\substack{\star\\\star}$). □

O valor da extensão g no extremo a é irrelevante devido ao Exercício 2.136. Em particular, o argumento acima mostra que se uma função $f\colon (a,b) \to \mathbb{R}$ limitada é contínua a menos de um conjunto de medida nula, então podemos estendê-la arbitrariamente a uma função $g\colon [a,b] \to \mathbb{R}$ Riemann-integrável, cuja integral (própria) de Riemann independe dos valores $g(a)$ e $g(b)$ e, por tal razão, pode ser atribuída como a integral de f. Talvez mais importante, a proposição anterior evidencia que a integração imprópria em funções definidas em intervalos da forma $(a,b]$ só tem "graça" para funções *ilimitadas*, uma vez que o caso limitado se reduz ao cenário da integração de Riemann usual.

Exercício 2.128 $(^{\star}_{\star})$**.** Seja $f\colon (0,1] \to \mathbb{R}$ a função dada por $f(x) := \frac{1}{x^2}$. Mostre que a integral imprópria de f diverge. ■

Teorema 2.4.35 (Critério de Cauchy)**.** *Seja $f\colon (a,b] \to \mathbb{R}$ uma função nas condições da Observação 2.4.33. Então a integral imprópria de f converge se, e somente se, para todo $\eta > 0$ existe $\varepsilon_0 > 0$ tal que para quaisquer $\varepsilon, \varepsilon' \in \mathbb{R}$ com $0 < \varepsilon, \varepsilon' < \varepsilon_0$ ocorre*

$$\left| \int_{a+\varepsilon}^{a+\varepsilon'} f(t)\,\mathrm{d}t \right| < \eta.$$

Demonstração. A convergência da integral imprópria de f equivale à existência do limite $\lim\limits_{\varepsilon \to 0^+} \varphi(\varepsilon)$, com $\varphi(\varepsilon)$ definido como em (2.17). Por sua vez, o limite acima existe se, e somente se, a rede correspondente $(\varphi(\varepsilon))_{\varepsilon \in (0,b-a)}$ é de Cauchy. Desse modo, o resultado segue ao se observar

$$\left| \int_{a+\varepsilon}^{a+\varepsilon'} f(t)\,\mathrm{d}t \right| = |\varphi(\varepsilon) - \varphi(\varepsilon')|.$$ □

Exercício 2.129 $(^{\star}_{\star})$**.** Observe! ■

Apesar de eventualmente útil, o critério acima pode ser bastante melhorado no caso de funções positivas. Começamos adaptando a noção de convergência absoluta, conhecida do contexto de séries.

Seja $f\colon (a,b] \to \mathbb{R}$ uma função nas condições anteriores. Como $f_{[c,b]}$ é integrável para cada $c \in (a,b]$, segue que a restrição de $|f|$ também é integrável em $[c,b]$ e, consequentemente, também temos o direito de calcular a integral imprópria de $|f|$, a saber

$$\int_a^b |f(t)|\,\mathrm{d}t := \lim_{\varepsilon \to 0^+} \int_{a+\varepsilon}^b |f(t)|\,\mathrm{d}t.$$

Dizemos que a integral imprópria de f **converge absolutamente** se a integral imprópria de $|f|$ convergir. Assim como ocorre com séries, temos o singelo

Corolário 2.4.36. *Seja $f\colon (a,b] \to \mathbb{R}$ uma função nas condições anteriores. Se a integral imprópria de f converge absolutamente, então a integral imprópria de f converge.*

Demonstração. Em geral temos

$$\left| \int_{a+\varepsilon}^{a+\varepsilon'} f(t)\,\mathrm{d}t \right| \leq \int_{a+\varepsilon}^{a+\varepsilon'} |f(t)|\,\mathrm{d}t \leq \left| \int_{a+\varepsilon}^{a+\varepsilon'} |f(t)|\,\mathrm{d}t \right|,$$

donde o resultado desejado segue da definição de convergência absoluta aliada ao teorema anterior. □

Observação 2.4.37 (Feita em aula por alguém[†] na minha turma de Análise II em 2019)**.** A hipótese de que f satisfaz as condições da Observação 2.4.33 é imprescindível: a função $g\colon (0,1] \to \mathbb{R}$ que faz $g(x) = -1$ se $x \notin \mathbb{Q}$ e $g(x) = 1$ caso contrário, não pode nem mesmo ter as integrais da forma $[\varepsilon, 1]$ definidas (por ser globalmente descontínua), mas $|g|$ é constante! △

Como no caso da convergência absoluta de séries, integrais absolutamente convergentes podem ser usadas para determinar a convergência da integral de outras funções de modo rápido e prático. Por exemplo:

Proposição 2.4.38. *Uma função $f\colon (a,b] \to [0,+\infty)$ nas condições anteriores tem sua integral convergente se, e somente se, existe $K > 0$ tal que $\int_{a+\varepsilon}^{b} f(t)\,\mathrm{d}t < K$ para todo $\varepsilon \in (0, b-a)$.*

Demonstração. Note que por ocorrer $f \geq 0$, a função $\varphi\colon (0, b-a) \to \mathbb{R}$ dada por $\varphi(\varepsilon) := \int_{a+\varepsilon}^{b} f(t)\,\mathrm{d}t$ é decrescente (por quê?)*. Logo, o limite

$$\int_a^b f(t)\,\mathrm{d}t := \lim_{\varepsilon \to 0^+} \varphi(\varepsilon)$$

existe em $\mathbb{R}$ se, e somente se, φ for limitada (certo?)$^{\star}_{\star}$. □

Naturalmente, as mesmas ideias utilizadas para funções da forma $(a,b] \to \mathbb{R}$ podem ser usadas para dar sentido a integrais (impróprias) de funções da forma $[a,b) \to \mathbb{R}$. Em particular, se uma função $f\colon (a,b) \to \mathbb{R}$ é tal que f é Riemann-integrável em cada subintervalo fechado $[c,d] \subset (a,b)$, podemos escolher $\gamma \in (a,b)$ e utilizar os dois tipos de integração imprópria, sobre $(a,\gamma]$ e $[\gamma, b)$, a fim de definir[‡]

$$\int_a^b f(t)\,\mathrm{d}t := \int_a^{\gamma} f(t)\,\mathrm{d}t + \int_{\gamma}^{b} f(t)\,\mathrm{d}t.$$

O próximo passo agora é inevitável: funções definidas em intervalos da forma $(-\infty, b]$, $[a, +\infty)$,... podem ter suas integrais impróprias definidas e investigadas desde que as restrições em cada subintervalo fechado sejam Riemann-integráveis. A Seção 2.5 traz algumas discussões sobre isso. Para encerrar:

Exercício 2.130 $(^{\star}_{\star})$**.** Seja $f\colon [1,+\infty) \to \mathbb{R}$ a função dada por $f(x) := \frac{1}{x^2}$. Mostre que a integral imprópria de f converge (ou prove que estou errado). Dica: não estou errado. ■

O material desta subseção foi baseado principalmente no Capítulo 8 de Djairo Guedes [9], com toques de medida nula extraídos de [24]. Se precisar de uma discussão aprofundada, confira o maravilhoso *Theories of Integration* [21] de Kurtz e Swartz.

[†] Meu palpite é de que tenha sido o Rígille Menezes e, muito provavelmente, por eu não ter explicitado sob quais hipóteses a função f deveria ser tomada.

[‡] Esta não é a única forma de definir tal tipo de integral. Com isso dito, note que para fazer sentido, precisa-se mostrar que o limite em questão independe da escolha de $\gamma \in (a,b)$ (cf. Exercício 2.148).

§2 O TFC revisitado

Suponha que você conheça duas funções reais, digamos $F, f\colon [a,b] \to \mathbb{R}$, tais que a função F é contínua em $[a,b]$ e diferenciável em (a,b) e, mais ainda, tenha-se $F'(x) = f(x)$ para todo $x \in (a,b)$. Neste caso, pelo que discutimos neste capítulo, parece seguro afirmar que

$$\int_a^b f(t)\,\mathrm{d}t = F(b) - F(a),$$

certo?

Resposta: errado.

A função que está *sendo integrada* acima é f, e não F! Para que a identidade acima tenha alguma chance de fazer sentido no presente contexto, a função $f\colon [a,b] \to \mathbb{R}$ *precisa* ser Riemann-integrável antes de qualquer outra coisa ou, equivalentemente, limitada e contínua a menos de um conjunto de medida nula. "*Mas o T.F.C. não provou, justamente, que a identidade anterior é verdadeira?*".

Resposta: não. Atente-se às hipóteses.

No T.F.C., mostramos que se $g\colon [a,b] \to \mathbb{R}$ é contínua, então g é Riemann-integrável e sua integral em $[a,b]$, que é um número real, pode ser calculada por meio da expressão $G(b) - G(a)$ para qualquer função diferenciável $G\colon [a,b] \to \mathbb{R}$ satisfazendo $G' = g$. Em particular, a função

$$\widetilde{G}(x) := \int_a^x g(t)\,\mathrm{d}t$$

cumpre tal tarefa no lugar de G. Por outro lado, na pergunta que abre a atual subsubseção, a função f, que ocuparia o lugar de g acima, não veio de fábrica com hipóteses de continuidade!

Exemplo 2.4.39. Se $F\colon [a,b] \to \mathbb{R}$ é contínua em $[a,b]$ mas tem derivada ilimitada em (a,b), então f não tem chances de ser Riemann-integrável, já que funções Riemann-integráveis são limitadas para começo de conversa. Quer um exemplo frustrante? Tome $F\colon [0,1] \to \mathbb{R}$ dada por $F(x) := \sqrt{x}$ e defina $f\colon [0,1] \to \mathbb{R}$ fazendo

$$f(x) := \begin{cases} 0, & \text{se } x = 0 \\ F'(x), & \text{caso contrário} \end{cases}.$$

Por construção, $F'(x) = f(x)$ para todo $x \in (0,1)$, mas f não é Riemann-integrável, já que

$$f(x) = \frac{1}{2\sqrt{x}}$$

para todo $x \in (0,1]$, e $f(x) \to +\infty$ se $x \to 0^+$. ▲

Exercício 2.131 ($^{\star}_{\star}$)**.** Mostre que $\displaystyle\int_0^1 \frac{1}{2\sqrt{x}}\,\mathrm{d}x = 1$ em sentido impróprio. ■

Observação 2.4.40. É possível encontrar $F\colon [a,b] \to \mathbb{R}$ contínua em $[a,b]$ e diferenciável em (a,b) tal que f' é limitada mas, ainda assim, f' não seja Riemann-integrável em $[a,b]$. Infelizmente, a descrição de tais exemplos exige mais paciência do que eu tenho agora. Caso queira saber mais a respeito, procure pela "função de Volterra" (cf. [35]). △

A boa notícia é que ao assumir a Riemann-integrabilidade de F', podemos afirmar que

$$\int_a^b F'(t)\,\mathrm{d}t = F(b) - F(a),$$

o que costuma ser chamado de *Segundo Teorema Fundamental do Cálculo* em algumas referências (cf. [1]). Note que tal afirmação generaliza o item (III) do Teorema 2.4.0, em que se pedia F' contínua.

Demonstração. Para qualquer partição $\mathcal{P} := (p_0, \ldots, p_n)$ de $[a,b]$ temos

$$F(b) - F(a) = \sum_{i=1}^{n} F(p_i) - F(p_{i-1}).$$

Como F é diferenciável em (a,b), o T.V.M. assegura que, para cada $i = 1, \ldots, n$, existe $t_i \in [p_{i-1}, p_i]$ satisfazendo

$$F'(t_i) = \frac{F(p_i) - F(p_{i-1})}{p_i - p_{i-1}}.$$

Acabou. □

Exercício 2.132 $\binom{\star}{\star}$**.** Complete a demonstração acima. Dica: use a hipótese de F' ser Riemann-integrável para garantir que qualquer sub-rede de $\left(\sum_{(\mathcal{P},T)} F'\right)_{(\mathcal{P},T)\in \mathrm{Par}_{\mathcal{R}}\,[a,b]}$ converge (cf. Exercício 1.61). ■

Com isso dito, o aspecto que realmente merece ênfase no T.F.C. decorre do item (II) em seu enunciado. Novamente: se $f\colon [a,b] \to \mathbb{R}$ é contínua, então existe $F\colon [a,b] \to \mathbb{R}$ contínua, diferenciável em (a,b) e satisfazendo $F'(c) = f(c)$ para todo $c \in (a,b)$. Explicitamente,

$$F(x) = \int_a^x f(t)\,\mathrm{d}t$$

para todo $x \in [a,b]$.

Na prática, $F\colon [a,b] \to \mathbb{R}$ é uma solução para a equação $U' = f$, em que U é a *função incógnita*. Dado que $F + c$ ainda satisfaz a equação para qualquer $c \in \mathbb{R}$, podemos resolver, na verdade, o *sistema*

$$\begin{cases} U' = f \\ U(a) = c \end{cases} \tag{2.18}$$

posto que $(F + c)(a) = F(a) + c = c$. Em outras palavras: estamos resolvendo *equações diferenciais*! Nesse sentido, o T.F.C. mostra que a integração de Riemann fornece um método *efetivo* para resolver sistemas da forma (2.18) em que f é contínua.

Essa é uma das formas de adentrar no mundo das *Equações Diferenciais* — mas esta é uma jornada que você precisará fazer por meio de outros textos. Aqui, chegamos ao fim.

2.5 Exercícios adicionais

Exercício 2.133 $\left(\substack{\star\\\star}\right)$**.** Mostre que existe $p > 0$ tal que $\cos(p) = 0$. Dica: se não fosse o caso, teríamos $\cos(x) > 0$ para todo $x \geq 0$ e (por algum motivo que você precisa descobrir!), $\operatorname{sen}(1) > 0$, acarretando

$$\cos(1) - \cos(x) = \int_1^x \operatorname{sen}(t)\,\mathrm{d}t \geq \int_1^x \operatorname{sen}(1)\,\mathrm{d}t = (x-1)\operatorname{sen}(1)$$

para todo $x > 1$, mas isto contraria a limitação da função cosseno. ■

Exercício 2.134 $(\star)$**.** Mostre que o conjunto $\{x > 0 : \cos(x) = 0\}$ tem menor elemento. Dica: trata-se de um conjunto fechado, não vazio e limitado inferiormente. ■

Observação 2.5.0. O menor elemento do conjunto acima é o que se costuma denotar por $\frac{\pi}{2}$... Sim, é isso mesmo: encontramos o π. △

Exercício 2.135 $\left(\substack{\star\\\star}\right)$**.** Mostre que

$$\left|\int_a^b f(t)\,\mathrm{d}t\right| \leq \int_a^b |f(t)|\,\mathrm{d}t \leq (b-a)\|f\|_\infty$$

para qualquer $f\colon [a,b] \to \mathbb{R}$ Riemann-integrável. ■

Exercício 2.136 $\left(\substack{\star\\\star}\right)$**.** Seja $f\colon [a,b] \to \mathbb{R}$ função Riemann-integrável. Mostre que se $g\colon [a,b] \to \mathbb{R}$ é tal que $\{x \in [a,b] : f(x) \neq g(x)\}$ é finito, então g é Riemann-integrável e $\int_a^b f(t)\,\mathrm{d}t = \int_a^b g(t)\,\mathrm{d}t$. Dica: com $h\colon [a,b] \to \mathbb{R}$ dada por $h(x) \coloneqq 0$ se $x \neq c \in [a,b]$ e $h(c) \coloneqq 1$, mostre $\int_a^b h(t)\,\mathrm{d}t = 0$. ■

Exercício 2.137 $\left(\substack{\star\\\star}\right)$**.** Sejam $f, g, h\colon [a,b] \to \mathbb{R}$ funções tais que $f \leq g \leq h$. Mostre que se f e h são Riemann-integráveis e $\int_a^b f(t)\,\mathrm{d}t = \int_a^b h(t)\,\mathrm{d}t$, então g é Riemann-integrável. Qual o valor da integral de g em $[a,b]$? ■

Exercício 2.138 $\left(\substack{\star\\\star\star}\right)$**.** Mostre que nenhum intervalo aberto não vazio tem medida nula. Dica: se (a,b) tivesse medida nula, então $[a,b]$ também teria e, desse modo, seria possível cobrir $[a,b]$ com finitos intervalos não vazios da forma (a_i, b_i), digamos que com i variando de 0 a n, satisfazendo

$$\sum_{i \leq n} b_i - a_i \leq b - a.$$

Isto lhe parece possível? Observação: cuidado para não confundir *comprimento* de intervalo com "medida de conjunto", conceito que nem definimos ainda[†]. ■

Exercício 2.139 $\left(\substack{\star\\\star\star}\right)$**.** Sejam $f\colon [a,b] \to \mathbb{R}$ e $g\colon [a,b] \to \mathbb{R}$ funções limitadas tais que o conjunto $\mathcal{N} \coloneqq \{x \in [a,b] : f(x) \neq g(x)\}$ tenha medida nula.

a) Se f for Riemann-integrável, é possível assegurar que g também é Riemann-integrável? Dica: pelo tom da pergunta, você já deveria suspeitar que a resposta é não.

[†] Até agora, apenas a expressão "medida nula" tem significado no texto.

b) Mostre que se f e g são Riemann-integráveis, então $\displaystyle\int_a^b f(t)\,\mathrm{d}t = \int_a^b g(t)\,\mathrm{d}t$. Dica: primeiro, prove que se $h\colon [a,b] \to \mathbb{R}$ é função Riemann-integrável não negativa tal que o conjunto $\{x \in [a,b] : h(x) \neq 0\}$ tem medida nula, então toda partição $\mathcal{P}$ de $[a,b]$ admite uma *tag* T satisfazendo $\sum_{(\mathcal{P},T)} h = 0$, acarretando $\displaystyle\int_a^b h(t)\,\mathrm{d}t = 0$; depois, note que $f - g \leq |f-g|$ e $g - f \leq |f-g|$. ■

Exercício 2.140 $(\star)$**.** Considere $F\colon [0,1] \to \mathbb{R}$ dada por

$$F(x) := \begin{cases} 0, & \text{se } x < \frac{1}{2} \\ 1, & \text{se } x \geq \frac{1}{2} \end{cases}.$$

Mostre que F é diferenciável em todo ponto $x \neq \frac{1}{2}$, com $F'(x) = 0$. O item (III) do T.F.C. permanece válido aqui? ■

Exercício 2.141 (?!)**.** Para $f\colon [a,b] \to \mathbb{R}$ função Riemann-integrável e $c, d \in [a,b]$ com $c \leq d$, costuma-se definir

$$\int_d^c f(t)\,\mathrm{d}t := -\int_c^d f(t)\,\mathrm{d}t.$$

Por que isto faz *sentido*? Dica: $\displaystyle -\sum_{i=1}^n f(t_i)(p_i - p_{i-1}) = \sum_{i=1}^n f(t_i)(p_{i-1} - p_i)$. ■

Exercício 2.142 $(^{\star}_{\star})$**.** Para uma função Riemann-integrável $f\colon [a,b] \to \mathbb{R}$, mostre que

$$\lim_{c \to a^+} \int_c^b f(t)\,\mathrm{d}t = \int_a^b f(t)\,\mathrm{d}t.$$

Dica: na verdade, o resultado permanece válido para qualquer noção de integração satisfazendo as condições (i), ...,(vi) enunciadas no começo da Seção 2.4. ■

Exercício 2.143 $(^{\star}_{\star})$**.** Seja $f\colon [a,b] \to \mathbb{R}$ limitada. Mostre que se $f|_{[c,b]}$ é Riemann-integrável para cada $c \in (a,b]$, então f é Riemann-integrável, e sua integral é dada pela identidade no exercício anterior. ■

Exercício 2.144 $(^{\star}_{\star})$**.** Mostre que a regra

$$(f,g) \mapsto \int_a^b f(t)g(t)\,\mathrm{d}t$$

define um produto interno em $\mathcal{C}[a,b]$. ■

Exercício 2.145 $(^{\star}_{\star})$**.** O resultado anterior permanece válido ao substituir $\mathcal{C}[a,b]$ por $\mathcal{R}[a,b]$? ■

Exercício 2.146 $(\star)$**.** Exiba uma função Riemann-integrável $f\colon [a,b] \to \mathbb{R}$ satisfazendo $f \neq 0$ mas $\int_a^b f(t)\,\mathrm{d}t = 0$. ■

Exercício 2.147 $(^{\star}_{\star\star})$**.** Para $f\colon [a,b] \to \mathbb{R}$ função Riemann-integrável, mostre que

$$\int_a^b |f(t)|\,\mathrm{d}t = 0$$

se, e somente se, $\{x \in [a,b] : f(x) = 0\}$ é denso em $[a,b]$. ■

Exercício 2.148 $(\substack{\star\\\star})$**.** Seja $f\colon (a,b) \to \mathbb{R}$ limitada tal que f é Riemann-integrável em cada subintervalo fechado $[c,d] \subseteq (a,b)$. Mostre que a identidade

$$\int_a^\gamma f(t)\,\mathrm{d}t + \int_\gamma^b f(t)\,\mathrm{d}t = \int_a^\delta f(t)\,\mathrm{d}t + \int_\delta^b f(t)\,\mathrm{d}t$$

se verifica para quaisquer $\gamma, \delta \in (a,b)$. ■

Exercício 2.149 $(\substack{\star\\\star})$**.** Para $f\colon [a,b] \to \mathbb{R}$ positiva e Riemann-integrável, mostre que

$$0 \leq \int_a^x f(t)\,\mathrm{d}t \leq \int_a^b f(t)\,\mathrm{d}t$$

para todo $x \in [a,b]$. Dica: mais uma vez, isto vale para qualquer noção abstrata de integração. ■

Exercício 2.150 (Critério de Darboux — $(\substack{\star\\\star})$)**.** Seja $f\colon [a,b] \to \mathbb{R}$ uma função limitada. Mostre que é Darboux-integrável se, e somente se, para todo $\varepsilon > 0$ existe uma partição $\mathcal{P}$ de $[a,b]$ tal que $U(f,\mathcal{P}) - L(f,\mathcal{P}) < \varepsilon$. ■

Exercício 2.151 $(\substack{\star\\\star})$**.** Mostre que

$$\int_1^x \frac{1}{t}\,\mathrm{d}t = \ln x$$

para todo $x > 0$. Reobtenha as propriedades elementares do logaritmo por meio desta descrição. Aproveite para revisar (e demonstrar) todas as regras de integração que te fizeram decorar em Cálculo. Dica: confira um livro de Cálculo, mas não se esqueça da regra para *mudança de variáveis* (cf. item (a) do Exercício 2.155). ■

Exercício 2.152 (Oscilações — $(\substack{\star\\\star\star})$)**.** Defina a **oscilação** de uma função $f\colon [a,b] \to \mathbb{R}$ em $x \in [a,b]$ como

$$o_x(f) := \inf\{o_f(x-\delta, x+\delta) : \delta > 0\}$$

onde $o_f(A) := \sup_{s,t\in A} |f(s) - f(t)|$ para qualquer $A \subseteq \mathbb{R}$.

a) Prove que f é contínua em x se, e somente se, $o_x(f) = 0$.

b) Mostre que f é descontínua em x se, e somente se, $o_x(f) > 0$.

c) Mostre que o conjunto $N \subseteq [a,b]$ dos pontos de descontinuidade de f é reunião dos conjuntos N_k, onde $N_k := \{x \in [a,b] : o_x(f) > \frac{1}{2^k}\}$, para $k \in \mathbb{N}$.

d) Mostre que se $N_k \cap [c,d] \neq \emptyset$, então $\sup_{[c,d]} f - \inf_{[c,d]} f \geq \frac{1}{2^k}$.

e) Dada uma partição $\mathcal{P} := (a_0, \ldots, a_n)$ de $[a,b]$, e indicando por J a coleção dos índices i tais que $[a_{i-1}, a_i] \cap N_k \neq \emptyset$, mostre que

$$\frac{1}{2^k}\sum_{i\in J}(a_i - a_{i-i}) \leq \sum_{i\in J}\Big(\sup_{[a_{i-1},a_i]} f - \inf_{[a_{i-1},a_i]} f\Big)(a_i - a_{i-1}) \leq U(f,\mathcal{P}) - L(f,\mathcal{P}).$$

Dica: para a última desigualdade, lembre-se de que $\sup_A f \geq \inf_A f$ para todo subconjunto $A \subseteq \mathrm{dom}\,(f)$.

f) Conclua, sem usar o critério de Lebesgue, que se f é Riemann-integrável, então cada N_k tem medida nula. Dica: lembre-se de que a condição de Riemann-integrabilidade permite obter uma partição $\mathcal{P}$ como no item anterior, mas satisfazendo $|U(f,\mathcal{P}) - L(f,\mathcal{P})| < \dfrac{\varepsilon}{2^k}$ (confira a desigualdade (2.14) na página 276).

g) Compare o que se fez acima com a prova apresentada para a "ida" do Critério de Lebesgue. ■

Exercício 2.153 $(^{\star}_{\star})$**.** Mostre que se $f\colon [a,b] \to \mathbb{R}$ é Riemann-integrável e $g\colon X \to \mathbb{R}$ é função contínua, então $g \circ f\colon [a,b] \to \mathbb{R}$ é Riemann-integrável sempre que $\operatorname{im}(f) \subseteq X$. ■

Exercício 2.154 $(^{\star}_{\star\star})$**.** Mostre que o resultado anterior pode ser falso se ambas forem meramente Riemann-integráveis. Dica: defina

$$f(x) := \begin{cases} 0, & \text{se } x \notin \mathbb{Q} \\ \frac{1}{n}, & \text{se } x = \frac{m}{n} \in \mathbb{Q} \text{ com } m \text{ e } n \text{ sem divisores comuns} \end{cases},$$

e $g(x) := 1$ para todo $x \in (0,1)$, com $g(0) := 0$. ■

Observação 2.5.1. Pode acontecer de $g \circ f$ não ser Riemann-integrável mesmo com g Riemann-integrável e f contínua, mas não conheço contraexemplo simples. △

Exercício 2.155 (Mudanças de variáveis — $(^{\star}_{\star\star})$)**.** Seja $\varphi\colon [a,b] \to \mathbb{R}$ diferenciável e assuma $\operatorname{im}(\varphi) = [c,d]$, com $\varphi(a) = c$ e $\varphi(b) = d$. Considere ainda uma função Riemann-integrável $f\colon [c,d] \to \mathbb{R}$.

a) Mostre (mais uma vez) que se φ' é contínua em (a,b) e f é contínua, então

$$\int_a^b f(\varphi(t))\varphi'(t)\,\mathrm{d}t = \int_c^d f(u)\,\mathrm{d}u. \tag{2.19}$$

b) Mostre que se φ é estritamente crescente e φ' é Riemann-integrável, então a identidade (2.19) ainda é satisfeita. Dica: para mostrar que $f \circ \varphi$ é Riemann-integrável, note que φ é invertível (cf. Corolário 2.2.6) e, além disso, existe uma constante $K > 0$ tal que $|\varphi^{-1}(y) - \varphi^{-1}(z)| \leq K|y - z|$ para quaisquer $y, z \in [c,d]$ (cf. Exercícios 2.82 e 2.84) pois daí, sob tais condições, $\varphi^{-1}[\mathcal{N}]$ tem medida nula sempre que $\mathcal{N}$ tem medida nula (veja a demonstração do Corolário 2.2.6); para a identidade em si, imite o truque com o TVM usado no Exercício 2.132 e apele para W-sub-redes em seguida (cf. Observação 2.4.27 e a Figura 2.4 a seguir). ■

$$\begin{tikzcd}
\operatorname{Par}_{\mathcal{R}}[c,d] \arrow[rr, "\sum_{\bullet} f"] & & \mathbb{R} \\
\mathscr{C} \arrow[u, "h"] \arrow[urr, "\sum_{\bullet}(f\circ\varphi)\cdot\varphi'"'] & &
\end{tikzcd}$$

Figura 2.4: Para cada partição $\mathcal{P} := (p_0, \ldots, p_n)$ de $[a,b]$, escolhe-se uma *tag* $T_{\mathcal{P}} := (t_1, \ldots, t_n)$ onde $\varphi'(t_i)(p_i - p_{i-1}) = \varphi(p_i) - \varphi(p_{i-1})$ para todo i. Daí, ao considerar a família (cofinal em $\operatorname{Par}_{\mathcal{R}}[a,b]$) $\mathscr{C} := \{(\mathcal{P}, T_{\mathcal{P}}) : \mathcal{P}$ é partição de $[a,b]\}$, a função h leva uma partição de Riemann $(\mathcal{P}, T_{\mathcal{P}}) \in \mathscr{C}$ à partição $\varphi(\mathcal{P}) := (\varphi(p_0), \ldots, \varphi(p_n))$ com a *tag* $\varphi(T) := (\varphi(t_1), \ldots, \varphi(t_n))$. Note que h está bem definida pois φ é estritamente crescente, enquanto a condição de Willard se verifica (com respeito às ordens do tipo $\preceq$) pois também existe $K' > 0$ tal que $|\varphi(y) - \varphi(z)| \leq K'|y - z|$ para quaisquer $y, z \in [c,d]$. A comutatividade do diagrama decorre do modo como $T_{\mathcal{P}}$ foi definida.

Exercício 2.156 $(\star)$**.** Seja $f\colon [-a,a] \to \mathbb{R}$ uma função Riemann-integrável.

a) Mostre que se f é *ímpar*, então $\displaystyle\int_{-a}^{a} f(t)\,\mathrm{d}t = 0$. Dica: tanto neste quanto no próximo item, lembre-se do Corolário 1.3.26; alternativamente, use o item (a) do exercício anterior.

b) Mostre que se f é *par*, então $\displaystyle\int_{-a}^{a} f(t)\,\mathrm{d}t = 2\int_{0}^{a} f(t)\,\mathrm{d}t$. ■

Exercício 2.157 $(^{\star}_{\star})$**.** Mostre que $\ln(2) = \displaystyle\sum_{n=1}^{\infty} \frac{1}{n2^n}$. Dica: perceba que

$$\ln(2) = \int_0^1 \frac{1}{2-t}\,\mathrm{d}t,$$

e depois expresse a função na ~~segunda~~ integral como uma série uniformemente convergente de funções que você deveria saber integrar. ■

Exercício 2.158 $(^{\star}_{\star\star})$**.** Seja $f\colon [a,b] \to [0,+\infty)$ contínua. Mostre que

$$\lim_{n\to\infty} \sqrt[n]{\int_a^b f^n(t)\,\mathrm{d}t} = \|f\|_\infty.$$

Dica: já passou da hora de você provar que $\lim\limits_{n\to\infty} \sqrt[n]{r} = 1$ para qualquer $r > 0$, não acha? De qualquer forma, assumindo isso, perceba que para $\varepsilon > 0$ qualquer, existe $N \in \mathbb{N}$ tal que $n \geq N$ acarreta

$$(1-\varepsilon)(\|f\|_\infty - \varepsilon) \leq \sqrt[n]{\int_a^b f^n(t)\,\mathrm{d}t} \leq (1+\varepsilon)\|f\|_\infty,$$

o que por sua vez permite tornar $\left|\sqrt[n]{\int_a^b f^n(t)\,\mathrm{d}t} - \|f\|_\infty\right|$ tão pequeno quanto desejado. ■

Exercício 2.159 $(^{\star}_{\star})$**.** Seja $(f_n)_{n\in\mathbb{N}}$ uma sequência de funções Riemann-integráveis da forma $[a,b] \to \mathbb{R}$ com $f(x) := \displaystyle\sum_{n=0}^{\infty} f_n(x) \in \mathbb{R}$ para todo $x \in [a,b]$. Mostre que f é Riemann-integrável, com

$$\int_a^b f(t)\,\mathrm{d}t = \int_a^b \sum_{n=0}^{\infty} f_n(t)\,\mathrm{d}t = \sum_{n=0}^{\infty} \int_a^b f_n(t)\,\mathrm{d}t,$$

em qualquer um dos casos a seguir.

a) $\displaystyle\sum_{n=0}^{\infty} f_n$ converge uniformemente, i.e., $\displaystyle\sum_{i=0}^{n} f_i \to_{\mathrm{u}} f$.

b) f_n é contínua e tal que $f_n \geq 0$ para todo $n \in \mathbb{N}$, com f contínua. ■

Exercício 2.160 ($^{\star}_{\star\star}$)**.** Mostre que o resultado anterior falha sem as hipóteses em a) ou b). Dica: para um contraexemplo bastante interessante, admita[†] que $\displaystyle\sum_{n=1}^{\infty} \frac{\operatorname{sen}(nx)}{n} = \frac{\pi - x}{2}$ para todo $x \in (0, 2\pi)$; para algo mais modesto, pense em escadas. ■

Exercício 2.161 ($^{\star}_{\star}$)**.** Como integrar séries de potências? ■

Exercício 2.162 (Integração por partes — ($^{\star}_{\star\star}$))**.** Para $f, g\colon [a, b] \to \mathbb{R}$ diferenciáveis, suponha que f' e g' sejam Riemann-integráveis. Mostre que

$$\int_a^b f(t)g'(t)\,\mathrm{d}t = f(b)g(b) - f(a)g(a) - \int_a^b f'(t)g(t)\,\mathrm{d}t.$$

Dica: use o critério de Lebesgue para assegurar a boa definição das integrais e deixe o T.F.C. fazer o resto. ■

Exercício 2.163 ($^{\star}_{\star\star}$)**.** Enuncie e demonstre um teste de comparação para integração imprópria. Dica: imite o teste de comparação para séries. ■

Exercício 2.164 ($^{\star}_{\star}$)**.** Mostre que $\displaystyle\int_1^{+\infty} \frac{\cos(t)}{t^2}\,\mathrm{d}t$ converge absolutamente. Dica: mostre que $\displaystyle\int_1^{+\infty} \frac{1}{t^2}\,\mathrm{d}t$ converge absolutamente. ■

Exercício 2.165 ($^{\star}_{\star\star}$)**.** Mostre que $\displaystyle\int_\pi^{+\infty} \frac{\operatorname{sen}(t)}{t}\,\mathrm{d}t$ converge. Dica: use o exercício anterior para calcular $\displaystyle\int_\pi^b \operatorname{sen}(t)\, t\,\mathrm{d}t$. ■

Exercício 2.166 ($^{\star}_{\star\star}$)**.** Mostre que $\displaystyle\int_0^{+\infty} \left|\frac{\operatorname{sen}(t)}{t}\right| \mathrm{d}t = +\infty$. Dica: avalie a integral de $\left|\frac{\operatorname{sen}(x)}{x}\right|$ no intervalo $[0, 2\pi n]$. ■

Exercício 2.167 ($^{\star}_{\star}$)**.** Mostre que $\displaystyle\int_0^1 \Delta_{a,b,h}(t)\,\mathrm{d}t = \frac{(b-a)h}{2}$, onde $0 \leq a < b \leq 1$, $h \geq 0$ e

$$\Delta_{a,b,h}(x) := \begin{cases} 0, & \text{se } 0 \leq x \leq a \text{ ou } b \leq x \leq 1 \\ \dfrac{2h}{b-a}(x-a), & \text{se } a \leq x \leq \dfrac{a+b}{2} \\ \dfrac{2h}{b-a}(b-x), & \text{se } \dfrac{a+b}{2} \leq x \leq b \end{cases}$$

para cada $x \in [0, 1]$. Sugestão: faça um desenho e perceba que você calculou a área do triângulo de vértices $(a, 0)$, $\left(\frac{a+b}{2}, h\right)$ e $(b, 0)$. ■

Exercício 2.168 ($^{\star}_{\star}$)**.** Exiba uma sequência $(f_n)_n$ de funções contínuas da forma $[0, 1] \to \mathbb{R}$ satisfazendo $f_n \to_{L_1} 0$, i.e., $f_n \to 0$ no espaço $(\mathcal{C}[0, 1], \|\cdot\|_1)$ (cf. Definição 2.1.26), mas tal que $f_n \not\to_{\mathrm{p}} 0$. Dica: lembre-se de que $f_n \to_{L_1} 0$ significa $\displaystyle\lim_{n\to\infty} \int_0^1 |f_n(t)|\,\mathrm{d}t = 0$; além disso, se $a_n \to c$ e $b_n \to c$, então $\displaystyle\lim_{n\to\infty} \frac{(b_n - a_n)h}{2} = 0$ para qualquer $h \in \mathbb{R}$. ■

[†] Mas não tente provar, já que isso usa técnicas que fogem do escopo deste material.

Exercício 2.169 $(\star)$**.** A função $\mathrm{Id}\colon (\mathcal{C}[0,1], \|\cdot\|_1) \to (\mathcal{C}[0,1], \|\cdot\|_\infty)$ é contínua? ■

Exercício 2.170 $(\substack{\star\\\star})$**.** Usando a mesma notação da Definição 2.1.26 para funções Riemann-integráveis, mostre que $\|\cdot\|_1$ não é uma norma sobre $\mathcal{R}[a,b]$. ■

Exercício 2.171 $(\substack{\star\\\star})$**.** Mostre que se $(f_n)_{n\in\mathbb{N}}$ é uma sequência de funções Riemann-integráveis em $[a,b]$ tal que $f_n \to_{L_1} f$ para $f\colon [a,b] \to \mathbb{R}$ Riemann-integrável, então $\|f_n\|_1 \to \|f\|_1$. ■

Exercício 2.172 $(\substack{\star\\\star\star})$**.** Mostre que $(\mathcal{C}[0,1], \|\cdot\|_1)$ não é um espaço de Banach. Dica: encare a figura a seguir até que ela te encare de volta. ■

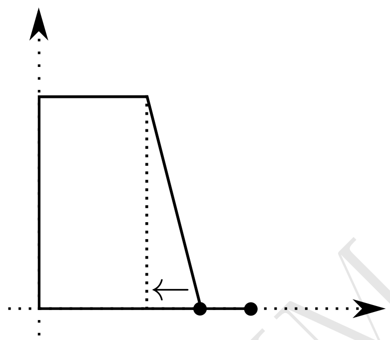


Figura 2.5: Conforme as funções contínuas se aproximam do que seria o gráfico (descontínuo) do retângulo, suas integrais se aproximam da área do retângulo.

Observação 2.5.2. Cuidado com o argumento: ao tentar solucionar o último problema, é tentador obter uma sequência $(T_n)_n$ de Cauchy em $(\mathcal{C}[0,1], \|\cdot\|_1)$ que converge pontualmente para uma função *descontínua* T e achar que isto basta, mas não é o caso. A princípio, poderia haver *outra* função contínua $S\colon [0,1] \to \mathbb{R}$ tal que $T_n \to_{L_1} S$ mas $T_n \not\to_{\mathrm{p}} S$. "*Mas o limite não era único??*". Resposta: sim, mas se a noção de convergência muda, então os limites também podem mudar. Se fosse verdade que "$T_n \to_{L_1} S \Rightarrow T_n \to_{\mathrm{p}} S$", então o argumento realmente funcionaria... mas o Exercício 2.168 já foi pensado, justamente, para aniquilar este miguezinho preguiçoso. Se precisar de uma dica sobre como proceder, confira a demonstração da Proposição 3.1.6. △

Capítulo 3

Tópicos complementares

Este capítulo trás tópicos complementares ao que se discutiu nos capítulos anteriores, principalmente nas seções extras, de modo a permitir que o material seja usado como *suporte* para outras disciplinas além de Análise I.

3.0 Derivadas em espaços euclidianos

Embora as questões de convergência e continuidade em *espaços euclidianos* (os $\mathbb{R}^n$'s da vida) se reduzam naturalmente a problemas em $\mathbb{R}$ (cf. Exercício 1.146), aumentar a dimensão dos espaços afeta outros aspectos pertinentes a Análise, mas que ainda não foram tratados com cuidado no texto.

Sugestão: Revise Álgebra Linear antes de avançar.

3.0.0 Derivadas direcionais e parciais, mas desta vez com calma

A Subseção 1.8.1 §1 discutiu muito rapidamente o problema de *derivar* funções entre espaços normados (cf. Definição 1.8.33): para X e Y espaços normados e um subconjunto aberto $S \subseteq X$, uma função $f: S \to Y$ é diferenciável em $a \in S$ se existe função $\Phi: S \to \mathcal{L}_c(X, Y)$, contínua em a, tal que $\Phi(x)(x-a) = f(x) - f(a)$ para todo $x \in S$. Neste caso, embora a *função de inclinação* Φ não seja única em geral (cf. Exercício 1.176)$^{\star}_{\star}$, a transformação linear contínua $\Phi(a): X \to Y$ é única, em virtude da identidade

$$\Phi(a)(v) = \lim_{t \to 0} \frac{f(a+tv) - f(a)}{t} := \frac{\partial f}{\partial v}(a) \tag{3.0}$$

provada na Proposição 1.8.34, que em certo sentido assegura que o "valor" de $\Phi(a)(v)$ não depende de Φ, mas sim do limite acima, que chamaremos de **derivada v-direcional de f no ponto a**. É por conta de tal unicidade que faz sentido chamar $\Phi(a)$ de *a* derivada de f em a, denotada por $f'(a)$.

Observação 3.0.0. A primeira coisa a notar, principalmente para acalmar quem já se deparou com o assunto, é a *suposição* de que a transformação linear $f'(a): X \to Y$ é contínua: no caso em que $X := \mathbb{R}^m$ para algum $m \in \mathbb{N} \setminus \{0\}$, o Corolário 2.0.29 assegura a continuidade automática da derivada pela finitude da dimensão, hipótese inexistente na Definição 1.8.33.

A segunda coisa a notar, também para quem já se deparou com diferenciabilidade neste contexto, é a *ausência* do *resto tendendo a zero*. Mas não se preocupe: ele está aqui (cf. Exercício 1.176). △

A fórmula (3.0) indica que o trabalho de calcular derivadas direcionais se resume a determinar limites em espaços normados, o que é relativamente simples quando os espaços em questão são euclidianos.

Para uma *base de Hamel*[†] $\mathcal{B}$ do espaço normado X, cada vetor $x \in X$ se escreve de forma única como combinação linear dos vetores de $\mathcal{B}$, o que permite escrever $x := (x_b)_{b\in\mathcal{B}}$, onde $x_b \in \mathbb{R}$ são os únicos escalares reais envolvidos na combinação linear

$$x = \sum_{b\in\mathcal{B}} x_b b,$$

o que só faz sentido pois $\{x_b : b \in \mathcal{B} \text{ e } x_b \neq 0\}$ é finito por hipótese. Com isso em mente, para $v \in \mathcal{B}$, temos $v = (\delta_{v,b})_{b\in\mathcal{B}}$, onde $\delta_{v,b} \in \{0,1\}$ é o *delta de Kronecker*, com $\delta_{v,b} = 1$ se, e somente se, $v = b$, resultando em

$$\frac{f(a+tv) - f(a)}{t} = \frac{1}{t}\left(f\left((a_b)_{b\in\mathcal{B}} + (t\delta_{v,b})_{b\in\mathcal{B}}\right) - f\,(a_b)_{b\in\mathcal{B}}\right). \tag{3.1}$$

Na prática, isto significa que para determinar $\dfrac{\partial f}{\partial v}(a)$, basta "derivar" a *expressão* de f na ~~coordenada~~ *direção de v* e tratar todas as outras como "constantes".

Exemplo 3.0.1. Para $f\colon \mathbb{R}^2 \to \mathbb{R}$ dada por $f(x,y) := x^2 + y^3$, com $v := (1,0)$ temos

$$\begin{aligned}\frac{\partial f}{\partial v}(a,b) &:= \lim_{t\to 0}\frac{f((a,b)+tv) - f(a,b)}{t} = \lim_{t\to 0}\frac{f(a+t,b) - f(a,b)}{t}\\ &= \lim_{t\to 0}\frac{(a+t)^2 + b^3 - (a^2+b^3)}{t} = \lim_{t\to 0}\frac{(a+t)^2 - a^2}{t} = 2a\end{aligned}$$

para qualquer $(a,b) \in \mathbb{R}^2$. Já para $F\colon \mathbb{R}^2 \to \mathbb{R}^2$ dada pela regra $F(x,y) := (f(x,y), f(y,y))$, temos

$$\begin{aligned}\frac{\partial F}{\partial v}(a,b) &:= \lim_{t\to 0}\frac{F((a,b)-tv) - F(a,b)}{t}\\ &= \lim_{t\to 0}\frac{((a+t)^2+b^3, b^2+b^3) - (a^2+b^3, b^2+b^3)}{t} = (2a, 0)\,,\end{aligned}$$

(certo?)*. ▲

Note que no primeiro caso, essencialmente consideramos a função $f_1\colon \mathbb{R} \to \mathbb{R}$ dada por $f_1(x) := f(x,b)$, em que de fato "b" é uma *constante*. Já na segunda derivada parcial, precisamos usar, além de f_1, a função $f_2\colon \mathbb{R} \to \mathbb{R}$ dada por $f_2(x) := f(b,b)$, cuja derivada é zero justamente por ser a função constante com valor $f(b,b)$. Porém, no dia a dia, ninguém tem paciência para usar o alfabeto inteiro, de modo que o mais comum é escrever

$$\frac{\partial f}{\partial v}(x,y) = 2x \quad \text{e} \quad \frac{\partial F}{\partial v}(x,y) = (2x,0)\,.$$

Exercício 3.0 ($\star$)**.** Calcule $\dfrac{\partial f}{\partial v}(x,y)$ e $\dfrac{\partial F}{\partial v}(x,y)$ para $v := (0,1)$. ■

[†]Subconjunto linearmente independente maximal de X ou, equivalentemente, um subconjunto linearmente independente de X tal que todo vetor de X se escreve como combinação linear (finita) dos vetores de $\mathcal{B}$.

De modo geral, se $f\colon S \to Y$ é diferenciável em $a \in S$, então sua derivada $f'(a)\colon X \to Y$ é uma transformação linear e, como tal, fica completamente determinada pelas imagens que assume em vetores de uma base de X fixada. Se X tem uma base finita, digamos $\{e_1, \ldots, e_n\}$, então para $v \in X$ existem únicos escalares $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ satisfazendo $v = \sum_{i \leq n} \alpha_i e_i$ e, por linearidade,

$$f'(a)(v) = \sum_{i \leq n} \alpha_i f'(a)(e_i) := \sum_{i \leq n} \alpha_i \frac{\partial f}{\partial e_i}(a), \tag{3.2}$$

onde cada $\dfrac{\partial f}{\partial e_i}(a)$ é um vetor em Y.

Exercício 3.1 ($\star$)**.** Refaça o exercício anterior com $v := (2,3)$ e compare com (3.2). ■

Agora, se Y também tem dimensão finita, digamos $Y := \mathbb{R}^m$, então para cada $j \leq m$ tem-se $\pi_j \circ f\colon S \to \mathbb{R}$, a função que a cada $s \in S$ associa a j-ésima coordenada de $f(s) := (f_1(s), \ldots, f_m(s))$. Fazendo $f_j = \pi_j \circ f$, a definição explícita da derivada e_i-direcional de f_j no ponto a diz que

$$\frac{\partial f_j}{\partial e_i}(a) := \lim_{t \to 0} \frac{f_j(a + te_i) - f_j(a)}{t},$$

secretamente, um limite de função da forma $I \to \mathbb{R}$, com $I \subseteq \mathbb{R}$ um intervalo.

De fato, como na Proposição 1.8.34, existe $r > 0$ tal que $a + te_i \in S$ sempre que $t \in (-r, r) := I$, de modo que ao definir o caminho diferenciável $\gamma\colon I \to S$ dado por $\gamma(t) := a + te_i$, resulta

$$(f_j \circ \gamma)'(0) := \lim_{t \to 0} \frac{f_j(\gamma(t)) - f_j(\gamma(0))}{t} = \lim_{t \to 0} \frac{f_j(a + te_i) - f(a)}{t} = \frac{\partial f_j}{\partial e_i}(a).$$

Em particular, pelo que sabemos de convergência em $\mathbb{R}^m$ (cf. Proposição 1.1.24, Exercício 1.146, etc.), obtemos

$$\frac{\partial f}{\partial e_i}(a) = \left(\frac{\partial f_1}{\partial e_i}(a), \ldots, \frac{\partial f_m}{\partial e_i}(a)\right)$$

e, consequentemente, chamando por $u_1, \ldots, u_m$ os vetores *da* base de Y,

$$f'(a)(v) = \sum_{i \leq n} \sum_{j \leq m} \alpha_i \frac{\partial f_j}{\partial e_i}(a) u_j. \tag{3.3}$$

Portanto, *se* f é diferenciável em a, *então* a transformação linear $f'(a)$ fica completamente determinada pela *matriz*† $J_f(a)$ composta pelas derivadas direcionais acima, tradicionalmente chamadas de **derivadas parciais** quando os vetores que determinam a *direção* pertencem às bases canônicas dos espaços envolvidos.

Observação 3.0.2 (A Jacobiana)**.** É importante se lembrar de como uma matriz $A := (a_{i,j})$ de ordem $m \times n$ e coeficientes reais determina uma transformação linear $T_A\colon \mathbb{R}^n \to \mathbb{R}^m$, pelo menos com respeito às bases canônicas de $\mathbb{R}^n$ e $\mathbb{R}^m$. Essencialmente, dado um vetor $x := (x_1, \ldots, x_n) \in \mathbb{R}^n$, define-se

$$T_A(x) := \left(\sum_{j=1}^{n} a_{1,j} x_j, \sum_{j=1}^{n} a_{2,j} x_j, \ldots, \sum_{j=1}^{n} a_{m,j} x_j\right) \in \mathbb{R}^m.$$

†Xingada de **Jacobiana** de f em a.

Matricialmente, a fórmula acima corresponde a

$$\begin{pmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{m,1} & \dots & a_{m,n} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{1,1}x_1 + \ldots + a_{1,n}x_n \\ \vdots \\ a_{m,1}x_1 + \ldots + a_{m,n}x_n \end{pmatrix}$$

a menos de transposição. Vetorialmente, T_A é a transformação linear que associa o i-ésimo vetor da base canônica de $\mathbb{R}^n$ à (transposta da) i-ésima coluna da matriz A, o que faz sentido pois A tem n colunas, e cada coluna tem m entradas.

No caso da Jacobiana $J_f(a)$ da função $f\colon S \to \mathbb{R}^m$ no ponto a, o coeficiente $a_{i,j}$ da matriz correspondente é $\dfrac{\partial f_j}{\partial e_i}(a)$, i.e., a i-ésima derivada parcial da j-ésima função coordenada, o que faz sentido pois as direções são determinadas pelo domínio de f, enquanto as funções coordenadas são determinadas por seu codomínio. Portanto, a derivada de f em a, se existir, é a transformação linear $f'(a)\colon \mathbb{R}^n \to \mathbb{R}^m$ que associa o i-ésimo vetor da base canônica de $\mathbb{R}^n$ ao vetor

$$\frac{\partial f}{\partial e_i}(a) = \left(\frac{\partial f_1}{\partial e_i}(a), \ldots, \frac{\partial f_m}{\partial e_i}(a)\right),$$

o que matricialmente corresponde a usar (as transpostas de) tais vetores como colunas de uma matriz. △

3.0.1 Critérios de diferenciabilidade

Tudo muito bom, tudo muito bem. Mas na hora de "pegar" uma função pelo pé e decidir se ela é ou não diferenciável, a definição pura não é tão prática, o que sugere procurar por critérios que permitam simplificar a tarefa. Para ilustrar com bastante calma, tratemos de um caso particular.

Exemplo 3.0.3. Vamos estudar o *comportamento* da função $f\colon \mathbb{R}^2 \to \mathbb{R}$ dada por

$$f(x,y) := \begin{cases} \dfrac{x^2 y}{x^2+y^2}, & \text{se } (x,y) \neq (0,0) \\ 0, & \text{se } (x,y) = (0,0) \end{cases}.$$

Para aquecer os motores, tratemos da continuidade[†]. Com tudo o que você já viu até aqui, a continuidade de f fora da origem *precisa* ser óbvia[‡]. A origem é um pouco mais problemática, mas não tanto: após observar que a desigualdade

$$\frac{x^2}{x^2+y^2} \leq 1$$

se verifica para quaisquer $(x,y) \in \mathbb{R}^2 \setminus \{(0,0)\}$, torna-se um mero problema de malícia (que será seu)[⋆] mostrar que para todo $\varepsilon > 0$ existe $\delta > 0$ tal que $|f(x,y)| < \varepsilon$ sempre que $\|(x,y)\| < \delta$, e a sugestão aqui é tomar $\|\cdot\|$ como a norma do máximo.

[†]Você não *precisa* verificar se uma função é contínua num ponto antes de lidar com diferenciabilidade, afinal de contas, diferenciabilidade implica em continuidade. Conforme se destacou no texto, isto foi feito aqui apenas como aquecimento.

[‡]Mas se não for, não se preocupe: basta fazer seus argumentos num rascunho, na privacidade do seu lar, sem contar para ninguém.

Com os motores aquecidos, tratemos da diferenciabilidade. Novamente, não há problemas fora da origem. De fato, como as funções $\pi_1\colon (x,y)\mapsto x$ e $\pi_2\colon (x,y)\mapsto y$ são diferenciáveis em todos os pontos de $\mathbb{R}^2$ (certo?)*, resulta que as funções $F\colon (x,y)\mapsto x^2y$ e $G\colon (x,y)\mapsto x^2+y^2$ também são diferenciáveis (cf. Exercício 1.174 e 1.175). Logo, pela Regra da Cadeia, a função f é diferenciável em todos os pontos $(x,y)\neq(0,0)$ pois, em tais situações, $x^2+y^2\neq 0$, de modo que a diferenciabilidade de $\iota\colon z\mapsto \dfrac{1}{z}$ fora de 0 assegura que $H\colon (x,y)\mapsto \dfrac{1}{x^2+y^2}$ é diferenciável fora da origem. Explicitamente, as regras de diferenciação estabelecidas nos exercícios supracitados garantem, para $(x,y)\neq(0,0)$ e $(\alpha,\beta)\in\mathbb{R}^2$ quaisquer,

$$\begin{aligned} f'(x,y)(\alpha,\beta) &= (H\cdot F)'(x,y)(\alpha,\beta) = H'(x,y)(\alpha,\beta)\cdot F(x,y) + F'(x,y)(\alpha,\beta)\cdot H(x,y)\\ &= (\iota\circ G)'(x,y)(\alpha,\beta)\cdot x^2y + \frac{1}{x^2+y^2}\cdot F'(x,y)(\alpha,\beta) \end{aligned}$$

de modo que resta entender quem "são" as transformações lineares $(\iota\circ G)'(x,y)$ e $F'(x,y)$. Vejamos:

$$F'(x,y)(\alpha,\beta) = \alpha F'(x,y)(1,0)+\beta F'(x,y)(0,1) = \alpha\frac{\partial F}{\partial e_1}(x,y)+\beta\frac{\partial F}{\partial e_2}(x,y) = 2xy\alpha + x^2\beta,$$
$$G'(x,y)(\alpha,\beta) = \alpha G'(x,y)(1,0)+\beta G'(x,y)(0,1) = \alpha\frac{\partial G}{\partial e_1}(x,y)+\beta\frac{\partial G}{\partial e_2}(x,y) = 2x\alpha + 2y\beta$$

e, pela Regra da Cadeia,

$$(\iota\circ G)'(x,y)(\alpha,\beta) = \left(\iota'(x^2+y^2)\circ G'(x,y)\right)(\alpha,\beta) = -\frac{1}{(x^2+y^2)^2}\cdot(2x\alpha+2y\beta)$$

pois $\iota'(t)(z) = -\dfrac{z}{t^2}$ para quaisquer $t\neq 0$ e $z\in\mathbb{R}$ (percebeu?)$^\star_\star$. Precisa que seja mais explícito? Pois bem: no ponto $(x,y) \coloneqq (2,3)$, por exemplo, $f'(2,3)$ é a transformação linear $\mathbb{R}^2\to\mathbb{R}$ dada pela regra

$$(\alpha,\beta)\mapsto -\frac{2\cdot 2\alpha + 2\cdot 3\beta}{(2^2+3^2)^2}\cdot 2^2\cdot 3 + \frac{1}{2^2+3^2}\cdot(2\cdot 2\cdot 3\alpha + 2^2\beta) = \frac{108\alpha - 20\beta}{169}. \tag{3.4}$$

Exercício 3.2 (!!)**.** Plote isso tudo num Geogebra da vida. ∎

Alternativamente, de modo matricial, bastaria calcular

$$\begin{aligned} \frac{\partial f}{\partial e_1}(x,y) &= \frac{2xy(x^2+y^2) - 2x\cdot x^2y}{(x^2+y^2)^2} = \dots,\\ \frac{\partial f}{\partial e_2}(x,y) &= \frac{x^2(x^2+y^2)-2y\cdot x^2y}{(x^2+y^2)^2} = \dots \end{aligned}$$

e, após fazer $(x,y)\coloneqq(2,3)$ nas expressões acima, obter a Jacobiana de f em $(2,3)$,

$$J_f(2,3) = \begin{pmatrix} \dfrac{108}{169} \\[2ex] \dfrac{-20}{169} \end{pmatrix},$$

que induz exatamente a transformação linear indicada em (3.4). Faça as contas! $(\star)$

E quanto ao ponto $(0,0)$? Seria f diferenciável em $(0,0)$? Aqui, o primeiro método não funciona justamente por culpa do zero. A abordagem matricial, por outro lado, parece bem mais simples, embora as derivadas direcionais precisem ser recalculadas por definição (por quê?)*: por valer

$$\frac{\partial f}{\partial\,(a,b)}(0,0) := \lim_{t\to 0}\frac{f(ta,tb)-f(0,0)}{t} = \lim_{t\to 0}\frac{t^3a^2b}{t^3(a^2+b^2)} = \frac{a^2b}{a^2+b^2}$$

para todo $(a,b)\neq(0,0)$, resulta em particular que

$$\frac{\partial f}{\partial e_1}(0,0) = \frac{1^2\cdot 0}{1^2+0^2} = 0 = \frac{0^2\cdot 1}{0^2+1^2} = \frac{\partial f}{\partial e_2}(0,0),$$

sugerindo que a derivada de f em $(0,0)$ seja a transformação linear nula — por ser induzida pela Jacobiana

$$J_f(0,0) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Só há um problema: f não é diferenciável em $(0,0)$. Com efeito, se $T\colon \mathbb{R}^2\to\mathbb{R}$ fosse a derivada de f em $(0,0)$, deveria valer

$$T(a,b) = \frac{\partial f}{\partial(a,b)}(0,0)$$

para qualquer $(a,b)\in\mathbb{R}^2$. E daí? Ora, já sabemos como calcular o lado direito da expressão acima, e o resultado não é uma função linear (convença-se disso!)$^{\star}_{\star}$. ▲

Entre outras coisas, o "problema" verificado no exemplo anterior revela a distinção fundamental entre a derivada $f'(a)$ e as derivadas parciais e direcionais: enquanto as últimas "aproximam" o comportamento de f ao longo da *curva* $t\mapsto a+tv$, a derivada $f'(a)\colon X\to Y$ é uma aproximação *linear* de f *em torno de* a.

De fato, por valer $\lim\limits_{x\to a}\|\Phi(x)-f'(a)\| = 0$, conclui-se que ao declarar

$$r(x) := f(x)-f(a)-f'(a)(x-a)$$

para cada $x\in S$, ocorre $\lim\limits_{x\to a}\frac{r(x)}{\|x-a\|} = 0$, uma vez que

$$|r(x)| = |\Phi(x)(x-a) - f'(a)(x-a)| \leq \|\Phi(x)-f'(a)\|\cdot\|x-a\|$$

(percebeu?)*. Grosso modo, se v estiver próximo do vetor 0, então ainda mais próximo de 0 estará o vetor $f(a+v)-f(a)-f'(a)(v)$ ou, em outras "palavras", $f(a+v)\approx f(a)+f'(a)(v)$. Graficamente, $f'(a)$ descreve um ~~plano~~ *subespaço vetorial* que, a menos de translação, é "tangente" ao gráfico da função f no ponto $(a,f(a))$.

Exercício 3.3 ($^{\star}_{\star\star}$)**.** Nas condições anteriores:

a) Mostre que $W := \{(v,f'(a)(v)) : v\in\mathbb{R}^n\}$ é subespaço vetorial de $\mathbb{R}^n\times\mathbb{R}^m$ cuja dimensão é $d := n+\dim\operatorname{im}(f'(a))$;

b) Mostre que $(a,f(a))\in W+(a,f(a))$, onde $W+(a,f(a)) := \{w+(a,f(a)) : w\in W\}$.

c) Seja $N := L+x$, onde $L\subseteq\mathbb{R}^n\times\mathbb{R}^m$ é subespaço vetorial cuja dimensão é d. Mostre que se existe uma sequência injetiva $(t_n)_n$ em $S=\operatorname{dom}(f)$ tal que $t_n\to a$ e $f(t_n)\in N$ para todo $n\in\mathbb{N}$, então $W+(a,f(a)) = N$. Compare com o item (b) do Exercício 1.170. ■

Em todo caso, embora graficamente a exigência de derivadas parciais não assegure diferenciabilidade, a situação muda se adicionarmos a hipótese de que as derivadas parciais são contínuas.

Teorema 3.0.4. *Sejam $S \subseteq \mathbb{R}^n$ um subconjunto aberto e $f\colon S \to \mathbb{R}^m$ uma função. Suponha que para cada $a \in S$, todas as derivadas parciais de f em a existam. Se, para todo $i \leq n$, a função $\dfrac{\partial f}{\partial e_i}(\cdot)\colon S \to \mathbb{R}^m$ é contínua, então f é diferenciável.*

Demonstração. Por simplicidade, vamos escrever $\partial_i f$ em vez de $\frac{\partial f}{\partial e_i}$ e, para facilitar a compreensão do argumento[†], cuidaremos primeiro do caso $n := 2$ com $m := 1$. Lembre-se de que o propósito é definir uma função $\Phi\colon S \to \mathcal{L}(\mathbb{R}^2, \mathbb{R})$, contínua em a e satisfazendo $\Phi(x)(x-a) = f(x) - f(a)$ para todo $x \in S$. A ideia é escrever

$$f(x) - f(a) = f(x_1, x_2) - f(a_1, a_2) = f(x_1, x_2) - f(a_1, x_2) + f(a_1, x_2) - f(a_1, a_2),$$

considerando implicitamente as funções $g_1(t) := f(t, x_2)$ e $g_2(t) := f(a_1, t)$, para daí apelar ao T.V.M., que assegura pontos ξ_x, ψ_x entre x_1 e a_1 e entre x_2 e a_2, respectivamente, tais que

$$\frac{g_1(x_1) - g_1(a_1)}{x_1 - a_1} = g_1'(\xi_x) \quad \text{e} \quad \frac{g_2(x_2) - g_2(a_2)}{x_2 - a_2} = g_2'(\psi_x).$$

Se conseguirmos justificar o *armengue* acima, poderemos fazer

$$\Phi(x)(\alpha, \beta) := g_1'(\xi_x)\alpha + g_2'(\psi_x)\beta, \tag{3.5}$$

para todo $x \neq a$, que satisfaz $\Phi(x)(x-a) = f(x) - f(a)$ por construção (verifique! ($\star\star$))[‡].

Vejamos... Como S é aberto e $a \in S$, existe $r > 0$ tal que $x := (x_1, x_2) \in S$ sempre que $|x_i - a_i| < r$ para $i \in \{1, 2\}$. Fixando um x dessa forma, note que $(z, x_2) \in S$ se $z \in (a_1 - r, a_1 + r)$, de modo que ao considerar $g_1\colon (a_1 - r, a_1 + r) \to \mathbb{R}$ como acima, resulta

$$g_1'(z) := \lim_{t \to 0} \frac{g_1(z+t) - g_1(z)}{t} = \lim_{t \to 0} \frac{f(z+t, x_2) - f(z, x_2)}{t} = \partial_1 f(z, x_2)$$

para qualquer $z \in (a_1 - r, a_1 + r)$, pois $\partial_1 f(p)$ existe para todo $p \in S$, por hipótese! Analogamente, a função $g_2\colon (a_2 - r, a_2 + r) \to \mathbb{R}$ também é diferenciável em todos os pontos do domínio, com $g_2'(z) = \partial_2 f(x_1, z)$. Portanto, o T.V.M. assegura a existência dos dois pontos ξ_x e ψ_x desejados.

Agora, fazendo $\Phi(x)$ como em (3.5) para $x \in Q := B(a_1, r) \times B(a_2, r) \subseteq S$ com $x \neq a$ e $\Phi(a)(\alpha, \beta) = \partial_1 f(a)\alpha + \partial_2 f(a)\beta$, mostraremos que $\Phi\colon Q \to \mathcal{L}(\mathbb{R}^2, \mathbb{R})$ é contínua em a, o que resolve o problema (cf. Exercício 3.4). É justamente nesta parte final do argumento que a hipótese sobre a continuidade das derivadas parciais será usada.

Para evitar contas desnecessárias, considere $\mathbb{R}^2$ munido da norma $\|\cdot\|_1$ (cf. Exercício 1.5), o que não acarreta perda de generalidade (cf. Teorema 2.0.27). Como as derivadas parciais são contínuas, para $\varepsilon > 0$ existe $\delta \in \mathbb{R}$ com $0 < \delta < r$ tal que $|\partial_i f(x) - \partial_i f(a)| < \frac{\varepsilon}{2}$ para todo x satisfazendo $\|x - a\|_1 < \delta$ e $i \in \{1, 2\}$ (certo?)*.

Em particular, tomando x com $\|x - a\|_1 < \frac{\delta}{2}$ garantimos $|x_i - a_i| < \frac{\delta}{2}$ para $i \in \{0, 1\}$ (por quê?)*, donde segue que $|\xi_x - a_1| < \frac{\delta}{2}$ e $|\psi_x - a_2| < \frac{\delta}{2}$ (por quê?!)** e, consequentemente,

$$\|\,(\xi_x, x_2) - (a_1, a_2)\,\|_1 < \delta \quad \text{e} \quad \|\,(x_1, \psi_x) - (a_1, a_2)\,\|_1 < \delta. \tag{3.6}$$

[†] Adaptado do texto de Hairer e Wanner [14].
[‡] Atenção: cuidado para não confundir $x, a \in \mathbb{R}^2$ com suas coordenadas reais $x_1, x_2, a_1, a_2 \in \mathbb{R}$.

Por que isto importa? Pelo seguinte: para qualquer $(\alpha, \beta) \in \mathbb{R}^2$,

$$|\Phi(x)(\alpha,\beta) - \Phi(a)(\alpha,\beta)| \leq |\partial_1 f(\xi_x, x_2) - \partial_1 f(a)||\alpha| + |\partial_2 f(x_1, \psi_x) - \partial_2 f(a)||\beta| < \varepsilon \| (\alpha,\beta) \|_1,$$

onde a última desigualdade se deve tanto a (3.6) quanto a escolha do δ, mostrando que $\|\Phi(x) - \Phi(a)\| \leq \varepsilon$ sempre que $\|x - a\|_1 < \frac{\delta}{2}$ (cf. Exercício 1.162).

O mesmo raciocínio se aplica para o caso $n \in \mathbb{N} \setminus \{0\}$ e $m := 1$, exceto que, desta vez, o "telescópio" deve ser maior[†]:

$$\begin{aligned} f(x) - f(a) = f(x_1, \ldots, x_n) &- f(x_1, \ldots, x_{n-1}, a_n) \\ &+ f(x_1, \ldots, x_{n-1}, a_n) - f(x_1, \ldots, x_{n-2}, a_{n-1}, a_n) \\ &+ \ldots + f(x_1, a_2, \ldots, a_n) - f(a_1, \ldots, a_n). \end{aligned}$$

Para o caso $m \in \mathbb{N} \setminus \{0\}$, faremos a seguinte

⌜**Afirmação.** *Para um subconjunto aberto $S \subseteq \mathbb{R}^n$, uma função $f\colon S \to \mathbb{R}^m$ é diferenciável em $a \in S$ se, e somente se, $f_j := \pi_j \circ f\colon S \to \mathbb{R}$ é diferenciável em a para cada $j \leq m$.*

Demonstração. A ida segue da Regra da Cadeia (cf. Teorema 1.8.36). Tratemos da volta. Para cada $j \leq m$ existe $\Phi_j\colon S \to \mathcal{L}(\mathbb{R}^n, \mathbb{R})$ satisfazendo $\Phi_j(x)(x - a) = f_j(x) - f_j(a)$ com Φ_j contínua em a, o que nos permite definir $\Phi\colon S \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ fazendo

$$\Phi(x)(v) := (\Phi_1(x)(v), \ldots, \Phi_m(x)(v)) \in \mathbb{R}^m$$

para todo $v \in \mathbb{R}^m$. É um exercício simples de Álgebra Linear verificar que $\Phi(x)$ é realmente linear. Quanto à identidade desejada, temos

$$\begin{aligned} \Phi(x)(x - a) &= (\Phi_1(x)(x - a), \ldots, \Phi_m(x)(x - a)) \\ &= (f_1(x) - f_1(a), \ldots, f_m(x) - f_m(a)) = f(x) - f(a). \end{aligned}$$

Finalmente, para a continuidade de Φ em a, convém considerar $\mathbb{R}^m$ com a norma $\| \cdot \|_1$, pois daí resulta

$$\|\Phi(x)(v) - \Phi(a)(v)\|_1 := \sum_{j=1}^{m} |\Phi_j(x)(v) - \Phi_j(a)(v)| \leq \sum_{j=1}^{m} \|\Phi_j(x) - \Phi_j(a)\| \|v\|$$

para qualquer $v \in \mathbb{R}^n$, e daí não é difícil usar a continuidade de cada Φ_j para mostrar que Φ é contínua. Os detalhes ficam por sua conta $(^{\star}_{\star\star})$. ⌟

De volta ao caso geral, com $m, n \in \mathbb{N} \setminus \{0\}$, para provar que $f\colon S \to \mathbb{R}^m$ é diferenciável basta mostrar que cada $f_j\colon S \to \mathbb{R}$ é diferenciável, mas isto segue do caso $m := 1$: já sabemos que $\partial_i f(x) = (\partial_i f_1(x), \ldots, \partial_i f_j(x))$ para todo x, de modo que se $\partial_i f$ é contínua para todo $i \leq n$, então também são contínuas as derivadas parciais $\partial_i f_j$ para quaisquer $i \leq n$ e $j \leq m$. Você pode cuidar do restante. □

Exercício 3.4 $(^{\star}_{\star})$**.** Para X e Y espaços normados, $a \in U \subseteq S \subseteq X$ com U e S abertos, e $f\colon S \to Y$ uma função, mostre que f é diferenciável em a se, e somente se, $f|_U$ é diferenciável em a. Dica: para a volta, será útil lembrar que para quaisquer $x, a \in X$ distintos e $w \in Y$, existe uma transformação linear $T\colon X \to Y$ tal que $T(x - a) = w$. ■

[†]E é um excelente exercício cuidar dos detalhes $(^{\star}_{\star\star})$. Dica: faça $g_i(t) := f(x_1, \ldots, x_{i-1}, t, a_{i+1}, \ldots, a_n)$.

Exemplo 3.0.5. No cansativo Exemplo 3.0.3, vimos que as derivadas parciais da função $f(x,y) := \dfrac{x^2y}{x^2+y^2}$ num ponto (x,y) fixado são dadas pelas regras a seguir:

$$\frac{\partial f}{\partial e_1}(x,y) = \begin{cases} \dfrac{2xy^3}{(x^2+y^2)^2}, & \text{se } (x,y) \neq (0,0) \\ 0, & \text{se } (x,y) = (0,0) \end{cases},$$

$$\frac{\partial f}{\partial e_2}(x,y) = \begin{cases} \dfrac{x^4 - y^2x^2}{(x^2+y^2)^2}, & \text{se } (x,y) \neq (0,0) \\ 0, & \text{se } (x,y) = (0,0) \end{cases}.$$

Segue que ao encarar tais regras como funções do tipo $\mathbb{R}^2 \to \mathbb{R}$, pelo menos uma delas é descontínua na origem, caso contrário f seria diferenciável em $(0,0)$. Será um trabalho seu verificar que ambas são descontínuas $(^{\star}_{\star\star})$. ▲

Exercício 3.5 $(\star)$**.** Calcule explicitamente a derivada da função F no Exemplo 3.0.1. Dica: não se esqueça de que $F'(a)$ é uma transformação linear de $\mathbb{R}^2$ em $\mathbb{R}^2$ para cada $a \in \mathbb{R}^2$ ou, se preferir, é uma matriz de ordem 2×2 . ■

Observação 3.0.6 (Importante)**.** O último teorema serve *apenas* como critério afirmativo, no sentido de detectar funções que <u>são</u> diferenciáveis. Uma função diferenciável *pode* ter derivadas parciais descontínuas.

Considere $f\colon \mathbb{R}^2 \to \mathbb{R}$ dada por

$$f(x,y) := \begin{cases} x^2 \operatorname{sen}\left(\dfrac{1}{y}\right), & \text{se } y \neq 0 \\ 0, & \text{caso contrário} \end{cases}.$$

Novamente, é *claro* que f é contínua em todos os pontos de seu domínio e é diferenciável nos pontos da forma $(x,y) \neq (0,0)$ (mas você irá fazer as contas para ter certeza, não é mesmo?)$^{\star}_{\star}$. Para verificar a diferenciabilidade na origem, precisamos obter uma função $\Phi\colon \mathbb{R}^2 \to \mathcal{L}_c(\mathbb{R}^2,\mathbb{R})$, contínua em $(0,0)$ e tal que $\Phi(x,y)(x,y) = f(x,y)$ para qualquer par $(x,y) \in \mathbb{R}^2$. Em particular, como as derivadas parciais de f na origem são nulas (certo?)*, segue que $\Phi(0,0)$ *precisará* ser a transformação linear nula.

Ora, para (x,y) com $y \neq 0$, vamos definir

$$\Phi(x,y)(h,k) := x \cdot \operatorname{sen}\left(\frac{1}{y}\right) \cdot h,$$

para qualquer $(h,k) \in \mathbb{R}^2$, enquanto $\Phi(x,0)\colon \mathbb{R}^2 \to \mathbb{R}$ será a transformação linear nula. Dessa forma, é claro que $\Phi(x,y)(x,y) = f(x,y)$ para todo $(x,y) \in \mathbb{R}^2$. Para verificar a continuidade de Φ na origem, podemos munir $\mathbb{R}^2$ com a norma da soma $\|\cdot\|_1$, pois assim

$$|\Phi(x,y)(h,k)| = |x| \cdot \left|\operatorname{sen}\left(\frac{1}{y}\right)\right| \cdot |h| \leq |x| \cdot |h| \leq \|(h,k)\|_1 \cdot \|(x,y)\|_1$$

sempre que $y \neq 0$, e daí não é difícil concluir que $\|\Phi(x,y)\| \leq \|(x,y)\|_1$ para *qualquer* par $(x,y) \in \mathbb{R}^2$, donde segue a continuidade.

Todavia, as derivadas parciais de f são funções da forma $\mathbb{R}^2 \to \mathbb{R}$ descontínuas na origem. Determinar explicitamente tais derivadas parciais e constatar que ambas são descontínuas serão problemas seus $(^{\star}_{\star})$. △

3.0.2 Derivação iterada

No último teorema, as hipóteses asseguraram a existência de $f'(a)$ para cada $a \in S$, o que define uma nova função[†]

$$f' \colon S \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m),$$

que a cada $x \in S$ associa a derivada de f em x. Em geral, quando $f'(a)$ existe para todo $a \in S$, chamamos a função f' de **derivada de** f.

Exercício 3.6 $\binom{\star}{\star}$**.** Nas condições do Teorema 3.0.4, mostre que $f' \colon S \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ é contínua. Conclua que f é diferenciável e tem derivada contínua se, e somente se, as derivadas parciais de f existem e são contínuas. Dica: lembre-se de que $f'(a)(v)$ é dada pela expressão em (3.3), e daí use a continuidade das derivadas parciais para controlar $\|f'(x) - f'(a)\|$, onde a última expressão é com respeito à norma de operadores. ■

Como $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ é um espaço normado, faz sentido perguntar: o que significaria dizer que a derivada f' é diferenciável em $a \in S$? A rigor, nada novo sob o sol: deve existir $\Phi \colon S \to \mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m))$ satisfazendo $\Phi(x)(x-a) = f'(x) - f'(a)$ para todo $x \in S$, com Φ contínua em a. Contudo, a natureza de $\mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m))$ merece atenção especial.

Se $B \in \mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m))$, então para cada $u \in \mathbb{R}^n$ temos $B(u) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$, i.e., $B(u) \colon \mathbb{R}^n \to \mathbb{R}^m$ é uma transformação linear que, por sua vez, pode ser aplicada a outro vetor $v \in \mathbb{R}^n$ de modo a resultar em $B(u)(v) \in \mathbb{R}^m$. Parece então razoável definir

$$\begin{aligned} \widehat{B} \colon \mathbb{R}^n \times \mathbb{R}^n &\to \mathbb{R}^m \\ (u, v) &\mapsto B(u)(v). \end{aligned}$$

A função $\widehat{B}$ é um exemplo da seguinte

Definição 3.0.7. Para espaços vetoriais X, Y e Z, uma função $\varphi \colon X \times Y \to Z$ é dita **bilinear** se, para quaisquer $\widetilde{x} \in X$ e $\widetilde{y} \in Y$ fixados, as funções

$$\begin{aligned} \varphi(\widetilde{x}, \cdot) \colon Y &\to Z \\ y &\mapsto \varphi(\widetilde{x}, y) \end{aligned} \qquad \text{e} \qquad \begin{aligned} \varphi(\cdot, \widetilde{y}) \colon X &\to Z \\ x &\mapsto \varphi(x, \widetilde{y}) \end{aligned}$$

são lineares. Vamos indicar por $\mathcal{L}_2(X \times Y, Z)$ a coleção das funções bilineares da forma $X \times Y \to Z$. ¶

Exercício 3.7 $\binom{\star}{\star}$**.** Para uma função $\varphi \colon X \times Y \to Z$, defina $\check{\varphi} \colon X \to Z^Y$ a função que a cada $x \in X$ associa $\varphi(x, \cdot) \colon Y \to Z$ como acima.

a) Mostre que a função φ é bilinear se, e somente se, $\check{\varphi}(x) \in \mathcal{L}(Y, Z)$ para todo $x \in X$ e $\check{\varphi} \in \mathcal{L}(X, \mathcal{L}(Y, Z))$.

b) Mostre que a correspondência $\varphi \mapsto \check{\varphi}$ determina uma bijeção entre as funções bilineares da forma $X \times Y \to Z$ e as funções lineares da forma $X \to \mathcal{L}(Y, Z)$.

c) Em particular, mostre que a inversa da bijeção acima faz $\psi \mapsto \widehat{\psi}$ para cada transformação linear $\psi \in \mathcal{L}(X, \mathcal{L}(Y, Z))$, onde $\widehat{\psi}(x, y) := \psi(x)(y)$.

d) Só para constar: mostre que $\mathcal{L}_2(X \times Y, Z)$ é espaço vetorial, e prove que as bijeções acima são lineares. ■

[†]Cuidado para não confundir f' com a função de inclinação Φ na definição de diferenciabilidade: ambas são da forma $S \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$, mas em geral, $f'(x)(x-a) \neq f(x) - f(a)$ (pense a respeito)$^{\star}_{\star}$.

Portanto, de um ponto de vista algébrico, podemos abandonar a exigência de Φ associar a cada $x \in S$ uma transformação linear $\Phi(x)\colon \mathbb{R}^n \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ e, em vez disso, tratar $\Phi(x)$ como uma função bilinear $\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^m$. Há, porém, a questão da continuidade de Φ, que exige uma norma em $\mathcal{L}_2(\mathbb{R}^n \times \mathbb{R}^n, \mathbb{R}^m)$. É claro que por ser um espaço vetorial com dimensão finita[†], quaisquer duas normas *nele* são topologicamente equivalentes — mas em vez de usar isso como desculpa para considerar qualquer norma preguiçosa, esta será nossa justificativa para adotar uma norma feita sob medida.

Proposição 3.0.8. *Para espaços normados X, Y e Z, seja $\varphi\colon X \times Y \to Z$ uma função bilinear. Sob tais condições, são equivalentes:*

(i) *φ é contínua em $(0,0) \in X \times Y$;*

(ii) *existe $R > 0$ tal que $\|\varphi(x,y)\| \leq R\|x\|\|y\|$ para quaisquer $x \in X$ e $y \in Y$;*

(iii) *$\check{\varphi}(x)\colon Y \to Z$ é contínua para todo x e $\check{\varphi}\colon X \to \mathcal{L}_c(Y,Z)$ é contínua;*

(iv) *φ é contínua.*

Demonstração. Para (i) $\Rightarrow$ (ii), note que por φ ser contínua, existe $\delta > 0$ de tal forma que $\|\varphi(u,v)\| < 1$ sempre que $u \in X$ e $v \in Y$ satisfazem $\|u\| < \delta$ e $\|v\| < \delta$, respectivamente (certo?)[‡], de modo que para quaisquer $x \in X$ e $y \in Y$ não nulos temos

$$\left\|\varphi\left(\frac{\delta}{2\|x\|}x, \frac{\delta}{2\|y\|}y\right)\right\| < 1 \Rightarrow \|\varphi(x,y)\| < \frac{4}{\delta^2}\|x\|\|y\|.$$

Supondo (ii), para $y \in Y$ qualquer vale $\|\check{\varphi}(x)(y)\| = \|\varphi(x,y)\| \leq (R\|x\|)\|y\|$, donde segue que $\check{\varphi}(x)\colon Y \to Z$ é contínua para cada $x \in X$ fixado. Para a segunda parte, basta mostrar que $\check{\varphi}$ é contínua em $0 \in X$: para $x_n \to 0$ e $\varepsilon > 0$, existe $N \in \mathbb{N}$ tal que $\|x_n\| < \frac{\varepsilon}{R}$ para todo $n \geq N$, acarretando

$$\|\check{\varphi}(x_n)(y)\| = \|\varphi(x_n,y)\| \leq R\|x_n\|\|y\| < \varepsilon\|y\|$$

para todo $y \in Y$ e, portanto, $\|\check{\varphi}(x_n)\| < \varepsilon$.

Assumindo (iii), para quaisquer $x \in X$ e $y \in Y$ se verifica $\|\varphi(x,y)\| \leq \|\check{\varphi}\|\|x\|\|y\|$ (por quê?)$^{\star}_{\star}$. Logo, para $(x_n, y_n) \to (x,y)$ em $X \times Y$ temos

$$\begin{aligned}\|\varphi(x_n,y_n) - \varphi(x,y)\| &= \|\varphi(x_n,y_n) - \varphi(x_n,y) + \varphi(x_n,y) - \varphi(x,y)\| \\ &\leq \|\check{\varphi}\|\|x_n\|\|y_n - y\| + \|\check{\varphi}\|\|x_n - x\|\|y\|,\end{aligned}$$

donde é fácil conluir que $\varphi(x_n,y_n) \to \varphi(x,y)$ em Z (conclua!)$^{\star}_{\star}$. Por fim, a implicação (iv) $\Rightarrow$ (i) é automática. $\square$

Assim como no caso das transformações lineares contínuas entre espaços normados, a continuidade de funções bilineares entre espaços normados também permite definir uma norma no espaço em que elas *vivem*: se $\varphi\colon X \times Y \to Z$ é uma função bilinear contínua, definimos

$$\|\varphi\| := \inf \underbrace{\{R > 0 : \|\varphi(x,y)\| \leq R\|x\|\|y\| \text{ para quaisquer } x \in X \text{ e } y \in Y\}}_{T_\varphi}.$$

[†]Qual a dimensão de $\mathcal{L}_2(\mathbb{R}^n \times \mathbb{R}^n, \mathbb{R}^m)$? $(^{\star}_{\star})$.

[‡]Estamos usando *mui implicitamente* a topologia produto em $X \times Y$.

Posto que a prova de que a definição análoga para transformações lineares determina uma norma foi deixada como exercício (cf. Exercício 1.162), é justo apresentar aqui a prova de que $\|\cdot\|$ é uma norma em $\mathcal{L}_{2,c}(X \times Y, Z)$, o espaço das funções bilineares contínuas:

✓ evidentemente $\|\varphi\| \geq 0$;

✓ como $\|\varphi(x,y)\| \leq \|\varphi\|\|x\|\|y\|$ para quaisquer $x \in X$ e $y \in Y$, segue que se $\|\varphi\| = 0$, então $\varphi = 0$;

✓ como $|\lambda| T_\varphi = T_{\lambda\varphi}$ para qualquer $\lambda \neq 0$, segue que $|\lambda|\|\varphi\| = \|\lambda\varphi\|$;

✓ por valer $\|(\varphi+\psi)(x,y)\| \leq (\|\varphi\| + \|\psi\|)(\|x\|\|y\|)$ para quaisquer $x \in X$ e $y \in Y$, resulta $\|\varphi+\psi\| \leq \|\varphi\| + \|\psi\|$.

Exercício 3.8 $({}^{\star}_{\star})$**.** Complete os detalhes omitidos. ■

Exercício 3.9 $(\star)$**.** Mostre que se X e Y são espaços normados com dimensão finita, então qualquer função bilinear $\varphi\colon X \times Y \to Z$ é contínua. Dica: $\check{\varphi}(x)$ e $\check{\varphi}$ devem ser contínuas, certo? ■

Uma vez estabelecidos todos esses conceitos, tratemos de um problema um pouco pior: se $f'\colon S \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ for diferenciável em a, como *descrever* a função bilinear $f''(a)\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^m$? Infelizmente, é difícil escapar de matrizes aqui.

Observação 3.0.9 (Notação)**.** A fim de não carregar desnecessariamente as notações a seguir, vamos manter a convenção para derivadas parciais usada na prova do Teorema 3.0.4: $\partial_i f$ em vez de $\dfrac{\partial f}{\partial e_i}$. Em tempo, fica o alerta de que é bastante comum escrever $\dfrac{\partial f}{\partial x_i}$ para fazer referência à derivada parcial na direção do i-ésimo vetor da base canônica, mesmo nas situações em que o termo "x_i" tenha outro significado ao longo do argumento. △

Secretamente, a identidade (3.3) indica que para $x \in S$, a matriz da transformação linear $f'(x)$ é

$$\begin{pmatrix} \partial_1 f_1(x) & \dots & \partial_n f_1(x) \\ \vdots & \ddots & \vdots \\ \partial_1 f_m(x) & \dots & \partial_n f_m(x) \end{pmatrix},$$

o que indica como interpretar f' como função da forma $S \to \mathbb{R}^{mn}$: enquanto as *funções coordenadas* de f eram $f_1, \dots, f_m\colon S \to \mathbb{R}$, a derivada f' tem mn funções coordenadas, a saber,

$$\begin{aligned} \partial_i f_j(\cdot)\colon S &\to \mathbb{R} \\ x &\mapsto \partial_i f_j(x) \end{aligned}$$

conforme $(i,j) \in \{1, \dots, n\} \times \{1, \dots, m\}$. Consequentemente:

Proposição 3.0.10. *Supondo $S \subseteq \mathbb{R}^n$ aberto e $f\colon S \to \mathbb{R}^m$ diferenciável, a derivada f' é diferenciável em $a \in S$ se, e somente se, $\partial_i f_j$ é diferenciável em a para quaisquer $i \leq n$ e $j \leq m$.*

Demonstração. Segue da Afirmação provada na parte final do Teorema 3.0.4. □

Exemplo 3.0.11. Para $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ dada por $f(x,y) := (x^2+xy, y^2)$, temos f diferenciável com derivada f' contínua[†] pois $\partial_1 f(x,y) = (2x+y, 0)$ e $\partial_2 f(x,y) = (x, 2y)$ existem e são contínuas[‡] para qualquer $(x,y) \in \mathbb{R}^2$ (cf. Exercício 3.6).

Ao reaplicar o raciocínio para a função f', chega-se à conclusão de que f'' também é continuamente diferenciável, já que

$$\partial_1\left(\partial_1 f\right)(x,y) = (2,0)\,,\ \partial_2\left(\partial_1 f\right)(x,y) = (1,0)\,,$$

$$\partial_1\left(\partial_2 f\right)(x,y) = (1,0) \text{ e } \partial_2\left(\partial_2 f\right)(x,y) = (2,0)$$

existem e são contínuas para todo ponto $(x,y) \in \mathbb{R}^2$. A simetria não foi coincidência. ▲

Teorema 3.0.12 (Clairut-Schwarz). *Sejam $S \subseteq \mathbb{R}^n$ um aberto e $f\colon S \to \mathbb{R}^m$ uma função diferenciável com derivada continuamente diferenciável*[††]*. Em tais condições,*

$$\frac{\partial}{\partial e_j}\left(\frac{\partial f}{\partial e_i}\right)(a) = \frac{\partial}{\partial e_i}\left(\frac{\partial f}{\partial e_j}\right)(a)$$

para todo $a \in S$.

Demonstração. Vamos começar assumindo $m := 1$, $a := 0$ e, evidentemente, $i \neq j$. Com $\alpha := \partial_i\left(\partial_j f\right)(a)$ e $\beta := \partial_j\left(\partial_i f\right)(a)$, mostraremos que $|\alpha - \beta| < \varepsilon$ (ou quase isso) para qualquer $\varepsilon > 0$ fixado. Pela continuidade das segundas derivadas parciais, existe $\delta > 0$ tal que

$$\left|\partial_j\left(\partial_i f\right)(x) - \alpha\right| < \varepsilon \quad \text{e}\left|\partial_i\left(\partial_j f\right)(x) - \beta\right| < \varepsilon$$

sempre que $\|x\| < 2\delta$. Note que podemos supor $B(0, 2\delta) \subseteq S$ (por quê?)*. O primeiro *Katzensprung* é usar o Teorema Fundamental do Cálculo! Como? Assim: pelo T.F.C., temos

$$\int_0^\delta \partial_i f(te_i + \delta e_j)\,\mathrm{d}t = f(\delta e_i + \delta e_j) - f(\delta e_j)$$

pois $g(x) := f(xe_i + \delta e_j)$ é tal que

$$\begin{aligned} g'(x) &= \lim_{t\to 0}\frac{g(t+x) - g(x)}{t} = \lim_{t\to 0}\frac{f((t+x)e_i + \delta e_j) - f(xe_i + \delta e_j)}{t} \\ &= \lim_{t\to 0}\frac{f(te_i + (xe_i + \delta e_j)) - f(xe_i + \delta e_j)}{t} := \partial_i f(xe_i + \delta e_j) \end{aligned}$$

e, analogamente,

$$f(\delta e_i) - f(0) = \int_0^\delta \partial_i f(te_i)\,\mathrm{d}t.$$

Agora, ao escrever $\Gamma := f(\delta e_i + \delta e_j) - f(\delta e_i) - f(\delta e_j) + f(0)$, resulta

$$\Gamma = \int_0^\delta \left(\partial_i f(te_i + \delta e_j) - \partial_i f(te_i)\right)\mathrm{d}t.$$

O segundo *Katzensprung* é apelar para o T.V.M.! Ao fixar $t \in (0,\delta)$ e definir a função real $h(x) := \partial_i f(te_i + x\delta e_j)$, tem-se h diferenciável no intervalo $(0,\delta)$ e, pelo T.V.M., existe $\widetilde{t} \in (0,\delta)$ tal que $h(\delta) - h(0) = \delta h'\left(\widetilde{t}\right)$, o que se traduz em

$$\partial_i f(te_i + \delta e_j) - \partial_i f(te_i) = \delta\partial_j\left(\partial_i f\right)(te_i + \widetilde{t}e_j). \tag{3.7}$$

[†] O que costuma ser chamado de "*continuamente diferenciável*".
[‡] Como funções da forma $\mathbb{R}^2 \to \mathbb{R}^2$.
[††] O que costuma ser chamado de "*duas vezes continuamente diferenciável*"...

O terceiro *Katzensprung* é conectar tudo isso! Ao considerar $\mathbb{R}^n$ com a norma da soma, tem-se $\|te_i + \tilde{t}e_j\| = |t| + |\tilde{t}| < 2\delta$, donde a forma como tomamos δ garante $\left|\partial_j\left(\partial_i f\right)(te_i + \tilde{t}e_j) - \alpha\right| < \varepsilon$. Ao multiplicar os dois lados dessa última desigualdade por δ, obtemos

$$|(\partial_i f(te_i + \delta e_j) - \partial_i f(te_i)) - \delta\alpha| = \left|\delta\partial_j\left(\partial_i f\right)(te_i + \tilde{t}e_j) - \delta\alpha\right| < \varepsilon\delta,$$

onde a primeira igualdade se deve a (3.7). Ao "integrar" as funções acima no intervalo $[0, \delta]$, a monotonicidade da integração de Riemann assegura

$$\gamma := \int_0^\delta |(\partial_i f(te_i + \delta e_j) - \partial_i f(te_i)) - \delta\alpha| \,\mathrm{d}t \leq \varepsilon\delta^2$$

e, consequentemente,

$$\left|\Gamma - \delta^2\alpha\right| = \left|\int_0^\delta (\partial_i f(te_i + \delta e_j) - \partial_i f(te_i)) \,\mathrm{d}t - \int_0^\delta \delta\alpha \,\mathrm{d}t\right| \leq \gamma.$$

Ao repetir o argumento trocando as posições de i e j, chega-se a

$$|\Gamma - \delta^2\beta| \leq \varepsilon\delta^2,$$

donde finalmente resulta

$$|\delta^2\alpha - \delta^2\beta| \leq |\delta^2\alpha - \Gamma| + |\Gamma - \delta^2\beta| \leq 2\varepsilon\delta^2$$

e, portanto, $|\alpha - \beta| \leq 2\varepsilon$, como desejado. O restante fica por sua conta. □

Exercício 3.10 ($^{\star}_{\star\star}$)**.** Complete os detalhes omitidos na demonstração anterior, observando, entre outras coisas, os domínios de definição das funções g e h. Além disso, mostre que a partir do caso $m := 1$ e $a := 0$ pode-se provar o caso geral. ■

O Teorema de Clairut-Schwarz é essencial no estudo dos pontos *pontos críticos* de funções de classe $\mathcal{C}^2$, i.e., funções continuamente diferenciáveis cujas derivadas também são continuamente diferenciáveis. Porém, isto não será abordado aqui, pois eu detesto matrizes. Ainda assim, se você quiser muito estudar isso, poderá achar o seguinte lema bastante útil:

Lema 3.0.13. *Seja $S \subseteq \mathbb{R}^n$ um aberto. Se $F\colon S \to \mathbb{R}^n$ é diferenciável em $a \in S$ e $\det F'(a) \neq 0$, então existe $\delta > 0$ tal que $F(x) \neq F(a)$ sempre que $0 < \|x - a\| < \delta$.*

Demonstração. Por hipótese, existe $\Phi\colon S \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$, função inclinação de F em a. Como a função determinante $\det\colon \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n) \to \mathbb{R}$ é contínua[†] e $\Phi(a) = F'(a)$, temos $\det \circ \Phi$ contínua em a com $\det(\Phi(a)) \neq 0$. Logo, por conservação de sinal, existe $\delta > 0$ tal que $\det(\Phi(x)) \neq 0$ para todo $x \in S$ satisfazendo $0 < \|x - a\| < \delta$. Acabou: por Álgebra Linear, sabemos que em tais situações, a transformação linear $\Phi(x)$ é bijetora e, em particular, injetora, de modo que para tais valores de x temos $x \neq a$, acarretando

$$0 \neq \Phi(x)(x - a) = F(x) - F(a),$$

como desejado. □

[†]Se você revisou Álgebra Linear, então deve se lembrar das fórmulas para expansão do determinante, e daí não é difícil se convencer de que a função determinante é contínua. Alternativamente: ao interpretar det como uma função "n-linear" $\mathbb{R}^n \times \ldots \times \mathbb{R}^n \to \mathbb{R}$, generalize o Exercício 3.9.

3.0.3 Comentários adicionais

Num universo ideal, este texto traria mais tópicos de Análise no $\mathbb{R}^n$. Porém, minha falta de paciência não é justificativa para que eu prive você de saber da existência de pelo menos dois desses tópicos.

- Embora *derivação implícita* seja um método amplamente empregado em Cálculo I, sua justificativa formal foge do escopo da Análise I, e requer o Teorema da Função Implícita/Inversa. Com isso dito, é possível provar tais resultados por meio da formulação de Carathéodory para derivadas, com roteiros bem menos desgastantes do que os tradicionais. Caso tenha se interessado, confira o artigo "Fréchet vs. Carathéodory", de Ernesto Acosta G. e Cesar Delgado G. (1994).

- Integrais de Riemann-Stieltjes podem ser formuladas por meio de redes [37], o que por sua vez simplifica muitos resultados envolvendo integrais de linha e formas diferenciáveis. Embora eu não conheça referências que discutam especificamente integrais de linha com tal abordagem, não se trata de uma tarefa de tradução árdua — inclusive, eu tenho isso anotado *em algum lugar*.

3.1 Integração à moda Lebesgue — mas sem medida

Antes que você apresente uma objeção ao título, permita-me esclarecer: *medidas* são, sem dúvida, ferramentas fundamentais e indispensáveis na formação de qualquer pessoa que pretenda se profissionalizar em Matemática. Mesmo que sua área de atuação não seja Análise (ou, neste caso específico, Probabilidade), os espaços de medida possuem uma importância comparável à dos espaços topológicos e, *eventualmente*, podem se revelar úteis na resolução de algum problema. Justamente por ser um tema de tamanha relevância, é natural que se invista tempo e esforço para estudá-lo de forma apropriada.

Contudo, zelar pelo estudo correto da Teoria da Medida não constitui impeditivo para que se aborde um de seus temas mais importantes de maneira independente: *a integral de Lebesgue*. Nesta seção, vamos introduzir uma *noção de integrabilidade* equivalente à noção de Lebesgue estudada nos cursos de Teoria da Medida, mas faremos isso sem passar pela tortuosa discussão envolvendo "álgebras" de conjuntos, construções e propriedades de medida, mensurabilidade de funções, etc. Em vez disso, usaremos a boa e velha integral de Riemann e algumas poucas ferramentas a mais de Análise Funcional[†]. A referência básica para tal tratamento é o sétimo capítulo do "*Analysis in Euclidean Space*", de Kenneth Hoffman [17], onde você poderá encontrar os desdobramentos e generalizações do pouco que será apresentado aqui[‡].

3.1.0 Como completar a integral de Riemann

Enquanto a noção de integral imprópria discutida no capítulo anterior permitiu estender a integração de Riemann para funções ilimitadas, aqui faremos algo um pouco mais modesto, mas que a longo prazo é igualmente satisfatório.

[†]Essa abordagem, embora conhecida no meio, não costuma ser empregada com frequência — talvez em virtude da importância dada às noções de medida, que permitem desenvolver a integral de Lebesgue de forma mais "ampla"

[‡]E caso queira um tratamento mais generalista, recomendo o ~~subversivo e~~ maravilhoso texto de Barry Simon, *A comprehensive course in Analysis, Part 1* [40].

Definição 3.1.0. O **suporte de uma função** $f\colon \mathbb{R} \to \mathbb{R}$, indicado por $\operatorname{supp}(f)$, é o menor fechado de $\mathbb{R}$ em que f não se anula[†]. Explicitamente: $\operatorname{supp}(f) := \overline{\{x \in \mathbb{R} : f(x) \neq 0\}}$. Diremos que f **tem suporte compacto** se... bem... o suporte de f for compacto[‡], e denotaremos por $\mathcal{C}_C(\mathbb{R})$ a coleção das funções contínuas com suporte compacto. ¶

Funções em $\mathcal{C}_C(\mathbb{R})$ lembram as funções em $\mathcal{C}[a,b]$, mas elas são distintas: ao tentar encarar uma função contínua $f\colon [a,b] \to \mathbb{R}$ como uma função $\mathbb{R} \to \mathbb{R}$ em $\mathcal{C}_C(\mathbb{R})$, precisamos *estender* o gráfico de f, o que pode ser feito de diversas formas. Por outro lado, se $g \in \mathcal{C}_C(\mathbb{R})$, então ao fazer $a_g := \min \operatorname{supp}(g)$ e $b_g := \max \operatorname{supp}(g)$, segue que $g|_{[a_g,b_g]} \in \mathcal{C}[a_g, b_g]$, um processo canônico em certo sentido, que permite estender a integral de Riemann para funções em $\mathcal{C}_C(\mathbb{R})$.

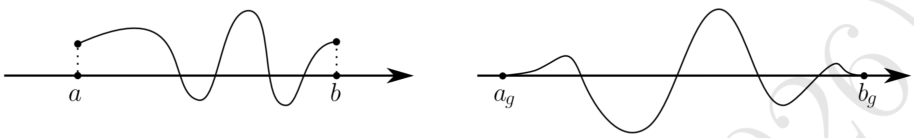


Figura 3.0: Função em $\mathcal{C}[a,b]$ vs. função em $\mathcal{C}_C(\mathbb{R})$.

Definição 3.1.1. Para $f \in \mathcal{C}_C(\mathbb{R})$, o número real

$$\int_{\mathbb{R}} f := \int_{a_f}^{b_f} f(t)\,\mathrm{d}t$$

será chamado de (a) **integral de Riemann de** f. ¶

A definição acima faz sentido pois a restrição de f ao intervalo (compacto) definido por seu suporte é contínua. Além disso, uma vez que para quaisquer $\alpha, \beta, x, y \in \mathbb{R}$ se verifica

$$\alpha x + \beta y \neq 0 \Rightarrow x \neq 0 \text{ ou } y \neq 0,$$

resulta que $\operatorname{supp}(\alpha f + \beta g) \subseteq \operatorname{supp}(f) \cup \operatorname{supp}(g)$ para quaisquer funções $f, g \in \mathcal{C}_C(\mathbb{R})$ (por quê?)*. Dado que a função nula tem suporte (vazio!) compacto, o argumento acima mostra que $\mathcal{C}_C(\mathbb{R})$ é um legítimo espaço vetorial real, por ser subespaço de $\mathcal{C}(\mathbb{R})$. O que isso tem a ver com a integral acima?

Proposição 3.1.2. *A função*

$$\begin{aligned} \int_{\mathbb{R}} (\cdot)\colon \mathcal{C}_C(\mathbb{R}) &\to \mathbb{R} \\ f &\mapsto \int_{\mathbb{R}} f \end{aligned}$$

é um ***funcional linear*** *(i.e., transformação linear com imagem contida em* $\mathbb{R}$*).*

Demonstração. Primeiro, observe que

$$\int_{\mathbb{R}} f = \int_a^b f(t)\,\mathrm{d}t$$

[†]Escreve-se "supp" em vez de "sup" não apenas para evitar conflitos com supremos, mas também por conta do inglês (*support*).

[‡]Evidentemente, a definição faz sentido para qualquer função da forma $X \to \mathbb{R}$ em que se possa dizer o que significa *ser compacto*.

para quaisquer $a, b \in \mathbb{R}$ satisfazendo $a \leq a_f$ e $b_f \leq b$ (certo?)*. Daí, para $f, g \in \mathcal{C}_C(\mathbb{R})$ e $\alpha, \beta \in \mathbb{R}$, existem $a, b \in \mathbb{R}$ com $a < b$ e $[a_f, b_f] \cup [a_g, b_g] \subseteq [a, b]$, resultando em

$$\int_{\mathbb{R}} (\alpha f + \beta g) = \int_a^b \alpha f(t) + \beta g(t)\,\mathrm{d}t = \alpha \int_a^b f(t)\,\mathrm{d}t + \beta \int_a^b g(t)\,\mathrm{d}t = \alpha \int_{\mathbb{R}} f + \beta \int_{\mathbb{R}} g,$$

como desejado. □

Como $\mathcal{C}_C(\mathbb{R})$ é subespaço vetorial de $(\mathcal{B}(X), \|\cdot\|_\infty)$, é tentador assumir que o Corolário 2.1.24 também permaneça válido para essa extensão da integral de Riemann. Porém, a vida não é tão boa:

Exemplo 3.1.3 (Importante). *A integral de Riemann*

$$\int_{\mathbb{R}} (\cdot) \colon \mathcal{C}_C(\mathbb{R}) \to \mathbb{R}$$

não é contínua com a norma do supremo. Para cada $n \in \mathbb{N} \setminus \{0\}$, considere $T_n \colon \mathbb{R} \to \mathbb{R}$ a função contínua que *parametriza* o *triângulo* de vértices $(-n, 0)$, $\left(0, \frac{1}{n}\right)$ e $(n, 0)$: tem-se $T_n \in \mathcal{C}_C(\mathbb{R})$ para todo $n > 0$, com $T_n \to_{\mathrm{u}} 0$ (certo?)*, mas

$$\int_{\mathbb{R}} T_n = \frac{2n}{2} \cdot \frac{1}{n} = 1 \not\to 0,$$

mostrando que $\int_{\mathbb{R}}$ não preserva sequências convergentes, i.e., é descontínua. ▲

Embora seja um resultado "negativo", há uma informação qualitativa importante escondida no exemplo acima: os suportes das funções T_n's se tornam grandes à medida em que n aumenta.

Exercício 3.11 ($\substack{\star\\\star}$)**.** Considere uma sequência $(f_n)_n$ de funções em $\mathcal{C}_C(\mathbb{R})$, e suponha que exista $K \subseteq \mathbb{R}$ compacto com $\operatorname{supp}(f_n) \subseteq K$ para todo n. Em tais condições, mostre que se $f_n \to_{\mathrm{u}} f$, então $\int_{\mathbb{R}} f_n \to \int_{\mathbb{R}} f$. ■

O exercício acima sugere como definir uma topologia em $\mathcal{C}_C(\mathbb{R})$ segundo a qual a integral de Riemann seja contínua. Há, porém, um caminho mais curto e concreto:

Proposição 3.1.4. *A função $\|\cdot\|_1 \colon \mathcal{C}_C(\mathbb{R}) \to \mathbb{R}$, que a cada $f \in \mathcal{C}_C(\mathbb{R})$ associa o número*

$$\|f\|_1 := \int_{\mathbb{R}} |f|,$$

*é uma norma em $\mathcal{C}_C(\mathbb{R})$, chamada de **norma L_1**.*

Demonstração. Evidente?[†] □

Exercício 3.12 ($\star$)**.** Demonstre a proposição anterior. ■

Proposição 3.1.5. *A função*

$$\int_{\mathbb{R}} (\cdot) \colon\ (\mathcal{C}_C(\mathbb{R}), \|\cdot\|_1) \to \mathbb{R}$$

é contínua.

[†]Perguntar não é afirmar. ;)

Demonstração. Para $f \in \mathcal{C}_C(\mathbb{R})$ qualquer temos $-|f| \leq f \leq |f|$ e daí

$$\left|\int_{\mathbb{R}}(f)\right| \leq \int_{\mathbb{R}} |f| := \|f\|_1,$$

(certo?)* donde o resultado segue do item (c) na Proposição 1.6.17, com $r := 1$. □

Bacana! Encontramos uma norma em $\mathcal{C}_C(\mathbb{R})$ que permite tratar a integral de Riemann como uma função contínua. Para encerrar o dia de bom humor, restaria apenas provar que tal norma é completa, não é mesmo?

Então...

Proposição 3.1.6. *O espaço $(\mathcal{C}_C(\mathbb{R}), \|\cdot\|_1)$ não é de Banach.*

Demonstração. Para cada $n \in \mathbb{N}$, considere $T_n \in \mathcal{C}_C(\mathbb{R})$ uma função cujo gráfico seja um *trapézio* de vértices $\left(-\frac{1}{2^n}, 0\right)$, $(0,1)$, $(1,1)$ e $\left(1+\frac{1}{2^n}, 0\right)$: para $x < -\frac{1}{2^n}$ ou $x > 1 + \frac{1}{2^n}$ coloca-se $T_n(x) := 0$, enquanto que para os demais casos define-se $T_n(x)$ por partes; os detalhes ficam por sua conta[†]. Se suas parametrizações estiverem todas certas, ou se você lembrar da fórmula para calcular a área de um trapézio, você vai perceber $(^\star_\star)$ que

$$\int_{\mathbb{R}} T_n = \left(2 + \frac{1}{2^{n-1}}\right)\frac{1}{2} \quad \text{e} \quad \|T_m - T_n\|_1 = \frac{1}{2}\left|\frac{1}{2^{m-1}} - \frac{1}{2^{n-1}}\right|$$

para quaisquer $m, n \in \mathbb{N}$, donde segue que $(T_n)_n$ é de Cauchy em $(\mathcal{C}_C(\mathbb{R}), \|\cdot\|_1)$.

Todavia, não há $T \in \mathcal{C}_C(\mathbb{R})$ tal que $T_n \to_{\|\cdot\|_1} T$, pois uma função com tal propriedade não pode ser contínua, como a intuição geométrica já denunciava[‡]. Vamos provar isso e, para tanto, assuma que exista tal função T.

O primeiro passo é observar que $T(x) = 1$ para todo $x \in [0,1]$. De fato, temos

$$\begin{aligned}\int_0^1 |T(x) - 1|\,\mathrm{d}x &\leq \int_0^1 |T(x) - T_n(x)|\,\mathrm{d}x + \int_0^1 |T_n(x) - 1|\,\mathrm{d}x \\ &= \int_0^1 |T(x) - T_n(x)|\,\mathrm{d}x \leq \int_{\mathbb{R}} |T - T_n| = \|T - T_n\|_1 \to 0\end{aligned}$$

(certo?)$^\star_\star$, mostrando que $|T(x) - 1| = 0$ para todo $x \in [0,1]$ (cf. Exercício 2.127). Por outro lado, para qualquer $\alpha \in \mathbb{R}$ com $\alpha > 1$, existe $n \in \mathbb{N}$ com $\frac{1}{2^n} < \alpha - 1$, de modo que

$$\begin{aligned}\int_1^\alpha |T(x)|\,\mathrm{d}x &\leq \int_1^\alpha |T(x) - T_n(x)|\,\mathrm{d}x + \int_1^\alpha T_n(x)\,\mathrm{d}x \\ &\leq \|T - T_n\|_1 + \int_1^{1+\frac{1}{2^n}} T_n(x)\,\mathrm{d}x = \|T - T_n\|_1 + \frac{1}{2^{n-1}} \to 0\end{aligned}$$

mostrando que $|T(x)| = 0$ para todo $x \in [1, \alpha]$. Absurdo (certo?)*. □

Exercício 3.13 $(^\star_\star)$**.** Com as notações da demonstração anterior, mostre que também se verifica $T(x) = 0$ para todo $x \in (-\infty, 0]$. ■

[†]Se precisar de inspiração, confira o Exercício 2.167.

[‡]Conforme n aumenta, os "trapézios" ficam cada vez mais próximos de um "quadrado", que não pode ser *parametrizado* por uma função contínua em $\mathcal{C}_C(\mathbb{R})$.

É chegada a hora de atacar o problema da *completude*: como *corrigir* um espaço normado que não é de Banach?

Quando X é um espaço métrico *incompleto*, isto é, no qual existem sequências de Cauchy divergentes, entende-se por *completamento de* X qualquer espaço métrico Y completo que contenha (uma cópia de) X e de tal forma que X seja denso em Y. A densidade aqui serve a *dois* propósitos: um *filosófico*, pois a densidade assegura que quaisquer dois completamentos são *idênticos*[†], e outro *prático*, pois a densidade assegura que funções uniformemente contínuas definidas em X e com imagens em outro espaço métrico completo admitem (única!) extensão contínua.

Por sua vez, há essencialmente duas formas clássicas de *completar* espaços métricos: o primeiro consiste em considerar a família das sequências de Cauchy do espaço *quocientadas* por uma relação de equivalência *marota*, para então definir uma métrica completa no conjunto das classes de equivalência; o segundo, menos abstrato, consiste em encontrar um espaço métrico completo que contenha uma cópia do espaço que se deseja completar — e daí o *completamento* será o fecho desta cópia no espaço maior (cf. Exercício 1.257). O subcaso dos espaços normados, que nos interessa, tem o agravante de que precisamos estender a *norma* (e não meramente uma métrica) para um espaço vetorial maior. A fim de diminuir o uso de relações de equivalência ~~e não forçar as barras~~[‡] e, ao mesmo tempo, apresentar mais técnicas de Análise Funcional, vamos seguir com o segundo método.

Proposição 3.1.7. *Para X e Y espaços normados, considere $\mathcal{L}_c(X,Y)$ o espaço vetorial das transformações lineares contínuas da forma $X \to Y$. Em tais condições, se Y é espaço de Banach, então $\mathcal{L}_c(X,Y)$ é espaço de Banach com a norma de operadores (cf. Exercício 1.162).*

Demonstração. A prova é quase simples: se $(T_n)_{n\in\mathbb{N}}$ é sequência de Cauchy em $\mathcal{L}_c(X,Y)$, então $(T_n(x))_{n\in\mathbb{N}}$ é de Cauchy em Y para todo $x \in X$, o que permite definir

$$T(x) \coloneqq \lim_{n\to\infty} T_n(x).$$

A linearidade de T segue então das propriedades operatórias dos limites em espaços normados, enquanto a continuidade resulta da continuidade das normas e da monotonicidade dos limites em $\mathbb{R}$: pelo modo como se definiu T, tem-se

$$\|T(x)\| = \left\|\lim_{n\to\infty} T_n(x)\right\| = \lim_{n\to\infty} \|T_n(x)\| \leq \lim_{n\to\infty} \|T_n\|\|x\| = \|x\| \cdot \lim_{n\to\infty} \|T_n\|.$$

Os detalhes serão problema seu. □

Exercício 3.14 ($^{\star}_{\star}$). Complete os detalhes omitidos. Em particular, na demonstração acima, por que é lícito considerar $\lim_{n\to\infty} \|T_n\|$? Dica: $|\|u\| - \|v\|| \leq \|u - v\|$ para quaisquer u e v num espaço normado. ■

Corolário 3.1.8. *O espaço $\mathcal{L}_c(X,\mathbb{R})$ é de Banach para qualquer espaço normado X.*

[†] Ou isomorfos na *categoria dos espaços métricos*. Explicitamente, se (Y,ρ) e (Y',ρ') são completamentos de (X,d), então as inclusões $i\colon X \to Y$ e $i'\colon X \to Y'$ admitem únicas extensões contínuas $I\colon Y' \to Y$ e $I'\colon Y \to Y'$ (cf. Teorema 2.1.3), que devem ser inversas uma da outra (pois $I \circ I'$ e $\mathrm{Id}_{Y'}$ coincidem no denso X, enquanto $I' \circ I$ e Id_Y coincidem no denso X); além disso, para $a, b \in Y$, existem sequências $(a_n)_n$ e $(b_n)_n$ em X com $a_n \to a$, $b_n \to b$ e $\rho(a,b) = \rho(\lim_n a_n, \lim_n b_n) = \lim_n \rho(a_n, b_n) = \lim_n d(a_n, b_n)$, donde segue que $\rho'(I'(a), I'(b)) = \rho(a,b)$. Os detalhes ficam por sua conta ($_{\star\star}^{\star}$).

[‡] Foi uma piada.

Para simplificar um pouco mais as notações, vamos omitir o corpo $\mathbb{R}$ e escrever apenas $\mathcal{L}_c(X)$ para indicar o **dual topológico**† de X com a norma de operadores. Por sua vez, o **bidual topológico** de X será indicado por $\mathcal{L}_c(\mathcal{L}_c(X))$. Sua introdução aqui se justifica pelo seguinte

Lema 3.1.9. *Para cada $x \in X$, a função*

$$\begin{aligned} \widehat{x}\colon \mathcal{L}_c(X) &\to \mathbb{R} \\ \varphi &\mapsto \varphi(x) \end{aligned}$$

é um funcional linear contínuo, i.e., $\widehat{x} \in \mathcal{L}_c(\mathcal{L}_c(X))$. Além disso, verifica-se $\|\widehat{x}\| \leq \|x\|$.

Demonstração. Para $\varphi, \psi \in \mathcal{L}_c(X)$ e $\alpha, \beta \in \mathbb{R}$ temos

$$\widehat{x}(\alpha\varphi + \beta\psi) \coloneqq (\alpha\varphi + \beta\psi)(x) = \alpha\varphi(x) + \beta\psi(x) = \alpha\widehat{x}(\varphi) + \beta\widehat{x}(\psi),$$

mostrando que $\widehat{x}$ é linear. Para a continuidade, observe que

$$|\widehat{x}(\varphi)| = |\varphi(x)| \leq \|\varphi\|\|x\|,$$

donde segue que $\widehat{x}$ é contínua (cf. item (iii) da Proposição 1.6.17, com $r \coloneqq \|x\|$). □

Exercício 3.15 ($^{\star}_{\star}$)**.** Mostre que a função $\widehat{(\cdot)}\colon X \to \mathcal{L}_c(\mathcal{L}_c(X))$, que a cada $x \in X$ associa $\widehat{x} \in \mathcal{L}_c(\mathcal{L}_c(X))$, é um funcional linear contínuo. Observação: cuidado para não se confundir! $\widehat{x}$ é a função que a cada $\varphi \in \mathcal{L}_c(X)$ associa o número $\varphi(x)$, enquanto $\widehat{(\cdot)}$ é a função que a cada x associa a função $\widehat{x}$! ■

Se mostrarmos que $\|\widehat{x}\| = \|x\|$, *teremos botado as mãos* num espaço de Banach (o bidual $\mathcal{L}_c(\mathcal{L}_c(X))$) que contém uma cópia de X (a família $\widehat{X} \coloneqq \{\widehat{x} : x \in X\} \subseteq \mathcal{L}_c(\mathcal{L}_c(X))$), o que nos permitirá obter o completamento de X como o fecho de $\widehat{X}$ no bidual, um legítimo espaço de Banach em que X é denso por definição — e aqui, se sua atenção estiver afiada, você certamente previu que o próximo exercício seria

Exercício 3.16 ($^{\star}_{\star}$)**.** Mostre que se S é subespaço vetorial de um espaço normado E, então $\overline{S}$ é subespaço vetorial de E. Dica: propriedades operatórias de limites em espaços normados. ■

Nos aproximamos de um resultado monumental em Análise Funcional.

Lema 3.1.10. *Sejam X um espaço normado, $S \subseteq X$ um subespaço vetorial de X com $S \neq X$, $R > 0$ uma constante e $\varphi\colon S \to \mathbb{R}$ um funcional linear. Se $|\varphi(s)| \leq R\|s\|$ para todo $s \in S$, então para $z \in X \setminus S$ existe um funcional linear $\Phi\colon S \oplus \langle z \rangle \to \mathbb{R}$ tal que $\Phi(s) = \varphi(s)$ para todo $s \in S$ e $|\Phi(x)| \leq R\|x\|$ para todo $x \in S \oplus \langle z \rangle$.*

Demonstração. Por definição, os elementos de $S \oplus \langle z \rangle$ se escrevem de forma única como uma soma da forma $s + \lambda z$, para $\lambda \in \mathbb{R}$. Assim, por linearidade, busca-se uma constante *apropriada* $c_z \in \mathbb{R}$ a fim de definir $\Phi(s + \lambda z) \coloneqq \varphi(s) + \lambda c_z$. Apropriada em que sentido? Neste: deve-se ter

$$-\varphi(s) - R\|s + \lambda z\| \leq \lambda c_z \leq R\|s + \lambda z\| - \varphi(s)$$

†Como contraponto ao *dual algébrico* de X, cujos elementos são *todos* os funcionais lineares de X em $\mathbb{R}$, sejam eles contínuous ou não. É comum denotar tais espaços por “$X^{\star}$”, “X^{*}”, etc.

para quaisquer $s \in S$ e $\lambda \in \mathbb{R}$. Em particular, para $\lambda > 0$, isto equivale a

$$\varphi\left(-\frac{s}{\lambda}\right) - R\left\|-\frac{s}{\lambda} - z\right\| \leq c_z \leq R\left\|\frac{s}{\lambda} + z\right\| - \varphi\left(\frac{s}{\lambda}\right),$$

o que serve de dica para determinar c_z. Com efeito, para $u, v \in S$ quaisquer, temos

$$\varphi(u) + \varphi(v) = \varphi(u + v) \leq R\|u + v\| = R\|u - z + z + v\| \leq R\|u - z\| + R\|z + v\|,$$

mostrando que $\varphi(u) - R\|u - z\| \leq R\|z + v\| - \varphi(v)$. Fixando-se $v_0 \in S$, resulta

$$c \coloneqq \sup_{u \in S} (\varphi(u) - R\|u - z\|) \leq R\|z + v_0\| - \varphi(v_0),$$

e, pela arbitrariedade do v_0 tomado, obtém-se

$$c \leq \inf_{v \in S} (R\|z + v\| - \varphi(v)) \coloneqq C.$$

Para encerrar, basta escolher $c_z \in [c, C]$. Os detalhes ficam por sua conta $(^{\star}_{\star})$. □

Teorema 3.1.11 (da extensão de Hahn-Banach)**.** *Com X, S e φ como no lema anterior, existe um funcional linear $\widetilde{\varphi}\colon X \to \mathbb{R}$, contínuo, tal que $\widetilde{\varphi}(s) = \varphi(s)$ para todo $s \in S$ e $\|\widetilde{\varphi}\| = \|\varphi\|$.*

Demonstração. No lema anterior, R assegura que φ é contínuo, o que permite substituir R por $\|\varphi\|$, de modo que o funcional Φ lá obtido passa a satisfazer $\|\Phi\| = \|\varphi\|$: a desigualdade $\|\Phi\| \leq \|\varphi\|$ é o conteúdo do lema, enquanto $\|\varphi\| \leq \|\Phi\|$ deveria ser óbvia (discorda?)*. Com isso, a ideia da prova do presente teorema é quase simples: usar o lema repetidamente, enquanto existirem vetores fora do domínio. Como o modo usual de formalizar essa ideia faz uso do Lema de Zorn, o restante da prova pode ser ignorado se você ainda não se sentir confortável com esse tipo de argumento. □

Restante da prova (*opcional*). Para cada par (N, ψ), onde $N \subseteq X$ é um subespaço vetorial e $\psi\colon N \to \mathbb{R}$ é um funcional linear, diremos que (N, ψ) é *maroto* se $S \subseteq N$, $\psi|_S = \varphi$ e $|\psi(x)| \leq \|\varphi\|\|x\|$ para todo $x \in N$. Daí, a família $\mathbb{P} \coloneqq \{(N, \psi) : (N, \psi) \text{ é maroto}\}$, parcialmente ordenada pela relação $\preceq$ onde $(N_0, \psi_0) \preceq (N_1, \psi_1)$ se, e somente se, $N_0 \subseteq N_1$ e $\psi_1|_{N_0} = \psi_0$, é uma *ordem* que satisfaz as hipóteses do Lema de Zorn. De fato, note que $(S, \varphi) \in \mathbb{P}$, mostrando que $\mathbb{P} \neq \emptyset$. Além disso, se $\mathcal{C} \subseteq \mathbb{P}$ *for* uma cadeia, então $\bigcup \mathcal{C} \in \mathbb{P}$ limita superiormente todos os pares $(N, \psi) \in \mathcal{C}$. Logo, o Lema de Zorn garante um par $(P, \Phi) \in \mathbb{P}$ maximal com respeito à relação $\preceq$. Note que se mostrarmos que $P = X$, seguirá que Φ é contínuo e tal que $\|\Phi\| = \|\varphi\|$ (por quê?)$^{\star}_{\star}$. Tratemos, então, da igualdade restante: ora, se não fosse este o caso, então o lema anterior daria $(P', \Phi_c) \in \mathbb{P}$ com $(P, \Phi) \prec (P', \Phi_c)$, contrariando a maximalidade do par maroto (P, Φ). □

Corolário 3.1.12. *Com as notações do Lema 3.1.9, vale $\|\widehat{x}\| = \|x\|$ para todo $x \in X$.*

Demonstração. Seja $S \coloneqq \langle x \rangle$ o subespaço gerado por x em X e considere o funcional linear contínuo $\varphi\colon S \to \mathbb{R}$ dado por $\varphi(\lambda x) \coloneqq \lambda\|x\|$. Note que $\|\varphi\| = 1$ (note!)$^{\star}_{\star}$. Pelo Teorema de Hahn-Banach, existe $\widetilde{\varphi} \in \mathcal{L}_c(X)$ tal que $\widetilde{\varphi}(\lambda x) = \lambda\|x\|$ para todo $\lambda \in \mathbb{R}$, com $\|\widetilde{\varphi}\| = \|\varphi\|$. Acabou:

$$\widehat{x}(\widetilde{\varphi}) \coloneqq \widetilde{\varphi}(x) = \|x\| \Rightarrow \|x\| = |\widehat{x}(\widetilde{\varphi})| \leq \|\widehat{x}\|\|\widetilde{\varphi}\| = \|\widehat{x}\|\|\varphi\| = \|\widehat{x}\|.$$ □

3.1.1 Funções Lebesgue integráveis e suas integrais

Definição 3.1.13. Denotaremos por $L^1(\mathbb{R})$ o completamento do espaço vetorial normado $(\mathcal{C}_C(\mathbb{R}), \|\cdot\|_1)$. Explicitamente, trata-se do fecho de $\widehat{\mathcal{C}_C(\mathbb{R})} := \left\{\widehat{f} : f \in \mathcal{C}_C(\mathbb{R})\right\}$ no espaço de Banach $\mathcal{L}_c(\mathcal{L}_c(\mathcal{C}_C(\mathbb{R})))$. ¶

Por definição, para $f \in \mathcal{C}_C(\mathbb{R})$, $\widehat{f} \colon \mathcal{L}_c(\mathcal{C}_C(\mathbb{R})) \to \mathbb{R}$ é a transformação linear que associa a cada funcional $\varphi \colon \mathcal{C}_C(\mathbb{R}) \to \mathbb{R}$ o número real $\widehat{f}(\varphi) := \varphi(f)$. Além disso, pelo último corolário,

$$\left\|\widehat{f}\right\| := \inf\left\{R > 0 : \left|\widehat{f}(\varphi)\right| \leq R\|\varphi\| \text{ para todo } \varphi \in \mathcal{L}_c(\mathcal{C}_C(\mathbb{R}))\right\} = \|f\|_1 := \int_{\mathbb{R}} |f|,$$

de modo que um habitante de $L^1(\mathbb{R})$ é um funcional linear contínuo $\Phi \colon \mathcal{L}_c(\mathcal{C}_C(\mathbb{R})) \to \mathbb{R}$ para o qual existe uma sequência de funções $(f_n)_n$ em $\mathcal{C}_C(\mathbb{R})$ satisfazendo

$$\|\Phi - \widehat{f_n}\| \to 0,$$

onde "$\|\cdot\|$" indica a norma de operadores em $\mathcal{L}_c(\mathcal{L}_c(\mathcal{C}_C(\mathbb{R})))$. Finalmente, como toda transformação linear contínua entre espaços normados é uniformemente contínua (cf. Exercício 2.48), o Teorema da Extensão (cf. Teorema 2.1.3 e Exercício 2.43) assegura que a integral de Riemann, enquanto funcional linear (cf. Definição 3.1.1 e Proposição 3.1.5), admite uma única extensão contínua

$$\int : L^1(\mathbb{R}) \to \mathbb{R},$$

que também é linear em virtude da densidade de $\widehat{\mathcal{C}_C(\mathbb{R})}$ em $L^1(\mathbb{R})$ (certo?)$^{\star}_{\star}$. Esta é, precisamente, a *integral de Lebesgue*!

Maravilhoso, não é mesmo?

Por um lado sim... No entanto, há um problema *óbvio* (?) em tudo o que se fez acima: intuitivamente, os elementos em $L^1(\mathbb{R})$ não *se parecem* com funções da forma $\mathbb{R} \to \mathbb{R}$ e, por conseguinte, não há qualquer pista concreta que ajude a interpretar $\displaystyle\int \Phi$ como algo que remeta a uma integral clássica, mesmo nos casos em que $\Phi \in \widehat{\mathcal{C}_C(\mathbb{R})}$.

Para resolver isso, *identificaremos* os habitantes de $L^1(\mathbb{R})$ com funções *especiais* do tipo $\mathbb{R} \to \mathbb{R}$. Explicitamente, definiremos um espaço vetorial de funções da forma $\mathbb{R} \to \mathbb{R}$, que será indicado por $\mathscr{L}^1(\mathbb{R})$, juntamente com uma sobrejeção linear $\pi \colon \mathscr{L}^1(\mathbb{R}) \to L^1(\mathbb{R})$, o que permitirá interpretar um *funcional* $\Phi \in L^1(\mathbb{R})$ como uma *classe* de funções reais em $\mathscr{L}^1(\mathbb{R})$: a saber, todas as funções $f \in \mathscr{L}^1(\mathbb{R})$ tais que $\pi(f) = \Phi$. Embora possa parecer abstrato numa primeira leitura, a ferramenta mais sofisticada que usaremos são sequências de Cauchy *especiais*.

Definição 3.1.14. Uma sequência $(x_n)_n$ num espaço normado satisfazendo

$$\sum_{n=0}^{\infty} \|x_n - x_{n+1}\| < +\infty$$

será chamada de **rápida**. ¶

Note que para $m, n \in \mathbb{N}$ com $n > m$, temos

$$\|x_m - x_n\| \leq \sum_{j=m}^{n} \|x_j - x_{j+1}\| \leq \sum_{j=m}^{\infty} \|x_j - x_{j+1}\| \to 0$$

se $m \to \infty$ (cf. Exercício 1.92, item c), mostrando assim que $(x_n)_n$ é de Cauchy (você já deve ter usado um truque parecido, por exemplo, no Exercício 1.239). Por que isso importa? Um primeiro ponto está escondido no

Lema 3.1.15. *Toda sequência de Cauchy num espaço normado admite uma subsequência rápida.*

Demonstração. Fixada uma sequência de Cauchy $(x_n)_n$ num espaço normado, vamos descrever uma subsequência rápida $(x_{n_k})_k$ recursivamente e, para isso, precisamos apenas de uma série $\sum_{n=0}^{\infty} r_n$ (absolutamente) convergente de números reais positivos para servir como *parâmetro*.

Primeiro, tomamos $n_0 \geq 0$ tal que $\|x_m - x_n\| < r_0$ sempre que $m, n \geq n_0$. Escolhemos então $n_1 > n_0$ tal que $\|x_m - x_n\| < r_1$ sempre que $m, n \geq n_1$. Procedendo desta forma recursivamente, resulta que

$$\|x_{n_{k+1}} - x_{n_k}\| \leq r_k$$

para todo k e, portanto,

$$\sum_{k=0}^{\infty} \|x_{n_{k+1}} - x_{n_k}\| \leq \sum_{k=0}^{\infty} r_k < +\infty. \qquad \square$$

Observação 3.1.16. Embora toda sequência rápida seja de Cauchy, lembre-se: *nem toda sequência de Cauchy é rápida!* De fato, fazendo $x_n = \frac{(-1)^n}{n}$ para todo $n \in \mathbb{N} \setminus \{0\}$, resulta que $(x_n)_n$ é de Cauchy por ser convergente, mas $(x_n)_n$ não é rápida (certo?)$^{\star}_{\star}$. $\triangle$

Definição 3.1.17. Diremos que $f\colon \mathbb{R} \to \mathbb{R}$ é **Lebesgue-integrável** se existirem uma sequência rápida $(f_n)_n$ de funções em $(\mathcal{C}_C(\mathbb{R}), \|\cdot\|_1)$ e um subconjunto $S \subseteq \mathbb{R}$ com medida nula (cf. Definição 2.4.15) tais que $\lim\limits_{n\to\infty} f_n(x) = f(x)$ para todo $x \in \mathbb{R} \setminus S$. O conjunto das funções Lebesgue-integráveis será indicado por $\mathscr{L}^1(\mathbb{R})$. ¶

Lema 3.1.18. *$\mathscr{L}^1(\mathbb{R})$ é subespaço vetorial de $\mathbb{R}^{\mathbb{R}}$.*

Demonstração. Claramente a função nula é habitante de $\mathscr{L}^1(\mathbb{R})$. Agora, para funções $f, g \in \mathscr{L}^1(\mathbb{R})$, existem sequências rápidas $(f_n)_n$ e $(g_n)_n$ em $\mathcal{C}_C(\mathbb{R})$ tais que $\lim\limits_{n\to\infty} f_n = f$ e $\lim\limits_{n\to\infty} g_n = g$ *a menos de conjuntos de medida nula*, ou seja: existem conjuntos $S, T \subseteq \mathbb{R}$, ambos com medida nula, tais que $\lim\limits_{n\to\infty} f_n(x) = f(x)$ e $\lim\limits_{n\to\infty} g_n(y) = g(y)$ sempre que $x \in \mathbb{R} \setminus S$ e $y \in \mathbb{R} \setminus T$, respectivamente. Consequentemente, para $\alpha \in \mathbb{R}$ qualquer e $z \in \mathbb{R} \setminus (S \cup T)$ temos

$$\lim_{n\to\infty} f_n(z) + \alpha g_n(z) = f(z) + \alpha g(z).$$

Uma vez que $S \cup T$ tem medida nula (lembra?)*, conclui-se que $f + \alpha g \in \mathscr{L}^1(\mathbb{R})$ pois $(f_n + \alpha g_n)_n$ ainda é uma sequência rápida em $\mathcal{C}_C(\mathbb{R})$, haja vista que

$$\sum_{n=0}^{\infty} \|f_n + \alpha g_n - (f_{n+1} + \alpha g_{n+1})\| \leq \sum_{n=0}^{\infty} \|f_n - f_{n+1}\| + |\alpha| \sum_{n=0}^{\infty} \|g_n - g_{n+1}\| < +\infty. \qquad \square$$

Agora, se $f \in \mathscr{L}^1(\mathbb{R})$, então *existe* uma sequência rápida $(f_n)_n$ que testemunha isso. Como $(f_n)_n$ é de Cauchy em $(\mathcal{C}_C(\mathbb{R}), \|\cdot\|_1)$, segue que $\left(\widehat{f_n}\right)_n$ é de Cauchy em $L^1(\mathbb{R})$ e, portanto, existe $\lim\limits_{n\to\infty} \widehat{f_n}$ em $L^1(\mathbb{R})$, o que permite *definir*

$$\pi(f) := \lim_{n\to\infty} \widehat{f_n}. \tag{3.8}$$

E eis onde o *problema* poderia estar: a definição de $\pi(f)$ depende da testemunha $(f_n)_n$ escolhida? Como veremos após uma longa cadeia de resultados técnicos, a resposta é um feliz e sonoro NÃO.

Definição 3.1.19. Para um subconjunto $S \subseteq \mathbb{R}$, o ínfimo

$$m^*(S) := \inf \left\{ \sum_{n=0}^{\infty} \mu(I_n) : \{I_n\}_n \text{ é coleção de intervalos com } S \subseteq \bigcup_{n\in\mathbb{N}} I_n \right\}$$

será chamado de **medida exterior** de S. ¶

Exercício 3.17 ($^{\star}_{\star}$)**.** Pratique.

a) Mostre que $m^*(S) \in [0, +\infty]$ para qualquer $S \subseteq \mathbb{R}$.

b) Mostre que $m^*(\emptyset) = 0$.

c) Mostre que S tem medida nula (cf. Definição 2.4.15) se, e somente se, $m^*(S) = 0$.

d) Mostre que $b - a \leq m^*([a, b])$ sempre que $a, b \in \mathbb{R}$ forem tais que $a \leq b$. Dica: cf. Exercício 2.138.

e) Mostre que se $S \subseteq T$, então $m^*(S) \leq m^*(T)$ e $m^*(S \cup T) \leq m^*(S) + m^*(T)$.

f) Conclua que $m^*((a, b)) = m^*([a, b]) = b - a$ e $m^*(\mathbb{R}) = +\infty$. ■

O penúltimo item no exercício anterior é um caso particular do próximo exercício, que por sua vez é a generalização natural da Proposição 2.4.19 ("reunião enumerável de conjuntos de medida nula tem medida nula"):

Exercício 3.18 ($_{\star\star}^{\star}$)**.** Mostre que se $S \subseteq \bigcup_{n\in\mathbb{N}} S_n$, então

$$m^*(S) \leq \sum_{n=0}^{\infty} m^*(S_n).$$

Dica: imite a prova da Proposição 2.4.19. ■

O que tudo isso tem a ver com a discussão sobre integrais?

Lema 3.1.20 (Chebyshev)**.** *Para $f \in \mathcal{C}_C(\mathbb{R})$ com $f \geq 0$, verifica-se*

$$\int_{\mathbb{R}} f \geq \alpha \cdot m^*(\{x \in \mathbb{R} : f(x) \geq \alpha\})$$

para qualquer $\alpha > 0$.

Demonstração. Pela definição de suporte, o subconjunto $S_\alpha := \{x \in \mathbb{R} : f(x) \geq \alpha\}$ está contido no compacto $\operatorname{supp}(f)$, o que nos permite assumir a existência de $R > 0$ tal que $S_\alpha \subseteq [-R, R]$. Agora, para uma partição $\mathcal{P} := (p_0, \ldots, p_n)$ de $[-R, R]$, obteremos uma *tag* $T_\mathcal{P}$ de P satisfazendo

$$\sum_{(\mathcal{P},T_\mathcal{P})} f \geq \alpha \cdot m^*(S_\alpha).$$

Note que ao fazer isso, teremos a desigualdade desejada no enunciado (por quê?!)[†].

A definição de $T_\mathcal{P}$ é muito simples: se o intervalo $J_i := (p_{i-1}, p_i)$ intercepta S_α, escolhemos $t_i \in J_i \cap S_\alpha$ qualquer; se $J_i \cap S_\alpha = \emptyset$, escolhemos $t_i \in J_i$ qualquer! Disso resulta

$$\sum_{(\mathcal{P},T_\mathcal{P})} f \geq \sum_{J_i \cap S_\alpha \neq \emptyset} f(t_i)\mu(J_i) \geq \alpha \sum_{J_i \cap S_\alpha \neq \emptyset} \mu(J_i) \geq \alpha \cdot m^*(S_\alpha),$$

onde a última desigualdade se deve ao fato de $\{J_i : J_i \cap S_\alpha \neq \emptyset\}$ formar uma cobertura por intervalos abertos para $S_\alpha \setminus \{p_0, \ldots, p_n\}$ (certo?)*. □

O lema acima permitirá traduzir a condição de medida nula em termos de sequências de funções com suporte compacto especiais que, por sua vez, serão fundamentais na demonstração de que a função π independe da escolha de representantes.

Lema 3.1.21. *Para $S \subseteq \mathbb{R}$, as seguintes afirmações são equivalentes:*

(i) *S tem medida nula;*

(ii) *existe uma sequência crescente $(f_n)_n$ em $\mathcal{C}_C(\mathbb{R})$ tal que $\left(\int_\mathbb{R} f_n\right)_n$ é sequência limitada em $\mathbb{R}$ com $\lim\limits_{n\to\infty} f_n(x) = +\infty$ para todo $x \in S$.*

Demonstração. Supondo S com medida nula, para cada $j \in \mathbb{N}$ existe uma cobertura $\{I_{j,n}\}_{n\in\mathbb{N}}$ para S composta por intervalos abertos de $\mathbb{R}$ satisfazendo

$$\sum_{n=0}^{\infty} \mu(I_{j,n}) < \frac{1}{2^j}.$$

Agora, para cada par (j, n) de números naturais, não é difícil se convencer de que existe uma função não negativa $f_{j,n} \in \mathcal{C}_c(\mathbb{R})$ tal que $f_{j,n}(x) \geq 1$ para qualquer x em $I_{j,n}$ e $\int_\mathbb{R} f_{j,n} \leq 2\mu(I_{j,n})$: se você ainda não se acostumou a parametrizar "trapézios", comece $(^\star_\star)$! Daí, para $n \in \mathbb{N}$ qualquer, defina

$$f_n := \sum_{i=0}^{n} \sum_{j+k=i} f_{j,k}.$$

Como cada f_n é uma soma finita de funções em $\mathcal{C}_C(\mathbb{R})$, tem-se $f_n \in \mathcal{C}_C(\mathbb{R})$, com $(f_n)_{n\in\mathbb{N}}$ crescente em virtude de todas as $f_{j,k}$ serem não negativas. Quanto à sequência das integrais, observe que

$$\int_\mathbb{R} f_n = \sum_{i=0}^{n} \sum_{j+k=i} \int_\mathbb{R} f_{j,k} \leq \sum_{i=0}^{n} \sum_{j+k=i} 2\mu(I_{j,k}) = 2\sum_{i=0}^{n} \sum_{j=0}^{n-i} \mu(I_{i,j}) \leq 2\sum_{i=0}^{n} \frac{1}{2^i} \leq 4.$$

[†]Dica: confira o Exercício 1.61 $(^\star_\star)$.

Para encerrar esta direção, note que se $x \in S$, então para cada $i \in \mathbb{N}$ existe $j_i \in \mathbb{N}$ tal que $x \in I_{i,j_i}$. Logo, para $M \in \mathbb{R}$ com $M > 0$, basta fixar $l \in \mathbb{N}$ com $l > M$ e escolher $N \in \mathbb{N}$ satisfazendo $N > \max\{i + j_i : i \leq l\}$, pois daí

$$f_n(x) \geq \sum_{i=0}^{l} f_{i,j_i}(x) = l > M$$

para qualquer $n \geq N$, mostrando que $f_n(x) \to +\infty$. Os detalhes ficam por sua conta.

Tratemos da recíproca. Nesta altura do campeonato, deve ser claro que as condições impostas em (ii) são hereditárias para subsequências e, além disso, asseguram a existência de $L_x \in [-\infty, +\infty]$ e $l \in \mathbb{R}$ tais que

$$f_n(x) \to L_x \text{ para cada } x \in \mathbb{R} \quad \text{e} \quad \int_{\mathbb{R}} f_n \to l.$$

Como a sequência das integrais é de Cauchy, o método apresentado na Proposição 3.1.15 permite extrair uma subsequência rápida, digamos $\left(\int_{\mathbb{R}} f_{n_k}\right)_k$, e escreveremos $g_k := f_{n_k}$ por simplicidade. Em particular, para $a \in (0,1)$, a subsequência $(g_n)_n$ pode ser tomada de forma a satisfazer

$$\|g_{n+1} - g_n\| := \int_{\mathbb{R}} |g_{n+1} - g_n| = \int_{\mathbb{R}} g_{n+1} - g_n < (a^n)^2, \tag{3.9}$$

(certo?)† para qualquer n. É agora que a mágica acontece.

Seja $R := \sum_{n=0}^{\infty} a^n$. Para $\varepsilon > 0$ fixado, o subconjunto

$$S_n := \left\{ x \in \mathbb{R} : g_{n+1}(x) - g_n(x) \geq \frac{Ra^n}{\varepsilon} \right\}$$

é tal que $S \subseteq \bigcup_{n \in \mathbb{N}} S_n$: de fato, se $x \notin S_n$ para todo n, então

$$\begin{aligned} g_n(x) &= g_0(x) + (g_1(x) - g_0(x)) + \ldots + (g_n(x) - g_{n-1}(x)) \\ &< g_0(x) + \frac{R}{\varepsilon}(a^0 + \ldots + a^{n-1}) \leq g_0(x) + \frac{R^2}{\varepsilon} < +\infty, \end{aligned}$$

acarretando $x \notin S$ (lembre-se da condição final na hipótese (ii)!).

Ao aliar (3.9) com a desigualdade de Chebyshev aplicada ao subconjunto S_n, obtemos

$$(a^n)^2 > \int_{\mathbb{R}} g_{n+1} - g_n \geq \frac{Ra^n}{\varepsilon} \cdot m^*(S_n) \Rightarrow m^*(S_n) \leq \frac{a^n \varepsilon}{R}$$

e, consequentemente,

$$m^*(S) \leq m^*\left(\bigcup_{n \in \mathbb{N}} S_n\right) \leq \sum_{n=0}^{\infty} m^*(S_n) \leq \sum_{n=0}^{\infty} \frac{\varepsilon a^n}{R} = \varepsilon,$$

donde segue que S tem medida nula. □

Exercício 3.19 ($^{\star}_{\star\star}$). Complete os detalhes omitidos na demonstração. ■

†Lembre-se de que na Proposição 3.1.15, bastava recorrer a uma série $\sum r_n$ absolutamente convergente. Neste caso, temos $0 < a^2 < 1$ e, portanto, $\sum (a^2)^n$ converge absolutamente.

Proposição 3.1.22. *Se $(f_n)_n$ é sequência rápida de funções em $(\mathcal{C}_C(\mathbb{R}), \|\cdot\|_1)$, então existe um conjunto de medida nula S tal que $\lim_{n\to\infty} f_n(x) \in \mathbb{R}$ para todo $x \in \mathbb{R} \setminus S$.*

Demonstração. Ao definir $g_n := \sum_{j=0}^{n} |f_{j+1} - f_j|$ para todo n, note que $(g_n)_n$ satisfaz a condição (ii) do lema anterior com respeito ao subconjunto

$$S := \left\{ x \in \mathbb{R} : \lim_{n\to\infty} g_n(x) = +\infty \right\},$$

ou, posto de outra forma: o subconjunto S definido acima tem medida nula. Ou ainda: ao definir $(g_n)_n$ da forma acima, $\lim_{n\to\infty} g_n(x)$ existe em $\mathbb{R}$ *para quase todo* x. Independentemente da forma pela qual você prefira expresar o fato anterior, o importante é notar que este conjunto S *serve*: afinal de contas, se $x \notin S$, então $(f_n(x))_n$ é uma sequência rápida de números reais e, portanto, converge em $\mathbb{R}$ (certo?)*. $\square$

A partir daqui, expressões do tipo "a propriedade $\mathcal{P}(x)$ vale quase sempre", "a propriedade $\mathcal{P}(x)$ vale a menos de um conjunto de medida nula", "a propriedade $\mathcal{P}(x)$ vale para quase todo ponto", etc. serão abreviadas como "a propriedade $\mathcal{P}(x)$ vale **q.t.p.**" (**q**uase **t**odo **p**onto), e indicam justamente a existência de um subconjunto S com medida nula tal que $\mathcal{P}(x)$ vale para qualquer $x \in \mathbb{R} \setminus S$. Em particular, uma expressão do tipo "$f = g$ q.t.p" para funções $f, g \colon \mathbb{R} \to \mathbb{R}$ indica que $f(x) = g(x)$ q.t.p. Com tais convenções, por exemplo, o enunciado da última proposição se escreve assim: "se $(f_n)_n$ é sequência rápida... então $\lim_{n\to\infty} f_n(x) \in \mathbb{R}$ q.t.p.".

A batalha se aproxima de seu ápice.

Lema 3.1.23. *Sejam $(f_n)_n$ e $(g_n)_n$ sequências crescentes de funções não negativas em $\mathcal{C}_C(\mathbb{R})$. Se $\lim_{n\to\infty} f_n \geq \lim_{n\to\infty} g_n$ q.t.p., então*

$$\lim_{n\to\infty} \int_{\mathbb{R}} f_n \geq \lim_{n\to\infty} \int_{\mathbb{R}} g_n.$$

Demonstração. Para cada $x \in \mathbb{R}$, vamos indicar por $f(x)$ e $g(x)$, respectivamente, os limites em $(-\infty, +\infty]$ das sequências $(f_n(x))_n$ e $(g_n(x))_n$, de modo que a hipótese no enunciado se traduz em dizer que o conjunto $S := \{x \in \mathbb{R} : f(x) < g(x)\}$ tem medida nula. Logo, o Lema 3.1.21 assegura uma sequência crescente de funções em $\mathcal{C}_C(\mathbb{R})$, digamos $(h_n)_n$, cujas integrais são limitadas por um número $M > 0$ e tais que $\lim_{n\to\infty} h_n(x) = +\infty$ sempre que $f(x) < g(x)$. Evidentemente, não há perda de generalidade em supor $M = 1$ (concorda?)* e, pela demonstração apresentada no lema em questão, é possível assegurar $h_n \geq 0$ para todo n.

Com isso, para $\varepsilon > 0$ fixado e escrevendo $h(x)$ para indicar o limite em $[0, +\infty]$ da sequência $(h_n(x))_n$, observe que a desigualdade $f(x) + \varepsilon h(x) \geq g(x)$ se verifica para qualquer $x \in \mathbb{R}$: se $x \notin S$, então $f(x) \geq g(x)$ e daí $f(x) + \varepsilon h(x) \geq g(x)$; se $x \in S$, então $h(x) = +\infty$ e, por conseguinte, $f(x) + \varepsilon h(x) = +\infty \geq g(x)$. Em particular, para $m \in \mathbb{N}$ fixado, $g(x) \geq g_m(x)$ para todo x, resultando em

$$f + \varepsilon h \geq g_m.$$

A ideia agora é usar o suporte compacto de g_m, que podemos assumir estar contido num intervalo $[a, b]$ de $\mathbb{R}$, juntamente com a continuidade da integral de Riemann em $\mathcal{C}[a, b] \to \mathbb{R}$ com respeito à convergência uniforme, a fim de obter a desigualdade desejada. Como fazer isso?

Para cada n, definimos a função $\varphi_n := (g_m - f_n - \varepsilon h_n)^+$, i.e., a parte positiva da função $g_m - f_n - \varepsilon h_n$ (cf. Exercício 2.124). Por conta da desigualdade anterior, para $x \in \mathbb{R}$ qualquer temos

$$\begin{aligned}\lim_{n\to\infty} \varphi_n(x) &= \lim_{n\to\infty} \max\{0, g_m(x) - f_n(x) - \varepsilon h_n(x)\} = \\ &= \max\{0, \lim_{n\to\infty} g_m(x) - f_n(x) - \varepsilon h_n(x)\} = \max\{0, g_m(x) - f(x) - \varepsilon h(x)\} = 0,\end{aligned}$$

com $(\varphi_n(x))_n$ decrescente pois tanto $(f_n)_n$ quanto $(h_n)_n$ são crescentes. Logo, o Teorema de Dini (cf. Exercício 2.109) garante que $(\varphi|_{[a,b]})_n \to_{\mathrm{u}} 0$, i.e., $\varphi_n \to 0$ uniformemente no intervalo $[a, b]$. Consequentemente, existe $N \in \mathbb{N}$ tal que

$$\varphi_n(x) \leq \frac{\varepsilon}{b-a}$$

para quaisquer $n \geq N$ e $x \in [a, b]$. Em particular,

$$f_n(x) + \varepsilon h_n(x) \geq g_m(x) - \frac{\varepsilon}{b-a}.$$

Acabou. Duvida? Veja: para $n \geq N$,

$$\begin{aligned}\left(\int_{\mathbb{R}} f_n\right) + \varepsilon &\geq \int_{\mathbb{R}} f_n + \varepsilon \int_{\mathbb{R}} h_n \geq \int_a^b f_n(t) + \varepsilon h_n(t)\,\mathrm{d}t \\ &\geq \int_a^b g_m(t) - \frac{\varepsilon}{b-a}\,\mathrm{d}t = \int_a^b g_m(t)\,\mathrm{d}t - \varepsilon = \int_{\mathbb{R}} g_m - \varepsilon,\end{aligned}$$

logo

$$\lim_{n\to\infty} \int_{\mathbb{R}} f_n + 2\varepsilon \geq \int_{\mathbb{R}} g_m,$$

donde a arbitrariedade de ε e m permitem obter o resultado desejado. □

Exercício 3.20 ($\star$). Complete os detalhes da prova anterior. Em particular: o que garante a desigualdade

$$\int_a^b f_n(t)\,\mathrm{d}t \leq \int_{\mathbb{R}} f_n,$$

implicitamente utilizada na parte final da demonstração? ■

Teorema 3.1.24. *Se $(h_n)_n$ é sequência rápida de funções não negativas em $(\mathcal{C}_C(\mathbb{R}), \|\cdot\|_1)$ tais que*

$$\lim_{n\to\infty} h_n(x) = 0 \quad q.t.p.,$$

então

$$\lim_{n\to\infty} \int_{\mathbb{R}} h_n = 0.$$

Demonstração. O limite das integrais existe pois $\left(\int_{\mathbb{R}} h_n\right)_n$ é uma sequência rápida em $\mathbb{R}$ (percebeu?)[†] e, portanto, é (de Cauchy, logo) convergente. Em particular, $0 \leq \lim_{n\to\infty} \int_{\mathbb{R}} h_n$. A fim de simplificar o que vem pela frente, vamos assumir $h_0 = 0$, o que evidentemente não afeta o resultado original.

[†]Cf. Proposição 3.1.5 ($\star$).

Agora, para $n \in \mathbb{N}$, colocamos

$$\alpha_n := (h_{n+1} - h_n)^- \quad \text{e} \quad \beta_n := (h_{n+1} - h_n)^+,$$

de modo que $\alpha_i - \beta_i = -(h_{i+1} - h_i)$, $\alpha_i + \beta_i = |h_{i+1} - h_i|$ e

$$\sum_{i=0}^{n} \alpha_i - \beta_i = -h_{n+1} \quad (\text{verifique!})^{\star}_{\star}.$$

Para encerrar (pois é!), definimos

$$f_n := \sum_{i=0}^{n} \alpha_n \quad \text{e} \quad g_n := \sum_{i=0}^{n} \beta_n,$$

que satisfazem as hipóteses do lema anterior (certo?)$^{\star}_{\star}$, o que permite concluir que

$$\lim_{n\to\infty} \int_{\mathbb{R}} f_n \geq \lim_{n\to\infty} \int_{\mathbb{R}} g_n, \tag{3.10}$$

e, consequentemente (?!),

$$0 \leq \lim_{n\to\infty} \int_{\mathbb{R}} f_n - g_n = \lim_{n\to\infty} \int_{\mathbb{R}} -h_{n+1} \Rightarrow \lim_{n\to\infty} \int_{\mathbb{R}} h_n \leq 0, \tag{3.11}$$

justamente a desigualdade que faltava. □

Exercício 3.21 ($^{\star}_{\star\star}$). Complete os detalhes da demonstração. Em particular: o que assegura a implicação (3.10) ⇒ (3.11) (haja vista que *poderia* ocorrer $\lim_{n\to\infty} \int_{\mathbb{R}} g_n = +\infty$)? Dica: convença-se de que

$$\sum_{n=0}^{\infty} \int_{\mathbb{R}} |h_n - h_{n+1}| = \lim_{n\to\infty} \int_{\mathbb{R}} f_n + \lim_{n\to\infty} \int_{\mathbb{R}} g_n,$$

o que anula a observação entre parênteses. ■

Corolário 3.1.25. *A função $\pi\colon \mathscr{L}^1(\mathbb{R}) \to L^1(\mathbb{R})$ definida em* (3.8) *(pág.* *321**) está bem definida. Mais precisamente: se $(f_n)_n$ e $(g_n)_n$ são sequências rápidas de funções em $\mathcal{C}_C(\mathbb{R})$ tais que*

$$\lim_{n\to\infty} f_n = \lim_{n\to\infty} g_n \quad q.t.p.$$

então $\lim_{n\to\infty} \widehat{f_n} = \lim_{n\to\infty} \widehat{g_n}$ em $L^1(\mathbb{R})$.

Demonstração. Fazendo $h_n := |f_n - g_n| \in \mathcal{C}_C(\mathbb{R})$ para todo n, obtemos $(h_n)_n$ rápida e $\lim_{n\to\infty} h_n = 0$ q.t.p (de acordo?)$^{\star}_{\star}$, de modo que o teorema anterior acarreta

$$\|h_n\|_1 := \int_{\mathbb{R}} h_n \to 0, \tag{3.12}$$

e isso basta para assegurar $\lim_{n\to\infty} \widehat{f_n} = \lim_{n\to\infty} \widehat{g_n}$. De fato, com $\Phi := \lim_{n\to\infty} \widehat{f_n}$ e $\Psi := \lim_{n\to\infty} \widehat{g_n}$, temos

$$\|\Phi - \Psi\| \leq \|\Phi - \widehat{f_n}\| + \|\widehat{f_n} - \widehat{g_n}\| + \|\Psi - \widehat{g_n}\| = \|\Phi - \widehat{f_n}\| + \|h_n\|_1 + \|\Psi - \widehat{g_n}\|,$$

o que permite concluir que $\|\Phi - \Psi\| < \varepsilon$ para qualquer $\varepsilon > 0$ escolhido (certo?)*. □

Exercício 3.22 ($^{\star}_{\star\star}$)**.** Prove que $\pi\colon \mathscr{L}^1(\mathbb{R}) \to L^1(\mathbb{R})$ é linear e sobrejetora. Dica: para a sobrejetividade, observe que se $\Phi \in L^1(\mathbb{R})$, então existe uma sequência $(f_n)_n$ em $\mathcal{C}_C(\mathbb{R})$ tal que

$$\lim_{n\to\infty} \widehat{f_n} = \Phi,$$

donde segue que $(f_n)_n$ é de Cauchy em $(\mathcal{C}_C(\mathbb{R}), \|\cdot\|_1)$ e, portanto, admite uma subsequência rápida; use então a Proposição 3.1.22 para cozinhar $f \in \mathscr{L}^1(\mathbb{R})$ tal que $\pi(f) = \Phi$. ∎

Observação 3.1.26 (**Opcional**: funções definidas q.t.p.)**.** A definição do espaço $\mathscr{L}^1(\mathbb{R})$ foi propositalmente simplificada: na prática, uma função de $\mathscr{L}^1(\mathbb{R})$ não *precisa* ter a reta real como domínino, basta que ela esteja definida no complementar de um subconjunto de medida nula! No entanto, *encarar* $\mathscr{L}^1(\mathbb{R})$ explicitamente desta forma traz um problema algébrico: como operar tais funções entre si?

A princípio, a resposta parece bastante clara, embora envolva definições mais complicadas para soma e produto de funções: para $f\colon S \to \mathbb{R}$ e $g\colon T \to \mathbb{R}$ funções em $\mathscr{L}^1(\mathbb{R})$, por exemplo, $f + g$ passaria a ser a função que faz $x \mapsto f(x) + g(x)$ para todo $x \in S \cap T$, com algo análogo para $f \cdot g$. Parece ótimo, mas há uma "falha": embora a função nula seja o elemento neutro aditivo, não há simétricos!

Um dos modos razoáveis de resolver o problema consiste, justamente, em considerar $f \equiv g$ sempre que $f(x) = g(x)$ q.t.p.. No quociente $\mathscr{L}^1(\mathbb{R})/\equiv$, as operações acima definem uma estrutura de espaço vetorial. Porém, ao repetir os argumentos utilizados nesta seção com apenas um pouco mais de cuidado, chega-se à conclusão de que tal quociente é *isomorfo* a $L^1(\mathbb{R})$, exatamente o mesmo comportamento que obtemos ao exigir que as funções tenham $\mathbb{R}$ como domínio[†]: basta aliar o exercício anterior com o próximo. △

Exercício 3.23 ($^{\star}_{\star\star}$)**.** Seja $f\colon \mathbb{R} \to \mathbb{R}$ uma função qualquer.

a) Mostre que se $f = 0$ q.t.p., então $f \in \mathscr{L}^1(\mathbb{R})$ e $\pi(f) = 0$. Dica: você precisa de uma sequência rápida que testemunhe $f \in \mathscr{L}^1(\mathbb{R})$.

b) Mostre que se $f \in \mathscr{L}^1(\mathbb{R})$ e $\pi(f) = 0$, então $f = 0$ q.t.p. Dica:

- as hipóteses asseguram uma sequência rápida $(f_n)_n$ em $\mathcal{C}_C(\mathbb{R})$ com $\lim\limits_{n\to\infty} f_n = f$ q.t.p. e $\lim\limits_{n\to\infty} \widehat{f_n} = 0$ em $L^1(\mathbb{R})$;
- perceba que $\displaystyle\int_{\mathbb{R}} |f_n| \to 0$ e daí, por Chebyshev, mostre que para $\varepsilon > 0$ e $n \in \mathbb{N}$ quaisquer, o conjunto $E_{n,\varepsilon} := \{x \in \mathbb{R} : |f_n(x)| \geq \varepsilon\}$ é tal que
$$\lim_{n\to\infty} m^*(E_{n,\varepsilon}) = 0;$$
- use o item anterior na verificação de que
$$F_\varepsilon := \{x \in \mathbb{R} : \{m \in \mathbb{N} : |f_m(x)| \geq \varepsilon\} \text{ é infinito}\}$$
tem medida nula (sugestão: $F_\varepsilon = \displaystyle\bigcap_{n=0}^{\infty} \bigcup_{j\geq n} E_{n,\varepsilon}$);

[†]E isto, por sua vez, é inofensivo: se a única coisa que interessa numa função $f \in \mathscr{L}^1(\mathbb{R})$ são seus valores a menos de conjuntos de medida nula, basta "completar" o domínio de f com valores arbitrários.

- enfim, observe que se S é o conjunto de medida nula em que $\lim_{n\to\infty} f_n(x) = f(x)$ **não** acontece, então
$$\lim_{n\to\infty} f_n(x) = 0$$
para todo $x \in \mathbb{R} \setminus \left(S \cup \bigcup_{n\in\mathbb{N}} F_{\frac{1}{2^n}}\right)$.

c) Se você já estudou o Teorema do Isomorfismo em Álgebra Linear: conclua que $L^1(\mathbb{R})$ é o quociente de $\mathscr{L}^1(\mathbb{R})$ pela relação de equivalência $\equiv$ em que $f, g \in \mathscr{L}^1(\mathbb{R})$ são declaradas equivalentes se, e somente se, $f = g$ q.t.p. ■

Após essa longa batalha para definir $\mathscr{L}^1(\mathbb{R})$ apropriadamente, a integral de Lebesgue de uma função em $\mathscr{L}^1(\mathbb{R})$ se define do modo mais natural possível:

Definição 3.1.27 (Integral de Lebesgue)**.** Para $f \in \mathscr{L}^1(\mathbb{R})$, vamos escrever
$$\int f := \lim_{n\to\infty} \int_{\mathbb{R}} f_n,$$
para qualquer sequência rápida $(f_n)_n$ em $(\mathcal{C}_C(\mathbb{R}), \|\cdot\|_1)$ tal que $\lim_{n\to\infty} f_n = f$ q.t.p.. O número $\int f$ será chamado de a **integral de Lebesgue de** f. ¶

A *naturalidade* da definição acima se reflete no fato de que ela surge (naturalmente...) de um *diagrama comutativo*:

$$\begin{tikzcd}
\mathscr{L}^1(\mathbb{R}) \arrow[r, "\pi"] & L^1(\mathbb{R}) \arrow[dr, dashed, "\int(\cdot)"] & \\
\mathcal{C}_C(\mathbb{R}) \arrow[u, "i"] \arrow[r, equal, "\widehat{(\cdot)}"] \arrow[rr, bend right, "\int_{\mathbb{R}}(\cdot)"'] & \widehat{\mathcal{C}_C(\mathbb{R})} \arrow[u, "i"] \arrow[r] & \mathbb{R}
\end{tikzcd}$$

Explicitamente, como a integral de Riemann é um funcional linear (uniformemente!) contínuo, segue que existe uma única extensão (linear!) contínua $\int : L^1(\mathbb{R}) \to \mathbb{R}$. Assim, para estender tal definição para as funções "de verdade" em $\mathscr{L}^1(\mathbb{R})$, basta compor com a projeção π. Se este argumento categórico lhe parecer ofensivo, basta encarar o Teorema 3.1.24 até que ele encare de volta: se duas sequências rápidas $(f_n)_n$ e $(g_n)_n$ em $\mathcal{C}_C(\mathbb{R})$ são tais que
$$\lim_{n\to\infty} f_n = \lim_{n\to\infty} g_n \quad \text{q.t.p.}$$
então
$$0 \leq \lim \left| \int_{\mathbb{R}} f_n - g_n \right| \leq \lim_{n\to\infty} \int_{\mathbb{R}} |f_n - g_n| = 0,$$
o que basta para garantir
$$\lim_{n\to\infty} \int_{\mathbb{R}} f_n = \lim_{n\to\infty} \int_{\mathbb{R}} g_n,$$
haja vista que ambos os limites são números reais (complete os detalhes!)$^{\star}_{\star}$.

Em particular,
$$\int f = \int_{\mathbb{R}} f$$

para qualquer função $f \in \mathcal{C}_C(\mathbb{R})$ ou, verbalmente: a integral de Riemann de uma função contínua com suporte compacto coincide com sua integral de Lebesgue.

E para quê isso tudo? Okay, a discussão sobre a falta de completude de $(\mathcal{C}_C(\mathbb{R}), \|\cdot\|_1)$ já adianta que $\mathcal{C}_C(\mathbb{R}) \subsetneq \mathscr{L}^1(\mathbb{R})$, i.e., necessariamente existem funções Lebesgue integráveis que não são Riemann integráveis (pense a respeito)*. No entanto, a mera capacidade de *integrar* mais funções não justifica todo esse trabalho. A completude de $L^1(\mathbb{R})$, por outro lado, é a justificativa verdadeiramente promissora, mas discutir isso exigiria pelo menos mais um extenso capítulo — e eu já não aguento mais. Caso se interesse numa discussão honesta das vantagens sobre a integral de Lebesgue, confira [23].

Exercícios adicionais

Exercício 3.24 $({}^{\star}_{\star})$**.** Mostre que se $f \in \mathcal{C}_C(\mathbb{R})$, então f é uniformemente contínua. Dica: existe $M > 0$ tal que $\operatorname{supp}(f) \subseteq [-M, M]$; lembre-se então de que a restrição de f ao compacto $[-M-1, M+1]$ já é uniformemente contínua[†]. ■

Exercício 3.25 (For fun — $({}^{\star}_{\star\star})$)**.** Para X espaço métrico e $f\colon X \to \mathbb{R}$ contínua com suporte compacto, mostre que f é uniformemente contínua. Cuidado: em espaços métricos, bolas fechadas podem não ser compactas, o que atrapalha adaptar diretamente a sugestão do exercício anterior. Dica: em vez de adaptar, observe que para $\varepsilon > 0$, $J := \{x \in X : |f(x)| \geq \varepsilon\}$ é fechado em X, contido no *interior* do suporte de f, o que permite encontrar $\delta' > 0$ tal que $J[\delta'] := \bigcup_{x \in J} B[x, \delta'] \subseteq K$ (cf. Lema 2.4.21). ■

Exercício 3.26 $(\star)$**.** Mostre que

$$\int_{-\infty}^{+\infty} f(t)\,\mathrm{d}t = \int_{\mathbb{R}} f$$

para qualquer $f \in \mathcal{C}_C(\mathbb{R})$, i.e., a integral imprópria de uma função contínua $f\colon \mathbb{R} \to \mathbb{R}$ com suporte compacto coincide com a integral de Riemann em qualquer intervalo que contenha seu suporte[‡]. ■

Exercício 3.27 $({}^{\star}_{\star})$**.** Seja $f \in \mathscr{L}^1(\mathbb{R})$.

a) Mostre que $|f| \in \mathscr{L}^1(\mathbb{R})$;

b) Mostre que

$$\begin{aligned} \|\cdot\|_1\colon \mathscr{L}^1(\mathbb{R}) &\to \mathbb{R} \\ g &\mapsto \int |g| \end{aligned}$$

é uma *seminorma*, i.e.: satisfaz todos os axiomas de norma, exceto "$\|f\|_1 = 0 \Rightarrow f = 0$".

[†]Mas tome cuidado: para $\varepsilon > 0$, a continuidade uniforme em $[-M-1, M+1]$ assegura $\delta > 0$ tal que $|f(x) - f(y)| < \varepsilon$ sempre que $x, y \in [-M-1, M+1]$ satisfazem $|x - y| < \delta$, de modo que ao estender as considerações para $\mathbb{R}$, precisa-se considerar as situações em que $x \in [-M-1, M+1]$ mas y não, o que exige *adaptar* o δ.

[‡]Embora não seja surpresa para ninguém, é importante deixar registrado, até porque isto mostra que é possível partir diretamente da integração de Riemann em sentido próprio para a integração de Lebesgue — o que pode ser útil se você estiver sem tempo.

c) (+⋆) A função $\pi\colon \mathscr{L}^1(\mathbb{R}) \to L^1(\mathbb{R})$ é contínua? Dica: como no caso dos espaços normados, basta testar se π é contínua no vetor (função) nula; alternativamente, observe que vale $\|\pi(f)\| = \|f\|_1$, onde $\|\cdot\|$ indica a norma de $L^1(\mathbb{R})$ herdada de $\mathcal{L}_c(\mathcal{L}_c(\mathcal{C}_C(\mathbb{R})))$ (cf. Definição 3.1.13).

d) Mostre que $(\mathcal{C}_C(\mathbb{R}), \|\cdot\|_1)$ é denso em $(\mathscr{L}^1(\mathbb{R}), \|\cdot\|_1)$, e que este último é um espaço *pseudo*métrico completo, i.e., no qual toda sequência de Cauchy converge[†]. Dica: para a primeira parte, *lembre-se* (prove!) que se

$$\lim_{m,n\to\infty} a_{m,n} = L \quad \text{e} \quad \lim_{m\to\infty} a_{m,n} = L_n \quad \text{para cada } n \in \mathbb{N},$$

então $\lim\limits_{n\to\infty} L_n = L$, onde o primeiro limite se refere à rede $(a_{m,n})_{(m,n)\in\mathbb{N}\times\mathbb{N}}$ com respeito ao conjunto dirigido $\mathbb{N}\times\mathbb{N}$ com a ordem produto usual (cf. Exercício 1.68); para a segunda parte, observe que se $(f_n)_n$ é uma sequência rápida em $\mathscr{L}^1(\mathbb{R})$ e $g_n \in \mathcal{C}_C(\mathbb{R})$ é tal que $\|f_n - g_n\|_1 < \frac{1}{2^n}$ (que existe pela densidade!), então $(g_n)_n$ é rápida em $(\mathcal{C}_C(\mathbb{R}), \|\cdot\|_1)$ e qualquer $f \in \mathscr{L}^1(\mathbb{R})$ satisfazendo $\|f - g_n\|_1 \to 0$ será tal que $\|f - f_n\|_1 \to 0$. ■

3.2 Compacidade e convergência uniforme

Nesta altura da vida, você já não precisa de motivações para se importar por compacidade (mas se precisar, pense em problemas de maximização/minimização + Teorema de Weierstrass). Com isso dito, o Lema de Riesz nos ensinou que bolas fechadas não são compactas em espaços normados com dimensão infinita (cf. Corolário 2.0.25), enquanto o Exercício 1.245 estabeleceu que $\mathcal{C}(X)$ tem dimensão infinita para qualquer espaço métrico infinito X. Isto traz um problema: sem o Teorema de Heine-Borel, como reconhecer os subconjuntos compactos de $(\mathcal{C}[a,b], \|\cdot\|_\infty)$, por exemplo?

Observação 3.2.0. Pelo restante desta seção, X denotará um espaço métrico compacto, essencialmente pelo fato de que os argumentos não se alteram se assumirmos $X := [0,1]$, por exemplo[‡]. Além disso, para não carregar a notação, $\mathcal{C}(X)$ será sempre considerado com a norma do supremo. △

Em vista do Teorema 2.0.22, sabemos que uma família de funções $K \subseteq \mathcal{C}(X)$ é compacta com respeito à topologia induzida pela norma do supremo se, e somente se, K for completo com a métrica induzida e totalmente limitado. Posto que $\mathcal{C}(X)$ já é completo, isto se resume a pedir que K seja *fechado* (cf. Exercício 1.257) e totalmente limitado, o que nos deixa com o dilema de investigar os subconjuntos *totalmente limitados* de $\mathcal{C}(X)$, uma vez que a limitação total se preserva no fecho (em virtude do próximo exercício).

Exercício 3.28 ($^\star_\star$)**.** Para um espaço métrico M, mostre que $\overline{L}$ é totalmente limitado sempre que $L \subseteq M$ for totalmente limitado. Dica: tome uma sequência $(z_n)_n$ em $\overline{L}$ e, para cada n, fixe uma sequência $(x_{n,m})_m$ em L satisfazendo $x_{n,m} \to z_n$; escolhendo termos x_{n,m_n} de maneira esperta para cada n, use uma subsequência de Cauchy de $(x_{n,m_n})_n$ para obter uma subsequência de Cauchy de $(z_n)_n$. ■

[†]Sim, é a mesma definição de espaço métrico, exceto que em vez de uma métrica, utiliza-se uma *pseudométrica*, que por sua vez é uma métrica que não satisfaz "$d(x,y) = 0 \Rightarrow x = 0$". Nesta altura do campeonato, você já deve ter percebido que seminormas induzem pseudométricas.

[‡]A situação muda se assumirmos X meramente compacto de Hausdorff, sem metrizabilidade: embora o resultado permaneça verdadeiro *mutatis mutandis*, a demonstração é mais complicada.

Exemplo 3.2.1 (Limitação total $\neq$ limitação)**.** Não custa lembrar: limitação total implica limitação, mas não vale a volta. Com efeito, se L é subconjunto totalmente limitado de X, então $\overline{L}$ é compacto e, portanto, limitado. Por outro lado, o Lema de Riesz já mostrou que as bolas fechadas (logo, limitadas) de espaços de Banach com dimensão infinita não são compactas — e, portanto, não podem ser totalmente limitadas. ▲

De volta ao problema de caracterizar os subespaços totalmente limitados de $\mathcal{C}(X)$: certamente são subconjuntos limitados segundo a norma $\|\cdot\|_\infty$, mas a questão é perceber qual condição adicional deve ser imposta para recuperar a limitação total. Observe que se $K \subseteq \mathcal{C}(X)$ é totalmente limitado, então para qualquer $\varepsilon > 0$ e $x \in X$ pode-se encontrar um aberto $V \subseteq X$ em torno de x tal que $|f(x) - f(y)| < \varepsilon$ para todo $y \in V$ e toda $f \in K$:

- ✓ pela afirmação estabelecida ao longo da demonstração do Teorema 2.0.22, existem $f_0, \ldots, f_n \in K$ tais que $K \subseteq \bigcup_{j \leq n} B[f_j, \varepsilon]$;
- ✓ a continuidade de cada f_j permite encontrar um aberto $V \subseteq X$ em torno de x satisfazendo $|f_j(x) - f_j(y)| < \varepsilon$ para todo $y \in V$ e *todo* $j \leq n$;
- ✓ ajustando os ε's e lembrando que $f \in B[f_j, r]$ se, e somente se, $\|f - f_j\|_\infty \leq r$, o resultado segue.

Exercício 3.29 $\left(^{\star}_{\star}\right)$**.** Complete o argumento anterior. ■

Em certo sentido, a condição verificada acima indica que a *continuidade* das funções em K *independe* da função considerada, no sentido de que o aberto em torno de x por meio do qual se controla a função é o mesmo para toda $f \in K$.

Definição 3.2.2. Diz-se que $K \subseteq \mathcal{C}(X)$ é **equicontínua no ponto** x se para todo $\varepsilon > 0$ existe um aberto $V \subseteq X$ com $x \in V$ e $|f(x) - f(y)| < \varepsilon$ para quaisquer $y \in V$ e $f \in K$. A menção ao ponto costuma ser suprimida se K for equicontínua em *todos* os pontos. ¶

Exemplo 3.2.3. Há exemplos banais de famílias equicontínuas:

(i) $K := \{f\}$ para alguma f contínua;

(ii) $K := K' \cup K''$, com K' e K'' equicontínuas;

(iii) $K \subseteq K'$ com K' equicontínua.

Para um caso menos preguiçoso, mas não tanto: fixe $\alpha \in \mathbb{R}$ com $\alpha > 0$ e para cada $t \in [0, \alpha)$, seja $f_t \colon [0,1] \to \mathbb{R}$ a função dada por $f_t(x) := tx$; em tais condições, $K_\alpha := \{f_t : t \in [0, \alpha)\}$ é equicontínua. De fato, para $\varepsilon > 0$ fixado, basta tomar $\delta < \frac{\varepsilon}{\alpha}$, pois se $x, y \in [0,1]$ satisfazem $|x - y| < \delta$, então $|f_t(x) - f_t(y)| \leq t|x - y| < \varepsilon$; em particular, note que o δ escolhido também independe do ponto x fixado, caso em que se diz que a família é **uniformemente equicontínua**. ▲

Exercício 3.30 $\left(^{\star}_{\star\star}\right)$**.** Mostre que se X é métrico compacto e $K \subseteq \mathcal{C}(X)$ é equicontínua, então K é uniformemente equicontínua. Dica: imite a demonstração do Teorema de Heine-Cantor. ■

É justamente ao adicionar equicontinuidade à noção usual de limitação que se recupera a limitação total.

Teorema 3.2.4 (Arzelà-Ascoli)**.** *Para X métrico e compacto, uma família $K \subseteq \mathcal{C}(X)$ é totalmente limitada se, e somente se, K é limitada e equicontínua.*

Demonstração. As discussões anteriores já deram conta da "ida". Tratemos da "volta". Como X também é totalmente limitado, para cada $n \in \mathbb{N}$ existe um subconjunto finito $F_n \subseteq X$ tal que $X = \bigcup_{x \in F_n} B_X\left[x, \frac{1}{2^n}\right]$, o que resulta num subconjunto enumerável (e denso, certo?) $F \coloneqq \bigcup_{n \in \mathbb{N}} F_n$, que pode ser reescrito como $F = \{x_n : n \in \mathbb{N}\}$. O truque agora é aplicar um argumento de diagonalização à moda de Cantor a fim de obter uma subsequência de Cauchy de uma sequência $(f_n)_n$ de funções tomadas em K.

Como a família K é limitada com respeito à norma $\|\cdot\|_\infty$, segue que para qualquer $x \in X$, $(f_n(x))_n$ é uma sequência limitada de números reais. Em particular, $(f_n(x_0))_n$ é limitada em $\mathbb{R}$ e, portanto, admite uma subsequência convergente $(f_{n,0}(x_0))_n$. Novamente, $(f_{n,0}(x))_n$ é limitada para todo $x \in X$ e, em particular, para $x \coloneqq x_1$, $(f_{n,0}(x_1))_n$ tem uma subsequência convergente $(f_{n,1}(x_1))_n$. Procedendo desta forma, obtêm-se sequências $(f_{n,k})_n$ para cada $k \in \mathbb{N}$, de tal forma que $(f_{n,k+1})_n$ é subsequência de $(f_{n,k})_n$ e $(f_{n,k}(x_k))_n$ converge em $\mathbb{R}$, para cada $k \in \mathbb{N}$. Para encerrar, basta mostrar que $(f_{k,k})_k$ é subsequência de Cauchy de $(f_n)_n$.

Primeiro, note que $(f_{k,k})_k$ é, de fato, uma subsequência de $(f_n)_n$: por definição de subsequência, o modo como as subsequências foram tomadas garante, para cada k, uma função $h_k \colon \mathbb{N} \to \mathbb{N}$ estritamente crescente tal que $(f_{n,k})_n = \left(f_{h_k(n)}\right)_n$ e $f_{h_{k+1}(n)} = f_{h_k(h_{k+1}(n))}$, de modo que $(f_{k,k})_k$ corresponde à subsequência de $(f_n)_n$ induzida pela função estritamente crescente $\varphi \colon \mathbb{N} \to \mathbb{N}$ que faz $k \mapsto h_0(h_1(h_2(\ldots(h_k(k)))))$ para todo $k \in \mathbb{N}$. O esquema a seguir pode te ajudar a se convencer.

$$
\begin{array}{cccccccccccccc}
(f_n)_n & = f_0 & \widehat{f_1} & f_2 & f_3 & f_4 & f_5 & f_6 & f_7 & f_8 & \widehat{f_9} & f_{10} & f_{11} & \cdots \\
 & & \| & & \| & & \| & & \| & & \| & & \| & \\
(f_{n,0})_n & = & \widehat{f_{0,0}} & & f_{1,0} & & f_{2,0} & & f_{3,0} & & f_{4,0} & & f_{5,0} & \cdots \\
 & & & & & & \| & & & & \| & & \| & \\
(f_{n,1})_n & = & & & & & f_{0,1} & & & & \widehat{f_{1,1}} & & f_{1,2} & \cdots \\
 & & & & & & & & & & & & \| & \\
(f_{n,2})_n & = & & & & & & & & & & & f_{2,0} & \cdots \\
 & & & & & & & & & & & & & \\
\vdots & & & & & & & & & & & & \vdots & \ddots
\end{array}
$$

Figura 3.1: Note que $f_{0,0} \coloneqq f_{n_0}$ para algum $n_0 \geq 0$ (no caso, $n_0 \coloneqq 1$). Como $(f_{n,1})_n$ é subsequência de $(f_{n,0})_n$, temos $f_{1,1} = f_{m_1,0}$ para algum $m_1 > m_0$ (no caso, $m_0 \coloneqq 2$ e $m_1 \coloneqq 4$) e, por sua vez, $f_{m_1,0} = f_{n_{m_1}}$, para algum $n_{m_1} \geq m_1 > m_0 \geq n_0$ (no caso, $n_{m_1} = 9$). E assim por diante.

Além disso, como $\{f_{k,k} : k \in \mathbb{N}\} \subseteq K$ com K equicontínuo, segue que $\{f_{k,k} : k \in \mathbb{N}\}$ também é equicontínua, e mais ainda: por X ser compacto, o exercício anterior assegura que a equicontinuidade é uniforme. Logo, existe $N \in \mathbb{N}$ tal que $|f_{k,k}(x) - f_{k,k}(y)| < \varepsilon$ para todo $k \in \mathbb{N}$ e quaisquer $x, y \in X$ com $d(x,y) < \frac{1}{2^N}$.

Após se convencer de que $(f_{k,k})_{k\geq k'}$ é subsequência de $(f_{n,k'})_n$ para todo $k' \in \mathbb{N}$ (faça isso!)$^{\star}_{\star}$, não será difícil perceber que para todo $z \in F_N$, $(f_{k,k}(z))_k$ converge em $\mathbb{R}$: $z = x_{k'}$ para algum k', e $(f_{n,k'}(x_{k'}))_n$ converge em $\mathbb{R}$, por construção. Logo, por F_N ser finito e $(f_{k,k}(z))_k$ ser de Cauchy para cada $z \in F_N$, existe um índice $M \in \mathbb{N}$ com $|f_{k,k}(z) - f_{l,l}(z)| < \varepsilon$ para quaisquer $k, l \geq M$ e $z \in F_N$ (percebeu?)$^{\star}_{\star}$.

Acabou: para $x \in X$ qualquer, existe pelo menos um $z \in F_N$ com $d(x,z) < \frac{1}{2^N}$ (por quê?)*, de modo que para $k, l \geq M$,

$$|f_{k,k}(x) - f_{l,l}(x)| \leq |f_{k,k}(x) - f_{k,k}(z)| + |f_{k,k}(z) - f_{l,l}(z)| + |f_{l,l}(z) - f_{l,l}(z)| < 3\varepsilon,$$

donde a arbitrariedade do x tomado acarreta $\|f_{k,k} - f_{l,l}\|_\infty \leq 3\varepsilon$. □

Exercício 3.31 ($^{\star}_{\star}$)**.** Ajuste os ε's na argumentação anterior e conclua a demonstração.■

Observação 3.2.5. É relativamente comum encontrar o seguinte enunciado para o Teorema de Arzelà-Ascoli: *para X espaço métrico e compacto, $\overline{K} \subseteq \mathcal{C}(X)$ é compacto se, e somente se, K é família equicontínua e $K(x) := \{f(x) : f \in K\}$ é limitado em $\mathbb{R}$ para todo $x \in X$.*

Tal versão é equivalente à que foi apresentada aqui†. Por um lado, se K é $\|\cdot\|_\infty$-limitado, então $K(x)$ é limitado para todo $x \in X$. Por outro lado, se K é (uniformemente) equicontínua e $K(x)$ é limitado em $\mathbb{R}$ para cada $x \in X$, então K é $\|\cdot\|_\infty$-limitado: pela equicontinuidade uniforme (garantida por X ser compacto), existe $\delta > 0$ tal que $|f(x) - f(y)| < 1$ para toda $f \in K$ e quaisquer $x, y \in X$ com $d(x,y) < \delta$, daí não é difícil utilizar a compacidade de X para obter um limitante superior para o conjunto $\{|f(x)| : x \in X \text{ e } f \in K\}$. △

Exercício 3.32 ($^{\star}_{\star\star}$)**.** Complete o argumento anterior. ■

Corolário 3.2.6 (Forma clássica de Arzelà-Ascoli)**.** *Se $(f_n)_n$ é sequência de funções contínuas da forma $[a,b] \to \mathbb{R}$ tal que $(f_n)_n$ é limitada na norma do supremo e $\{f_n : n \in \mathbb{N}\}$ é equicontínua, então $(f_n)_n$ admite subsequência que converge uniformemente.*

Demonstração. Segue do teorema anterior, posto que $\overline{\{f_n : n \in \mathbb{N}\}}$ será fechado e totalmente limitado e, portanto, um espaço métrico compacto. □

Moralmente, o Teorema de Arzelà-Ascoli deve ser encarado como uma versão do Teorema de Heine-Borel-Lebesgue que caracteriza os compactos de $\mathbb{R}^n$: no caso euclidiano clássico, os compactos são caracterizados pela *dobradinha* “fechado+limitado”, enquanto que para espaços de funções, Arzelà-Ascoli apenas acrescenta a condição de equicontinuidade. Nesse sentido, as aplicações usuais são análogas: garantir a existência de subsequências convergentes a fim de encontrar soluções de equações.

Exemplo 3.2.7 (Cf. Exercício 2.78)**.** Considere o problema de mostrar que *toda função polinomial é fechada*, i.e., se $p\colon \mathbb{R} \to \mathbb{R}$ é polinomial e $F \subseteq \mathbb{R}$ é fechado, então $p[F]$ é fechado. Ora, na prática, isto consiste em tomar $y \in \overline{p[F]}$ e mostrar que existe $x \in F$ com $p(x) = y$, ou seja: precisamos solucionar a equação “$p(x) = y$”.

†Adaptada de Efe Ok. [31].

Certamente existe $(x_n)_n$ em F com $p(x_n) \to y$, de modo que se $(x_n)_n$ convergisse ou, pelo menos, tivesse uma subsequência convergente, a continuidade de p asseguraria uma solução para a equação: se $x_{n_k} \to x$, então $x \in F$ (pois F é fechado) e $p(x_{n_k}) \to p(x)$. O problema é que $(x_n)_n$ poderia ser ilimitada, certo? Errado: se fosse, então haveria uma subsequência $\left(x_{n_j}\right)_j$ convergindo para $\pm\infty$ e, por p ser polinomial, resultaria $p(x_{n_j}) \not\to y$. Portanto (e este é o ponto importante!), $(x_n)_n$ tem uma subsequência convergente, justamente pela caracterização de compacidade em $\mathbb{R}$.

Se em vez de $p\colon \mathbb{R} \to \mathbb{R}$ tivéssemos uma função da forma $P\colon \mathcal{C}([0,1]) \to \mathcal{C}([0,1])$, por exemplo, e procurássemos uma função g num subconjunto K satisfazendo $P(g) = f$, teríamos uma equação envolvendo funções. Neste caso, se P fosse contínua e K fosse compacto, poderíamos apelar para subsequências de forma semelhante ao que se fez no parágrafo anterior — e, justamente em tais contextos funcionais, costuma ser mais simples verificar limitação uniforme e equicontinuidade. ▲

Exercícios adicionais

Exercício 3.33 $(\star)$**.** Seja $(f_n)_{n\in\mathbb{N}}$ uma sequência de funções contínuas da forma $[a,b] \to \mathbb{R}$. Mostre que se $f_n \to_{\mathrm{u}} f$, então $\{f_n : n \in \mathbb{N}\}$ é uniformemente equicontínua. ■

Exercício 3.34 $({}^{\star}_{\star})$**.** Seja $(f_n)_{n\in\mathbb{N}}$ uma sequência (uniformemente) equicontínua de funções da forma $[a,b] \to \mathbb{R}$. Mostre que se $(f_n(d))_{n\in\mathbb{N}}$ é de Cauchy para todo ponto d num denso enumerável de $[a,b]$, então $(f_n)_{n\in\mathbb{N}}$ converge uniformemente. Dica: você já fez isso. ■

Exercício 3.35 $({}^{\star}_{\star\star})$**.** Seja $(f_n)_n$ uma sequência de funções da forma $[a,b] \to \mathbb{R}$, todas diferenciáveis em (a,b), e tais que as derivadas sejam uniformemente limitadas. Mostre que se $(f_n(x))_{n\in\mathbb{N}}$ é limitada para algum $x \in [a,b]$, então $(f_n)_n$ tem subsequência que converge uniformemente. Dica: T.V.M. ■

O próximo resultado lida com o problema da *aproximação*: dada uma função contínua $f\colon [a,b] \to \mathbb{R}$, é possível aproximá-la por funções contínuas bem comportadas?

Teorema 3.2.8 (Stone-Weierstrass)**.** *Seja X um espaço compacto de Hausdorff. Se $\mathcal{A} \subseteq \mathcal{C}_u(X)$ é uma álgebra fechada e que separa pontos, então $\mathcal{A} = \mathcal{C}_u(X)$.*

No enunciado acima, uma família $\mathcal{A}$ de funções da forma $X \to \mathbb{R}$ é chamada de **álgebra** se $\mathcal{A}$ contém todas as funções constantes e, para quaisquer $f, g \in \mathcal{A}$ ocorrer $f + g \in \mathcal{A}$ e $f \cdot g \in \mathcal{A}$. Por sua vez, $\mathcal{A}$ **separa pontos** se para quaisquer $x, y \in X$ distintos existe $h \in \mathcal{A}$ satisfazendo $h(x) \neq h(y)$. Finalmente, a condição de fechamento sobre $\mathcal{A}$ se refere à topologia da convergência uniforme em $\mathcal{C}_u(X)$.

Ingredientes

Antes de proceder com a prova do teorema, que depende de três lemas simples, convém observar como o enunciado acima se relaciona com sua versão clássica. A primeira coisa a notar é que a condição de fechamento no enunciado de *Stone-Weierstrass* pode ser retirada desde que se mude a conclusão: qualquer álgebra de funções que separa pontos é *densa* em $\mathcal{C}_u(X)$, o que segue do

Lema 3.2.9. *Sejam X um espaço topológico e $\mathcal{A} \subseteq \mathcal{C}_u(X)$ uma álgebra de funções. Se as funções de $\mathcal{A}$ são limitadas, então $\mathcal{B} := \overline{\mathcal{A}}$ ainda é uma álgebra de funções.*

Demonstração. É claro que $\mathcal{B}$ contém todas as funções constantes. Agora, se $f, g \in \mathcal{B}$, então $f + g$ e $f \cdot g$ pertencem a $\mathcal{B}$: de fato, existem sequências $(f_n)_n$ e $(g_n)_n$ em $\mathcal{A}$, com $f_n \to f$ e $g_n \to g$ uniformemente; como $f_n + g_n, f_n \cdot g_n \in \mathcal{A}$ e $f_n + g_n \to f + g$ e $f_n \cdot g_n \to f \cdot g$ uniformemente, o resultado segue. □

Exercício 3.36 ($^{\star}_{\star}$)**.** Assumindo a validade do Teorema 3.2.8, mostre que se $f\colon [a,b]\to\mathbb{R}$ é contínua, então existe uma sequência $(p_n)_n$ de funções polinomiais da forma $[a,b]\to\mathbb{R}$ tais que $p_n\to f$ uniformemente. Dica: além do lema anterior, observe que a função identidade, que separa pontos trivialmente, é polinomial. ■

Caso queira ver algumas aplicações desse teorema de aproximação, confira a subsubseção de exercícios que fecha esta seção. Por aqui, vamos provar o Teorema de Stone-Weierstrass em três passos, sendo que o primeiro já foi feito: o Teorema de Dini (cf Exercício 2.109). Em certo sentido, o segundo passo é uma dose *não homeopática* do próprio resultado que buscamos provar[†].

Lema 3.2.10. *Existe uma sequência $(w_n)_n$ de polinômios tal que $w_n(x)\to\sqrt{x}$ uniformemente em* $[0,1]$.

Demonstração. Para cada $x\in[0,1]$, sejam $w_0(x) \coloneqq 0$, $w_1(x)\coloneqq w_0(x)+\frac{1}{2}(x-w_0^2(x))$ e, supondo a função polinomial $w_n(x)$ definida para $n>0$, considere

$$w_{n+1}(x)\coloneqq w_n(x)+\frac{1}{2}(x-w_n^2(x)). \tag{3.13}$$

Para aquecer os motores, vamos ver que $w_n(x)\leq\sqrt{x}$ para qualquer $n\in\mathbb{N}$ e $x\in[0,1]$. Isto é claro para $n\coloneqq 0$. Supondo a afirmação válida para $n>0$, note que

$$\sqrt{x}-w_{n+1}(x)=\sqrt{x}-w_n(x)-\frac{1}{2}(x-w_n^2(x))=\sqrt{x}-w_n(x)-\frac{1}{2}(\sqrt{x}-w_n(x))(\sqrt{x}+w_n(x))=$$
$$=(\sqrt{x}-w_n(x))\left(1-\frac{1}{2}(\sqrt{x}+w_n(x))\right).$$

Como $0\leq x\leq 1$, tem-se $0\leq\sqrt{x}\leq 1$, donde a hipótese de indução dá tanto $\sqrt{x}-w_n(x)\geq 0$ quanto $\sqrt{x}+w_n(x)\leq 2$ e, consequentemente, $1-\dfrac{1}{2}(\sqrt{x}+w_n(x))\geq 0$. Logo, $\sqrt{x}-w_{n+1}(x)\geq 0\cdot 0=0$, mostrando que $w_{n+1}(x)\leq\sqrt{x}$ em $[0,1]$, como afirmado.

Isso garante que $x-w_n^2(x)\geq 0$ para todo n e, por conseguinte, a sequência $(w_n(x))_n$ é crescente[‡] para cada ponto $x\in[0,1]$. Logo, existe $y\in[0,1]$ com $w_n(x)\to y$. Ao fazer $n\to+\infty$ em (3.13), resulta

$$y=y+\frac{1}{2}(x-y^2)\Rightarrow 0=\frac{1}{2}(x-y^2)\Rightarrow y^2=x\Rightarrow$$
$$\Rightarrow y=|y|=\sqrt{y^2}=\sqrt{x},$$

i.e., $w_n(x)\to\sqrt{x}$ para todo $x\in[0,1]$. A uniformidade da convergência segue então pelo Teorema de Dini. □

Lema 3.2.11 (Terceiro passo)**.** *Seja $\mathcal{A}$ uma álgebra de funções contínuas e limitadas definidas em X. Se $\mathcal{A}$ é fechada em $\mathcal{C}_u(X)$, então* $\max(f,g),\min(f,g)\in\mathcal{A}$ *para quaisquer $f,g\in\mathcal{A}$.*

Demonstração. O primeiro *Katzensprung* consiste em notar que

$$\min(f,g)=\frac{1}{2}(f+g-|f-g|)\quad\text{e}\quad\max(f,g)=\frac{1}{2}(f+g+|f-g|).$$

[†]Se fosse homeopática seria inútil.
[‡]Reveja a definição recursiva dos polinômios w_n em caso de dúvidas!

Por conta disso, é suficiente mostrar que $|f| \in \mathcal{A}$ sempre que $f \in \mathcal{A}$. O segundo *Katzensprung* é ainda mais legal: podemos supor $|f| \leq 1$. De fato, se $|f| > 1$, a hipótese de que as funções de $\mathcal{A}$ são limitadas assegura um $c > 0$ com $|f| \leq c$, de modo que a suposição se aplica a $\frac{1}{c}f \in \mathcal{A}$. Finalmente, o terceiro *Katzensprung* usa o lema anterior: como $|f| \leq 1$ e $f \in \mathcal{A}$, tem-se $f^2 \in \mathcal{A}$ e $w_n \circ f^2 \in \mathcal{A}$ (vide o exercício a seguir), com $w_n \circ f^2 \to \sqrt{f^2} = |f|$ uniformemente (pelo lema anterior), donde a pertinência $|f| \in \mathcal{A}$ segue por $\mathcal{A}$ ser fechada em $\mathcal{C}_u(X)$. □

Exercício 3.37 ($^\star_\star$)**.** Sejam $\mathcal{A}$ um álgebra de funções e $p\colon \mathbb{R} \to \mathbb{R}$ uma função polinomial. Mostre que $p \circ g \in \mathcal{A}$ para qualquer $g \in \mathcal{A}$. ■

Demonstração do Teorema de Stone-Weierstrass. Pela hipótese de $\mathcal{A}$ ser fechada por convergência uniforme, basta mostrar que para $f\colon X \to \mathbb{R}$ contínua e $\varepsilon > 0$ fixados, existe $f_\varepsilon \in \mathcal{A}$ com $\|f_\varepsilon - f\|_\infty \leq \varepsilon$. De fato, com tal resultado *em mãos*, ao fazer $g_n := f_{\frac{1}{2^n}}$ para cada $n \in \mathbb{N}$ resulta $g_n \to f$ em $\mathcal{C}_u(X)$ e, portanto, $f \in \overline{\mathcal{A}} = \mathcal{A}$.

Não é difícil perceber que a função f_ε desejada, na prática, deve tão somente satisfazer

$$f(x) - \varepsilon < f_\varepsilon(x) < f(x) + \varepsilon \tag{3.14}$$

para todo $x \in X$. A compacidade de X será essencial nesta busca, que se inicia a seguir.

Fixados $a, b \in X$ pontos distintos, a hipótese de que $\mathcal{A}$ separa pontos dá uma função $h \in \mathcal{A}$ com $h(a) \neq h(b)$. Daí, para

$$g := \frac{1}{h(b) - h(a)}(h(x) - h(a)),$$

deve-se ter $g \in \mathcal{A}$, $g(a) = 0$ e $g(b) = 1$, bem como $f_{a,b} := (f(b) - f(a))g + f(a) \in \mathcal{A}$, já que $f(b) - f(a)$ e $f(a)$ são constantes. Enfim, por valerem as identidades $f_{a,b}(a) = f(a)$ e $f_{a,b}(b) = f(b)$, resulta que os abertos

$$U_{a,b} := \{x \in X : f_{a,b}(x) < f(x) + \varepsilon\} \quad \text{e} \quad V_{a,b} := \{x \in X : f_{a,b}(x) > f(x) - \varepsilon\}$$

são tais que $a, b \in U_{a,b} \cap V_{a,b}$. Em certo sentido, a função $f_{a,b}$ faz em $U_{a,b} \cap V_{a,b}$ o que gostaríamos que f_ε fizesse em todo o X. O restante da prova consiste em usar a compacidade para cozinhar a função f_ε a partir das funções $f_{a,b}$'s.

Primeiro, para $b \in X$ fixado, a família $\{U_{a,b} : a \in X \setminus \{b\}\}$ é uma cobertura por abertos para X. Logo, a compacidade de X *fornece* $a_0, \ldots, a_n \in X$ tais que $X = U_{a_0,b} \cup \ldots \cup U_{a_n,b}$. Da compacidade de X e da continuidade das funções em $\mathcal{A}$ segue que todas são limitadas, o que permite usar o lema anterior, por indução, a fim de estabelecer

$$f_b := \min(f_{a_0,b}, \ldots, f_{a_n,b}) \in \mathcal{A}.$$

Note que para $x \in X$ qualquer, existe $i \leq n$ com $x \in U_{a_i,b}$, o que garante a desigualdade $f_b(x) \leq f_{a_i,b}(x) < f(x) + \varepsilon$, "metade" do que se estipulou em (3.14).

Por sua vez, ocorre $b \in \bigcap_{i \leq n} V_{a_i,b} := V_b$ e, para cada $y \in V_b$, $f(y) - \varepsilon < f_{a_i,b}(y)$, donde segue que $f(y) - \varepsilon < f_b(y)$. Nossos problemas estariam resolvidos se ocorresse $V_b = X$. Embora isso não aconteça, pode-se remediar a situação, apelando-se mais uma vez para a compacidade.

Repetindo o procedimento anterior para cada $b \in X$, obtêm-se $b_0, \ldots, b_m \in X$ satisfazendo a identidade $X = V_{b_0} \cup \ldots \cup V_{b_m}$. Novamente, o lema anterior assegura $f_\varepsilon := \max(f_{b_0}, \ldots, f_{b_m}) \in \mathcal{A}$, que é digna de tal abreviação: de fato, por termos $f_{b_i}(x) < f(x) + \varepsilon$ para cada $i \leq m$, deve ocorrer $f_\varepsilon(x) < f(x) + \varepsilon$; por outro lado, dado $x \in X$ qualquer, existe b_j com $x \in V_{b_j}$, e daí $f(x) - \varepsilon < f_{b_j}(x) \leq f_\varepsilon(x)$, como desejado. □

Exercício 3.38 ($^{\star}_{\star\star}$)**.** Complete os detalhes omitidos na demonstração anterior. ■

Exercício 3.39 ($^{\star}_{\star}$)**.** Mostre que $\mathcal{C}[a,b]$ tem um subconjunto denso enumerável. Dica: quantos polinômios com coeficientes racionais existem? ■

Exercício 3.40 ($^{\star}_{\star\star}$)**.** Sejam $f, g\colon [a,b] \to \mathbb{R}$ funções contínuas tais que

$$\int_a^b t^n f(t)\,\mathrm{d}t = \int_a^b t^n g(t)\,\mathrm{d}t$$

para todo $n \in \mathbb{N}$. Mostre que $f = g$. Dica: se $\displaystyle\int_a^b t^n h(t)\,\mathrm{d}t = 0$ para todo n e $(p_n)_{n\in\mathbb{N}}$ é sequência de polinômios com $p_n \to_{\mathrm{u}} h$ em $[a,b]$ então o Corolário 2.1.24 assegura

$$\int_a^b h^2(t)\,\mathrm{d}t = \lim_{n\to\infty} \int_a^b p_n(t)h(t)\,\mathrm{d}t;$$

qual o valor de $\displaystyle\int_a^b p_n(t)h(t)\,\mathrm{d}t$? ■

Lista de símbolos e siglas

$-A$	reflexão de A em torno da origem, 77
$-\infty$ e $+\infty$	pontos no infinito de um corpo estendido, 80
$[-\infty, +\infty]_{\mathbb{K}}$	corpo estendido a partir do corpo ordenado $\mathbb{K}$, 80
$[-\infty, \beta)$	intervalo na reta estendida, fechado em $-\infty$ e aberto em β, 80
$(\alpha, +\infty]$	intervalo na reta estendida, aberto em α e fechado em $+\infty$, 80
(a, b)	intervalo aberto em a e b, 80
$[a, b]$	intervalo fechado em a e b, 82
$[a, b)$	intervalo fechado em a e aberto em b, 82
$(a, b]$	intervalo aberto em a e fechado em b, 82
$\mathbb{R}$	conjunto dos números reais, 92
$\mathrm{seq}\,(\mathbb{R})$	conjunto das sequências finitas de números reais, 95
$\sum_{i\leq n} f_i$ ou $\sum_{i=0}^{n} f_i$	somatório, 96
$\prod_{i\leq n} f_i$ ou $\prod_{i=0}^{n} f_i$	produtório, 96
$\sqrt{r}$	raiz quadrada de $r \geq 0$, 102
$\lVert\cdot\rVert$	norma num espaço vetorial, 103
$\lVert\cdot\rVert_\infty$	norma do supremo (ou do máximo), 103
$\lim_{x\to p} f(x)$	limite de $f(x)$ quando x tende a p, 104
$\lim_{n\to\infty} x_n$	limite da sequência $(x_n)_n$, 104
$(x_d : n \in \mathbb{D})$ ou $(x_d)_{d\in\mathbb{D}}$ ou $(x_d)_d$	rede, 106
$x_d \to L$	$(x_d)_d$ converge para L, 106
$\lim_{d\in\mathbb{D}} x_d$ ou $\lim_{d\in\mathbb{D}} x_d$	limite da rede $(x_d)_{d\in\mathbb{D}}$, 106
$\lVert\cdot\rVert_2$	norma euclidiana, 107
$\lVert\cdot\rVert_1$	norma da soma, 108
$\sum_{n=0}^{\infty} x_n$, $\sum_{n\in\mathbb{N}} x_n$ ou $\sum x_n$	série determinada pela sequência $(x_n)_{n\in\mathbb{N}}$, 114
$[-\infty, +\infty]$	reta estendida, 116
$\lim_{d\in\mathbb{D}} x_d$ ou $\lim_{d\in\mathbb{D}} x_d$	limite da rede $(x_d)_{d\in\mathbb{D}}$, 117
$\lim_{x\to p^-} f(x)$	limite lateral à esquerda de p, 120
$\lim_{x\to p^+} f(x)$	limite lateral à direita de p, 120
$B_d(x, r)$ ou $B(x, r)$	d-bola aberta de centro x e raio r, 121
$\mathrm{Par}_{\mathcal{R}}\,[a, b]$	conjunto das partições de Riemann do intervalo $[a, b]$, 136
$\sum_{(\mathcal{P},T)} f$	soma de Riemann de f associada à partição $(\mathcal{P}, T)$, 136
$\lVert\mathcal{P}\rVert$	norma da partição $\mathcal{P}$ de um intervalo $[a, b]$, 137

f^+	parte positiva de f, 281
f^-	parte negativa de f, 281
π	pi, 290
$J_f(a)$	Jacobiana de f no ponto a, 300
f'	derivada de f, 307
$\mathcal{L}_2(X \times Y, Z)$	conjunto das funções bilineares da forma $X \times Y \to Z$, 307
$\mathcal{L}_{2,c}(X \times Y, Z)$	espaço das transformações bilineares contínuas, 309
$\partial_i f(a)$	derivada parcial de f no ponto a na direção do i-ésimo vetor da base canônica, 309
$\operatorname{supp}(f)$	suporte da função f, 313
$\mathcal{C}_C(X)$	funções contínuas $X \to \mathbb{R}$ com suporte compacto, 313
$\mathcal{L}_c(X)$	dual topológico de X, 317
$\mathscr{L}^1(\mathbb{R})$	espaço das funções reais Lebesgue-integráveis, 320
q.t.p	quase todo ponto, 324
$\int f$	integral de Lebesgue de f, 328

Referências Bibliográficas

[0] E. Acosta and G. Delgado. Fréchet vs. Carathéodory. *American Mathematical Monthly*, 101:332–338, 1994.

[1] T. Apostol. *Mathematical Analysis*. Addison-Wesley, 2 edition, 1981.

[2] S. Arora, H. Browne, and D. Daners. An alternative approach to Fréchet derivatives. *Journal of the Australian Mathematical Society*, 111(2):202–220, 2021.

[3] A. F. Beardon. *Limits: a new approach to Real Analysis*. Undergraduate Texts in Mathematics. Springer, 1997.

[4] N. Bourbaki. *General Topology, part 1*. Addison-Wesley, London, 1966.

[5] N. Bourbaki. *General Topology, part 2*. Addison-Wesley, London, 1966.

[6] H. Brézis. *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. Universitext. Springer, 1 edition, 2010.

[7] L. Bukovský. *The structure of the real line*. Monografie Matematyczne 71. Birkhäuser Basel, 1 edition, 2011.

[8] P. L. Clark. Honors Calculus, 2014. Notas de aula, University of Georgia.

[9] D. G. de Figueiredo. *Análise I*. LTC, 2 edition, 1996.

[10] R. A. dos Santos Fajardo. *A Teoria dos Conjuntos e os Fundamentos da Matemática*. EDUSP, 1 edition, 2025.

[11] R. Engelking. *General Topology: Revised and completed edition*. Sigma series in pure mathematics. Heldermann Verlag, Berlin, 1989.

[12] J. Ferreirós. *Labyrinth of thought: A history of set theory and its role in modern mathematics*. Birkhäuser Basel, 2ª edition, 2007.

[13] R. Gordon. Riemann integration in Banach spaces. *Rocky Mountain Journal of Mathematics*, 21(3), 1991.

[14] E. Hairer and G. Wanner. *Analysis by its History*. Undergraduate texts in mathmatics. Readings in mathmatics. Springer, 2008.

[15] J. F. Hall. *Completeness of Ordered Fields*. California Polytechnic State University, 2010. Monografia de graduação.

[16] J. D. Hamkins. *Lectures on the Philosophy of Mathematics*. MIT Press, 2021.

[17] K. Hoffman. *Analysis in Euclidean Space.* Prentice-Hall, 1975.

[18] S. Johar. *The Big Book of Real Analysis: From Numbers to Measures.* Springer, 2024.

[19] V. J. Katz and K. H. Parshall. *Taming the Unknown: A History of Algebra from Antiquity to the Early Twentieth Century.* Princeton University Press, 2014.

[20] H. Keisler. *Elementary Calculus. An Infinitesimal Approach.* Dover, 2 edition, 2000.

[21] D. S. Kurtz and C. W. Swartz. *Theories of integration - The integrals of Riemann, Lebesgue, Henstock-Kurzweil, and McShane.* Series in Real Analysis. World Scientific, 2004.

[22] S. Lang. *Undergraduate Analysis.* Undergraduate Texts in Mathematics. Springer, 2 ed edition, 1997.

[23] A. D. Lewis. Should we fly in the Lebesgue-designed airplane? – the correct defence of the lebesgue integral. *arXiv*, 2023.

[24] E. L. Lima. *Análise Real, Volume 1: Funções de uma Variável.* IMPA, 2006.

[25] E. L. Lima. *Curso de Análise, Volume 1.* IMPA, 14 edition, 2017.

[26] E. L. Lima. *Curso de Análise, Volume 2.* IMPA, 14 edition, 2017.

[27] P. Maddy. *Defending the Axioms: On the Philosophical Foundations of Set Theory.* Oxford University Press, USA, 2011.

[28] R. M. Mezabarba. Fundamentos de Topologia Geral, 2023. manuscrito, disponível em https://github.com/mezabarbarm/Fund_Top_Geral.

[29] R. M. Mezabarba. Um curso fechado e limitado de análise real, 2023. manuscrito, disponível em https://github.com/mezabarbarm/AnalysisZero.

[30] R. M. Mezabarba. *Teoria dos conjuntos: uma introdução maliciosa.* Number 26 in Textuniversitários. LF Editorial, São Paulo, 1 edition, 2025.

[31] E. A. Ok. *Real Analysis with Economic Applications.* Princeton University, 2007.

[32] C. C. Pugh. *Real Mathematical Analysis.* Undergraduate Texts in Mathematics. Springer, 2 edition, 2015.

[33] M. Raman-Sundström. A pedagogical history of compactness. *The American Mathematical Monthly*, 122(7):619–635, 2015.

[34] T. Roque. *História da Matemática - Uma Visão Crítica, Desfazendo Mitos e Lendas.* Zahar, 2012.

[35] C. R. Rosentrater. *Varieties of Integration.* MAA Press, 2015.

[36] W. Rudin. *Principles of Mathematical Analysis.* McGraw Hill, 3 edition, 1976.

[37] E. Schechter. *Handbook of Analysis and Its Foundations.* Academic Press, 1996.

[38] R. Shakarchi. *Problems and solutions for undergraduate analysis.* Undergraduate Texts in Mathematics. Springer, 1 edition, 1998.

[39] S. Shapiro, editor. *The Oxford Handbook of Philosophy of Mathematics and Logic.* Oxford Handbooks in Philosophy. Oxford University, 2005.

[40] B. Simon. *Real Analysis. A Comprehensive Course in Analysis, Part 1.* American Mathematical Society, 2015.

[41] T. Sonar. *3000 Years of Analysis: Mathematics in History and Culture.* Birkhäuser, 2021.

[42] T. Tao. *Analysis I.* Texts and Readings in Mathematics. Springer, 3 edition, 2016.

[43] S. Willard. *General Topology.* Addison-Wesley, New York, 1970. Reprinted in 2004 by Dover.

Índice Remissivo