
ADONAI SANT'ANNA
MATEMÁTICA PANDÊMICA
Como ocorre a contagiosa disseminação de ZFC

Matemática Pandêmica

como ocorre a contagiosa disseminação de ZFC

MATEMÁTICA PANDÊMICA

COMO OCORRE A CONTAGIOSA DISSEMINAÇÃO DE ZFC

ADONAI S. SANT'ANNA

Este arquivo foi obtido no sítio

https://matematicapandemica.blog/.

Versão expandida de notas de aula usadas pelo autor em atividades remotas do Departamento de Matemática da Universidade Federal do Paraná (DMAT/UFPR) durante os dois primeiros anos da pandemia de COVID-19.

Público-alvo: aqueles que não se contentam com a usual metodologia doutrinária de ensino de matemática.

Críticas e sugestões são bem-vindas para o aprimoramento deste documento. Contatos com o autor podem ser feitos por e-mail no endereço abaixo.

adonaisantanna@gmail.com

Quarta versão

Janeiro de 2022

Dedicado a todos os J. L., verdadeiros guias para este livro.

Prefácio

Esta é a quarta e última versão de Matemática Pandêmica. A proposta é mostrar que aritmética, cálculo diferencial e integral, álgebra linear, geometria euclidiana e teoria probabilidades podem ser percebidos simplesmente como o estudo de casos muito particulares de conjuntos. Logo, para conhecer esses temas, é desejável familiaridade com alguma teoria usual de conjuntos. Escolhemos a de Zermelo-Fraenkel por ser a mais popular, além de ser suficientemente forte para lidar com as necessidades de ferramentas elementares para desenvolver e aplicar matemática.

Aproveitamos a oportunidade para apontar múltiplos possíveis caminhos para desenvolver matemática, a qual é um dos ramos do conhecimento mais importantes em termos de impacto social, econômico, tecnológico, cultural e filosófico.

Este livro é disponibilizado gratuitamente em formato PDF, com diversos recursos de navegabilidade. Basta o leitor clicar nos trechos em azul para ser imediatamente levado ao item citado. Mas foi planejado para funcionar também em forma impressa, caso o leitor escolha esta opção. Isso porque Partes, Seções, Teoremas, Proposições, Definições e Exemplos são numerados. Além disso, há um extenso índice remissivo para auxiliar na busca por informações.

Sumário
Índice
Rede

Sumário

PARTE 1

Introdução

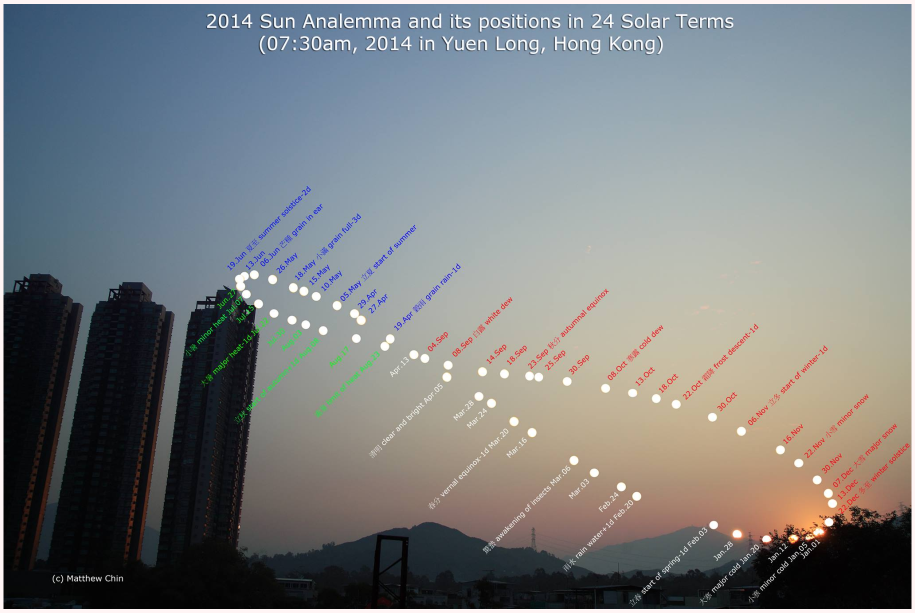
Sumário
Índice
Rede

Este livro é aquilo que eu gostaria de ter estudado em meu primeiro ano como aluno de licenciatura em matemática. Pena que eu não sabia disso em 1983.

Não é um texto que trata exaustivamente sobre os temas abordados. Longe disso, apenas proponho este livro como um primeiro contato com assuntos básicos indispensáveis a uma compreensão mínima sobre matemática. Mas é necessário um primeiro contato honestamente fundamentado para que o aluno tenha condições de buscar autonomia em seus estudos. Se consigo atingir meu propósito, apenas os leitores poderão responder. Mas esta é a minha intenção aqui.

A abordagem usual em aulas de matemática mais parece um processo de doutrinação do que um exame crítico de conhecimentos científicos. Aquele que inicia seus estudos de matemática deve perceber que essa ciência é tema de debates acalorados. Debate e doutrinação são incompatíveis entre si. Além disso, deve perceber também que existem múltiplas formas de fazer matemática. Em algumas dessas formas, por exemplo, o argumento de *redução ao absurdo* é legítimo, enquanto em outras não é. Em certas formulações de cálculo diferencial e integral há infinitesimais, enquanto em outras esse conceito simplesmente não é sequer formulável. Em algumas formulações de álgebra linear todo espaço vetorial tem base, enquanto em outras isso não acontece. Em teorias usuais de conjuntos, funções têm nomes, como f e g. Em cálculo lambda isso não ocorre e nem pode ocorrer. Matemática é uma área do conhecimento muito flexível e com considerável tolerância para ideias novas e críticas a ideias antigas. Mas, sem conhecimento, não há espaço para a criatividade. O estudo de matemática deve estar focado na direção da criatividade e da análise crítica, não de procedimentos eficazes para a aprovação em exames.

Quando nossos ancestrais contemplaram as estrelas pela primeira vez – observando padrões de movimentos de corpos celestes, bem como as relações entre o céu e as estações do ano – naquele momento nascia a matemática. Naqueles tempos remotos matemática era um processo de abstração que se caracterizava pela identificação de padrões na natureza. Ninguém olhou para o céu pensando que aquela vastidão de complexos padrões era um ótimo ponto de partida para a criação de vestibulares e concursos públicos.



Analema marcando a posição do Sol às 7h30, em Hong Kong

Fonte: EarthSky.

Daí a necessidade de um estudo introdutório justificado e bem fundamentado para a matemática! O estudo criticamente fundamentado da matemática é imprescindível para que alunos aprendam a não levar muito a sério professores e autores. Matemática é um fenômeno humano que paradoxalmente transcende as idiossincrasias humanas.

Neste contexto, o presente documento é destinado àqueles que desejam respostas claras, honestas e motivadoras para pelo menos alguns dos procedimentos elementares e usuais da matemática. Quaisquer erros aqui cometidos são responsabilidade minha, não da matemática.

Seguem algumas questões básicas tratadas neste livro.

i: Há alguma justificativa para as famosas *regras de sinais* da multiplicação entre inteiros? Ver Seção 30.

ii: O que é, afinal, a *unidade imaginária dos números complexos*? Ver Seção 40.

iii: O que é um *número*? Ver início da Parte 4.

iv: O que é $5^{\sqrt{2}}$? Ver Seção 66.

v: Qual é a diferença entre equação e função? Ver Seção 43.

vi: O que são soluções de uma equação? Ver Seção 10, bem como inúmeros exemplos importantes ao longo de todo o livro.

vii: Se o conjunto vazio não tem elementos, como pode estar contido em qualquer conjunto? Ver Teorema 3.6.

viii: O que é um *conjunto*? Se um conjunto é uma coleção de objetos, o conjunto vazio pode ser interpretado como uma coleção de selos sem um único selo? Se fosse o caso, qualquer pessoa é um colecionador de selos, mesmo não tendo uma única peça que justifique sua suposta coleção! Ver Parte 3.

ix: O que é *infinito*? O infinito é algo que não acaba? Qual é o critério a ser usado para responder se algo acaba ou não? Paciência eterna? Ver Seção 33.

x: O que tem a ver logaritmos com teoria de grupos? Ver Seção 68.

xi: O que é um ponto no plano euclidiano? Se um ponto não tem largura, altura ou profundidade, então a cor vermelha é um ponto? Afinal, a cor vermelha não tem altura, largura ou profundidade. Ver Parte 7.

xii: Mais importante, o que tem a ver matemática com a vida de cada um de nós?

Seção 1

Matemática como fenômeno humano

Sumário
Índice
Rede

Começamos respondendo parcialmente à última questão da Seção anterior. Matemática é uma das atividades de maior impacto social na história da humanidade, independentemente de convicções políticas, religiosas ou pessoais. Matemática não precisa de bandeiras, hinos, eleições, brasões, regulamentações ou decretos para se impor. Isso porque matemática é naturalmente uma boa ideia. Métodos matemáticos são empregados com grande sucesso nas seguintes áreas.

- física: para uma melhor compreensão sobre o universo onde vivemos, via geometria diferencial, espaços de Hilbert, funções especiais e outros.
- tecnologia: para a concepção de equipamentos, métodos e materiais.
- artes: no desenvolvimento de novas técnicas artísticas baseadas em processos iterativos, *splines*, fractais e outros conceitos.
- medicina: na criação de novas drogas, equipamentos e métodos de investigação, incluindo modelos matemáticos de proliferação de agentes infecciosos.
- paleontologia: via métodos de datação.
- sociologia: na concepção de modelos que permitam antecipar o futuro de civilizações.
- psicologia: via teoria das decisões ou cognição quântica, entre outros.
- linguística: através de métodos estatísticos ou gramáticas gerativas.
- economia: via teoria dos jogos e pesquisa operacional.
- mercado de ações: via sistemas ergódicos.
- cálculo de prêmios de seguros: via matemática atuarial.
- administração de empresas: via análise multivariada de dados.
- sistemas de segurança civil e militar: via criptografia.

- ⌨ ciência da computação: via cálculo lambda, máquinas de Turing, teoria de categorias e séries de Fourier, entre outros.
- ☣ guerra: via equações diferenciais, teoria das decisões, teoria dos jogos, pesquisa operacional, criptografia e muitas outras ferramentas usadas também em tempos de paz.
- ⚽ muitas outras áreas: música estocástica, cinema, arquitetura, história, química, geografia, esportes, logística em geral, inteligência artificial etc.

Assim como ocorre com a música, matemática é naturalmente cultivada por comunistas e capitalistas, crentes e ateus, homossexuais e heterossexuais, conservadores e liberais, vendedores e compradores, especuladores e empresários, índios e europeus, empáticos e psicopatas, militares e civis, nazistas e judeus, estudiosos e leigos.

No entanto, matemática não se limita a aplicações imediatas para lidar com problemas do mundo real. A atividade matemática atingiu um nível de amadurecimento que lhe proporcionou a qualidade de objeto de estudo por mérito próprio. A compreensão da matemática enquanto legítimo campo de estudos é condição indispensável para antecipar novas aplicações no futuro. Daí a ênfase neste documento sobre parte dos fundamentos desta ciência formal!

Compreender minimamente a atividade matemática é uma condição necessária (apesar de não suficiente) para o amadurecimento de uma visão sensata e bem informada sobre o mundo onde vivemos.

Dois ingredientes − desde que tratados com certo cuidado − definem matemática: linguagem e lógica.

Linguagem é um instrumento de comunicação [49] (ao lado de outros, como gestos e pantomimas) que serve ao propósito de veicular ideias e sentimentos. Faz parte da natureza humana a veiculação de ideias e sentimentos.

Lógica permite concatenar ideias veiculadas pelo emprego de linguagens. Tal concatenação é realizada por meio de inferências. Faz parte da natureza humana a concatenação de ideias.

Logo, matemática espelha dois aspectos profundos sobre a natureza humana:
as necessidades de comunicação e de inferir consequências a partir daquilo que é comunicado.

A partir do momento em que naturalmente empregamos linguagens e inferências em nosso cotidiano, isso significa que naturalmente usamos matemática todos os dias. Mas o estudo sistemático de matemática em si é algo que exige muito mais, além de experiências cotidianas.

Em uma primeira aproximação sobre linguagens, essas podem ser divididas em dois grupos:

I aquelas que estão naturalmente comprometidas com uma semântica e

II aquelas que não estão.

A língua portuguesa se enquadra no item I. Com efeito, para fins de mera ilustração, o termo 'cadeira' é usualmente interpretado como uma cadeira no mundo real. Mais do que isso, não é usual interpretar a palavra 'cadeira' como um sorvete derretido de baunilha ou um sentimento de repulsa a aranhas. Neste sentido, há uma certa rigidez na dimensão semântica de uma linguagem como o português.

É claro que nem todos os termos da língua portuguesa podem ser interpretados como objetos do mundo real. Exemplos triviais são as palavras 'unicórnio' e 'lobisomem'. No entanto, ainda permanece invariante o compromisso de associar termos da língua portuguesa a coisas, lugares, épocas, sentimentos, intuições ou ideias que transcendem a própria linguagem, os quais são os significados dos termos. Ainda que poetas como Fernando Pessoa consigam explorar certas liberdades, como na frase 'O mito é o nada que é tudo', pessoas são compelidas a associarem uma frase da língua portuguesa a potenciais significados, mesmo que tais significados não sejam necessariamente compartilhados por duas ou mais pessoas.

Neste contexto significados não podem ser confundidos com sinônimos. Uma palavra da língua portuguesa pode ser um sinônimo de outra no sentido de que, pelo menos em certos contextos de caráter pragmático, elas compartilham um mesmo significado.

Matemática, lógica formal e ciência da computação, não obstante, são ramos do conhecimento que demandam o emprego de *linguagens formais* não comprometidas com qualquer contraparte semântica. Em outras palavras, o emprego de língua portuguesa é, no mínimo, insuficiente para lidar com a matemática exigida hoje em dia. Logo, não é surpreendente que pessoas sem treino matemático percebam com estranheza essa atividade humana. Mas, assim como os modos de pensar de mulheres causam estranheza diante de certos homens, isso não muda o fato de que mulheres são seres humanos. Matemáticos, com o perdão dos platonistas, são apenas criaturas que exploram certos aspectos da natureza humana nem sempre contemplados por não-matemáticos.

Enquanto aspectos significativos da língua portuguesa podem ser reduzidos ao estudo de morfemas, porção significativa da matemática pode ser reduzida ao estudo de *objetos matemáticos*. Os objetos matemáticos mais amplamente estudados e usados na literatura especializada são *conjuntos*. Logo, a compreensão sobre teoria de conjuntos é um passo natural para a devida apreciação da matemática, pelo menos diante das atuais visões sobre o que é essa ciência. Isso justifica o fato de que este livro inicia o estudo de matemática a partir da teoria de Zermelo-Fraenkel, a mais popular teoria de conjuntos.

Para ilustrar essas primeiras considerações a respeito das linguagens formais da matemática, considere a equação

$$x = 1 + \frac{x}{2}.$$

Uma equação é um caso particular de *fórmula* de uma certa linguagem formal (a qual não é comprometida com qualquer semântica em particular). Fórmulas de uma linguagem formal são afirmações feitas (no contexto da linguagem formal) sobre certos objetos matemáticos. Se os objetos de estudo são conjuntos, os termos x, 1, 2 e $\frac{x}{2}$ (que ocorrem na equação acima) são conjuntos. O símbolo $+$ na equação acima corresponde a uma operação entre conjuntos, a qual produz novos conjuntos. Portanto, $1 + \frac{x}{2}$ também é um conjunto. Mas a fórmula em questão (a qual pode ser entendida como uma afirmação sobre os conjuntos x, 1, 2 e $\frac{x}{2}$) é desprovida de significado. Não há compromisso com qualquer contraparte semântica. Esse fato confere significativa liberdade à matemática. Temos a liberdade de interpretar a fórmula acima de várias maneiras:

I: um tijolo pesa um quilo mais meio tijolo (neste caso x é interpretado como o peso de um tijolo em quilogramas);

II: a idade de Alfredo é igual a um ano somado da metade da idade de Alfredo (neste caso x é interpretado como a idade de Alfredo em anos);

III: o número de pessoas na sala é igual a uma pessoa mais a metade do número total de pessoas da sala (neste caso x é interpretado como número de pessoas na sala);

entre muitos outros possíveis exemplos. Por conta disso, a matemática encontra ampla aplicabilidade em múltiplas áreas do conhecimento.

É justamente a renúncia da matemática ao mundo real que a torna tão útil no mundo real.

Outra vantagem do descompromisso de linguagens formais com semântica reside no fato de que linguagens naturais, como o português, são preocupantemente ambíguas para fins científicos. Por exemplo, o verbo 'ser' pode expressar

UMA PREDICAÇÃO: como na frase 'Ernst Zermelo é inteligente';

UMA IDENTIDADE: como na frase 'Ernst Zermelo é o criador do Axioma da Escolha;

UMA EXISTÊNCIA: como na frase 'Ernst Zermelo é';

UMA INCLUSÃO DE CLASSE: como na frase 'Ernst Zermelo é um matemático';

entre outras possibilidades. Observar, por exemplo, a dificuldade para discernir PREDICAÇÃO de INCLUSÃO DE CLASSE.

Ambiguidades são nocivas para a atividade científica, uma vez que frequentemente a ciência se vê obrigada a lidar com situações não familiares à maioria das pessoas. Logo, é necessária clareza de ideias, antes de avançarmos na atividade científica.

Sem compromisso com semântica, linguagens formais não abrem espaço para ambiguidades no sentido acima colocado. No entanto, ainda resiste uma certa ambiguidade muito mais sutil, mesmo em certas linguagens formais. Discutimos sobre isso na Seção 111.

Com relação à lógica, matemática emprega diferentes formas de inferência, as quais viabilizam relações entre fórmulas. Lógica permite inferir novas fórmulas a partir de fórmulas anteriormente conhecidas. No caso da equação

$$x = 1 + \frac{x}{2},$$

é possível inferir que $x = 2$, desde que lógica e linguagem sejam claramente definidas com antecedência.

Logo, seguindo os exemplos acima, o tijolo pesa dois quilogramas, Alfredo tem dois anos de idade e a sala conta com duas pessoas. Neste livro, porém, estamos interessado em apenas um tipo particular de inferência: as dedutivas. Detalhes são apresentados nas próximas seções.

Para uma visão mais ampla sobre outras formas de inferência em matemática e demais áreas do conhecimento, recomendamos o livro de Ian Hacking [20] sobre indução e probabilidades. Na obra citada o autor estimula o leitor com uma lista de sete problemas com enunciados perfeitamente compreensíveis mesmo entre aqueles sem treino matemático. São problemas cujas soluções desafiam aquilo que normalmente se assume como senso comum. Um deles, referente a testemunhos de eventos extraordinários, está reproduzido na Parte 9.

Senso comum não é um bom ponto de partida para uma visão racional de mundo. Com efeito, senso comum reflete uma visão compartilhada entre segmentos sociais. Racionalidade, porém, não é democrática.

Como enfatizado por Bertrand Russell, um dos aspectos essenciais da racionalidade é não ter qualquer certeza inquestionável.

Seção 2

Bastam linguagem e lógica?

Sumário
Índice
Rede

Ainda que pelo menos alguns ramos da matemática estejam suficientemente definidos pelos ingredientes 'lógica' e 'linguagem formal', a prática social de *fazer*, *aplicar*, *justificar*, *questionar*, *filosofar*, *especular*, *cultivar*, *divulgar* e até mesmo *financiar* matemática demanda muito mais. Uma pessoa que tenha uma nova ideia matemática precisa convencer pessoas qualificadas sobre a relevância, a originalidade e a validade de sua proposta. A estratégia social mais comum e confiável para convencer pessoas sobre novas ideias matemáticas é a veiculação de artigos científicos em periódicos especializados. Não pretendemos explorar este delicado ponto aqui. Mas é imprescindível que o leitor compreenda que matemática é uma atividade social. Sem trocas de ideias não há matemática alguma.

Ilustramos a seguir em que sentido a prática matemática exige muito mais do que linguagem e lógica. No final do século 19 Georg Cantor teve uma ideia: introduzir o conceito de conjunto para qualificar o que é o infinito. Com o passar de décadas de pesquisas, matemáticos do mundo todo perceberam a elegância, o alcance, a originalidade, a aplicabilidade e, especialmente, a necessidade das ideias de Cantor. Antes dele, Bernardus Bolzano teve ideias semelhantes. Mas a proposta dele seguia uma estratégia com severas limitações. A proposta de Bolzano foi esquecida [57], enquanto a de Cantor triunfou, apesar das duras críticas nas primeiras décadas após o primeiro artigo dele sobre o tema. É assim que funciona a prática social da matemática: ideias seguidas de discussões. O nascer de ideias é um aspecto não matemático da prática social da matemática. O mesmo é cabível para a análise crítica de novas ideias. Toda nova ideia é uma ilha cercada por antigas concepções. É necessário sair da ilha para apreciá-la como um todo.

Tanto Bolzano quanto Cantor foram os primeiros a perceber a importância de qualificação do infinito. Mas Bolzano sustentou sua proposta em uma visão muito difícil de colocar em prática. Neste sentido, as ideias de Cantor foram mais felizes e, por conseguinte, mais facilmente aceitáveis. Apesar disso, até mesmo a proposta original de Cantor enfrentou forte resistência.

A questão do financiamento da matemática é ainda mais ardilosa, uma vez que ela depende de decisões políticas de governantes e empresários, os quais não são necessariamente familiarizados com os poderosos efeitos da matemática a curto, médio e longo prazo sobre comunidades, sociedades, nações e o mundo onde vivemos. Daí a importância da divulgação da matemática para um público leigo. Tal discussão, porém, escapa dos propósitos desta obra.

O foco deste livro é o emprego de uma única linguagem formal e uma única lógica, para fins de fundamentação de vastos ramos da matemática, como aritmética, álgebra, álgebra linear, topologia, probabilidades, geometria, cálculo diferencial e integral, equações diferenciais e muito mais.

Neste livro é discutida de maneira sucinta a teoria de conjuntos de Zermelo-Fraenkel (ZF), bem como a aplicabilidade da mesma em alguns dos ramos mencionados acima. A teoria ZF é a mais popular entre as formalizações atualmente conhecidas para as ideias originais de Cantor, o criador da teoria de conjuntos.

SEÇÃO 3

Requisitos para a leitura

SUMÁRIO
ÍNDICE
REDE

São três os resquisitos indispensáveis para a compreensão dos assuntos aqui abordados.

- Saber ler.
- Ter a mente aberta.
- Dispor-se a dialogar com pares.

Leitura não é uma atividade fácil. Isso porque ela exige senso crítico e, consequentemente, uma boa dose de cultura. Não há senso crítico, por exemplo, entre aqueles que assumem o que o autor quis dizer sem que ele tenha de fato dito. Não há senso crítico também onde certezas estão alojadas, seja por convicções pessoais ou por conta de submissão ao doutrinamento promovido em escolas. Para fins de ilustração, se o leitor tem certeza de que $0 + 5 = 5$, vale observar que apenas na Seção 29 conseguimos provar isso. Tal prova é feita no Teorema 4.2 e consome uma redação de trinta e duas linhas de justificativas. Ademais, não há senso crítico onde domina a ignorância. A não familiaridade com cultura científica e filosófica é um terreno árido onde dificilmente podem brotar questionamentos pertinentes. Senso crítico é o exercício de enunciar questionamentos pertinentes. A avaliação da pertinência de uma pergunta, no entanto, é um processo subjetivo. Logo, não é fácil uma pessoa responder a si mesma se é capaz de ler e levantar questões relevantes a partir do que leu. Para ilustrar exemplos de questões pertinentes, ver Seção 46.

Mente aberta é a qualidade de saber lidar com incertezas como, por exemplo, o problema de estabelecer o que é pertinente no estudo de matemática. Um dos principais obstáculos contra o aprendizado de matemática reside em preconceitos intelectuais. Aquilo que alguém julga saber pode oferecer espantosa resistência contra novos aprendizados. Para citar um exemplo simples, teoria de conjuntos *não é* uma teoria sobre coisas chamadas conjuntos. Teoria de conjuntos é um corpo do conhecimento que trata de dois predicados binários, conhecidos como igualdade e pertinência. Conjuntos, no contexto de ZF, são apenas termos de uma linguagem formal. Do ponto de vista matemático, não faz diferença alguma chamar os termos de ZF de conjuntos, unicórnios ou cenouras. O que está em jogo são as relações entre pertinência e igualdade, os alicerces da teoria de conjuntos ZF. Teoria de conjuntos não visa o estudo de cenouras. Teoria de conjuntos é um assunto que demanda modos de pensar abstratos e refinados, algo que só pode ser conquistado com um alinhamento entre aptidão, disposição e paciência. Retomamos esse assunto na Parte 11, a qual só pode ser apreciada após um detalhado estudo sobre ZF na Parte 3.

O *diálogo com pares* se refere à troca de ideias matemáticas com pessoas que compartilham os mesmos interesses e com dedicação destacada à matemática. Uma vez que a *mente aberta* é um fenômeno emergente entre grupos de pessoas que compartilham a mesma busca por melhores ferramentas para a compreensão do mundo onde vivemos, cada um dos requisitos acima está emaranhado com os demais. A busca pelo conhecimento não é uma aventura que possa ser realizada em solitude. Ciência é um fenômeno social sinérgico.

Jules Henri Poincaré desenvolveu uma extensa obra de enorme impacto para os fundamentos de teorias físicas, como a relatividade restrita e a mecânica celeste, bem como teorias matemáticas, como topologia, álgebra e equações diferenciais. Também foi um grande filósofo e um brilhante escritor.

No livro *La Valeur de la Science* (publicado em 1905 e traduzido para vários idiomas), Poincaré afirma:

O matemático nasce, não se cria.

Neste contexto é importante o leitor não confundir práticas institucionais de ensino de matemática com matemática. Adestrar alunos a se tornarem mímicos de atitudes tipicamente encontradas entre matemáticos é algo muito diferente de estudar e fazer matemática. Daí a importância dos requisitos acima.

Seção 4

Diferenciais desta obra

Sumário
Índice
Rede

A meta principal deste livro é introduzir conceitos básicos típicos de um primeiro ano de estudos de graduação em matemática, física e áreas afins, como engenharia e matemática industrial. No entanto, a abordagem adotada segue algumas diferenças em relação à literatura padrão:

i: Linguagem e lógica são explicitados, para fins de fundamentação, bem como seus papeis na prática matemática.

ii: É destacada a existência de outras maneiras para desenvolver matemática, além das mais usuais, como análise infinitesimal suave, lógica intuicionista, lógica paraconsistente, entre outras.

iii: É salientado que até hoje não se sabe se a fundamentação usual via teoria de conjuntos de Zermelo-Fraenkel é consistente.

iv: São propostos exercícios que visam promover mudanças na definição usual de limite de função real, com o propósito de compreender melhor esse importante conceito.

v: É destacada a importância de cálculo diferencial e integral para definir seno, co-seno, logaritmo e exponencial, entre outras funções de uso corrente.

vi: As interpretações geométricas de seno e co-seno são exibidas como teoremas a partir da definição dada por soluções de uma equação diferencial.

vii: São qualificados os conceitos de *definição*, *teorema*, *demonstração*, *metateorema*, *premissa*, *hipótese*, *argumento*, *axioma*, *postulado*, entre outros comumente empregados na literatura especializada.

viii: É explicitado o poder da pertinência em teoria de conjuntos, mostrando como esse conceito consegue qualificar números *naturais*, *inteiros*, *racionais*, *irracionais*, *reais*, *algébricos*, *transcendentes* e *complexos*, bem como fundamentar cálculo diferencial e integral, geometria euclidiana, geometria analítica, álgebra linear, espaços métricos e espaços de probabilidades. Ou seja, números reais são conjuntos, números naturais são conjuntos, pontos no plano euclidiano são conjuntos, vetores são conjuntos, matrizes são conjuntos, relações são conjuntos, funções são conjuntos, espaços amostrais são conjuntos, probabilidades são conjuntos etc.

ix: A definição de Carathéodory para função real diferenciável é abordada aqui, mas na forma de um teorema.

x: É salientado o papel de *verdade* e, consequentemente, *falsidade*, em matemática.

xi: Mostra-se claramente por que uma probabilidade condicional não é uma probabilidade.

xii: É discutido o emprego de nomes em matemática.

Além disso, muitos exemplos são dados e detalhadamente justificados, para fins de ilustração.

Mas o principal diferencial deste texto é o tom provocativo. Como disse o matemático britânico Bertrand Russell [42],

> *Matemática pode ser definida como o assunto no qual jamais sabemos sobre o que estamos falando ou se o que estamos dizendo é verdadeiro ou não.*

Seção 5

Metodologia

Sumário
Índice
Rede

O método aqui adotado para exposição dos temas é inspirado na Teoria das Histórias, como apresentada por Robert McKee [37] (apesar de Aristóteles, em sua obra *Poética*, já ter se ocupado do tema dois milênios atrás). Na visão de McKee, a maioria das histórias de cinema, teatro, televisão e literatura conta com a estrutura de uma *arquitrama* dividida em três atos. No primeiro ato a personagem principal é apresentada, com suas características inerentes e seu atual estado. No segundo ato algo acontece com a personagem principal, exercendo pressão sobre ela. Essa pressão deve revelar o caráter da personagem principal. Caráter, por definição, é a forma como alguém reage diante de pressão. Pressão se refere a eventos que antagonizam com as características inerentes e o estado em que se encontra a personagem. Finalmente, no terceiro ato deve ocorrer a resolução da história, ou seja, a solução final que a personagem apresenta para a pressão iniciada no segundo ato.

Nossa personagem principal aqui é a teoria de conjuntos ZF. Neste contexto, o primeiro ato consiste nas Partes 2 e 3. A partir da Parte 4 exercemos pressão sobre ZF, para avaliar sua capacidade de lidar com as práticas matemáticas necessárias para o cotidiano de matemáticos, físicos, engenheiros e demais interessados em usá-la. Com relação ao terceiro ato, este ainda está em andamento na aventura humana que busca conhecer ZF e outras formas de fundamentos da matemática que complementam ou até antagonizam ZF.

Neste contexto, matemática é tratada aqui não como uma arquitrama, mas uma *minitrama*, na acepção de McKee em sua grande obra. Detalhes podem ser avaliados pelo leitor no livro citado.

Um dos erros mais graves no ensino de matemática é o foco sobre alunos e professores, uma vez que este foco deveria estar direcionado à matemática. Com efeito, em narrativas não interessa quem está narrando ou acompanhando a história, mas apenas a história. No entanto, essa é uma extensa discussão que não é contemplada aqui.

Não obstante, a metodologia aqui adotada pode ser facilmente mal interpretada. Isso porque o leitor pode ficar com a impressão de que matemática é edificada a partir de conceitos básicos (como ZF), na direção de conceitos mais sofisticados (como toda a matemática que pode ser fundamentada em ZF). Porém, não é assim que matemática (ou até mesmo o estudo de matemática) funciona.

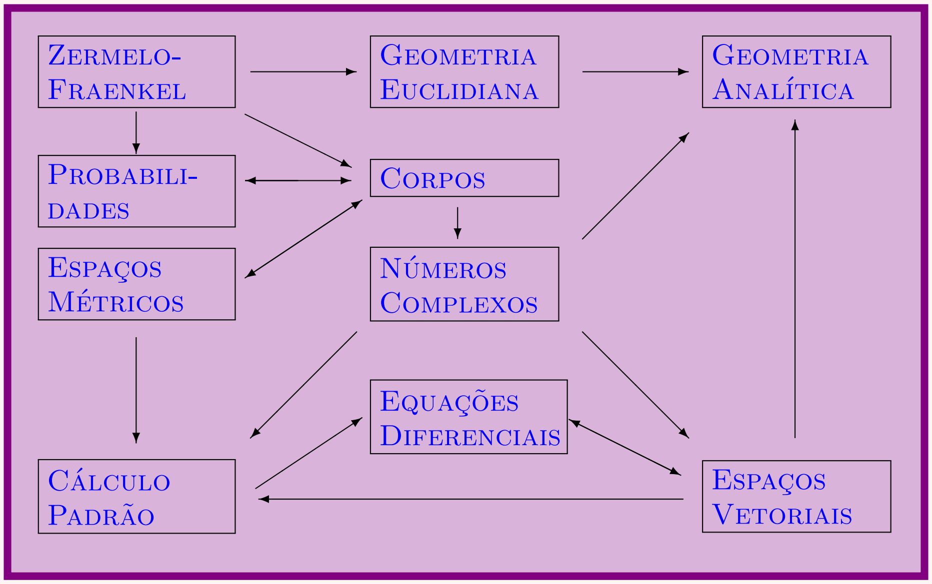
Assim como as arquitramas de McKee contam com tramas paralelas que se emaranham com a história principal, conferindo uma dinâmica que opera em rede sobre a personagem principal, algo análogo acontece com a prática matemática.

Consideremos, para fins de ilustração, o caso do cálculo diferencial e integral padrão, assunto tratado nesta obra. Para demonstrarmos certos resultados sobre funções trigonométricas aplicadas a números reais, é altamente conveniente conhecermos números complexos. A teoria ZF qualifica claramente o conceito de número complexo. No entanto, limites, derivadas e integrais de funções trigonométricas só podem ser definidos no contexto de espaços métricos, assunto este que pode ser qualificado em ZF sem sabermos o que são números complexos. Logo, devemos necessariamente estudar espaços métricos antes de limites, derivadas e integrais de funções trigonométricas? A resposta é claramente negativa, pelo menos do ponto de vista de opções disponíveis para o aprendizado de matemática. Basta examinarmos a literatura padrão de cálculo diferencial e integral. Na maioria dos livros jamais são conceituados números reais ou complexos, e nem mesmo espaços métricos.

Este é o principal problema no estudo de matemática a partir de livros. Livros apresentam conteúdos que são lidos sequencialmente, da página n para a página $n+1$, como se a matemática pudesse ser conhecida de maneira linear, lendo página por página e fazendo exercícios.

Matemática, porém, não conta com qualquer estrutura hierárquica de pré-requisitos que permita avançar do básico a um nível avançado, passando em algum momento por temas de nível intermediário de sofisticação. Uma pessoa pode conhecer muito bem aspectos profundos de equações diferenciais sem se dar conta dos axiomas de ZF que sustentam a matemática de equações diferenciais.

Por conta disso, apresentamos a seguir uma estrutura em rede dos temas abordados neste texto.



Estrutura em rede dos assuntos deste livro

O leitor pode julgar seus estudos deste livro como bem sucedidos se, ao término da leitura e da solução dos exercícios propostos, puder avaliar criticamente a rede ilustrada na imagem acima.

A rede acima representada conta com dez nós, os quais correspondem aos principais assuntos aqui tratados. As flechas sugerem a influência de um nó sobre outros. Neste contexto, o nó Zermelo-Fraenkel é o único do qual apenas partem flechas e nenhuma flecha chega até ele. Neste sentido, a rede estabelece que ZF é o ponto de partida para definir os demais nós.

Digamos que alguém levante a seguinte questão:

Equação de reta, no plano cartesiano, é dada por definição ou teorema?

A resposta depende do contexto em que equação de reta é apresentada em geometria analítica. Na Seção 77 a equação de reta é usada para *definir* retas quaisquer em um *modelo* de geometria euclidiana plana conhecido como plano cartesiano. Neste sentido o nó Geometria Analítica da rede acima é construído a partir dos nós Geometria Euclidiana e Números Complexos. Porém, no estudo de modelos de espaços vetoriais, a equação de reta surge como *teorema* (Teorema 8.30). Logo, o nó Geometria Analítica da rede acima é construído a partir dos nós Geometria Euclidiana, Números Complexos e Espaços Vetoriais.

Apenas para citar mais um exemplo, o nó Números Complexos corresponde não apenas ao estudo dos números complexos, mas também à investigação de diversos conjuntos que os complexos são capazes de 'copiar', como os reais, os racionais, os irracionais, os inteiros e os naturais. Neste contexto, as flechas informam que ZF permite definir *corpos*, assim como também permite conceituar naturais, inteiros, racionais e reais, além dos complexos. Uma vez definidos os reais, as flechas indicam que ZF permite qualificar espaços métricos. Uma vez que os nós Espaços Métricos e Corpos são conectados por uma flecha de duplo sentido, isso mostra que o estudo de espaços métricos permite uma compreensão mais ampla sobre os próprios reais (casos particulares de corpos) usados para defini-los.

De forma alguma está sendo sugerido que a rede acima corresponde à maneira como matemática deve ser tratada. Trata-se apenas de uma visão resumida sobre como os assuntos aqui explorados estão conectados entre si. Por exemplo, se houvesse neste livro alguma discussão sobre espaços métricos probabilísticos, haveria na rede acima uma flecha conectando os nós Espaços Métricos e Probabilidades.

Levando em conta o alerta já feito sobre a linearidade imposta pelo formato 'livro', deve ficar evidente ao leitor o grande desafio que é a apresentação desta rede no formato de um livro. Este é um dos grandes desafios que autores enfrentam. Por consequência, quem deve pagar o esforço final é evidentemente o leitor. Ou seja, não é fácil estudar matemática.

Seção 6

Signos usados neste livro

Sumário
Índice
Rede

Nesta obra utilizamos alguns signos para destacar certas partes do texto.

destaca exercícios recomendados ao leitor.

destaca informações que podem ser obtidas consultando outras fontes.

destaca endereço eletrônico (e-mail).

destaca informações históricas.

destaca Seção que pode ser ‘cortada’ (ignorada) sem prejuízo óbvio ao restante da leitura.

destaca que a leitura deve ser interrompida para fins de reflexão.

Também usamos retângulos coloridos para contrastar certos trechos do livro. A meta é facilitar ao leitor a eventual busca por definições, exemplos, proposições e axiomas. Levando em conta que não empregamos qualquer notação para sinalizar conclusão de demonstrações e provas, essas caixas coloridas devem auxiliar na imediata localização dos pontos onde começam e terminam as argumentações que justificam teoremas e proposições.

Definições são posicionadas em retângulos cinzas.

Exemplos são posicionados em retângulos azuis.

Proposições e teoremas são posicionados em retângulos verdes.

Provas e demonstrações são posicionadas em retângulos com outro tom de verde.

Axiomas próprios de ZFC são posicionados em retângulos amarelos.

PARTE 2

Linguagem e lógica

$\aleph$

$=$ $\in$

Nesta segunda parte qualificamos e desenvolvemos uma linguagem formal específica e uma lógica tradicionalmente conhecida como *lógica clássica*. Para facilitar a visão intuitiva dos conceitos aqui explorados, promovemos analogias com noções elementares sobre ciência da computação.

Seção 7 Linguagem $\mathfrak{S}$

Sumário
Índice
Rede

Os conteúdos aqui discutidos sobre linguagens formais e lógica são uma adaptação da famosa obra de Elliott Mendelson [38]. No entanto, no livro citado o autor não discute sobre a Teoria de Zermelo-Fraenkel (ZF). Por motivos profissionais e pessoais, Mendelson optou tratar do sistema de von Neumann-Bernays-Gödel (NBG).

Ambas ZF e NBG são teorias formais amplamente conhecidas, apesar de ZF ser obviamente mais popular. Referências interessantes aos axiomas de ZF são [28] e [8]. O livro de Thomas Jech [28] não é adequado a iniciantes, mas é perfeito para quem já tem familiaridade com teorias formais e deseja conhecer com alguma profundidade teoria de modelos, incluindo universos de von Neumann, conjuntos construtíveis, modelos de permutações e *forcing*. Já a obra de Tim Button [8] é dirigida a filósofos da matemática, com a vantagem de tratar também de ZF_2, ou seja, a versão de segunda ordem de Zermelo-Fraenkel.

ZF é uma *teoria formal axiomática*. Toda teoria formal axiomática exige uma linguagem formal e uma lógica. Nesta Seção tratamos da linguagem de ZF.

Toda linguagem demanda um *vocabulário*, ou seja, uma coleção de símbolos. Chamamos a linguagem da teoria ZF de $\mathfrak{S}$ (letra S na fonte Fraktur).

O vocabulário de $\mathfrak{S}$ é formado pelos seguintes símbolos:

- *Variáveis*: x_1, x_2, x_3, $\cdots$. Eventualmente variáveis podem ser abreviadas por letras latinas minúsculas em itálico como x, y, z, r, s, $\cdots$ ou até mesmo letras gregas minúsculas como α (alfa), β (beta), γ (gama), δ (delta), ε (épsilon), $\cdots$, por uma questão de mera conveniência. Outros símbolos podem ser empregados para designar variáveis, desde que seja explicitado ser o caso.
- *Constantes*: c_1, c_2, c_3, $\cdots$. Eventualmente constantes podem ser abreviadas por símbolos especiais, conforme a conveniência. Exemplos que são explorados ao longo do texto são os símbolos $\varnothing$ (vazio), $\{\varnothing\}$ (unitário vazio), $\{\{\varnothing\}\}$ (unitário unitário vazio), ω (omega), $\{\omega\}$ (unitário omega), entre muitos outros. Observar que as letras gregas π e ω ocupam uma posição privilegiada entre as constantes. São as únicas letras gregas não usadas aqui para denotar variáveis.
- Dois *predicados binários*: $=$ (igualdade) e $\in$ (pertinência).
- Cinco *conectivos lógicos*: $\neg$ (negação), $\wedge$ (conjunção), $\vee$ (disjunção), $\Rightarrow$ (condicional), $\Leftrightarrow$ (bicondicional).
- Dois *quantificadores lógicos*: $\forall$ (universal) e $\exists$ (existencial).
- Dois *símbolos auxiliares*: (,), chamados de 'abre parênteses' e 'fecha parênteses', respectivamente.

Os símbolos especiais usados para certas constantes são conceituados oportunamente neste documento.

Vale a pena notar que, diferentemente da língua portuguesa, a linguagem formal $\mathfrak{S}$ emprega um vocabulário com uma infinidade de símbolos.

Uma *sentença* de $\mathfrak{S}$ é qualquer sequência finita de símbolos do vocabulário de $\mathfrak{S}$.

Exemplo 2.1. i: *$x_4x_1x_4\forall()c_{30}\neg$ é uma sentença de $\mathfrak{S}$; com efeito, esta é uma sequência com oito ocorrências de símbolos de $\mathfrak{S}$;*

ii: *$x_1 = x_1$ é uma sentença de $\mathfrak{S}$; com efeito, esta é uma sequência com três ocorrências de símbolos de $\mathfrak{S}$;*

iii: *$x_2\mathfrak{S}\neg$ não é uma sentença de $\mathfrak{S}$; com efeito, o símbolo $\mathfrak{S}$ não está na lista de símbolos da linguagem $\mathfrak{S}$.*

Intuitivamente falando, pedimos ao leitor para imaginar um teclado de computador com infinitas teclas, uma para cada símbolo de $\mathfrak{S}$. Uma sentença qualquer de $\mathfrak{S}$ pode ser escrita digitando aleatoriamente esse teclado, sem atenção alguma além de digitar apenas as teclas. No momento em que a digitação encerrar, teremos então uma sentença de $\mathfrak{S}$.

Observar que a linguagem formal $\mathfrak{S}$ aqui edificada é a *linguagem-objeto*, no sentido de ser uma linguagem sobre a qual está sendo dito algo a respeito. No entanto, está sendo empregada uma outra linguagem para falar a respeito de $\mathfrak{S}$. Essa outra linguagem é o que se chama de *metalinguagem*. Este mesmo parágrafo foi escrito na metalinguagem usada aqui para discutir sobre a linguagem-objeto $\mathfrak{S}$. Neste contexto, a linguagem-objeto $\mathfrak{S}$ é uma linguagem formal, enquanto a metalinguagem usada para tratar de $\mathfrak{S}$ não é. Logo, expressões até aqui empregadas para descrever $\mathfrak{S}$, como 'vocabulário', 'coleção de símbolos', 'sequência finita de símbolos', entre outros, são termos metalinguísticos com significados implicitamente assumidos. A metalinguagem aqui usada não é formal. Logo, mesmo o estudo de linguagens formais da matemática exige o emprego de linguagens que não são formais.

Como foi dito na Seção 1, a língua portuguesa é insuficiente para fazer matemática, a qual demanda linguagens formais que prescindem de significados. No entanto, sem uma linguagem não formal como português, inglês ou francês, não parece ser possível compreender algo como $\mathfrak{S}$.

O próximo passo para edificar $\mathfrak{S}$ é estabelecer quais sentenças de $\mathfrak{S}$ são *fórmulas*. O papel de fórmulas é explicitar uma *sintaxe* para $\mathfrak{S}$. Com efeito, matemáticos não estão interessados em sentenças quaisquer de $\mathfrak{S}$, ou seja, sequências quaisquer dos símbolos que constituem o vocabulário de $\mathfrak{S}$.

Se fizermos uma analogia entre $\mathfrak{S}$ e uma linguagem de programação de computadores, não basta conhecermos os símbolos da última, se quisermos efetivamente criar um programa de computador. É necessário conhecermos a sintaxe da linguagem. Caso contrário, qualquer tentativa de fazer o programa funcionar fracassará, se houver algum erro de sintaxe. Ou seja, apenas digitar aleatoriamente símbolos de uma linguagem de programação não produz necessariamente um programa de computador.

Mas, antes de estabelecer a sintaxe de $\mathfrak{S}$, é necessário qualificar o que é um *termo* de $\mathfrak{S}$. Isso porque fórmulas devem agir como 'afirmações a respeito de termos'. Segue abaixo.

Variáveis e constantes são os únicos *termos* de $\mathfrak{S}$.

Ou seja, os únicos símbolos de $\mathfrak{S}$ chamados de termos são as variáveis e as constantes.

Exemplo 2.2. i: *x_3 é um termo de $\mathfrak{S}$, uma vez que x_3 é uma variável;*

ii: *$\forall$ não é um termo de $\mathfrak{S}$; com efeito, o símbolo $\forall$ não é variável e nem constante.*

iii: *$=$ não é um termo de $\mathfrak{S}$; com efeito, o símbolo $=$ não é variável e nem constante.*

Os objetos de estudo de ZF são, neste primeiro momento, os termos de $\mathfrak{S}$, ou seja, variáveis e constantes de $\mathfrak{S}$. Na literatura especializada tais termos são comumente chamados de *conjuntos*, pelo menos no contexto da linguagem $\mathfrak{S}$ que está sendo construída aqui. Mas o leitor deve ser advertido. Uma vez que $\mathfrak{S}$ não é comprometida com qualquer contraparte semântica, os termos de $\mathfrak{S}$ não devem ser interpretados como 'conjuntos' nas acepções usualmente empregadas na língua portuguesa. A terminologia *conjunto* é tão somente um nome para os termos de $\mathfrak{S}$, livre de significado. Comentário análogo vale para os demais símbolos de $\mathfrak{S}$ (aqueles que não são termos, como $\forall$, $\neg$, $\Rightarrow$ e outros).

Retornando à analogia com programas de computador, estes também empregam linguagens desprovidas de significado. Por conta disso, uma mesma linguagem de programação pode ser concebida para criar desde jogos eletrônicos, para fins de entretenimento, até *softwares* que gerenciam estoques de supermercados.

Agora podemos finalmente introduzir a sintaxe de $\mathfrak{S}$:

I Se u e v são termos de $\mathfrak{S}$, então as sentenças

$$u = v$$

e

$$u \in v$$

são *fórmulas atômicas* de $\mathfrak{S}$.

II Toda fórmula atômica de $\mathfrak{S}$ é *fórmula* de $\mathfrak{S}$.

III Se $\mathcal{A}$ e $\mathcal{B}$ são fórmulas de $\mathfrak{S}$ e u é uma variável, então as sentenças $\neg(\mathcal{A})$, $(\mathcal{A} \wedge \mathcal{B})$, $(\mathcal{A} \vee \mathcal{B})$, $(\mathcal{A} \Rightarrow \mathcal{B})$, $(\mathcal{A} \Leftrightarrow \mathcal{B})$, $\forall u(\mathcal{A})$ e $\exists u(\mathcal{A})$ são fórmulas de $\mathfrak{S}$.

IV Apenas as sentenças de $\mathfrak{S}$ que seguem os itens acima são fórmulas de $\mathfrak{S}$.

As fórmulas atômicas $u = v$ e $u \in v$ se lêem, respectivamente, 'u é igual a v' (ou, 'u é idêntico a v') e 'u pertence a v' (ou 'u é elemento de v').

A fórmula $\neg(\mathcal{A})$ se lê 'não $\mathcal{A}$' ou 'a negação de $\mathcal{A}$'.

$(\mathcal{A} \wedge \mathcal{B})$ se lê '$\mathcal{A}$ e $\mathcal{B}$'.

$(\mathcal{A} \vee \mathcal{B})$ se lê '$\mathcal{A}$ ou $\mathcal{B}$'.

$(\mathcal{A} \Rightarrow \mathcal{B})$ se lê 'se $\mathcal{A}$, então $\mathcal{B}$' (ou '$\mathcal{A}$ implica em $\mathcal{B}$').

$(\mathcal{A} \Leftrightarrow \mathcal{B})$ se lê '$\mathcal{A}$ se, e somente se, $\mathcal{B}$' ou '$\mathcal{A}$ é equivalente a $\mathcal{B}$'.

$\forall u(\mathcal{A})$ se lê 'para todo u, $\mathcal{A}$'.

$\exists u(\mathcal{A})$ se lê 'existe u tal que $\mathcal{A}$'.

Fórmulas de $\mathfrak{S}$ que não são fórmulas atômicas são *fórmulas moleculares*.

Na fórmula molecular $\forall x(\mathcal{A})$ dizemos que $\mathcal{A}$ é o *escopo* do quantificador universal $\forall x$. Além disso, x tem *ocorrência ligada* em $\forall x(\mathcal{A})$. Qualquer ocorrência de x no escopo de $\forall x(\mathcal{A})$ também é ligada.

Variáveis que não são ligadas são variáveis de *ocorrências livres*.

Exemplo 2.3. i: *A fórmula $\forall x(x = y)$ tem duas ocorrências ligadas de x e uma ocorrência livre de y;*

ii: *a fórmula $\forall y(x = y)$ tem duas ocorrências ligadas de y e uma ocorrência livre de x;*

iii: *na fórmula $\forall x(\forall y(x = y))$ todas as ocorrências de x e de y são ligadas;*

iv: *na fórmula $x = y$ todas as ocorrências de x e de y são livres.*

Observar que os símbolos $\mathcal{A}$ e $\mathcal{B}$, na definição de fórmulas, são abreviações metalinguísticas de fórmulas de $\mathfrak{S}$, uma vez que $\mathcal{A}$ e $\mathcal{B}$ não fazem parte do vocabulário de $\mathfrak{S}$.

É usual se referir à fórmula atômica $x = y$ como *equação*. Neste sentido, toda equação é tão somente um caso particular de fórmula atômica.

A sintaxe de $\mathfrak{S}$ deixa claro que toda fórmula de $\mathfrak{S}$ é uma sentença de $\mathfrak{S}$, mas nem toda sentença de $\mathfrak{S}$ é uma fórmula de $\mathfrak{S}$.

De agora em diante, para fins de abreviação, fórmulas de $\mathfrak{S}$ e sentenças de $\mathfrak{S}$ são chamadas simplesmente de fórmulas e sentenças, respectivamente.

Exemplo 2.4. *As seguintes sentenças são fórmulas:*

i: $\forall x(x = x)$;

ii: $\exists x(\forall y(\neg(y \in x)))$;

iii: $\forall x(\neg(x = x))$.

Observar que, nos exemplos acima, estão sendo empregadas as abreviações usuais para variáveis.

Justificando item ii do Exemplo acima: x e y abreviam variáveis; logo, x e y são termos; logo, item i da sintaxe de $\mathfrak{S}$ garante que $y \in x$ é fórmula atômica; logo, item ii da sintaxe de $\mathfrak{S}$ garante que

$y \in x$ é fórmula; logo, item III da sintaxe de $\mathfrak{S}$ garante que $\neg(y \in x)$ é fórmula; logo, item III da sintaxe de $\mathfrak{S}$ garante que $\forall y(\neg(y \in x))$ é fórmula; logo, item III da sintaxe de $\mathfrak{S}$ garante que $\exists x(\forall y(\neg(y \in x)))$ é fórmula.

Todas as ocorrências de x e de y nos três itens do Exemplo 2.4 são ligadas.

O leitor deve ter observado que a sintaxe de $\mathfrak{S}$ é uma definição recursiva de fórmula, no sentido de que o item III pode ser aplicado quantas vezes forem necessárias para verificar se uma sentença é fórmula. O critério de parada dessa definição recursiva é garantido pelo fato de que toda sentença de $\mathfrak{S}$ deve ser uma sequência finita de símbolos de $\mathfrak{S}$.

Exemplo 2.5. *As seguintes sentenças não são fórmulas:*

I: $= x \Rightarrow$;

II: $\exists\forall(x = x)$;

III: $(x = y)$;

IV: $\exists x \wedge \exists y(x = y)$.

Justificando item I do último exemplo: o predicado binário $=$ exige as ocorrências de um termo imediatamente à esquerda e um termo imediatamente à direita de $=$. Porém, não há qualquer ocorrência de termo à esquerda de $=$.

Justificando item II: apesar de $x = x$ ser fórmula, a sentença $\exists\forall(x = x)$ não é uma fórmula, uma vez que o item III da definição de fórmula exige a ocorrência de uma variável imediatamente à direita do quantificador universal $\forall$ e de uma variável imediatamente à direita do quantificador existencial $\exists$.

Justificando item III: Apesar de $x = y$ ser fórmula atômica (e, portanto, fórmula), item IV da definição de fórmula garante que $(x = y)$ não é fórmula. Com efeito, se $\mathcal{A}$ é fórmula, então $(\mathcal{A})$ não é fórmula.

Justificando item IV: Apesar de x ser uma abreviação para uma variável, item III da definição de fórmula exige que $\exists x$ seja seguido imediatamente à direita pela sentença $(\mathcal{A})$, onde $\mathcal{A}$ é uma fórmula; no entanto não é o que acontece com a sentença $\wedge\exists y(x = y)$; logo, item IV da definição de fórmula garante que essa sentença não é fórmula.

Seção 8

Definindo definições

Sumário
Índice
Rede

Uma prática comum é o emprego de *definições explícitas abreviativas* na linguagem $\mathfrak{S}$. Apesar de tais definições explícitas abreviativas serem matematicamente desnecessárias, elas são extremamente úteis para facilitar a escrita e a leitura de abreviações metalinguísticas de fórmulas. Um dos talentos exigidos de qualquer matemático é a capacidade de ler, escrever e usar fórmulas, bem como refletir e discutir sobre elas.

Em [41] Alonzo Church ressalta que definições servem ao propósito de introduzir novas notações, por uma questão de mera conveniência. Neste sentido, existem vários tipos de definições. Detalhes em [43]. Mas as definições mais frequentemente empregadas neste texto são as *explícitas abreviativas*.

Uma *definição explícita abreviativa* em $\mathfrak{S}$ é uma sentença metalinguística da forma

$$definiendum \vdots definiens.$$

O símbolo $\vdots$ é uma notação metalinguística cujo propósito é separar o *definiendum* (termo a ser definido) do *definiens* (fórmula da linguagem $\mathfrak{S}$ que qualifica o que o *definiendum* está abreviando).

Observar que não há qualquer circularidade envolvida na definição de definições explícitas em $\mathfrak{S}$. Isso porque definimos na metalinguagem o que são abreviações metalinguísticas da linguagem formal $\mathfrak{S}$. Temos, dessa forma, mais um exemplo das virtudes de discriminação entre linguagem-objeto e metalinguagem.

Seguem dois exemplos de definições explícitas que são usadas com frequência aqui e no restante da literatura especializada:

$$x \neq y \vdots \neg(x = y)$$

$$x \notin y \vdots \neg(x \in y)$$

Ou seja, apesar de $x \neq y$ não ser uma fórmula de $\mathfrak{S}$, é uma abreviação metalinguística para a fórmula $\neg(x = y)$. Comentário análogo vale para $x \notin y$.

Vale a pena observar que a sequência de símbolos $x \neq y$ tem *comprimento* 3, no sentido de que há três ocorrências de símbolos nela. Em contrapartida, o *definiens* correspondente tem comprimento 6. Com efeito, a fórmula $\neg(x = y)$ conta com seis ocorrências de símbolos, o dobro do *definiendum*. Comentário análogo vale para a definição de $x \notin y$. Isso deixa patente a grande vantagem do emprego de definições explícitas abreviativas: economia para a escrita de fórmulas. Desenvolver um tema como cálculo diferencial e integral sem o uso de definições torna o assunto intelectualmente indigesto, exaustivo e não produtivo. Justamente por isso muitas outras definições são introduzidas ao longo de todo este texto.

É uma prática comum se referir a abreviações metalinguísticas como fórmulas, desde que sejam definidas nos moldes acima. Essa prática é o que se chama de *abuso de linguagem*. Apesar de abuso de linguagem não ser justificável formalmente (uma vez que abreviações metalinguísticas não fazem parte do vocabulário de $\mathfrak{S}$), ela facilita a discussão sobre aspectos formais da matemática. Matemáticos em geral não perdem tempo com formalismo. Mas é indispensável o rigor. *Rigor*, neste contexto, significa 'a capacidade de reescrever abreviações metalinguísticas como fórmulas da linguagem $\mathfrak{S}$'. *Formalismo* significa 'escrever apenas de acordo com a sintaxe da linguagem formal, sem o emprego de abreviações metalinguísticas'.

Eventualmente podemos substituir o símbolo metalinguístico $\vdots$ pelo símbolo metalinguístico 'sss' (abreviação para 'se, e somente se,'). Logo, as definições acima para $\neq$ e $\notin$ poderiam ter sido escritas também como

$$x \neq y \text{ sss } \neg(x = y)$$

e

$$x \notin y \text{ sss } \neg(x \in y).$$

Apelando novamente à nossa analogia da linguagem $\mathfrak{S}$ com uma linguagem de programação, definições explícitas abreviativas funcionam como 'sub-rotinas' que 'chamam' uma fórmula a partir de um 'nome'. Ou seja, a 'sub-rotina' $x \neq y$ apenas 'chama' a fórmula $\neg(x = y)$, toda vez que ela ocorre em uma sentença.

Muitos abusos de linguagem ocorrem com frequência em cálculo diferencial e integral e álgebra linear, temas que aqui estudamos. Por conta disso, não são raros os alunos que se sentem confusos, por conta de tais abusos. Exemplo clássico é a afirmação de que

$$\lim_{x\to a} f(x) = \infty,$$

em certas situações envolvendo funções reais.

A experiência mostra que muitos alunos tratam o símbolo ∞ como um termo. Mas este não é o caso. Discutimos em detalhes sobre o assunto na Seção 45. Um maior detalhamento sobre definições é encontrado na Seção 14.

✍ Escrever novos exemplos de duas sentenças de $\mathfrak{S}$ que são fórmulas e de duas sentenças que não são fórmulas. Cada exemplo deve ser justificado de forma circunstanciada.

Seção 9

Lógica

Sumário
Índice
Rede

A *lógica* da teoria de conjuntos ZF é definida por *axiomas* e *regras de inferência dedutiva*.

Axiomas (também conhecidos como *postulados*) de ZF são fórmulas selecionadas para compor a *lista de axiomas* de ZF. Neste contexto, 'fórmulas de $\mathfrak{S}$' e 'fórmulas de ZF' são tratados aqui como sinônimos.

Apesar da aparente circularidade no conceito de axioma, o fato é que um axioma é tão somente uma fórmula que faz parte da lista de axiomas de ZF. Obviamente a lista de axiomas de ZF poderia, em princípio, ser dada pela totalidade de fórmulas de $\mathfrak{S}$. Mas, neste caso, ZF seria uma teoria formal inútil, conforme o leitor deve perceber mais adiante. Ou seja, todo axioma de ZF é uma fórmula de ZF, mas nem toda fórmula de ZF é um axioma de ZF. Apesar disso, o leitor também perceberá que a lista de axiomas de ZF é formada por uma quantia não finita de fórmulas de $\mathfrak{S}$.

Uma *regra de inferência dedutiva* (ou *argumento dedutivo*) R é uma relação

$$R(\mathcal{F}_1, \mathcal{F}_2, \mathcal{F}_3, \cdots, \mathcal{F}_n)$$

entre n fórmulas $\mathcal{F}_1$, $\mathcal{F}_2$, $\mathcal{F}_3$, $\cdots$, $\mathcal{F}_n$, de modo que $\mathcal{F}_n$ é única.

Por abuso de linguagem nos referimos a regras de inferência dedutiva simplesmente como *regras de inferência* ou *argumentos*.

Qualquer regra de inferência $R(\mathcal{F}_1, \mathcal{F}_2, \mathcal{F}_3, \cdots, \mathcal{F}_n)$ se lê '$\mathcal{F}_n$ é *consequência imediata* de $\mathcal{F}_1$, $\mathcal{F}_2$, $\cdots$, $\mathcal{F}_{n-1}$ via R'. No caso particular em que $n = 3$, dizemos que R é um *silogismo*.

Observar que qualquer regra de inferência R em ZF é um símbolo que não faz parte do vocabulário de ZF. Linguagens não têm o poder de promover inferências.

A lista de axiomas de ZF é dividida em dois grupos de fórmulas:

axiomas lógicos e *axiomas próprios*.

Tal lista de axiomas lógicos e axiomas próprios de ZF é apresentada a partir de alguns parágrafos abaixo.

Seguindo a analogia com programação de computadores, os axiomas lógicos e as regras de inferência de ZF operam como um 'sistema operacional', enquanto os axiomas próprios de ZF funcionam como um *software* especializado que é executado sob a gerência do sistema operacional. Por conta disso, podemos considerar variações de ZF (acrescentando, omitindo ou até alterando axiomas próprios), mas sempre sob os mesmos princípios ditados pelos axiomas lógicos e pelas regras de inferência.

A teoria de conjuntos ZF conta com apenas duas regras de inferência:

Modus Ponens (abreviada como M)
e
Generalização (abreviada como G).

Se $\mathcal{P}$ e $\mathcal{Q}$ são fórmulas e x é uma variável, então

$$M(\mathcal{P}, (\mathcal{P} \Rightarrow \mathcal{Q}), \mathcal{Q}) \ \text{ e } \ G(\mathcal{P}, \forall x(\mathcal{P})).$$

Neste caso Modus Ponens M se lê como '$\mathcal{Q}$ é consequência imediata de $\mathcal{P}$ e de $(\mathcal{P} \Rightarrow \mathcal{Q})$'. Logo, Modus Ponens é um exemplo de silogismo, uma vez que é uma regra de inferência que envolve três ocorrências de fórmulas. Generalização G, por sua vez, se lê como '$\forall x(\mathcal{P})$ é consequência imediata de $\mathcal{P}$'. Logo, Generalização não é um silogismo, uma vez que é um argumento que envolve apenas duas ocorrências de fórmulas.

Do ponto de vista intuitivo, regras de inferência permitem *deduzir* novas fórmulas a partir de fórmulas anteriores. No caso de Modus Ponens, é possível inferir a nova fórmula $\mathcal{Q}$ a partir das fórmulas $\mathcal{P}$ e $(\mathcal{P} \Rightarrow \mathcal{Q})$. Desta maneira Modus Ponens confere um caráter dedutivo ao conectivo condicional $\Rightarrow$. Isso justifica a leitura de $(\mathcal{P} \Rightarrow \mathcal{Q})$ como 'se $\mathcal{P}$, então $\mathcal{Q}$'.

A importância de axiomas e regras de inferência na edificação de ZF é discutida na próxima Seção. Por enquanto basta dizer que axiomas e regras de inferência são indispensáveis para ZF. Isso porque ZF deve expressar ideias a partir de seus axiomas e permitir inferir novas ideias não explicitamente expressas por seus axiomas.

Se $\mathcal{A}$, $\mathcal{B}$ e $\mathcal{C}$ são fórmulas de $\mathfrak{S}$, então os axiomas lógicos de ZF são as seguintes fórmulas:

L1 $(\mathcal{A} \Rightarrow (\mathcal{B} \Rightarrow \mathcal{A}))$;

L2 $((\mathcal{A} \Rightarrow (\mathcal{B} \Rightarrow \mathcal{C})) \Rightarrow ((\mathcal{A} \Rightarrow B) \Rightarrow (\mathcal{A} \Rightarrow \mathcal{C})))$;

L3 $(((\neg\mathcal{B}) \Rightarrow (\neg\mathcal{A})) \Rightarrow (((\neg\mathcal{B}) \Rightarrow \mathcal{A}) \Rightarrow \mathcal{B}))$;

L4 $\forall x(\mathcal{A}(x) \Rightarrow \mathcal{A}(t))$, se t for um *termo livre para x em $\mathcal{A}(x)$*, ou seja, nenhuma ocorrência livre de x em $\mathcal{A}(x)$ está no escopo de quantificador $\forall y$ com y ocorrendo em t;

L5 $(\forall x(\mathcal{A} \Rightarrow \mathcal{B}) \Rightarrow (\mathcal{A} \Rightarrow \forall x(\mathcal{B})))$, se x não tem ocorrências livres em $\mathcal{A}$.

Os axiomas lógicos L1 e L2 estabelecem como deve 'funcionar' o conectivo condicional $\Rightarrow$. Por exemplo, axioma L1 diz, intuitivamente falando, que 'se temos a fórmula $\mathcal{A}$, então qualquer fórmula $\mathcal{B}$ implica em $\mathcal{A}$'. Axioma L3 estabelece relações entre os conectivos condicional $\Rightarrow$ e negação $\neg$.

Os demais conectivos lógicos (conjunção $\wedge$, disjunção $\vee$ e bicondicional $\Leftrightarrow$) não têm ocorrência alguma entre os axiomas lógicos de ZF por conta de um fato muito simples: são matematicamente desnecessários. Com efeito, poderíamos ter definido tais conectivos como se segue:

I: $(\mathcal{A} \wedge \mathcal{B}) \vdots \neg(\mathcal{A} \Rightarrow \neg(B))$;

II: $(\mathcal{A} \vee \mathcal{B}) \vdots (\neg(\mathcal{A}) \Rightarrow \mathcal{B})$;

III: $(\mathcal{A} \Leftrightarrow \mathcal{B}) \vdots ((\mathcal{A} \Rightarrow \mathcal{B}) \wedge (\mathcal{B} \Rightarrow \mathcal{A}))$;

sendo $\mathcal{A}$ e $\mathcal{B}$ fórmulas.

Em outras palavras, $(\mathcal{A} \wedge \mathcal{B})$ ($\mathcal{A}$ e $\mathcal{B}$) equivale a dizer que 'não é o caso de $\mathcal{A}$ implicar na negação de $\mathcal{B}$'; $(\mathcal{A} \vee \mathcal{B})$ ($\mathcal{A}$ ou $\mathcal{B}$) equivale a dizer que 'a negação de $\mathcal{A}$ implica em $\mathcal{B}$'; e $(\mathcal{A} \Leftrightarrow \mathcal{B})$ ($\mathcal{A}$ se, e somente se, $\mathcal{B}$) equivale a dizer que '$\mathcal{A}$ implica em $\mathcal{B}$, e $\mathcal{B}$ implica em $\mathcal{A}$'.

Neste texto escolhemos a incorporação dos conectivos lógicos conjunção, disjunção e bicondicional ao vocabulário de $\mathfrak{S}$ por motivos meramente pedagógicos. Neste contexto, podemos acrescentar aos axiomas lógicos de ZF as seguintes fórmulas:

$$(\mathcal{A} \wedge \mathcal{B}) \Leftrightarrow \neg(\mathcal{A} \Rightarrow \neg(B));$$

$$(\mathcal{A} \vee \mathcal{B}) \Leftrightarrow (\neg(\mathcal{A}) \Rightarrow \mathcal{B});$$

e

$$(\mathcal{A} \Leftrightarrow \mathcal{B}) \Leftrightarrow ((\mathcal{A} \Rightarrow \mathcal{B}) \wedge (\mathcal{B} \Rightarrow \mathcal{A})).$$

Eventualmente pares de parênteses podem ser omitidos (se não houver risco de ambiguidade na leitura das fórmulas) desde que o rigor seja seguido, conforme discussão anterior sobre a diferença entre formalismo e rigor.

Axiomas lógicos L4 e L5 estabelecem as 'relações' entre o quantificador universal $\forall$ e o conectivo condicional $\Rightarrow$.

O quantificador existencial $\exists$ não tem ocorrência alguma entre os axiomas lógicos de ZF porque ele pode ser definido a partir do quantificador universal como se segue:

$$\exists x(\mathcal{A}) \vdots \neg(\forall x(\neg(\mathcal{A}))),$$

sendo $\mathcal{A}$ uma fórmula.

Logo, analogamente à discussão sobre os conectivos lógicos conjunção, disjunção e bicondicional, podemos acrescentar como axioma lógico de ZF a seguinte fórmula:

$$\exists x(\mathcal{A}) \Leftrightarrow \neg(\forall x(\neg(\mathcal{A}))),$$

desde que $\mathcal{A}$ seja uma fórmula.

Uma discussão mais detalhada sobre os axiomas lógicos de ZF está além dos propósitos introdutórios deste documento. Informações complementares, com uma abordagem muito didática e matematicamente rigorosa, podem ser encontradas em [38]. No entanto, na próxima Seção há uma discussão que deve ajudar o leitor a desenvolver uma visão intuitiva sobre axiomas e, particularmente, axiomas lógicos de ZF.

Finalmente, observar que, entre os axiomas lógicos de ZF, não há uma única ocorrência do predicado pertinência $\in$. Os axiomas onde ocorrem tal predicado são os axiomas próprios de ZF, os quais são discutidos a partir da Seção 18, na Parte 3 deste livro.

Seção 10

O papel de axiomas e regras de inferência

Sumário
Índice
Rede

Axiomas são casos especiais de fórmulas. Regras de inferência permitem inferir novas fórmulas a partir de fórmulas anteriores, em uma dada sequência finita de fórmulas. O princípio por trás desses conceitos consiste na seguinte proposta: obter fórmulas novas, a partir de axiomas e regras de inferência, chamadas de *teoremas*. Matemáticos são caçadores de teoremas.

Definição 2.1. *Uma* demonstração *em ZF é uma sequência finita de fórmulas*

$$\mathcal{F}_1, \mathcal{F}_2, \cdots, \mathcal{F}_n$$

de $\mathfrak{S}$ de modo que cada fórmula $\mathcal{F}_i$ dessa sequência é um axioma de ZF ou uma consequência imediata de fórmulas anteriores via o emprego de uma regra de inferência de ZF. Um teorema *$\mathcal{T}$ de ZF é a última fórmula de uma demonstração em ZF. Neste caso dizemos que $\mathcal{F}_1$, $\mathcal{F}_2$, ..., $\mathcal{F}_n$ é uma demonstração de $\mathcal{T}$ (sendo que $\mathcal{F}_n$ é a fórmula $\mathcal{T}$).*

Proposição 2.1. *Todo axioma de ZF é teorema de ZF.*

Prova: Seja $\mathcal{A}$ um axioma de ZF. Logo, a sequência finita $\mathcal{A}$ (formada por uma única fórmula) satisfaz a definição de demonstração em ZF. Como $\mathcal{A}$ é a última fórmula da sequência $\mathcal{A}$, então $\mathcal{A}$ é teorema de ZF.

Observar que a proposição acima não é um teorema de ZF, uma vez que foi formulada na metalinguagem aqui empregada para discutirmos sobre ZF. Proposições, no contexto do estudo de teorias formais, são conhecidas também como *metateoremas*.

Proposição 2.2. *Todo teorema em ZF admite infinitas demonstrações.*

Prova: Seja $\mathcal{T}$ um teorema de ZF. Logo, existe demonstração $\mathcal{F}_1$, $\mathcal{F}_2$, ..., $\mathcal{F}_n$ em ZF de modo que $\mathcal{F}_n$ é a fórmula $\mathcal{T}$. Logo, a sequência $\mathcal{F}_1$, $\mathcal{F}_2$, ..., $\mathcal{F}_n$, $\mathcal{F}_n$ também é uma demonstração de $\mathcal{T}$. Analogamente, a sequência $\mathcal{F}_1$, $\mathcal{F}_2$, ..., $\mathcal{F}_n$, $\mathcal{F}_n$, $\mathcal{F}_n$ é uma demonstração de $\mathcal{T}$. Podemos repetir esse procedimento para definir novas demonstrações de $\mathcal{T}$ quantas vezes quisermos.

Se $\mathcal{T}$ é teorema em ZF, denotamos isso como

$$\vdash_{\mathrm{ZF}} \mathcal{T}$$

ou

$$\vdash_{ZF} \mathcal{T}.$$

Se $\mathcal{T}$ não é teorema em ZF, denotamos isso como

$$\nvdash_{\mathrm{ZF}} \mathcal{T}$$

ou

$$\nvdash_{ZF} \mathcal{T}.$$

Para ilustrarmos um exemplo de demonstração não trivial, considere o seguinte enunciado.

Teorema 2.1. *Se $\mathcal{A}$ é uma fórmula de ZF, então*

$$\vdash_{ZF} (\mathcal{A} \Rightarrow \mathcal{A}).$$

Demonstração: $(\mathcal{A} \Rightarrow ((\mathcal{A} \Rightarrow \mathcal{A}) \Rightarrow \mathcal{A}))$, $((\mathcal{A} \Rightarrow ((\mathcal{A} \Rightarrow \mathcal{A}) \Rightarrow \mathcal{A})) \Rightarrow ((\mathcal{A} \Rightarrow (\mathcal{A} \Rightarrow \mathcal{A})) \Rightarrow (\mathcal{A} \Rightarrow \mathcal{A})))$, $((\mathcal{A} \Rightarrow (\mathcal{A} \Rightarrow \mathcal{A})) \Rightarrow (\mathcal{A} \Rightarrow \mathcal{A}))$, $(\mathcal{A} \Rightarrow (\mathcal{A} \Rightarrow \mathcal{A}))$, $(\mathcal{A} \Rightarrow \mathcal{A})$.

A demonstração acima é uma sequência de cinco fórmulas de ZF, separadas por vírgulas (neste caso a vírgula é um símbolo auxiliar metalinguístico).

A primeira das cinco fórmulas é o axioma L1, onde a fórmula $\mathcal{B}$ de L1 foi substituída pela fórmula $(\mathcal{A} \Rightarrow \mathcal{A})$. A segunda é o axioma L2, onde $\mathcal{B}$ foi substituída por $(\mathcal{A} \Rightarrow \mathcal{A})$ e $\mathcal{C}$ foi substituída por $\mathcal{A}$. A terceira fórmula da demonstração é consequência imediata das fórmulas dos passos 1 e 2 via Modus Ponens. A quarta é novamente o axioma L1, onde substituímos $\mathcal{B}$ por $\mathcal{A}$. Finalmente, o último passo é consequência imediata dos passos 3 e 4 via Modus Ponens.

Do ponto de vista intuitivo, o enunciado acima estabelece que toda fórmula de ZF implica nela mesma. Ou seja, se $\mathcal{A}$ é uma fórmula de ZF, então $(\mathcal{A} \Rightarrow \mathcal{A})$ é um teorema de ZF, independentemente de $\mathcal{A}$ ser teorema de ZF ou não. Por exemplo, a sentença $x = y$ é uma fórmula de ZF. Logo, $(x = y \Rightarrow x = y)$ é um teorema de ZF. Analogamente, $(x \neq y \Rightarrow x \neq y)$ é outro teorema de ZF.

Observar que a fórmula $(\mathcal{A} \Rightarrow \mathcal{A})$ não é um axioma de ZF. No entanto, é um teorema de ZF, desde que $\mathcal{A}$ seja fórmula. A meta do matemático, neste contexto, é estabelecer quais fórmulas de ZF são teoremas e quais não são.

Neste contexto, uma *conjectura* em ZF é uma fórmula $\mathcal{A}$ sobre a qual acredita-se ser um teorema (ou pelo menos algum grupo de matemáticos crê nisso), ainda que ninguém a tenha demonstrado. A partir do momento em que uma demonstração é exibida na qual $\mathcal{A}$ é a última fórmula da demonstração, tal fórmula deixa de ser uma conjectura e passa a ser um teorema.

Os fatos colocados acima justificam a afirmação anterior de que, uma versão de ZF onde todas as possíveis fórmulas são axiomas, seria inútil. Se todas as fórmulas de ZF fossem axiomas, logo, todas as fórmulas seriam teoremas. Logo, não haveria discriminação entre fórmulas que são teoremas e aquelas que não são. Logo, não haveria necessidade alguma de regras de inferência. Logo, em particular, ZF jamais poderia ser aplicada para lidar com problemas do mundo real. Com efeito, existem fenômenos que ocorrem no mundo real e aqueles que não ocorrem. Os fenômenos que ocorrem no mundo real devem ser, de algum modo, mapeados por teoremas de ZF.

Uma teoria formal como ZF não é um luxo intelectual. Há nesta teoria algo inerentemente pragmático no que se refere a potenciais aplicações tanto em matemática quanto em ciências nas quais a matemática se mostra relevante.

O fato de que nem todas as fórmulas de ZF são teoremas torna essa teoria um objeto de estudo matemático e filosófico. Por exemplo, se, em algum sentido, for possível enunciar um conceito de *verdade* (ver Seções 15 e 111), é possível provar a existência de fórmulas verdadeiras de ZF que não são teoremas?

O estudo mais detalhado dos axiomas lógicos de ZF demanda um esforço que vai muito além dos propósitos deste livro, como já foi dito acima. Por conta disso, interessa apenas saber que, se $\mathcal{P}$, $\mathcal{Q}$ e $\mathcal{R}$ forem fórmulas quaisquer (teoremas ou não), então as seguintes fórmulas são teoremas de ZF:

1. $(\mathcal{P} \wedge \mathcal{Q}) \Rightarrow \mathcal{P}$.
2. $(\mathcal{P} \wedge \mathcal{Q}) \Rightarrow \mathcal{Q}$.
3. $\mathcal{P} \Rightarrow \mathcal{P}$.
4. $\mathcal{P} \Rightarrow (\mathcal{P} \vee \mathcal{Q})$.
5. $\mathcal{Q} \Rightarrow (\mathcal{P} \vee \mathcal{Q})$.
6. $\neg\neg\mathcal{P} \Leftrightarrow \mathcal{P}$. *Princípio da Dupla Negação.*
7. $(\mathcal{P} \Rightarrow \mathcal{Q}) \Leftrightarrow (\neg\mathcal{Q} \Rightarrow \neg\mathcal{P})$.
8. $\mathcal{P} \vee \neg\mathcal{P}$. *Princípio do Terceiro Excluído.*
9. $(\mathcal{P} \wedge \mathcal{Q}) \Leftrightarrow (\mathcal{Q} \wedge \mathcal{P})$. Conjunção é comutativa.
10. $(\mathcal{P} \vee \mathcal{Q}) \Leftrightarrow (\mathcal{Q} \vee \mathcal{P})$. Disjunção é comutativa.
11. $(\mathcal{P} \Leftrightarrow \mathcal{Q}) \Leftrightarrow (\neg\mathcal{P} \Leftrightarrow \neg\mathcal{Q})$.
12. $(\mathcal{P} \Leftrightarrow \mathcal{Q}) \Leftrightarrow ((\mathcal{P} \Rightarrow \mathcal{Q}) \wedge (\mathcal{Q} \Rightarrow \mathcal{P}))$.
13. $((\mathcal{P} \vee \mathcal{Q}) \vee \mathcal{R}) \Leftrightarrow (\mathcal{P} \vee (\mathcal{Q} \vee \mathcal{R}))$. Disjunção é associativa.
14. $((\mathcal{P} \wedge \mathcal{Q}) \wedge \mathcal{R}) \Leftrightarrow (\mathcal{P} \wedge (\mathcal{Q} \wedge \mathcal{R}))$. Conjunção é associativa.
15. $(\mathcal{P} \wedge (\mathcal{Q} \vee \mathcal{R})) \Leftrightarrow ((\mathcal{P} \wedge \mathcal{Q}) \vee (\mathcal{P} \wedge \mathcal{R}))$. Distributividade da conjunção em relação à disjunção.
16. $(\mathcal{P} \vee (\mathcal{Q} \wedge \mathcal{R})) \Leftrightarrow ((\mathcal{P} \vee \mathcal{Q}) \wedge (\mathcal{P} \vee \mathcal{R}))$. Distributividade da disjunção em relação à conjunção.
17. $\neg(\mathcal{P} \vee \mathcal{Q}) \Leftrightarrow (\neg\mathcal{P} \wedge \neg\mathcal{Q})$.

Obviamente há uma infinidade de outros teoremas, além desses. O que escrevemos aqui é apenas para fins de ilustração e futura referência em trechos que ocorrem adiante neste texto.

A fórmula $(\neg\mathcal{Q} \Rightarrow \neg\mathcal{P})$ no item 7 acima é chamada de *contrapositiva* de $(\mathcal{P} \Rightarrow \mathcal{Q})$. Por conta do Princípio da Dupla Negação, a fórmula $(\mathcal{P} \Rightarrow \mathcal{Q})$ também é a contrapositiva de $(\neg\mathcal{Q} \Rightarrow \neg\mathcal{P})$.

Exemplo 2.6. i: *$(\neg(x \neq y) \Leftrightarrow x = y)$ é o Teorema 6 da lista acima, onde a fórmula $\mathcal{P}$ é $x = y$; logo, por Generalização,*

$$\forall x(\neg(x \neq y) \Leftrightarrow x = y)$$

é teorema de ZF. Aplicando Generalização novamente,

$$\forall y(\forall x(\neg(x \neq y) \Leftrightarrow x = y))$$

é mais um teorema de ZF;

ii: *$(x = y \vee x \neq y)$ é o Teorema 8 da lista acima, onde a fórmula $\mathcal{P}$ é $x = y$; logo, por Generalização,*

$$\forall x(x = y \vee x \neq y)$$

é outro teorema de ZF.

A igualdade $=$ deve satisfazer a duas condições, no sentido de serem fórmulas que são teoremas:

i: $\forall x(x = x)$;

ii: $x = y \Rightarrow (\mathcal{P}(x, x) \Rightarrow \mathcal{P}(x, y))$, onde $\mathcal{P}(x, y)$ é uma fórmula obtida a partir de $\mathcal{P}(x, x)$ por substituição de pelo menos uma ocorrência de x por y (desde que y seja livre para x em $\mathcal{P}(x, x)$, ou seja, nenhuma ocorrência livre de x em $\mathcal{P}(x, x)$ está no escopo de quantificador $\forall z$ com z ocorrendo em y).

O teorema i sobre igualdade é chamado de *reflexividade da igualdade*. Já o teorema ii é conhecido como *substitutividade da igualdade*. O importante aqui é perceber que qualquer termo t só pode ser igual a ele mesmo. Quando se escreve $x = y$, essa fórmula atômica apenas diz que o mesmo termo x é chamado também de y.

A partir da reflexividade da igualdade e da substitutividade da igualdade é possível provar que a igualdade é simétrica e transitiva. Ou seja,

$$\forall x \forall y(x = y \Rightarrow y = x)$$

e

$$\forall x \forall y \forall z((x = y \land y = z) \Rightarrow x = z)$$

são teoremas de ZF. Demonstrações desses dois últimos resultados para situações muito mais amplas do que aquelas aqui colocadas podem ser encontradas em [38].

Aqui cabe um esclarecimento. Como já definido, equações são fórmulas atômicas

$$x = y.$$

Soluções de uma equação, se existirem, são todos os termos x e y tais que a equação $x = y$ é teorema. Por conta disso, a reflexividade da igualdade garante que todo termo x é solução da equação

$$x = x.$$

Mas, obviamente, estamos interessado em outras equações, bem como na determinação das suas soluções, conforme se vê adiante, no restante da leitura.

Seção 11

Esquemas de teoremas

Sumário
Índice
Rede

O leitor mais crítico deve ter observado algo de 'errado' no Teorema 2.1. Se Teorema 2.1 é de fato um teorema de ZF, então por que o emprego da sentença metalinguística "Se $\mathcal{A}$ é uma fórmula de ZF, então..."? Isso ocorre porque, rigorosamente falando, Teorema 2.1 é um *esquema de teoremas*. Um teorema de fato de ZF é o seguinte:

Teorema 2.2. $(x = y \Rightarrow x = y)$.

Uma possível demonstração do teorema acima é feita exatamente como na demonstração do Teorema 2.1, substituindo a fórmula $\mathcal{A}$ por $x = y$. No entanto, $x = y$ não é a única possível fórmula de ZF. O que foi feito na 'demonstração' do Teorema 2.1 foi uma infinidade de demonstrações, uma para cada possível fórmula $\mathcal{A}$ de ZF. Apresentamos a seguir uma pequena lista com alguns deles:

$$\vdash_{ZF} (x = y \Rightarrow x = y),$$
$$\vdash_{ZF} (x \neq y \Rightarrow x \neq y),$$

$$\vdash_{ZF} (x \in y \Rightarrow x \in y),$$
$$\vdash_{ZF} ((x = y \wedge x \neq y) \Rightarrow (x = y \wedge x \neq y)).$$

Uma vez que $(x = y \Rightarrow x = y)$ é teorema de ZF, é possível aplicar Generalização para obter o novo teorema

$$\forall x(x = y \Rightarrow x = y).$$

Aplicando Generalização mais uma vez se obtém o novo teorema

$$\forall y \forall x(x = y \Rightarrow x = y).$$

Seção 12

Metateorema da Dedução

Sumário
Índice
Rede

Nesta Seção é qualificado, no contexto de ZF, o conceito de *premissa*, o qual é sinônimo de *hipótese*. Para uma definição aplicável a uma vasta gama de teorias formais, ver [38].

Definição 2.2. *Seja* Γ *um conjunto (na acepção da metalinguagem aqui empregada) de fórmulas de ZF. Dizemos que uma fórmula* $\mathcal{T}$ *é* consequência de Γ *em ZF sss (abreviação para 'se, e somente se') existe sequência finita de fórmulas*

$$\mathcal{F}_1, \mathcal{F}_2, \cdots, \mathcal{F}_n$$

tal que $\mathcal{F}_n$ *é* $\mathcal{T}$ *e cada passo da sequência é um axioma de ZF ou uma fórmula de* Γ *ou consequência imediata de fórmulas anteriores via o emprego de uma regra de inferência de ZF.*

Γ *é chamado de* conjunto de premissas. *Cada fórmula de* Γ *é chamada de* premissa *ou* hipótese. *Denotamos isso por*

$$\Gamma \vdash_{ZF} \mathcal{T}.$$

A visão intuitiva da definição acima é desenvolvida melhor a partir dos metateoremas que seguem abaixo.

Proposição 2.3. *Sejam* Γ *e* Δ *conjuntos quaisquer de fórmulas de ZF e* $\mathcal{T}$ *uma fórmula de ZF. Se* $\Gamma \vdash_{ZF} \mathcal{T}$ *então*

$$\Gamma \cup \Delta \vdash_{ZF} \mathcal{T}.$$

✍ A prova da proposição acima é imediata, bastando aplicar a Definição 2.2. Fica como sugestão de exercício para o leitor. Obviamente a recíproca da Proposição 2.3 (se $\Gamma \cup \Delta \vdash_{\text{ZF}} T$ então $\Gamma \vdash_{\text{ZF}} T$) não é uma proposição. Consegue encontrar contra-exemplo para a recíproca da Proposição 2.3?

O que Proposição 2.3 estabelece é o seguinte: se uma fórmula $\mathcal{T}$ é consequência de um conjunto de premissas, então não faz diferença alguma acrescentar novas premissas; $\mathcal{T}$ continuará sendo consequência do novo conjunto de hipóteses. Esse tipo de resultado ajuda a desenvolver intuições sobre o papel de premissas. Matemáticos sempre estão interessados na 'menor quantia' possível de hipóteses não triviais para provar que uma fórmula é consequência de tais hipóteses.

Consequência imediata da Proposição 2.3 é a seguinte.

PROPOSIÇÃO 2.4. *Se $\mathcal{T}$ é teorema de ZF e Γ é um conjunto de fórmulas, então $\Gamma \vdash_{ZF} \mathcal{T}$.*

Ou seja, qualquer teorema é consequência de qualquer conjunto de hipóteses. Novamente a recíproca não é uma proposição. O fato de uma fórmula ser consequência de um conjunto de premissas não implica necessariamente que tal fórmula é teorema. No entanto, se existe demonstração para $\Gamma \vdash_{\text{ZF}} \mathcal{T}$ tal que não ocorra uma única fórmula de Γ, então $\mathcal{T}$ é teorema de ZF.

PROPOSIÇÃO 2.5 (DEDUÇÃO). *Sejam Γ um conjunto de fórmulas de ZF e $\mathcal{H}$ e $\mathcal{T}$ fórmulas de ZF. Então*

$$\Gamma \cup \{\mathcal{H}\} \vdash_{ZF} \mathcal{T} \text{ sss } \Gamma \vdash_{ZF} \mathcal{H} \Rightarrow \mathcal{T}.$$

Essa última proposição é o célebre *Metateorema da Dedução*, devido ao francês Jacques Herbrand (1930). Ele garante que, se $\mathcal{H}$ é uma hipótese de um conjunto Π de premissas e $\mathcal{T}$ é consequência de tal conjunto Π de hipóteses, então $\mathcal{H} \Rightarrow \mathcal{T}$ é consequência de um conjunto de premissas que tem todas as hipóteses de Π, exceto $\mathcal{H}$. Obviamente, $\Pi = \Gamma \cup \{\mathcal{H}\}$.

Sua demonstração foge do escopo da proposta deste documento. Mas o resultado em si é de grande importância, uma vez que o Metateorema da Dedução justifica a prática das *demonstrações condicionais*, aquelas nas quais são assumidas hipóteses Γ para derivar um resultado $\mathcal{T}$. Em particular, $\{\mathcal{H}\} \vdash_{\text{ZF}} \mathcal{T}$ é equivalente a $\vdash_{\text{ZF}} \mathcal{H} \Rightarrow \mathcal{T}$.

✍ Se $\mathcal{H} \Rightarrow \mathcal{T}$ é uma fórmula de ZF, então $\mathcal{T} \Rightarrow \mathcal{H}$ é chamada de *recíproca* de $\mathcal{H} \Rightarrow \mathcal{T}$. O fato de $\mathcal{H} \Rightarrow \mathcal{T}$ ser teorema de ZF não implica que sua recíproca é teorema. Consegue justificar isso?

Escrevemos $\nvdash_{\text{ZF}} \mathcal{T}$ para dizer que $\mathcal{T}$ não é teorema de ZF e $\Gamma \nvdash_{\text{ZF}} \mathcal{T}$ para dizer que $\mathcal{T}$ não é consequência do conjunto de premissas Γ em ZF.

Exemplo 2.7. i: $\{\mathcal{A}\} \vdash_{ZF} \mathcal{B} \Rightarrow \mathcal{A}$, *sendo* $\mathcal{A}$ *e* $\mathcal{B}$ *fórmulas de ZF; a prova deste resultado é feita a partir do axioma lógico L1 e do Metateorema da Dedução 2.5;*

ii: ✍ $\{\mathcal{A}\} \vdash_{ZF} \mathcal{A}$*; consegue justificar este resultado?*

Quando o conjunto de hipóteses conta com uma única fórmula, podemos omitir o emprego de chaves. Logo, item ii acima pode ser reescrito como

$$\mathcal{A} \vdash_{\text{ZF}} \mathcal{A}.$$

Em outras palavras, está escrito acima que, se $\mathcal{A}$ é uma hipótese, então $\mathcal{A}$ é consequência dela mesma.

Proposição 2.6. *Se* $\mathcal{T}$ *é teorema de ZF e* $\mathcal{H}$ *é uma fórmula, então*

$$\vdash_{ZF} \mathcal{H} \Rightarrow \mathcal{T}.$$

✍ A prova pode ser feita a partir do Metateorema da Dedução e da Proposição 2.4. Recomendamos ao leitor que faça como exercício.

Seção 13

Princípio da Explosão

Sumário
Índice
Rede

✂ ℹ Apesar do Princípio da Explosão ser usado na discussão sobre o Paradoxo de Russell na Seção 22, não há prejuízo óbvio ao ignorar esta discussão.

Se considerarmos apenas axiomas L1, L2 e L3 da Seção 9 (ou seja, ZF sem os axiomas próprios e sem os lógicos L4 e L5), o Metateorema de Kalmár [38] garante como resultado secundário que, se uma fórmula $\mathcal{F}$ é teorema, então $\neg(\mathcal{F})$ não é teorema. Ou seja, os demais axiomas lógicos L4 e L5 devem ser consistentes com este resultado.

Em particular, sabendo que o Princípio do Terceiro Excluído

$$\mathcal{P} \vee \neg\mathcal{P}$$

é teorema, então a negação

$$\neg(\mathcal{P} \vee \neg\mathcal{P})$$

não é teorema.

No entanto, de acordo com as fórmulas 6 e 17 da Seção 10, a fórmula

$$\neg(\mathcal{P} \vee \neg\mathcal{P}) \Leftrightarrow (\neg\mathcal{P} \wedge \mathcal{P})$$

é teorema. Isso implica que $\neg\mathcal{P} \wedge \mathcal{P}$ não é teorema.

✍ Por outro lado, se $\mathcal{P}$ e $\mathcal{Q}$ são fórmulas de ZF, logo, $(\mathcal{P} \wedge \neg\mathcal{P}) \Rightarrow \mathcal{Q}$ é teorema de ZF (consegue provar?).

A fórmula $(\mathcal{P} \wedge \neg\mathcal{P})$ é uma *contradição* '$\mathcal{P}$ e não $\mathcal{P}$', no sentido de que $\neg(\mathcal{P} \wedge \neg\mathcal{P})$ é teorema. Logo, de acordo com o Metateorema da Dedução,

$$(\mathcal{P} \wedge \neg\mathcal{P}) \vdash_{\mathrm{ZF}} \mathcal{Q}.$$

Equivalentemente isso pode ser escrito como

$$\{\mathcal{P}, \neg\mathcal{P}\} \vdash_{\mathrm{ZF}} \mathcal{Q}.$$

Este é o *Princípio da Explosão*: a partir de um conjunto Γ de hipóteses cuja conjunção é uma contradição, qualquer fórmula $\mathcal{Q}$ de ZF é consequência de Γ.

Em outras palavras, contradições permitem inferir qualquer fórmula. Se, em particular, os axiomas próprios de ZF produzirem algum teorema $\mathcal{T}$ tal que $\neg(\mathcal{T})$ também é teorema, então todas as fórmulas de ZF são teoremas. Até hoje não se sabe se esse fenômeno altamente inconveniente ocorre ou não em ZF.

Matemáticos também operam sob o comando de crenças pessoais. A sensação dominante é que muito provavelmente ZF é consistente. Com efeito, até hoje não foi encontrada qualquer inconsistência. Mas, claro, essa fé não é cega. Se alguém conseguir exibir uma inconsistência em ZF, a teoria deverá ser reescrita.

O Princípio da Explosão motivou a edificação de outras lógicas chamadas de *paraconsistentes* [9], nas quais tal princípio não vale. Logo, lógicas paraconsistentes não são equivalentes à lógica clássica, apesar de poderem ser percebidas como uma extensão da mesma. Detalhes na referência citada.

Seção 14

Ainda sobre definições

Sumário
Índice
Rede

✂ Agora que conhecemos um pouco a respeito de linguagem e lógica, podemos detalhar mais a respeito de definições, dando continuidade à Seção 8.

Definições explícitas abreviativas devem ser:

- *Elimináveis* e
- *Conservativas*.

O critério de *eliminabilidade* estabelece que, em qualquer definição explícita abreviativa, podemos substituir o *definendum* pelo *definiens*. Por exemplo, considere a definição do símbolo metalingístico $\neq$, introduzido na Seção 8. Ao escrevermos $x \neq y$, podemos substituir tal abreviação pela fórmula correspondente na linguagem $\mathfrak{S}$, a saber, $\neg(x = y)$.

O critério de eliminabilidade reforça a economia de notação na prática matemática, como já discutido. Uma vez que ZF é uma teoria de fundamentação para assuntos como cálculo diferencial e integral, conceitos sofisticados, como *integral de Riemann* (Seção 58), podem ser introduzidos com considerável economia de notação graças a uma estratégica lista de definições abreviativas. Escrever os conceitos de limite, derivada e integral, sem o emprego de tais definições dadas anteriormente (na medida em que este texto evolui), é obviamente possível (usando única e exclusivamente os símbolos do vocabulário da linguagem $\mathfrak{S}$ de ZF); mas não é um recurso amigável para fins de escrita e leitura de matemática. Ou seja, a prática matemática deve levar com consideração limitações cognitivas humanas.

O segundo critério estabelece que toda definição explícita abreviativa deve ser conservativa, ou seja, não deve permitir a formulação de novos teoremas que não poderiam ser obtidos sem a definição. Por exemplo, digamos que alguém proponha o que se segue, como definição para o símbolo metalinguístico $\boxplus$: dadas as fórmulas $\mathcal{A}$, $\mathcal{B}$

e $\mathcal{C}$, então

$$(\mathcal{A} \boxplus \mathcal{B}) \Leftrightarrow \mathcal{C} \vdots (\mathcal{A} \wedge \mathcal{B}) \Rightarrow \mathcal{C}.$$

Claramente o critério de eliminabilidade é satisfeito para quaisquer ocorrências de $(\mathcal{A} \boxplus \mathcal{B}) \Leftrightarrow \mathcal{C}$. Mesmo nos casos em que há ocorrência apenas de $\mathcal{A} \boxplus \mathcal{B}$, podemos substituir $\mathcal{A} \boxplus \mathcal{B}$ por qualquer fórmula $\mathcal{C}$ (por conta da bicondicional $\Leftrightarrow$). No entanto, de acordo com a Seção 10, as fórmulas $(\mathcal{A} \wedge \mathcal{B}) \Rightarrow \mathcal{A}$ e $(\mathcal{A} \wedge \mathcal{B}) \Rightarrow \mathcal{B}$ são teoremas. Logo, $(\mathcal{A} \boxplus \mathcal{B}) \Leftrightarrow \mathcal{A}$ e $(\mathcal{A} \boxplus \mathcal{B}) \Leftrightarrow \mathcal{B}$ são teoremas. Portanto, $\mathcal{A} \Leftrightarrow \mathcal{B}$ é teorema. Se $\mathcal{B}$ for a fórmula $\neg\mathcal{A}$, isso implica em uma contradição $\mathcal{A} \Leftrightarrow \neg\mathcal{A}$, a qual é um novo teorema, no sentido de que, antes da suposta definição de $\boxplus$, a fórmula $\mathcal{A} \Leftrightarrow \neg\mathcal{A}$ não era teorema. Portanto, apesar da estrutura

$$\textit{definiendum} \vdots \textit{definens},$$

isso não é uma definição explícita abreviativa para o símbolo $\boxplus$.

Resumidamente, definições explícitas abreviativas devem apenas *abreviar* fórmulas, desde que não sejam equivalentes a novos postulados de ZF, uma vez que apenas novos postulados podem ser responsáveis por novos teoremas. No caso acima, a suposta definição para o símbolo $\boxplus$ introduz a fórmula

$$((\mathcal{A} \boxplus \mathcal{B}) \Leftrightarrow \mathcal{C}) \Leftrightarrow ((\mathcal{A} \wedge \mathcal{B}) \Rightarrow \mathcal{C})$$

como novo postulado à lógica de ZF, o qual é inconsistente com os demais postulados lógicos. Mais detalhes podem ser encontrados em [43].

Seção 15

Verdade

Sumário
Índice
Rede

A essa altura o leitor já deve ter observado que, em momento algum, foram qualificados os conceitos de *verdade* e *falsidade*. O estudo de ZF pode ser promovido sem jamais mencionar algo como verdade ou falsidade. O que interessa em ZF é se uma dada fórmula é teorema ou não. No entanto, é perfeitamente possível (e extremamente útil) qualificar a afirmação 'a fórmula $\mathcal{A}$ é verdadeira'. Uma discussão sobre este problema e sua relevância é colocada na Seção 111.

Seção 16

Resumo da ópera

Sumário
Índice
Rede

O que sabemos até aqui pode ser resumido da seguinte maneira.

- A partir do vocabulário de $\mathfrak{S}$ são definidas sentenças, as quais são apenas sequências finitas de símbolos do vocabulário de $\mathfrak{S}$.
- Entre as sentenças de $\mathfrak{S}$ são selecionadas aquelas que são fórmulas. Isso é feito graças a uma sintaxe.
- Entre as fórmulas de $\mathfrak{S}$ são selecionadas aquelas que são os axiomas de ZF. Com efeito, ZF não é definida apenas por uma linguagem, mas por uma lógica também.
- Entre as fórmulas de $\mathfrak{S}$ são selecionados os teoremas de ZF. Teoremas são obtidos a partir de axiomas e/ou regras de inferência. Teoremas cujas respectivas demonstrações não demandam o emprego de qualquer regra de inferência são chamados de *triviais*. Logo, todo axioma de ZF é um teorema trivial, conforme Proposição 2.1. Teoremas, cujas possíveis demonstrações sempre empregam pelo menos uma regra de inferência, são chamados de *não triviais*.
- Matemáticos que trabalham com ZF estão interessados prioritariamente nos teoremas não triviais de ZF.

Com relação ao último item acima, notar que há também interesse no estudo dos próprios axiomas de ZF, pelo menos de um ponto de vista metalinguístico. Ver Seção 111.

Como foi dito anteriormente, matemáticos estão mais interessados em rigor do que formalismo. Neste contexto, as demonstrações realizadas na prática matemática não seguem *ipsis litteris* a Definição 2.1. No lugar disso, demonstrações típicas de ZF (aquelas que são comumente encontradas na literatura especializada) são simplesmente sequências finitas de afirmações, as quais podem ser formalmente transcritas nos moldes da Definição 2.1.

Obviamente, não há procedimento efetivo para decidir se uma demonstração, nesta acepção mais relaxada, é rigorosa ou não. Justamente por conta disso que erros humanos são muito comuns entre matemáticos. Daí a necessidade de troca de ideias entre pares, como salientado na Seção 3. Ao longo de todo o restante deste texto adotamos a prática matemática comum de que demonstrações devem ser rigorosas, mas não necessariamente formais.

Seção 17

Notas históricas

SUMÁRIO
ÍNDICE
REDE

A linguagem $\mathfrak{S}$ aqui empregada é um caso particular de *Cálculo Predicativo de Primeira Ordem* [38], o qual atingiu um considerável amadurecimento nas mãos de Gottlob Frege [22]. No entanto, há outras linguagens formais com expressividade muito maior, como os *Cálculos de Ordem Superior*. Um exemplo bem conhecido é a teoria ZF em sua versão de *segunda ordem* [8] (conhecida como ZF_2), a qual conta com uma linguagem diferente da linguagem $\mathfrak{S}$ aqui discutida.

DAVID HILBERT EM 1912

Fonte: Wikipedia.

A clara distinção entre cálculos de primeira ordem e de ordem superior somente tomou forma na segunda década do século 20, graças principalmente a David Hilbert e colaboradores. Para uma ampla visão histórica do nascimento de lógica e fundamentos nos moldes do que hoje se entende sobre o tema, ver o extraordinário livro de Jean van Heijenoort [22].

PARTE 3

O que faz a pertinência

Nesta terceira parte a ênfase é sobre os axiomas próprios de ZF, bem como os requisitos para a fundamentação de certos ramos da matemática que encontram ampla aplicabilidade.

Seção 18
O primeiro axioma próprio de ZF

Sumário
Índice
Rede

Os axiomas próprios de ZF se referem explicitamente ao predicado binário $\in$, no sentido de como ele se relaciona com conectivos lógicos, quantificadores lógicos e a igualdade. Segue nesta, e nas próximas Seções, a lista de todos os postulados próprios de ZF.

Cada axioma próprio de ZF tem um nome:

- *Extensionalidade*,
- *Vazio*,
- *Par*,
- *Potência*,
- *União*,
- *Separação*,

- *Infinito*,
- *Regularidade* e
- *Substituição*.

Uma variação de ZF é discutida mais adiante, chamada de ZFC. Ela conta com os mesmos axiomas de ZF e um postulado a mais chamado de *Escolha*.

Segue o primeiro postulado próprio de ZF.

ZF1 - Extensionalidade:

$$\forall x \forall y \forall z((z \in x \Leftrightarrow z \in y) \Rightarrow x = y).$$

Os termos de ZF são chamados de *conjuntos*. A origem histórica do termo *conjunto* deriva da interpretação pretendida de que ZF deve capturar pelo menos parte das ideias originais de Georg Cantor, autor de um corpo do conhecimento chamado *Mengenlehre* (teoria de conjuntos, em tradução livre do alemão).

O Axioma da Extensionalidade de ZF afirma o seguinte: se x e y são conjuntos que compartilham os mesmos elementos z, então x é idêntico a y.

De um ponto de vista intuitivo, o Axioma da Extensionalidade estabelece que um conjunto x é identificado única e exclusivamente pelos conjuntos z tais que $z \in x$, ou seja, por seus elementos. A recíproca do Axioma da Extensionalidade é teorema, como se percebe a seguir.

Teorema 3.1. $\forall x \forall y \forall z(x = y \Rightarrow (z \in x \Leftrightarrow z \in y))$.

Demonstração: Sabemos que

$$z \in x \Leftrightarrow z \in x$$

é teorema (cuja demonstração pode ser exibida usando apenas os axiomas lógicos de ZF e Modus Ponens, de maneira análoga ao Teorema 2.1).

Logo,

$$z \in x \Leftrightarrow z \in x$$

é consequência de qualquer premissa (Proposição 2.4), em particular, $x = y$.

Portanto,

$$x = y \Rightarrow (z \in x \Leftrightarrow z \in x)$$

é teorema (Proposição 2.6). Mas, de acordo com a substitutividade da igualdade, podemos substituir qualquer ocorrência livre de x por y na fórmula $z \in x \Leftrightarrow z \in x$, de modo que a nova fórmula é teorema. Logo,

$$x = y \Rightarrow (z \in x \Leftrightarrow z \in y)$$

é teorema.

Aplicando Generalização, temos

$$\forall z(x = y \Rightarrow (z \in x \Leftrightarrow z \in y)).$$

Aplicando novamente temos

$$\forall y \forall z(x = y \Rightarrow (z \in x \Leftrightarrow z \in y)).$$

Aplicando Generalização mais uma vez temos

$$\forall x \forall y \forall z(x = y \Rightarrow (z \in x \Leftrightarrow z \in y)).$$

Isso conclui a prova.

Ou seja, o Axioma da Extensionalidade, em parceria com o Teorema 3.1, estabelece que a fórmula $x = y$ é equivalente a afirmar que x e y compartilham os mesmos elementos. Essa é uma informação de extraordinária importância sobre o predicado binário de pertinência $\in$. Tal predicado é necessário e suficiente para identificar um conjunto.

Se o leitor se interessar por uma compreensão mais aprofundada sobre o Axioma da Extensionalidade, no artigo [1] há uma proposição que prova o seguinte resultado:

$$\nvdash_{\text{ZF}-\{\text{Extensionalidade}, \text{Infinito}\}} \text{Extensionalidade}.$$

Por um lado, ZF $-\{$Extensionalidade, Infinito$\}$ é uma teoria formal com os mesmos postulados de ZF, exceto o Axioma da Extensionalidade e o Axioma do Infinito (este último é discutido na Seção 23). Por outro, a proposição acima simplesmente diz que o Axioma da Extensionalidade não é teorema em uma teoria que conta com todos os axiomas de ZF, exceto Extensionalidade e Infinito. Tal resultado é de enorme importância. Com efeito, isso significa que apenas o predicado de igualdade $=$ não é o bastante para identificar conjuntos.

Logo, aprendemos uma valiosa lição:

O Axioma da Extensionalidade desempenha papel indispensável para a identificação de conjuntos.

Seção 19

Quantificador $\exists!$

Sumário
Índice
Rede

Antes de prosseguirmos com os demais postulados de ZF, é útil a introdução de uma nova abreviação metalinguística.

Seja $\mathcal{A}$ uma fórmula de ZF. Logo,

$$\exists! x(\mathcal{A}(x)) \vdots \exists x \forall y(\mathcal{A}(y) \Leftrightarrow y = x).$$

A abreviação $\exists! x(\mathcal{A}(x))$ se lê 'existe um único x tal que $\mathcal{A}(x)$'. A ideia intuitiva é simples: existe um x tal que $\mathcal{A}(x)$ e, para qualquer y tal que $\mathcal{A}(y)$, temos que $y = x$.

i Existem outras formas para definir o quantificador $\exists!$. Mas o conceito dado acima basta para nossos propósitos.

Seção 20

Existem Conjuntos?

Sumário
Índice
Rede

O Axioma da Extensionalidade não garante a existência de conjuntos. Apenas garante que, se existirem conjuntos, sabemos como identificá-los a partir da pertinência $\in$. O primeiro postulado a garantir que pelo menos um conjunto existe é o que se segue.

ZF2 - Vazio:
$$\exists x \forall y(y \notin x).$$

Observar atentamente o quantificador existencial acima, bem como a maneira como ele opera em 'parceria' com o quantificador universal. Este postulado garante a *existência* de um conjunto x tal que *nenhum* conjunto y pertence a ele. O próximo teorema ilustra como os postulados de ZF trabalham em 'parceria' uns com os outros.

TEOREMA 3.2. *O conjunto x do Axioma do Vazio é único.*

DEMONSTRAÇÃO: O Axioma da Extensionalidade pode ser reescrito como

$$\forall x \forall y \forall z((z \notin x \Leftrightarrow z \notin y) \Rightarrow x = y).$$

Ver Teoremas 7 e 12 da lista de 17 teoremas da Seção 10, para saber como provar essa última fórmula.

Ou seja, a fórmula acima é teorema de ZF.

Seja x o conjunto cuja existência é garantida pelo Axioma do Vazio, i.e., para todo z temos que $z \notin x$. Supor que existe outro conjunto y (ou seja, $y \neq x$) que também satisfaz o Axioma do Vazio. Logo, para todo z temos $z \notin y$. Isso significa que

$$\forall z(z \notin x \Leftrightarrow z \notin y).$$

Mas, de acordo com o Axioma da Extensionalidade (na forma como está reescrito acima), isso implica em $y = x$ ($\perp$).

O símbolo $\perp$ usado ao final da demonstração acima (conhecido como *falsum*) é o que se chama de *contradição* (neste caso, a contradição sinalizada por $\perp$ é $y \neq x \wedge y = x$). Uma vez que $\mathcal{P} \vee \neg\mathcal{P}$ é teorema para qualquer fórmula $\mathcal{P}$, se $\neg\mathcal{P}$ garante uma contradição, então $\mathcal{P}$ deve ser teorema. Uma vez que a negação da tese acima produz uma contradição, então deve valer a tese como teorema. A tese em questão pode ser escrita formalmente como se segue:

$$\exists! x(\forall y(y \notin x)).$$

Caso o leitor não saiba, a expressão 'i.e.' (usada na última prova) abrevia '*id est*' que, em latim, se traduz como 'isto é'.

✍ Um exercício interessante é escrever formalmente o Teorema 3.2 usando apenas os quantificadores $\forall$ e $\exists$, de acordo com a Seção 19. Obviamente, o que legitima tal demonstração é a hipótese de que ZF é consistente (ou seja, a hipótese de que não existe fórmula $\mathcal{A}$ tal que ambas $\mathcal{A}$ e $\neg\mathcal{A}$ são teoremas de ZF), algo que até hoje não se sabe se é o caso.

Como já dito anteriormente, o teorema $\mathcal{P} \vee \neg\mathcal{P}$ é conhecido como *Princípio do Terceiro Excluído*. Este legitima as demonstrações *reductio ad absurdum* (redução ao absurdo, em tradução livre do latim): se a negação de uma tese (a qual é tão somente uma fórmula) implica em contradição, então a tese é teorema.

Notar também que a fórmula

$$\forall x \forall y \forall z((z \notin x \Leftrightarrow z \notin y) \Rightarrow x = y),$$

a qual é equivalente ao Axioma da Extensionalidade, viabiliza outra visão intuitiva a respeito da identificação de conjuntos. Assim como conjuntos são identificados por seus elementos, equivalentemente conjuntos são também identificados pelos termos que *não são* seus elementos.

Uma vez que acabamos de provar que conjunto vazio é único, este é uma constante de ZF. Por conta disso é usual a adoção de um símbolo especial para tal constante: $\varnothing$. Ou seja,

$$\forall y(y \notin \varnothing).$$

Aqui cabe uma oportuna observação de caráter histórico, filosófico, matemático e didático, em relação à técnica empregada para provar Teorema 3.2.

A experiência em sala de aula revela que muitos alunos encontram dificuldade para compreender e aceitar a técnica de demonstração por redução ao absurdo. Pois bem, isso não é exclusividade de alunos. Alguns matemáticos, justamente por conta de suas experiências profissionais, também criticam esse método de demonstração.

No início do século 20, Luitzen Egbertus Jan Brouwer não aceitava demonstrações por redução ao absurdo. Uma vez que ela é sustentada pelo Princípio do Terceiro Excluído, na visão de Brouwer a fórmula $\mathcal{P} \vee \neg\mathcal{P}$ só pode ser teorema se existir uma demonstração para $\mathcal{P}$ ou uma demonstração para $\neg\mathcal{P}$, de modo que qualquer demonstração de uma não pode depender do 'fracasso' de outra, por conta de uma contradição. Provar que a negação $\neg\mathcal{P}$ de uma tese $\mathcal{P}$ implica em uma contradição, não garante que $\mathcal{P}$ é teorema, segundo a postura filosófica de Brouwer. Com efeito, se a negação de uma tese implica em uma contradição, apenas foi provado que tal negação da tese implica em uma contradição, nada além disso. Por conta dessa visão, este conhecido matemático holandês rejeitava a lógica clássica usada hoje para edificar ZF (entre muitas outras teorias).

Em oposição à lógica clássica, Brouwer introduziu a *Lógica Intuicionista*, na qual o Princípio do Terceiro Excluído não é teorema.

ℹ Hoje em dia existem diversos sistemas formais que empregam lógica intuicionista, incluindo uma versão intuicionista de ZF [5]. Essa última referência é um livro não publicado de John Bell, mas gratuitamente disponível em pdf na internet. Até onde sabemos, não há livros publicados sobre o tema.

ℹ Outro exemplo de teoria fundamentada em lógica intuicionista é a *Análise Infinitesimal Suave*. Esta última permite desenvolver uma forma de cálculo diferencial e integral na qual todas as funções são contínuas, algo que não ocorre no Cálculo Diferencial e Integral Padrão (ver Definição 5.23). Outrossim, demonstrações por redução absurdo não são aplicáveis em análise infinitesimal suave. Se o leitor estiver interessado, no livro de John Bell [4] há uma excelente e sucinta exposição sobre o tema, onde derivadas e integrais podem ser definidas sem a necessidade de limites. No cálculo padrão derivadas e integrais são casos especiais de limites.

No entanto, a motivação de Brouwer era meramente filosófica, apesar de hoje encontrar grande repercussão em matemática e até mesmo em física teórica. Neste livro adotamos lógica clássica.

Em lógica clássica o Princípio do Terceiro Excluído é teorema. Portanto, demonstrações por redução ao absurdo podem ser empregadas para a obtenção de teoremas. Essas informações devem ajudar o leitor a perceber que existem muitas formas para desenvolver matemática. Neste livro apenas tangenciamos uma dessas formas, a qual é a mais usual.

Se um aluno encontra dificuldade para aceitar a técnica de redução ao absurdo para a demonstração de certos teoremas, basta que este mesmo aluno tenha consciência de que o Princípio do Terceiro Excluído é tão somente uma consequência dos axiomas lógicos de ZF. Logo, qualquer objeção dessa natureza é de caráter filosófico, não matemático.

Garantir a existência de um único conjunto em ZF, a saber, o vazio, é insuficiente para a prática matemática. Logo, precisamos de mais postulados.

ZF3 - Par:

$$\forall x \forall y \exists z \forall t (t \in z \Leftrightarrow (t = x \vee t = y)).$$

O Axioma do Par garante a existência de outros conjuntos z (chamados de *pares*) além de $\varnothing$ (observar o quantificador existencial $\exists z$).

O Axioma do Par diz o seguinte: dados x e y, *existe* z cujos elementos são x ou y. Por exemplo, uma vez que é garantida a existência do conjunto vazio $\varnothing$, o Axioma do Par garante a existência de um z tal que $t \in z$ se, e somente se, $t = \varnothing \vee t = \varnothing$ (aqui os termos x e y do Axioma do Par assumem os valores $\varnothing$ e $\varnothing$). Neste caso o Axioma da Extensionalidade garante que $z \neq \varnothing$, uma vez que $\varnothing \in z$ mas $\varnothing \notin \varnothing$.

Neste momento se mostra útil a introdução de símbolos auxiliares metalinguísticos novos: $\{$ e $\}$ (chamados de *chaves*).

Se z é um par com elementos x e y, denotamos isso por

$$z = \{x, y\},$$

desde que $x \neq y$.

Se $x = y$, escrevemos

$$z = \{x\}$$

ou

$$z = \{y\}.$$

O Axioma da Extensionalidade garante que $\{x, y\} = \{y, x\}$. Também garante (na forma de teorema) que, dados x e y, o par $z = \{x, y\}$ (ou $z = \{x\}$) é único. Se o par z conta com um único elemento, ele é chamado de *singleton* ou *unitário*.

Exemplo 3.1. *Sejam $x = \varnothing$ e $y = \varnothing$. Logo, $z = \{\varnothing\}$. Neste caso z é um* singleton.

Exemplo 3.2. *Sejam $x = \varnothing$ e $y = \{\varnothing\}$. Logo, $z = \{\varnothing, \{\varnothing\}\}$. Com efeito, a existência de $\varnothing$ é garantida pelo Axioma do Vazio, enquanto a existência de $\{\varnothing\}$ é garantida pela aplicação do Axioma do Par no* Exemplo *anterior. Observar que, de acordo com o Axioma de Extensionalidade, $\{\varnothing, \{\varnothing\}\} = \{\{\varnothing\}, \varnothing\}$.*

O emprego de chaves { e } como novos símbolos auxiliares motiva uma notação alternativa para o conjunto vazio, a saber, $\{\}$. Apesar desta ser uma notação bastante comum na literatura, ela não é empregada aqui.

Teorema 3.3. *Se x é um conjunto unitário e $x = y$, então y é unitário.*

Demonstração: Se x é unitário, $\exists a(x = \{a\})$. Supor que y não é unitário. Logo, existe pelo menos um elemento t em y tal que $t \neq a$. Logo, $t \notin x$. Logo, o Axioma da Extensionalidade garante que $x \neq y$. $\perp$

O Axioma do Par garante a existência de uma infinidade de conjuntos. Basta aplicá-lo repetidas vezes a partir do conjunto vazio. No entanto, cada um dos conjuntos obtidos a partir de Par e Vazio conta com, no máximo, dois elementos. Para fins de fundamentação da prática matemática isso é muito pouco. Daí a necessidade de mais postulados! Mas, antes de proseguirmos com novos axiomas, segue uma definição muito útil: o conceito de *par ordenado*.

Definição 3.1 (Kuratowski).

$$(a, b) \vdots \{\{a\}, \{a, b\}\}.$$

Na definição abreviativa acima não está sendo introduzida qualquer abreviação metalinguística para uma fórmula de ZF, mas uma abreviação metalinguística para um termo denotado por (a, b). Obviamente tal manobra pode ser adaptada para a seguinte forma:

$$t = (a, b) \vdots t = \{\{a\}, \{a, b\}\}$$

ou

$$t = (a, b) \vdots \exists x \exists y(x \in t \wedge y \in t \wedge a \in x \wedge a \in y \wedge b \in y),$$

sendo t obtido por repetidas aplicações do Axioma do Par.

Observar que (a, b) é um conjunto, uma vez que a e b são conjuntos. O termo (a, b) é chamado de *par ordenado*. Esse nome se justifica pelo próximo teorema.

Teorema 3.4. *$(a, b) = (c, d)$ se, e somente se, $a = c$ e $b = d$.*

DEMONSTRAÇÃO: Uma vez que o teorema é dado por uma bicondicional, a demonstração é dividida em duas partes. A conjunção do final de ambas as partes é exatamente o teorema.

Parte $\Leftarrow$. De acordo com a definição de Kuratowski,

$$(a, b) = \{\{a\}, \{a, b\}\}$$

e

$$(c, d) = \{\{c\}, \{c, d\}\}.$$

Se $a = c$ e $b = d$, o Axioma da Extensionalidade garante que $(a, b) = (c, d)$.

Parte $\Rightarrow$. Essa segunda parte da demonstração deve ser dividida em duas possíveis situações:

I: o caso em que $a = b$ e

II: o caso em que $a \neq b$.

Se $a = b$, temos que

$$(a, b) = (a, a) = \{\{a\}\}.$$

Logo, o par ordenado (a, b) é unitário. Mas Teorema 3.3 garante que (c, d) é unitário. Logo, $(c, d) = \{\{c\}\}$, sendo $c = d$. Logo, $\{\{a\}\} = \{\{c\}\}$. O Axioma da Extensionalidade garante que $a = c$. Neste caso $b = d$ é consequência da transitividade da igualdade. ✍ O restante da demonstração fica a cargo do leitor interessado.

EXEMPLO 3.3. *O par ordenado $(\varnothing, \{\varnothing\})$ é diferente de $(\{\varnothing\}, \varnothing)$. Com efeito,*

$$(\varnothing, \{\varnothing\}) = \{\{\varnothing\}, \{\varnothing, \{\varnothing\}\}\}$$

e

$$(\{\varnothing\}, \varnothing) = \{\{\{\varnothing\}\}, \{\varnothing, \{\varnothing\}\}\}.$$

Apesar de ambos os conjuntos compartilharem um elemento em comum, a saber, $\{\varnothing, \{\varnothing\}\}$, o termo $\{\varnothing\}$ pertence ao primeiro par ordenado mas não ao segundo. Logo, o Axioma da Extensionalidade garante que $(\varnothing, \{\varnothing\}) \neq (\{\varnothing\}, \varnothing)$.

A definição de par ordenado, introduzida por Kazimierz Kuratowski, motiva nova nomenclatura. Qualquer par obtido pelo Axioma do Par é chamado de *par não ordenado*. Isso porque, por exemplo,

$$\{\varnothing, \{\varnothing\}\} = \{\{\varnothing\}, \varnothing\},$$

apesar de

$$(\varnothing, \{\varnothing\}) \neq (\{\varnothing\}, \varnothing).$$

Neste contexto, pares ordenados são casos particulares de pares não ordenados. O que permite estabelecer relevância na 'ordenação' de um par ordenado é o fato de ZF ser uma teoria com igualdade. Essa foi a ideia genial de Kuratowski!

A definição de par ordenado não foi uma conquista fácil em lógica-matemática. Outras propostas, muito mais complicadas, antecederam a ideia de Kuratowski. Detalhes em [48].

Seção 21

Potência, união arbitrária e união finitária

Sumário
Índice
Rede

Os axiomas do Vazio e do Par não garantem a existência de conjuntos suficientes para a prática matemática. Logo, precisamos de novos postulados. Mas, antes disso, as seguintes definições são úteis.

Definição 3.2. *Sejam x e y conjuntos. Logo,*

i: $x \subseteq y \vdots \forall t(t \in x \Rightarrow t \in y)$; *lemos $x \subseteq y$ como 'x é* subconjunto *de y' ou 'x* está contido *em y';*

ii: $x \subset y \vdots x \subseteq y \wedge x \neq y$; *lemos $x \subset y$ como 'x é* subconjunto próprio *de y';*

iii: $x \not\subseteq y \vdots \neg(x \subseteq y)$;

iv: $x \not\subset y \vdots \neg(x \subset y)$.

Ou seja, x é subconjunto de y sss todo elemento t de x é elemento de y. Além disso, x é subconjunto próprio de y sss x é subconjunto de y e x é diferente de y.

Exemplo 3.4. i: $\{\varnothing\} \subseteq \{\varnothing, \{\varnothing\}\}$;

ii: $\{\varnothing\} \subset \{\varnothing, \{\varnothing\}\}$;

III: $\{\varnothing, \{\varnothing\}\} \not\subseteq \{\varnothing\}$;

IV: $\{\varnothing\} \subseteq \{\varnothing\}$;

V: $\{\varnothing\} \not\subset \{\varnothing\}$.

Teorema 3.5. *Todo conjunto é subconjunto de si mesmo.*

Demonstração: Formalmente, o teorema estabelece que

$$\forall x(x \subseteq x).$$

De acordo com a Definição 3.2, devemos provar que

$$\forall x \forall t(t \in x \Rightarrow t \in x).$$

Mas $t \in x \Rightarrow t \in x$ é teorema em ZF (de acordo com Teorema 2.1). Logo, aplicando Generalização duas vezes, temos $\forall x \forall t(t \in x \Rightarrow t \in x)$.

Em particular, foi provado acima que $\varnothing \subseteq \varnothing$. Notar também que, apesar de não termos ainda à nossa disposição outros conjuntos, além de vazio e pares, o último teorema diz o seguinte: quaisquer outros postulados que garantam a existência de novos conjuntos devem ser tais que todo conjunto é subconjunto de si mesmo.

No entanto, o próximo teorema mostra que vazio não é subconjunto apenas dele mesmo.

Teorema 3.6. *O conjunto vazio é subconjunto de qualquer conjunto.*

Demonstração: Formalmente, o teorema estabelece que

$$\forall x(\varnothing \subseteq x).$$

Supor que $\neg\forall x(\varnothing \subseteq x)$. Logo, $\exists x(\varnothing \not\subseteq x)$. Logo, existe t tal que $t \in \varnothing \wedge t \notin x$. $\perp$.

Mais uma vez redução ao absurdo foi usada como técnica de demonstração, uma vez que empregamos aqui a lógica clássica. A ideia intuitiva da prova acima é a seguinte. Supor que a tese não é teorema, ou seja, não é teorema a afirmação de que o conjunto vazio é subconjunto de todo e qualquer conjunto. Isso é equivalente a afirmar que existe pelo menos um conjunto x tal que $\varnothing$ não é subconjunto de x. Mas isso, de acordo com a definição de subconjunto, é equivalente a afirmar que existe pelo menos um t que pertence a $\varnothing$ de modo que t não pertence a x. Não obstante, $\varnothing$ é um conjunto que não admite elemento algum. Logo, a negação da tese garante uma contradição (a existência de um t tal que $t \in \varnothing$, sendo que nenhum t pertence a vazio). Logo, o Princípio do Terceiro Excluído (um dos teoremas de ZF) garante que a tese é necessariamente teorema.

Uma confusão frequente entre aprendizes de matemática reside na diferença entre $\in$ e $\subseteq$. O último teorema é uma ótima oportunidade para evitar tal desconforto desnecessário. Basta observar que $\varnothing \notin \varnothing$, uma vez que termo algum pertence a vazio. No entanto, $\varnothing \subseteq \varnothing$, uma vez que vazio é subconjunto de qualquer conjunto, incluindo, obviamente, o próprio vazio. Ambas as fórmulas

$$\varnothing \notin \varnothing \quad \text{e} \quad \varnothing \subseteq \varnothing$$

são teoremas de ZF. Isso implica que as fórmulas $\varnothing \in \varnothing$ e $\varnothing \nsubseteq \varnothing$ não são teoremas de ZF (se ZF for consistente, claro).

Agora podemos finalmente introduzir o Axioma da Potência.

ZF4 - Potência:

$$\forall x \exists y \forall t (t \in y \Leftrightarrow t \subseteq x).$$

O conjunto y acima é chamado de *potência* de x. Se x é um conjunto qualquer, sua potência y (cuja existência é garantida por ZF4) é o conjunto cujos elementos são todos os subconjuntos t de x.

✍ O Axioma da Extensionalidade garante que a potência y de qualquer conjunto x é única (mais um teorema que sugerimos ao leitor demonstrar). Por conta disso, usualmente a potência y de x é denotada por

$$y = \wp(x).$$

Em particular, se c é uma constante de ZF, então $\wp(c)$ também é uma constante de ZF.

Exemplo 3.5. i: *Se* $x = \varnothing$, *então* $\wp(x) = \{\varnothing\}$*;*

ii: *se* $x = \{\varnothing\}$, *então* $\wp(x) = \{\varnothing, \{\varnothing\}\}$*;*

iii: $\wp(\{\varnothing, \{\varnothing\}\}) = \{\varnothing, \{\varnothing\}, \{\{\varnothing\}\}, \{\varnothing, \{\varnothing\}\}\}$.

O último Exemplo ilustra o fato de que ZF admite uma infinidade de constantes, a saber, $\varnothing$, $\wp(\varnothing)$, $\wp(\wp(\varnothing))$ e assim por diante.

Observar que, se x tem n elementos, então $\wp(x)$ tem 2^n elementos (um simples problema de análise combinatória). No item iii acima, o conjunto $\{\varnothing, \{\varnothing\}\}$ tem dois elementos, enquanto $\wp(\{\varnothing, \{\varnothing\}\})$ tem 2^2 elementos, ou seja, 4.

Aqui cabe mais uma observação. Essa conta 2^n, para o número de elementos da potência de um conjunto com n elementos, está sendo feita aqui no contexto da metalinguagem usada para falarmos sobre a linguagem-objeto $\mathfrak{S}$ empregada para edificar ZF. No entanto, é possível qualificar com precisão o que é o *'número' de elementos de um conjunto*. Isso se faz a partir da noção de *cardinalidade* de um conjunto. No entanto, este é outro assunto que escapa de nossos propósitos para um texto meramente introdutório. Para detalhes sobre cardinalidade de um conjunto, ver [28]. Para um estudo muito mais avançado sobre o tema, ver [30].

Graças aos quatro primeiros axiomas de ZF, podemos garantir agora a existência de uma nova infinidade de conjuntos, incluindo aqueles que contam com 2^n elementos (1, 2, 4, 8, 16, ...). Para efeitos práticos, isso significa que o Axioma da Potência garante a existência de conjuntos tais que os demais postulados anteriores não conseguem garantir. O item iii do Exemplo 3.5 exibe um conjunto cuja existência não pode ser garantida apenas a partir dos axiomas que antecedem ZF4.

O próximo é o Axioma da União.

ZF5 - União:

$$\forall x \exists y \forall z (z \in y \Leftrightarrow \exists w (z \in w \wedge w \in x)).$$

Chamamos y de *união arbitrária dos termos w que pertencem a x* e denotamos isso como

$$y = \bigcup_{w \in x} w.$$

Em outras palavras, dado um conjunto x, os elementos de y (cuja existência é garantida por ZF5) são os termos z que pertencem a w, para cada w que pertence a x. Novamente o Axioma da Extensionalidade garante que, para cada x, a união arbitrária $y = \bigcup_{w \in x} w$ é única.

Exemplo 3.6. *Seja* $x = \{\varnothing, \{\varnothing\}, \{\{\varnothing\}\}, \{\varnothing, \{\varnothing\}\}\}$. *Logo,*
$$\bigcup_{w \in x} w = \{\varnothing, \{\varnothing\}\}.$$

✍ É obviamente interessante que o leitor crie seus próprios exemplos.

Se $x = \{r, s\}$, denotamos abreviadamente $\bigcup_{w \in x} w$ como $r \cup s$. Neste caso, a união arbitrária é chamada de *união finitária*.

Teorema 3.7. *Se* x *é o par* $\{r, s\}$, *então*
$$\forall t(t \in r \cup s \Leftrightarrow (t \in r \vee t \in s)).$$

✍ A prova deste último fica a cargo do leitor interessado.

No ensino básico o teorema acima é comumente apresentado como definição para união, da seguinte maneira:

$r \cup s$ é o conjunto dos t tais que t pertence a r ou s.

Mas, para os propósitos da matemática a união finitária é insuficiente. Exemplos ilustrativos são apresentados oportunamente.

Teorema 3.8. *União finitária tem elemento neutro, é associativa e é comutativa.*

Formalmente, o teorema acima estabelece que
$$\forall x(x \cup \varnothing = x),$$
ou seja, $\varnothing$ é o elemento neutro mencionado,
$$\forall x \forall y \forall z(x \cup (y \cup z) = (x \cup y) \cup z)$$
e
$$\forall x \forall y(x \cup y = y \cup x),$$
respectivamente.

✍ As demonstrações desses resultados ficam a cargo do leitor interessado. Se o Teorema 3.7 for provado, a demonstração deste último se torna praticamente imediata a partir da lista de 17 teoremas da Seção 10.

Seção 22

Separação

Sumário
Índice
Rede

Considere as seguintes fórmulas (observar o emprego do plural!), conhecidas historicamente como *Esquema da Compreensão*. Se $\mathcal{P}(y)$ é uma fórmula cujas ocorrências de y são livres, então

$$\exists x \forall y (y \in x \Leftrightarrow \mathcal{P}(y)).$$

Nos primórdios da teoria de conjuntos, antes do trabalho de Ernst Zermelo (um dos criadores de ZF), o Esquema da Compreensão era empregado para definir um conjunto x por um predicado monádico $\mathcal{P}(y)$ (ou seja, uma fórmula com ocorrências livres de y). Para cada fórmula $\mathcal{P}$ temos um axioma. Daí o nome *Esquema* da Compreensão!

Ou seja, abreviadamente, o conjunto x, cuja existência era garantida por uma fórmula $\mathcal{P}$ que seus elementos y devem satisfazer, era denotado como

$$x = \{y \mid \mathcal{P}(y)\}$$

(lê-se 'o conjunto x dos elementos y tais que $\mathcal{P}(y)$'). Neste sentido, em particular, o conjunto x de todos os conjuntos pode ser definido como $x = \{y \mid y = y\}$. Com efeito, todo conjunto y é idêntico a si mesmo. Afirmar que y é um conjunto, neste contexto, equivale a afirmar $y \in x$. Como caso especial, temos que $x \in x$.

No entanto, apliquemos o Esquema da Compreensão para definir um outro conjunto x da seguinte maneira:

$$x = \{y \mid y \notin y\}.$$

Neste caso, o predicado monádico $P(y)$ é $y \notin y$. Se $x \in x$, então x deve satisfazer a fórmula em questão. Logo, $x \notin x$. Se $x \notin x$, então x deve pertencer a x, uma vez que satisfaz a fórmula em questão. Logo, $x \in x$. Resumidamente, temos que, neste caso, $x \in x$ e $x \notin x$.

Este é o célebre *Paradoxo de Russell* (1901), o qual mostra que o Esquema da Compreensão é inconsistente com os demais postulados de ZF (uma vez que $x \in x \wedge x \notin x$). Logo, o Princípio da Explosão (Seção 13) garante que, em uma teoria formal com os axiomas ZF1~ZF5 + Esquema da Compreensão, qualquer fórmula é teorema. Tal resultado é obviamente indesejável.

Para evitar essa antinomia, uma possível solução é a adoção do *Esquema de Separação de Zermelo*, como se segue.

ZF6$_{\mathcal{F}}$ - Separação: Se $\mathcal{F}(y)$ é uma fórmula onde não há ocorrências livres de x, então:

$$\forall z \exists x \forall y (y \in x \Leftrightarrow y \in z \wedge \mathcal{F}(y)).$$

O conjunto x do postulado acima (cuja existência é garantida pelo Esquema de Separação) é usualmente denotado por

$$x = \{y \in z \mid \mathcal{F}(y)\}.$$

Neste contexto, a existência de um conjunto x, cujos elementos são termos y tais que $\mathcal{F}(y)$, depende da existência de um conjunto z tal que os termos y pertencem a z. Comumente z é chamado de *conjunto universo*, o qual pode ser qualquer conjunto cuja existência é garantida pelos axiomas de ZF.

Novamente o Axioma da Extensionalidade garante que o conjunto $\{y \in z \mid \mathcal{F}(y)\}$ é único, desde que seja dado o conjunto z, bem como a fórmula $\mathcal{F}$. Ademais, se $\mathcal{F}$ é equivalente a uma fórmula $\mathcal{G}$ (ou seja, $\mathcal{F} \Leftrightarrow \mathcal{G}$), então

$$\{y \in z \mid \mathcal{F}(y)\} = \{t \in z \mid \mathcal{G}(t)\}.$$

O Esquema de Separação permite, entre outras coisas, definir a *diferença* entre conjuntos.

Definição 3.3. *Dados os conjuntos x e y, a* diferença entre x e y *é dada por*

$$x - y = \{t \in x \mid t \notin y\}.$$

Exemplo 3.7. *Sejam $x = \{\varnothing, \{\varnothing\}, \{\{\varnothing\}\}, \{\varnothing, \{\varnothing\}\}\}$ e $y = \bigcup_{w \in x} w$, ou seja, $y = \{\varnothing, \{\varnothing\}\}$. Logo,*

$$x - y = \{\{\{\varnothing\}\}, \{\varnothing, \{\varnothing\}\}\}.$$

Com efeito, os elementos de $x - y$ são aqueles que pertencem a x mas não a y. Analogamente, $y - x = \varnothing$; isso porque $y \subseteq x$.

O Exemplo acima deixa claro que diferença entre conjuntos é não comutativa. Em outras palavras, não é teorema a fórmula

$$x - y = y - x.$$

Importante destacar que a terminologia '*Esquema da Compreensão*' admite outras acepções, além daquela colocada nesta Seção. Em ZF_2 (ZF de segunda ordem), por exemplo, há um esquema de axiomas conhecido pelo mesmo nome, mas que não tem relação alguma com o que foi discutido acima.

Moral da História: Para definir um conjunto x cujos elementos y devem satisfazer a uma fórmula $\mathcal{F}$ é necessário qualificar um conjunto universo z tal que cada y de x pertence a z. Caso contrário, a teoria de conjuntos em tela seria inconsistente. Ou seja, no contexto de $ZF6_{\mathcal{F}}$, $x \subseteq z$. É claro que se, em particular, o conjunto universo z for vazio, para qualquer fórmula $\mathcal{F}$ teremos $x = \varnothing$.

✍ Mostrar por que a adoção do Esquema de Separação no lugar do Esquema de Compreensão evita o Paradoxo de Russell.

Procurar na literatura por outras soluções que evitam o Paradoxo de Russell sem o emprego do Esquema de Separação.

Seção 23

Usando união finitária

Sumário
Índice
Rede

Nesta Seção são dados os primeiros passos para edificar os números naturais a partir dos axiomas de ZF.

Definição 3.4. $S(x) \vdots x \cup \{x\}$. *Lemos* $S(x)$ *como* 'sucessor de x'.

Em outras palavras, t pertence ao sucessor de x sss $t \in x$ ou $t = x$.

Exemplo 3.8. i: $S(\varnothing) = \varnothing \cup \{\varnothing\} = \{\varnothing\}$;

ii: $S(S(\varnothing)) = S(\{\varnothing\}) = \{\varnothing\} \cup \{\{\varnothing\}\} = \{\varnothing, \{\varnothing\}\}$;

iii: $S(S(S(\varnothing))) = S(\{\varnothing, \{\varnothing\}\}) = \{\varnothing, \{\varnothing\}\} \cup \{\{\varnothing, \{\varnothing\}\}\} = \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\}$.

Observar que $\varnothing$ tem zero elementos, $S(\varnothing)$ tem somente um elemento, $S(S(\varnothing))$ tem dois elementos e assim por diante. ✍ Além disso, não existe x tal que $S(x) = \varnothing$ (consegue provar por redução ao absurdo?).

O sucessor $S(x)$ de um conjunto x tem, além de todos os elementos de x, o próprio x como elemento. Ou seja, para qualquer x temos que $x \subset S(x)$. Esse conceito é essencial para que sejamos capazes de finalmente definir números naturais no âmbito de ZF. Mas, uma coisa é definir número natural; outra é definir o conjunto de todos os números naturais. Para que essa distinção essencial seja percebida, introduzimos a seguir o Axioma do Infinito.

ZF7 - Infinito:

$$\exists x(\varnothing \in x \wedge \forall y(y \in x \Rightarrow S(y) \in x)).$$

O Axioma do Infinito garante a existência de pelo menos um conjunto x que satisfaz a conjunção das seguintes fórmulas:

- $\varnothing$ pertence a x;
- se y pertence a x, então o sucessor de y também pertence a x.

Isso produz algo como um 'efeito dominó', no seguinte sentido: uma vez que $\varnothing$ pertence a x, e $\varnothing$ pertencer a x implica que o sucessor de $\varnothing$ pertence a x, então o sucessor de $\varnothing$ também pertence a x; uma vez que o sucessor de $\varnothing$ pertence a x, e o sucessor de $\varnothing$ pertencer a x implica que o sucessor do sucessor de $\varnothing$ pertence a x, então o sucessor do sucessor de $\varnothing$ também pertence a x; e assim por diante. Ou seja, Modus Ponens está sendo usado indefinidamente.

Obviamente o emprego de chaves para denotar sucessor de vazio, sucessor do sucessor de vazio, sucessor do sucessor do sucessor de vazio (e assim por diante) se mostra extremamente inconveniente, além de esteticamente repulsivo. Para contornar tal dificuldade é de interesse o emprego de abreviações metalinguísticas que facilitem a vida do matemático.

Considere um alfabeto D cujos símbolos são $0, 1, 2, 3, 4, 5, 6, 7, 8$ e 9, ordenados por *ordem lexicográfica* de acordo com o sistema decimal usual. Tal ordem lexicográfica (análoga à ordem alfabética de dicionários) é definida da seguinte maneira:

i: 0 é menor do que 1, 1 é menor do que 2, 2 é menor do que 3, 3 é menor do que 4, 4 é menor do que 5, 5 é menor do que 6, 6 é menor do que 7, 7 é menor do que 8, 8 é menor do que 9;

II: os dez símbolos do sistema decimal usual podem ser concatenados para formar sequências finitas como, por exemplo, 1234;

III: se uma sequência finita de símbolos de D conta com mais de um símbolo, então o primeiro é sempre diferente de 0;

IV: se $x_1x_2\cdots x_n$ e $y_1y_2\cdots y_n$ são sequências de n símbolos de D em cada uma, então $x_1x_2\cdots x_n$ é menor do que $y_1y_2\cdots y_n$ sss (i) x_1 é menor do que y_1 ou (ii) $x_1 = y_1$ e x_2 é menor do que y_2, ou (iii) $x_1 = y_1$, $x_2 = y_2$ e x_3 é menor do que y_3, e assim por diante, até o caso em que $x_1 = y_1$, $x_2 = y_2$, $\cdots$, $x_{n-1} = y_{n-1}$ e x_n é menor do que y_n;

V: se $x_1x_2\cdots x_m$ e $y_1y_2\cdots y_n$ são sequências de m e de n símbolos de D, respectivamente, tal que m é menor do que n, então, $x_1x_2\cdots x_m$ é menor do que $y_1y_2\cdots y_n$.

Cada símbolo de D é chamado de *dígito*.

Exemplo 3.9. 1234 *é menor do que* 1244*; com efeito, os dois primeiros dígitos são respectivamente iguais em cada uma, mas o terceiro dígito da primeira é menor do que o terceiro dígito da segunda.*

Logo, podemos adotar as seguintes abreviações:

$$0 \vdots \varnothing; \qquad 1 \vdots S(\varnothing); \qquad 2 \vdots S(S(\varnothing)); \qquad 3 \vdots S(S(S(\varnothing)))$$

e assim por diante.

Em outras palavras, se n denota um termo definido a partir de D e $n+1$ é o termo seguinte pela ordem lexicográfica, então $n+1 = S(n)$, onde $0 = \varnothing$. Ou seja,

$$0 \vdots \varnothing; \qquad 1 \vdots S(0); \qquad 2 \vdots S(S(0)); \qquad 3 \vdots S(S(S(0)))$$

e assim por diante.

Essa codificação a partir do alfabeto D permite dispensar o emprego de chaves, além de se identificar com práticas comuns de notação para números naturais.

Agora fica mais fácil definir e exemplificar adição e multiplicação entre certos conjuntos que pertencem ao x do Axioma do Infinito.

Observar também que

$$\forall x(x \in S(x)).$$

Uma das consequências disso é que, em particular, 0, 1 e 2 são os únicos elementos de 3. Analogamente, $5 = \{0, 1, 2, 3, 4\}$.

Agora podemos finalmente definir certas operações.

Definição de Adição:

I: $+(m, 0) = m$;

II: $+(m, S(n)) = S(+(m, n))$.

Lê-se $+(m, n)$ como '*adição* de m com n' ou '*adição* entre m e n'.

Exemplo 3.10.

$$+(2,3) = S(+(2,2)) = S(S(+(2,1))) = S(S(S(+(2,0)))) = S(S(S(2))) = S(S(3)) = S(4) = 5.$$

Foi provado no Exemplo acima que a adição entre 2 e 3 é 5.

A notação mais usual para adição $+$ segue abaixo.

$m + n \because + (m, n)$.

Se $m + n = p$, dizemos que p é a *soma* das *parcelas* m e n.

Definição 3.5. *Qualquer termo x, cuja existência é garantida pelo Axioma do Infinito ZF6, é chamado de* conjunto indutivo. *Usando o Esquema de Separação é possível definir o conjunto ω dos* números naturais*:*

$$\omega = \{t \in z \mid \forall w(w \text{ é indutivo } \Rightarrow t \in w)\},$$

sendo z um conjunto indutivo.

Ou seja, ω é o conjunto cujos elementos são $\varnothing$, $S(\varnothing)$, $S(S(\varnothing))$ e assim por diante, denotados abreviadamente por 0, 1, 2 etc. Esse conjunto ω é chamado de *conjunto dos números naturais*. Cada elemento de ω é um *número natural*.

A razão para definir ω a partir do Axioma do Infinito e do Esquema de Separação é a seguinte: o Axioma do Infinito é consistente com a existência de conjuntos indutivos diferentes de ω.

Por exemplo, considere o conjunto indutivo x cujos elementos são os números naturais e, além deles, os conjuntos

$$\{\{\varnothing\}\},\ S(\{\{\varnothing\}\}),\ S(S(\{\{\varnothing\}\}))$$

e assim por diante.

Claramente existem termos pertencentes a x que não são números naturais, a saber, $\{\{\varnothing\}\}$, $S(\{\{\varnothing\}\})$, $S(S(\{\{\varnothing\}\}))$ etc. Analogamente, podem existir muitos outros conjuntos indutivos diferentes de ω. No entanto, a definição de conjunto indutivo dada acima garante que, se x é indutivo, então $\omega \subseteq x$. Dessa maneira, o emprego do Esquema de Separação na Definição 3.5 garante a definição do conjunto ω cujos elementos são aqueles que ocorrem obrigatoriamente em todos os conjuntos indutivos. Tais elementos são exatamente os números naturais. O Axioma da Extensionalidade permite provar que ω é único. Logo, ω é uma constante de ZF.

O Axioma da Extensionalidade permite provar a unicidade do conjunto vazio, da potência de um conjunto qualquer, da união arbitrária sobre um conjunto qualquer, de um par qualquer, mas não de conjuntos indutivos. Daí a necessidade das considerações feitas no último parágrafo!

✍ Notar também que a definição de adição dada acima viola o critério de eliminabilidade introduzido na Seção 14, se aplicarmos essa adição sobre termos pertencentes a um conjunto indutivo x diferente de ω. Portanto, o que foi introduzido como adição é uma definição explícita abreviativa somente para os termos pertencentes a ω. Consegue provar isso?

Adição entre naturais permite definir o que é um natural m menor ou igual a um natural n.

Definição 3.6. *Sejam a e b naturais. Logo,*

$$a \leq b \;\vdots\; \exists c(c \in \omega \wedge b = a + c).$$

Lemos $a \leq b$ como 'a é menor ou igual a b'.

Exemplo 3.11. i: $2 \leq 5$. *Com efeito,* $5 = 2 + 3$*;*

ii: $2 \leq 2$. *Com efeito,* $2 = 2 + 0$.

Por abuso de notação, é usual escrever $m \leq n \leq p$ como abreviação para $m \leq n \wedge n \leq p$.

Considere a seguinte fórmula:

$$\exists! x(\varnothing \in x \wedge \forall y(y \in x \Rightarrow S(y) \in x)).$$

Chamemos a fórmula acima de ZF7MALUCO.

✍ Uma teoria que tivesse todos os axiomas de ZF e, além disso, a fórmula ZF7MALUCO, seria inconsistente, no sentido de que haveria alguma fórmula $\mathcal{F}$ nessa nova teoria tal que ambas $\mathcal{F}$ e $\neg\mathcal{F}$ seriam teoremas. Consegue provar isso?

✍ Para dificultar um pouco mais, considere uma teoria que tivesse todos os axiomas de ZF, mas com ZF7MALUCO substituindo ZF7. Consegue provar que essa teoria também seria inconsistente?

DEFINIÇÃO DE MULTIPLICAÇÃO:

I: $\cdot(0, n) = 0$;

II: $\cdot(S(m), n) = \cdot(m, n) + n$.

Lê-se $\cdot(m, n)$ como 'a *multiplicação* de m com n' ou 'a *multiplicação* entre m e n'. Obviamente, levando em consideração comentários anteriores, devemos assumir que m e n são naturais, i.e., elementos de ω.

EXEMPLO 3.12.

$$\cdot(3,2) = \cdot(2,2) + 2 = (\cdot(1,2) + 2) + 2 =$$
$$((\cdot(0,2) + 2) + 2) + 2 = ((0 + 2) + 2) + 2 =$$
$$(2 + 2) + 2 = 4 + 2 = 6.$$

Foi provado, no EXEMPLO acima, que a multiplicação entre 3 e 2 é 6. A notação mais usual para a multiplicação $\cdot$ entre números naturais é a que segue abaixo.

$mn \vdots \cdot (m, n)$.

No entanto, é também usual denotar a multiplicação acima por $m \cdot n$. Se $mn = p$, dizemos que p é o *produto* dos *fatores* m e n.

Existem propriedades algébricas de estratégica importância para as operações de adição $+$ e multiplicação $\cdot$ em ω. Discutimos sobre isso na Seção 29.

✍ Observar que não existe x tal que $S(x) = \omega$. Em outras palavras, ω não é sucessor de conjunto algum.

O Esquema de Separação, em parceria com os postulados União e Potência, permite definir um conceito útil para a matemática:

Definição 3.7.

$$x \times y \,\vdots\, \{(a,b) \in \wp(\wp(x \cup y)) \mid a \in x \wedge b \in y\}.$$

O termo $x \times y$ se lê '*produto cartesiano* de x por y'. Observar que, no emprego do Esquema da Separação acima, o conjunto universo é $\wp(\wp(x \cup y))$.

Exemplo 3.13. *Sejam* $x = \{0,3,1\}$ *e* $y = \{1,2\}$. *Logo,*

$$x \cup y = \{0,1,2,3\}.$$

Portanto,

$$\wp(x \cup y) = \{\varnothing, \{0\}, \{1\}, \{2\}, \{3\}, \{0,1\}, \{0,2\}, \{0,3\}, \{1,2\}, \{1,3\}, \{2,3\}, \{0,1,2\}, \{0,1,3\}, \{0,2,3\}, \{1,2,3\}, x \cup y\}.$$

O termo $\wp(\wp(x \cup y))$ *conta com 65.536 elementos. Entre eles, temos os seguintes:*

$$\{\{0\},\{0,1\}\}, \{\{0\},\{0,2\}\}, \{\{1\}\}, \{\{1\},\{1,2\}\}, \{\{3\},\{1,3\}\}$$

e $\{\{3\},\{2,3\}\}$.

Mas estes são exatamente os pares ordenados $(0,1)$, $(0,2)$, $(1,1)$, $(1,2)$, $(3,1)$ *e* $(3,2)$, *respectivamente. Logo,*

$$x \times y = \{(0,1),(0,2),(1,1),(1,2),(3,1),(3,2)\}.$$

Exemplo 3.14. *Sejam* $x = \{0,3,1\}$ *e* $y = \{1,2\}$. *Logo,*

$$y \times x = \{(1,0),(1,1),(1,3),(2,0),(2,1),(2,3)\}.$$

Observar que $x \times y \neq y \times x$, *pelo menos neste exemplo.*

Logo, produto cartesiano é não comutativo.

Exemplo 3.15. *Sejam* $x = \{0,3,1\}$ *e* $y = \{1,2\}$. *Logo,*

$$(x \times y) \times x \neq x \times (y \times x).$$

Com efeito, $((0,1),2) \in (x \times y) \times x$, *mas* $((0,1),2) \notin x \times (y \times x)$.

No último Exemplo acima fica claro que produto cartesiano é não associativo. Logo, o emprego de parênteses é necessário, no caso de produtos cartesianos envolvendo três ou mais conjuntos.

O número de maneiras de parentesar o produto cartesiano

$$x_0 \times x_1 \times x_2 \times \cdots \times x_n,$$

entre $n+1$ ocorrências de termos, é o *Número de Catalan* [51]

$$C_n = \frac{(2n)!}{(n+1)!n!}.$$

Se o leitor não recorda o que é o *fatorial* $n!$ de um número natural n, aqui vai.

Definição 3.8. i $0! = 1$;

ii $(n+1)! = (n+1)n!$.

Exemplo 3.16. $5!$, *de acordo com o item* ii *da definição de fatorial, é igual a* $5(4!)$. *Aplicando novamente o item* ii, *para calcular* $4!$, *temos que* $5! = 5(4(3!))$. *Aplicando de novo, até chegarmos a* $0!$ *(este é o critério de parada), temos* $5! = 5(4(3(2!))) = 5(4(3(2(1!)))) = 5(4(3(2(1(0!))))) = 120$, *uma vez que item* i *diz que* $0! = 1$.

Exemplo 3.17. Sobre Número de Catalan. *Sabemos que*

$$C_3 = \frac{(2(3))!}{(3+1)!3!}.$$

Logo, $C_3 = 5$. *Logo, é possível parentesar*

$$x_0 \times x_1 \times x_2 \times x_3$$

de cinco maneiras distintas, cada uma produzindo um produto cartesiano diferente dos demais (se todos os conjuntos envolvidos são diferentes de $\varnothing$*).*

As cinco maneiras mencionadas no último Exemplo são as seguintes:

i: $((x_0 \times x_1) \times x_2) \times x_3$;

ii: $(x_0 \times x_1) \times (x_2 \times x_3)$;

iii: $x_0 \times ((x_1 \times x_2) \times x_3)$;

iv: $(x_0 \times (x_1 \times x_2)) \times x_3$;

v: $x_0 \times (x_1 \times (x_2 \times x_3))$.

No entanto, é uma prática comum introduzir a seguinte definição, a qual é aplicada frequentemente em matemática e contorna a dificuldade de lidar com a não associatividade de produto cartesiano:

> Definição 3.9. *Seja x um conjunto. Logo,*
>
> i: $x^2 = x \times x$ *(lê-se 'x 2');*
>
> ii: $x^{n+1} = x \times x^n$ *(lê-se 'x $n + 1$');*
>
> *onde n é um número natural diferente de 0.*

A definição acima evita qualquer ambiguidade no cálculo de, por exemplo,

$$x^3 = x \times x \times x$$

(lê-se 'x 3').

Neste caso,

$$x^3 = x \times x^2 = x \times (x \times x).$$

Os elementos de x^n (lê-se 'x n') são chamados de n-uplas ordenadas

$$(a_1, a_2, \cdots, a_n),$$

onde cada a_i $(1 \leq i \leq n)$ é elemento de x. Em particular, cada 3-upla ordenada (também chamada de *tripla ordenada*) de x^3 é o termo

$$(a_1, a_2, a_3) = (a_1, (a_2, a_3)),$$

onde a_1, a_2 e a_3 são elementos de x. Isso significa que a Definição 3.9 permite generalizar a definição de Kuratowski para par ordenado. Uma n-upla ordenada $(a_1, a_2, \cdots, a_n)$ é tão somente elemento de x^n para algum conjunto x tal que cada a_i (onde $1 \leq i \leq n$) pertence a x. Em outras palavras, toda n-upla ordenada é um caso particular de par ordenado. Uma vez que todo par ordenado é um caso particular de par não ordenado, então toda n-upla ordenada é um par não ordenado.

A não comutatividade de produto cartesiano permite introduzir os conceitos de relação e função, como vemos nas Seções 25 e 29.

✍ Exibir conjunto x diferente de ω e diferente de $\varnothing$ de modo que x não seja sucessor de conjunto algum. Dica: pelo menos um exemplo já foi apresentado nesta Seção!

Seção 24

Substituição, Regularidade e Escolha

Sumário
Índice
Rede

✂ O objetivo desta Seção é encerrar os axiomas próprios de ZF.

Os axiomas de Substituição e de Regularidade de ZF não são necessários para provar os resultados de interesse para aqueles que estão focados em temas do cotidiano da maioria dos matemáticos, como cálculo diferencial e integral, análise matemática, espaços métricos, equações diferenciais, álgebra linear, análise funcional, topologia, álgebra, probabilidades, teoria dos números, geometria diferencial, teoria de reticulados, matemática *fuzzy*, geometria euclidiana, geometrias não euclidianas, geometria absoluta, geometria projetiva, entre outros temas. Isso ocorre apesar de todas essas áreas poderem ser fundamentadas com os axiomas até aqui apresentados. Ou seja, os resultados mais populares de tais áreas do conhecimento podem ser escritos como teoremas de ZF, bastando os axiomas ZF1∼ZF7. Logo, o leitor não é prejudicado se ignorar esta Seção.

No entanto, se o leitor estiver interessado em questões ligadas aos *fundamentos da matemática* (como epistemologia e metodologia da matemática), esses postulados desempenham papel estratégico e necessário.

> ZF8$_{\mathcal{F}}$ - Substituição: Seja $\mathcal{F}(x, y)$ uma fórmula onde todas as ocorrências de x e y são livres; logo,
>
> $$\forall x \exists ! y \mathcal{F}(x, y) \Rightarrow \forall z \exists w \forall t (t \in w \Leftrightarrow \exists s (s \in z \wedge \mathcal{F}(s, t))).$$

Substituição (o qual não pode ser confundido com a substitutividade da igualdade) é um esquema de axiomas. Com efeito, há um axioma para cada fórmula $\mathcal{F}$, desde que $\mathcal{F}$ satisfaça as condições sintáticas acima impostas.

O *Esquema da Substituição* (como também é conhecido) é aplicável somente a fórmulas $\mathcal{F}(x, y)$ tais que, para qualquer conjunto x existe um único y tal que $\mathcal{F}(x, y)$. É exatamente isso que está escrito antes da primeira ocorrência da condicional $\Rightarrow$ em ZF8$_{\mathcal{F}}$. Exemplos de fórmulas $\mathcal{F}(x, y)$ desse tipo são os seguintes:

I: $y = x$;

II: $y = \wp(x)$;

III: $y = \bigcup_{w \in x} w$;

IV: $y = S(x)$;

V: $y = \wp(x) \cup S(x)$;

VI: $(x = \varnothing \Rightarrow y = \{\varnothing\}) \wedge (x \neq \varnothing \Rightarrow y = \{\{\varnothing\}\})$;

entre uma infinidade de outros. Neste caso, o termo y é chamado de *imagem* de x pela fórmula $\mathcal{F}$.

Exemplo de fórmula $\mathcal{F}(x, y)$ que não atende às exigências do Esquema de Substituição: $y \subseteq x$. Com efeito, um mesmo conjunto x pode admitir mais de um subconjunto y (basta que x seja diferente de $\varnothing$). Outro exemplo de fórmula $\mathcal{F}(x, y)$ que não atende às exigências do Esquema de Substituição: $y \neq x$.

O Esquema da Substituição estabelece o seguinte, desde que a fórmula $\mathcal{F}(x, y)$ atenda às exigências já mencionadas: dado um conjunto z, existe um conjunto w tal que, cada elemento t de w é imagem de um termo s pertencente a z pela fórmula $\mathcal{F}(s, t)$. Em particular, se o conjunto z é $\{\varnothing\}$ e a fórmula $\mathcal{F}(x, y)$ é aquela do exemplo VI dado acima, então $w = \{\{\varnothing\}\}$.

Ou seja, o Esquema da Substituição permite garantir a existência de conjuntos w a partir de conjuntos z e fórmulas. É um papel semelhante ao do Esquema de Separação. No entanto, no caso de Substituição, o conjunto w não é necessariamente subconjunto de z.

Na Seção 111 é provado que, graças ao Esquema da Substituição, o Axioma do Par é desnecessário em ZF. Tradicionalmente, o Axioma do Par é mantido por motivos didáticos.

ℹ Foi mencionado anteriormente que $\varnothing$ e ω são exemplos de conjuntos que não são sucessores de qualquer outro conjunto. Pois bem. Há uma generalização dos números naturais que permite conceituar *ordinais*, no sentido de que todo natural é um ordinal. Mas em ZF existem outros ordinais além dos elementos de ω. O próprio ω é um ordinal, bem como o sucessor de ω, o sucessor do sucessor de ω e assim por diante. Usando o Esquema de Substituição é possível provar a existência de outros ordinais λ tais que λ não é sucessor de conjunto algum. Esse resultado permite provar diversos teoremas com impacto profundo até mesmo em ramos como teoria da medida e, consequentemente, em análise matemática [52]. Mas este é um assunto que vai muito além da proposta deste livro.

O próximo postulado desempenha um papel radicalmente diferente dos demais. Para facilitar a sua leitura, introduzimos nova abreviação metalinguística.

Definição 3.10. *$x \cap y = \{r \in x \cup y \mid r \in x \wedge r \in y\}$, sendo x e y conjuntos.*

Ou seja, $x \cap y$ é o conjunto dos termos que pertencem a ambos x e y.

O símbolo $\cap$ é chamado de *interseção finitária* ou, simplesmente, *interseção*. Lê-se $x \cap y$ como 'x interseção y' ou 'interseção de x com y'. Para o conceito de *interseção arbitrária*, ver Definição 9.3.

Exemplo 3.18. I: *Sejam $x = \{2, 3\}$ e $y = \{3, 4\}$; portanto, $x \cap y = \{3\}$;*

II: *sejam $x = \{2, 3\}$ e $z = \{4, 5\}$; logo, $x \cap z = \varnothing$.*

Se x e y são conjuntos tais que $x \cap y = \varnothing$, dizemos que x e y são *disjuntos*. Item II do Exemplo acima ilustra um caso de conjuntos disjuntos.

Teorema 3.9. *Sejam x, y e z conjuntos. Logo*

I: *$x \cap y = y \cap x$;*

II: *$x \cap (y \cap z) = (x \cap y) \cap z$;*

III: *$x \cap \varnothing = \varnothing$;*

IV: *$x \cap (y \cup z) = (x \cap y) \cup (x \cap z)$.*

✍ A demonstração é recomendada como exercício ao leitor.

Agora podemos enunciar o próximo postulado próprio de ZF.

ZF9 - Regularidade: $\forall x(x \neq \varnothing \Rightarrow \exists y(y \in x \wedge x \cap y = \varnothing))$.

Também conhecido como *Axioma da Boa Fundação*, o *Axioma da Regularidade* garante que qualquer conjunto x não vazio admite pelo menos um elemento y que não compartilha qualquer elemento em comum com x. O objetivo deste axioma não é garantir a existência de conjuntos, mas proibir a existência de termos x onde ocorram cadeias infinitas de pertinência como

$$x \in y \in x \in y \in x \cdots .$$

O mesmo postulado também impede a existência de conjuntos que pertençam a si mesmos.

Graças ao Axioma da Regularidade é possível introduzir uma definição alternativa para par ordenado (diferente daquela devida a Kuratowski):

$$(a, b) = \{a, \{a, b\}\}.$$

Dessa maneira, um par de chaves se mostra desnecessário. Portanto, a definição de par ordenado devida a Kuratowski pode ser usada tanto em ZF quanto em variações de ZF que abrem mão do Axioma de Regularidade.

Uma das vantagens mais significativas do Axioma da Regularidade é o fato de que ele permite definir o conceito de *rank* de um conjunto. No entanto, novamente este é um tema que vai além dos propósitos desta pequena obra.

Finalmente, os axiomas próprios ZF1∼ZF9 encerram todos os postulados próprios de ZF.

Uma variação de ZF, conhecida como ZFC (a letra C se refere à palavra ‘Choice’ em inglês, a qual se traduz como ‘Escolha’) acrescenta o seguinte postulado.

ZF10 - Escolha:

$$\forall x(\forall y \forall z((y \in x \wedge z \in x \wedge y \neq z) \Rightarrow (y \neq \varnothing \wedge y \cap z = \varnothing)) \Rightarrow$$
$$\exists y \forall z(z \in x \Rightarrow \exists w(y \cap z = \{w\}))).$$

O *Axioma da Escolha* afirma o seguinte: dado um conjunto x cujos elementos são conjuntos não vazios e sem quaisquer elementos em comum, então existe um *conjunto escolha* y tal que cada elemento de y é um, e apenas um, elemento de cada elemento de x.

Bertrand Russell introduziu uma analogia para facilitar a compreensão do Axioma da Escolha: considere uma gaveta com uma quantia infinita de pares de meias, de modo que cada par de meias é facilmente discernível de todos os demais; neste caso o Axioma da Escolha permite definir uma nova gaveta que terá uma, e apenas uma, meia de cada par da primeira gaveta.

Russell fez a analogia com pares do meias por conta de um fato simples: o pé esquerdo é indiscernível do pé direito em qualquer par de meias. Isso significa que não é possível estabelecer qualquer critério para a escolha de uma meia de cada par (algo bem diferente de uma gaveta de sapatos). Logo, é isso o que o Axioma da Escolha faz! Ele permite *escolher* elementos quaisquer de conjuntos dados sem estabelecer qualquer critério. Apenas escolhe, como em um ato de inquestionável *livre arbítrio*.

O Axioma da Escolha, introduzido por Ernst Zermelo em 1904, provocou enorme debate entre matemáticos do início do século passado. Parte das críticas era sustentada pelo caráter não construtivo deste postulado, no sentido de o mesmo não estabelecer critérios de escolha. Parte das críticas ocorria por conta de resultados contra-intuitivos que eram consequências do Axioma da Escolha, como o Teorema de Banach-Tarski. Hoje se sabe que tal postulado apenas permite desenvolver novas formas de matemática. Atualmente ele exerce enorme impacto sobre a matemática, como os seguintes resultados:

i: Todo conjunto admite uma *boa ordem*. Relações de boa ordem sobre um conjunto x são relações de ordem total $\leq$ (ver Seção 25) tais que qualquer subconjunto de x admite um menor elemento relativamente a $\leq$.

ii: O Princípio de Partição (PP) é teorema de ZFC. PP é uma fórmula envolvendo funções. Detalhes na Seção 112.

iii: O Teorema de Tychonov, o qual é aplicado no estudo de topologia geral.

iv: Todo espaço vetorial não trivial admite base (ver Seção 96), o qual é um resultado de análise funcional.

entre centenas de outros. No entanto, a maioria desses resultados está fora do escopo dos interesses deste texto.

Em 1938 Kurt Gödel provou que, se ZF for consistente, então ZFC é consistente. Em 1963 Paul Cohen provou que

$$\nvdash_{ZF} \text{Escolha}$$

e

$$\nvdash_{ZF} \neg(\text{Escolha}),$$

ou seja, nem o Axioma da Escolha ou a sua negação são teoremas em ZF. A revolucionária técnica criada por Cohen para garantir tal resultado rendeu a ele a única Medalha Fields destinada a uma contribuição em lógica. Detalhes na Seção 111.

Há na literatura muitas outras variações de ZF, além de ZFC, sendo que algumas delas contam com impacto significativo sobre a prática matemática.

Seção 25

Relações

Sumário
Índice
Rede

Para que sejamos capazes de introduzir números inteiros, racionais, reais e complexos, precisamos qualificar o que são relações.

Definição 3.11.

$$r \textit{ é uma } \text{relação} \vdots\, \exists x \exists y (r \subseteq x \times y).$$

Lemos 'r é uma relação com *domínio* x e *co-domínio* y' ou, simplesmente, 'r é uma relação de x em y'. A notação usual é

$$r : x \to y.$$

Lembrar que produto cartesiano é uma operação não comutativa, conforme Exemplo 3.14.

Esse fato permite a discriminação entre domínio e co-domínio de uma relação r, uma vez que toda relação é subconjunto de um produto cartesiano. Alguns autores se referem ao co-domínio de uma relação como *contradomínio*.

Exemplo 3.19. *Sejam $x = \{1, 2\}$ e $y = \{2, 3\}$. São exemplos de relações de x em y os seguintes conjuntos:*

i: *$\varnothing$; com efeito, vazio é subconjunto de qualquer conjunto, de acordo com o Teorema 3.6; logo,*

$$\varnothing : x \to y$$

é relação para quaisquer x e y;

ii: *$x \times y$; com efeito, todo conjunto é subconjunto de si mesmo, de acordo com o Teorema 3.5;*

iii: *$\{(1, 2)\}$; com efeito, todo elemento de $\{(1, 2)\}$ pertence a $\{2, 3\} \times \{2, 3\}$;*

iv: $\{(1,2),(1,3)\}$;

v: $\{(1,2),(2,2)\}$;

vi: $\{(1,3),(2,2)\}$.

✍ *Para os conjuntos x e y aqui sugeridos, há 16 possíveis relações de x em y. Listamos aqui apenas seis delas. Cabe ao leitor listar outros exemplos.*

Definição 3.12. *Uma* relação r em x *é qualquer subconjunto de $x \times x$.*

Ou seja, uma relação em x é uma relação cujo domínio é idêntico ao seu co-domínio. Em particular, a relação

$$r = \{(a,b) \in \wp(\wp(x)) \mid a \in x \wedge b \in x \wedge a = b\}$$

é a *diagonal do conjunto x*.

Exemplo 3.20. *A diagonal de ω é o conjunto*

$$d = \{(m,n) \in \omega \times \omega \mid m = n\}.$$

Definição 3.13. *Uma relação r em x é:*

i: reflexiva *sss* $\forall a(a \in x \Rightarrow (a,a) \in r)$;

ii: simétrica *sss* $\forall a \forall b((a,b) \in r \Rightarrow (b,a) \in r)$;

iii: transitiva *sss* $\forall a \forall b \forall c(((a,b) \in r \wedge (b,c) \in r) \Rightarrow (a,c) \in r)$.

Exemplo 3.21. i: ✍ *A diagonal de qualquer conjunto é uma relação reflexiva, simétrica e transitiva; consegue provar isso?*

ii: ✍ *Seja r uma relação em ω tal que*

$$(a,b) \in r \text{ sss } a + b = 2n,$$

para algum $n \in \omega$; logo, r é reflexiva, simétrica e transitiva. Consegue provar isso? Observar que, neste caso, r é o conjunto

$$\{(0,0),(0,2),(0,4),\cdots,(1,1),(1,3),(1,5),\cdots,$$
$$(2,0),(2,2),(2,4),\cdots\}.$$

Se r é uma relação em x, então

$$arb \vdots (a,b) \in r.$$

A notação introduzida acima é muito comum em matemática, como ilustrado no próximo parágrafo.

Sejam a e b elementos de ω. Logo, $\leq$ é uma relação em ω tal que

$$a \leq b \vdots \exists c(c \in \omega \wedge b = a + c).$$

É interessante contrastar essa última definição com a Definição 3.6. Isso porque $\leq$ é o conjunto

$$\{(0,0),(0,1),(0,2),\cdots,(1,1),(1,2),(1,3),\cdots,$$
$$(2,2),(2,3),(2,4),\cdots\}.$$

Exemplo 3.22. i: $2 \leq 2$, *ou seja,* $(2,2) \in \leq$. *Com efeito,* $2 = 2 + 0$.

ii: $2 \leq 5$, *ou seja,* $(2,5) \in \leq$. *Com efeito,* $5 = 2 + 3$*;*

iii: ✍ $\neg(5 \leq 2)$, *ou seja,* $(5,2) \notin \leq$. *Com efeito, não existe natural* c *tal que* $2 = 5 + c$. *Cabe ao leitor provar este último resultado por redução ao absurdo.*

A relação $\leq$ em ω é reflexiva e transitiva, apesar de não ser simétrica. No entanto, a diagonal d de ω é subconjunto próprio de $\leq$, i.e.,

$$d \subset \leq .$$

Isso significa que uma relação simétrica pode estar contida em uma relação não simétrica.

Sejam a e b elementos de ω. Logo,

$$a < b \vdots a \leq b \wedge a \neq b.$$

Observar que

$$<=\{(0,1),(0,2),(0,3),\cdots,(1,2),(1,3),(1,4),\cdots,$$
$$(2,3),(2,4),(2,5),\cdots\}.$$

A relação $<$ em ω é transitiva, mas não reflexiva e nem simétrica.

Exemplo 3.23. i: $\neg(2 < 2)$; *com efeito, apesar de* $2 \leq 2$, $2 = 2$. *Este resultado prova a não reflexividade de* $<$ *em* ω.

ii: $2 < 5$; *com efeito,* $2 \leq 5$ *e* $2 \neq 5$.

Relações encontram ampla aplicabilidade em matemática. Entre elas, estão as célebres *relações de equivalência*:

Definição 3.14. *Uma relação* r *em* x *é de* equivalência *sss* r *é reflexiva, simétrica e transitiva.*

Se r é uma relação de equivalência em x, lemos arb como 'a é equivalente a b relativamente a r', ou simplesmente, 'a é equivalente a b', se não houver risco de confusão.

Exemplo 3.24. i: *A diagonal de qualquer conjunto* x *é uma relação de equivalência; logo, se* a *pertence a* x*, então apenas* a *é equivalente a* a *em relação à diagonal de* x*;*

ii: *seja* r *uma relação em* ω *tal que* $(a, b) \in r$ *sss* $a + b = 2n$*, para algum* $n \in \omega$*; logo,* r *é uma relação de equivalência; além disso, o natural* 2 *é equivalente a* 0, 2, 4, 6 *etc., enquanto o natural* 3 *é equivalente a* 1, 3, 5, *etc., relativamente a* r.

✍ Prove todas as afirmações dos dois últimos Exemplos.

Seção 26

Classes de Equivalência e Partições

Sumário
Índice
Rede

O primeiro exemplo de aplicação de relações de equivalência é dado na Seção 30. Mas, antes disso, são necessárias mais informações.

Definição 3.15. *Seja* $\sim$ *uma relação de equivalência em* x. *Logo,*

$$[a] = \{t \in x \mid t \sim a\}.$$

Chamamos $[a]$ *de* classe de equivalência *de* x *relativamente a* $\sim$ *(ou apenas* classe de equivalência*, se não houver risco de confusão). Qualquer elemento* $b \in [a]$ *é chamado de* representante *da classe de equivalência* $[a]$.

Observar que está sendo usado o Esquema de Separação para definir a classe de equivalência $[a]$ de x relativamente a $\sim$.

Exemplo 3.25. *Seja r a relação de equivalência em ω definida por*

$$(a, b) \in r \text{ sss } a + b = 2n,$$

para algum $n \in \omega$. Logo,

$$[0] = \{0, 2, 4, 6, \cdots\}$$

e

$$[1] = \{1, 3, 5, 7, \cdots\}.$$

Observar que

$$[0] = [2] = [4] = [2n],$$

para qualquer natural n, enquanto

$$[1] = [3] = [5] = [2n + 1],$$

para qualquer natural n.

A classe de equivalência $[0]$ é chamada de conjunto dos naturais pares, *enquanto $[1]$ é o conjunto dos naturais* ímpares.

O leitor deve observar que, no Exemplo acima, o conjunto ω dos números naturais conta com apenas duas classes de equivalência relativamente a r: $[0]$ e $[1]$. Ademais,

$$[0] \cap [1] = \varnothing$$

e

$$[0] \cup [1] = \omega.$$

Exemplo 3.26. *Como caso particular do* Exemplo 3.24, *a diagonal d de ω é uma relação de equivalência em ω. Logo, para qualquer natural n, temos*

$$[n] = \{n\}.$$

Neste caso, cada classe de equivalência é um singleton.

No Exemplo acima o conjunto ω dos números naturais conta com uma infinidade de classes de equivalência relativamente à diagonal. Além disso, se $m \neq n$, então

$$[m] \cap [n] = \varnothing.$$

De maneira análoga ao caso anterior, temos

$$\bigcup_{n\in\omega}[n] = \omega.$$

Nos dois últimos Exemplos mostramos classes de equivalência distintas que são disjuntas, bem como o fato de que a união arbitrária de todas elas é o próprio conjunto ω. Tal fenômeno é onipresente em relações de equivalência, como mostramos adiante.

Para o conceito de *partição* de um conjunto, dado a seguir, é empregada a Definição 3.10. Partições são essenciais para compreendermos classes de equivalência.

Definição 3.16. *Seja x um conjunto. Dizemos que p é uma* partição *de x sss*

i: $\forall t(t \in p \Rightarrow (t \neq \varnothing \land t \subseteq x))$;

ii: $\forall r\forall s((r \in p \land s \in p) \Rightarrow (r = s \lor r \cap s = \varnothing))$; *e*

iii: $\bigcup_{t\in p} t = x$.

Ou seja, uma partição p de um conjunto x é um conjunto de subconjuntos de x (i.e., $p \subset \wp(x)$) tal que

i: cada elemento de p é não vazio;

ii: dois elementos distintos de p têm interseção vazia, ou seja, são disjuntos; e

iii: a união arbitrária de todos os elementos de p é igual a x.

Exemplo 3.27. i: *Seja r uma relação em ω tal que*

$$(a, b) \in r \text{ sss } a + b = 2n,$$

para algum $n \in \omega$. Logo,

$$p = \{[0], [1]\}$$

é uma partição de ω, conforme Exemplo 3.25. *Afinal, $p \subset \wp(\omega)$, $[0] \neq \varnothing$, $[1] \neq \varnothing$, $[0] \cap [1] = \varnothing$ e $[0] \cup [1] = \omega$.*

ii: *Seja d a diagonal de ω. Logo,*

$$p = \{y \in \wp(\omega) \mid y \text{ é unitário}\}$$

é uma partição de ω (ver Exemplo 3.26*). Com efeito, $p \subset \wp(\omega)$, uma vez que p é o conjunto de todos os singletons $\{n\}$, onde n é um natural. Observar que cada elemento de*

p é uma classe de equivalência $[n]$ relativamente à diagonal d de ω. Além disso, cada singleton $[n] = \{n\}$ é diferente de $\varnothing$; se $[m] \neq [n]$, então $[m] \cap [n] = \varnothing$; e $\bigcup_{[n]\in p}[n] = \omega$.

Nos dois últimos Exemplos partições foram definidas como conjuntos de classes de equivalência. Isso não é uma mera coincidência, como se percebe nos próximos dois teoremas.

Teorema 3.10. *Seja $\sim$ uma relação de equivalência em x. Logo,*

$$p = \{y \in \wp(x) \mid \exists r(r \in x \wedge y = [r])\}$$

é uma partição de x, onde $[r] = \{t \in x \mid t \sim r\}$.

Demonstração: A definição de classe de equivalência garante que cada uma delas é subconjunto não vazio de x, o que satisfaz item i da Definição 3.16. A reflexividade de $\sim$ garante que $\bigcup_{[r]\in p}[r] = x$ (item iii da Definição 3.16). Se $z \in [r] \cap [s]$, então $r \sim z$ e $s \sim z$. Logo, pela transitividade e pela simetria de $\sim$, temos $r \sim s$, o que implica em $[r] = [s]$ (item ii da Definição 3.16).

Teorema 3.11. *Toda partição p de qualquer conjunto x define uma relação de equivalência.*

Demonstração: Basta definir $\sim$ como se segue:

$$r \sim s \text{ sss ambos } r \text{ e } s \text{ pertencem ao mesmo } y,$$

sendo $y \in p$. Reflexividade é imediata. Se $r \in y$ e $s \in y$, então $s \in y$ e $r \in y$. Logo, temos simetria. Finalmente, se $r \in y$ e $s \in y$, e $s \in y$ e $t \in y$, então $r \in y$ e $t \in y$ (transitividade).

Se x é um conjunto e $\sim$ é uma relação de equivalência em x, então a partição

$$x/\!\sim\; = \{y \in \wp(x) \mid \exists r(r \in x \wedge y = [r])\}$$

é também conhecida como o *quociente de x por $\sim$*.

Por abuso de notação, é usual escrever $x/\!\sim\; = \{[r] \mid r \in x\}$ (de forma alguma isso significa que está sendo usado o Esquema de Compreensão, uma vez que esta é apenas uma notação abusiva, mas muito frequente na literatura). Ou seja, o quociente de um conjunto por uma relação de equivalência é apenas uma partição induzida pela relação de equivalência. O que legitima tal definição é o Teorema 3.10. O que garante a unicidade de $x/\sim$, para cada x, é o Axioma de Extensionalidade.

Definição 3.17. *Uma relação r em x é de* ordem parcial *sss r é reflexiva, transitiva e antissimétrica, sendo que antissimetria se traduz formalmente pela fórmula*

$$\forall a \forall b(((a,b) \in r \wedge a \neq b) \Rightarrow (b,a) \notin r).$$

Exemplo 3.28. i: ✍ *A diagonal de um conjunto x é também conhecida como* igualdade em x *(a qual não pode ser confundida com o predicado binário $=$ do alfabeto da linguagem de ZF). A igualdade em qualquer conjunto x é uma relação de equivalência e uma relação de ordem parcial.*

ii: ✍ *A relação $\leq$ em ω é de ordem parcial.*

Definição 3.18. *Uma relação de ordem parcial r em um conjunto x é de* ordem total *sss*

$$\forall a \forall b((a \in x \wedge b \in x) \Rightarrow ((a,b) \in r \vee (b,a) \in r)).$$

Exemplo 3.29. i: *A relação $\leq$ em ω é de ordem total; com efeito, para quaisquer naturais m e n temos $m \leq n$ ou $n \leq m$;*

ii: *A relação $<$ em ω não é de ordem total; com efeito $\neg(2 < 2)$.*

Seção 27

Resumo da ópera

Sumário

Índice

Rede

O que vimos nesta Parte pode ser resumido da seguinte maneira.

- Os axiomas próprios de ZF servem a dois propósitos: *garantir a existência de certos conjuntos* e *descrever propriedades do predicado de pertinência* $\in$.

- Entre os axiomas *reguladores* (aqueles que apenas descrevem propriedades de $\in$) estão o Axioma da Extensionalidade e o Axioma da Regularidade. O primeiro estabelece critérios para a identificação de conjuntos. O segundo proíbe cadeias infinitas de pertinência. Entre os axiomas *existenciais* (aqueles que garantem a existência de conjuntos) estão todos os demais.
- O ponto de partida para a existência de uma hierarquia de conjuntos é o Axioma do Vazio. A partir deste, os Axiomas Par, Potência, União, Separação, Substituição e Infinito permitem construir uma infinidade de outros conjuntos. O Axioma da Escolha desempenha papel de destaque neste processo, por conta de seu caráter não construtivo.
- Uma vez definida tal hierarquia de conjuntos, é possível dar os primeiros passos para uma fundamentação de números naturais, incluindo as operações de adição e multiplicação entre naturais. Mas ainda falta examinar as propriedades algébricas de tais operações, algo que é feito na próxima parte.
- A linguagem de ZF permite qualificar relações. Relações de equivalência particionam conjuntos, e partições induzem relações de equivalência.
- A meta é fundamentar vastas porções da matemática. Isso justifica a formulação de ZF, do ponto de vista social.

Seção 28

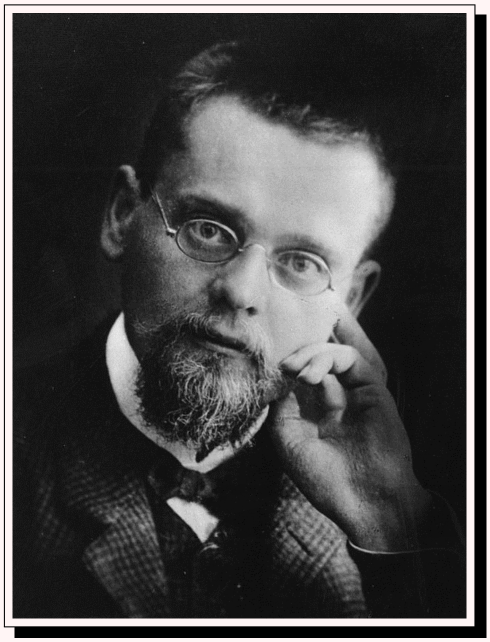
Notas históricas

Sumário
Índice
Rede

Para encerrar essa breve introdução a ZFC, vale mencionar que os axiomas dessa teoria são devidos a Ernst Zermelo, com exceção de *Substituição* e *Regularidade*. O Esquema de Substituição foi uma contribuição de Abraham Fraenkel, enquanto o Axioma de Regularidade foi proposto por Thoralf Skolem e John von Neumann. Fraenkel colaborou também com as primeiras discussões sobre a independência do Axioma da Escolha.

Uma fórmula $\mathcal{F}$ é *independente* dos axiomas de ZF se, e somente se, $\nvdash_{ZF} \mathcal{F}$. Fraenkel estudou uma variação de ZF, conhecida como

ZFU (alguns autores se referem a essa teoria como ZFA), na qual ele esboçou as primeiras ideias para provar que $\nvdash_{ZFU} AE$, onde AE denota o Axioma da Escolha em ZFU [16].



Ernst Zermelo, no início do século 20

Fonte: Wikipedia.

PARTE 4

Números naturais, inteiros e racionais

Não há, em matemática, o conceito de número. Mas há números naturais, números inteiros e números racionais, os quais são qualificados nesta quarta parte.

Seção 29

Aritmética

Sumário
Índice
Rede

Certas relações são de especial interesse, além daquelas já discutidas. São as *funções*.

> Definição 4.1. *Seja*
>
> $$r : x \to y$$
>
> *uma relação.*
>
> *Neste caso são explicitados domínio x e co-domínio y de r. Dizemos que r é uma* função *de x em y sss para todo a pertencente a x existe apenas um b pertencente a y tal que $(a, b) \in r$.*

Funções são casos especiais de relações e, portanto, conjuntos.

Exemplo 4.1. *Se* $r = \{(1,2),(2,3),(3,4),(4,5),(5,6),(6,7)\}$, r *é uma relação? É uma função?*

Se existem x *e* y *tais que* $r \subseteq x \times y$, *então* r *é uma relação* $r : x \to y$. *Logo,* r *é uma relação com domínio e co-domínio* $\{1,3\}$*? Não! Com efeito,* $r \not\subseteq \{1,3\} \times \{1,3\}$.

O mesmo conjunto r *é uma relação com domínio e co-domínio dados pelo conjunto* $\{1,2,3,4,5,6,7\}$*? Sim! Com efeito,*

$$r \subseteq \{1,2,3,4,5,6,7\} \times \{1,2,3,4,5,6,7\}.$$

É uma função com domínio e co-domínio dados por

$$\{1,2,3,4,5,6,7\}?$$

Não! Com efeito, apesar de r *ser subconjunto de*

$$\{1,2,3,4,5,6,7\} \times \{1,2,3,4,5,6,7\},$$

o termo 7 *pertence ao domínio de* r *mas não existe termo* b *pertencente ao co-domínio de* r *tal que* $(7,b)$ *pertença a* r.

O mesmo conjunto r *é uma função com domínio* $\{1,2,3,4,5,6\}$ *e co-domínio* $\{1,2,3,4,5,6,7,8,9\}$*? Sim! Com efeito,*

$$r \subseteq \{1,2,3,4,5,6\} \times \{1,2,3,4,5,6,7,8,9\}$$

e, além disso, cada elemento a *do domínio* $\{1,2,3,4,5,6\}$ *corresponde a um e apenas um* b *pertencente a* $\{1,2,3,4,5,6,7,8,9\}$ *tal que* $(a,b) \in r$.

Se $r : x \to y$ é uma função e $(a,b) \in r$, dizemos que 'b é a *imagem* de a pela função r'. Neste caso é usual a notação $r(a) = b$.

Como já vimos, se

$$r : x \to y$$

é uma relação e $(a,b) \in r$, podemos escrever arb.

Mas, apesar de toda função ser um caso particular de relação, no caso em que r é uma função, no lugar de arb escreve-se $b = r(a)$.

Um conceito muito usual é o que se segue.

Se $r : x \to y$ é uma função, dizemos que $s : z \to y$ é uma *restrição* de r sss $z \subseteq x$ e, para todo a pertencente a z, temos $s(a) = r(a)$.

Em outras palavras, uma função s é restrição de uma função r sss $s \subseteq r$. Em particular, toda função é restrição de si mesma. Revisitamos este conceito de maneira mais circunstanciada na Definição 4.12.

A operação adição $+$ entre números naturais, introduzida na Seção 23, pode ser definida como uma função

$$+ : \omega \times \omega \to \omega$$

tal que $+(m, n) = m + n$ (observar que $+(m, n)$ é uma notação abreviada para $+((m, n))$, ou seja, a imagem de (m, n) pela função $+$). Analogamente, a operação multiplicação $\cdot$ entre números naturais introduzida na mesma Seção pode ser definida como uma função

$$\cdot : \omega \times \omega \to \omega$$

tal que $\cdot(m, n) = mn$.

Vale observar também que, por exemplo, a função

$$\oplus : \omega \times \{2\} \to \omega,$$

tal que $\oplus(m, 2) = m + 2$, é uma restrição de $+$. Intuitivamente falando, essa nova função adiciona a cada natural m o número natural 2.

Lembrando que, por definição, $m+0 = m$, e $m+S(n) = S(m+n)$, provamos a seguir alguns teoremas importantes.

TEOREMA 4.1. $0 + 0 = 0$.

DEMONSTRAÇÃO: De acordo com a definição de adição, $m + 0 = m$. Se $m = 0$, então $0 + 0 = 0$.

TEOREMA 4.2. *Se m é um número natural, então $0+m = m$.*

DEMONSTRAÇÃO: Usamos aqui uma técnica de demonstração conhecida como *indução infinita*, a qual permite empregar Modus Ponens para obter uma infinidade de teoremas. Para isso é necessário dividir a demonstração em duas etapas. Na primeira devemos provar que

$$0 + S(0) = S(0),$$

lembrando que $S(0) = 1$ (observar que já foi provado acima que $0 + 0 = 0$). Na segunda etapa devemos demonstrar que a fórmula

$$0 + S(n) = S(n)$$

implica na fórmula

$$0 + S(S(n)) = S(S(n)).$$

Dessa maneira cria-se um 'efeito dominó' no seguinte sentido: Se $0 + 1 = 1$ é teorema (de acordo com a primeira etapa) e a fórmula $0 + 1 = 1$ implica na fórmula $0 + 2 = 2$ (de acordo com a segunda etapa), então $0+2 = 2$ é teorema. Mas, se a fórmula $0 + 2 = 2$ implica na fórmula $0 + 3 = 3$ (novamente usando a segunda etapa), então $0 + 3 = 3$ é teorema, e assim por diante.

Ou seja, Modus Ponens é aplicada ao longo de todos os números naturais, produzindo uma infinidade de teoremas (um para cada natural m).

Agora podemos finalmente iniciar a prova.

Etapa 1: $0 + S(0) = S(0 + 0)$, de acordo com a definição de adição. Mas $0+0 = 0$, de acordo com o Teorema 4.1. Logo, $S(0 + 0) = S(0)$, de acordo com a substitutividade da igualdade. Logo, a transitividade da igualdade garante que $0 + S(0) = S(0)$.

Etapa 2: Supor que $0 + S(n) = S(n)$. De acordo com a definição de adição, $0 + S(S(n)) = S(0 + S(n))$. Mas, como assumimos por hipótese que $0 + S(n) = S(n)$, então $0 + S(S(n)) = S(S(n))$. Logo, para qualquer natural m temos $0 + m = m$.

Em particular, $0 + 5 = 5$, como foi anunciado na Seção 3.

Levando em conta que na célebre obra *Principia Mathematica* (de Bertrand Russell e Alfred North Whitehead) foram consumidas mais de 360 páginas para provar que $1 + 1 = 2$, parece que estamos com uma certa vantagem aqui. A demonstração de Russell e Whitehead não é feita no contexto de ZF. A teoria formal explorada neste grande clássico da literatura é a *teoria de tipos*, a qual emprega uma linguagem e uma lógica diferentes daquelas de ZF.

Teorema 4.3. *Se m e n são números naturais, então*

$$m + n = n + m.$$

✍ A prova fica como exercício para o leitor, a qual também pode ser feita por indução infinita. Este último teorema garante a comutatividade da adição entre números naturais.

Teorema 4.4. *Se m, n e p são números naturais, então*

$$(m+n)+p = m+(n+p).$$

✍ A prova fica como exercício para o leitor, a qual também pode ser feita por indução infinita. Este último teorema garante a associatividade da adição entre números naturais.

Observar que a associatividade da adição entre naturais é uma propriedade facilitadora para fazer contas envolvendo adição. Com efeito, levando em conta que $+$ é uma operação binária (é aplicável sobre *duas* ocorrências de termos), como calcular $m+n+p$ ou $m+n+p+q$, entre outras possibilidades? De acordo com o Teorema 4.4, não importa se calculamos

$$m+(n+(p+q))$$

ou

$$(m+n)+(p+q),$$

sempre é obtida exatamente a mesma soma. Ou seja, a associatividade da adição entre naturais dispensa o emprego de parênteses para operar com três ou mais números naturais, ainda que $+$ seja uma operação binária.

O próximo teorema se refere à multiplicação entre naturais.

Teorema 4.5. *Se m, n e p são números naturais, então*

$$0m = 0,$$

$$1m = m,$$

$$mn = nm,$$

$$(mn)p = m(np)$$

e

$$m(n+p) = mn + mp.$$

✍ A demonstração fica como exercício para o leitor.

Por conta do último teorema, uma convenção comum é a seguinte:

$$mn + p = (mn) + p.$$

Ou seja, diante de uma notação abusiva caracterizada por falta de ocorrências de pares de parênteses, deve-se priorizar a multiplicação sobre a adição.

Outra convenção usual é a seguinte, para qualquer natural $m \neq 0$:

I: $m^0 = 1$;

II: $m^{n+1} = m \cdot m^n$, onde n é um natural.

Exemplo 4.2. $5^4 = 5 \cdot 5^3$, *de acordo com item* II*; logo,* $5^4 = 5 \cdot 5 \cdot 5^2$, *de acordo com o mesmo item; logo,* $5^4 = 5 \cdot 5 \cdot 5 \cdot 5^1$, *novamente de acordo com item* II*; finalmente,* $5^4 = 5 \cdot 5 \cdot 5 \cdot 5 \cdot 5^0$, *que é igual a 625, uma vez que item* I *garante que* $5^0 = 1$*. Nesta demonstração tiramos proveito do fato da multiplicação entre naturais ser associativa.*

Moral da História: A adição $+$ entre números naturais é comutativa, associativa e admite elemento neutro (0). A multiplicação $\cdot$ entre números naturais é comutativa, associativa e admite elemento neutro (1). Também temos como teorema a *distributividade* da multiplicação em relação à adição, ou seja, $m(n + p) = mn + mp$. Tais propriedades algébricas de adição e multiplicação entre números naturais permitem qualificar o que é aritmética.

Aritmética é o estudo da tripla ordenada $(\omega, +, \cdot)$.

A tripla ordenada $(\omega, +, \cdot)$ permite definir números primos e compostos, bem como todos os resultados conhecidos na literatura sobre o tema. Por exemplo, diversos sistemas de criptografia são definidos a partir de $(\omega, +, \cdot)$, eventualmente exigindo outras ferramentas. Um conjunto como $(\omega, +, \cdot)$ é capaz de garantir as bases para a segurança em transações bancárias realizadas no mundo todo [32].

Definição 4.2. *Um número natural* n *é* primo *sss*

$$n \neq 0 \wedge n \neq 1 \wedge \forall p \forall q ((p \in \omega \wedge q \in \omega \wedge p \neq n \wedge q \neq n) \Rightarrow n \neq pq).$$

Se $n \neq 0 \wedge n \neq 1$*, dizemos que* n *é* composto *sss* n *não é primo.*

Exemplo 4.3. I: 5 *é primo; afinal, não há fatoração* $pq = 5$*, onde ambos* p *e* q *são naturais diferentes de* 5*;*

II: 6 *é composto, uma vez que* $6 = 2(3)$*.*

A título de curiosidade, a soma dos quadrados dos sete primeiros primos é 666. Com efeito,

$$2^2 + 3^2 + 5^2 + 7^2 + 11^2 + 13^2 + 17^2 = 666.$$

Além disso, a soma dos primeiros 36 naturais diferentes de 0 é 666. Isso se escreve usualmente como

$$\sum_{k=1}^{36} k = 1 + 2 + 3 + \cdots + 35 + 36 = 666.$$

Um dos resultados mais conhecidos e úteis da aritmética é o *Teorema Binomial para Naturais*, como se segue.

Teorema 4.6. *Sejam a, b e n naturais, onde $n \neq 0$. Logo,*

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k},$$

onde

$$\sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$$

é a adição dos termos $\binom{n}{k} a^k b^{n-k}$, com k variando de 0 a n, e

$$\binom{n}{k} = \frac{n!}{k!(n-k)!},$$

onde $n-k$ é um natural p tal que $n = p + k$ (observar que k é inevitavelmente menor ou igual a n).

Demonstração: Demonstramos esse importante resultado por indução infinita, de maneira análoga à prova do Teorema 4.2.

Primeira Etapa: Provar que $(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$ é teorema para $n = 1$. Ou seja, devemos provar que

$$(a+b)^1 = \sum_{k=0}^{1} \binom{1}{k} a^k b^{1-k}.$$

Por um lado, $(a+b)^1 = a + b$. Por outro,

$$\sum_{k=0}^{1} \binom{1}{k} a^k b^{1-k} = \binom{1}{0} a^0 b^{1-0} + \binom{1}{1} a^1 b^{1-1}.$$

Mas o último termo é igual a

$$\frac{1!}{0!(1-0)!}a^0b^{1-0} + \frac{1!}{1!(1-1)!}a^1b^{1-1},$$

o qual é idêntico a $b+a$. Como adição entre naturais é comutativa, isso encerra a Primeira Etapa.

Segunda Etapa: Devemos provar que

$$(a+b)^n = \sum_{k=0}^{n}\binom{n}{k}a^kb^{n-k}$$

implica em

$$(a+b)^{n+1} = \sum_{k=0}^{n+1}\binom{n+1}{k}a^kb^{n+1-k}.$$

Observar que cada parcela do somatório que antecede a condicional acima, envolvendo o fator a^jb^l, é tal que $j+l=n$. Mas $(a+b)^{n+1} = (a+b)(a+b)^n$. Logo, $(a+b)^{n+1} = a(a+b)^n + b(a+b)^n$. Portanto,

$$(a+b)^{n+1} = a\sum_{k=0}^{n}\binom{n}{k}a^kb^{n-k} + b\sum_{k=0}^{n}\binom{n}{k}a^kb^{n-k}.$$

Ou seja, agora cada parcela da adição dos somatórios do lado direito da igualdade acima, envolvendo fatores a^jb^l, é tal que $j+l=n+1$. Logo,

$$(a+b)^{n+1} = \sum_{k=0}^{n+1}\binom{n+1}{k}a^kb^{n+1-k}.$$

✍ Se o leitor não se convenceu da última parte da demonstração acima, observar que

$$\binom{n}{k} + \binom{n}{k-1} = \binom{n+1}{k},$$

cuja demonstração pode ser um interessante exercício.

Se o leitor não se convenceu com a definição de *somatório*

$$\sum_{k=0}^{n} z_k,$$

introduzida no último teorema, essa pode ser escrita como se segue:

I:
$$\sum_{k=0}^{1} z_k = z_0 + z_1;$$

II:
$$\sum_{k=0}^{n+1} z_k = \sum_{k=0}^{n} z_k + z_{n+1}.$$

Apesar do Teorema Binomial para Naturais ser um resultado da aritmética, ele pode ser estendido de modo a repercutir em áreas como cálculo diferencial e integral, conforme vemos na Seção 49. Esse é um dos aspectos mais marcantes da matemática: o surpreendente alcance dos resultados mais relevantes.

Seção 30

Inteiros

Sumário
Índice
Rede

Nesta Seção iniciamos as primeiras aplicações de relações de equivalência.

Eventualmente relações podem ser definidas sobre relações, como se faz a seguir. Afinal, toda relação é um conjunto.

Definição 4.3. *Sejam* (m, n) *e* (p, q) *elementos da relação*
$$\omega \times \omega$$
em ω. *Logo,*

$(m, n) \sim (p, q) \;\vdots\; m + q = n + p.$

$(m, n) \not\sim (p, q) \;\vdots\; \neg((m, n) \sim (p, q)).$

Exemplo 4.4. i: $(5, 2) \sim (7, 4)$*; isso porque* $5 + 4 = 2 + 7$*;*

ii: $(7, 4) \sim (32, 29)$*; com efeito,* $7 + 29 = 4 + 32$*;*

iii: $(5, 2) \sim (32, 29)$*;*

iv: $(5, 2) \not\sim (2, 5)$*; com efeito,* $5 + 5 \neq 2 + 2$.

Notar que $\omega \times \omega$ é uma relação em ω, e $\sim$ é uma relação em $\omega \times \omega$. Neste momento é importante não confundir uma relação em $\omega \times \omega$

com qualquer subconjunto de ω^4. Com efeito,

$$\omega^4 = \omega \times (\omega \times (\omega \times \omega)),$$

de acordo com Definição 3.9.

Porém, $\sim$ é subconjunto próprio de

$$(\omega \times \omega) \times (\omega \times \omega).$$

Logo, $\sim$ não é subconjunto de ω^4.

Teorema 4.7. *A relação $\sim$ em $\omega \times \omega$ da Definição 4.3 é de equivalência.*

✍ A demonstração deste último resultado fica a cargo do leitor. Resumidamente, tanto reflexividade quanto simetria de $\sim$ são consequências da comutatividade da adição $+$ entre naturais. Com relação à transitividade de $\sim$, essa pode ser facilmente provada se o leitor enunciar e demonstrar um teorema de *cancelamento* de termos para a adição de naturais. Tal teorema de cancelamento diz o seguinte: dados m, n e p naturais, então

$$m + n = m + p \Leftrightarrow n = p.$$

Uma vez que toda relação de equivalência define uma partição (Teorema 3.10), há aqui a oportunidade para introduzir números inteiros. As classes de equivalência de $\omega \times \omega$ relativamente a $\sim$ são denotadas como se segue.

Definição 4.4.

$$\underline{+n} = [(n, 0)] = \{(a, b) \in \omega \times \omega \mid (a, b) \sim (n, 0)\}$$

$$\underline{-n} = [(0, n)] = \{(a, b) \in \omega \times \omega \mid (a, b) \sim (0, n) \wedge n \neq 0\}$$

A classe de equivalência $\underline{+n}$ se lê '*inteiro positivo n*'. A classe de equivalência $\underline{-n}$ se lê '*inteiro negativo n*'.

Exemplo 4.5. i: *O inteiro positivo zero é $\underline{0} = [(0, 0)] = \{(0, 0), (1, 1), (2, 2), (3, 3), \cdots\}$;*

ii: *o inteiro positivo um é $\underline{+1} = [(1, 0)] = \{(1, 0), (2, 1), (3, 2), (4, 3), \cdots\}$;*

iii: *o inteiro negativo um é $\underline{-1} = [(0, 1)] = \{(0, 1), (1, 2), (2, 3), (3, 4), \cdots\}$;*

iv: *o inteiro positivo dois é* $\underline{+2} = [(2,0)] = \{(2,0),(3,1),(4,2),(5,3),\cdots\}$*;*

v: *o inteiro negativo dois é* $\underline{-2} = [(0,2)] = \{(0,2),(1,3),(2,4),(3,5),\cdots\}$.

Um número inteiro é uma classe de equivalência de pares ordenados de números naturais relativamente a $\sim$. Um inteiro positivo (ver sinal $+$) tem como representante um par ordenado (m,n) onde $m \geq n$ (isso equivale a afirmar que $n \leq m$). Um inteiro negativo (ver sinal $-$) tem como representante um par ordenado (m,n) onde $m < n$. Um inteiro estritamente positivo é um inteiro positivo diferente de $\underline{0}$. Eventualmente podemos omitir o sinal $+$ entre inteiros positivos.

O emprego das notações $\underline{+n}$ e $\underline{-n}$ serve ao propósito de enfatizar que nenhum inteiro é natural e nenhum natural é inteiro. Por exemplo, o natural 0 é o conjunto vazio, enquanto o inteiro positivo zero é o conjunto $\underline{0} = [(0,0)] = \{(0,0),(1,1),(2,2),(3,3),\cdots\}$. Logo, de acordo com o Axioma da Extensionalidade, $0 \neq \underline{0}$.

Para definirmos operações de adição e multiplicação entre inteiros, basta, portanto, definirmos operações sobre representantes quaisquer de inteiros. Essa é a enorme vantagem do emprego de classes de equivalência! Para operar entre inteiros não há necessidade alguma de definir operações entre classes de equivalência. Definir operações entre representantes de classes de equivalência induz operações entre as próprias classes de equivalência.

Definição 4.5. *Se* (m,n) *e* (p,q) *são representantes quaisquer de inteiros, então*

$$(m,n) + (p,q) = (m+p, n+q)$$

e

$$(m,n) \cdot (p,q) = (mp+nq, mq+np).$$

Se

$$(m,n) + (p,q) = (r,s),$$

dizemos que (r,s) é a *soma* das *parcelas* (m,n) e (p,q). Se

$$(m,n) \cdot (p,q) = (r,s),$$

dizemos que (r,s) é o *produto* dos *fatores* (m,n) e (p,q). O mesmo se diz sobre os respectivos inteiros com representantes (m,n), (p,q) e (r,s).

Em outras palavras, em virtude do que foi dito acima, se $\underline{x}$ e $\underline{y}$ são inteiros, positivos ou negativos, então $\underline{x} + \underline{y} = \underline{z}$ sss (m, n) e (p, q) forem representantes de $\underline{x}$ e $\underline{y}$, respectivamente, e $(m, n) + (p, q)$ for representante de $\underline{z}$. Situação análoga ocorre com a multiplicação entre inteiros.

Exemplo 4.6. *Como calcular* $\underline{4} + \underline{-2}$*?*

Basta escolhermos representantes quaisquer dos inteiros $\underline{4}$ *e* $\underline{-2}$ *e aplicarmos a Definição 4.5.*

Por exemplo, um dos representantes de $\underline{+4}$ *é* $(5, 1)$*, e um dos representantes de* $\underline{-2}$ *é* $(16, 18)$*. Logo,*

$$(5, 1) + (16, 18) = (5 + 16, 1 + 18) = (21, 19).$$

Mas $(21, 19)$ *é representante de* $\underline{+2}$*. Com efeito,* $(21, 19) \sim (2, 0)$ *(ver Definição 4.4), uma vez que* $21 + 0 = 19 + 2$.

Logo, $\underline{4} + \underline{-2} = \underline{2}$.

Exemplo 4.7. *Como calcular* $\underline{4} \cdot \underline{-2}$*? Basta usar a mesma estratégia do* Exemplo *anterior. Ou seja,*

$$(4, 0) \cdot (1, 3) = (4.1 + 0.3, 4.3 + 0.1) = (4 + 0, 12 + 0) = (4, 12).$$

Mas $(4, 12)$ *é representante de* $\underline{-8}$*, uma vez que* $(4, 12) \sim (0, 8)$ *e* $(0, 8)$ *é representante de* $\underline{-8}$*, de acordo com a Definição 4.4.*

Importante observar que a operação de adição entre naturais é uma função

$$+ : \omega \times \omega \to \omega,$$

enquanto a adição entre inteiros é uma função

$$+ : ((\omega \times \omega)/\sim) \times (\omega \times \omega)/\sim) \to ((\omega \times \omega)/\sim)$$

induzida pela Definição 4.5. Logo, são funções diferentes. Mais do que isso, nenhuma é restrição da outra.

Do ponto de vista formal isso significa que tais funções deveriam ser denotadas por símbolos diferentes. Mas, como já foi dito anteriormente, matemáticos estão mais interessados em rigor do que formalismo. Do ponto de vista do rigor, naturalmente se sabe que adição entre naturais é uma função e adição entre inteiros é outra. Comentário análogo vale para a multiplicação entre naturais e a multiplicação entre inteiros.

Teorema 4.8. *$\underline{0}$ é neutro aditivo.*

Demonstração: Seja (m, n) um representante de um inteiro qualquer. Uma vez que todo representante de zero inteiro é um par ordenado (p, p), onde p é natural, então basta aplicar a Definição 4.5. Logo,

$$(m, n) + (p, p) = (m + p, n + p).$$

Mas $(m, n) \sim (m + p, n + p)$, uma vez que

$$m + n + p = n + m + p,$$

graças à comutatividade e à associatividade da adição entre naturais. Logo,

$$(m + p, n + p)$$

e

$$(m, n)$$

são representantes do mesmo inteiro.

Demonstração análoga para o caso da adição entre zero inteiro e um inteiro qualquer. Logo, $\underline{0}$ é neutro aditivo.

Em outras palavras, no teorema acima foi provado que, se $\underline{+n}$ ou $\underline{-n}$ são inteiros, então

$$\underline{+n} + \underline{0} = \underline{+n},$$
$$\underline{-n} + \underline{0} = \underline{-n},$$
$$\underline{0} + \underline{+n} = \underline{+n}$$

e

$$\underline{0} + \underline{-n} = \underline{-n}.$$

Para evitar notação sobrecarregada, eventualmente podemos nos referir a inteiros simplesmente por letras latinas minúsculas em itálico, i.e., sem a barra embaixo. Ou seja, foi provado acima que, para qualquer inteiro p,

$$p + \underline{0} = \underline{0} + p = p$$

(observar que a última notação é uma abreviação metalinguística equivalente a afirmar que $p + \underline{0} = p \wedge \underline{0} + p = p$).

Teorema 4.9. *Todo inteiro admite simétrico aditivo. Ou seja, se p é um inteiro, então existe inteiro q tal que, tanto $p+q$ quanto $q + p$ resulta no neutro aditivo $\underline{0}$.*

Demonstração: Seja (m, n) um representante qualquer de um inteiro p. Se (n, m) é representante de um inteiro q, então

$$(m, n) + (n, m) = (m + n, n + m).$$

Uma vez que adição entre naturais é comutativa,

$$(m + n, n + m) = (m + n, m + n).$$

Logo, este último par ordenado é representante de $\underline{0}$, o qual é neutro aditivo. Portanto, todo inteiro p (com representante (m, n)) admite simétrico aditivo q (com representante (n, m)).

✍ Observar que $\underline{0}$ é o único inteiro cujo simétrico aditivo é ele mesmo. Consegue provar isso?

É uma prática comum denotar o simétrico aditivo de um inteiro p por $-p$. Neste texto a mesma notação é empregada. Mas é preciso cuidado: não confundir o sinal $-$, usado na definição de inteiros, com simétrico aditivo $-p$ de p. Isso porque, eventualmente, $-p$ pode ser um inteiro estritamente positivo.

O último teorema é de importância vital para compreender a diferença entre naturais e inteiros. Todo inteiro admite simétrico aditivo. No entanto, 0 é o único natural que admite simétrico aditivo relativamente à adição entre naturais. Por exemplo, não existe natural n tal que $n + 2$ ou $2 + n$ seja igual a 0 (o neutro aditivo entre os naturais).

Para uma definição precisa do conceito de *simétrico* relativamente a uma operação binária qualquer (não apenas adição ou multiplicação), ver Seção 68.

Graças à existência de simétrico aditivo entre inteiros, é possível definir uma nova operação a partir da adição entre inteiros. A *subtração*

$$p - q$$

entre inteiros é a adição

$$p + (-q),$$

ou seja, a adição do inteiro p com o simétrico aditivo de q. Obviamente não é possível definir conceito equivalente entre naturais.

Teorema 4.10. *Existe neutro multiplicativo entre os inteiros.*

Demonstração: Seja (m, n) um representante de um inteiro qualquer. Logo,

$$(m, n) \cdot (1, 0) = (m.1 + n.0, m.0 + n.1) =$$
$$(m + 0, 0 + n) = (m, n).$$

Demonstração análoga para o caso de $(1, 0) \cdot (m, n)$. Logo, $\underline{1}$ é neutro multiplicativo, uma vez que $(1, 0)$ é representante de $\underline{1}$.

✍ Recomendamos ao leitor provar este último teorema usando outro representante para o inteiro $\underline{1}$.

Teorema 4.11. *O neutro aditivo entre os inteiros é* absorvente multiplicativo*. Ou seja, se p é um inteiro, então $p \cdot \underline{0} = \underline{0} \cdot p = \underline{0}$.*

Demonstração: Seja (m, n) um representante de um inteiro qualquer. Um representante do neutro aditivo entre os inteiros é $(0, 0)$. Logo,

$$(m, n) \cdot (0, 0) = (m(0) + n(0), m(0) + n(0)).$$

Mas este último é o par ordenado $(0, 0)$, uma vez que a multiplicação entre naturais garante trivialmente que o natural 0 é absorvente multiplicativo. Demonstração análoga para $\underline{0} \cdot p = \underline{0}$. Logo, $\underline{0}$ é absorvente multiplicativo.

Obviamente, a demonstração acima poderia ser feita a partir de qualquer outro representante de $\underline{0}$. Optamos pelo par ordenado $(0, 0)$ para destacar que o próprio natural 0 é absorvente multiplicativo entre os naturais.

Outros teoremas podem ser demonstrados:

i: a adição entre inteiros é comutativa e associativa;

ii: a multiplicação entre inteiros é comutativa e associativa;

iii: se p, q e r são inteiros, então $p(q + r) = pq + pr$.

Moral da História: Todas as propriedades algébricas da adição e da multiplicação entre naturais ocorrem também para a adição e multiplicação entre inteiros. No entanto, os inteiros contam com uma propriedade algébrica não replicada entre os naturais, a saber, a existência de simétricos aditivos. Esta é a relevante diferença entre naturais e inteiros!

Teorema 4.12. *A multiplicação entre um inteiro estritamente positivo e um inteiro negativo é um inteiro negativo.*

Demonstração: Um representante de um inteiro estritamente positivo $\underline{+m}$ qualquer é o par ordenado $(m, 0)$, onde $m \neq 0$.

Um representante de um inteiro negativo qualquer $\underline{-q}$ é o par ordenado $(0, q)$, onde $q \neq 0$. Logo, a multiplicação entre eles é simplesmente

$$(m, 0) \cdot (0, q) = (m(0) + 0(q), mq + 0(0)) = (0, mq).$$

Mas este último é representante de um inteiro negativo.

As demais *regras de sinais* (tão propagadas no ensino médio, mas sem justificativa alguma!) podem ser demonstradas de maneira análoga:

i negativo multiplicado por estritamente positivo é negativo,

ii negativo multiplicado por negativo é estritamente positivo,

iii positivo multiplicado por positivo é positivo.

✍ Recomendamos ao leitor que prove esses últimos três teoremas.

Apesar de nenhum natural ser inteiro, como já foi discutido acima, ainda é possível *copiar* os naturais entre os inteiros. Para tanto, basta observar os seguintes teoremas:

- A adição entre inteiros positivos é *fechada* nos inteiros positivos, ou seja, se p e q são inteiros positivos, então $p + q$ é um inteiro positivo.
- A multiplicação entre inteiros positivos é fechada nos inteiros positivos, ou seja, se p e q são inteiros positivos, então $p \cdot q$ é um inteiro positivo.
- A adição entre inteiros positivos é comutativa, associativa e admite neutro aditivo.
- A multiplicação entre inteiros positivos é comutativa, associativa e admite neutro multiplicativo.

- Entre os inteiros positivos temos como teorema a distributividade da multiplicação.
- *Não é teorema* a seguinte afirmação: 'para todo inteiro positivo existe simétrico aditivo que seja inteiro positivo'. Com efeito, basta provar que o simétrico aditivo de qualquer inteiro estritamente positivo é um inteiro negativo.

Ou seja, os inteiros positivos contam com as mesmas propriedades algébricas dos naturais, no que se refere às respectivas operações de adição e multiplicação.

Vale a pena notar que, em momento algum, foram definidas operações de adição ou multiplicação entre um natural e um inteiro, ou entre um inteiro e um natural. Não há necessidade disso justamente porque os inteiros positivos podem replicar os naturais.

Apesar de alguns autores afirmarem irresponsavelmente que todo número natural é inteiro, o que se mostra aqui é que os inteiros positivos copiam os naturais. Nada além disso. Mais detalhes na Seção 41.

Observar também que, entre os inteiros, não é teorema a seguinte afirmação: 'todo inteiro admite *simétrico multiplicativo*'. Se existisse, o simétrico multiplicativo de um inteiro p, deveria ser um inteiro q tal que $pq = \underline{1}$, sendo $\underline{1}$ o neutro multiplicativo entre os inteiros. Obviamente o neutro multiplicativo dos inteiros admite ele mesmo como simétrico multiplicativo. Analogamente, o simétrico aditivo do neutro multiplicativo (ou seja, $\underline{-1}$) também admite como simétrico multiplicativo ele mesmo, uma vez que $\underline{-1} \cdot \underline{-1} = \underline{1}$. Mas nenhum outro inteiro conta com essa propriedade algébrica.

Considere, para fins de ilustração, o inteiro $\underline{2}$. Supor que ele admite simétrico multiplicativo com representante (p, q). Logo,

$$(2, 0) \cdot (p, q) = (2p + 0q, 2q + 0p) = (2p, 2q).$$

Para que o resultado $(2p, 2q)$ seja representante do neutro multiplicativo é necessário que

$$(2p, 2q) = (n + 1, n)$$

para pelo menos algum n natural. No entanto, ambos $2p$ e $2q$, independentemente dos valores de p e q, são naturais pares. Logo, é necessário que ambos n e $n+1$ sejam pares. Mas, se n é par, então $n+1$ é ímpar. Se n é ímpar, então $n+1$ é par. Isso é uma contradição!

O fato de não haver simétrico multiplicativo para todo e qualquer inteiro serve como motivação para a definição dos *números racionais*. A proposta é a seguinte:

> *Como definir um conjunto x e duas operações ($+$ e $\cdot$), fechadas em x, de modo que este novo conjunto x consiga copiar os inteiros e os naturais e ainda admitir a existência de simétrico multiplicativo para todos os termos pertencentes a x?*

Este problema é resolvido na próxima Seção.

Entre os inteiros é possível definir relações de ordem total $\geq$ (maior ou igual) e $\leq$ (menor ou igual) como se segue:

Definição 4.6. *Sejam r e s inteiros. Logo,*

- $r \geq \underline{0} \vdots r$ *é inteiro positivo;*
- $r \geq s \vdots r + (-s) \geq \underline{0}$;
- $r \leq s \vdots s \geq r$.

Além disso,

$r < s$ sss $r \leq s \wedge r \neq s$; e

$r > s$ sss $r \geq s \wedge r \neq s$.

Exemplo 4.8. $\underline{5} > \underline{2}$. *Com efeito,* $\underline{5} + \underline{-2} = \underline{3}$*; Uma vez que* $\underline{3}$ *é um inteiro positivo, então* $\underline{5} + \underline{-2} \geq \underline{0}$*. Uma vez que* $\underline{5} \neq \underline{2}$*, então* $\underline{5} > \underline{2}$.

Para encerrar essa discussão, o conjunto dos números inteiros é denotado por $\mathbb{Z}$. Em outras palavras,

$$\mathbb{Z} = (\omega \times \omega)/\sim .$$

Estudar os números inteiros significa estudar o conjunto $(\mathbb{Z}, +, \cdot)$.

Seção 31

Racionais

Sumário
Índice
Rede

Assim como os inteiros foram definidos a partir dos naturais, os racionais são definidos a partir dos inteiros, novamente usando classes de equivalência.

Definição 4.7. *Sejam* (a, b) *e* (c, d) *elementos de*

$$\mathbb{Z} \times (\mathbb{Z} - \{\underline{0}\}).$$

Logo,

$(a, b) \approx (c, d) \,\vdots\, ad = bc.$

$(a, b) \not\approx (c, d) \,\vdots\, \neg((a, b) \approx (c, d)).$

Observar que ad é uma multiplicação entre inteiros. O mesmo vale para bc.

Por abuso de notação, de agora em diante omitimos a barra __ sob cada inteiro. Dessa maneira, para evitar confusão, sempre qualificamos se n denota um inteiro ou um natural. Nesta Seção, quaisquer pares ordenados (m, n) são tais que ambos m e n são inteiros.

Exemplo 4.9. i: $(1, 2) \approx (2, 4)$; *com efeito,* $1(4) = 2(2)$;

ii: $(2, 4) \approx (3, 6)$; *com efeito,* $2(6) = 4(3)$;

iii: $(1, 2) \approx (3, 6)$;

iv: $(1, 2) \not\approx (2, 1)$; *com efeito,* $1(1) \neq 2(2)$.

É claro que poderíamos ter escrito, por exemplo, $1 \cdot 4$ ou $2 \cdot 2$ no lugar de $1(4)$ e $2(2)$, respectivamente. Mas é uma boa ideia o leitor se habituar com diferentes possíveis notações. O que realmente está em jogo aqui são os conceitos envolvidos e como lidar com eles.

Teorema 4.13. *A relação* $\approx$ *em* $\mathbb{Z} - \{0\}$ *(ver Definição 4.7) é de equivalência.*

✍ A prova, muito simples, fica a cargo do leitor. Com relação ao termo $\mathbb{Z} - \{0\}$, ver a definição de *diferença* entre conjuntos ao final da Seção 22 (a qual nada tem a ver com diferença entre inteiros, introduzida na Seção 30). O termo $\mathbb{Z}-\{0\}$ é simplesmente o conjunto de todos os inteiros diferentes de 0.

Definição 4.8. *Um* número racional *é o termo*

$$[(0,1)] = \{(a,b) \in \mathbb{Z} \times (\mathbb{Z} - \{0\}) \mid (a,b) \approx (0,1)\}$$

ou qualquer classe de equivalência

$$[(m,n)] = \{(a,b) \in (\mathbb{Z} - \{0\}) \times (\mathbb{Z} - \{0\}) \mid (a,b) \approx (m,n)\}$$

pertencente a $(\mathbb{Z} - \{0\})/\approx$.

Oportunamente é provado adiante que o racional $[(0,1)]$ é neutro aditivo relativamente à operação de adição a ser definida abaixo. O conjunto $[(0,1)]$ obviamente *não* é uma classe de equivalência pertencente a $(\mathbb{Z} - \{0\})/\approx$. A razão para tal manobra é a seguinte: mantemos neste texto a prática usual de não definir divisão por zero, onde 'zero' é como se lê o racional $[(0,1)]$. Uma vez que a definição de divisão depende da existência de simétricos multiplicativos (assim como a definição de subtração entre inteiros depende da existência de simétricos aditivos), dessa maneira garantimos que divisão por zero não é definida.

No entanto, é perfeitamente possível definir racionais (ou até mesmo números reais e números complexos) de maneira a permitir divisão por zero. Detalhes podem ser encontrados em [53].

O racional $[(0,1)]$ é denotado por

$$\frac{0}{1}.$$

Os demais racionais $[(m,n)]$ são denotados por

$$\frac{m}{n}.$$

Se $\frac{m}{n}$ é um racional, chamamos m de *numerador* e n de *denominador*. Logo, a Definição 4.8 não permite a existência de racionais com denominador 0. Uma vez que racionais são conjuntos de pares ordenados, o que permite discriminar numerador de denominador em um racional é a definição de par ordenado de Kuratowski.

Para fins de notação, supor que $[(m,n)]$ seja um racional tal que ambos m e n são inteiros estritamente positivos ou ambos negativos. Neste caso, $\frac{m}{n}$ dispensa qualquer sinal dos inteiros m e n. Se m é um inteiro negativo e n é um inteiro estritamente positivo, basta novamente escrever $\frac{m}{n}$, mas explicitando o sinal negativo de m. Porém, se m é um inteiro estritamente positivo e n é um inteiro negativo, observar que $(-m, -n)$ também pertence a $[(m, n)]$. Com efeito, $(m, n) \approx (-m, -n)$, uma vez que $m(-n) = n(-m)$. Neste caso, fica mais conveniente representar o racional $[(m, n)]$ com o sinal negativo de $-m$ (lembrar que $-m$ é um inteiro negativo se m é um inteiro estritamente positivo). Essa convenção se mostra consistente com resultados colocados adiante.

Exemplo 4.10.

$$\frac{1}{2} = [(1,2)] = \{(a,b) \in \mathbb{Z} \times (\mathbb{Z} - \{0\}) \mid (a,b) \approx (1,2)\}.$$

$$\frac{-1}{2} = [(1,-2)] = \{(a,b) \in \mathbb{Z} \times (\mathbb{Z} - \{0\}) \mid (a,b) \approx (1,-2)\}.$$

$$\frac{2}{1} = [(2,1)] = \{(a,b) \in \mathbb{Z} \times (\mathbb{Z} - \{0\}) \mid (a,b) \approx (6,3)\}.$$

$$\frac{11}{23} = [(-11,-23)] =$$
$$\{(a,b) \in \mathbb{Z} \times (\mathbb{Z} - \{0\}) \mid (a,b) \approx (-22,-46)\}.$$

Apesar do racional zero $\frac{0}{1}$ não ser uma classe de equivalência pertencente a $(\mathbb{Z} - \{0\})/\approx$, fica fácil perceber que a interseção entre $\frac{0}{1}$ e qualquer classe de equivalência de $(\mathbb{Z} - \{0\})/\approx$ é o conjunto vazio. Logo, não há risco de confusão (no sentido de confundir o racional $\frac{0}{1}$ com os demais). Os elementos de $\frac{0}{1}$ também são chamados de *representantes* de $\frac{0}{1}$. Isso por conta do fato de que

$$\frac{0}{1} = [(0,1)] = [(0,-1)] = [(0,2)] = [(0,-2)] = \cdots = [(0,n)],$$

para qualquer n inteiro diferente de 0.

Uma vez que racionais são definidos a partir de classes de equivalência, basta usar representantes para definir operações de adição $+$ e multiplicação $\cdot$, de maneira análoga àquilo que foi feito na Seção anterior.

Definição 4.9. *Sejam m, n, p e q inteiros. Logo,*

$$(m, n) + (p, q) = (mq + np, nq)$$

e

$$(m, n) \cdot (p, q) = (mp, nq).$$

Se $(m, n)+(p, q) = (r, s)$, dizemos que (r, s) é a *soma* das *parcelas* (m, n) e (p, q). Se $(m, n) \cdot (p, q) = (r, s)$, dizemos que (r, s) é o *produto* dos *fatores* (m, n) e (p, q). O mesmo se diz sobre os respectivos racionais com representantes (m, n), (p, q) e (r, s).

Exemplo 4.11. i: *Como calcular*

$$\frac{1}{3} + \frac{-2}{5}?$$

Basta escolher representantes quaisquer de cada racional envolvido e usar a Definição 4.9. O par ordenado $(3, 9)$ *é um dos representantes de* $\frac{1}{3}$*. O par ordenado* $(2, -5)$ *é um dos representantes do racional* $\frac{-2}{5}$*. Logo,*

$$(3, 9) + (2, -5) = (3.(-5) + 9(2), 9(-5)) =$$
$$(-15 + 18, -45) = (3, -45).$$

Mas $(3, -45)$ *é representante do racional* $\frac{-1}{15}$*. Com efeito,* $3(15) = -45(-1)$*. Logo,*

$$\frac{1}{3} + \frac{-2}{5} = \frac{-1}{15}.$$

ii: *Como calcular*

$$\frac{1}{3} \cdot \frac{-2}{5}?$$

Basta usar a mesma estratégia do item acima:

$$(3, 9).(2, -5) = (3.2, 9.(-5)) = (6, -45).$$

Mas $(6, -45)$ *é representante de*

$$\frac{-2}{15}.$$

✍ De agora em diante, por abuso de notação, todo racional $\frac{n}{1}$ é denotado simplesmente por n. Essa notação é conveniente, uma vez que racionais $\frac{n}{1}$ copiam os inteiros (consegue provar isso?). Se n for um inteiro positivo, esses mesmos racionais copiam os naturais (consegue provar isso?).

Teorema 4.14. *Adição + entre racionais é comutativa e associativa.*

Demonstração: Sejam (m, n) e (p, q) representantes quaisquer de racionais. Logo,

$$(m, n) + (p, q) = (mq + np, nq).$$

Mas

$$(mq + np, nq) = (pn + qm, qn),$$

uma vez que adição e multiplicação entre inteiros são comutativas. No entanto,

$$(pn + qm, qn) = (p, q) + (m, n),$$

de acordo com a Definição 4.9. Logo, a transitividade da igualdade garante que

$$(m, n) + (p, q) = (p, q) + (m, n).$$

Isso prova a comutatividade da adição entre racionais. ✍

A demonstração da associatividade fica como exercício para o leitor.

Teorema 4.15. *Multiplicação · entre racionais é comutativa e associativa.*

✍ A prova fica como exercício para o leitor.

Teorema 4.16. *O racional 0 é neutro aditivo e absorvente multiplicativo. O racional 1 é neutro multiplicativo.*

Demonstração: Seja (m, n) um representante de um racional qualquer. Então,

I:

$$(m, n) + (0, 1) = (m(1) + n(0), n(1)).$$

Mas

$$(m(1) + n(0), n(1)) = (m, n).$$

Isso prova que o racional 0 (com representante escolhido $(0, 1)$) é neutro aditivo.

II: Além disso,

$$(m, n) \cdot (0, 1) = (m(0), n(1)).$$

Mas

$$(m(0), n(1)) = (0, n).$$

Uma vez que $(0, n)$ é representante do racional 0, isso prova que o mesmo é absorvente multiplicativo.

III: Outrossim,

$$(m, n) \cdot (1, 1) = (m(1), n(1)).$$

Mas

$$(m(1), n(1)) = (m, n).$$

Isso prova que o racional 1 (com representante escolhido $(1, 1)$) é neutro multiplicativo.

✍ As operações $+$ e $\cdot$ entre racionais contam com as mesmas propriedades algébricas de adição e multiplicação entre inteiros (incluindo distributividade, a qual, naturalmente, precisa ser demonstrada pelo leitor).

Porém, entre os racionais há uma propriedade algébrica nova:

Teorema 4.17. *Todo racional diferente de* 0 *admite simétrico multiplicativo.*

Demonstração: Seja (a, b) um representante de um racional diferente de 0, ou seja, a é diferente do inteiro 0 (lembrar que b jamais é o inteiro 0, de acordo com a definição de número racional). Logo, $(a, b) \cdot (b, a) = (ab, ba)$. Mas (ab, ba) é representante do racional 1, uma vez que a multiplicação entre inteiros é comutativa. Logo, (b, a) é representante do simétrico multiplicativo do racional com representante (a, b).

Exemplo 4.12. I: *$\frac{1}{2}$ é simétrico multiplicativo de 2, assim como 2 é simétrico multiplicativo de $\frac{1}{2}$;*

II: *$\frac{-3}{4}$ é simétrico multiplicativo de $\frac{-8}{6}$;*

III: *o simétrico multiplicativo do simétrico multiplicativo do racional r é o próprio r, desde que r seja diferente do racional 0;*

IV: *o simétrico aditivo do neutro multiplicativo entre os racionais é -1.*

✍ Usualmente o simétrico multiplicativo de um racional r diferente de 0 é denotado por r^{-1}. Analogamente àquilo que é feito entre inteiros, o simétrico aditivo de um racional r é denotado por $-r$ (naturalmente, o leitor precisa saber provar que todo racional admite simétrico aditivo). Logo, a *subtração* entre racionais é definida de maneira análoga à subtração entre inteiros.

Por conta do último teorema é possível definir uma nova operação binária entre os racionais, a partir da multiplicação entre racionais.

Se r e s são racionais, a *divisão* entre r e s é

$$r/s = r(s^{-1}),$$

desde que $s \neq 0$.

Podemos também nos referir a r/s como a *divisão de r por s*.

Exemplo 4.13. *A divisão entre 5 e 3 é o racional $\frac{5}{3}$.*

Entre os racionais é possível definir uma relação de ordem total $\leq$.

$r < 0 \vdots$ qualquer representante (a, b) de r é tal que a e b não compartilham o mesmo sinal. Caso contrário, dizemos que $r > 0$.

$r < s \vdots r - s < 0$, sendo $r - s = r + (-s)$, onde $-s$ é o simétrico aditivo de s.

$r \leq s \vdots r < s \vee r = s$.

✍ Cabe ao leitor provar que $\leq$ é uma relação de ordem total entre os racionais.

O conjunto dos racionais é denotado por $\mathbb{Q}$.

Seção 32

Bijetividade e composição de funções

Sumário
Índice
Rede

Nesta Seção apresentamos conceitos estratégicos sobre funções. Uma vez que funções são essenciais para a prática matemática, há a necessidade de conhecê-las melhor.

Se o leitor está se perguntando por que não estamos tratando sobre reais nesta Seção, por enquanto a mensagem é a seguinte: paciência. Não há caminho fácil para os números reais. Mas vamos chegar lá!

TEOREMA 4.18. *Seja d a diagonal de um conjunto x. Logo d é uma função $d : x \to x$.*

DEMONSTRAÇÃO: Se d é a diagonal de x, então todo elemento de d é um par ordenado (a, a), onde $a \in x$. Logo, $d \subseteq x \times x$. Isso prova que d é uma relação em x. Além disso, para todo a pertencente a x, existe um único b pertencente a x tal que (a, b) pertence a d; tal b é simplesmente a. Em outras palavras, d satisfaz à definição de função $d : x \to x$, onde $d(a) = a$ para todo a pertencente a x.

O teorema acima motiva o conceito de *função identidade*.

DEFINIÇÃO 4.10. *Se x é um conjunto não vazio, então a* função identidade *em x é a diagonal de x. Ela é denotada por $\mathbb{I}_x$. Ou seja,*

$$\mathbb{I}_x : x \to x$$

é uma função tal que

$$\mathbb{I}_x(a) = a$$

para todo a pertencente a x.

DEFINIÇÃO 4.11. *Seja $f : a \to b$ uma função. Dizemos que f é* injetora *(ou* injetiva*) sss*

$$\forall x \forall y((x \in a \wedge y \in a \wedge x \neq y) \Rightarrow f(x) \neq f(y)).$$

Em outras palavras, elementos distintos do domínio de uma função injetora correspondem a imagens distintas. Equivalentemente,

$$f(x) = f(y) \Rightarrow x = y,$$

no caso em que f é injetora.

EXEMPLO 4.14. I: *Seja $f : \mathbb{Q} \to \mathbb{Q}$ uma função tal que*

$$f(x) = 3x;$$

se $x \neq y$, então $3x \neq 3y$; logo, $f(x) \neq f(y)$; logo, f é injetora;

ii: *seja* $g : \mathbb{Q} \to \mathbb{Q}$ *uma função tal que*

$$g(x) = x^2;$$

$g(-1) = g(1)$*; logo,* g *não é injetora;*

iii: *toda função identidade é injetora.*

Definição 4.12. *Seja*

$$f : x \to y$$

uma função qualquer.

Uma função

$$g : z \to y$$

é restrição de f ao domínio z *sss* $z \subseteq x$ *e* $g \subseteq f$.

Na Seção 29 foi introduzido o conceito de restrição de uma função. A única novidade aqui é a menção explícita ao domínio da restrição. É uma prática comum não mencionar o domínio da restrição se o contexto já deixa essa questão clara.

Exemplo 4.15. *Como visto no item* ii *do* Exemplo 4.14*, se* $g : \mathbb{Q} \to \mathbb{Q}$ *é uma função tal que*

$$g(x) = x^2,$$

então g *não é injetora.*

No entanto, g *admite uma infinidade de restrições injetoras. Uma delas, por exemplo, é a função*

$$h : \{x \in \mathbb{Q} \mid x \geq 0\} \to \mathbb{Q}$$

dada por $h(x) = x^2$*. Neste caso,* $h \subset g$ *e* h *é injetora.*

✍ *Recomendamos que o leitor crie outros exemplos de restrições injetoras de* g*. Exemplos interessantes podem envolver domínios que incluam tanto racionais positivos quanto racionais negativos.*

Definição 4.13. *Seja*

$$f : a \to b$$

uma função. Dizemos que f *é* sobrejetora *(ou* sobrejetiva*) sss*

$$\forall z(z \in b \Rightarrow \exists x(x \in a \wedge f(x) = z)).$$

Em outras palavras, f é uma função sobrejetora sss todo elemento do co-domínio de f é imagem de um termo do domínio de f.

Exemplo 4.16. i: *Seja $f : \mathbb{Q} \to \mathbb{Q}$ uma função tal que*

$$f(x) = 3x;$$

para todo z pertencente ao co-domínio $\mathbb{Q}$ existe x pertencente ao domínio $\mathbb{Q}$ tal que $f(x) = z$; basta fazer $x = z\frac{1}{3}$:

$$f\left(z\frac{1}{3}\right) = 3\left(z\frac{1}{3}\right) = z;$$

logo, f é sobrejetora;

ii: *seja $g : \mathbb{Q} \to \mathbb{Q}$ uma função tal que*

$$g(x) = x^2;$$

não existe x pertencente a $\mathbb{Q}$ tal que $g(x) = -1$, uma vez que não existe racional x tal que $x^2 = -1$; logo, g não é sobrejetora;

iii: *toda função identidade é sobrejetora.*

Definição 4.14. *Sejam $f : a \to b$ e $g : b \to c$ funções. A* composição de g com f *é a função $g \circ f : a \to c$ tal que*

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

para todo x pertencente a a.

Exemplo 4.17. *Sejam $f : \mathbb{Q} \to \mathbb{Q}$ e $g : \mathbb{Q} \to \mathbb{Q}$ tais que*

$$f(x) = 2x \quad \text{e} \quad g(x) = x + 2.$$

Logo, $g \circ f : \mathbb{Q} \to \mathbb{Q}$ é dada por

$$(g \circ f)(x) = g(f(x)) = g(2x) = (2x) + 2 = 2x + 2,$$

enquanto $f \circ g : \mathbb{Q} \to \mathbb{Q}$ é dada por

$$(f \circ g)(x) = f(g(x)) = f(x + 2) = 2(x + 2) = 2x + 4.$$

Esse exemplo deixa claro que composição é uma operação binária não comutativa. Com efeito, o Axioma da Extensionalidade garante que $f \circ g \neq g \circ f$

Importante notar que o fato de existir a composição $f \circ g$ não implica necessariamente que existe a composição $g \circ f$. Por exemplo, sejam $g : x \to y$ e $f : y \to z$ funções tais que $x \neq y$, $y \neq z$ e $x \neq z$.

Neste caso, existe $f \circ g$, mas não $g \circ f$, uma vez que o domínio de f coincide com o co-domínio de g, mas o domínio de g não coincide com o co-domínio de f.

Outra questão importante é que composição $\circ$ não é uma função

$$\circ : \mathfrak{h} \times \mathfrak{h} \to \mathfrak{h},$$

onde $\mathfrak{h}$ é o conjunto de todas as funções e $\circ(f, g) = f \circ g$. Isso porque o Esquema de Separação não permite definir o conjunto $\mathfrak{h}$ de todas as funções. Com efeito, funções são casos particulares de relações, as quais são casos particulares de conjuntos. Uma vez que não há o conjunto de todos os conjuntos em ZF, logo não é possível escolher um conjunto universo que permita definir $\mathfrak{h}$ através do Esquema de Separação.

ℹ Não obstante, alguns autores se referem à composição $\circ$ como uma 'função' com domínio $\mathfrak{h} \times \mathfrak{h}$ e co-domínio $\mathfrak{h}$. Este é um abuso de linguagem, no contexto de ZF. Em certas teorias de conjuntos como NBG, é possível qualificar $\mathfrak{h}$ como uma *classe própria* e, então, garantir que $\circ$ também é uma classe própria. Isso porque, diferentemente de ZF e ZFC, em NBG nem todos os termos são conjuntos. Em NBG todos os termos são *classes*. Entre as classes há aquelas que são conjuntos, enquanto as classes que não são conjuntos são chamadas de *classes próprias*. Mas este é um assunto que extrapola os objetivos deste livro.

Teorema 4.19. *Composição é uma operação associativa, i.e.,*

$$(f \circ g) \circ h = f \circ (g \circ h),$$

se todas as composições envolvidas existirem.

✍ A demonstração é imediata.

Definição 4.15. *Uma função $f : a \to b$ admite* inversa *sss existe $g : b \to a$ tal que*

$$g \circ f = \mathbb{I}_a$$

e

$$f \circ g = \mathbb{I}_b,$$

sendo $\mathbb{I}_a$ e $\mathbb{I}_b$ as funções identidade sobre a e b, respectivamente (ver Teorema 4.18 e o parágrafo que segue a sua demonstração). Denota-se a inversa g de f, quando existe, por f^{-1}.

Não confundir $f^{-1}(x)$ com $(f(x))^{-1}$. O primeiro caso se refere à imagem de x pela função inversa de f. O segundo se refere ao simétrico multiplicativo da imagem de x por f, pelo menos no caso em que $f(x)$ é um racional diferente do neutro aditivo.

Exemplo 4.18. *Seja $f : \mathbb{Q} \to \mathbb{Q}$ tal que*

$$f(x) = 3x.$$

Logo, $f^{-1} : \mathbb{Q} \to \mathbb{Q}$ é tal que

$$f^{-1}(x) = x\frac{1}{3}.$$

Com efeito,

$$(f^{-1} \circ f)(x) = f^{-1}(f(x)) = f^{-1}(3x) = 3x\frac{1}{3} = x,$$

ou seja, $f^{-1} \circ f = \mathbb{I}_{\mathbb{Q}}$; além disso,

$$(f \circ f^{-1})(x) = f(f^{-1}(x)) = f\left(x\frac{1}{3}\right) = 3\left(x\frac{1}{3}\right) = x,$$

ou seja, $f \circ f^{-1} = \mathbb{I}_{\mathbb{Q}}$.

Observar que a definição de função inversa não oferece qualquer procedimento efetivo para a determinação de f^{-1}, caso esta exista. Em outras palavras, a definição em si não 'ensina como calcular f^{-1}'. Apenas 'ensina' como *verificar* se uma dada g é inversa de f.

A definição de função inversa garante que, se $(x, y) \in f$, então $(y, x) \in f^{-1}$. Esse fato é importante para o próximo teorema, o qual estabelece que a inversa da inversa de uma função, quando existe, é a própria função.

Teorema 4.20. *Se f admite inversa f^{-1}, então $(f^{-1})^{-1} = f$.*

Demonstração: Seja $f : a \to b$ uma função com inversa

$$f^{-1} : b \to a.$$

Se $f(x) = y$ para algum $x \in a$, então $(x, y) \in f$. Logo, $(y, x) \in f^{-1}$. Logo, $(x, y) \in (f^{-1})^{-1}$. A partir da mesma estratégia, se $(x, y) \in (f^{-1})^{-1}$, então $(y, x) \in f^{-1}(x, y)$. Logo, $(x, y) \in f$. Ou seja, o Axioma da Extensionalidade garante que $f = (f^{-1})^{-1}$.

Definição 4.16. *$f : a \to b$ é* bijetora *(ou* bijetiva*) sss f é injetora e sobrejetora.*

A definição de bijetividade (também chamada de *bijeção*) dada acima é necessária por pelo menos dois motivos:

i: injetividade e sobrejetividade são propriedades independentes uma da outra; isso porque podem existir funções sobrejetoras injetoras, funções sobrejetoras não injetoras, funções injetoras não sobrejetoras e funções que não são nem injetoras e nem sobrejetoras;

ii: há uma estreita relação entre funções bijetoras e aquelas que admitem inversa, como se percebe no próximo teorema.

Teorema 4.21. *Uma função $f : a \to b$ admite inversa sss f é bijetora.*

Demonstração: Uma vez que este teorema envolve uma bicondicional, a prova é dividida em duas partes. Isso porque bicondicional é uma conjunção de duas condicionais.

Parte $\Rightarrow$. Se $f : a \to b$ admite inversa $f^{-1} : b \to a$, é necessário provar que f é sobrejetora e injetora. Seja $y \in b$, de modo que $f^{-1}(y) = x$. Então

$$f(x) = f(f^{-1}(y)) = (f \circ f^{-1})(y) = \mathbb{I}_b(y) = y.$$

Logo, f é sobrejetora. Agora, sejam x_1 e x_2 elementos do domínio a de f, tais que $f(x_1) = f(x_2)$. Se a fórmula $f(x_1) = f(x_2)$ implicar na fórmula $x_1 = x_2$, provamos a injetividade de f. Sejam $y = f(x_1)$ e $x = f^{-1}(y)$. Logo,

$$x_2 = \mathbb{I}_a(x_2) = (f^{-1} \circ f)(x_2) = f^{-1}(f(x_2)) =$$
$$f^{-1}(f(x_1)) = f^{-1}(y) = x.$$

No entanto,

$$x_1 = \mathbb{I}_a(x_1) = (f^{-1} \circ f)(x_1) = f^{-1}(f(x_1)) = f^{-1}(y) = x.$$

Logo, a transitividade da igualdade garante que $x_1 = x_2$. Isso conclui a prova da primeira parte.

Parte $\Leftarrow$. Se $f : a \to b$ é bijetiva, precisamos apenas provar que ela admite inversa. Seja $r : b \to a$ uma relação definida da seguinte maneira: uma vez que f é sobrejetora, para qualquer y pertencente a b existe x pertencente a a tal que $f(x) = y$; logo, basta fazer $r(y) = x$; uma vez que f é injetora, tal x é único; logo, r é uma função; além disso,

$$f \circ r = \mathbb{I}_b$$

e

$$r \circ f = \mathbb{I}_a.$$

Ou seja, $r = f^{-1}$.

Em outras palavras, bijetividade e inversibilidade são conceitos equivalentes. Funções que admitem inversa são também conhecidas como *inversíveis*.

Teorema 4.22. *Se f é uma função bijetora, sua inversa f^{-1} também é.*

✍ A prova fica a cargo do leitor.

Teorema 4.23. *A composição entre funções injetoras, quando existe, é uma função injetora. Ademais, a composição entre funções sobrejetoras, quando existe, é uma função sobretora.*

✍ Ou seja, a composição entre funções bijetoras, quando existe, é uma função bijetora. A prova deste último teorema fica por conta do leitor.

Seja $f : x \to y$ uma função. Logo,

$$f \circ \mathbb{I}_x$$

e

$$\mathbb{I}_y \circ f$$

sempre existem, independentemente da função f. Além disso,

$$f \circ \mathbb{I}_x = f$$

e

$$\mathbb{I}_y \circ f = f.$$

Isso significa que funções identidade operam como *elementos neutros* relativamente à composição, desde que seja tomado cuidado com o domínio (o qual coincide com o co-domínio) de cada função identidade. Se $f : x \to y$ admite inversa f^{-1}, então f^{-1} é um *simétrico* *composicional* de f, uma vez que $f^{-1} \circ f = \mathbb{I}_x$ e $f \circ f^{-1} = \mathbb{I}_y$. No caso em que

$$f^{-1} \circ f = \mathbb{I}_x,$$

diz-se que f^{-1} é a *inversa à esquerda* de f. No caso em que

$$f \circ f^{-1} = \mathbb{I}_y,$$

diz-se que f^{-1} é a *inversa à direita* de f. Logo, f é inversível sss f admite inversa à esquerda e à direita.

✍ Para uma função admitir inversa à direita, basta ser injetiva. Para admitir inversa à esquerda, basta ser sobrejetiva. Recomendamos ao leitor provar essas duas últimas afirmações.

Seção 33

Conjuntos infinitos

Sumário
Índice
Rede

Um *ordinal finito* é qualquer elemento de ω (ver Definição 3.5), e apenas elementos de ω são ordinais finitos. A motivação para esse conceito reside no fato de que ordinais podem ser estendidos para outros, além dos ordinais finitos. Detalhes podem ser encontrados em [28].

Exemplo 4.19. i: 2022 *é um ordinal finito;*

ii: ω *não é um ordinal finito, uma vez que* $\omega \notin \omega$*;*

iii: $S(\omega)$ *não é um ordinal finito;*

iv: ✍ *o sucessor de um ordinal finito é um ordinal finito, uma vez que* ω *é indutivo. Recomendamos ao leitor que prove isso.*

Definição 4.17. *Um conjunto* x *é equipotente a* y *sss existe bijeção* $f : x \to y$*. Denotamos isso por* $x \sim y$*.*

Não confundir a notação $\sim$ para equipotência entre conjuntos com a mesma notação empregada na Seção 30 para definir inteiros a partir de naturais.

A ideia intuitiva por trás da última definição é milenar, muito anterior ao advento das teorias de conjuntos. Matemáticos aprenderam a 'contar' fazendo correspondências *um-para-um*. Uma bijeção $f : x \to y$ é uma 'correspondência' de *cada* elemento de x a um, e *apenas um*, elemento de y, e de *cada* elemento de y a um, e *apenas um*, elemento de x. Tal 'correspondência' é possível graças ao fato da linguagem $\mathfrak{S}$ aqui empregada contar com a igualdade $=$. Onde há igualdade, há a negação dela (pelo menos sob os cânones da lógica clássica), para garantir a discernibilidade dos elementos de x, bem como dos elementos de y. Logo, por exemplo, o conjunto

$$x = \{\varnothing, \{\varnothing\}, \{\{\varnothing\}\}\}$$

tem três elementos porque x é equipotente ao natural 3. Com efeito, uma vez que o natural 3 é o conjunto

$$3 = \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\}$$

(conforme Seção 23), é possível definir uma bijeção $f : x \to 3$, dada, digamos, por

$$f(\varnothing) = \varnothing, \quad f(\{\varnothing\}) = \{\varnothing\}, \quad f(\{\{\varnothing\}\}) = \{\varnothing, \{\varnothing\}\}.$$

O mérito da ideia acima reside no fato de que é possível qualificar que um conjunto x tem três elementos, ainda que x seja diferente do ordinal finito 3.

Se n é um ordinal finito, um conjunto x tem n elementos sss x for equipotente a n.

Teorema 4.24. *Equipotência entre conjuntos é reflexiva, simétrica e transitiva.*

Demonstração: i: Todo conjunto x é equipotente a si mesmo. Com efeito, basta definir $f : x \to x$ tal que $f(a) = a$ para todo a pertencente a x. Tal f é bijetora. Logo, $x \sim x$.

ii: Se $x \sim y$ então existe bijeção $f : x \to y$. Logo, existe inversa de f dada por $f^{-1} : y \to x$, a qual é bijetora (Teorema 4.22). Logo, $y \sim x$.

iii: Se $x \sim y$ e $y \sim z$, então existem $f : x \to y$ e $g : y \to z$ bijetoras. Logo, $g \circ f : x \to z$ é bijetora (Teorema 4.23). Logo, $x \sim z$.

Por conta do Esquema da Separação, não existe em ZF o conjunto de todos os conjuntos. Isso porque, para definir o conjunto de todos os conjuntos via Separação é necessário um conjunto universo, o qual deveria ser o conjunto de todos os conjuntos. Mas essa estratégia é uma *circularidade*, no sentido de que não permite discernir *definiendum* de *definiens*. Logo, equipotência entre conjuntos não é uma relação no sentido da Seção 25. No entanto, é usual se referir a $\sim$ como uma relação de equivalência no sentido do Teorema 4.24, por conta da reflexividade, simetria e transitividade de $\sim$.

Definição 4.18. *Um conjunto x é* finito *sss x é equipotente a um ordinal finito. Caso contrário, dizemos que x é* infinito.

Teorema 4.25. *Todo ordinal finito é um conjunto finito.*

A prova deste último teorema é imediata. Se n é um ordinal finito, basta definir $f : n \to n$ como $f(a) = a$, para todo $a \in n$. Isso porque a diagonal de qualquer ordinal finito é uma bijeção.

A recíproca do último teorema não é teorema. Com efeito, basta exibir um conjunto finito que não seja um ordinal finito. Por exemplo, $x = \{3, 4\}$ não é um ordinal finito. Para provar que x é finito, considere a função $f : x \to 2$ tal que $f(3) = 0$ e $f(4) = 1$. Logo, $\{3, 4\} \sim 2$.

✍ O conjunto ω é infinito. Com efeito, seja n um ordinal finito. Logo, qualquer $f : \omega \to n$ é não injetora. Igualmente, qualquer função $g : n \to \omega$ é não sobrejetora. Consegue provar esses resultados?

Definição 4.19. *Um conjunto x é* Dedekind-infinito *sss existe $y \subset x$ tal que $y \sim x$. Caso contrário, x é* Dedekind-finito.

Exemplo 4.20. *ω é um conjunto Dedekind-infinito. Com efeito, seja*

$$p = \{n \in \omega \mid \exists m(m \in \omega \wedge n = 2m)\}.$$

O termo p dado é o conjunto dos naturais pares e, portanto, subconjunto próprio de ω. Seja agora

$$f : \omega \to p$$

dada por $f(n) = 2n$. Se $m \neq n$, então $2m \neq 2n$. Logo, $f(m) \neq f(n)$. Logo, f é injetora. Além disso, todo natural par é o dobro de um natural, o que garante que f é sobrejetora. Logo, $p \sim \omega$.

Exemplo 4.21. ✍ *O conjunto* $x = \{3, 4\}$ *é Dedekind-finito. Com efeito, se* $y \subset x$*, então* $y = \{3\}$*, ou* $y = \{4\}$ *ou* $y = \varnothing$*. Em qualquer um dos casos não há bijeção* $f : x \to y$*.*

Nos primórdios dos estudos sobre teoria de conjuntos, alguns matemáticos acreditavam que conjuntos infinitos e conjuntos Dedekind-infinitos eram conceitos equivalentes. Mas, com o tempo, foi percebido que este não é necessariamente o caso, especialmente em formulações de ZF nas quais o Axioma da Escolha não é teorema. Via Teoria de Modelos (Seção 111) é possível provar a existência de conjuntos infinitos que são Dedekind-finitos. Mas este é um assunto que está fora do escopo deste livro.

Definição 4.20. *Seja* x *um conjunto não vazio. Um conjunto* m *pertencente a* x *é* maximal relativamente à inclusão *(ou simplesmente* maximal*, se não houver risco de confusão) sss*

$$\forall r((r \in x \wedge m \subseteq r) \Rightarrow m = r).$$

Em outras palavras, o maximal m de x não está contido em qualquer outro elemento de x além dele mesmo.

Conjuntos quaisquer podem ter um único maximal, nenhum, ou vários maximais, conforme ilustrado a seguir.

Exemplo 4.22. *Seja* x *um conjunto não vazio. Logo,* $\wp(x)$ *admite um único maximal, a saber,* x*.*

Exemplo 4.23. *Seja* x *um conjunto não vazio. Seja também* y *o conjunto*

$$y = \{r \in x \mid r \text{ é singleton}\},$$

o qual foi definido usando o Esquema de Separação. Neste caso, todo elemento de y *é maximal.*

Exemplo 4.24. ✍ *Seja*

$$y = \{r \in \wp(\omega) \mid r \text{ é finito}\},$$

onde ω *é o conjunto dos naturais. Neste caso,* y *não admite qualquer maximal. Em termos mais gerais, se* x *é infinito e* y *é o conjunto de todos os subconjuntos finitos de* y*, então* y *não admite qualquer maximal. Sugerimos que o leitor prove isso.*

Uma das possíveis aplicações do conceito de maximal de um conjunto é a definição de base para um espaço vetorial qualquer, conforme Seção 97.

I: Prove que o conjunto $x = \{3, 4, 5\}$ é finito e Dedekind-finito;

II: prove que o conjunto dos números naturais ímpares é infinito e Dedekind-infinito.

Seção 34

Preliminares para os reais

Sumário
Índice
Rede

O leitor pode ignorar esta discussão e avançar para Seção 35, sem prejuízo significativo para o que vem adiante. O objetivo aqui é apenas motivar os mais sedentos pelo conhecimento.

Até o presente momento foi mostrado como ZF permite edificar números naturais, inteiros e racionais. Naturais são construídos a partir do conjunto vazio e da operação monádica Sucessor, em parceria com o Axioma do Infinito. Inteiros são definidos como classes de equivalência de pares ordenados de naturais. Racionais são definidos como classes de equivalência de pares ordenados de inteiros. No entanto, qualquer tentativa de definir números reais como classes de equivalência de pares ordenados de racionais está fadada ao fracasso. Apresentamos aqui um esboço da prova deste resultado, o qual é dividido em duas partes.

Na *primeira parte* provamos que é impossível existir bijeção entre ω e o conjunto dos números reais. Ainda que a definição de número real não tenha sido dada até este momento, qualquer que seja a definição, ela deve ser consistente com a representação de números reais na notação decimal usual dada a seguir:

$$i_n i_{n-1} i_{n-2} \cdots i_2 i_1 i_0, d_0 d_1 d_2 \cdots$$

onde cada i_j e cada d_k é um dos dez símbolos 0, 1, 2, 3, 4, 5, 6, 7, 8 e 9, exceto possivelmente i_n. Isso porque, no caso da sequência finita $i_n i_{n-1} i_{n-2} \cdots i_2 i_1 i_0$ contar com mais de uma ocorrência de símbolos, então $i_n \neq 0$.

Exemplo 4.25. i: $1945,00689$ *é a representação de um número real na notação decimal usual; com efeito,* $i_3 = 1$, $i_2 = 9$, $i_1 = 4$, $i_0 = 5$, $d_0 = 0$, $d_1 = 0$, $d_2 = 6$, $d_3 = 8$, $d_4 = 9$ *e os demais* d_n *são 0, com* $n > 4$;

ii: $0,3333\cdots$ *é a representação de um número real na notação decimal usual; com efeito,* $i_0 = 0$ *e todos os* d_n, *onde* n *é um natural, são iguais a 3.*

Observar que empregamos, no Exemplo acima, uma *linguagem infinitária* para representar números reais em base decimal usual. *Linguagens infinitárias* são aquelas que admitem sentenças de comprimento não finito, enquanto sentenças de *linguagens finitárias* sempre são sequências finitas de símbolos da linguagem. O item ii do último Exemplo ilustra uma sentença de comprimento não finito. Um dos aspectos mais fascinantes de ZF é o fato desta teoria formal empregar uma linguagem *finitária* (conforme Seção 7) que permite conceituar números reais (como é mostrado na Seção 39).

Ademais, os números reais devem contar com relações de ordem total $\leq$ (menor ou igual) e $\geq$ (maior ou igual) análogas às relações de ordem total entre inteiros e racionais, de modo que os reais sejam capazes de copiá-los.

Neste contexto, qualquer número real maior ou igual ao real 0 (o qual deve ser neutro aditivo) e menor ou igual a 1 (o qual deve ser neutro multiplicativo) pode ser representado da seguinte maneira:

$$0, d_0 d_1 d_2 d_3 d_4 \cdots$$

onde cada d_k é um dos dez dígitos do sistema decimal usual, para todo k natural.

Exemplo 4.26. i: $0,00689$;

ii: $0,3333\cdots$; *neste caso* d_k *é igual a 3, para cada* k *natural.*

Item ii *é um caso particular daquilo que é conhecido como* dízima periódica.

Agora, seja $[0,1]$ o conjunto de todos os números reais maiores ou iguais a 0 e menores ou iguais a 1. Supor que $[0,1]$ e ω são equipotentes, i.e., existe uma bijeção

$$f : \omega \to [0,1].$$

Logo, a cada natural de ω corresponde um e apenas um real do conjunto $[0,1]$; e cada real deste conjunto corresponde a um apenas um natural de ω. Podemos representar tal bijeção da seguinte maneira:

$$\begin{array}{l}
0 \longrightarrow 0, d_{00}d_{01}d_{02}d_{03}d_{04}d_{05}\cdots \\
1 \longrightarrow 0, d_{10}d_{11}d_{12}d_{13}d_{14}d_{15}\cdots \\
2 \longrightarrow 0, d_{20}d_{21}d_{22}d_{23}d_{24}d_{25}\cdots \\
3 \longrightarrow 0, d_{30}d_{31}d_{32}d_{33}d_{34}d_{35}\cdots \\
4 \longrightarrow 0, d_{40}d_{41}d_{42}d_{43}d_{44}d_{45}\cdots \\
5 \longrightarrow 0, d_{50}d_{51}d_{52}d_{53}d_{54}d_{55}\cdots \\
\qquad\qquad\vdots
\end{array}$$

sendo que cada d_{ij} é um dos dez símbolos do sistema decimal.

Neste contexto, cada natural n corresponde a um real

$$0, d_{n0}d_{n1}d_{n2}d_{n3}d_{n4}d_{n5}\cdots$$

pertencente a $[0,1]$, no sentido de que

$$f(n) = 0, d_{n0}d_{n1}d_{n2}d_{n3}d_{n4}d_{n5}\cdots.$$

No caso particular em que o real correspondente a um n natural é 0, temos d_{nk} igual a 0, para todo k natural. No caso particular em que o real correspondente a um natural m é 1, temos d_{mk} igual a 9, para todo k natural. Com efeito, a dízima periódica $0,999\cdots$ é igual à dízima periódica $0,333\cdots$ multiplicada por 3. No entanto, $0,333\cdots = \frac{1}{3}$. Mas, $\frac{1}{3}$ multiplicado por 3 é 1. Ou seja, $0,999\cdots$ e 1 são apenas notações distintas para o mesmo número real, a saber, o neutro multiplicativo entre reais.

Agora considere o seguinte número real r do conjunto $[0,1]$:

$$r = 0, r_0r_1r_2r_3r_4r_5\cdots$$

sendo que cada r_i é igual a $9 - d_{ii}$, para cada natural i.

Ou seja,

- se $d_{ii} = 9$, então $r_i = 0$;
- se $d_{ii} = 8$, então $r_i = 1$;
- se $d_{ii} = 7$, então $r_i = 2$;
- se $d_{ii} = 6$, então $r_i = 3$;
- se $d_{ii} = 5$, então $r_i = 4$;

- se $d_{ii} = 4$, então $r_i = 5$;
- se $d_{ii} = 3$, então $r_i = 6$;
- se $d_{ii} = 2$, então $r_i = 7$;
- se $d_{ii} = 1$, então $r_i = 8$;
- se $d_{ii} = 0$, então $r_i = 9$.

Logo, r_i é sempre diferente de d_{ii}.

Neste caso, r será diferente de

$$0, d_{00}d_{01}d_{02}d_{03}d_{04}d_{05}\cdots,$$

uma vez que $r_0 \neq d_{00}$. Analogamente, r será diferente de

$$0, d_{10}d_{11}d_{12}d_{13}d_{14}d_{15}\cdots,$$

uma vez que $r_1 \neq d_{11}$. De maneira análoga, r será diferente de cada

$$d_{n0}d_{n1}d_{n2}d_{n3}d_{n4}d_{n5}\cdots,$$

uma vez que cada r_i é diferente de d_{ii}.

Isso significa que r é diferente de toda e qualquer imagem $f(n)$. Logo, qualquer função injetora $f : \omega \to [0, 1]$ jamais pode ser sobrejetora. Com efeito, sempre restará pelo menos um real r pertencente a $[0, 1]$ que não é igual a $f(n)$ para natural n algum do domínio de f. Na verdade é possível provar que existe uma infinidade de reais r diferentes de todo e qualquer $f(n)$. Mas basta exibir um r de $[0, 1]$ que não é igual a qualquer $f(n)$, para garantir que f não é sobrejetora. Logo, f não pode ser bijetora, como foi inicialmente assumido.

Se nenhuma função $f : \omega \to [0, 1]$ pode ser bijetora, então ω não é equipotente ao conjunto $[0, 1]$ de números reais entre 0 e 1, incluindo 0 e 1. Uma consequência imediata disso é que ω não é equipotente ao próprio conjunto dos números reais, uma vez que $[0, 1]$ deve ser subconjunto do conjunto dos números reais.

Na *segunda parte* da prova é mostrado que, qualquer tentativa de construir os reais a partir de pares ordenados de racionais implica que o conjunto de números reais deve ser, na melhor das hipóteses, equipotente a ω. Uma vez que isso contradiz o que foi provado na primeira parte, logo, não é possível definir números reais a partir de pares ordenados de racionais.

Em primeiro lugar, é possível provar que existe bijeção

$$f : \omega \times \omega \to \omega$$

(a qual garante uma bijeção $f^{-1} : \omega \to \omega \times \omega$), ou seja, ω é equipotente a $\omega \times \omega$. Com efeito, considere f da seguinte maneira (apenas esboço a definição de f):

I: $f(0, n) = 2n$; dessa maneira teremos $f(0, 0) = 0$, $f(0, 1) = 2$, $f(0, 2) = 4$ e assim por diante, cobrindo todos os naturais pares;

II: uma vez que restaram apenas os naturais ímpares para serem imagens de elementos do domínio $\omega \times \omega$ via a bijeção f, fazemos $f(1, 0) = 1$ (o primeiro ímpar) e, para os demais $f(1, n)$, 'pulamos' sempre um ímpar, de modo que $f(1, 1) = 5$ (pulamos o 3), $f(1, 2) = 9$ (pulamos o 7), $f(1, 3) = 13$ (pulamos o 11) e assim por diante;

III: ainda resta uma infinidade de ímpares para serem imagens de elementos de $\omega \times \omega$ (os ímpares 'pulados' no passo anterior); uma vez que o primeiro ímpar 'pulado' foi 3, fazemos $f(2, 0) = 3$ e, para os demais $f(2, n)$ novamente 'pulamos' um ímpar por vez, entre aqueles que ainda não são imagens de algum (m, n); de modo que $f(2, 1) = 11$ (pulamos o 7), $f(2, 2) = 19$ (pulamos o 15), $f(2, 3) = 27$ (pulamos o 23) e assim por diante;

IV: repetimos o processo por *indução infinita*, de modo a cobrir todos os ímpares. Logo, f é uma bijeção.

Em segundo lugar, ω é equipotente a $\mathbb{Z}$. Com efeito, basta considerar a seguinte bijeção $f : \omega \to \mathbb{Z}$ dada por $f(0) = 0$, $f(1) = -1$, $f(2) = 1$, $f(3) = -2$, $f(4) = 2$, $f(5) = -3$, $f(6) = 3$ e assim por diante.

Outro resultado espantoso é o fato de ω ser equipotente a $\mathbb{Q}$.

Antes de provar isso, vale ressaltar que todas as técnicas aqui usadas podem ser empregadas para provar também que $\mathbb{Q}$ é equipotente a $\mathbb{Q} \times \mathbb{Q}$. Uma vez que equipotência é transitiva, todos esses resultados apontam para o fato de que $\mathbb{Q} \times \mathbb{Q}$ é equipotente a ω. Logo, qualquer tentativa de estabelecer uma bijeção entre o conjunto dos números reais e $\mathbb{Q} \times \mathbb{Q}$ deve fracassar, no sentido de que tal bijeção simplesmente não existe.

Observar que a existência de tal bijeção é indispensável, uma vez que eventuais partições de $\mathbb{Q} \times \mathbb{Q}$ devem ser definidas por classes de equivalência $[(r, s)]$ (onde r e s são racionais) de modo que cada uma delas corresponde a um e apenas um número real.

Com relação à demonstração de que ω é equipotente a $\mathbb{Q}$, considere $f : \omega \to \mathbb{Q}$ dada como se segue: $f(0) = 0$ e as demais imagens $f(n)$ são dadas de acordo com a tabela abaixo, na qual estão representados todos os racionais diferentes de 0.

$$\begin{array}{ccccccccc}
\frac{+1}{1} & \frac{-1}{1} & \frac{+1}{2} & \frac{-1}{2} & \frac{+1}{3} & \frac{-1}{3} & \frac{+1}{4} & \frac{-1}{4} & \cdots \\
\frac{+2}{1} & \frac{-2}{1} & \frac{+2}{2} & \frac{-2}{2} & \frac{+2}{3} & \frac{-2}{3} & \frac{+2}{4} & \frac{-2}{4} & \cdots \\
\frac{+3}{1} & \frac{-3}{1} & \frac{+3}{2} & \frac{-3}{2} & \frac{+3}{3} & \frac{-3}{3} & \frac{+3}{4} & \frac{-3}{4} & \cdots \\
\frac{+4}{1} & \frac{-4}{1} & \frac{+4}{2} & \frac{-4}{2} & \frac{+4}{3} & \frac{-4}{3} & \frac{+4}{4} & \frac{-4}{4} & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{array}$$

Argumento da diagonal de Cantor

Seguindo as flechas da esquerda para a direita, conforme os sentidos indicados,

$$f(1) = \frac{+1}{1},\ f(2) = \frac{-1}{1},\ f(3) = \frac{+2}{1},\ f(4) = \frac{+1}{2},\ f(5) = \frac{-2}{1},$$

$$f(6) = \frac{+3}{1},\ f(7) = \frac{-1}{2},\ f(8) = \frac{-3}{1},\ f(9) = \frac{+4}{1},\ f(10) = \frac{+1}{3}$$

e assim por diante.

O cuidado a ser tomado é evitar imagens repetidas, para garantir a injetividade de f. Afinal, por exemplo,

$$\frac{+2}{4} = \frac{+1}{2}.$$

Toda vez que ocorrer um racional repetido, basta ignorá-lo e ir para o próximo na diagonal correspondente, para definir f.

As técnicas usadas acima para provar a equipotência dos racionais com os naturais e a não equipotência dos naturais com os reais são conhecidas na literatura como o *argumento da diagonal de Cantor*. Isso porque essas técnicas foram concebidas por Georg Cantor, e publicadas em 1891.

Seção 35

Sequências

Sumário
Índice
Rede

Sequências são casos especiais de funções.

Definição 4.21. *x é uma* sequência *sss x é uma função com domínio de racionais que copia ω.*

Por abuso de notação chamamos esse domínio de ω. Eventualmente, por questão de conveniência, podemos omitir o natural 0 do domínio de uma sequência. No caso especial de uma sequência x, usualmente $x(n)$ (a imagem de n por x) é denotada por x_n.

Uma *sequência racional* é uma sequência cujas imagens são números racionais.

Exemplo 4.27. i: *$x : \omega \to \mathbb{Q}$ tal que $x_n = 7$.*

Neste caso, $x_0 = 7$, $x_1 = 7$, $x_2 = 7$, $\cdots$. Observar que $(0,7) \in x$, $(1,7) \in x$ e assim por diante;

ii: *$y : \omega - \{0\} \to \mathbb{Q}$ tal que $y_n = \frac{1}{n}$.*

Neste caso, $y_1 = 1$, $y_2 = \frac{1}{2}$, $y_3 = \frac{1}{3}$, $\cdots$; ou seja $(1,1) \in y$, $(2, \frac{1}{2}) \in y$, $(3, \frac{1}{3}) \in y$ e assim por diante.

De agora em diante, por questão de conveniência, são empregados *quantificadores relativizados*, os quais são amplamente empregados em Cálculo Diferencial e Integral Padrão.

Seja $\mathcal{P}$ uma fórmula. Logo:

$$\forall \varepsilon > 0(\mathcal{P}) \vdots \forall \varepsilon(\varepsilon > 0 \Rightarrow \mathcal{P});$$

$$\exists \delta > 0(\mathcal{P}) \vdots \exists \delta(\delta > 0 \wedge \mathcal{P}).$$

Lê-se $\forall \varepsilon > 0(\mathcal{P})$ como ‘para todo ε maior do que zero, $\mathcal{P}$’. Lê-se $\exists \delta > 0(\mathcal{P})$ como ‘existe δ maior do que zero tal que $\mathcal{P}$’. Obviamente quantificadores relativizados representam economia de notação. Adotamos essa convenção para que este texto fique em sintonia com práticas comuns encontradas em livros de cálculo diferencial e integral e análise matemática, entre outros.

No primeiro caso da última definição, o quantificador universal está relativizado à fórmula $\varepsilon > 0$. No segundo caso, o quantificador existencial está relativizado à fórmula $\delta > 0$.

Na próxima definição pretende-se capturar a seguinte ideia: uma sequência racional x_n *converge* para um racional L se, e somente se, independentemente de qualquer valor racional para ε estritamente positivo, as imagens x_n ficam confinadas ao intervalo aberto

$$(L - \varepsilon, L + \varepsilon),$$

desde que n seja suficientemente grande.

O intervalo aberto em questão é apenas o conjunto de todos os racionais r tais que $L - \varepsilon < r < L + \varepsilon$ (isso é uma abreviação para a fórmula $L - \varepsilon < r \wedge r < L + \varepsilon$). Uma vez que a terminologia ‘confinadas’ e ‘suficientemente grande’ é vaga, há a necessidade de traduzir essa ideia na linguagem de ZF, como se segue:

Definição 4.22.

$$x_n \to L \vdots \forall \varepsilon > 0 \, \exists \delta > 0 (n > \delta \Rightarrow |x_n - L| < \varepsilon).$$

Lê-se $x_n \to L$ como ‘x_n converge *para L’.*

É uma prática comum se referir a uma sequência x como x_n, se não houver risco de confusão.

Seguem algumas observações.

i: Todos os termos envolvidos na última definição são racionais: L, ε, δ, 0, n, x_n.

ii: Se x é racional, então $|x| = x$ se $x \geq 0$, e $|x| = -x$ se $x < 0$ (lê-se $|x|$ como ‘*valor absoluto* de x’).

Por exemplo, $|5| = |-5| = 5$.

✍ É teorema em ZF a seguinte fórmula:

$$|a + b| \leq |a| + |b|,$$

para quaisquer a e b racionais (recomendamos provar esse resultado).

iii: O termo $|x_n - L|$ é uma *distância* entre x_n e L. Aqui cabe um breve comentário: no estudo de *espaços métricos* (Seção 86) qualifica-se o que é a distância entre um termo a e um termo b; neste sentido é possível provar que, de fato, $|x_n - L|$ é uma *distância* entre x_n e L; no entanto, para os propósitos deste texto, basta saber que $|x_n - L|$ captura a ideia intuitiva do que deve ser a distância entre duas ocorrências de racionais; por exemplo, a distância entre $\frac{-1}{3}$ e $\frac{3}{5}$ é

$$|\frac{-1}{3} - \frac{3}{5}| = |\frac{3}{5} - \frac{-1}{3}| = \frac{14}{15}.$$

iv: Nem toda sequência racional x_n converge para algum racional L, como é ilustrado em alguns exemplos adiante.

v: O valor racional L é chamado de *limite da sequência* x_n.

vi: Excepcionalmente estão sendo usadas letras latinas maiúsculas em itálico, na definição de sequência convergente, por um motivo de caráter pragmático: faz parte da literatura padrão esse tipo de notação.

A definição de sequência racional convergente (ou seja, com limite L) dada acima captura exatamente a interpretação pretendida que foi anteriormente sugerida. O valor racional estritamente positivo ε define, para efeitos práticos, o que é *confinar* x_n ao intervalo aberto $(L - \varepsilon, L + \varepsilon)$. Com efeito, a fórmula

$$|x_n - L| < \varepsilon$$

é equivalente à fórmula

$$x_n \in (L - \varepsilon, L + \varepsilon).$$

O valor racional estritamente positivo δ define, para os mesmos efeitos práticos, o que são naturais *suficientemente grandes*: são aqueles n tais que $n > \delta$.

Logo, afirmar que a sequência racional x_n converge para o racional L é equivalente a afirmar o seguinte: dado um intervalo de confinamento

$$(L - \varepsilon, L + \varepsilon),$$

é necessário exibir um δ racional estritamente positivo tal que todo n maior do que δ garante que

$$x_n \in (L - \varepsilon, L + \varepsilon).$$

Ou seja, se δ não puder ser arbitrário, deve depender única e exclusivamente de ε. Garantir que uma sequência racional x_n tem limite é equivalente a exibir δ nas condições impostas pela definição.

Exemplo 4.28. *A sequência racional $x_n = 7$ converge para 7.*

Neste caso, o valor de δ pode ser qualquer racional estritamente positivo, uma vez que qualquer $n > \delta$ implica que

$$x_n \in (7 - \varepsilon, 7 + \varepsilon).$$

Isso porque, independentemente do valor estritamente positivo de ε, 7 (a imagem de qualquer n via x) sempre pertence ao intervalo aberto $(7 - \varepsilon, 7 + \varepsilon)$.

Uma extensão deste resultado é o tema do próximo teorema.

No teorema abaixo adota-se uma notação bastante comum na literatura para *sequências constantes* (aquelas cujas imagens x_n têm todas o mesmo valor) $x_n = c$, a saber, c.

Teorema 4.26.

$$c \to c.$$

Demonstração: Seja $x_n = c$, onde c é racional. Logo, devemos provar que

$$\forall \varepsilon > 0 \; \exists \delta > 0 (n > \delta \Rightarrow |c - c| < \varepsilon),$$

o que é equivalente a

$$\forall \varepsilon > 0 \; \exists \delta > 0 (n > \delta \Rightarrow 0 < \varepsilon).$$

Mas já temos como hipótese que $\varepsilon > 0$ (ver Teorema 2.1). Logo, qualquer δ racional maior do que 0 satisfaz a condicional da definição. Com efeito, basta perceber que $n > \delta$ é apenas uma hipótese a mais (ver Proposição 2.3).

Teorema 4.27.

$$\frac{1}{n} \to 0.$$

Demonstração: Devemos provar que

$$\forall \varepsilon > 0 \; \exists \delta > 0 \left(n > \delta \Rightarrow \left| \frac{1}{n} - 0 \right| < \varepsilon \right).$$

Podemos reescrever isso como

$$\forall \varepsilon > 0 \; \exists \delta > 0 \left(n > \delta \Rightarrow \left| \frac{1}{n} \right| < \varepsilon \right),$$

que é equivalente a

$$\forall\varepsilon > 0\, \exists\delta > 0 \left(n > \delta \Rightarrow \frac{1}{n} < \varepsilon\right),$$

uma vez que $n > \delta$ e $\delta > 0$, o que faz de n estritamente positivo. Essa última é equivalente a

$$\forall\varepsilon > 0\, \exists\delta > 0 \left(n > \delta \Rightarrow \frac{1}{\varepsilon} < n\right).$$

Finalmente, isso equivale a

$$\forall\varepsilon > 0\, \exists\delta > 0 \left(n > \delta \Rightarrow n > \frac{1}{\varepsilon}\right).$$

Se escolhermos

$$\delta = \frac{1}{\varepsilon},$$

teremos uma fórmula implicando nela mesma; isso, de acordo com Teorema 2.1, é um teorema.

Com relação à última demonstração, notar que qualquer δ' maior do que $\delta = \frac{1}{\varepsilon}$ também garante que

$$\left(n > \delta' \Rightarrow \left|\frac{1}{n} - 0\right| < \varepsilon\right)$$

é teorema. Portanto, $\delta = \frac{1}{\varepsilon}$ não é o único possível valor para δ que garante a demonstração de que $\frac{1}{n}$ converge para 0. Mas, levando em conta que a definição de sequência racional convergente exige que exista pelo menos um δ que satisfaça o *definiens*, a prova acima é suficiente.

EXEMPLO 4.29. *Supor* $\varepsilon = \frac{1}{1000}$. *Neste caso,* $\delta = 1000$. *Todo* n *maior do que* 1000 *garante que a distância entre* $\frac{1}{n}$ *e* 0 *é menor do que* $\frac{1}{1000}$.

Supor $\varepsilon = \frac{1}{3}$. *Neste caso,* $\delta = 3$. *Todo* n *maior do que* 3 *garante que a distância entre* $\frac{1}{n}$ *e* 0 *é menor do que* $\frac{1}{3}$.

TEOREMA 4.28. $\nexists L((-1)^n \to L)$.

DEMONSTRAÇÃO: As imagens de $(-1)^n$ são -1 e 1. Se, e.g., $\varepsilon = \frac{1}{10}$, nenhum δ racional maior do que 0 poderá satisfazer a definição.

Para o leitor não familiarizado com a expressão 'e.g.', esta abrevia '*exempli gratia*', a qual se traduz como 'por exemplo', do latim.

✍ Provar que

$$\frac{2}{n^2} \to 0.$$

Para resolver o exercício acima proposto, demonstrar os seguintes teoremas.

- O produto entre racionais é um racional.
- O quadrado r^2 de um racional r (ou seja, $r^2 = rr$) é um racional.
- Se r é um racional, então existe racional s tal que $s > r$.
- Se, para qualquer $\varepsilon > 0$, existe $\delta > 0$ tal que

  $$n > \delta \Rightarrow |x_n - L| < \varepsilon,$$

  e $\delta' > \delta$, então

  $$n > \delta' \Rightarrow |x_n - L| < \varepsilon.$$

Feito isso, temos o que se segue:

$$\frac{2}{n^2} \to 0 \text{ sss } \forall \varepsilon > 0 \, \exists \delta > 0 \left(n > \delta \Rightarrow |\frac{2}{n^2} - 0| < \varepsilon \right).$$

Logo,

$$\frac{2}{n^2} \to 0 \text{ sss } \forall \varepsilon > 0 \, \exists \delta > 0 \left(n > \delta \Rightarrow n^2 > \frac{2}{\varepsilon} \right).$$

Logo,

$$\frac{2}{n^2} \to 0 \text{ sss } \forall \varepsilon > 0 \, \exists \delta > 0 \left(n^2 > \delta^2 \Rightarrow n^2 > \frac{2}{\varepsilon} \right).$$

✍ Ou seja, basta escolher δ tal que $\delta^2 > \frac{2}{\varepsilon}$. Os demais detalhes ficam a cargo do leitor.

Uma das vantagens da introdução de números reais (a ocorrer na Seção 39) é que a demonstração do teorema

$$\frac{2}{n^2} \to 0$$

se torna extraordinariamente mais simples, se $\frac{2}{n^2}$ é uma sequência cujas imagens são números reais.

Seção 36

Sequências de Cauchy

Sumário
Índice
Rede

Uma notação muito comum para sequências racionais convergentes é a seguinte.

Definição 4.23.

$$\lim_{n\to\infty} x_n = L \vdots x_n \to L.$$

$\lim_{n\to\infty} x_n = L$ se lê como '*limite de x_n, com n tendendo a infinito, é L*'.

Observar que o símbolo ∞ não corresponde a termo algum da linguagem de ZF. Trata-se tão somente de um símbolo metalinguístico que serve ao propósito de destacar a condição $n > \delta$ na definição de sequência racional convergente. Neste sentido, uma sequência de racionais é convergente sss existe L tal que

$$\lim_{n\to\infty} x_n = L.$$

STOP Levando em conta que muitos alunos insistem em tratar ∞ como um termo, recomendamos que o leitor diga, diante do espelho, a seguinte frase: 'infinito não é um termo'. Repetir o procedimento cinco vezes consecutivas.

O fato de que ∞ não é um termo implica, entre outras coisas, que não são termos sentenças como '$\infty + \infty$', '$\infty - \infty$', '$\infty + 7$' etc.

Entre sequências racionais é possível definir operações de adição, multiplicação, subtração e divisão:

Definição 4.24. *Sejam x, y e z sequências racionais. Logo,*

i: *$x + y = z$ sss $z_n = x_n + y_n$, para todo $n \in \omega$;*

ii: *$x - y = z$ sss $z_n = x_n - y_n$, para todo $n \in \omega$;*

iii: *$xy = z$ sss $z_n = x_n y_n$, para todo $n \in \omega$; e*

iv: *$x/y = z$ sss $z_n = x_n/y_n$, para todo $n \in \omega$, desde que $y_n \neq 0$.*

Exemplo 4.30. *Sejam x e y sequências dadas por $x_n = 2n$ e $y_n = 7$; logo, xy é uma sequência z dada por $z_n = 14n$ e $x + y$ é uma sequência w dada por $w_n = 2n + 7$.*

Importante perceber que a adição de sequências racionais é definida a partir da adição de imagens das mesmas sequências. Consequentemente, a adição de sequências racionais é definida a partir da adição de racionais. Isso implica que as propriedades algébricas de adição entre racionais são replicadas na adição de sequências racionais. Por exemplo, a adição se sequências racionais é comutativa, associativa e admite simétrico aditivo, bem como neutro aditivo. Comentário análogo vale para as demais operações acima definidas.

O próximo teorema expressa o fato de que o limite da soma de sequências racionais é a soma dos limites das mesmas, caso estes existam.

Teorema 4.29. $(x_n \to L \wedge y_n \to M) \Rightarrow (x_n + y_n \to L + M)$.

Demonstração: Temos, por hipótese, a conjunção de duas fórmulas, a saber,

i: $\forall \varepsilon > 0\ \exists \delta' > 0 (n > \delta' \Rightarrow |x_n - L| < \varepsilon)$ e

ii: $\forall \varepsilon > 0\ \exists \delta'' > 0 (n > \delta'' \Rightarrow |y_n - M| < \varepsilon)$.

Uma vez que a definição de sequência racional convergente exige que sejam considerados todos os ε racionais estritamente positivos, não há problema algum em assumir o mesmo ε para ambas as fórmulas i e ii. No entanto, a partir do momento em que x e y são sequências racionais quaisquer, é possível que δ' seja eventualmente diferente de δ''. Isso justifica o emprego dos rótulos δ' e δ''.

Uma vez que

$$|x_n + y_n - (L + M)| \leq |x_n - L| + |y_n - M|$$

(ver item (ii) das Observações logo após a Definição 4.22 na Seção 35), se δ for o maior valor entre δ' e δ'' (ou igual a ambos no caso em que $\delta' = \delta''$), então

$$\forall \varepsilon > 0\ (n > \delta \Rightarrow |(x_n + y_n) - (L + M)| \leq$$
$$|x_n - L| + |y_n - M| < \varepsilon + \varepsilon = 2\varepsilon).$$

Levando em conta que ε é arbitrário (desde que seja racional estritamente positivo), o fator 2 em 2ε é irrelevante. Ou seja, a existência de δ para o caso $x_n + y_n$ é garantida pela hipótese assumida no teorema de que δ' e δ'' existem.

Os teoremas a seguir são bastante úteis para a prova do Teorema 4.32, o qual se refere a limite de uma multiplicação entre funções reais.

Teorema 4.30. *Seja x uma sequência de racionais. Logo,*

$$x_n \to L \ \ sss \ \ (x_n - L) \to 0.$$

Demonstração: Basta usar a definição de sequência racional convergente e observar que

$$|x_n - L| = |(x_n - L) - 0|.$$

✍ Observar que o termo L, que ocorre em

$$(x_n - L) \to 0$$

no último teorema, é uma abreviação para a sequência constante $y_n = L$, enquanto o termo L que ocorre em

$$x_n \to L$$

é um número racional.

Logo, temos aqui mais um exemplo de notação abusiva. Um exercício que sempre se revela interessante é escrever formalmente, usando apenas o vocabulário de $\mathfrak{S}$, enunciados de teoremas que, na literatura, são escritos com abusos de linguagem.

Ou seja, como já foi discutido anteriormente, toda definição explícita abreviativa é matematicamente supérflua (eliminável).

Teorema 4.31. *Se $x_n \to L$, então $cx_n \to cL$.*

Demonstração: Se a constante c for 0, a prova é trivial, de acordo com o Teorema 4.26. Agora consideremos o caso em que $c \neq 0$. Temos, por hipótese,

$$\forall \varepsilon > 0 \, \exists \delta > 0 (n > \delta \Rightarrow |x_n - L| < \varepsilon).$$

Logo,

$$\forall \varepsilon > 0 \, \exists \delta > 0 (n > \delta \Rightarrow |c||x_n - L| < |c|\varepsilon).$$

Ou seja,

$$\forall\varepsilon > 0\ \exists\delta > 0(n > \delta \Rightarrow |cx_n - cL| < |c|\varepsilon).$$

Levando em conta que ε é arbitrário (desde que seja racional estritamente positivo), o fator $|c|$ é irrelevante.

O teorema a seguir estabelece que o limite do produto entre sequências racionais é o produto entre os limites das mesmas, caso estes existam.

Teorema 4.32. $(x_n \to L \wedge y_n \to M) \Rightarrow (x_n \cdot y_n \to L \cdot M)$.

Demonstração: Temos, por hipótese, a conjunção de duas fórmulas, a saber,

i: $\forall\varepsilon > 0\ \exists\delta' > 0(n > \delta' \Rightarrow |x_n - L| < \varepsilon)$ e

ii: $\forall\varepsilon > 0\ \exists\delta'' > 0(n > \delta'' \Rightarrow |y_n - M| < \varepsilon)$.

Uma vez que

$$|(x_n - L)(y_n - M) - 0| =$$
$$|(x_n - L)(y_n - M)| = |x_n - L| \cdot |y_n - M|,$$

se escolhermos δ como o maior valor entre δ' e δ'', então

$$\forall\varepsilon > 0(n > \delta \Rightarrow |(x_n - L)(y_n - M) - 0| < \varepsilon \cdot \varepsilon = \varepsilon^2.$$

Levando em conta que ε é arbitrário (desde que seja racional estritamente positivo), a condição $|(x_n - L)(y_n - M) - 0| < \varepsilon^2$ (desde que n seja maior do que δ) é equivalente a

$$(x_n - L)(y_n - M) \to 0.$$

No entanto,

$$x_n \cdot y_n = (x_n - L)(y_n - M) + Mx_n + Ly_n - LM.$$

Logo, usando Teoremas 4.26, 4.29, 4.30 e 4.31, temos que

$$\lim_{n\to\infty}(x_n \cdot y_n) =$$
$$\lim_{n\to\infty}((x_n - L)(y_n - M)) + \lim_{n\to\infty}(Mx_n) + \lim_{n\to\infty}(Ly_n) + \lim_{n\to\infty}(-LM) =$$
$$0 + ML + LM - LM = LM.$$

A transitividade da igualdade encerra a demonstração.

O próximo teorema trata do limite da diferença entre sequências racionais.

TEOREMA 4.33. $(x_n \to L \wedge y_n \to M) \Rightarrow (x_n - y_n \to L - M)$.

Finalmente,

TEOREMA 4.34.

$$(x_n \to L \wedge y_n \to M \wedge M \neq 0) \Rightarrow (x_n/y_n \to L/M).$$

✍ As provas dos dois últimos ficam como sugestões de exercícios ao leitor.

EXEMPLO 4.31. *Se* $x_n = 7$ *e* $y_n = \frac{1}{n}$, *então* $x_n \to 7$ *e* $y_n \to 0$; *logo,*

$$x_n + y_n \to 7$$

e

$$x_n \cdot y_n \to 0.$$

Como o leitor deve ter observado, Teoremas 4.26 (limite de sequência constante), 4.27 (limite de $\frac{1}{n}$), 4.29 (limite da soma), 4.32 (limite do produto), 4.33 (limite da diferença) e 4.34 (limite da razão) oferecem poderosas ferramentas para o efetivo cálculo de limites de sequências racionais. Obviamente a definição de sequência racional convergente não é 'amigável' para fins de cálculos, até porque tal definição não oferece explicitamente qualquer procedimento efetivo para determinar limites (caso existam). Neste momento deve ficar claro o papel altamente relevante de teoremas. Teoremas, neste caso, representam considerável *economia de pensamento*.

Uma possível crítica em relação aos teoremas até aqui provados é a seguinte: como garantir que não pode haver ambiguidade no cálculo de limite? Em particular, se o limite da sequência constante $x_n = c$ é a própria constante c (Teorema 4.26), como garantir que o limite não pode ser também um valor racional d diferente de c? Pois bem, o próximo teorema garante que jamais pode ocorrer tal ambiguidade para sequência alguma que admite limite.

TEOREMA 4.35. *O limite de uma sequência x de racionais, se existe, é único.*

DEMONSTRAÇÃO: Temos, por hipótese, que existe L racional tal que $x_n \to L$. Supor que existe $L' \neq L$ tal que L' é racional e $x_n \to L'$. Logo, basta escolher

$$\varepsilon = \frac{1}{2}|L - L'|,$$

ou seja, a metade da distância entre L e L'. Neste caso não existe racional x_n tal que

$$x_n \in (L-\varepsilon, L+\varepsilon) \wedge x_n \in (L'-\varepsilon, L'+\varepsilon).$$

Logo, não há δ tal que $n > \delta$ implique em $x_n \in (L-\varepsilon, L+\varepsilon)$ e $x_n \in (L'-\varepsilon, L'+\varepsilon)$. $\perp$

Ou seja, o último teorema foi demonstrado por redução ao absurdo. Esta é uma técnica muito usual em teoremas de unicidade, como o caso do Teorema 3.2.

Um conceito relacionado ao de sequência racional convergente é o de *sequência racional de Cauchy*. A ideia intuitiva é a seguinte: uma sequência racional x_p de Cauchy é aquela em que imagens x_m ficam confinadas ao intervalo $(x_n - \varepsilon, x_n + \varepsilon)$ e imagens x_n ficam confinadas ao intervalo $(x_m - \varepsilon, x_m + \varepsilon)$ na medida em que ambos m e n se tornam arbitrariamente grandes. Em outras palavras, x_p é de Cauchy sss suas imagens x_m e x_n 'se aproximam cada vez mais umas das outras', na medida em que se aumentam os valores de m e n. Neste sentido, uma sequência racional de Cauchy não é necessariamente convergente.

Definição 4.25. *x_p é de Cauchy sss*

$$\forall\varepsilon > 0\ \exists\delta > 0((m > \delta \wedge n > \delta) \Rightarrow |x_m - x_n| < \varepsilon).$$

Exemplo 4.32. *Seja x a sequência dada por*

$$x_0 = 2$$

e

$$x_{n+1} = \left(x_n + \frac{2}{x_n}\right)/2.$$

Logo,

$$x_1 = \frac{3}{2},\ x_2 = \frac{17}{12},\ x_3 = \frac{577}{408},\ x_4 = \frac{665857}{470832}, \cdots.$$

A sequência x do exemplo acima é definida recursivamente, no seguinte sentido:

I: x_0 é igual a 2;

ii: sabendo que $x_{n+1} = (x_n + \frac{2}{x_n})/2$, logo, $x_1 = (x_0 + \frac{2}{x_0})/2$;

iii: logo, $x_2 = (x_1 + \frac{2}{x_1})/2$; e assim por diante.

Observar que, no Exemplo acima, as distâncias entre as imagens x_m e x_n se tornam cada vez menores, na medida em que m e n aumentam.

Exemplo 4.33.

$$|x_0 - x_1| = \frac{1}{2},\ |x_1 - x_2| = \frac{1}{12},$$
$$|x_2 - x_3| = \frac{1}{408},\ |x_3 - x_4| = \frac{1}{470832}, \cdots .$$

É possível provar que a sequência acima é de Cauchy. Não fazemos tal demonstração neste livro, a qual pode ser feita por indução infinita mas é bastante árdua. No entanto, é fácil provar que a sequência dada no último Exemplo não é convergente. A partir de sua definição, observar que

$$2x_{n+1} = x_n + \frac{2}{x_n}.$$

Supor que existe racional L tal que $L = \lim_{n\to\infty} x_n$. Logo,

$$\lim_{n\to\infty}(2x_{n+1}) = \lim_{n\to\infty}(x_n + \frac{2}{x_n}),$$

por conta do Teorema 4.35. Além disso, $\lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} x_n$, se existir $\lim_{n\to\infty} x_n$. Afinal, $n > \delta \Rightarrow n + 1 > \delta$. Logo,

$$\lim_{n\to\infty} 2 \cdot \lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} x_n + \lim_{n\to\infty} \frac{2}{x_n},$$

por conta dos Teoremas 4.29 e 4.32. Isso implica em

$$2L = L + \frac{2}{L}.$$

Finalmente,

$$L = \frac{2}{L},$$

o que implica em

$$L^2 = 2.$$

Porém, não existe racional L tal que $L^2 = 2$. Com efeito, se L for racional, então

$$L = \frac{p}{q},$$

sendo p e q inteiros e $q \neq 0$. Logo, a substitutividade da igualdade garante que

$$\frac{p^2}{q^2} = 2,$$

o que implica em

$$p^2 = 2q^2.$$

Mas, uma vez que p e q são inteiros, então p e q são primos ou compostos (ou 1) (ver Definição 4.2). Logo, p^2 conta com uma quantia par de fatores primos, enquanto $2q^2$ conta com uma quantia ímpar de fatores primos (lembrar que 2 é primo). Isso é uma contradição com o Teorema Fundamental da Aritmética (o qual estabelece que qualquer fatoração de um natural em primos é única, a menos de arranjos dos fatores)! Consequentemente, L não é racional. Logo,

$$\nexists L \left(L = \lim_{n \to \infty} x_n \right).$$

Apesar de sequências de Cauchy não serem necessariamente convergentes, o fato é que toda sequência racional convergente é de Cauchy, como se mostra a seguir.

Teorema 4.36. *Se x é uma sequência racional convergente, então é de Cauchy.*

Demonstração: Temos, por hipótese, que existe racional L tal que $x_n \to L$. Logo,

$$\forall \varepsilon > 0 \; \exists \delta > 0 (n > \delta \Rightarrow |x_n - L| < \varepsilon).$$

Logo,

$$\forall \varepsilon > 0 \; \exists \delta > 0 (m > \delta \Rightarrow |x_m - L| < \varepsilon).$$

Uma vez que

$$|x_m - x_n| = |(x_m - L) - (x_n - L)| \leq$$
$$|x_m - L| + |x_n - L|,$$

logo,

$$\forall \varepsilon > 0 \; \exists \delta > 0 ((m > \delta \wedge n > \delta) \Rightarrow |x_m - x_n| < 2\varepsilon).$$

Sequências de racionais constituem uma ótima ferramenta para definir números reais a partir de racionais, como se vê na próxima Parte. Observar que isso não conflita com a discussão na Seção 34, uma vez que a proposta não é o emprego de pares ordenados de racionais, mas algo muito mais rico: sequências racionais de Cauchy. Isso ajuda a ilustrar o enorme poder de funções.

Seção 37

Resumo da ópera

Sumário
Índice
Rede

Esta quarta parte pode ser resumida como se segue.

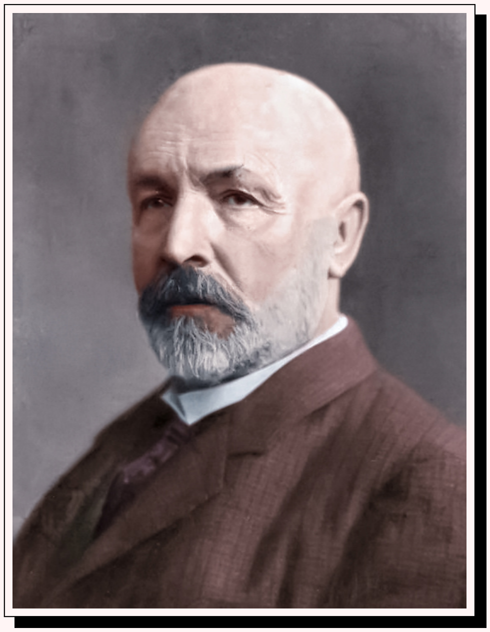
- ZF permite conceituar, no contexto de sua linguagem, números naturais, inteiros e racionais.
- A lógica de ZF permite conhecer diversas propriedades algébricas das operações de adição e multiplicação entre naturais, inteiros e racionais.
- Naturais e suas operações usuais são definidos a partir de um conjunto indutivo em particular, denotado por ω.
- Inteiros são classes de equivalência de pares ordenados de naturais, enquanto racionais são classes de equivalência de pares ordenados de inteiros.
- O que diferencia naturais de inteiros e racionais são as propriedades algébricas das operações de adição e multiplicação. Propriedades algébricas da adição entre naturais são preservadas entre os inteiros. Mas os últimos contam com a existência de simétrico aditivo, algo que não ocorre entre naturais. As propriedades algébricas de adição e multiplicação entre inteiros são preservadas entre os racionais. Mas os últimos contam com a existência de simétrico multiplicativo (exceto para o neutro aditivo), algo que não acontece entre inteiros ou naturais.
- Uma vez que reais não podem ser definidos como classes de equivalência de pares ordenados de racionais, alguns conceitos são desenvolvidos na linguagem de ZF para contornar essa dificuldade. Entre esses conceitos, sequências racionais de Cauchy são de interesse estratégico a ser explorado na próxima Parte.

Seção 38

Notas históricas

Sumário
Índice
Rede

Durante a transição do século 19 para o século 20 houve extensas discussões sobre as ideias originais de Cantor. Leopold Kronecker chegou a dizer que o *infinito* (na acepção da Definição 4.18) de seu ex-aluno Georg Cantor era filosofia ou religião, mas não matemática.



Georg Cantor, no início do século 20

Fonte: Wikipedia.

Afirmando que Cantor era um corruptor das novas gerações de matemáticos, Kronecker exerceu severas interferências na carreira de seu ex-aluno, impedindo-o de se tornar professor na Universidade de Berlim. No entanto, esse tipo de resistência não era novidade. Na Grécia Antiga, por exemplo, os números irracionais eram aqueles sobre os quais nada se falava. Por isso o nome! Irracionais eram números 'ilógicos'. Preconceito é uma inevitável condição humana, mesmo quando o assunto é matemática.

PARTE 5

Números reais e complexos

Nesta quinta parte finalmente iniciamos os primeiros passos na direção de cálculo diferencial e integral padrão.

Seção 39 — Reais

Sumário
Índice
Rede

Definimos números reais como certas classes de equivalência de sequências de Cauchy de racionais. Antes, porém, precisamos introduzir uma nova relação.

Definição 5.1. *Sejam x_n e y_n sequências de racionais. Logo,*

$$x_n \equiv y_n \vdots (x_n - y_n) \to 0.$$

Lemos $x_n \equiv y_n$ como 'x_n é equivalente *a y_n'.*

Seja ω uma notação abusiva para a cópia do conjunto dos números naturais em $\mathbb{Q}$. Se

$$r = \{t \in \wp(\omega \times \mathbb{Q}) \mid t \text{ é sequência}\}$$

é o conjunto de todas as sequências racionais, então $\equiv$ define uma relação em r.

EXEMPLO 5.1. I: *$x_n = \frac{1}{n}$ e $y_n = \frac{2}{n^2}$ são equivalentes, ou seja,*

$$\frac{1}{n} \equiv \frac{2}{n^2}.$$

Com efeito,

$$\left(\frac{1}{n} - \frac{2}{n^2}\right) \to 0.$$

II: *$x_n = \frac{1}{n}$ e $z_n = 5$ não são equivalentes, uma vez que*

$$\left(\frac{1}{n} - 5\right) \to -5$$

e $-5 \neq 0$.

III: *Seja v uma sequência racional tal que $v_n = (-1)^n$. Seja w uma sequência racional tal que*

$$w_n = \begin{cases} n^2 & \textit{se } n \leq 10 \\ (-1)^n & \textit{se } n > 10. \end{cases}$$

Logo, para todo $n > 10$ temos $v_n = w_n$. Isso implica que

$$(v_n - w_n) \to 0.$$

Portanto, $v \equiv w$.

✍ É obviamente recomendável que o leitor prove o item III acima.

TEOREMA 5.1. *A relação $\equiv$ da Definição 5.1 é de equivalência.*

DEMONSTRAÇÃO: Uma vez que $x_n - x_n = 0$, de acordo com Teorema 4.26,

$$(x_n - x_n) \to 0.$$

Logo, $x_n \equiv x_n$, o que prova que $\equiv$ é reflexiva.

Se $x_n \equiv y_n$, então $(x_n - y_n) \to 0$ (Definição 5.1). Mas

$$y_n - x_n = (-1)(x_n - y_n).$$

Logo,

$$(y_n - x_n) \to (-1) \lim_{n\to\infty} (x_n - y_n),$$

de acordo com os Teoremas 4.26 e 4.32. Logo, $(y_n - x_n) \to 0$, uma vez que o racional 0 é absorvente multiplicativo. Logo, $y_n \equiv x_n$, o que prova a simetria de $\equiv$.

Se $x_n \equiv y_n$ e $y_n \equiv z_n$, então $(x_n - y_n) \to 0$ e $(y_n - z_n) \to 0$.

Logo,
$$((x_n - y_n) + (y_n - z_n)) \to 0,$$
de acordo com o Teorema 4.29. Mas,
$$((x_n - y_n) + (y_n - z_n)) = (x_n - z_n).$$
Logo, a substitutividade da igualdade garante que
$$(x_n - z_n) \to 0.$$

Isso implica em $x_n \equiv z_n$, o que prova a transitividade de $\equiv$.

Consequentemente, a relação $\equiv$ definida sobre o conjunto r das sequências racionais é de equivalência.

Se r é o conjunto das sequências racionais e c é o conjunto das sequências racionais de Cauchy, então $c \subset r$ e novamente $\equiv$ define uma relação de equivalência, desta vez sobre c. Notar que item III do Exemplo 5.1 prova que, de fato, c é subconjunto próprio de r. Com efeito, as sequências v_n e w_n daquele item não são de Cauchy, apesar de serem sequências racionais equivalentes entre si.

Teorema 5.1, em parceria com Teorema 3.10, permite finalmente definir *números reais*, bem como reais *racionais* e reais *irracionais*.

Definição 5.2. *Seja c o conjunto das sequências racionais de Cauchy. Logo,*
$$\mathbb{R} = c/\equiv$$
é o conjunto dos números reais.

Cada elemento de $c/\equiv$ é chamado de número real. *Se qualquer representante x_n de $[x_n]$ (onde $[x_n]$ é uma classe de equivalência pertencente a $c/\equiv$) é uma sequência de Cauchy convergente, então $[x_n]$ é um número real* racional. *Caso contrário, $[x_n]$ é um número real* irracional.

Lembrar que $c/\equiv$ é o quociente do conjunto das sequências racionais de Cauchy pela relação de equivalência $\equiv$ (ver parágrafo imediatamente após a demonstração do Teorema 3.11).

O conjunto $\mathbb{R}$ dos números reais é também conhecido como o *corpo dos números reais*. Existem outros *corpos* além de $\mathbb{R}$. Detalhes na Seção 96.

✍ Em $\mathbb{R}$ toda sequência de Cauchy é convergente. Recomendamos que o leitor prove isso.

Se os reais r e s têm, respectivamente, representantes x_n e y_n, então $r+s$ (*adição* entre reais) é um real com representante x_n+y_n, e $r\cdot s$ (ou, simplesmente, rs, a *multiplicação* entre reais) é um real com representante $x_n\cdot y_n$.

Exemplo 5.2. i: *Seja x a sequência racional dada por*

$$x_0=2 \text{ e } x_{n+1}=\left(x_n+\frac{2}{x_n}\right)/2;$$

logo,

$$x_1=\frac{3}{2},\ x_2=\frac{17}{12},\ x_3=\frac{577}{408},\ x_4=\frac{665857}{470832},\ \cdots .$$

Este é o mesmo Exemplo 4.32*, apresentado na Seção 36. Neste caso, x é de Cauchy, mas não convergente entre os racionais (como já discutido). Isso significa que x é representante de um real $r=[x_n]$ irracional, a saber, um real r tal que $r^2=2$. Este número real é usualmente denotado por $\sqrt{2}$. Para que o leitor tenha uma ideia melhor sobre os demais representantes de $\sqrt{2}$, ver o próximo item.*

ii: *Seja y a sequência racional dada por*

$$y_0=5 \text{ e } y_{n+1}=\left(y_n+\frac{2}{y_n}\right)/2;$$

logo,

$$y_1=\frac{27}{10},\ y_2=\frac{929}{540},\ y_3=\frac{1446241}{1003320},\ \cdots .$$

Neste caso, $x_n\equiv y_n$, apesar de $x_n\neq y_n$. Observar que, em notação decimal, $x_1-y_1=1,2$, $x_2-y_2=0,303$, $x_3-y_3=0,0272$, $\cdots$.

Ambas as sequências x e y são de Cauchy, porém não convergentes. Afinal, analogamente à discussão na Seção 36, se x ou y convergissem, deveriam convergir para um racional L tal que $L^2=2$, o que não pode ser o caso.

No entanto, $[x_n]$, a qual é igual a $[y_n]$, é o número real r tal que $r^2=2$, ou seja, $\sqrt{2}$. ✍ *Outros exemplos de representantes de $\sqrt{2}$ podem ser dados pelo leitor.*

Se $x_n + y_n = z_n$, dizemos que z_n é a *soma* das *parcelas* x_n e y_n. Se $x_n \cdot y_n = z_n$, dizemos que z_n é o *produto* dos *fatores* x_n e y_n. O mesmo se diz sobre os respectivos reais com representantes x_n, y_n e z_n.

Exemplo 5.3. $\sqrt{2}(\sqrt{2}) = 2$.

Se r é um real racional, empregamos a mesma notação introduzida para racionais.

Obviamente, as propriedades algébricas de adição e multiplicação entre racionais induzem as mesmas propriedades para a adição $+$ e a multiplicação $\cdot$ entre números reais. Logo, são teoremas as seguintes fórmulas:

I: a adição entre reais é comutativa e associativa;

II: a adição entre reais admite neutro aditivo (denotado por 0) e simétrico aditivo para qualquer real r (denotado por $-r$);

III: a multiplicação entre reais é comutativa e associativa;

IV: a multiplicação entre reais admite neutro multiplicativo (denotado por 1) e simétrico multiplicativo para qualquer real r diferente de 0 (denotado por r^{-1});

V: o neutro aditivo é absorvente multiplicativo;

VI: a multiplicação é distributiva em relação à adição.

A nova propriedade algébrica entre números reais, inexistente entre racionais, é o fato de que sequências e Cauchy e sequências convergentes são conceitos equivalentes em $\mathbb{R}$.

Entre os números reais há uma relação de ordem total $\leq$:

$r < 0 \;\vdots\;$ para todo representante x_n de r há δ tal que $n > \delta \Rightarrow x_n < 0$, sendo a última a relação de ordem $<$ entre racionais.

$r < s \;\vdots\; r - s < 0$ (lembrar que $r - s = r + (-s)$, sendo que $-s$ é o simétrico aditivo de s)

$r \leq s \;\vdots\; r < s \vee r = s$.

$s \geq r \;\vdots\; r \leq s$

$s > r \;\vdots\; r < s$

Graças às relações de ordem $\leq$ e $<$ em $\mathbb{R}$, é possível introduzir conceitos muito úteis para os estudos da Seção 44:

- Um *intervalo aberto* (a, b) é o conjunto
$$\{x \in \mathbb{R} \mid a < x \wedge x < b\};$$
- Um *intervalo fechado* $[a, b]$ é o conjunto
$$\{x \in \mathbb{R} \mid a \leq x \wedge x \leq b\};$$
- Um *intervalo fechado degenerado* $[a, b]$ é um intervalo fechado tal que $a = b$;
- Um *intervalo fechado não degenerado* é um intervalo fechado que não é degenerado;
- Um *intervalo aberto à esquerda e fechado à direita* $(a, b]$ é o conjunto
$$\{x \in \mathbb{R} \mid a < x \wedge x \leq b\};$$
- Um *intervalo fechado à esquerda e aberto à direita* $[a, b)$ é o conjunto
$$\{x \in \mathbb{R} \mid a \leq x \wedge x < b\};$$
- Uma *vizinhança* de um número real r é qualquer intervalo aberto (a, b) tal que $r \in (a, b)$.

EXEMPLO 5.4. I: *$(3, 8)$ é uma vizinhança de 5, mas não de 3;*

II: ✍ *todo número real r admite uma vizinhança (a, b) (consegue provar isso?).*

✍

I: Exibir a classe de equivalência de sequências de Cauchy de racionais correspondente ao real $\sqrt{5}$;

II: provar que nenhum representante x_n de $\sqrt{5}$ é convergente em $\mathbb{Q}$;

III: provar que, para quaisquer reais a, b e c tais que $a < b$ e $b < c$, temos que $(a, b) \cap (b, c) = \varnothing$. Este último é de importância estratégica para a compreensão de limites de funções reais, a serem discutidos na Seção 44.

Análise na reta é o estudo da tripla ordenada $(\mathbb{R}, +, \cdot)$, incluindo funções com domínios e co-domínios contidos em $\mathbb{R}$.

SEÇÃO 40

Complexos

SUMÁRIO
ÍNDICE
REDE

Lembrando que $\mathbb{R}^2$ é o conjunto $\mathbb{R} \times \mathbb{R}$ dos pares ordenados de números reais, podemos agora introduzir o que são complexos.

DEFINIÇÃO 5.3. *O* corpo $\mathbb{C}$ *dos* números complexos *é o conjunto* $\mathbb{C} = (\mathbb{R}^2, +, \cdot)$*, onde*

- $+ : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}^2$ *é a função dada por*
$$+((a,b),(c,d)) = (a,b) + (c,d) = (a+c, b+d) \quad e$$
- $\cdot : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}^2$ *é a função dada por*
$$\cdot((a,b),(c,d)) = (a,b) \cdot (c,d) = (ac - bd, ad + bc).$$

Cada $(a,b) \in \mathbb{R}^2$ é um *número complexo*. A função $+$ é chamada de *adição* de complexos, enquanto $\cdot$ é a *multiplicação* de complexos.

Se $(m,n) + (p,q) = (r,s)$, dizemos que (r,s) é a *soma* das *parcelas* (m,n) e (p,q).

Se $(m,n) \cdot (p,q) = (r,s)$, dizemos que (r,s) é o *produto* dos *fatores* (m,n) e (p,q).

EXEMPLO 5.5. *A adição entre o complexo* $(5,-2)$ *e o complexo* $(\sqrt{3}, 0)$ *é o complexo* $(5,-2) + (\sqrt{3}, 0) = (5+\sqrt{3}, -2)$.

A multiplicação entre o complexo $(5,-2)$ *e o complexo* $(\sqrt{3},0)$ *é o complexo* $(5,-2) \cdot (\sqrt{3}, 0) = (5\sqrt{3} - (-2)0, 5(0) + (-2)\sqrt{3})$. *Ou seja,* $(5,-2) \cdot (\sqrt{3}, 0) = (5\sqrt{3}, -2\sqrt{3})$.

TEOREMA 5.2. *A adição entre complexos é comutativa. Formalmente, isso se traduz como*
$$(a,b) + (c,d) = (c,d) + (a,b),$$
onde (a,b) *e* (c,d) *são complexos.*

Demonstração:

$$(a,b)+(c,d)=(a+c,b+d)=$$
$$(c+a,d+b)=(c,d)+(a,b).$$

Teorema 5.3. *A multiplicação entre complexos é comutativa. Formalmente, isso se traduz como*

$$(a,b)\cdot(c,d)=(c,d)\cdot(a,b),$$

onde (a,b) e (c,d) são complexos.

Demonstração:

$$(a,b)\cdot(c,d)=(ac-bd,ad+bc)=$$
$$(ca-db,da+cb)=(c,d)\cdot(a,b).$$

Teorema 5.4. *Existe neutro multiplicativo entre os complexos. Ademais, ele é único. Formalmente, isso se traduz como*

$$\exists! c\ \exists! d((c,d)\in\mathbb{R}^2\wedge\forall a\forall b((a,b)\in\mathbb{R}^2\Rightarrow(a,b)\cdot(c,d)=(a,b))).$$

Demonstração: Basta fazer $(c,d)=(1,0)$. Com efeito,

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

Ou seja, $(1,0)$ é neutro multiplicativo.

Para provar a unicidade do neutro multiplicativo, supor que existe outro. ✍ Cabe ao leitor verificar que essa hipótese produz uma contradição. Portanto, o par ordenado (c,d) mencionado é apenas $(1,0)$.

Outra maneira para demonstrar o último teorema é a seguinte. Uma vez que $(a,b)\cdot(c,d)=(ac-bd,ad+bc)$, basta provar que

$$(ac-bd,ad+bc)=(a,b)\text{ sss }c=1\wedge d=0.$$

Graças à comutatividade da multiplicação, $(1,0)\cdot(c,d)=(c,d)$.

Teorema 5.5. *Existe neutro aditivo entre os complexos. Além disso, ele é único. Formalmente, isso se traduz como*

$$\exists! c\ \exists! d((c,d)\in\mathbb{R}^2\wedge\forall a\forall b((a,b)\in\mathbb{R}^2\Rightarrow(a,b)+(c,d)=(a,b))).$$

Demonstração:

$$(a, b) + (0, 0) = (a + 0, b + 0) = (a, b).$$

Ou seja, $(0, 0)$ é neutro aditivo. Para provar a unicidade do neutro aditivo, supor que existe outro. ✍ Cabe ao leitor verificar que essa hipótese produz uma contradição. Portanto, o par ordenado (c, d) mencionado é $(0, 0)$.

Teorema 5.6. *Todo complexo admite simétrico aditivo. Formalmente, isso se traduz como*

$$\forall a \forall b((a, b) \in \mathbb{R}^2 \Rightarrow \exists c\, \exists d((c, d) \in \mathbb{R}^2 \wedge (a, b) + (c, d) = (0, 0))),$$

sendo $(0, 0)$ *o neutro aditivo do Teorema 5.5.*

Demonstração:

$$(a, b) + (-a, -b) = (a + (-a), b + (-b)) = (0, 0).$$

Logo, $(-a, -b)$ é simétrico aditivo de (a, b), onde $-a$ e $-b$ são os simétricos aditivos dos reais a e b, respectivamente. Portanto, o par ordenado (c, d) mencionado é $(-a, -b)$. Notar que, para cada (a, b) complexo, $(-a, -b)$ é único.

Em particular, $(-1, 0)$ é o simétrico aditivo do neutro multiplicativo entre os complexos. Essa informação se revela particularmente relevante para discernirmos complexos de reais.

✍ É teorema a seguinte fórmula: a multiplicação entre complexos é associativa. Recomendamos que o leitor prove isso. Esse fato facilita bastante o cálculo dado pela seguinte definição.

Definição 5.4. *Se* (a, b) *é um complexo diferente do neutro aditivo e* n *é um natural, então*

i: $(a, b)^0 = (1, 0)$;

ii: $(a, b)^{n+1} = (a, b) \cdot (a, b)^n$.

Exemplo 5.6. $(0, 1)^3 = (0, 1) \cdot (0, 1)^2 = (0, 1) \cdot (0, 1) \cdot (0, 1)^1 = (0, 1) \cdot (0, 1) \cdot (0, 1) \cdot (0, 1)^0 = (0, 1) \cdot (0, 1) \cdot (0, 1) \cdot (1, 0) = (0, -1)$; *o leitor escolhe a ordem em que deseja realizar as operações de multiplicação, uma vez que multiplicação entre complexos é comutativa e associativa.*

Lemos $(a,b)^n$ como '(a,b) *elevado a* n'. Em particular, $(a,b)^2$ se lê também como '(a,b) *ao quadrado*' e $(a,b)^3$ se lê também como '(a,b) *ao cubo*'.

Entre os complexos existe uma propriedade algébrica que não ocorre entre os reais, os racionais, os inteiros ou os naturais, conforme o próximo teorema.

Teorema 5.7. *Existe um complexo cujo quadrado é o simétrico aditivo do neutro multiplicativo.*

Demonstração: $(0,1) \cdot (0,1) = (0.0 - 1.1, 0.1 + 1.0) = (-1,0)$.

O simétrico aditivo do neutro multiplicativo entre os reais é -1. No entanto, não existe real r tal que $r^2 = -1$, sendo $r^2 = r \cdot r$. Comentário análogo vale para os racionais e os inteiros. Entre os naturais, em particular, o simétrico aditivo do neutro multiplicativo sequer existe.

O complexo $(0,1)$ cujo quadrado $(0,1)^2$ é o simétrico aditivo do neutro multiplicativo $(1,0)$ (ou seja, $(-1,0)$) é conhecido como *unidade imaginária*. Comumente abrevia-se $(0,1)$ pelo símbolo i. Ou seja,

$$i = (0,1).$$

Teorema 5.8. *Os complexos da forma* $(e,0)$ *copiam os números reais.*

Demonstração: Basta observar que

$$(a,0) \cdot (c,0) = (ac - 0.0, a.0 + 0.c) = (ac, 0)$$

e

$$(a,0) + (c,0) = (a + c, 0 + 0) = (a + c, 0).$$

Logo, adição $a + c$ entre reais é copiada por $(a,0) + (c,0)$. Resultado análogo vale para multiplicação.

Este último teorema justifica a prática comum de abreviar complexos $(a,0)$ como a. Neste contexto, se $z = (a,b)$ é um complexo qualquer, então

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

(basta fazer as contas para confirmar). Abreviadamente, isso corresponde a afirmar que

$$z = a + bi,$$

onde a e b são complexos que copiam reais (lembrar que $a \cdot 1 = a$) e i é a unidade imaginária.

Logo,

$$(a + bi) + (c + di) = (a + c) + (b + d)i$$

e

$$(a + bi)(c + di) = (ac - bd) + (ad + bc)i,$$

de acordo com a Definição 5.3.

Se $z = (a, b)$ é um complexo, chamamos a de *parte real* de z, e b de *parte imaginária* do complexo z. Essa convenção é consistente com o fato de que complexos $(a, 0)$ copiam os reais, enquanto complexos $(0, b)$ contam com uma propriedade algébrica não replicável pelos reais, por consequência do Teorema 5.7.

Uma vez que complexos são definidos como pares ordenados (ver Teorema 3.4) de reais, um complexo z é igual a um complexo z' sss a parte real de z for igual à parte real de z' e a parte imaginária de z for igual à parte imaginária de z'. Obviamente essa última afirmação é um teorema.

A partir de agora adotamos a notação abreviada a para complexos $(a, 0)$ e i para a unidade imaginária $(0, 1)$. Logo, bi abrevia o complexo $(0, b)$, enquanto $a + bi$ abrevia (a, b). Neste contexto, são teoremas as seguintes fórmulas (lembrar que multiplicação entre complexos é associativa):

$$i^0 = 1,\ i^1 = i,\ i^2 = -1,\ i^3 = -i,\ i^4 = 1,\ i^5 = i \text{ e assim por diante.}$$

Ou seja,

$$i^{4n} = 1, \quad i^{4n+1} = i, \quad i^{4n+2} = -1, \quad i^{4n+3} = -i,$$

onde n é um natural.

Esses resultados são usados na Seção 57.

Observar que, apesar dos complexos estenderem os reais em termos das operações algébricas de adição e multiplicação, eles não fazem o mesmo para a relação de ordem total $\leq$ entre reais. Com efeito, se r e s são reais tais que $r \neq 0$ ou $s \neq 0$, então $r^2 + s^2 > 0$.

No entanto, entre os complexos isso não é teorema. Por exemplo, $i^2 + 1^2 = 0$.

Pior ainda, $(2i)^2 + 1^2 = -3 < 0$.

Por conta disso, entre os complexos não é possível definir uma relação de ordem total $\leq$ que seja compatível com as operações de adição e multiplicação entre complexos, e que ainda seja uma cópia da relação usual $\leq$ entre reais.

Copiar a relação de ordem total $\leq$ dos reais entre complexos que copiam os reais é algo trivial:

$$(a, 0) \leq (b, 0) \text{ sss } a \leq b.$$

O problema sem solução é estender essa relação para todos os complexos.

Análise complexa é o estudo da tripla ordenada $(\mathbb{C}, +, \cdot)$ e das funções com domínio e co-domínio contidos em $\mathbb{C}$.

Seção 41

$\omega \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$?

Sumário
Índice
Rede

Um discurso usual na literatura diz que todo natural é um inteiro, todo inteiro é um racional, todo racional é um real e todo real é um complexo. Usualmente isso se traduz como

$$\omega \subset \mathbb{Z} \wedge \mathbb{Z} \subset \mathbb{Q} \wedge \mathbb{Q} \subset \mathbb{R} \wedge \mathbb{R} \subset \mathbb{C}.$$

No entanto, obviamente não é o caso aqui. O zero natural é o conjunto vazio, enquanto o zero inteiro é uma classe de equivalência de pares ordenados de naturais. Dadas as construções aqui exibidas, nenhum natural é inteiro, nenhum inteiro é racional, nenhum racional é real e nenhum real é complexo. Ou seja, a fórmula

$$\omega \not\subset \mathbb{Z} \wedge \mathbb{Z} \not\subset \mathbb{Q} \wedge \mathbb{Q} \not\subset \mathbb{R} \wedge \mathbb{R} \not\subset \mathbb{C}$$

é teorema.

Por outro lado, vimos que inteiros positivos copiam naturais; racionais $\frac{p}{q}$, tais que $q = 1$, copiam os inteiros; reais cujos representantes são sequências de Cauchy convergentes copiam os racionais; e complexos $(a, 0)$ copiam os reais (incluindo a ordem total $\leq$ entre reais).

Isso significa que complexos podem ser usados para copiar também racionais, inteiros e naturais. Logo, o que temos é o seguinte:

$$c(\omega) \subset c(\mathbb{Z}) \wedge c(\mathbb{Z}) \subset c(\mathbb{Q}) \wedge c(\mathbb{Q}) \subset c(\mathbb{R}) \wedge c(\mathbb{R}) \subset \mathbb{C},$$

onde $c(\omega)$, $c(\mathbb{Z})$, $c(\mathbb{Q})$ e $c(\mathbb{R})$ são, respectivamente, cópia dos naturais entre os complexos, cópia dos inteiros entre os complexos, cópia dos racionais entre os complexos e cópia dos reais entre os complexos. Apenas por abuso de notação que se afirma que $\omega \subset \mathbb{Z} \wedge \mathbb{Z} \subset \mathbb{Q} \wedge \mathbb{Q} \subset \mathbb{R} \wedge \mathbb{R} \subset \mathbb{C}$.

Na literatura especializada há muitos outros conjuntos numéricos, como os *quatérnions*, os *hiperreais*, os *hipercomplexos*, os *surreais*, os *perplexos*, os *transfinitos*, entre outros. As relações entre esses conjuntos não são óbvias. Por exemplo, a multiplicação entre quatérnions é não comutativa.

Seção 42

Funções reais

Sumário
Índice
Rede

Por enquanto voltamos a discutir sobre números reais, deixando os complexos de lado. Mais adiante fica evidente que o conhecimento sobre certas funções reais $-$ aquelas cujas imagens são apenas números reais $-$ depende de considerações sobre os complexos.

A definição recursiva a seguir é usual.

Definição 5.5. *Seja x um número real diferente do neutro aditivo. Logo,*

i: $x^0 = 1$*;*

ii: $x^{n+1} = x \cdot x^n$*, onde n é um real que copia um inteiro positivo.*

Exemplo 5.7.

$x^4 = x \cdot x^3 = x \cdot x \cdot x^2 = x \cdot x \cdot x \cdot x^1 = x \cdot x \cdot x \cdot x \cdot x^0 = x \cdot x \cdot x \cdot x$*;*

uma vez que multiplicação entre reais é associativa, não há necessidade de qualquer preocupação com parênteses.

O termo $\sqrt{x}$ (lê-se 'raiz quadrada de x') é uma abreviação para um real y não negativo tal que $y^2 = x$. Ou seja,

$$y = \sqrt{x} \vee y = -\sqrt{x} \text{ sss } y^2 = x.$$

Obviamente isso somente pode ser o caso se $x \geq 0$. Se $y^2 = x$, onde x é positivo, então $y = \sqrt{x}$ ou $y = -\sqrt{x}$, o que se abrevia como $y = \pm\sqrt{x}$.

É uma convenção adotar que $\sqrt{x} \geq 0$ se $x \geq 0$. Logo, $-\sqrt{x}$ é o simétrico aditivo de $\sqrt{x}$. Se $\sqrt{x}$ for estritamente positivo, então $-\sqrt{x}$ é um real negativo.

O termo $\sqrt[n]{x}$ (lê-se 'raiz n-ésima de x') é uma abreviação para um real y tal que $y^n = x$, onde $n \geq 2$ é um real que copia um natural. Se n for ímpar, então $\sqrt[n]{x}$ é definido para qualquer real x; se n for par, então $\sqrt[n]{x}$ está definido apenas para os reais x positivos. No caso particular $\sqrt[3]{x}$, lê-se 'raiz cúbica de x'.

O principal propósito, de agora em diante, é o estudo de funções

$$f : a \to b$$

tais que ambos a e b são subconjuntos de $\mathbb{R}$.

Qualquer função cujo co-domínio é subconjunto de $\mathbb{R}$ é dita uma *função real*. A Definição 5.5 para x^n e seu correspondente $\sqrt[n]{x}$ são úteis para o estudo de muitas funções reais.

Exemplo 5.8. i: *$f : \mathbb{R} \to \mathbb{R}$ tal que $f(x) = x$ (função identidade); observar que*

$$f = \{(x, y) \in \mathbb{R} \times \mathbb{R} \mid y = x\},$$

ou seja, $f = \{(x, x) \mid x \in \mathbb{R}\}$;

ii: *Sejam c um número real e $g : \mathbb{R} \to \mathbb{R}$ tal que $g(x) = c$ (função* constante*); observar que*

$$g = \{(x, c) \in \mathbb{R}^2 \mid x \in \mathbb{R}\};$$

iii: *Seja $h : d \to \mathbb{R}$ tal que $d \subseteq \mathbb{R}$ e*

$$h(x) = a_n x^n + a_{n-1}x^{n-1} + a_{n-2}x^{n-2} + \cdots + a_2x^2 + a_1x + a_0,$$

onde $a_0, a_1, \cdots, a_n$ são números reais e n é uma cópia de um número natural entre os reais. Função h é conhecida como função polinomial *de grau menor ou igual a n; o grau dessa função polinomial h é n se $a_n \neq 0$;*

iv: *A função identidade é uma função polinomial de grau* 1*;*

v: $j : \mathbb{R} \to \mathbb{R}$ *tal que*

$$j(x) = 4x - 2x^4 + \sqrt{2}$$

é uma função polinomial de grau 4*; neste caso* $a_0 = \sqrt{2}$, $a_1 = 4$, $a_2 = 0$, $a_3 = 0$ *e* $a_4 = -2$.

vi: $k : \mathbb{R} \to \mathbb{R}$ *tal que* $k(x) = c$*, onde* $\{c\}$ *é o conjunto escolha obtido por aplicação do Axioma da Escolha sobre o conjunto unitário* $\{\mathbb{R}\}$*. Essa função* k *é conhecida como* função escolha *de* $\mathbb{R}$*. Uma vez que o Axioma da Escolha foi usado apenas sobre um singleton, a função escolha aqui ilustrada é uma função real constante. Obviamente não sabemos qual foi a constante* c *'escolhida'.*

Na Seção 102 é usada uma função escolha para ilustrar exemplo de evento que, apesar de ter probabilidade zero, ocorre.

Funções polinomais são extremamente versáteis para expressar até mesmo funções não polinomiais, conforme se percebe a partir da Seção 54. Portanto, é de grande interesse conhecê-las.

Na próxima Seção há uma breve discussão sobre os *zeros* de funções polinomiais. Naturalmente, o assunto não é esgotado apenas com isso. Mas já é um começo.

Seção 43

Zeros de funções polinomiais

Sumário
Índice
Rede

Os *zeros* de uma função real f qualquer (polinomial ou não), com domínio $d \subseteq \mathbb{R}$, são os valores $r \in d$ tais que

$$f(r) = 0.$$

Exemplos são dados nos próximos parágrafos. Mas, antes, precisamos de algumas considerações básicas.

Existe uma estreita relação entre zeros de funções reais e certas equações. Para evitar possíveis confusões muito comuns entre alunos, é essencial que o leitor tenha consciência sobre a importante diferença entre funções e equações. Funções são casos particulares de conjuntos, conforme discutido na Seção 29. Logo, funções são termos de ZF (conforme Seção 7). Equações, por outro lado, são fórmulas atômicas da forma $u = v$, onde u e v são termos (conforme Seção 7). Uma vez que nenhuma fórmula é um termo e nenhum termo é uma fórmula, nenhuma equação é uma função e nenhuma função é uma equação.

Se o termo r é um zero da função f, nas condições acima colocadas, então $(r, 0)$ pertence a f. No entanto, cada elemento de f é um par ordenado $(r, f(r))$. Logo, determinar os zeros de f é equivalente a determinar os reais r tais que

$$(r, f(r)) = (r, 0).$$

Por conta do Teorema 3.4,

$$(r, f(r)) = (r, 0) \text{ se, e somente se, } f(r) = 0.$$

Mas a igualdade

$$f(r) = 0$$

é uma equação, uma fórmula atômica que estabelece uma igualdade entre a imagem de r, via f, e 0. Neste contexto, a equação $f(r) = 0$ é equivalente a outra fórmula atômica, a saber, $(r, 0) \in f$.

Ou seja, determinar os zeros de uma *função* real f implica em responder quais são os valores r que satisfazem a *equação* $f(r) = 0$.

Em outras palavras, determinar os zeros de uma *função* real f implica em responder quais são os valores r tais que a *equação* $f(r) = 0$ é teorema.

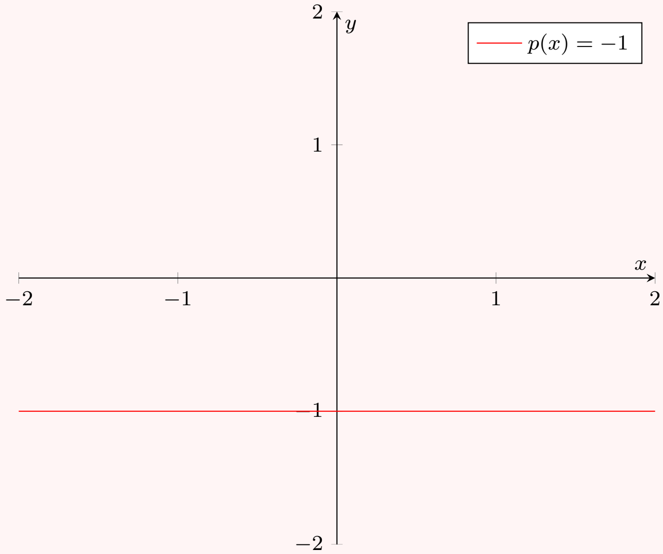
Polinomiais de grau 0

Seja $p : \mathbb{R} \to \mathbb{R}$ uma função polinomial de grau 0 qualquer, ou seja,

$$p(x) = \alpha.$$

Se a constante α for diferente de 0, então p não admite zero algum. Se a constante α for 0, então cada elemento do domínio $\mathbb{R}$ de p é um zero de p.

Nos próximos parágrafos promovemos uma breve discussão sobre zeros de funções polinomiais de grau maior do que 0. Deve ficar claro que polinomiais de grau 0 são as únicas polinomiais que podem admitir uma infinidade de zeros.



Na imagem acima temos uma representação gráfica de uma polinomial $p(x) = \alpha$ para o caso $\alpha = -1$.

Polinomiais de grau 1

Seja $f : \mathbb{R} \to \mathbb{R}$ uma função polinomial de grau 1 qualquer, ou seja,

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

sendo $\alpha \neq 0$. Determinar os zeros de f é equivalente a determinar os números reais x tais que $\alpha x + \beta = 0$ (ou seja, tais que a equação $\alpha x + \beta = 0$ é teorema, onde $\alpha \neq 0$). Neste caso,

$$x = -\frac{\beta}{\alpha}.$$

Ou seja,

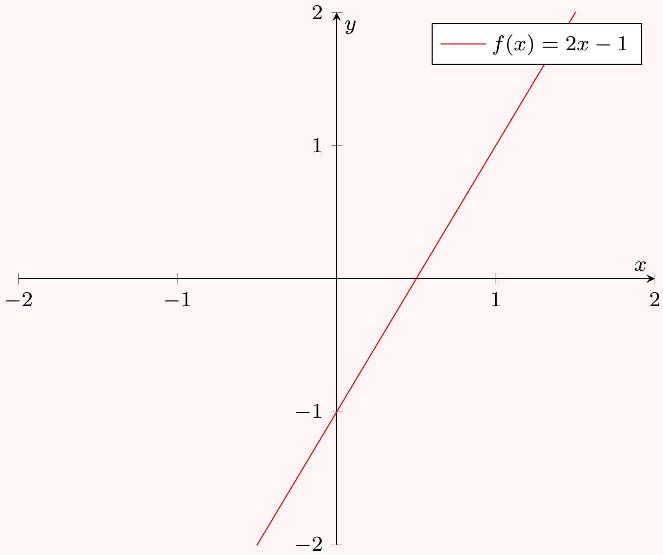
$$x = -\frac{\beta}{\alpha} \Leftrightarrow \alpha x + \beta = 0$$

é teorema. Isso equivale a afirmar que

$$x = -\frac{\beta}{\alpha}$$

é o único zero de f.

Observar que, de acordo com a discussão anterior, f é uma função e, portanto, um conjunto. Neste sentido, determinar os zeros de f é equivalente a determinar quais são os pares ordenados $(x, 0)$ tais que $(x, 0) \in f$. Já a equação $\alpha x + \beta = 0$ é uma fórmula atômica $u = v$ (onde u é o termo $\alpha x + \beta$ e v é o termo 0) usada como ferramenta para auxiliar na busca por zeros da função f. Comentário análogo vale para cada uma das funções polinomiais apresentadas nos próximos parágrafos.



Na imagem acima temos uma representação gráfica de uma polinomial $f(x) = \alpha x + \beta$ para o caso $\alpha = 2$ e $\beta = -1$.

Polinomiais de grau 2

Seja $g : \mathbb{R} \to \mathbb{R}$ uma função polinomial de grau 2 qualquer, ou seja,

$$g(x) = \alpha x^2 + \beta x + \gamma,$$

sendo $\alpha \neq 0$. Determinar os zeros de g é equivalente a definir os números reais x tais que $\alpha x^2 + \beta x + \gamma = 0$.

Mas $\alpha x^2 + \beta x + \gamma = 0$ equivale a $4\alpha^2 x^2 + 4\alpha\beta x + 4\alpha\gamma = 0$, uma vez que $\alpha \neq 0$. Logo,

$$4\alpha^2 x^2 + 4\alpha\beta x + 4\alpha\gamma + \beta^2 = \beta^2,$$

o que implica em $4\alpha^2 x^2 + 4\alpha\beta x + \beta^2 = \beta^2 - 4\alpha\gamma$.

A última equivale a $(2\alpha x + \beta)^2 = \beta^2 - 4\alpha\gamma$, a qual implica em $2\alpha x + \beta = \pm\sqrt{\beta^2 - 4\alpha\gamma}$, sendo que o símbolo $\pm$ serve ao propósito de destacar que existem até dois possíveis valores reais $2\alpha x + \beta$ tais que $(2\alpha x + \beta)^2 = \beta^2 - 4\alpha\gamma$.

Logo,

$$2\alpha x = -\beta \pm \sqrt{\beta^2 - 4\alpha\gamma},$$

o que implica em

$$x = \frac{-\beta \pm \sqrt{\beta^2 - 4\alpha\gamma}}{2\alpha}.$$

Logo, g admite um único zero sss $\beta^2 - 4\alpha\gamma = 0$.

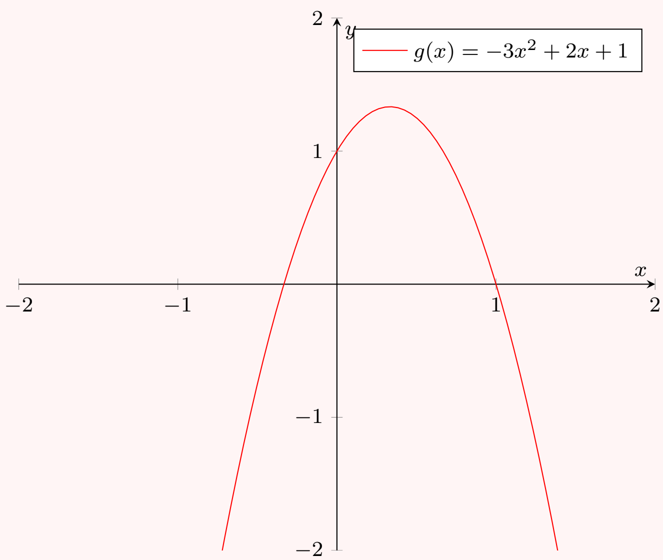
A mesma função polinomial g de grau 2 admite dois zeros se, e somente se, $\beta^2 - 4\alpha\gamma > 0$. Finalmente, g não admite zero algum sss $\beta^2 - 4\alpha\gamma < 0$.

Notar que os zeros da função g foram obtidos por meio de uma raiz quadrada envolvendo apenas os coeficientes α, β e γ.

Aparentemente Brasil é o único país do mundo a se referir à equação

$$x = \frac{-\beta \pm \sqrt{\beta^2 - 4\alpha\gamma}}{2\alpha}$$

como *fórmula de Bhāskara*, em homenagem ao famoso matemático e astrônomo indiano do século 12. A comunidade internacional se refere a ela como *fórmula quadrática*.



Na imagem acima temos uma representação gráfica de uma polinomial $g(x) = \alpha x^2 + \beta x + \gamma$ para o caso $\alpha = -3$, $\beta = 2$ e $\gamma = 1$.

Polinomiais de grau 3

Seja $h : \mathbb{R} \to \mathbb{R}$ uma função polinomial de grau 3 dada por

$$h(x) = \alpha x^3 + \beta x^2 + \gamma x + \delta,$$

sendo $\alpha \neq 0$. Estabelecer os zeros de h é equivalente a determinar os números reais x tais que $\alpha x^3 + \beta x^2 + \gamma x + \delta = 0$. A mudança de variáveis

$$x = t - \frac{\beta}{3\alpha}$$

permite reescrever a igualdade anterior como se segue:

$$\alpha t^3 + rt + s = 0,$$

sendo

$$r = \frac{3\alpha\gamma - \beta^2}{3\alpha} \quad \text{e} \quad s = \frac{2\beta^3 - 9\alpha\beta\gamma + 27\alpha^2\delta}{27\alpha^2}.$$

Observar que a igualdade $\alpha t^3 + rt + s = 0$ envolve um polinômio no qual o coeficiente real que multiplica t^2 é 0.

Os valores t que satisfazem a última equação são os mesmos que satisfazem

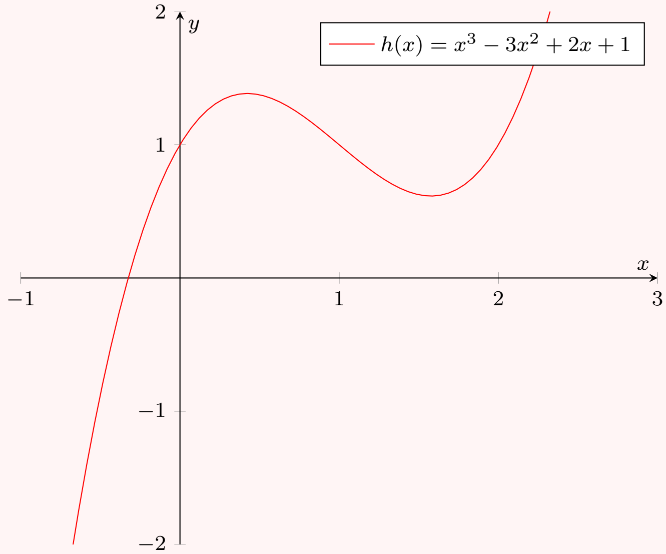
$$t^3 + \frac{r}{\alpha}t + \frac{s}{\alpha} = 0.$$

Logo, o problema pode ser resolvido por raízes $\sqrt[n]{\ }$, seguindo o *Método de Cardano* (em homenagem a Girolamo Cardano, polímata italiano do século 16). O Método de Cardano pode ser encontrado em inúmeras referências da literatura.

Uma vez obtidos os valores t que satisfazem a última equação em destaque (via Método de Cardano), o fato de que

$$x = t - \frac{\beta}{3\alpha}$$

permite obter os zeros da função polinomial h de grau 3, a qual admite pelo menos um zero real, independentemente dos valores de α, β, γ e δ. Em contrapartida, h admite no máximo três zeros reais.



Na imagem acima temos uma representação gráfica de uma polinomial $h(x) = \alpha x^3 + \beta x^2 + \gamma x + \delta$ para o caso $\alpha = 1$, $\beta = -3$, $\gamma = 2$ e $\delta = 1$.

Polinomiais de grau 4

Seja $i : \mathbb{R} \to \mathbb{R}$ uma função dada por

$$i(x) = \alpha x^4 + \beta x^3 + \gamma x^2 + \delta x + \varepsilon,$$

sendo $\alpha \neq 0$. Ou seja, i é uma função polinomial de grau 4. Determinar os zeros de i (aqui o símbolo i nada tem a ver com a unidade imaginária dos complexos!) é equivalente a determinar os números reais x tais que

$$\alpha x^4 + \beta x^3 + \gamma x^2 + \delta x + \varepsilon = 0.$$

A mudança de variáveis

$$x = t - \frac{\beta}{4\alpha}$$

permite reescrever a igualdade anterior como se segue:

$$\alpha t^4 + r t^2 + s t + u = 0.$$

sendo

$$r = \frac{-3\beta^2}{4\alpha} + \gamma, \quad s = \frac{\beta^3}{4\alpha^2} - \frac{\beta\gamma}{2\alpha} + \delta, \quad u = \frac{-3\beta^4}{256\alpha^3} + \frac{\gamma\beta^2}{16\alpha^2} - \frac{\beta\delta}{4\alpha} + \varepsilon.$$

Os valores de t que satisfazem a equação

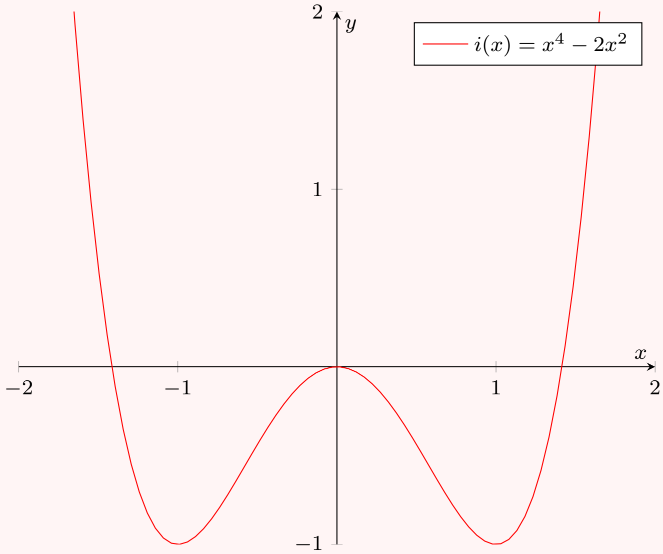
$$t^4 + \frac{r}{\alpha}t^2 + \frac{s}{\alpha}t + \frac{u}{\alpha} = 0$$

podem ser obtidos a partir do *Método de Ferrari* (em homenagem a Lodovico Ferrari, matemático italiano do século 16). O Método de Ferrari pode ser encontrado em diversas referências.

Uma vez obtidos os valores t que satisfazem a última equação (via Método de Ferrari), o fato de que

$$x = t - \frac{\beta}{4\alpha}$$

permite obter os zeros da função polinomial i de grau 4, a qual admite no máximo quatro zeros reais, podendo também não ter um único zero.



Na imagem acima temos uma representação gráfica de uma polinomial $i(x) = \alpha x^4 + \beta x^3 + \gamma x^2 + \delta x + \varepsilon$ para o caso $\alpha = 1$, $\beta = 0$, $\gamma = -2$, $\delta = 0$ e $\varepsilon = 0$.

Polinomiais de grau maior ou igual a 5

Seja $j : \mathbb{R} \to \mathbb{R}$ uma função dada por

$$j(x) = \alpha x^5 + \beta x^4 + \gamma x^3 + \delta x^2 + \varepsilon x + \zeta,$$

sendo $\alpha \neq 0$. Determinar os zeros de j é equivalente a definir os números reais x tais que

$$\alpha x^5 + \beta x^4 + \gamma x^3 + \delta x^2 + \varepsilon x + \zeta = 0.$$

O Teorema de Abel-Ruffini (1799, 1824) garante a impossibilidade de estabelecer soluções de equações polinomiais de grau maior ou igual a 5, para coeficientes reais arbitrários, em termos de raízes dos coeficientes que multiplicam cada monômio x^m. A Teoria de Galois (devida a Évariste Galois), a qual trata de conexões entre teoria de corpos e teoria de grupos, estende consideravelmente este resultado.

Para certos casos particulares de funções polinomiais, de grau maior ou igual a 5, ainda é possível determinar os zeros por meio de raízes dos coeficientes envolvidos. Um exemplo simples é a função $q : \mathbb{R} \to \mathbb{R}$ tal que

$$q(x) = x^8 - 1.$$

Neste caso, os zeros de q são os reais x tais que $x^8 = 1$, ou seja,

$$x = \pm\sqrt[8]{1},$$

o que equivale a $x = \pm 1$. Mas o Teorema de Abel-Ruffini impede que um método envolvendo raízes seja desenvolvido para toda e qualquer função polinomial de grau maior do que 4.

No entanto, ainda é possível obter zeros de funções polinomiais quaisquer (entre outras) via *aproximações* obtidas pelo truncamento de funções definidas recursivamente. Funções recursivas $\mathfrak{f}$ que podem ser programadas em máquinas são da forma

$$x_{n+1} = \mathfrak{f}(x_n),$$

onde $\mathfrak{f}(x_0) = \mathfrak{f}_0$. Essas funções são simplesmente sequências reais. O truncamento de $\mathfrak{f}$ para algum número natural m é necessário como critério de parada do algoritmo executado pela máquina, desde que existam condições de convergência para $\mathfrak{f}$. Detalhes sobre casos particulares de tais métodos implementáveis em máquinas são examinados na Seção 108.

Rudimentos de métodos numéricos definidos por funções recursivas são conhecidos há milênios, muito antes do advento do computador digital. Ver, por exemplo, o método babilônico para a obtenção da raiz quadrada de qualquer número real positivo, o qual também é discutido na Seção 108.

Observação final

Como última observação, vale a pena mencionar um fato de grande importância. O célebre *Teorema Fundamental da Álgebra* garante, como uma de suas consequências, que qualquer equação polinomial de grau n, ou seja, qualquer fórmula

$$a_n x^n + a_{n-1}x^{n-1} + a_{n-2}x^{n-2} + \cdots + a_3x^3 + a_2x^2 + a_1x + a_0 = 0,$$

onde $a_n \neq 0$, admite no máximo n valores reais x que satisfazem tal igualdade. Ou seja, ainda que não seja possível determinar – por raízes dos coeficientes a_n, a_{n-1}, $\cdots$, a_0 – todos os reais x que satisfazem essa equação (no caso de $n \geq 5$), pelo menos se sabe que há sempre uma quantia finita desses valores (quantia finita essa menor ou igual a n).

Logo, sejam $u : \mathbb{R} \to \mathbb{R}$ e $v : \mathbb{R} \to \mathbb{R}$ funções polinomiais quaisquer. O Teorema Fundamental da Álgebra garante, como outra consequência, que $u = v$ se, e somente se, os *coeficientes dos monômios de mesmo grau* de u e v forem idênticos. Se

$$u(x) = a_n x^n + a_{n-1}x^{n-1} + a_{n-2}x^{n-2} + \cdots + a_3x^3 + a_2x^2 + a_1x + a_0,$$

cada parcela $a_j x^j$ é chamada de *monômio de grau* j; além disso, o fator a_j é o *coeficiente* do monômio $a_j x^j$. Ou seja, a última afirmação é simplesmente a seguinte: se

$$u(x) = a_n x^n + a_{n-1}x^{n-1} + a_{n-2}x^{n-2} + \cdots + a_3x^3 + a_2x^2 + a_1x + a_0$$

e

$$v(x) = b_n x^n + b_{n-1}x^{n-1} + b_{n-2}x^{n-2} + \cdots + b_3x^3 + b_2x^2 + b_1x + b_0,$$

então $u(x) = v(x)$ para todo $x \in \mathbb{R}$ sss $a_j = b_j$ para todo j tal que $0 \leq j \leq n$.

A prova desse resultado pode ser feita por *reductio ad absurdum*. Com efeito, se houver algum j tal que $a_j \neq b_j$, então a igualdade $u(x) = v(x)$ passa a ser uma equação polinomial que somente pode ser satisfeita para uma quantia finita de possíveis valores reais x. Logo, não há igualdade entre $u(x)$ e $v(x)$ para todo x real.

Qualquer soma finita de monômios é um polinômio.

A demonstração do Teorema Fundamental da Álgebra está fora do escopo deste livro. No entanto, este é um resultado usado com muita frequência aqui e em diversas áreas da matemática.

Seção 44

Limite de função real

Sumário
Índice
Rede

O que se segue é fortemente relacionado com o conceito de sequência racional convergente. No caso desta Seção, porém, o foco é sobre funções reais. Também não há qualquer preocupação com valores arbitrariamente grandes. Estes são os primeiros passos na direção do *Cálculo Diferencial e Integral Padrão*.

Definição 5.6. *Seja $f : d \to \mathbb{R}$ uma função tal que $d \subseteq \mathbb{R}$ e $a \in (b, c)$ para algum intervalo aberto $(b, c) \subseteq d$. Logo,*

$$\lim_{x \to a} f(x) = L \vdots \forall \varepsilon > 0\ \exists \delta > 0\ (0 < |x - a| < \delta \Rightarrow |f(x) - L| < \varepsilon).$$

O termo $|x - a|$ é uma *distância* entre x e a, assim como $|f(x) - L|$ é uma *distância* entre $f(x)$ e L. A justificativa formal para essas afirmações é dada no Exemplo 8.44, Seção 88, Parte 8. Por enquanto, basta uma visão intuitiva sobre distâncias em $\mathbb{R}$.

O *definiens* na Definição 5.6 é equivalente à seguinte fórmula:

$$\forall \varepsilon > 0\ \exists \delta > 0\ (x \in (a - \delta, a + \delta) - \{a\} \Rightarrow f(x) \in (L - \varepsilon, L + \varepsilon)).$$

Observar que novamente estamos usando quantificadores relativizados, os quais foram introduzidos na Seção 35. A diferença é que agora estamos lidando com termos que são números reais.

Lê-se $\lim_{x \to a} f(x) = L$ como 'limite de $f(x)$, com x tendendo a a, é L'. A ideia intuitiva é a seguinte: afirmar

$$\lim_{x \to a} f(x) = L$$

equivale a dizer que, para toda vizinhança $(L - \varepsilon, L + \varepsilon)$ de L, deve existir uma vizinhança $(a - \delta, a + \delta)$ de a de modo que todo x pertencente a $(a - \delta, a + \delta)$, exceto o próprio a, admite uma imagem $f(x)$ pertencente a $(L - \varepsilon, L + \varepsilon)$.

É usual se referir ao conjunto $(L - \varepsilon, L + \varepsilon)$ como um *intervalo aberto centrado em L e com raio ε*. Analogamente, $(a - \delta, a + \delta)$ é um *intervalo aberto centrado em a e com raio δ*. Obviamente, todo intervalo aberto centrado em um real b, com raio γ real estritamente positivo, é uma vizinhança de b.

O conceito de vizinhança de um real, introduzido ao final da Seção 39, permite capturar a intuição de pontos 'próximos' de r. Neste contexto, em geral, quanto 'mais próximo' um x estiver de a, 'mais próximo' $f(x)$ está de L. Para efeitos práticos, quanto menor o valor de ε, menores os valores admissíveis para δ, caso o limite L exista. O único caso em que δ não depende de ε é aquele que envolve funções constantes, como se verifica no próximo teorema.

Teorema 5.9. *Seja $f : \mathbb{R} \to \mathbb{R}$ tal que $f(x) = c$. Logo, $\lim_{x\to a} f(x) = c$.*

Demonstração: Devemos provar que $\lim_{x\to a} c = c$. Logo, devemos provar que

$$\forall\varepsilon > 0\ \exists\delta > 0\ (0 < |x - a| < \delta \Rightarrow |c - c| < \varepsilon).$$

Mas essa última fórmula equivale a

$$\forall\varepsilon > 0\ \exists\delta > 0\ (0 < |x - a| < \delta \Rightarrow 0 < \varepsilon).$$

Qualquer δ real maior do que 0 satisfaz essa fórmula! Com efeito, basta aplicar Teorema 2.1 e Proposição 2.3. Afinal, $\varepsilon > 0 \Rightarrow \varepsilon > 0$ é teorema, independentemente de qualquer hipótese envolvendo δ.

O último teorema pode ser estendido para funções *localmente constantes*, i.e., para funções f tais que existe intervalo aberto (α, β) de modo que $f(x) = c$, para todo x pertencente a (α, β). Neste caso, $\lim_{x\to a} f(x) = c$ para todo a pertencente a (α, β). Obviamente a demonstração deste resultado exige um cuidado extra com relação ao valor de δ, a saber, δ pode ser qualquer real estritamente positivo menor ou igual ao menor dos dois valores a seguir: $a - \alpha$ e $\beta - a$.

Teorema 5.10. *Seja $f : \mathbb{R} \to \mathbb{R}$ tal que $f(x) = x$. Logo,*

$$\lim_{x\to a} f(x) = a.$$

Demonstração: Devemos provar que $\lim_{x\to a} x = a$. Logo, devemos provar que

$$\forall\varepsilon > 0\ \exists\delta > 0\ (0 < |x - a| < \delta \Rightarrow |x - a| < \varepsilon).$$

Faça $\delta = \varepsilon$. Com efeito, se $\delta = \varepsilon$, então a fórmula

$$0 < |x - a| < \delta \Rightarrow |x - a| < \varepsilon$$

é teorema. Isso porque a fórmula

$$|x-a| < \varepsilon \Rightarrow |x-a| < \varepsilon$$

é teorema, mesmo que $|x-a| < \varepsilon$ não seja teorema (ver Teorema 2.1 e Proposições 2.3 e 2.5).

O último teorema pode ser estendido para funções f tais que, localmente, se comportam como a função identidade, i.e., $f(x) = x$ para todo x pertencente a um intervalo aberto (α, β). Neste caso δ deve ser menor ou igual ao menor entre dois possíveis valores: $a - \alpha$ e $\beta - a$.

Sejam f e g funções reais que compartilham o mesmo domínio d tal que $d \subseteq \mathbb{R}$. Neste caso,

I $(f+g)$ é uma função com domínio d tal que

$$(f+g)(x) = f(x) + g(x),$$

para todo $x \in d$;

II $(f-g)$ é uma função com domínio d tal que

$$(f-g)(x) = f(x) - g(x),$$

para todo $x \in d$;

III (fg) é uma função com domínio d tal que

$$(fg)(x) = f(x)g(x),$$

para todo $x \in d$;

IV (f/g) é uma função com domínio d tal que

$$(f/g)(x) = f(x)/g(x),$$

desde que $g(x) \neq 0$ para todo x pertencente a d.

Observar que, na última definição, foram conceituadas adição, subtração, multiplicação e divisão entre funções reais, a partir de adição, subtração, multiplicação e divisão entre reais, respectivamente. Logo, as propriedades algébricas de adição e multiplicação entre funções reais são análogas àquelas entre reais, como comutatividade, associatividade e as demais.

Exemplo 5.9. *Sejam* $f : \mathbb{R} \to \mathbb{R}$ *e* $g : \mathbb{R} \to \mathbb{R}$ *funções tais que*

$$f(x) = x$$

e

$$g(x) = 5.$$

Logo, $f + g$ é uma função real com domínio $\mathbb{R}$ tal que

$$(f + g)(x) = x + 5.$$

Teorema 5.11. *Se f e g são funções que compartilham o mesmo domínio $d \subseteq \mathbb{R}$ e*

$$\lim_{x\to a} f(x) = L$$

e

$$\lim_{x\to a} g(x) = M,$$

então:

I: $\lim_{x\to a}(f + g)(x) = L + M$,

II: $\lim_{x\to a}(f - g)(x) = L - M$,

III: $\lim_{x\to a}(fg)(x) = LM$ *e*

IV: $\lim_{x\to a}(f/g)(x) = L/M$ *(se $M \neq 0$).*

Em particular,

$$\lim_{x\to a} cf(x) = cL,$$

se

$$\lim_{x\to a} f(x) = L.$$

✍ A demonstração do Teorema 5.11 é análoga às provas dos Teoremas 4.29, 4.32, 4.33 e 4.34. Basta fazer as adaptações necessárias. Recomendamos que o leitor faça isso como exercício.

Teorema 5.12. *Seja $p : \mathbb{R} \to \mathbb{R}$ uma função tal que*

$$p(x) = a_n x^n + a_{n-1}x^{n-1} + a_{n-2}x^{n-2} + \cdots + a_3x^3 + a_2x^2 + a_1x + a_0,$$

onde $a_0, a_1, \cdots, a_n$ são números reais e n é uma cópia de um inteiro positivo entre os reais. Então,

$$\lim_{x\to a} p(x) = p(a).$$

Demonstração: ✍ Basta usar os Teoremas 5.9, 5.10 e 5.11, uma vez que qualquer função polinomial é redutível a operações de adição e multiplicação envolvendo funções constantes e a função identidade. Lembrar também que essa prova tira proveito do fato de adição e multiplicação entre funções reais serem associativas. Com efeito, essas são definidas a partir de adição e multiplicação entre reais, como já mencionado. Essa prova exige paciência, mas é muito simples.

O último teorema pode ser estendido para funções *localmente polinomiais*, i.e., funções p tais que

$$p(x) = a_n x^n + a_{n-1}x^{n-1} + a_{n-2}x^{n-2} + \cdots + a_3x^3 + a_2x^2 + a_1x + a_0$$

para todo x pertencente a um intervalo aberto (α, β). Neste caso,

$$\lim_{x \to a} p(x) = p(a),$$

se a pertence a (α, β).

Uma função real f é *racional* sss

$$f(x) = \frac{p(x)}{q(x)},$$

sendo p e q funções polinomiais.

Não confundir funções racionais (aquelas cujas imagens são números racionais) com funções reais racionais, acima definidas.

Exemplo 5.10. i: *$f : \mathbb{R} \to \mathbb{R}$ tal que*

$$f(x) = \frac{x^3 - 2x}{x^2 + 4}$$

é uma função real racional; afinal, $m : \mathbb{R} \to \mathbb{R}$ dada por

$$m(x) = x^3 - 2x$$

e $n : \mathbb{R} \to \mathbb{R}$ dada por

$$n(x) = x^2 + 4$$

são funções polinomiais, e

$$f(x) = \frac{m(x)}{n(x)},$$

onde $n(x) \neq 0$ para todo x pertencente a $\mathbb{R}$;

ii: *toda função polinomial $p : \mathbb{R} \to \mathbb{R}$ é real racional, uma vez que*

$$p = \frac{p}{1}$$

e a função constante $g : \mathbb{R} \to \mathbb{R}$ dada por $g(x) = 1$ é polinomial de grau 0;

III: *A função $h : \mathbb{R} \to \mathbb{R}$ dada por $h(x) = |x|$ não é polinomial e nem real racional.*

Neste e no próximo parágrafo justificamos a afirmação feita no item III dado acima. Supor que $h : \mathbb{R} \to \mathbb{R}$, dada por

$$h(x) = |x|,$$

é polinomial. Logo, existem n, a_n, a_{n-1}, $\cdots$, a_0, tais que

$$h(x) = a_n x^n + a_{n-1} x^{n-1} + a_{n-2} x^{n-2} + \cdots + a_3 x^3 + a_2 x^2 + a_1 x + a_0.$$

No entanto,

$$h(x) = \begin{cases} x & \text{se } x \geq 0 \\ -x & \text{se } x < 0 \end{cases}$$

Logo, de acordo com a OBSERVAÇÃO FINAL da Seção 43, $h(x) = x$ e $h(x) = -x$ para todo x real. $\perp$.

✍ Aqui cabe outra observação. A função $k : d \to \mathbb{R}$ dada por

$$k(x) = |x|$$

é polinomial se todos os elementos de d forem reais positivos ou todos os elementos forem reais menores ou iguais a 0. Consegue provar?

O parágrafo acima deixa claro que o conceito de função polinomial depende do domínio da função.

EXEMPLO 5.11. *Seja $f : \mathbb{R} - \{0\} \to \mathbb{R}$ uma função tal que*

$$f(x) = \frac{|x|}{x}.$$

Neste caso, não existe L tal que

$$\lim_{x \to 0} f(x) = L.$$

Com efeito, se existisse o limite, então

$$\forall \varepsilon > 0\ \exists \delta > 0\ (x \in (0 - \delta, 0 + \delta) - \{0\} \Rightarrow f(x) \in (L - \varepsilon, L + \varepsilon)).$$

No caso em que

$$\varepsilon = \frac{1}{10},$$

não existe δ que satisfaça a condicional da definição. Isso porque $f(x)$ assume apenas os valores 1 e -1.

Além disso, a distância entre 1 *e* -1 *é* 2*, um número real bem maior do que*

$$\frac{1}{10}.$$

Logo, nem toda função admite limite.

Observar que a função do último EXEMPLO não é uma função real racional, uma vez que $|x|$ não define uma polinomial em $\mathbb{R} - \{0\}$.

Notar também que a função $f : \mathbb{R} - \{0\} \to \mathbb{R}$ tal que

$$f(x) = \frac{|x|}{x}$$

admite limite

$$\lim_{x \to a} f(x)$$

para qualquer a diferente de 0. Para provar isso, não esquecer que f é localmente polinomial em qualquer ponto a pertencente ao seu domínio.

DEFINIÇÃO 5.7. *Uma função real* f *é* contínua em um ponto *a sss*

$$\lim_{x \to a} f(x) = f(a).$$

EXEMPLO 5.12. I: *Toda função* $p : \mathbb{R} \to \mathbb{R}$ *polinomial é contínua em todos os pontos de seu domínio, conforme Teorema 5.12;*

II: $f : \mathbb{R} - \{0\} \to \mathbb{R}$ *tal que*

$$f(x) = \frac{|x|}{x}$$

não é contínua em 0*, apesar de ser contínua em todos os pontos de seu domínio.* ✍ *Fortemente recomendamos a demonstração deste teorema.*

TEOREMA 5.13. *O limite de uma função real, quando existe, é único.*

A prova deste último é análoga à do Teorema 4.35.

Seção 45

Estendendo limites

Sumário
Índice
Rede

Nesta Seção estendemos a Definição 5.6 sobre limite de função real, de modo a incluir doze outras definições de tipos especiais de limites. Tratam-se de conceitos úteis, por exemplo, na prova do importante Teorema 6.19, na Seção 61.

Definição 5.8. *Seja $f : d \to \mathbb{R}$ uma função tal que $d \subseteq \mathbb{R}$. Logo,*

$$\lim_{x\to a^+} f(x) = L \,\vdots\, \forall\varepsilon > 0\ \exists\delta > 0\ (a < x < a+\delta \Rightarrow |f(x)-L| < \varepsilon).$$

Lemos $\lim_{x\to a^+} f(x)$ como '*limite de $f(x)$, com x tendendo ao ponto a pela direita*'. A ideia é semelhante à Definição 5.6. A única diferença reside no fato de que aqui estamos interessado apenas nos reais x tais que

$$x \in (a, a+\delta).$$

Na Definição 5.6 estamos interessado em todos os x tais que

$$x \in (a-\delta, a) \cup (a, a+\delta).$$

A expressão 'tendendo pela direita' remete ao fato de que δ é estritamente positivo; portanto, $a+\delta > a$, o que implica que todo x pertencente ao intervalo aberto $(a, a+\delta)$ está 'à direita' de a.

Exemplo 5.13. *Seja $f : \mathbb{R} - \{0\} \to \mathbb{R}$ uma função tal que*

$$f(x) = \frac{|x|}{x}.$$

Neste caso, como já foi discutido anteriormente, não existe L tal que

$$\lim_{x\to 0} f(x) = L.$$

Com efeito, se existisse o limite, então

$$\forall\varepsilon > 0\ \exists\delta > 0\ (x \in (0-\delta, 0+\delta) - \{0\} \Rightarrow f(x) \in (L-\varepsilon, L+\varepsilon)).$$

No caso em que $\varepsilon = \frac{1}{10}$, não existe δ que satisfaça a condicional da definição. Isso porque $f(x)$ assume apenas os valores 1 e -1.

No entanto, existe $\lim_{x\to 0^+} f(x)$*. Além disso,* $\lim_{x\to 0^+} f(x) = 1$*. Com efeito,*

$$\forall \varepsilon > 0\ \exists \delta > 0\ (0 < x < 0 + \delta \Rightarrow |1 - 1| < \varepsilon).$$

Isso porque todo x *entre* 0 *e* $0+\delta$ *(excluindo* 0 *e* $0+\delta$*) é estritamente positivo, o que implica em* $|x| = x$ *e, portanto,* $f(x) = 1$*. Novamente Teorema 2.1 e Proposição 2.3 concluem a prova.*

Definição 5.9. *Seja* $f : d \to \mathbb{R}$ *uma função tal que* $d \subseteq \mathbb{R}$*. Logo,*

$$\lim_{x\to a^-} f(x) = L \vdots \forall \varepsilon > 0\ \exists \delta > 0\ (a-\delta < x < a \Rightarrow |f(x)-L| < \varepsilon).$$

Lemos $\lim_{x\to a^-} f(x)$ como '*limite de* $f(x)$*, com* x *tendendo ao ponto* a *pela esquerda*'. A ideia aqui é análoga à discussão acima, mas desta vez estamos lidando com reais x à esquerda de a.

Os termos

$$\lim_{x\to a^+} f(x) \quad \text{e} \quad \lim_{x\to a^-} f(x),$$

quando existem, são chamados de *limites laterais*.

Exemplo 5.14. ✍ *Seja* $f : \mathbb{R} - \{0\} \to \mathbb{R}$ *uma função tal que*

$$f(x) = \frac{|x|}{x}.$$

Neste caso, $\lim_{x\to 0^-} f(x) = -1$*. Cabe ao leitor justificar.*

Teorema 5.14. *Seja* f *uma função real. Logo,*

$$\lim_{x\to a} f(x) = L \ \ \textit{sss} \ \ \lim_{x\to a^+} f(x) = L \wedge \lim_{x\to a^-} f(x) = L.$$

✍ A discussão acima sobre limites laterais dá uma boa ideia de como provar esse último resultado. Por conta disso, deixamos a tarefa ao leitor. O teorema acima apenas afirma que limite existe sss os limites laterais existirem e forem coincidentes. Graças a esse teorema, os dois últimos Exemplos ficam bem mais fáceis de justificar. Afinal, uma mesma função admite limites laterais distintos. Logo, neste caso, ela não admite limite.

As próximas definições desta Seção são conhecidas como *limites envolvendo infinito*. Elas se dividem em dois grupos não necessariamente excludentes entre si, a saber, *limites infinitos* e *limites no infinito*.

> Definição 5.10. *Seja $f : d \to \mathbb{R}$ uma função tal que $d \subseteq \mathbb{R}$.*
>
> $$\lim_{x\to a^+} f(x) = \infty \vdots \forall\varepsilon > 0\ \exists\delta > 0\ (a < x < a + \delta \Rightarrow f(x) > \varepsilon).$$

Este é o primeiro caso de limite infinito. A definição acima segue uma notação abusiva que comumente confunde alunos. Isso porque no *definiendum* há uma igualdade na qual ocorre à sua direita o símbolo metalinguístico ∞. Mas é imprescindível que o leitor entenda que ∞ **não** é um termo da linguagem $\mathfrak{S}$ aqui usada. Além disso, ∞ também **não** abrevia termo algum de $\mathfrak{S}$. No entanto, a sentença metalinguística $\lim_{x\to a^+} f(x) = \infty$ abrevia a fórmula

$$\forall\varepsilon > 0\ \exists\delta > 0\ (a < x < a + \delta \Rightarrow f(x) > \varepsilon),$$

na qual não há uma única ocorrência do símbolo ∞.

Do ponto de vista intuitivo, a última definição captura a seguinte ideia: na medida em que x se aproxima de a pela direita, as imagens $f(x)$ se tornam arbitrariamente grandes. A ideia de imagens $f(x)$ se tornarem arbitrariamente grandes se caracteriza pela desigualdade

$$f(x) > \varepsilon,$$

a qual deve ser satisfeita para qualquer real ε estritamente positivo. Levando em conta que, nas condições acima ditadas, $f(x)$ não fica confinado a qualquer intervalo $(L - \varepsilon, L + \varepsilon)$, fica claro então que o *limite lateral* acima é um caso particular de limite que não existe. Justamente por isso que insistimos que a notação acima é abusiva. Todo limite infinito, como vemos nas próximas discussões, é um caso particular de limite que não existe.

Limites laterais são aqueles em que x pertence ao intervalo $(a, a+\delta)$ (limite lateral pela direita) ou $(a-\delta, a)$ (limite lateral pela esquerda).

A título de curiosidade, o símbolo ∞ foi introduzido em 1655, por John Wallis, em um tratado sobre seções cônicas. Contemporâneo de Isaac Newton, Wallis foi um dos responsáveis pela concepção do cálculo diferencial e integral.

EXEMPLO 5.15. *Seja $f : \mathbb{R} - \{0\} \to \mathbb{R}$ uma função dada por*

$$f(x) = \frac{1}{x}.$$

Logo,

$$\lim_{x \to 0^+} f(x) = \infty.$$

Com efeito, a fórmula

$$\forall \varepsilon > 0 \; \exists \delta > 0 \; \left(0 < x < 0 + \delta \Rightarrow \frac{1}{x} > \varepsilon \right)$$

(a qual é o definiens *na Definição 5.10) é equivalente à fórmula*

$$\forall \varepsilon > 0 \; \exists \delta > 0 \; \left(0 < x < \delta \Rightarrow x < \frac{1}{\varepsilon} \right),$$

uma vez que x é estritamente positivo. Logo, Teorema 2.1 e Proposição 2.3 garantem que basta fazer

$$\delta = \frac{1}{\varepsilon}.$$

DEFINIÇÃO 5.11. *Seja $f : d \to \mathbb{R}$ uma função tal que $d \subseteq \mathbb{R}$.*

$$\lim_{x \to a^+} f(x) = -\infty \vdots \forall \varepsilon > 0 \; \exists \delta > 0 \; (a < x < a + \delta \Rightarrow f(x) < -\varepsilon).$$

Este é um segundo exemplo de limite infinito. Neste caso, na medida em que x se aproxima pela direita de a, as imagens $f(x)$ se tornam arbitrariamente grandes em valor absoluto, mas com sinais negativos.

EXEMPLO 5.16. ✍ *Seja $f : \mathbb{R} - \{0\} \to \mathbb{R}$ dada por*

$$f(x) = \frac{-1}{x}.$$

Logo,

$$\lim_{x \to 0^+} f(x) = -\infty.$$

Recomendamos que o leitor faça a prova.

✍ O leitor pode escrever os conceitos de

$$\lim_{x \to a^-} f(x) = \infty$$

e

$$\lim_{x \to a^-} f(x) = -\infty,$$

bem como exemplificar.

As próximas duas definições são mais dois casos de limites infinitos.

Definição 5.12. *Seja $f : d \to \mathbb{R}$ uma função tal que $d \subseteq \mathbb{R}$.*

$$\lim_{x \to a} f(x) = \infty \vdots \forall \varepsilon > 0 \; \exists \delta > 0 \; (0 < |x - a| < \delta \Rightarrow f(x) > \varepsilon).$$

A ideia intuitiva é a seguinte: na medida em que x se aproxima de a (tanto pela esquerda quanto pela direita) as imagens $f(x)$ se tornam arbitrariamente grandes. Ou seja, mais um caso particular de limite que não existe.

Exemplo 5.17. *Seja $f : \mathbb{R} - \{0\} \to \mathbb{R}$ uma função dada por*

$$f(x) = \frac{1}{x^2}.$$

Logo,

$$\lim_{x \to 0} f(x) = \infty.$$

Com efeito, a fórmula

$$\forall \varepsilon > 0 \; \exists \delta > 0 \; \left(0 < |x - 0| < \delta \Rightarrow \frac{1}{x^2} > \varepsilon\right)$$

é equivalente à fórmula

$$\forall \varepsilon > 0 \; \exists \delta > 0 \; \left(0 < |x| < \delta \Rightarrow |x| < \frac{1}{\sqrt{\varepsilon}}\right).$$

Logo, basta fazer $\delta = \frac{1}{\sqrt{\varepsilon}}$.

Definição 5.13. *Seja $f : d \to \mathbb{R}$ uma função tal que $d \subseteq \mathbb{R}$.*

$$\lim_{x \to a} f(x) = -\infty \vdots \forall \varepsilon > 0 \; \exists \delta > 0 \; (0 < |x - a| < \delta \Rightarrow f(x) < -\varepsilon).$$

A ideia intuitiva é a seguinte: na medida em que x se aproxima de a (tanto pela esquerda quanto pela direita) os valores absolutos das imagens $f(x)$ se tornam arbitrariamente grandes mas com sinal negativo.

Exemplo 5.18. ✍ *Seja $f : \mathbb{R} - \{0\} \to \mathbb{R}$ uma função dada por*

$$f(x) = \frac{-1}{x^2}.$$

Logo,

$$\lim_{x \to 0} f(x) = -\infty.$$

Recomendamos ao leitor que prove isso.

TEOREMA 5.15. *Seja f uma função real. Logo,*

$$\lim_{x \to a} = \infty \;\; sss \;\; \lim_{x \to a^+} = \infty \wedge \lim_{x \to a^-} = \infty.$$

Analogamente,

$$\lim_{x \to a} = -\infty \;\; sss \;\; \lim_{x \to a^+} = -\infty \wedge \lim_{x \to a^-} = -\infty.$$

✍ A prova fica a cargo do leitor.

EXEMPLO 5.19. *Seja $f : \mathbb{R} - \{0\} \to \mathbb{R}$ uma função dada por $f(x) = \frac{1}{x}$. Logo, como discutido anteriormente,*

$$\lim_{x \to 0^+} f(x) = \infty \wedge \lim_{x \to 0^-} f(x) = -\infty.$$

Ou seja, além de não existir $\lim_{x \to 0} f(x)$, este é um caso de limite inexistente que não é limite infinito. Em outras palavras, todo limite infinito é um caso particular de limite inexistente. Mas nem todo limite inexistente é um limite infinito. ✍ *A propósito, essa função f é contínua em todos os pontos de seu domínio. Consegue provar isso?*

DEFINIÇÃO 5.14. *Seja $f : d \to \mathbb{R}$ uma função tal que $d \subseteq \mathbb{R}$.*

$$\lim_{x \to \infty} f(x) = L \;\; sss \;\; \forall \varepsilon > 0 \, \exists \delta > 0 \, (x > \delta \Rightarrow |f(x) - L| < \varepsilon).$$

Este é o primeiro caso de limite no infinito, pelo menos neste breve estudo sobre funções reais.

A ideia intuitiva aqui é a seguinte: na medida em que x se torna arbitrariamente grande (conceito esse dado por $x > \delta$), as imagens $f(x)$ ficam confinadas ao intervalo

$$(L - \varepsilon, L + \varepsilon),$$

para todo e qualquer real ε estritamente positivo.

EXEMPLO 5.20. *Seja $f : \mathbb{R} - \{0\} \to \mathbb{R}$ uma função dada por*

$$f(x) = \frac{1}{x}.$$

Logo,

$$\lim_{x\to\infty} f(x) = 0.$$

Com efeito, a fórmula

$$\forall\varepsilon > 0\ \exists\delta > 0\ \left(x > \delta \Rightarrow \left|\frac{1}{x} - 0\right| < \varepsilon\right)$$

é equivalente à fórmula

$$\forall\varepsilon > 0\ \exists\delta > 0\ \left(x > \delta \Rightarrow \left|\frac{1}{x}\right| < \varepsilon\right),$$

a qual é equivalente a

$$\forall\varepsilon > 0\ \exists\delta > 0\ \left(x > \delta \Rightarrow |x| > \frac{1}{\varepsilon}\right).$$

A última, por sua vez, é equivalente a

$$\forall\varepsilon > 0\ \exists\delta > 0\ \left(x > \delta \Rightarrow x > \frac{1}{\varepsilon}\right),$$

uma vez que x é estritamente positivo por conta da premissa $x > \delta$. Logo, basta fazer

$$\delta = \frac{1}{\varepsilon}.$$

Definição 5.15. *Seja $f : d \to \mathbb{R}$ uma função tal que $d \subseteq \mathbb{R}$.*

$$\lim_{x\to-\infty} f(x) = L \ \text{ sss } \ \forall\varepsilon > 0\ \exists\delta > 0\ (x < -\delta \Rightarrow |f(x) - L| < \varepsilon).$$

Exemplo 5.21. ✍ *Seja $f : \mathbb{R} - \{0\} \to \mathbb{R}$ dada por*

$$f(x) = \frac{1}{x}.$$

Logo,

$$\lim_{x\to-\infty} f(x) = 0.$$

O leitor deve justificar.

Definição 5.16. *Seja $f : d \to \mathbb{R}$ uma função tal que $d \subseteq \mathbb{R}$.*

$$\lim_{x\to\infty} f(x) = \infty \ \text{ sss } \ \forall\varepsilon > 0\ \exists\delta > 0\ (x > \delta \Rightarrow f(x) > \varepsilon).$$

Este é o primeiro caso de um limite infinito no infinito, entre funções reais. A ideia intuitiva é a seguinte: na medida em que x se torna arbitrariamente grande, as imagens $f(x)$ também assumem valores reais arbitrariamente grandes.

Exemplo 5.22. *Seja $f : \mathbb{R} \to \mathbb{R}$ dada por $f(x) = 6x$.*

Logo,

$$\lim_{x\to\infty} f(x) = \infty.$$

Com efeito, a fórmula

$$\forall\varepsilon > 0\ \exists\delta > 0\ (x > \delta \Rightarrow 6x > \varepsilon)$$

é equivalente a

$$\forall\varepsilon > 0\ \exists\delta > 0\ \left(x > \delta \Rightarrow x > \frac{\varepsilon}{6}\right).$$

Logo, basta fazer

$$\delta = \frac{\varepsilon}{6}.$$

✍ Para as três últimas definições de limites infinitos no infinito o próprio leitor pode criar seus exemplos.

Definição 5.17. *Seja $f : d \to \mathbb{R}$ uma função tal que $d \subseteq \mathbb{R}$. Logo,*

$$\lim_{x\to\infty} f(x) = -\infty \ \text{ sss } \ \forall\varepsilon > 0\ \exists\delta > 0\ (x > \delta \Rightarrow f(x) < -\varepsilon).$$

Definição 5.18. *Seja $f : d \to \mathbb{R}$ uma função tal que $d \subseteq \mathbb{R}$. Logo,*

$$\lim_{x\to-\infty} f(x) = \infty \ \text{ sss } \ \forall\varepsilon > 0\ \exists\delta > 0\ (x < -\delta \Rightarrow f(x) > \varepsilon).$$

Definição 5.19. *Seja $f : d \to \mathbb{R}$ uma função tal que $d \subseteq \mathbb{R}$. Logo,*

$$\lim_{x\to-\infty} f(x) = -\infty \ \text{ sss } \ \forall\varepsilon > 0\ \exists\delta > 0\ (x < -\delta \Rightarrow f(x) < -\varepsilon).$$

Seção 46

Mergulhando nas águas de limites

Sumário
Índice
Rede

✂ ✍ Esta Seção é um grande exercício.

A definição de limite de funções reais foi uma das grandes conquistas da matemática, introduzida por Karl Weierstraß e Augustin-Louis Cauchy no século 19. O padre católico Bernardus Bolzano teve ideias semelhantes muito antes de Weierstraß e Cauchy, mas a proposta dele passou completamente despercebida na época.

Não é uma tarefa fácil para pessoas em geral perceberem o que os quantificadores alternados $\forall\varepsilon > 0\ \exists\delta > 0$ realmente estão dizendo na definição de limite de função real. Uma ótima maneira para compreender a definição de limite é 'navegar' por ela, resolvendo exercícios, como aqueles já propostos até aqui. Outra, porém, é 'mergulhar' na definição, levantando a seguinte questão: o que aconteceria se a Definição 5.6 fosse diferente?

Uma vez que certas definições, como a de limite, estão socialmente consolidadas na comunidade matemática, não faz sentido propor qualquer alteração nelas. Mas, para fins de exercício investigativo, podemos propor novos conceitos inspirados na definição de limite.

Por exemplo, digamos que o matemático ficcional Dick Tate (*alter ego* de Dinah Mite) proponha o seguinte conceito:

$$\mathrm{limi}_{x\to a} f(x) = L \vdots \forall\varepsilon > 0\ \forall\delta > 0\ (0 < |x-a| < \delta \Rightarrow |f(x) - L| < \varepsilon,$$

onde lemos $\mathrm{limi}_{x\to a} f(x)$ como '*limitante* de $f(x)$ com x *tendencioso* a a'.

Neste caso, o limitante de $f(x)$ com x tendencioso a um real a existiria apenas para funções constantes $f : \mathbb{R} \to \mathbb{R}$ tais que $f(x) = c$. Além disso, temos que $\mathrm{limi}_{x\to a} f(x) = c$. Recomendamos ao leitor provar esse atípico teorema do misterioso senhor Tate.

Logo, o conceito de limitante seria algo trivial e completamente inútil.

Digamos agora que Hugh Jass, oponente de Dick Tate, proponha a seguinte definição:

$$\mathrm{limiar}_{x\to a} f(x) = L \vdots \exists\varepsilon > 0\ \exists\delta > 0\ (0 < |x-a| < \delta \Rightarrow |f(x)-L| < \varepsilon,$$

onde lemos $\mathrm{limiar}_{x\to a} f(x)$ como '*limiar* de $f(x)$ com x *se aproximando* de a'.

Nesta situação, o limiar de qualquer função $f : \mathbb{R} \to \mathbb{R}$ existiria para qualquer real a, algo bem mais abrangente do que os limitantes do senhor Tate. No entanto, o limiar de uma função real não seria único, tornando o conceito proposto por Jass como algo novamente inútil, uma vez que não poderia ser usado para definir o desejável conceito de derivada como caso particular de limiar. Uma das grandes vantagens da unicidade de limite radica no seu emprego para definir derivada de uma função, tema da próxima Seção.

Propomos ao leitor provar esses resultados sugeridos.

Outras propostas podem ser introduzidas para rivalizar (ou não) com a definição de limite. Esse tipo de atividade é um excelente exercício de criatividade, aparentemente nunca explorado em livros de cálculo diferencial e integral ou salas de aula.

Questões como 'o que aconteceria se a matemática fosse diferente' são altamente pertinentes para fins investigativos. Com efeito, tais questões podem provocar estimulantes discussões. Mas é necessário que os Tates e Jasses da vida não levem para o lado pessoal eventuais críticas que receberem às suas ideias.

Seção 47

Derivada

Sumário

Índice

Rede

Limites permitem definir *derivadas* e *integrais de Riemann*.

Derivadas são uma das ferramentas mais comumente empregadas para mapear fenômenos do mundo real. Isso porque derivadas capturam as ideias intuitivas de 'dinâmica', 'gradiente', 'velocidade', 'aceleração', 'taxa de variação', entre muitos outros. Neste sentido é uma prática comum a proposta de modelos matemáticos para descrever fenômenos físicos a partir de derivadas.

Integrais de Riemann, por sua vez, permitem lidar com os modelos propostos via derivadas, para que seja possível fazer previsões de longo termo. Detalhes são dados na medida em que avançamos por aqui. Exemplos são dados também adiante.

Definição 5.20. *Seja f uma função real. A* derivada *de f em relação a x* no ponto *a, se existir, é definida como*

$$\frac{d}{dx}f(x)\Big|_{x=a} = \lim_{h\to 0}\frac{f(a+h)-f(a)}{h}.$$

Obviamente, duas condições necessárias para

$$\frac{d}{dx}f(x)\Big|_{x=a}$$

existir são as seguintes:

I: $f(a)$ deve existir, ou seja, a é um elemento do domínio de f, e

II: $f(x)$ deve existir para todos os reais x pertencentes a algum intervalo aberto I tal que $a \in I$.

Com efeito, uma vez que

$$\frac{d}{dx} f(x)\Big|_{x=a}$$

é um limite com 'h tendendo a 0', isso corresponde a dizer que h pertence a um intervalo aberto

$$(0 - \delta, 0 + \delta),$$

exceto o ponto 0, de acordo com a definição de limite.

Logo, $a + h$ (termo usado na Definição 5.20) pertence ao conjunto

$$(a - \delta, a + \delta) - \{a\},$$

o qual é uma vizinhança de a (lembrar que δ é estritamente positivo) na qual se ignora o próprio ponto a.

No entanto, essas duas condições ($f(a)$ existe e $f(x)$ existe para qualquer x de uma vizinhança de a), apesar de necessárias, não são suficientes para garantir a existência de $\frac{d}{dx} f(x)\Big|_{x=a}$, conforme exemplificamos mais adiante.

Com relação à visão intuitiva da Definição 5.20, ela expressa a

taxa de variação de $f(x)$, em relação à variação de x, no ponto a.

Conforme discutido na Seção 1, matemática pode ser usada para *mapear* fenômenos do mundo real. Neste contexto, digamos que x seja interpretado como tempo em segundos, enquanto $f(x)$ é interpretado como a posição, em metros, de um corpo material ao longo de uma estrada retilínea. Diante desta interpretação é sugerido que a posição $f(x)$, em metros, do corpo material depende do tempo x em segundos: a cada instante x de tempo o corpo está em uma posição $f(x)$.

A *velocidade média* do corpo material, ao longo de dez segundos, pode ser obtida da seguinte maneira:

I: no instante, digamos, 7s (a letra s abrevia 'segundos') avalia-se a posição $f(7)$m (a letra m abrevia metros) do corpo;

ii: no instante $(7+10)$s (ou seja, 17s) avalia-se a posição $f(17)$m;

iii: a razão

$$\frac{f(17)-f(7)}{17-7}$$

é a velocidade média do corpo ao longo de dez segundos na estrada retilínia. Tal velocidade média é dada em metros por segundo, uma vez que temos uma divisão entre distância percorrida e intervalo de tempo transcorrido.

No contexto da Definição 5.20, a discussão acima corresponde ao termo

$$\frac{f(a+h)-f(a)}{h},$$

uma vez que $h=(a+h)-a$.

No entanto, Definição 5.20, no contexto do mapeamento proposto, não informa velocidade média do corpo material em questão, mas *velocidade instantânea*. Isso ocorre por conta do limite aplicado sobre a função

$$\frac{f(a+h)-f(a)}{h},$$

com 'h tendendo a 0'.

Em outras palavras, a derivada de f em relação a x, no ponto a, é o limite de uma função $g(h)$ com h tendendo a zero, sendo

$$g(h)=\frac{f(a+h)-f(a)}{h}.$$

Um carro viajando de Curitiba a São Paulo pode ter uma velocidade média de 50 quilômetros por hora ao longo de todo o tempo de viagem, ainda que em alguns trechos da estrada atinja a velocidade instantânea de 120 quilômetros por hora e, em outros, permaneça com uma velocidade instantânea de 0 quilômetros por hora (por conta de um engarrafamento). Velocidade instantânea, neste sentido, corresponde àquilo que é registrado no velocímetro do carro.

Retornando à Definição 5.20, uma vez que 'h tende a zero', isso corresponde ao fato de h pertencer ao intervalo aberto

$$(0-\delta, 0+\delta),$$

exceto possivelmente o ponto 0.

Logo, h pode assumir tanto valores reais estritamente positivos (à direita de 0) quanto valores reais negativos (à esquerda de 0). Se

$$\lim_{h\to 0} \frac{f(a+h)-f(a)}{h}$$

existir, porém, obviamente o valor de $\frac{d}{dx}f(x)\Big|_{x=a}$ deve ser um único real, estritamente positivo, negativo ou nulo, por conta do Teorema 5.13.

Observar que a possível interpretação de derivada de uma função real, em um ponto a, como velocidade instantânea, não é única. Se x for interpretado como tempo em segundos e $f(x)$ for interpretado como velocidade em metros por segundo,

$$\frac{d}{dx}f(x)\Big|_{x=a}$$

é uma aceleração (variação de velocidade em relação a tempo) instantânea em metros por segundo por segundo, no instante a.

Se x for associado com posição em metros numa reta vertical (altura) e $f(x)$ for associado com temperatura em graus Celsius, o mesmo valor

$$\frac{d}{dx}f(x)\Big|_{x=a}$$

corresponde a um gradiente de temperatura na altura vertical a.

Aplicações de derivadas para lidar com fenômenos físicos são amplamente documentadas e muito bem sucedidas há mais de três séculos, ajudando a moldar até mesmo a economia de nações, no que se refere a avanços tecnológicos.

No entanto, em momento algum é sugerido que o mapeamento matemático de fenômenos físicos implica que a posição de um automóvel numa estrada é uma função de tempo, no sentido do que se entende por funções em ZF. Assim como o mapa de uma cidade não é a cidade, a matemática opera tão somente como uma forma de retratar certos aspectos do universo onde todos vivemos.

Por outro lado, derivadas de funções reais em um dado ponto admitem interpretações fora do âmbito de aplicações no mundo real. Por conta disso, segue a próxima Seção, na qual conceitos muito elementares de Geometria Analítica Plana são explorados do ponto de vista de derivadas. Na Seção 49 continuamos a estudar derivadas através de alguns teoremas importantes. Ne Secão 63 usamos derivadas para mapear fenômenos físicos de decaimento radioativo.

Seção 48

Plano cartesiano

Nesta Seção estamos interessado apenas nas retas

$$r = \{(x, y) \in \mathbb{R}^2 \mid ax + by = c\}$$

tais que $b \neq 0$. Estas são chamadas de *retas não verticais*. Uma reta não vertical, portanto, é o conjunto de todos os pares ordenados $(x, y) \in \mathbb{R}^2$ tais que

$$y = \alpha x + \beta,$$

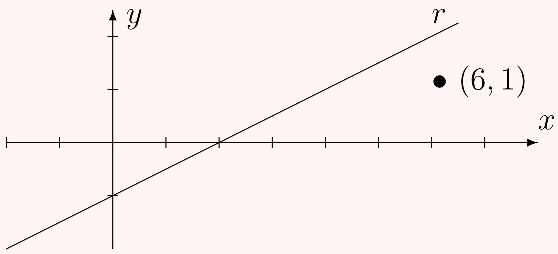
onde

$$\alpha = \frac{-a}{b} \ \text{ e } \ \beta = \frac{c}{b}.$$

O real α é chamado de *coeficiente angular* da reta r e β é chamado de *coeficiente linear* da reta.

Pontos e retas podem ser representados visualmente como se segue na próxima imagem.

Se (x, y) é um ponto de $\mathbb{R}^2$, chamamos x de *abscissa* do ponto e y de *ordenada*. Os valores x e y são chamados de *coordenadas* do ponto (x, y).



O conjunto de todas as possíveis abscissas de pontos de $\mathbb{R}^2$ está visualmente representado acima pelo eixo horizontal x. O conjunto de todas as possíveis ordenadas de pontos de $\mathbb{R}^2$ está representado pelo eixo vertical y. Logo, os eixos horizontal x e vertical y permitem identificar univocamente quaisquer coordenadas de quaisquer pontos de $\mathbb{R}^2$.

Sumário
Índice
Rede

As flechas representadas nos eixos servem para indicar que os reais do eixo x crescem para a direita, enquanto os reais do eixo y crescem para cima. A interseção entre esses eixos é o ponto $(0, 0)$. Na imagem acima há uma representação visual do ponto $(6, 1)$ e da reta r dada por

$$y = \frac{1}{2}x - 1.$$

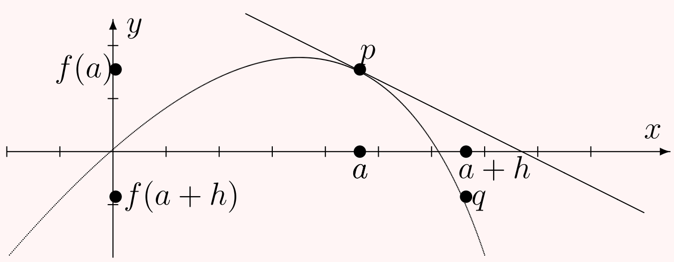
Observar que, se $x = 0$, então $y = -1$. Isso significa que o ponto $(0, -1)$ incide sobre r. Além disso, se $y = 0$, então $x = 2$, o que implica que o ponto $(2, 0)$ também incide sobre r. Há uma infinidade de outros pontos incidentes sobre r. Mas os pontos $(0, -1)$ e $(2, 0)$ bastam para definir r.

Observar também que o ponto $(6, 1)$ não incide sobre r. Com efeito, se $x = 6$, então $y = 2$. Logo, qualquer ponto com abscissa 6 incidente sobre r deve ter ordenada 2, o que não ocorre com o ponto $(6, 1)$.

Um ponto (x, y) de $\mathbb{R}^2$ incide sobre uma reta r sss $(x, y) \in r$.

Pois bem. Acontece que existe uma estreita relação entre derivadas de funções reais em um ponto e retas de $\mathbb{R}^2$.

Uma vez que toda função $f : \mathbb{R} \to \mathbb{R}$ é um subconjunto de $\mathbb{R}^2$, f também pode contar com uma representação visual de maneira análoga àquela da última imagem.



Na imagem acima a *Curva Bézier* ilustrada é uma representação visual de uma função $f : \mathbb{R} \to \mathbb{R}$.

Curvas de Bézier foram concebidas nos anos 1960 por Pierre Bézier, para o desenho de carros Renault. Hoje são amplamente utilizadas em computação gráfica. Aqui empregamos para ilustrar uma visão intuitiva sobre derivadas.

O ponto p ilustrado acima é o par ordenado $(a, f(a))$, o qual é um ponto pertencente a f. O ponto q (também pertencente a f) é o par ordenado $(a+h, f(a+h))$, para o caso particular em que h é estritamente positivo. Por conta disso que $a+h$ está à direita de a nesta ilustração.

Os pontos p e q definem um segmento de reta que pode ser interpretado como a hipotenusa de um triângulo retângulo no qual um dos catetos mede a distância de a até $a+h$, ou seja, $|h|$ (neste caso, $|h| = h$). Em contrapartida, o outro cateto mede a distância de $f(a)$ até $f(a+h)$, ou seja, $|f(a+h) - f(a)|$.

A razão

$$\frac{f(a+h) - f(a)}{h}$$

é o coeficiente angular de uma reta definida pelos pontos p e q. Essa reta definida por p e q (não representada visualmente para não sobrecarregar a imagem) intersecta a função f exatamente nos pontos p e q. Ao aplicar o limite

$$\lim_{h \to 0} \frac{f(a+h) - f(a)}{h},$$

o ponto $a+h$ se 'aproxima arbitrariamente' do ponto a. Uma vez que o limite, quando existe, é único, neste caso o limite acima é o coeficiente angular de uma reta que tangencia a função f no ponto a. Situação análoga ocorre para o caso em que h é negativo e, consequentemente, $a+h$ está à esquerda de a.

Neste contexto, retas tangentes a uma curva são casos 'limites' de retas secantes, sendo que uma reta secante a uma curva é aquela que intersecta a curva em pelo menos dois pontos.

Logo,

$$\left.\frac{d}{dx} f(x)\right|_{x=a},$$

se existir, pode ser interpretada como o coeficiente angular de uma reta que tangencia f no ponto $(a, f(a))$.

Observar que os valores $f(a+h) - f(a)$ e h não correspondem necessariamente a medidas de catetos de um triângulo retângulo com hipotenusa definida pelos pontos p e q, uma vez que tais valores podem eventualmente ser negativos. No entanto, os valores absolutos de $f(a+h) - f(a)$ e h são medidas de tais catetos. Essa questão é relevante para discussões sobre crescimento de decrescimento de funções, conforme vemos mais adiante.

Seção 49

Teoremas elementares sobre derivadas

Sumário
Índice
Rede

Aprendemos aqui a calcular algumas derivadas.

Teorema 5.16. *Seja $f : \mathbb{R} \to \mathbb{R}$ tal que $f(x) = c$ (função constante). Logo, a derivada de f em relação a x em qualquer ponto a do domínio de f é 0.*

Demonstração:

$$\frac{d}{dx}c\Big|_{x=a} = \lim_{h\to 0}\frac{c-c}{h} = \lim_{h\to 0}\frac{0}{h} = \lim_{h\to 0} 0 = 0.$$

Observar como se justifica a igualdade

$$\lim_{h\to 0}\frac{0}{h} = \lim_{h\to 0} 0$$

na última demonstração: levando em conta que h tende a 0, isso é equivalente a afirmar que h pertence a uma vizinhança $(0-\delta, 0+\delta)$ exceto o ponto 0. Logo, de fato $\frac{0}{h}$ é igual a 0, uma vez que h é diferente de 0.

Com relação à última igualdade na demonstração acima,

$$\lim_{h\to 0} 0 = 0,$$

foi aplicado o Teorema 5.9 sobre limite de função constante (aqui a constante é zero).

O último teorema pode ser estendido para funções *localmente constantes*, i.e., para funções f tais que existe intervalo aberto (α, β) de modo que $f(x) = c$ para todo x pertencente a (α, β). Neste caso,

$$\frac{d}{dx}f(x)\Big|_{x=a} = 0$$

para todo a pertencente a (α, β).

Teorema 5.17. *Seja $f : \mathbb{R} \to \mathbb{R}$ tal que*

$$f(x) = x$$

(função identidade em $\mathbb{R}$). Logo, a derivada de f em relação a x em qualquer ponto a do domínio de f é 1.

Demonstração:

$$\frac{d}{dx}x\Big|_{x=a} = \lim_{h\to 0}\frac{(a+h)-a}{h} = \lim_{h\to 0}\frac{h}{h} = \lim_{h\to 0} 1 = 1.$$

Observar que novamente estamos levando em conta que h é diferente de 0, uma vez que h pertence a uma vizinhança de 0, excluindo o próprio ponto 0. Por conta disso que temos

$$\lim_{h\to 0}\frac{h}{h} = \lim_{h\to 0} 1$$

na última demonstração.

Novamente aplicamos o Teorema 5.9 sobre limite de função constante, para o último passo. Neste caso a constante é 1.

O último teorema pode ser estendido para funções que se comportam localmente como a função identidade. Comentário análogo vale para teoremas que seguem nos próximos parágrafos.

Teorema 5.18. *Seja $f : \mathbb{R} \to \mathbb{R}$ tal que*

$$f(x) = x^2.$$

Logo, a derivada de f em relação a x em qualquer ponto a do domínio de f é $2a$.

Demonstração:

$$\frac{d}{dx}x^2\Big|_{x=a} = \lim_{h\to 0}\frac{(a+h)^2-a^2}{h} = \lim_{h\to 0}\frac{a^2+2ah+h^2-a^2}{h} =$$

$$\lim_{h\to 0}\frac{h(2a+h)}{h} = \lim_{h\to 0}(2a+h) = \lim_{h\to 0} 2a + \lim_{h\to 0} h =$$

$$2a + 0 = 2a.$$

Cabe ao leitor justificar cada passo da demonstração.

Teorema 5.19. *Seja $f : \mathbb{R} \to \mathbb{R}$ tal que*

$$f(x) = x^n,$$

onde n é um real que copia um inteiro estritamente positivo.

Logo, a derivada de f em relação a x em qualquer ponto a do domínio de f é

$$na^{n-1}.$$

Demonstração:

$$\begin{aligned}\frac{d}{dx}x^n\Big|_{x=a} &= \lim_{h\to 0}\frac{(a+h)^n - a^n}{h} = \\ &\lim_{h\to 0}\frac{1}{h}((a+h)^n - a^n) = \\ &\lim_{h\to 0}\frac{1}{h}((a^n + na^{n-1}h + \cdots + h^n) - a^n) = \\ &\lim_{h\to 0}\frac{1}{h}(na^{n-1}h + \cdots + h^n) = \\ &\lim_{h\to 0}\frac{1}{h}h(na^{n-1} + \cdots + h^{n-1}) = \\ &\lim_{h\to 0}(na^{n-1} + \cdots + h^{n-1}) = \\ &na^{n-1}.\end{aligned}$$

Notar que

$$na^{n-1} + \cdots + h^{n-1}$$

é um polinômio relativamente à variável h, de grau $n-1$.

Logo, na última igualdade empregamos o Teorema 5.12 sobre limite de funções polinomiais.

Na última demonstração foi utilizado o *binômio de Newton*, o qual é uma generalização do Teorema Binomial para Naturais (ver Teorema 4.6), no seguinte sentido: se a e b são reais e n é um real que copia naturais diferentes de 0, então

$$(a+b)^n = \sum_{k=0}^{n}\binom{n}{k}a^k b^{n-k}.$$

✍ Fortemente recomendamos que o leitor justifique passo a passo essa última prova.

Teorema 5.20. *Se*

$$\frac{d}{dx}f(x)\Big|_{x=a}$$

existe, então

$$\frac{d}{dx}cf(x)\Big|_{x=a} = c\left(\frac{d}{dx}f(x)\Big|_{x=a}\right),$$

onde c é um real.

DEMONSTRAÇÃO:

$$\left.\frac{d}{dx}cf(x)\right|_{x=a} = \lim_{h\to 0}\frac{cf(a+h)-cf(a)}{h} =$$

$$\lim_{h\to 0}\frac{c(f(a+h)-f(a))}{h} =$$

$$\lim_{h\to 0} c \cdot \lim_{h\to 0}\frac{f(a+h)-f(a)}{h} = c\left.\frac{d}{dx}f(x)\right|_{x=a}.$$

✍ Recomendamos que o leitor justifique cada igualdade da última demonstração.

DEFINIÇÃO 5.21. *Seja $f : d \to \mathbb{R}$ uma função tal que d é uma união arbitrária de intervalos abertos de reais. A* função derivada *f' de f, se existir, é dada por $f' : d \to \mathbb{R}$ tal que,*

$$\forall a\left(a \in d \Rightarrow f'(a) = \left.\frac{d}{dx}f(x)\right|_{x=a}\right).$$

Uma função real, nas condições acima, é diferenciável *sss admitir função derivada.*

Em outras palavras, se uma função f admite derivada em cada ponto de seu domínio, é possível definir a *função derivada* de f simplesmente como uma função f' tal que cada termo x do domínio de f tem como imagem $f'(x)$ a derivada de f no ponto x. O domínio de f' é o mesmo de f.

EXEMPLO 5.23. I: *se $f : \mathbb{R} \to \mathbb{R}$ é uma função dada por*

$$f(x) = c,$$

então $f' : \mathbb{R} \to \mathbb{R}$ dada por

$$f'(x) = 0$$

é a função derivada de f, por conta do Teorema 5.16;

II: *se $g : \mathbb{R} \to \mathbb{R}$ é uma função dada por*

$$g(x) = x^4,$$

então $g' : \mathbb{R} \to \mathbb{R}$ dada por

$$g'(x) = 4x^3$$

é a função derivada de g, por conta do Teorema 5.19.

As duas funções dadas no Exemplo acima são diferenciáveis.

A seguir mostramos como definir funções derivadas em intervalos fechados não degenerados.

Definição 5.22. *Seja* $f : [\alpha, \beta] \to \mathbb{R}$*, uma função, onde* $\alpha < \beta$*. A* função derivada f' *de* f*, se existir, é dada por*

$$f' : [\alpha, \beta] \to \mathbb{R}$$

tal que,

$$f'(a) = \begin{cases} \frac{d}{dx} f(x)\Big|_{x=a} & \text{se } a \in (\alpha, \beta) \\ \lim_{h\to 0^+} \frac{f(a+h)-f(a)}{h} & \text{se } a = \alpha \\ \lim_{h\to 0^-} \frac{f(a+h)-f(a)}{h} & \text{se } a = \beta \end{cases}$$

Neste caso dizemos que f *é* diferenciável *no intervalo fechado* $[\alpha, \beta]$.

Os termos

$$\lim_{h\to 0^+} \frac{f(a+h)-f(a)}{h}$$

e

$$\lim_{h\to 0^-} \frac{f(a+h)-f(a)}{h},$$

na última definição, são chamados de *derivada à direita* e *derivada à esquerda*, respectivamente.

Exemplo 5.24. ✍

I: *se* $f : [-\sqrt{2}, \sqrt{3}] \to \mathbb{R}$ *é uma função dada por* $f(x) = c$*, então* $f' : [-\sqrt{2}, \sqrt{3}] \to \mathbb{R}$ *dada por* $f'(x) = 0$ *é a função derivada de* f*, por conta do Teorema 5.16 e de sua extensão para derivadas laterais;*

II: *se* $g : [2, 3] \to \mathbb{R}$ *é uma função dada por* $g(x) = x^4$*, então* $g' : [2, 3] \to \mathbb{R}$ *dada por* $g'(x) = 4x^3$ *é a função derivada de* g*, por conta do Teorema 5.19 e de sua extensão para derivadas laterais.*

✍ Todos os teoremas para derivadas, examinados neste livro, podem ser generalizados para derivadas laterais. Por conta disso, não nos preocupamos com elas nas próximas demonstrações.

Teorema 5.21. *Se u e v compartilham o mesmo domínio e são diferenciáveis, então $(u+v)' = u' + v'$.*

Demonstração:

$$(u+v)' = \lim_{h\to 0} \frac{(u+v)(x+h) - (u+v)(x)}{h} =$$

$$\lim_{h\to 0} \frac{u(x+h) + v(x+h) - (u(x) + v(x))}{h} =$$

$$\lim_{h\to 0} \left(\frac{u(x+h) - u(x)}{h} + \frac{v(x+h) - v(x)}{h} \right) =$$

$$\lim_{h\to 0} \frac{u(x+h) - u(x)}{h} + \lim_{h\to 0} \frac{v(x+h) - v(x)}{h} = u' + v'$$

Ou seja, a derivada de uma adição é a adição de derivadas.

Teorema 5.22. *Se u e v compartilham o mesmo domínio e são diferenciáveis, então $(u-v)' = u' - v'$.*

Demonstração: Basta usar o Teorema 5.21 em parceria com o Teorema 5.20 para a constante $c = -1$. ✍ Com efeito, é teorema entre os reais a seguinte fórmula:

$$(-1)r = -r$$

para qualquer real r (consegue provar isso?).

Ou seja, a derivada de uma diferença é a diferença entre derivadas.

Uma consequência imediata dos dois últimos teoremas é que, se u e v são diferenciáveis, então $u+v$ e $u-v$ são diferenciáveis.

Para que possamos enunciar e provar teoremas sobre derivada de produto e derivada de uma razão, precisamos de mais informações.

Definição 5.23. *Seja f uma função real cujo domínio é uma união arbitrária de intervalos abertos de reais. Dizemos que f é* contínua *sss*

$$\lim_{x\to a} f(x) = f(a)$$

para todo a pertencente ao domínio de f.

Ou seja, grosso modo, uma função real é contínua se, e somente se, for contínua em todos os pontos de seu domínio.

Exemplo 5.25. i: *Teorema 5.12 garante que toda função polinomial*

$$p : \mathbb{R} \to \mathbb{R}$$

é contínua.

ii: *Seja $f : \mathbb{R} \to \mathbb{R}$ uma função dada por*

$$f(x) = \begin{cases} 7 & se\ x < 2 \\ 8 & se\ x > 2 \\ 9 & se\ x = 2 \end{cases}$$

Se $a < 2$, então existem α e β tais que $\alpha < 2$, $\beta < 2$ e $\alpha < a < \beta$; logo, f é localmente constante no intervalo aberto (α, β), no sentido de que $a \in (\alpha, \beta)$ e todo x pertencente a (α, β) é menor do que 2; portanto,

$$\lim_{x \to a} f(x) = f(a) = 7.$$

Se $a > 2$, novamente f é localmente constante e, portanto,

$$\lim_{x \to a} f(x) = f(a) = 8.$$

No entanto, não existe real L tal que

$$\lim_{x \to 2} f(x) = L$$

(recomendamos provar isso). Logo, não é teorema a fórmula

$$\lim_{x \to 2} f(x) = f(2),$$

apesar de $f(2) = 9$. Uma vez que 2 pertence ao domínio de f mas f não é contínua em 2, então f não é contínua.

iii: *A* função de Dirichlet *$g : \mathbb{R} \to \mathbb{R}$ dada por*

$$g(x) = \begin{cases} 1 & se\ x\ é\ real\ racional \\ 0 & se\ x\ é\ real\ irracional \end{cases}$$

é não contínua em todos os pontos *de seu domínio. Logo, este é um exemplo bastante radical de função não contínua.* ✍ *Recomendamos que o leitor prove isso. Dica: demonstrar que todo intervalo aberto (α, β) de números reais conta com reais racionais e reais irracionais pertencentes a ele.*

Teoremas 5.25 e 5.26 evidenciam profunda relação entre diferenciabilidade e continuidade. Mas, antes de explorarmos essa importante questão, precisamos dos próximos resultados.

TEOREMA 5.23. *A composição entre funções contínuas, quando existe, é uma função contínua.*

DEMONSTRAÇÃO: Sejam f e g funções reais tais que $f \circ g$ existe e g é contínua em a, enquanto f é contínua em $g(a)$. Logo,

$$\forall \varepsilon > 0\ \exists \delta > 0\ (0 < |x - a| < \delta \Rightarrow |g(x) - g(a)| < \varepsilon)$$

e

$$\forall \varepsilon > 0\ \exists \delta' > 0\ (0 < |g(x) - g(a)| < \delta' \Rightarrow |f(g(x)) - f(g(a))| < \varepsilon)$$

Portanto,

$$\forall \varepsilon > 0\ \exists \delta > 0\ (0 < |x - a)| < \delta \Rightarrow |f(g(x)) - f(g(a))| < \varepsilon).$$

Seguindo estratégia semelhante àquela empregada para diferenciabilidade em intervalos fechados, podemos estender o conceito de continuidade para intervalos fechados não degenerados.

DEFINIÇÃO 5.24. *Uma função real f é* contínua em um intervalo fechado $[a, b]$ *de números reais sss*

I: *f é contínua no intervalo aberto (a, b);*

II: $\lim_{x \to a^+} f(x) = f(a)$*;*

III: $\lim_{x \to b^-} f(x) = f(b)$*;*

Teorema 5.23, sobre composição de funções contínuas, pode ser estendido para funções contínuas em intervalos fechados.

O próximo resultado mostra que derivada de uma função em um dado ponto real qualquer pode ser definida de maneira diferente, porém equivalente.

TEOREMA 5.24. *Se existe $\frac{d}{dx} f(x)\Big|_{x=a}$, então*

$$\frac{d}{dx} f(x)\Big|_{x=a} = \lim_{x \to a} \frac{f(x) - f(a)}{x - a}.$$

DEMONSTRAÇÃO: Basta fazer $h = x - a$. Logo, $x = a + h$ e

$$\lim_{x \to a} \frac{f(x) - f(a)}{x - a} = \lim_{x \to a} \frac{f(a + h) - f(a)}{h}.$$

Mas, se x tende a a, isso equivale a afirmar que

$$x \in (a - \delta, a + \delta) - \{a\}.$$

Logo, $x - a \in (0 - \delta, 0 + \delta) - \{0\}$, o que é equivalente a $h \in (0 - \delta, 0 + \delta) - \{0\}$. Logo,

$$\lim_{x \to a} \frac{f(a + h) - f(a)}{h} = \lim_{h \to 0} \frac{f(a + h) - f(a)}{h} = \frac{d}{dx} f(x) \bigg|_{x=a}.$$

A transitividade da igualdade encerra a prova.

O próximo teorema é de extraordinária importância.

Teorema 5.25. *Toda função diferenciável é contínua.*

Demonstração: Se f é diferenciável, então existe a função derivada f' de f. Mas, de acordo com o Teorema 5.24,

$$f'(a) = \lim_{x \to a} \frac{f(x) - f(a)}{x - a}$$

para todo a pertencente ao domínio de f. No entanto,

$$\lim_{x \to a} \left(\frac{f(x) - f(a)}{x - a} (x - a) \right) =$$

$$\lim_{x \to a} (f(x) - f(a)) = \lim_{x \to a} f(x) - \lim_{x \to a} f(a),$$

uma vez que $x \neq a$. Além disso,

$$\lim_{x \to a} \left(\frac{f(x) - f(a)}{x - a} (x - a) \right) =$$

$$\lim_{x \to a} \frac{f(x) - f(a)}{x - a} \cdot \lim_{x \to a} (x - a) = f'(a) \cdot 0 = 0.$$

A transitividade da igualdade garante que

$$\lim_{x \to a} f(x) - \lim_{x \to a} f(a) = 0.$$

Logo,

$$\lim_{x \to a} f(x) = \lim_{x \to a} f(a) = f(a).$$

para todo a pertencente ao domínio de f, o que equivale a afirmar que f é contínua. Observar que a igualdade $\lim_{x \to a} f(a) = f(a)$ é consequência do Teorema 5.9 sobre limite de função constante.

A recíproca do último teorema não é teorema. Isso significa que existem funções contínuas não diferenciáveis.

Exemplo 5.26. *Seja $f : \mathbb{R} \to \mathbb{R}$ uma função dada por*

$$f(x) = |x|.$$

Para qualquer $a \in \mathbb{R}$ temos que

$$\lim_{x \to a} f(x) = f(a)$$

(✍ recomendamos que o leitor prove isso, dividindo a demonstração em três partes: para $a < 0$, para $a > 0$ e, finalmente, para $a = 0$). No entanto, apesar de existir

$$\frac{d}{dx} f(x)\Big|_{x=a}$$

para qualquer $a \neq 0$, não existe L real tal que

$$L = \frac{d}{dx} f(x)\Big|_{x=0}.$$

Logo, f é não diferenciável.

Este último exemplo revela que é impossível uma reta tangenciar a função f no ponto $(0, 0)$.

Teoremas 5.25 e 5.24 revelam algo de extraordinária importância:

Continuidade é uma condição essencial para diferenciabilidade.

Em virtude disso, temos o seguinte resultado.

Teorema 5.26. *Seja*

$$f : d \to \mathbb{R}$$

uma função tal que $f(a)$ existe e o ponto a pertence a algum intervalo aberto contido em d. Logo, f é diferenciável no ponto a sss existe uma função φ_a contínua no ponto a tal que

$$f(x) - f(a) = \varphi_a(x)(x - a).$$

Demonstração: De acordo com o Teorema 5.24, φ_a é uma função que aproxima continuamente os coeficientes angulares de retas que secam f nos pontos $(x, f(x))$ e $(a, f(a))$ (para $x \neq a$) do coeficiente angular da reta que tangencia f no ponto $(a, f(a))$.

Logo, φ_a é única para cada a, tal que $\varphi_a(a) = f'(a)$.

A primeira pessoa a perceber esse último resultado foi Constantin Carathéodory [33], em 1954. Carathéodory chegou a propor o teorema acima como definição para função diferenciável, apesar de aqui preferirmos a definição usual. No entanto, esse resultado simplifica consideravelmente certas demonstrações de cálculo diferencial e integral (como ocorre no Teorema 6.3, o qual é discutido mais adiante e permite calcular rapidamente a derivada de funções compostas). Observar também que a função φ_a acima mencionada depende do ponto a. Em outras palavras, para cada a existe uma φ_a.

Teorema 5.27. *Se u e v compartilham o mesmo domínio e são diferenciáveis, então*

$$(uv)' = u'v + uv'.$$

Demonstração:

$$(uv)' = \lim_{h\to 0} \frac{(uv)(x+h) - (uv)(x)}{h} =$$

$$\lim_{h\to 0} \frac{u(x+h)v(x+h) - u(x)v(x)}{h} =$$

$$\lim_{h\to 0} \frac{u(x+h)v(x+h) - u(x)v(x+h) + u(x)v(x+h) - u(x)v(x)}{h}.$$

Mas este último é igual a

$$\lim_{h\to 0} \frac{(u(x+h) - u(x))v(x+h) + u(x)(v(x+h) - v(x))}{h} =$$

$$\lim_{h\to 0} \frac{(u(x+h) - u(x))v(x+h)}{h} + \lim_{h\to 0} \frac{u(x)(v(x+h) - v(x))}{h} =$$

$$\lim_{h\to 0} \frac{u(x+h) - u(x)}{h} \cdot \lim_{h\to 0} v(x+h) + \lim_{h\to 0} u(x) \cdot \lim_{h\to 0} \frac{v(x+h) - v(x)}{h} =$$

$$u'(x)v(x) + u(x)v'(x),$$

uma vez que $\lim_{h\to 0} v(x+h) = v(x)$, por conta do fato de que toda função diferenciável é contínua (Teorema 5.25) e

$$\lim_{h\to 0} u(x) = u(x)$$

por conta do Teorema 5.9 sobre limite de função constante.

✍ Uma sugestão divertida de exercício é provar que

$$(uvw)' = u'vw + uv'w + uvw',$$

se u, v e w são funções reais diferenciáveis.

Obviamente fica muito fácil demonstrar o resultado acima se o leitor usar o Teorema 5.27. Consegue estender esse resultado para uma multiplicação entre n funções diferenciáveis, usando indução infinita para provar o teorema proposto? Esta é uma ótima oportunidade para perceber a versatilidade do Teorema 5.27.

TEOREMA 5.28. *Se u e v compartilham o mesmo domínio e são diferenciáveis, então*

$$\left(\frac{u}{v}\right)' = \frac{u'v - uv'}{v^2},$$

desde que $v(x) \neq 0$ para todo x pertencente ao domínio de v.

DEMONSTRAÇÃO:

$$\left(\frac{u}{v}\right)' = \lim_{h\to 0} \frac{\frac{u}{v}(x+h) - \frac{u}{v}(x)}{h} = \lim_{h\to 0} \frac{\frac{u(x+h)}{v(x+h)} - \frac{u(x)}{v(x)}}{h} =$$

$$\lim_{h\to 0} \frac{u(x+h)v(x) - u(x)v(x+h)}{h \cdot v(x) \cdot v(x+h)} =$$

$$\lim_{h\to 0} \frac{u(x+h)v(x) - u(x)v(x+h)}{h} \cdot \lim_{h\to 0} \frac{1}{v(x) \cdot v(x+h)}.$$

Temos assim um produto entre duas ocorrências de limites, onde o primeiro pode ser reescrito como

$$\lim_{h\to 0} \frac{u(x+h)v(x) - u(x)v(x) + u(x)v(x) - u(x)v(x+h)}{h}.$$

Daqui em diante a prova é muito semelhante com o que foi feito na demonstração do Teorema 5.27. ✍ Sugerimos que o leitor termine.

EXEMPLO 5.27. *Seja $f : \mathbb{R} \to \mathbb{R}$ tal que $f(x) = 5x^3 - 18x^2 + 16$. Então, $f'(x) = 15x^2 - 36x$, sendo f' uma função real com domínio $\mathbb{R}$.*

Teorema 5.28 estende o alcance do Teorema 5.19.

TEOREMA 5.29. *Seja $f : \mathbb{R} - \{0\} \to \mathbb{R}$ uma função dada por $f(x) = x^n$, onde n é um real que copia um inteiro negativo. Então a função derivada f' de f existe e é tal que*

$$f'(x) = nx^{n-1}.$$

Demonstração: Se $f(x) = x^n$ (onde n é uma cópia de um inteiro negativo), então

$$f(x) = \frac{1}{x^{-n}},$$

onde $-n$ é um real que copia um inteiro estritamente positivo. Logo, de acordo com Teoremas 5.28 e 5.19,

$$f'(x) = \frac{0 \cdot x^{-n} - 1 \cdot (-n) \cdot x^{-n-1}}{(x^{-n})^2} =$$

$$\frac{-1 \cdot (-n) \cdot x^{-n-1}}{x^{-2n}} = n \cdot x^{-n-1+2n}.$$

Consequentemente, $f'(x) = nx^{n-1}$.

Agora sabemos que, em notação abreviada,

$$(x^n)' = nx^{n-1}$$

para qualquer n real que copia um inteiro diferente de 0, independentemente de n ser estritamente positivo ou negativo. Esse resultado pode ser estendido ainda mais, como vemos no Teorema 5.31. Mas antes precisamos de um resultado preliminar.

Teorema 5.30. *Seja* $f : \mathbb{R} - \{0\} \to \mathbb{R}$ *uma função dada por* $f(x) = \sqrt[q]{x}$. *Logo,*

$$\lim_{x \to a} f(x) = \sqrt[q]{a}.$$

Demonstração: Se estamos assumindo que f é uma função, naturalmente estamos excluindo a possibilidade de q ser um inteiro par. Para obter resultado análogo no caso de q par, basta assumir

$$f : \{x \in \mathbb{R} \mid x > 0\} \to \mathbb{R}.$$

Supor que

$$\lim_{x \to a} \sqrt[q]{x} \neq \sqrt[q]{\lim_{x \to a} x},$$

ou seja, a negação da tese. Logo,

$$a = \lim_{x \to a} x = \lim_{x \to a} (\sqrt[q]{x})^q = \left(\lim_{x \to a} \sqrt[q]{x}\right)^q \neq$$

$$\left(\sqrt[q]{\lim_{x \to a} x}\right)^q = \lim_{x \to a} x = a, \text{ ou seja, } a \neq a. \perp$$

Teorema 5.31. *Seja $f : \mathbb{R} - \{0\} \to \mathbb{R}$ uma função dada por $f(x) = x^n$, onde n é um real que copia um racional diferente de 0. Então a função derivada f' de f existe e é tal que $f'(x) = nx^{n-1}$.*

Demonstração: Se estamos assumindo que f é uma função, naturalmente estamos excluindo a possibilidade de $n = \frac{p}{q}$ onde q é um real que copia um inteiro par. Para obter resultado análogo no caso de q par, basta assumir

$$f : \{x \in \mathbb{R} \mid x > 0\} \to \mathbb{R}.$$

De acordo com o Teorema 5.24,

$$\frac{d}{dx} x^n \bigg|_{x=a} = \lim_{x \to a} \frac{x^n - a^n}{x - a}.$$

Se $n = \frac{p}{q}$, onde p e q são reais que copiam inteiros tais que $q \neq 0$, então

$$\lim_{x \to a} \frac{x^n - a^n}{x - a} = \lim_{x \to a} \frac{x^{\frac{p}{q}} - a^{\frac{p}{q}}}{x - a} =$$

$$\lim_{x \to a} \left(\frac{x^{\frac{p}{q}} - a^{\frac{p}{q}}}{x - a} \cdot \frac{x^{\frac{1}{q}} - a^{\frac{1}{q}}}{x^{\frac{1}{q}} - a^{\frac{1}{q}}} \right) =$$

$$\lim_{x \to a} \left(\frac{x^{\frac{p}{q}} - a^{\frac{p}{q}}}{x^{\frac{1}{q}} - a^{\frac{1}{q}}} \cdot \frac{x^{\frac{1}{q}} - a^{\frac{1}{q}}}{x - a} \right).$$

Se fizermos $y = x^{\frac{1}{q}}$ e $b = a^{\frac{1}{q}}$ (ver Teorema 5.30, uma vez que este prova que $x^{\frac{1}{q}}$ é contínua), então o último limite é igual a

$$\lim_{y \to b} \left(\frac{y^p - b^p}{y - b} \cdot \frac{y - b}{y^q - b^q} \right) = \lim_{y \to b} \frac{y^p - b^p}{y - b} \cdot \lim_{y \to b} \frac{y - b}{y^q - b^q} =$$

$$\lim_{y \to b} \frac{y^p - b^p}{y - b} \Bigg/ \lim_{y \to b} \frac{y^q - b^q}{y - b}.$$

Mas no último termo acima temos exatamente a derivada de y^p em relação a y no ponto b (onde p é um inteiro e, portanto, podemos aplicar Teoremas 5.19 e 5.29), bem como a derivada de y^q em relação a y no ponto b (onde q é novamente um inteiro estritamente positivo ou negativo), de acordo com Teorema 5.24.

Logo, este último termo é

$$pb^{p-1}/(qb^{q-1}) = \frac{p}{q}b^{p-q} = na^{\frac{1}{q}(p-q)} = na^{\frac{p}{q}-1} = na^{n-1}.$$

Uma vez que isso vale para todo a do domínio de f, então

$$(x^n)' = nx^{n-1}.$$

Mais adiante, na Seção 66, é possível provar que tal resultado pode ser estendido para qualquer n real, desde que x seja estritamente positivo.

Função derivada $f'(x)$ de uma $f(x)$ é também conhecida como *derivada primeira* de $f(x)$. Derivada em relação a x *de ordem* $n+1$ de uma função real $f(x)$, se existir, é definida como

$$\frac{d^{n+1}}{dx^{n+1}}f(x) = \frac{d}{dx}\left(\frac{d^n}{dx^n}f(x)\right),$$

uma vez que já definimos derivada primeira.

Exemplo 5.28. i:

$$\frac{d^2}{dx^2}(x^3 - 2x^2 + 6x) = \frac{d}{dx}(3x^2 - 4x + 6) = 6x - 4;$$

ii:

$$\frac{d^3}{dx^3}(x^3 - 2x^2 + 6x) = \frac{d^2}{dx^2}(3x^2 - 4x + 6) =$$

$$\frac{d}{dx}(6x - 4) = 6.$$

Naturalmente, estamos assumindo que $x^3 - 2x^2 + 6x$ *abrevia uma função polinomial com domínio* $\mathbb{R}$.

Comumente derivadas de segunda e terceira ordem são denotadas por f'' e f''', respectivamente. Derivadas de ordem n, quando existem, são também denotadas como $f^{(n)}$.

Exemplo 5.29. *Se* $s : \mathbb{R} \to \mathbb{R}$ *é uma função tal que*

$$s(t) = -5t^2 + 20,$$

então $s'(t) = -10t$ *e* $s''(t) = -10$.

Com os devidos cuidados, s *pode ser interpretada como uma função que, localmente, descreve queda livre de um objeto próximo à superfície da Terra. Com efeito, basta interpretar* t *como tempo* *em segundos e $s(t)$ como a altura em metros em que o objeto se encontra relativamente ao solo.*

Se assumirmos que, a partir de um estado de repouso relativamente ao solo, o objeto é abandonado em queda livre no instante 0 segundo,

$$s(0) = 20$$

informa que esse objeto foi abandonado a vinte metros do solo. No mesmo instante $t = 0$, a velocidade $s'(0)$ em metros por segundo é zero.

Mas, na medida em que o tempo passa, a velocidade aumenta em valor absoluto. No instante $t = 1$, por exemplo, o objeto está a 15 metros do solo e com velocidade de -10 metros por segundo (cujo valor absoluto é 10 metros por segundo). O sinal negativo da velocidade indica a rota de colisão em direção ao ponto zero, o solo. Isso porque está em queda livre (sem resistência do ar ou outros agentes físicos).

Durante todo o tempo de queda, a aceleração $s''(t)$ é constante, no valor de menos dez metros por segundo por segundo. Mas este mapeamento da queda livre pela função s só é possível no intervalo aberto $(0, 2)$, uma vez que no instante 2 segundos o objeto atinge o solo. Afinal, $s(2) = 0$.

A partir da Seção 54 discutimos sobre funções circulares (seno, co-seno, secante, co-secante, tangente e cotangente), as quais não são polinomiais ou reais racionais. Antes, porém, é conveniente conhecermos um pouco sobre sequências reais e séries.

Seção 50

Sequências e séries

Sumário

Índice

Rede

De agora em diante, sempre que falarmos de sequências, estamos nos referindo a sequências reais (a não ser que seja avisado o contrário), ou seja, funções x cujas imagens x_n são números reais e cujos domínios são ω ou $\omega - \{0\}$.

Por abuso de notação estamos adotando a notação ω para designar os reais que copiam naturais. Neste primeiro momento trabalhamos com sequências cujos domínios são $\omega - \{0\}$, onde 0 é o neutro aditivo dos reais.

Definição 5.25. *Seja x uma sequência cujas imagens são x_n. A* soma parcial S_n *de x é definida como*

$$S_n = x_1 + x_2 + \cdots + x_n = \sum_{i=1}^{n} x_i.$$

Ou seja, toda soma parcial de uma sequência x é a soma das primeiras n imagens de x.

Deve ficar claro ao leitor que a definição de soma parcial é possível por conta da relação de ordem total $\leq$ entre os naturais, além do fato de que qualquer conjunto de números naturais admite menor elemento relativamente a $\leq$. Graças a isso é possível qualificar o que são as primeiras n imagens.

Exemplo 5.30. i: *se $x_n = 7$, então*

$$S_4 = 7 + 7 + 7 + 7 = 28.$$

ii: *se $x_n = \frac{1}{n}$, então*

$$S_5 = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5},$$

ou seja, $S_5 = \frac{137}{60}$. Em notação decimal, $S_5 = 2,2833333\cdots$;

iii: *se $x_n = \frac{1}{n}$, então*

$$S_{10^6} = S_{1\,000\,000} = 14,3927267\cdots;$$

além disso,

$$S_{10^{43}} < 100.$$

Consegue provar a última desigualdade?

Definição 5.26. *Dada uma sequência x_n, a* série

$$\sum x_n$$

é a sequência de somas parciais S_n de x_n.

Exemplo 5.31. i: *se $x_n = 7$, a série $\sum x_n$ é a sequência cujas imagens são 7, 14, 21, 28, $\cdots$; em outras palavras,*

$$\sum x_n = \{(1,7),(2,14),(3,21),(4,28),\cdots\};$$

ii: *se* $x_n = \frac{1}{n}$*, então a série* $\sum \frac{1}{n}$ *é uma sequência cujas primeiras imagens são*

$$1,\ \frac{3}{2},\ \frac{11}{6},\ \frac{50}{24},\ \frac{137}{60},\ \dots .$$

É imprescindível que o leitor não confunda somatório

$$\sum_{i=0}^{n} x_i$$

como uma série

$$\sum x_n.$$

No primeiro caso, temos uma soma de $n+1$ termos. No segundo caso temos uma sequência de somas parciais de uma dada sequência.

Intuitivamente falando, uma série opera como uma 'soma de infinitas parcelas'. No entanto, uma vez que adição de reais é uma operação binária, não é possível definir qualquer soma que envolva quantias não finitas de termos (apesar da associatividade da adição de reais). Para contornar essa dificuldade, introduz-se os conceitos de soma parcial (toda soma parcial é um somatório) e de sequência de somas parciais (funções com domínio ω ou $\omega - \{0\}$).

Definição 5.27. *Dada uma sequência* x_n*, sua série correspondente*

$$\sum x_n$$

converge *sss a sequência de somas parciais* S_n *de* x_n *converge. Caso contrário, dizemos que a série* diverge.

A série

$$\sum \frac{1}{n}$$

é chamada de *série harmônica*.

Como foi ilustrado no Exemplo 5.30, se somarmos as primeiras 10^{43} imagens (dez milhões de quintilhões de quintilhões imagens) da sequência $x_n = \frac{1}{n}$, sequer alcançamos a soma 100. Logo, é natural questionar se a série harmônica converge, ou seja, se há um limite para a sequência de somas parciais associadas a $x_n = \frac{1}{n}$. Como se percebe no próximo teorema, esse limite não existe.

Teorema 5.32. *A série harmônica diverge.*

Demonstração: A soma parcial S_n de $\frac{1}{n}$ é dada por

$$1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n}.$$

A soma parcial S_{2n} é dada por

$$1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} + \frac{1}{n+1} + \frac{1}{n+2} + \cdots + \frac{1}{2n}.$$

Logo,

$$S_{2n} - S_n = \frac{1}{n+1} + \frac{1}{n+2} + \cdots + \frac{1}{2n}.$$

Mas,

$$\frac{1}{n+1} + \frac{1}{n+2} + \cdots + \frac{1}{2n} > \frac{1}{2n} + \frac{1}{2n} + \cdots + \frac{1}{2n},$$

desde que no lado direito da última desigualdade existam n ocorrências de $\frac{1}{2n}$. No entanto,

$$\frac{1}{2n} + \frac{1}{2n} + \cdots + \frac{1}{2n} = n\frac{1}{2n} = \frac{1}{2}.$$

Logo, $S_{2n} - S_n > \frac{1}{2}$.

Lembremos agora que uma sequência S_n é de Cauchy sss

$$\forall \varepsilon > 0\ \exists \delta > 0((m > \delta \wedge n > \delta) \Rightarrow |S_m - S_n| < \varepsilon).$$

Para o caso em que $m = 2n$ e $\varepsilon = \frac{1}{2}$, temos que a condicional da definição de sequência de Cauchy jamais é satisfeita para qualquer δ real maior do que 0, uma vez que

$$|S_{2n} - S_n| > \frac{1}{2}.$$

Logo, a sequência de somas parciais S_n não é de Cauchy. Uma vez que toda sequência convergente é de Cauchy, então S_n não é convergente.

Esse último teorema, em parceria com o resultado $S_{10^{43}} < 100$, mostra que a divergência da série harmônica é muito lenta. Mas, a passos muito lentos,

$$\sum_{i=1}^{n} \frac{1}{i}$$

sempre ultrapassa qualquer número real. Basta escolher n suficientemente grande.

A contrapositiva do próximo teorema nos fornece uma condição suficiente para uma série ser divergente.

Teorema 5.33. *Se $\sum x_n$ converge, então $x_n \to 0$.*

Demonstração: Se $\sum x_n$ converge, então existe real L tal que

$$\lim_{n\to\infty} \sum_{i=0}^{n} x_i = L$$

Mas

$$\lim_{n\to\infty} \sum_{i=0}^{n} x_i = \lim_{n\to\infty} \left(x_n + \sum_{i=0}^{n-1} x_i \right) = \lim_{n\to\infty} x_n + \lim_{n\to\infty} \sum_{i=0}^{n-1} x_i.$$

Uma vez que

$$\lim_{n\to\infty} \sum_{i=0}^{n-1} x_i = L,$$

então

$$L = \lim_{n\to\infty} x_n + L.$$

Logo, $\lim_{n\to\infty} x_n = 0$.

Ou seja, se x_n não converge para zero, então $\sum x_n$ diverge.

Como vimos acima, a recíproca deste último teorema não é teorema. Com efeito, $\frac{1}{n} \to 0$, mas

$$\sum \frac{1}{n}$$

diverge.

Mesmo assim, Teorema 5.33 oferece uma condição necessária para uma série ser convergente, apesar de não ser suficiente.

Na literatura de cálculo diferencial e integral padrão e análise matemática há discussões pormenorizadas sobre sequências e séries reais, incluindo outros critérios de convergência de séries, além do Teorema 5.33. Para finalizar esta Seção, mencionamos um desses critérios, o qual é usado mais adiante.

Teorema 5.34. *Seja $\sum x_n$ a série real correspondente à sequência real x_n. Logo, se*

$$\lim_{n\to\infty} \left| \frac{x_{n+1}}{x_n} \right| < 1$$

então $\sum x_n$ é convergente.

Mais do que isso, é teorema que $\sum |x_n|$ converge se

$$\lim_{n\to\infty}\left|\frac{x_{n+1}}{x_n}\right| < 1.$$

A prova desse resultado pode ser encontrada em textos de análise matemática. Não é uma demonstração difícil, mas foge dos propósitos deste livro.

SEÇÃO 51

Resumo da ópera

SUMÁRIO
ÍNDICE
REDE

Esta quinta parte pode ser resumida como se segue.

- Números reais são classes de equivalência de sequências racionais de Cauchy. Se um representante qualquer de um real r é uma sequência de Cauchy convergente, r é um real racional (não confundir com racional). Se um representante qualquer de um real r é uma sequência de Cauchy não convergente, r é irracional.
- Entre os reais há operações de adição e multiplicação que preservam as propriedades algébricas de adição e multiplicação entre racionais. No entanto, os reais têm uma propriedade algébrica que não conta com equivalente entre racionais: toda sequência de Cauchy de reais é convergente.
- Complexos podem ser definidos como pares ordenados de reais. Entre os complexos há operações de adição e multiplicação que preservam as propriedades algébricas de adição e multiplicação entre reais. No entanto, os complexos contam com uma propriedade algébrica que não ocorre entre reais: a existência de um termo cujo quadrado é o simétrico aditivo do neutro multiplicativo.
- Uma vez que a linguagem $\mathfrak{S}$ de ZF finalmente permite qualificar naturais, inteiros, racionais, reais e complexos, estamos pronto para iniciar estudos de cálculo diferencial e integral padrão. Iniciamos isso com os conceitos de limite e derivada.

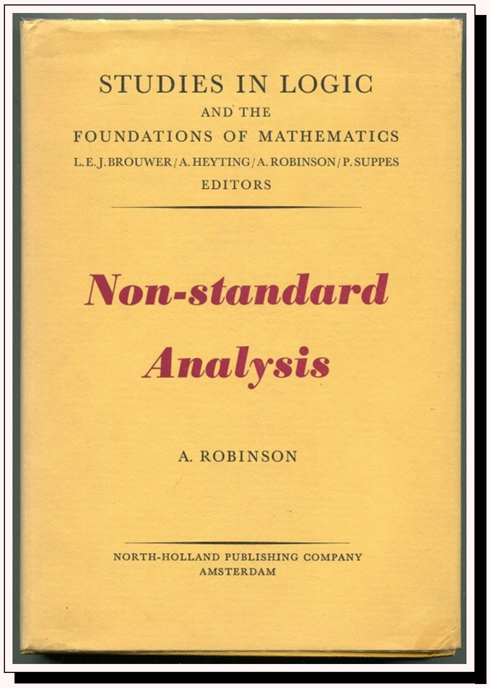
- Derivadas são casos especiais de limites, no contexto do cálculo padrão.
- As definições de limite e derivada de uma função real não são amigáveis para fins de cálculos. Por conta disso, os teoremas sobre limites e derivadas são indispensáveis para compreender o caráter epistemológico e metodológico do cálculo padrão.

Seção 52

Notas históricas

Sumário
Índice
Rede

O que define cálculo diferencial e integral é o Teorema Fundamental do Cálculo, assunto a ser discutido adiante. Neste sentido, as primeiras ideias intuitivas sobre cálculo diferencial e integral padrão nasceram com a obra de Isaac Newton, no século 17.



Capa da primeira edição do livro de Abraham Robinson

Fonte: Evening Star Books.

O movimento de análise matemática surgido no século 19 viabilizou os conceitos hoje estudados sobre limites e derivadas. Paralelamente ao trabalho de Newton, Gottfried Leibniz desenvolveu, de forma independente, ideias semelhantes, porém enfatizando o papel de *infinitesimais* para conceituar derivadas. Um infinitesimal ς é um termo estritamente positivo menor do que qualquer real estritamente positivo. Além disso, podem existir infinitesimais ς tais que $\varsigma > 0$ e $\varsigma^2 = 0$. Logo, infinitesimais não podem ser números reais.

Somente no século 20, graças ao trabalho de Abraham Robinson, infinitesimais foram formalizados de maneira clara, dando origem à *análise não standard*. Hoje se sabe que infinitesimais são casos particulares de hiperreais e surreais, os quais estendem os números reais. Tanto hiperreais quanto surreais conseguem copiar os reais.

PARTE 6

Funções circulares, exponenciais e logarítmicas

Aqui exploramos certas funções transcendentes, as quais são funções reais não racionais. Mas, antes, precisamos saber mais sobre sequências, séries e derivadas.

Seção 53

Equações diferenciais

Sumário
Índice
Rede

Um conceito importante é o de *operador diferencial*. Não pretendemos conceituar operadores diferenciais. Mas um caso particular perfeitamente útil para os nossos propósitos é o que se segue.

Seja $\mathfrak{f}$ um conjunto de funções $f : \mathbb{R} \to \mathbb{R}$ tais que, cada f admite derivada de ordem n. Logo, a função $\mathfrak{D} : \mathfrak{f} \to \mathfrak{f}$ dada por

$$\mathfrak{D}(f) = \sum_{i=0}^{n} \alpha_i f^{(i)}$$

é um *operador diferencial* definido sobre $\mathfrak{f}$, onde cada α_i é um número real e cada $f^{(i)}$ é uma derivada de f, em relação a x, de ordem i, se $1 \leq i \leq n$, e $f^{(0)} = f$.

Uma definição mais ampla deveria assumir que cada α_i é uma função

$$\alpha_i : \mathbb{R} \to \mathbb{R}.$$

Não obstante, isso ainda não cobre o espectro de todos os possíveis operadores diferenciais. A definição dada acima já é o bastante para os nossos propósitos.

Uma vez que f admite derivada de ordem n, naturalmente admite qualquer derivada de ordem i, desde que i seja menor ou igual a n. Isso é consequência imediata da definição de derivada de ordem superior.

Exemplo 6.1. *Seja $\mathfrak{f}$ o conjunto de todas as funções reais polinomiais*

$$p : \mathbb{R} \to \mathbb{R}.$$

Logo, $\mathfrak{D} : \mathfrak{f} \to \mathfrak{f}$, dada por

$$\mathfrak{D}(f) = f',$$

é um operador diferencial

$$\sum_{i=0}^{n} \alpha_i f^{(i)},$$

onde $n = 1$, $\alpha_0 = 0$ e $\alpha_1 = 1$. Isso implica que a derivada primeira de qualquer função polinomial é um operador diferencial.

Exemplo 6.2. *Seja $\mathfrak{f}$ o conjunto de todas as funções reais*

$$p : \mathbb{R} \to \mathbb{R}$$

que admitem derivada terceira. Logo,

$$\mathfrak{E} : \mathfrak{f} \to \mathfrak{f},$$

dada por

$$\mathfrak{E}(f) = 5f''' - 2f'' + f',$$

é um operador diferencial

$$\sum_{i=0}^{n} \alpha_i f^{(i)},$$

onde $n = 3$, $\alpha_0 = 0$, $\alpha_1 = 1$, $\alpha_2 = -2$ e $\alpha_3 = 5$. Isso implica que a derivada primeira de qualquer função de $\mathfrak{f}$ subtraída do dobro de sua derivada segunda e somada do quíntuplo de sua derivada terceira é um operador diferencial.

Exemplo 6.3. *Seja* $\mathfrak{f}$ *o conjunto de todas as funções reais polinomiais* $p : \mathbb{R} \to \mathbb{R}$. *Logo,*

$$\mathfrak{H} : \mathfrak{f} \to \mathfrak{f},$$

dada por

$$\mathfrak{H}(f) = 0,$$

é um operador diferencial

$$\sum_{i=0}^{n} \alpha_i f^{(i)},$$

onde $n = 1$, $\alpha_0 = 0$ *e* $\alpha_1 = 0$. *Logo, a função constante identicamente nula no espaço de funções polinomais, é um operador diferencial.*

Deve ficar claro que cada parcela de um operador diferencial

$$\mathfrak{D}(f) = \sum_{i=1}^{n} \alpha_i f^{(i)}$$

também define um operador diferencial. Além disso, cada $f^{(i)}$ é um operador diferencial, para $i \geq 1$. O somatório acima é chamado de *combinação linear* dos operadores diferenciais $f^{(i)}$. Logo, qualquer operador diferencial é uma combinação linear de operadores diferenciais.

✍ É teorema fácil de provar que, se $\mathfrak{D}$ é um operador diferencial, então

$$\mathfrak{D}(f + g) = \mathfrak{D}(f) + \mathfrak{D}(g)$$

e, além disso,

$$\mathfrak{D}(cf) = c\mathfrak{D}(f),$$

onde c é uma constante real.

Agora podemos definir o que é uma equação diferencial, pelo menos para os nossos modestos propósitos neste livro. Uma *equação diferencial* é uma equação $u = v$ onde ocorre pelo menos um operador diferencial em u ou v. A rigor, estamos tratando aqui apenas de equações diferenciais definidas sobre espaços de funções reais.

Exemplo 6.4. *Seja* $y : \mathbb{R} \to \mathbb{R}$ *uma função que admite derivada de qualquer ordem. Logo,*

$$y'' + y = 0$$

é uma equação diferencial. Com efeito, essa equação é equivalente a

$$\mathfrak{D}(y) = 0,$$

onde $\mathfrak{D} : \mathfrak{f} \to \mathfrak{f}$ *é o operador diferencial dado por*

$$\mathfrak{D}(y) = y'' - y.$$

Essa equação diferencial em especial é de grande interesse, como se percebe na Seção 54.

O estudo e a aplicação de equações diferenciais é a principal meta do cálculo diferencial e integral.

Seção 54

Séries de potências

Sumário
Índice
Rede

Uma função f é chamada de *classe* $\mathcal{C}^\infty$ sss f admite derivada de ordem n, para qualquer n inteiro estritamente positivo.

Seja $f : \mathbb{R} \to \mathbb{R}$ uma função tal que

$$f(x) = \sum a_n x^n.$$

Podemos representar f da seguinte maneira:

$$f(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \cdots + a_n x^n + \cdots ,$$

onde assumimos a convenção $x^0 = 1$ para qualquer real x.

Funções como essas são chamadas de *séries de potências*. Ou seja, funções f definidas por séries de potências são aquelas em que, cada x do domínio de f, admite uma imagem $f(x)$ dada pela série acima.

É usual escrever séries de potências como

$$\sum_{n=0}^{\infty} a_n x^n$$

ou

$$\sum_{n=k}^{\infty} a_n x^n,$$

sendo k um inteiro positivo.

Teorema 6.1. *Toda função definida por uma série de potências é de classe* $\mathcal{C}^\infty$.

Demonstração: Cada soma parcial S_n de $\sum a_n x^n$ é uma função polinomial. Uma vez que toda função polinomial admite derivada de qualquer ordem, então $\sum a_n x^n$ admite derivada de qualquer ordem.

Adotamos aqui o emprego de séries de potências para representar certas funções. Lembrar que, para todo x real, temos

$$a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \cdots + a_n x^n = b_0 + b_1 x + b_2 x^2 + b_3 x^3 + \cdots + b_n x^n$$

sss $a_i = b_i$ para todo i de 0 a n (ver discussão na Seção 43). Uma vez que séries de potências são sequências de somas parciais definidas por polinômios, então

$$\sum a_i x^i = \sum b_i x^i \ \text{ sss } \ a_i = b_i$$

para todo i natural.

Para fins de ilustração, consideremos a função $y : \mathbb{R} \to \mathbb{R}$ tal que

$$y''(x) + y(x) = 0$$

e

$$y(0) = 0$$

e

$$y'(0) = 1.$$

Esta função é conhecida como *seno*. Abreviamos $y(x)$ como $\operatorname{sen}(x)$, neste caso. Lemos $\operatorname{sen}(x)$ como 'seno de x'.

Se $y''(x) + y(x) = 0$, então $y''(x) = -y(x)$. Se existir série de potências para representar y, temos o seguinte:

$$y(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + a_4 x^4 + a_5 x^5 + a_6 x^6 + a_7 x^7 + \cdots$$

$$y'(x) = a_1 + 2a_2 x + 3a_3 x^2 + 4a_4 x^3 + 5a_5 x^4 + 6a_6 x^5 + 7a_7 x^6 + \cdots$$

$$y''(x) = 2a_2 + 3.2a_3 x + 4.3a_4 x^2 + 5.4a_5 x^3 + 6.5a_6 x^4 + 7.6a_7 x^5 + \cdots$$

sendo

$$-y(x) = -a_0 - a_1 x - a_2 x^2 - a_3 x^3 - a_4 x^4 - a_5 x^5 - a_6 x^6 - a_7 x^7 - \cdots$$

Logo,

$$a_2 = \frac{-a_0}{2},\ a_3 = \frac{-a_1}{3.2},\ a_4 = \frac{-a_2}{4.3} = \frac{a_0}{4.3.2},\ a_5 = \frac{-a_3}{5.4} = \frac{a_1}{5.4.3.2}, \cdots,$$

uma vez que os coeficientes dos monômios de grau m das somas parciais de $y''(x)$ devem ser iguais aos coeficientes dos monômios de grau m das somas parciais de $-y(x)$.

Ou seja,

$$a_2 = \frac{-a_0}{2},\ a_3 = \frac{-a_1}{3.2},\ a_4 = \frac{a_0}{4.3.2},$$

$$a_5 = \frac{a_1}{5.4.3.2},\ a_6 = \frac{-a_0}{6.5.4.3.2},\ a_7 = \frac{-a_1}{7.6.5.4.3.2}$$

e assim por diante.

Em outras palavras, a igualdade $y''(x) = -y(x)$ permite reduzir a infinidade de coeficientes a_m da série de potências de $y(x)$ a apenas dois coeficientes, a saber, a_0 e a_1. Conhecer os valores de a_0 e a_1 permite determinar todos os demais a_m.

Mas seno não é definida apenas por $y''(x) = -y(x)$. As condições de contorno também fazem parte da definição. Observar, por exemplo, que $y(0) = a_0$ e $y'(0) = a_1$. Logo, $a_0 = 0$ e $a_1 = 1$. Logo, $a_{\text{par}} = 0$. Além disso, cada $a_{\text{ímpar}}$ é diferente de 0, como se percebe a seguir:

$$a_1 = 1,\ a_3 = \frac{-1}{3!},\ a_5 = \frac{1}{5!},\ a_7 = \frac{-1}{7!},\ a_9 = \frac{1}{9!}$$

e assim por diante.

Basta agora substituir os valores dos coeficientes a_m na série de potências correspondente a $y(x)$. Portanto,

$$y(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots .$$

Ou seja,

$$\text{sen}(x) = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!},$$

sendo que $\sum_{n=0}^{\infty}$ denota apenas uma série $\sum$.

✍ A última fórmula pode ser facilmente demonstrada por indução infinita. Basta usar a definição de seno dada acima. Recomendamos como exercício, uma vez que a notação acima empregando reticências $(\cdots)$ não é uma prática matematicamente elegante.

Nessa discussão é imprescindível que o leitor perceba o seguinte: seno, por definição, é uma função $y : \mathbb{R} \to \mathbb{R}$ que satisfaz uma *equação diferencial sujeita a duas condições de contorno*. A equação diferencial é $y'' + y = 0$; as condições de contorno são as fórmulas $y(0) = 0$ e $y'(0) = 1$. Ou seja, seno é uma função y cuja derivada segunda é o simétrico aditivo de y, tal que seno de zero é zero e a derivada primeira de seno no ponto zero é um.

Para garantir a consistência do que fizemos até agora (admitindo que ZF é consistente), precisamos do próximo teorema.

Teorema 6.2.

$$\operatorname{sen}(x) = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}$$

converge para todo real x.

Demonstração: Aplicando Teorema 5.34, temos que

$$\lim_{n\to\infty} \left| \frac{(-1)^{n+1} \frac{x^{2(n+1)+1}}{(2(n+1)+1)!}}{(-1)^n \frac{x^{2n+1}}{(2n+1)!}} \right| = \lim_{n\to\infty} \frac{(2n+1)!}{(2n+3)!} |x|^2 =$$

$$\lim_{n\to\infty} \frac{1}{(2n+3)(2n+2)} |x|^2.$$

Mas, para qualquer x real, o último limite é 0, o qual é menor do que 1. Logo, Teorema 5.34 garante a convergência da série em questão.

Logo, a série de potências

$$\sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}$$

permite de fato definir uma função $y : \mathbb{R} \to \mathbb{R}$ que atende às condições impostas na definição de seno.

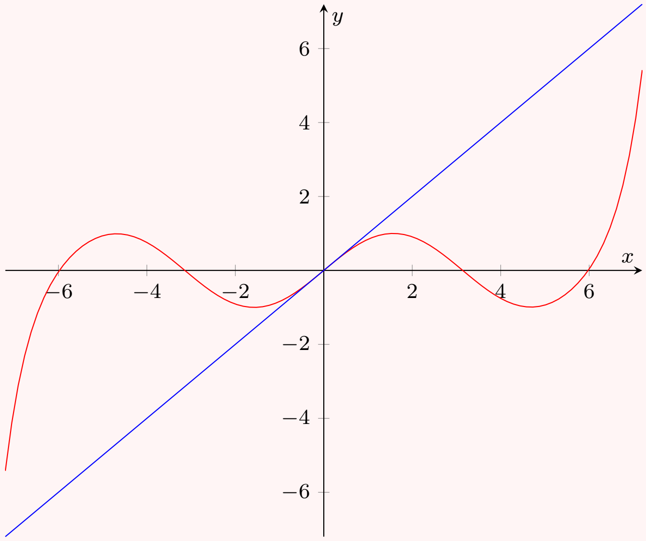
Para efeitos computacionais, é possível programar uma máquina para gerar aproximações de seno de um real x qualquer com a precisão desejada. Para isso basta truncar a série de potências acima. Na representação gráfica abaixo a função

$$f(x) = \sum_{n=0}^{0} (-1)^n \frac{x^{2n+1}}{(2n+1)!} = x,$$

em azul, é uma primeira aproximação de seno de x. A função

$$g(x) = \sum_{n=0}^{6} (-1)^n \frac{x^{2n+1}}{(2n+1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \frac{x^9}{9!} - \frac{x^{11}}{11!} + \frac{x^{13}}{13!},$$

em vermelho, é uma aproximação que trunca a série na sétima parcela.



A função *co-seno* é uma função $y : \mathbb{R} \to \mathbb{R}$ tal que

$$y''(x) + y(x) = 0,$$

$$y(0) = 1$$

e

$$y'(0) = 0.$$

Ou seja, co-seno de x, abreviada como $\cos(x)$, é definida a partir da mesma equação diferencial usada para conceituar seno. A diferença entre seno e co-seno reside única e exclusivamente nas condições de contorno $y(0)$ e $y'(0)$.

Usando as mesmas técnicas empregadas para representar seno por uma série de potências, é possível provar que

$$\cos(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \frac{x^8}{8!} - \cdots = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!},$$

onde $\cos(x) = y(x)$.

✍ É teorema a convergência da série acima. Recomendamos a prova deste resultado.

Uma vez que seno e co-seno podem ser representadas por sequências de somas parciais, onde cada soma parcial é um polinômio, então fica fácil demonstrar que

$$\frac{d}{dx}\operatorname{sen}(x) = \cos(x).$$

Analogamente, é fácil provar que

$$\frac{d}{dx}\cos(x) = -\text{sen}(x).$$

Por exemplo,

$$\frac{d}{dx}\text{sen}(x) = \frac{d}{dx}\left(x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots\right) =$$

$$1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots = \cos(x).$$

Uma vez que seno e co-seno foram definidas a partir da mesma equação diferencial, mudando apenas as condições de contorno, é natural questionar quais são as possíveis soluções para a mesma equação diferencial sob condições de contorno arbitrárias, não apenas aquelas usadas para definir seno e co-seno. Ou seja, considere uma função $y : \mathbb{R} \to \mathbb{R}$ tal que

$$y''(x) + y(x) = 0,$$

$$y(0) = \alpha$$

e

$$y'(0) = \beta,$$

onde α e β são reais quaisquer.

Se existir série de potências para representar y, temos o seguinte:

$$y(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 + a_5x^5 + \cdots$$

$$y'(x) = a_1 + 2a_2x + 3a_3x^2 + 4a_4x^3 + 5a_5x^4 + 6a_6x^5 + \cdots$$

$$y''(x) = 2a_2 + 3.2a_3x + 4.3a_4x^2 + 5.4a_5x^3 + 6.5a_6x^4 + 7.6a_7x^5 + \cdots$$

sendo

$$-y(x) = -a_0 - a_1x - a_2x^2 - a_3x^3 - a_4x^4 - a_5x^5 - a_6x^6 - a_7x^7 - \cdots$$

Logo,

$$a_2 = \frac{-a_0}{2},\ a_3 = \frac{-a_1}{3.2},\ a_4 = \frac{-a_2}{4.3} = \frac{a_0}{4.3.2},\ a_5 = \frac{-a_3}{5.4} = \frac{a_1}{5.4.3.2},\ \cdots.$$

Ou seja,

$$a_2 = \frac{-a_0}{2},\ a_3 = \frac{-a_1}{3.2},\ a_4 = \frac{a_0}{4.3.2},$$

$$a_5 = \frac{a_1}{5.4.3.2},\ a_6 = \frac{-a_0}{6.5.4.3.2},\ a_7 = \frac{-a_1}{7.6.5.4.3.2}$$

e assim por diante.

Mas $y(0) = \alpha$ e $y'(0) = \beta$. Logo, $a_0 = \alpha$ e $a_1 = \beta$. Logo,

$$y(x) = \alpha + \beta x - \frac{\alpha x^2}{2!} - \frac{\beta x^3}{3!} + \frac{\alpha x^4}{4!} + \frac{\beta x^5}{5!} - \frac{\alpha x^6}{6!} - \frac{\beta x^7}{7!} + \cdots$$

ou seja,

$$y(x) =$$

$$\alpha\left(1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \frac{x^8}{8!} - \cdots\right) + \beta\left(x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots\right)$$

Logo, $y(x) = \alpha\cos(x) + \beta\mathrm{sen}(x)$.

Observar que

$$y'(x) = -\alpha\mathrm{sen}(x) + \beta\cos(x).$$

Logo, $y(0) = \alpha$ e $y'(0) = \beta$.

A prova de que $y(x) = \alpha\cos(x) + \beta\mathrm{sen}(x)$ é solução da equação diferencial $y'' + y =$, sob condições de contorno $y(0) = \alpha$ e $y'(0) = \beta$, é de interesse no estudo de *álgebra linear*. Detalhes nas Seções 84 e 85.

Moral da história: A equação diferencial $y''(x) + y(x) = 0$ admite uma infinidade de soluções. Uma vez definidas as condições de contorno $y(0) = \alpha$ e $y'(0) = \beta$, temos uma única solução expressa por uma *combinação linear* de seno e co-seno (ou seja, a adição entre $\alpha\cos(x)$ e $\beta\mathrm{sen}(x)$). Em particular, se $\alpha = \beta = 0$, $y(x) = 0$. Logo, a função constante $y : \mathbb{R} \to \mathbb{R}$ dada por $y(x) = 0$ também é solução da equação diferencial dada, desde que as condições de contorno sejam $y(0) = 0$ e $y'(0) = 0$.

As funções seno e co-seno admitem interpretações geométricas no contexto de triângulos retângulos em geometria plana. Mas, para que sejamos capazes de contemplar esse fato, precisamos de algumas considerações dadas a seguir.

Seção 55

Derivada de composição de funções

Sumário
Índice
Rede

Apresentamos aqui uma poderosa técnica para cálculo de derivadas de composições não triviais de funções reais.

Teorema 6.3. *Sejam $u = u(v)$ e $v = v(x)$ funções reais diferenciáveis em relação a v e x, respectivamente, que admitem composição $u \circ v$. Logo, u é diferenciável em relação a x e*

$$\frac{d}{dx}u = \frac{d}{dv}u\frac{d}{dx}v.$$

Demonstração: Devemos provar o seguinte: se v é diferenciável no ponto a e u é diferenciável no ponto $v(a)$, então $u \circ v$ é diferenciável em a e

$$(u \circ v)'(a) = u'(v(a)) \cdot v'(a),$$

onde

$$u'(v(a)) = \frac{d}{dv}u(v)\Big|_{v=v(a)}$$

e as demais derivadas são em relação a x no ponto a. Isso porque

$$u(x) = u(v(x)).$$

De acordo com o Teorema 5.26, existe função contínua φ_a (nas condições impostas pelo teorema) tal que

$$v(x) - v(a) = \varphi_a(x - a).$$

Analogamente, existe função contínua $\varphi_{v(a)}$ tal que

$$u(x) - u(v(a)) = \varphi_{v(a)}(x - v(a)).$$

Logo,

$$(u \circ v)(x) - (u \circ v)(a) = u(v(x)) - u(v(a)) =$$

$$\varphi_{v(a)}(v(x)) \cdot (v(x) - v(a)) = (\varphi_{v(a)} \circ v) \cdot \varphi_a(x) \cdot (x - a).$$

Mas $(\varphi_{v(a)} \circ v) \cdot \varphi_a$ é contínua no ponto a, com valor

$$u'(v(a)) \cdot v'(a)$$

no ponto a.

Uma vez que isso vale para todo a do domínio de v (e, consequentemente, do domínio de $u \circ v$), então a prova está concluída.

Comumente o teorema acima é escrito como

$$\frac{du}{dx} = \frac{du}{dv}\frac{dv}{dx}.$$

Apesar do caráter mnemônico da fórmula acima, a qual parece sugerir uma 'simplificação' de dv com dv no lado direito da igualdade (restando apenas $\frac{du}{dx}$), obviamente este não é o caso. Com efeito, uma derivada como $\frac{dv}{dx}$ **não** é uma razão entre um real dv e um real dx. Uma vez que derivada é o limite de uma razão, Teorema 5.28 garante que o limite de uma razão é a razão entre limites, *desde que* o limite do denominador seja diferente de zero. No entanto, este não é o caso do conceito de derivada, de acordo com a Definição 5.20.

Exemplo 6.5. *Seja* $u : \mathbb{R} \to \mathbb{R}$ *uma função tal que* $u(x) = (3x^2 + 2x)^2$. *Como calcular* $\frac{du}{dx}$*?*

Técnica 1: sem empregar Teorema 6.3, temos que

$$u(x) = 9x^4 + 12x^3 + 4x^2;$$

logo,

$$\frac{du}{dx} = 36x^3 + 36x^2 + 8x;$$

Técnica 2: usando Teorema 6.3, podemos reescrever a função u *como uma composição, onde* $u(v) = v^2$ *e* $v(x) = 3x^2 + 2x$*; logo,*

$$\frac{du}{dx} = \frac{du}{dv}\frac{dv}{dx} =$$

$$2v(6x + 2) = 2(3x^2 + 2x)(6x + 2) = (6x^2 + 4x)(6x + 2) =$$
$$36x^3 + 12x^2 + 24x^2 + 8x = 36x^3 + 36x^2 + 8x.$$

Técnica 2 (ou seja, empregar Teorema 6.3) acima pode ser inconveniente para o exemplo dado. No entanto, ela se mostra muito eficiente para uma função $u : \mathbb{R} \to \mathbb{R}$ tal que $u(x) = (3x^2 + 2x)^{38}$. Afinal, não é uma boa ideia desenvolver $(3x^2 + 2x)^{38}$ em sua forma polinomial antes de derivar em relação a x. Portanto, mais uma vez percebemos o enorme poder de teoremas. A meta, num momento como esse, é sempre a mesma: economia de pensamento.

Logo, se $u(x) = (3x^2 + 2x)^{38}$, então

$$u'(x) = 38(6x + 2)(3x^2 + 2x)^{37}.$$

Exemplo 6.6. *Quanto é*

$$\frac{d}{dx}\text{sen}(x^2)?$$

Neste caso, $u(v) = \text{sen}(v)$ *e* $v(x) = x^2$. *Logo,*

$$\frac{d}{dx}\text{sen}(x^2) = \frac{du}{dv}\frac{dv}{dx} = \cos(v)2x = \cos(x^2)2x = 2x\cos(x^2).$$

EXEMPLO 6.7. *Quanto é*

$$\frac{d}{dx}\operatorname{sen}^2(x)?$$

Neste caso, $u(v) = v^2$ *e* $v(x) = \operatorname{sen}(x)$. *Logo,*

$$\frac{d}{dx}\operatorname{sen}^2(x) = \frac{du}{dv}\frac{dv}{dx} = 2v\cos(x) = 2\operatorname{sen}(x)\cos(x).$$

ℹ ✍ Em 1800 Louis François Antoine publicou em um livro de cálculo uma generalização do Teorema 6.3, hoje conhecida como *fórmula de Faà di Bruno*. Esta permite determinar a derivada de qualquer ordem n de uma composição de duas funções, sem a necessidade de calcular as derivadas de ordem anterior a n. Demonstrar a fórmula de Faà di Bruno por indução infinita pode ser uma tarefa um tanto exaustiva. Mas o leitor está convidado a pensar sobre o assunto.

SEÇÃO 56

Função exponencial

SUMÁRIO
ÍNDICE
REDE

Existe uma importante relação entre funções circulares (seno e co-seno) e função *exponencial*, a qual é uma das mais comumente empregadas em inúmeras aplicações, incluindo estatística. Antes de estudarmos isso, precisamos definir exponencial.

DEFINIÇÃO 6.1. Exponencial *é uma função* $y : \mathbb{R} \to \mathbb{R}$ *tal que*

$$y'(x) = y(x)$$

e

$$y(0) = 1.$$

Abreviamos $y(x)$ *como* $\exp(x)$. *Lemos* $\exp(x)$ *como 'exponencial de* x*'.*

Ou seja, exponencial é uma função y que é solução de uma equação diferencial ($y' = y$) com uma condição de contorno ($y(0) = 1$).

Se existir série de potências para representar $y(x) = \exp(x)$, temos:

$$y(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 + a_5x^5 + a_6x^6 + a_7x^7 + \cdots$$

$$y'(x) = a_1+2a_2x+3a_3x^2+4a_4x^3+5a_5x^4+6a_6x^5+7a_7x^6+8a_8x^7+\cdots.$$

Logo,

$$a_1 = a_0,\ a_2 = \frac{a_1}{2} = \frac{a_0}{2},\ a_3 = \frac{a_2}{3} = \frac{a_0}{3.2},\ a_4 = \frac{a_3}{4} = \frac{a_0}{4.3.2},\ \cdots .$$

Uma vez que $y(0) = 1$, temos $a_0 = 1$. Logo

$$\exp(x) = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots = \sum_{n=0}^{\infty} \frac{x^n}{n!}.$$

✍ Sugerimos ao leitor provar a fórmula acima por indução.

Observar que

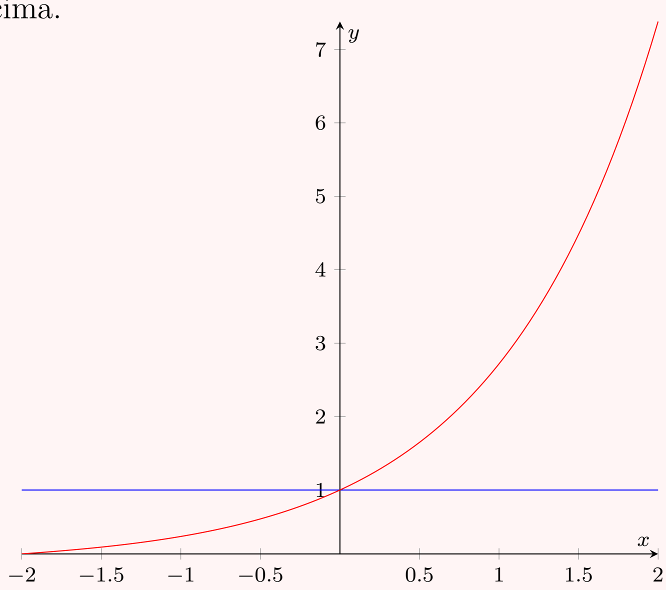
$$\frac{d}{dx}\left(1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots\right) = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots ,$$

ou seja,

$$\frac{d}{dx}\left(\sum_{n=0}^{\infty} \frac{x^n}{n!}\right) = \sum_{n=0}^{\infty} \frac{x^n}{n!},$$

o que confirma a condição $y' = y$.

Para efeitos computacionais, é possível programar uma máquina para gerar aproximações de exponencial de um real x qualquer com a precisão desejada. Para isso basta truncar a série de potências acima.



Na representação gráfica acima a função

$$f(x) = \sum_{n=0}^{0} \frac{x^n}{n!} = 1,$$

em azul, é uma primeira aproximação de exponencial de x. A função

$$g(x) = \sum_{n=0}^{7} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} + \frac{x^6}{6!} + \frac{x^7}{7!},$$

em vermelho, é uma aproximação que trunca a série na oitava parcela.

Teorema 6.4. *A série*

$$\sum_{n=0}^{\infty} \frac{x^n}{n!}$$

converge para todo x real.

Demonstração: Aplicando Teorema 5.34, temos que

$$\lim_{n\to\infty} \left| \frac{\frac{x^{n+1}}{(n+1)!}}{\frac{x^n}{n!}} \right| = \lim_{n\to\infty} |x| \frac{n!}{(n+1)!} = \lim_{n\to\infty} |x| \frac{1}{n+1} = 0.$$

Como $0 < 1$, então a série em questão converge.

Portanto, a série de potências acima de fato pode ser usada para representar $\exp(x)$.

Seção 57

Propriedades de funções circulares

As funções seno, co-seno e exponencial podem ser estendidas para o corpo $\mathbb{C}$ dos complexos. Detalhes em livros sobre funções complexas. Logo

$$\operatorname{sen}(ix) = ix - \frac{(ix)^3}{3!} + \frac{(ix)^5}{5!} - \frac{(ix)^7}{7!} - \cdots = ix + i\frac{x^3}{3!} + i\frac{x^5}{5!} + i\frac{x^7}{7!} + \cdots$$

Analogamente,

$$\cos(ix) = 1 - \frac{(ix)^2}{2!} + \frac{(ix)^4}{4!} - \frac{(ix)^6}{6!} + \frac{(ix)^8}{8!} - \cdots =$$

Sumário
Índice
Rede

$$1+\frac{x^2}{2!}+\frac{x^4}{4!}+\frac{x^6}{6!}+\frac{x^8}{8!}+\cdots,$$

onde i é uma abreviação para a unidade imaginária introduzida na Seção 40.

Ver Seção 40 para lidar com as potências n acima.

No entanto,

$$\exp(ix)=1+ix+\frac{(ix)^2}{2!}+\frac{(ix)^3}{3!}+\frac{(ix)^4}{4!}+\cdots=$$
$$1+ix-\frac{x^2}{2!}-i\frac{x^3}{3!}+\frac{x^4}{4!}+\cdots$$

Logo, podemos rearranjar os termos da série de potências como se segue:

$$\exp(ix)=1-\frac{x^2}{2!}+\frac{x^4}{4!}-\frac{x^6}{6!}+\cdots+i\left(x-\frac{x^3}{3!}+\frac{x^5}{5!}-\frac{x^7}{7!}+\cdots\right).$$

Ou seja,

$$\exp(ix)=\cos(x)+i\text{sen}(x)$$

A última fórmula é o célebre *Teorema de Euler*.

Teorema 6.5. *Sejam α e β números reais quaisquer. Logo,*
$$\exp(\alpha)\exp(\beta)=\exp(\alpha+\beta).$$

Demonstração: Considere a equação diferencial $u'=u$ com a condição de contorno $u(0)=\exp(\alpha)$. Se
$$y:\mathbb{R}\to\mathbb{R}$$
for uma função tal que
$$y(x)=\exp(\alpha)\exp(x),$$
então Teorema 5.20 garante que
$$y'(x)=\exp(\alpha)\exp(x),$$
ou seja,
$$y'=y.$$
Além disso,
$$y(0)=\exp(\alpha),$$
uma vez que Definição 6.1 diz que exponencial de 0 é 1.

Agora, seja $z : \mathbb{R} \to \mathbb{R}$ uma função dada por

$$z(x) = \exp(\alpha + x).$$

Teorema 6.3 garante que

$$z'(x) = \exp(\alpha + x),$$

uma vez que a derivada, em relação a x, de $\alpha + x$ é 1. Logo,

$$z' = z.$$

Além disso,

$$z(0) = \exp(\alpha).$$

Logo, ambas as funções

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

e

$$z(x) = \exp(\alpha + x)$$

satisfazem a mesma equação diferencial com a mesma condição de contorno.

Uma vez que $u' = u$, com $u(0) = \exp(\alpha)$, admite uma única solução para qualquer real α (✍ recomendamos ao leitor que prove isso), logo,

$$\exp(\alpha)\exp(x) = \exp(\alpha + x).$$

Para concluir a prova, basta fazer $x = \beta$.

No Teorema 6.21 (a ser examinado adiante) é provado, entre outras coisas, que exponencial $\exp(r)$ de qualquer real r é um real estritamente positivo. Logo, no contexto da demonstração acima, fica claro que a equação diferencial $y' = y$ com condição de contorno $y(0) = \gamma$, admite solução se $\gamma \geq 0$. Se $\gamma < 0$, não há solução alguma para o mesmo problema de contorno.

Lembrar que uma solução y de uma equação diferencial $\mathfrak{D}(y) = g$ (onde $\mathfrak{D}$ é um operador diferencial e g é uma função real) é uma função tal que a fórmula $\mathfrak{D}(y) = g$ é teorema de ZF. Naturalmente, $y' = y$ é um caso particular de $\mathfrak{D}(y) = g$.

Teorema 6.6. *Sejam β e γ reais quaisquer. Logo,*

$$\cos(\beta + \gamma) = \cos(\beta)\cos(\gamma) - \mathrm{sen}(\beta)\mathrm{sen}(\gamma)$$

e

$$\mathrm{sen}(\beta + \gamma) = \mathrm{sen}(\beta)\cos(\gamma) + \cos(\beta)\mathrm{sen}(\gamma).$$

Demonstração: De acordo com o Teorema de Euler,

$$\exp(i\beta + i\gamma) = \exp(i(\beta + \gamma)) =$$
$$\cos(\beta + \gamma) + i\mathrm{sen}(\beta + \gamma).$$

No entanto,

$$\exp(i\beta)\exp(i\gamma) = (\cos(\beta) + i\mathrm{sen}(\beta))(\cos(\gamma) + i\mathrm{sen}(\gamma))$$

que é igual a

$$\cos(\beta)\cos(\gamma) - \mathrm{sen}(\beta)\mathrm{sen}(\gamma) +$$
$$i(\mathrm{sen}(\beta)\cos(\gamma) + \cos(\beta)\mathrm{sen}(\gamma)).$$

Uma vez que

$$\exp(i\beta)\exp(i\gamma) = \exp(i\beta + i\gamma)$$

(uma generalização do Teorema 6.5 que pode ser demonstrada de forma análoga), comparando as partes reais e imaginárias, temos as duas igualdades a seguir:

$$\cos(\beta + \gamma) = \cos(\beta)\cos(\gamma) - \mathrm{sen}(\beta)\mathrm{sen}(\gamma)$$

e

$$\mathrm{sen}(\beta + \gamma) = \mathrm{sen}(\beta)\cos(\gamma) + \cos(\beta)\mathrm{sen}(\gamma).$$

✍ Observar que seno é uma função ímpar, i.e.,

$$\mathrm{sen}(-x) = -\mathrm{sen}(x).$$

Logo,

$$\mathrm{sen}(\beta - \gamma) = \mathrm{sen}(\beta)\cos(\gamma) - \cos(\beta)\mathrm{sen}(\gamma).$$

Teorema análogo pode ser obtido, notando que $\cos(-\gamma) = \cos(\gamma)$.

$$\cos(\beta - \gamma) = \cos(\beta)\cos(\gamma) + \mathrm{sen}(\beta)\mathrm{sen}(\gamma)$$

Ou seja, para compreender propriedades das funções seno e co-seno sobre reais, ajuda muito conhecer as mesmas sobre os complexos.

Definição 6.2. *Seja $f : d \to \mathbb{R}$ uma função, sendo $d \subseteq \mathbb{R}$. Dizemos que $F : d \to \mathbb{R}$ é uma* primitiva *de f sss*

$$\frac{dF(x)}{dx} = f(x).$$

EXEMPLO 6.8. I: *Seja* $f : \mathbb{R} \to \mathbb{R}$ *tal que*

$$f(x) = 6x^3 - 2x + 5;$$

logo, para cada $c \in \mathbb{R}$, $F : \mathbb{R} \to \mathbb{R}$, *tal que*

$$F(x) = \frac{6}{4}x^4 - x^2 + 5x + c,$$

é uma primitiva de f*;*

II: *seja* $y : \mathbb{R} \to \mathbb{R}$ *uma função tal que*

$$y(x) = \mathrm{sen}(x);$$

logo, para cada $c \in \mathbb{R}$, $Y : \mathbb{R} \to \mathbb{R}$, *tal que*

$$Y(x) = -\cos(x) + c,$$

é uma primitiva de y*.*

Se F e G são primitivas de f, então, para todo x pertencente ao domínio de f temos

$$F(x) = G(x) + c,$$

para alguma constante real c. Basta derivar em relação a x ambos os lados da igualdade acima. Logo,

$$\frac{d}{dx}F(x) = \frac{d}{dx}(G(x) + c) = \frac{d}{dx}G(x) + \frac{d}{dx}c = \frac{d}{dx}G(x).$$

Mas, por hipótese,

$$\frac{d}{dx}F(x) = f(x) \;\text{ e }\; \frac{d}{dx}G(x) = f(x).$$

Logo, primitivas F e G de uma mesma função real f diferem entre si apenas por uma constante real c. Esse fato repercute significativamente no estudo das funções seno e co-seno, entre outras. Para ilustrar essa última afirmação, ver a discussão a seguir.

Como vimos anteriormente,

$$\frac{d}{dx}\mathrm{sen}^2(x) = 2\mathrm{sen}(x)\cos(x).$$

Entretanto,

$$\frac{d}{dx}\cos^2(x) = -2\mathrm{sen}(x)\cos(x),$$

ou seja,

$$\frac{d}{dx}(-\cos^2(x)) = 2\mathrm{sen}(x)\cos(x).$$

Isso implica que uma primitiva qualquer de $2\text{sen}(x)\cos(x)$ pode ser tanto

$$\text{sen}^2(x) + d$$

quanto

$$-\cos^2(x) + e.$$

Logo, essas primitivas diferem entre si por uma constante, ou seja,

$$\text{sen}^2(x) = -\cos^2(x) + c.$$

Logo,

$$\text{sen}^2(x) + \cos^2(x) = c,$$

para todo real x.

Não obstante, essa constante c deve assumir o mesmo valor para todo x pertencente ao domínio de ambas as funções $\text{sen}(x)$ e $\cos(x)$. Se $x = 0$, temos $\text{sen}(x) = 0$ e $\cos(x) = 1$ (isso é consequência das definições das funções circulares). Logo, $c = 1$. Portanto, para qualquer x pertencente aos reais temos

$$\boxed{\text{sen}^2(x) + \cos^2(x) = 1.}$$

Se $|\text{sen}(x)|$ e $|\cos(x)|$ são medidas de catetos de um triângulo retângulo com hipotenusa medindo 1, a última igualdade se identifica com o Teorema de Pitágoras, o qual diz o seguinte: o quadrado da medida da hipotenusa de um triângulo retângulo é a soma dos quadrados das medidas dos catetos.

Lembrar que $|\text{sen}(x)|^2 = \text{sen}^2(x)$ e $|\cos(x)|^2 = \cos^2(x)$, para todo real x. Lembrar também que um *triângulo* é uma *curva poligonal fechada* com três *lados*, um *triângulo retângulo* é um triângulo em que um de seus *ângulos internos* é um *ângulo reto* e a *hipotenusa* de um triângulo retângulo é o lado com maior *medida*. Alguns detalhes podem ser vistos na Parte 7.

Uma vez que ambas as funções sen e cos têm periodicidade 2π ($\text{sen}(x + 2\pi) = \text{sen}(x)$ e $\cos(x + 2\pi) = \cos(x)$, por conta da interpretação geométrica acima), no contexto do círculo trigonométrico usual $\text{sen}(x)$ pode ser identificado como a razão entre a medida do cateto oposto a um ângulo agudo de medida x (em radianos) e a medida da hipotenusa de um triângulo retângulo. Resultado análogo vale para cos, dessa vez envolvendo um cateto adjacente ao ângulo de medida x.

Com relação à periodicidade 2π das funções seno e co-seno, lembrar que π é a razão entre o perímetro de uma circunferência e seu diâmetro. Tal definição se sustenta por um teorema da geometria plana que estabelece a invariância da razão entre o perímetro de uma circunferência e seu diâmetro. O valor aproximado, em notação decimal usual, é

$$\pi \approx 3,1415926535897932384626,$$

com vinte e duas casas decimais após a vírgula. O símbolo $\approx$ denota 'valor aproximado', com uma quantia finita de casas após a vírgula.

Na Seção 108 há uma breve discussão sobre como calcular rapidamente o valor de π com uma ótima precisão.

Para finalizar esta Seção, alguns conceitos usuais.

Definição 6.3. i:

$$\tan : \left\{ x \in \mathbb{R} \mid \forall n \left(n \in \mathbb{Z} \Rightarrow x \neq \frac{\pi}{2} + n\pi \right) \right\} \to \mathbb{R}$$

é a função dada por

$$\tan(x) = \frac{\operatorname{sen}(x)}{\cos(x)};$$

ii:

$$\cot : \left\{ x \in \mathbb{R} \mid \forall n \left(n \in \mathbb{Z} \Rightarrow x \neq n\pi \right) \right\} \to \mathbb{R}$$

é a função dada por

$$\cot(x) = \frac{\cos(x)}{\operatorname{sen}(x)};$$

iii:

$$\sec : \left\{ x \in \mathbb{R} \mid \forall n \left(n \in \mathbb{Z} \Rightarrow x \neq \frac{\pi}{2} + n\pi \right) \right\} \to \mathbb{R}$$

é a função dada por

$$\sec(x) = \frac{1}{\cos(x)};$$

iv:

$$\csc : \left\{ x \in \mathbb{R} \mid \forall n \left(n \in \mathbb{Z} \Rightarrow x \neq n\pi \right) \right\} \to \mathbb{R}$$

é a função dada por

$$\csc(x) = \frac{1}{\operatorname{sen}(x)};$$

Lemos essas funções, respectivamente, como *tangente*, *cotangente*, *secante* e *co-secante*.

As condições impostas aos domínios são necessárias por conta dos zeros de seno e co-seno.

Seção 58

Integral de Riemann

Sumário
Índice
Rede

Georg Friedrich Bernhard Riemann é o criador daquilo hoje conhecido como *integral de Riemann*. Os trabalhos deste matemático alemão sobre séries de Fourier inspiraram Georg Cantor a desenvolver as primeiras ideias sobre teoria de conjuntos, as quais foram também influenciadas pela obra de Bolzano.

Relembrando conceitos já vistos na Seção 39, um intervalo fechado $[a, b]$ é o conjunto

$$[a, b] = \{x \in \mathbb{R} \mid a \leq x \leq b\}.$$

Se $a \neq b$, dizemos que $[a, b]$ é não degenerado. Se $a = b$, $[a, b]$ é degenerado.

Já um intervalo aberto (a, b) é o conjunto

$$(a, b) = \{x \in \mathbb{R} \mid a < x < b\}.$$

Uma *partição* P de $[a, b]$ em n intervalos fechados é o conjunto

$$P = \{[a_i, a_{i+1}] \in \wp([a, b]) \mid i \in \omega \wedge 0 \leq i \leq n-1 \wedge$$
$$[a_0, a_1] \cup [a_1, a_2] \cup \cdots \cup [a_{n-1}, a_n] = [a, b]\},$$

sendo que cada elemento de P é um intervalo fechado não degenerado.

Usualmente P é referida simplesmente como *partição de* $[a, b]$.

Ou seja, uma partição P de um intervalo fechado $[a, b]$ é um conjunto de subconjuntos de $[a, b]$ (observar que o conjunto universo usado no emprego do Esquema da Separação é $\wp([a, b])$) tal que seus elementos são intervalos fechados $[a_i, a_{i+1}]$ que satisfazem a duas condições:

I: a união arbitrária de todos os elementos de P é o intervalo fechado $[a, b]$ e

II: a interseção entre dois elementos quaisquer de P, distintos entre si, é vazia ou um *singleton* $\{a_{i+1}\}$.

No caso particular em que o intervalo fechado $[a, b]$ é degenerado, obviamente ele admite uma única partição. Nas demais situações há uma infinidade de possíveis partições para um mesmo intervalo fechado não degenerado.

Aqui cabe observar algo importante. O contexto desta Seção é o conceito de integral de Riemann, a ser dado adiante. Neste sentido, o intervalo fechado $[a, b]$ é chamado de *domínio de integração*, enquanto P é a *partição do domínio de integração*. Ou seja, partição de um domínio de integração nada tem a ver com partição de um conjunto no sentido apresentado na Definição 3.16, Seção 26. Com efeito, dois elementos distintos de uma partição P de um domínio de integração $[a, b]$ podem ter interseção não vazia, como se percebe no exemplo abaixo. No contexto da Definição 3.16 dois elementos distintos de uma partição sempre têm interseção vazia.

Exemplo 6.9. *Uma possível partição de* $[-2, 7]$ *é*

$$P = \{[-2, -1], [-1, 2], [2, 6], [6, 7]\}.$$

Com efeito, $n = 4$*, onde*

$$a_0 = -2,\ a_1 = -1,\ a_2 = 2,\ a_3 = 6 \ \text{ e } a_4 = 7;$$

além disso,

$$[-2, -1] \cup [-1, 2] \cup [2, 6] \cup [6, 7] = [-2, 7],$$

lembrando que união finitária é associativa. Observar também que, por exemplo, $[-2, -1] \cap [6, 7] = \varnothing$*, enquanto* $[-2, -1] \cap [-1, 2]$ *é o* singleton $\{-1\}$*. Comentário análogo vale para quaisquer elementos tomados dois a dois a partir de* P*.*

Denotamos por Δx_i o real $a_{i+1} - a_i$, onde $0 \leq i \leq n - 1$. Neste sentido, cada intervalo fechado $[a_i, a_{i+1}]$ pertencente à partição P de $[a, b]$ é um conjunto de números reais, enquanto

$$\Delta x_i = a_{i+1} - a_i$$

é chamada de *medida* do intervalo $[a_i, a_{i+1}]$. Intuitivamente falando, essa medida corresponde ao 'comprimento' do intervalo fechado. Em particular, todo intervalo fechado degenerado tem medida zero.

Sobre o conceito de medida, ver Seção 103. Mas, para os atuais propósitos, essa noção intuitiva de medida é suficiente.

A *norma* de uma partição P é definida como

$$\| P \| = \max\{\Delta x_i\},$$

onde $\max\{\Delta x_i\}$ denota o real Δx_j tal que

$$\Delta x_j \geq \Delta x_i$$

para todo i de modo que $0 \leq i \leq n-1$. Ou seja, $\max\{\Delta x_i\}$ é o *máximo* valor entre todos os Δx_i.

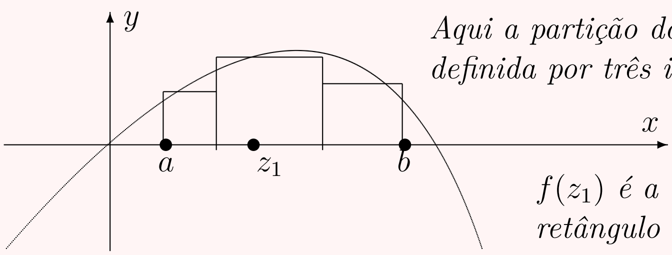
EXEMPLO 6.10. *Na partição P do* EXEMPLO *anterior temos*

$$\| P \| = 4.$$

Seja f uma função real definida sobre um intervalo não degenerado $[a, b]$. A *integral de Riemann* (ou, simplesmente, a *integral*) de f em relação a x em $[a, b]$ é dada por

$$\int_a^b f(x)dx = \lim_{\|P\| \to 0} \sum_i f(z_i)\Delta x_i,$$

sendo $\Delta x_i = a_{i+1} - a_i$ e $z_i \in (a_i, a_{i+1})$.



Aqui a partição do intervalo $[a, b]$ é definida por três intervalos fechados.

$f(z_1)$ é a altura do retângulo do meio.

Chamamos o intervalo fechado $[a, b]$ de *domínio de integração* da integral $\int_a^b f(x)dx$, enquanto o ponto a é chamado de *limite inferior de integração*, e b é chamado de *limite superior de integração*.

O somatório

$$\sum_i f(z_i)\Delta x_i$$

é chamado de *soma de Riemann*.

Na imagem acima é sugerida uma soma de Riemann

$$f(z_0)(a_1 - a_0) + f(z_1)(a_2 - a_1) + f(z_3)(a_3 - a_2),$$

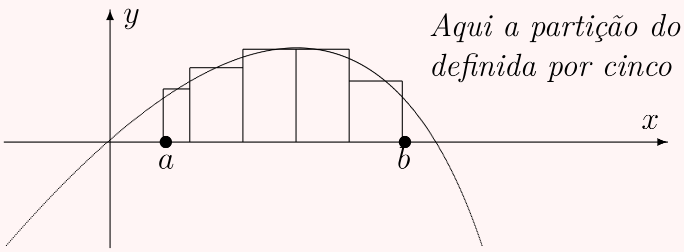
onde $a_0 = a$ e $a_3 = b$.

Logo, a soma de Riemann é uma função da partição P e da escolha de cada z_i. Cada partição e cada escolha de z_i, para cada i, corresponde a uma soma de Riemann. Por exemplo, para a mesma função sugerida na imagem acima e para o mesmo domínio de integração $[a, b]$ podemos ter a seguinte soma de Riemann sugerida na próxima imagem:

$$f(z_0)(a_1 - a_0) + f(z_1)(a_2 - a_1) + f(z_3)(a_3 - a_2)+$$

$$f(z_4)(a_4 - a_3) + f(z_5)(a_5 - a_4),$$

onde $a_0 = a$ e $a_5 = b$.

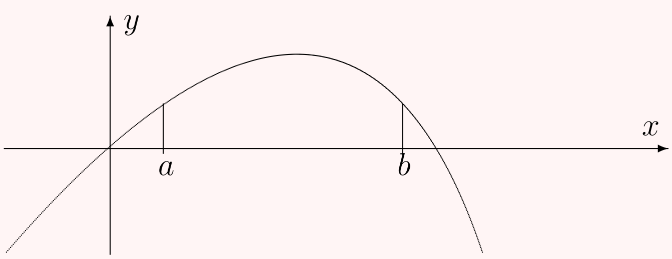


Aqui a partição do intervalo $[a, b]$ é definida por cinco intervalos.

No caso particular em que $f(x) \geq 0$ para todo x pertencente ao intervalo $[a, b]$ (como sugerido nas imagens acima), cada termo $f(z_i)\Delta x_i$ corresponde à área de um retângulo com base de medida Δx_i e altura de medida $f(z_i)$. Ainda neste caso,

$$\int_a^b f(x)dx$$

corresponde à área da região de $\mathbb{R}^2$ compreendida abaixo de $f(x)$, acima do eixo x e ladeada pelas retas verticais $x = a$ e $x = b$, como sugerido na imagem abaixo.



Com efeito, a integral

$$\int_a^b f(x)dx$$

é o limite da soma de Riemann

$$\sum_i f(z_i)\Delta x_i,$$

com a norma da partição P tendendo a zero.

Mas, se o máximo entre os Δx_i (ou seja, a norma de P) se torna arbitrariamente pequeno (ou seja, $\| P \| \in (0-\delta, 0+\delta)$), então o número n de elementos da partição P se torna arbitrariamente grande (ou seja, $n > \varepsilon$ para qualquer ε estritamente positivo.).

Observar que, enquanto cada Δx_i é uma medida de um segmento de reta $[a_i, a_{i+1}]$, $\int_a^b f(x)dx$ também é uma medida, pelo menos para o caso particular em que $f(x) \geq 0$ para todo $x \in [a, b]$. Mas, desta vez, trata-se da medida de uma região de $\mathbb{R}^2$.

Medidas de segmentos de reta são também conhecidas como *comprimentos*, enquanto medidas de regiões de $\mathbb{R}^2$ são chamadas de *áreas*.

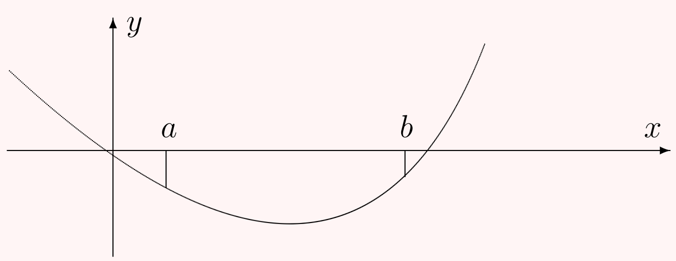
Com relação às retas verticais acima mencionadas, elas correspondem aos conjuntos

$$\{(a, y) \in \mathbb{R}^2 \mid y \in \mathbb{R}\}$$

e

$$\{(b, y) \in \mathbb{R}^2 \mid y \in \mathbb{R}\},$$

onde $b > a$.



Se

$$f(x) < 0,$$

para todo x pertencente ao domínio de integração $[a, b]$, então

$$\int_a^b f(x)dx$$

é um real negativo tal que

$$\left|\int_a^b f(x)dx\right|$$

é a área da região de $\mathbb{R}^2$ delimitada por $f(x)$, pelo eixo x e pelas retas $x = a$ e $x = b$.

Assim como derivada é um caso particular de limite, integral de Riemann também é. Uma vez que limites podem existir ou não, o mesmo ocorre com integrais. Se f admite integral em $[a, b]$, dizemos que f é *integrável* em $[a, b]$.

SEÇÃO 59

Teoremas básicos

SUMÁRIO
ÍNDICE
REDE

Seguem alguns resultados estratégicos.

TEOREMA 6.7. *Seja $f : [a, b] \to \mathbb{R}$ uma função integrável tal que $f(x) \geq 0$, para todo x pertencente a $[a, b]$. Então,*

$$\int_a^b f(x)dx \geq 0.$$

DEMONSTRAÇÃO: Se $[a, b]$ for um intervalo não degenerado, então

$$\int_a^b f(x)dx = \lim_{\|P\| \to 0} \sum_i f(z_i)\Delta x_i.$$

Porém, cada Δx_i é positivo. Além disso, por hipótese, cada $f(z_i)$ é positivo. Uma vez que o somatório de parcelas positivas é uma soma positiva, então o limite da soma de Riemann é positivo.

TEOREMA 6.8. *Seja $f : [a, b] \to \mathbb{R}$ uma função tal que*

$$f(x) = c,$$

sendo $a < b$. Então,

$$\int_a^b f(x)dx = c(b - a).$$

DEMONSTRAÇÃO:

$$\int_a^b c\,dx = \lim_{\|P\| \to 0} \sum_i c\Delta x_i = \lim_{\|P\| \to 0} c \sum_i \Delta x_i =$$

$$\lim_{\|P\|\to 0} c \cdot \lim_{\|P\|\to 0} \sum_i \Delta x_i = \lim_{\|P\|\to 0} c \cdot \lim_{\|P\|\to 0} (b-a) =$$
$$c(b-a).$$

Observar que, na última igualdade acima, foi usado o Teorema 5.9 sobre limite de função constante.

Exemplo 6.11. i:

$$\int_{-2}^{5} 7dx = 7(5-(-2)) = 7(5+2) = 49;$$

ii:

$$\int_{-2}^{5} -7dx = -7(5-(-2)) = -7(5+2) = -49.$$

Integral de Riemann não é um conceito elegante, como ocorre com *integral de Lebesgue*. Para detalhes sobre integração de Lebesgue, ver [29]. A obra citada exige como requisitos apenas conhecimentos básicos sobre limites, séries e derivadas de funções reais.

Por conta da falta de elegância de integrais de Riemann, dois casos especiais devem ser considerados para concluir a definição:

i: aqueles em que o limite inferior de integração é maior do que o limite superior de integração e

ii: aqueles nos quais o limite inferior de integração é idêntico ao limite superior de integração.

Para lidar com essas situações, a solução usual é incluir as fórmulas a seguir para finalizar a definição de integral de Riemann:

$$\int_{b}^{a} f(x)dx = -\int_{a}^{b} f(x)dx,$$
$$\int_{a}^{a} f(x)dx = 0.$$

Exemplo 6.12. *Uma vez que* $\int_{-2}^{5} 7dx = 49$*, então*

$$\int_{5}^{-2} 7dx = -49.$$

Ou seja, a permutação de limites de integração implica na troca de sinal de uma integral de Riemann. Além disso, a integral de Riemann de uma função real em um intervalo fechado degenerado é zero.

Observar que toda função é integrável em um intervalo degenerado.

Exemplo 6.13. *A* função de Dirichlet $g : \mathbb{R} \to \mathbb{R}$ *dada por*

$$g(x) = \begin{cases} 1 & \text{se } x \text{ é real racional} \\ 0 & \text{se } x \text{ é real irracional} \end{cases}$$

não é integrável se o domínio de integração $[a, b]$ *for um intervalo não degenerado.* ✍ *Consegue provar isso?*

Apesar da função de Dirichlet não ser integrável no sentido de Riemann, ela é integrável *à la* Lebesgue. Logo, além de integrais de Lebesgue serem mais elegantes, elas conseguem dar conta de situações não tratáveis via integração de Riemann.

Teorema 6.9. *Seja* f *uma função contínua em* $[a, b]$*. Logo, existe* z *pertencente a* (a, b) *tal que*

$$\int_a^b f(x)dx = f(z)(b - a).$$

O resultado acima é conhecido como o *Teorema do Valor Médio para Integrais*. Sua demonstração exige a aplicação de outros teoremas aqui ignorados, como o *Teorema do Valor Extremo* e o *Teorema do Valor Intermediário*, os quais consistem em um aprofundamento no estudo de funções contínuas. Por esse motivo omitimos aqui a sua prova. No entanto, Teorema 6.9 é bastante intuitivo, como se mostra a seguir.

Exemplo 6.14. i: *Seja* $f : \mathbb{R} \to \mathbb{R}$ *tal que* $f(x) = c$*; logo, qualquer* z *pertencente a* (a, b) *satisfaz o teorema acima;*

ii: *Seja* $g : \mathbb{R} \to \mathbb{R}$ *tal que* $g(x) = 2x$*; logo,*

$$\int_0^3 g(x)dx = 9;$$

com efeito, essa integral é a área de um triângulo com base de medida 3 e altura 6; logo,

$$f(1,5) \cdot (3 - 0) = 3 \cdot 3 = 9$$

garante que $z = 1,5$ *satisfaz o Teorema do Valor Médio para Integrais.*

No caso particular em que $f(x) > 0$ para todo x pertencente ao domínio de integração $[a, b]$, o Teorema do Valor Médio para Integrais afirma que existe retângulo cuja área é igual à integral

$$\int_a^b f(x)dx,$$

de modo que a base do retângulo mede $b-a$ e a altura mede $f(z)$, para algum z pertencente a (a,b), desde que f seja contínua em $[a,b]$.

Teorema 6.10. *Sejam f e g funções integráveis em $[a,b]$. Logo,*

i:
$$\int_a^b (f+g)(x)dx = \int_a^b f(x)dx + \int_a^b g(x)dx;$$

ii:
$$\int_a^b (f-g)(x)dx = \int_a^b f(x)dx - \int_a^b g(x)dx;$$

iii:
$$\int_a^b cf(x)dx = c\int_a^b f(x)dx,$$

onde c é uma constante real.

Demonstração: Provamos aqui apenas o item i, uma vez que os demais são demonstrados de maneira análoga. Começando com a definição de integral de Riemann, temos que

$$\int_a^b (f+g)(x)dx = \lim_{\|P\|\to 0}\sum_i (f+g)(z_i)\Delta x_i =$$

$$\lim_{\|P\|\to 0}\sum_i (f(z_i)+g(z_i))\Delta x_i = \lim_{\|P\|\to 0}\sum_i (f(z_i)\Delta x_i + g(z_i)\Delta x_i) =$$

$$\lim_{\|P\|\to 0}\left(\sum_i f(z_i)\Delta x_i + \sum_i g(z_i)\Delta x_i\right) =$$

$$\lim_{\|P\|\to 0}\sum_i f(z_i)\Delta x_i + \lim_{\|P\|\to 0}\sum_i g(z_i)\Delta x_i =$$

$$\int_a^b f(x)dx + \int_a^b g(x)dx.$$

Teorema 6.11. *Seja f uma função integrável em quaisquer intervalos fechados de $\mathbb{R}$. Logo,*

$$\int_a^c f(x)dx = \int_a^b f(x)dx + \int_b^c f(x)dx.$$

A prova deste último teorema é muito simples, bastando usar a definição de integral de Riemann e teoremas sobre limites e somatórios. No entanto, é uma prova que consome bastante tempo. Com efeito, devemos considerar todas as possibilidades envolvendo os limites de integração a, b e c:

$$a \leq b \leq c,\ a \leq c \leq b,\ b \leq c \leq a,$$
$$b \leq a \leq c,\ c \leq b \leq a \text{ e } c \leq a \leq b.$$

Além disso, deve ser levado em conta que as imagens de f podem mudar de sinal entre limites de integração ou nos próprios.

✍ Sugerimos que o leitor prove pelo menos para dois ou três casos.

Seção 60

Teorema Fundamental do Cálculo

Sumário
Índice
Rede

O teorema a seguir é um dos resultados de maior impacto social na história da humanidade, com implicações em matemática, física, engenharias, psicologia, medicina, ciências biológicas, artes áudio-visuais, estatística, ciência da computação, paleontologia, arqueologia, música estocástica, entre outras áreas do conhecimento.

Pode não ser algo comparável com a invenção da roda ou o domínio do fogo. Mas é um exemplo marcante das conquistas da ciência.

Teorema 6.12 (Teorema Fundamental do Cálculo). *Seja f uma função real contínua em $[a, b]$. Logo:*

I:
$$G(x) = \int_a^x f(t)dt \quad \text{é uma primitiva de } f(x);$$

II:
$$\int_a^b f(x)dx = F(b) - F(a), \quad \text{sendo } F(x) \text{ uma primitiva de } f(x).$$

Demonstração: (I) Devemos provar que

$$\frac{d}{dx}G(x) = f(x).$$

Logo, calculemos $\frac{d}{dx}G(x)$.

$$\frac{d}{dx}G(x) = \lim_{h\to 0}\frac{G(x+h)-G(x)}{h} =$$

$$\lim_{h\to 0}\frac{1}{h}(G(x+h)-G(x)) =$$

$$\lim_{h\to 0}\frac{1}{h}\left(\int_a^{x+h} f(t)dt - \int_a^x f(t)dt\right) =$$

$$\lim_{h\to 0}\frac{1}{h}\left(\int_a^{x+h} f(t)dt + \int_x^a f(t)dt\right) =$$

$$\lim_{h\to 0}\frac{1}{h}\int_x^{x+h} f(t)dt =$$

$$\lim_{h\to 0}\frac{1}{h}f(z)h,$$

sendo $z \in (x, x+h)$ ou $z \in (x+h, x)$.

Observando as quatro linhas de contas acima, justificamos cada uma a seguir.

Na primeira linha foi usada a definição de função derivada.

Na segunda foi empregada a definição de G (dada como hipótese).

Na terceira linha usamos a definição de integral de Riemann (para o caso de permutação de limites de integração), bem como o Teorema 6.11, o qual permite escrever certas somas de integrais como uma única integral.

Finalmente, a quarta linha faz uso do Teorema do Valor Médio para Integrais 6.9.

Logo, a transitividade da igualdade nos diz que

$$\frac{d}{dx}G(x) = \lim_{h\to 0} f(z).$$

No entanto,

$$\lim_{h\to 0} f(z) = \lim_{z\to x} f(z),$$

uma vez que z está entre x e $x+h$ (ou entre $x+h$ e x, se $h<0$).

Como f é contínua em $[a,b]$, então

$$\lim_{z\to x} f(z) = f(x).$$

Logo,

$$\frac{d}{dx}G(x) = f(x),$$

encerrando a demonstração da parte I.

(II) Se F é uma primitiva de f, então $F(x) = G(x) + C$. Logo,

$$F(a) = G(a) + C.$$

Mas $G(a) = 0$. Logo, $C = F(a)$. Logo,

$$F(b) = G(b) + F(a).$$

Mas $G(b) = \int_a^b f(t)dt$. Logo,

$$\int_a^b f(t)dt = F(b) - F(a),$$

encerrando a prova da parte II.

Denotamos $F(b) - F(a)$ por $F(x)\Big|_a^b$.

O Teorema Fundamental do Cálculo 6.12 estabelece:

I: uma inesperada relação entre integral de Riemann e derivada. Afinal, derivada é o limite de uma razão, enquanto integral é o limite de um somatório;

II: um critério simples para o cálculo de integrais. Com efeito, há vários teoremas que tornam o cálculo de derivadas bastante simples. Logo, determinar primitivas, como ocorre no item II do Teorema Fundamental do Cálculo, é algo muito mais simples do que calcular limite de uma soma de Riemann.

O fato de haver teoremas para derivadas de produtos torna o Teorema 6.12 um resultado muito bem-vindo, uma vez que não há teoremas não triviais sobre integral de produto entre funções. No entanto, os impactos mais significativos deste resultado são apreciados mais adiante.

O Teorema Fundamental do Cálculo não foi uma façanha conquistada 'do dia pra noite'. James Gregory enunciou e provou uma versão rudimentar deste resultado, utilizando argumentos de caráter essencialmente geométricos. Isaac Barrow demonstrou, a seguir, uma versão mais ampla. Por fim, Isaac Newton, aluno de Barrow, refinou o enunciado e a prova para uma versão mais próxima do que hoje se entende sobre o tema. Mas a forma como hoje se apresenta tal teorema, somente resultados conquistados no século 20 são capazes de justificá-lo.

Exemplo 6.15. i:

$$\int_2^5 x^3 dx = \left(\frac{x^4}{4} + C\right)\Bigg|_2^5 =$$

$$\frac{5^4}{4} + C - \left(\frac{2^4}{4} + C\right) = \frac{5^4}{4} - \frac{2^4}{4};$$

ii:

$$\int_a^b c\,dx = (cx + C)\Big|_a^b = cb + C - (ca + C) = c(b - a)$$

(no Teorema 6.8 o mesmo resultado foi provado independentemente do Teorema Fundamental do Cálculo);

iii:

$$\int_1^2 \operatorname{sen}(x)dx = -\cos(x)\Big|_1^2 =$$

$$-\cos(2) - (-\cos(1)) = \cos(1) - \cos(2);$$

iv:

$$\int_1^3 (3x^2 - 2x)dx = \left(\frac{3x^3}{3} - x^2\right)\Bigg|_1^3 =$$

$$\left(\frac{3(3)^3}{3} - 3^2\right) - \left(\frac{3(1)^3}{3} - 1^2\right) = 3^3 - 3^2 = 18.$$

Nos dois primeiros itens do Exemplo acima destacamos a constante C da primitiva da função integrada em relação x, no domínio de integração dado. No entanto, uma vez que uma integral de Riemann é uma diferença entre $F(b)$ e $F(a)$, claramente essa constante C não desempenha papel algum. Por conta disso que, nos dois últimos itens do Exemplo, omitimos qualquer consideração sobre tal constante.

A seguir relembramos funções reais ímpar e par.

Definição 6.4. *Seja $f : \mathbb{R} \to \mathbb{R}$ uma função. Dizemos que f é* ímpar *sss para todo $x \in \mathbb{R}$ temos*

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

Dizemos que f é par *sss para todo $x \in \mathbb{R}$ temos*

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

Exemplo 6.16. i: *$f : \mathbb{R} \to \mathbb{R}$, tal que*

$$f(x) = x^3,$$

é ímpar; com efeito, $(-x)^3 = -x^3$, o que implica em $f(-x) = -f(x)$.

ii: *$g : \mathbb{R} \to \mathbb{R}$, tal que*

$$g(x) = x^4,$$

é par; com efeito, $(-x)^4 = x^4$, o que implica em $g(-x) = g(x)$.

iii: *$h : \mathbb{R} \to \mathbb{R}$, tal que*

$$h(x) = x^3 + x^2,$$

não é par e nem ímpar.

Teorema 6.13. *Seja f uma função real integrável ímpar, com domínio $\mathbb{R}$. Logo,*

$$\int_{-a}^{a} f(x)dx = 0,$$

para qualquer real a.

Exemplo 6.17. i:

$$\int_{-3}^{3} (x^3 - x)dx = 0;$$

ii:

$$\int_{-4}^{4} sen(x)dx = 0.$$

Teorema 6.14. *Seja f uma função real integrável par, com domínio $\mathbb{R}$. Logo,*

$$\int_{-a}^{a} f(x)dx = 2\int_{0}^{a} f(x)dx,$$

para qualquer real a.

As provas dos dois últimos teoremas podem ser feitas sem dificuldade a partir da definição de integral de Riemann. Por conta dos limites de integração serem simétricos relativamente a zero, basta assumir partições de $[-a, a]$ que sejam simétricas em relação a zero. Ou seja, nem sempre o Teorema Fundamental do Cálculo é um agente facilitador para a demonstração de teoremas.

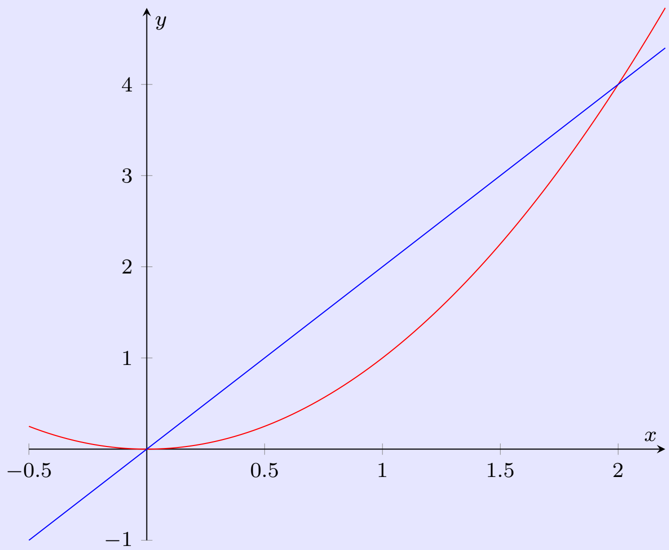
Exemplo 6.18. *Como calcular a área A da região R de $\mathbb{R}^2$ delimitada pelas funções $f : \mathbb{R} \to \mathbb{R}$ e $g : \mathbb{R} \to \mathbb{R}$ tais que*

$$f(x) = x^2$$

(em vermelho, na imagem abaixo) e

$$g(x) = 2x$$

(em azul)?



Temos que

$$f = \{(x, y) \in \mathbb{R}^2 \mid y = x^2\}$$

e

$$g = \{(x, y) \in \mathbb{R}^2 \mid y = 2x\}.$$

Logo,

$$f \cap g = \{(0, 0), (2, 4)\},$$

uma vez que $x^2 = 2x$ sss $x = 0$ ou $x = 2$.

✍ *Se $x \in [0, 2]$, então $g(x) \geq f(x)$. Logo,*

$$A = \int_0^2 g(x)dx - \int_0^2 f(x)dx = \int_0^2 2xdx - \int_0^2 x^2dx =$$

$$x^2\Big|_0^2 - \frac{x^3}{3}\Big|_0^2 = (2^2 - 0^2) - \left(\frac{2^3}{3} - \frac{0^3}{3}\right) =$$

$$4 - \frac{8}{3} = \frac{12}{3} - \frac{8}{3} = \frac{4}{3}.$$

Ou seja, no problema acima calculamos uma área a partir da diferença entre duas áreas. Observar, no Exemplo acima, que

$$A = \int_0^2 (g(x) - f(x))dx,$$

por conta do Teorema 6.10.

Questão: É interessante notar que a área A da região R exemplificada acima é invariante sob a ação de translações. Translações de uma região de $\mathbb{R}^2$ podem ser feitas para a direita ou para a esquerda, para cima ou para baixo, ou por combinações de deslocamentos horizontais com verticais. Neste sentido, translações horizontais são definidas por uma operação $x - \alpha$, enquanto translações verticais são dadas por uma operação $y + \beta$.

Logo, para representarmos uma translação qualquer da região da Questão acima, basta fazer

$$f(x) = (x - \alpha)^2 + \beta$$

e

$$g(x) = 2(x - \alpha) + \beta.$$

Se $\alpha = \beta = 0$, então $f(x) = x^2$ e $g(x) = 2x$ delimitam uma região R de $\mathbb{R}^2$ cuja área é $\frac{4}{3}$, conforme já discutido.

Se $\alpha > 0$, temos uma nova região à direita de R. Se $\alpha < 0$, temos uma nova região à esquerda de R. Se $\beta > 0$, temos uma nova região acima de R. Se $\beta < 0$, temos uma nova região abaixo de R.

Os valores de α e β podem ser interpretados como translações de R ao longo de $\mathbb{R}^2$. Valores não nulos de α produzem translações horizontais, enquanto valores não nulos de β produzem translações verticais. Seja qual for a translação, a área da região delimitada por f e g deve ser invariante, ou seja

$$(x - \alpha)^2 + \beta = 2(x - \alpha) + \beta \Rightarrow (x = \alpha \vee x = 2 + \alpha)$$

e

$$\int_\alpha^{2+\alpha} (2(x - \alpha) + \beta - ((x - \alpha)^2 + \beta))dx = \frac{4}{3}.$$

✍ Recomendamos que o leitor faça as contas. Recomendamos também que faça representações visuais do problema, para desenvolver as intuições correspondentes à Questão. Uma maneira rápida de fazer isso é atribuindo valores para α e β e criar representações gráficas através do serviço gratuito em https://www.geogebra.org/.

Princípios de invariância são um dos pilares da matemática, geralmente enunciados através da identidade =. Mas esta é uma questão de grande impacto epistemológico que escapa dos nossos propósitos.

SEÇÃO 61

Logaritmo natural

SUMÁRIO
ÍNDICE
REDE

Nesta Seção mostramos como usar integral de Riemann para definir *logaritmo natural*. Graças a isso seremos capazes de qualificar posteriormente o que são *logaritmos*.

DEFINIÇÃO 6.5. *Uma função $f : x \subseteq \mathbb{R} \to \mathbb{R}$ é* localmente crescente *no ponto $c \in x$ sss existe intervalo aberto I tal que $c \in I$, $I \subseteq x$ e, para quaisquer a e b pertencentes a I,*

$$a > b \Rightarrow f(a) > f(b).$$

Uma função $f : x \subseteq \mathbb{R} \to \mathbb{R}$ é localmente decrescente *no ponto c sss existe intervalo aberto I tal que $c \in I$, $I \subseteq x$ e, para quaisquer a e b pertencentes a I,*

$$a > b \Rightarrow f(a) < f(b).$$

EXEMPLO 6.19. I: *$f : \mathbb{R} \to \mathbb{R}$, tal que*

$$f(x) = x^3,$$

é localmente crescente em qualquer ponto c de $\mathbb{R}$; com efeito, se $a > b$, então $a^3 > b^3$, independentemente de qualquer intervalo aberto de reais onde a, b e c pertençam; logo,

$$f(a) > f(b).$$

II: *$g : \mathbb{R} \to \mathbb{R}$, tal que*

$$g(x) = -6x,$$

é localmente decrescente em qualquer ponto c de $\mathbb{R}$; com efeito, se $a > b$, então $-6a < -6b$; logo;

$$f(a) < f(b),$$

independentemente de qualquer intervalo aberto de reais onde a, b e c pertençam.

III: ✍ *$h : \mathbb{R} \to \mathbb{R}$, tal que*

$$h(x) = x^2,$$

é localmente crescente em qualquer real maior do que 0 *e localmente decrescente em qualquer real menor do que* 0*; no entanto, não é localmente crescente, nem localmente decrescente, em 0.*

Funções reais f localmente crescentes (localmente decrescentes) em qualquer ponto do domínio de f são chamadas de *globalmente crescentes* (*globalmente decrescentes*).

Funções globalmente crescentes são também chamadas de *crescentes*. Comentário análogo para as globalmente descrescentes. Obviamente, qualquer função crescente (decrescente) é injetora. Com efeito, se $a < b \Rightarrow f(a) < f(b)$, no caso de f crescente, então $a \neq b \Rightarrow f(a) \neq f(b)$. Prova semelhante para as decrescentes.

Teorema 6.15. *Uma função $f : d \subseteq \mathbb{R} \to \mathbb{R}$ é globalmente crescente se $f'(x) > 0$ para todo x pertencente a d.*

Demonstração: Sabemos, pelo Teorema 5.24, que

$$\frac{d}{dx} f\Big|_{x=a} = \lim_{x \to a} \frac{f(x) - f(a)}{x - a}$$

para qualquer a do domínio de f, se f for diferenciável. Considere uma vizinhança de a definida por $(a - \delta, a + \delta)$. Se $x \in (a - \delta, a + \delta)$ e $x \neq a$, então $x > a$ ou $x < a$. Se

$$\frac{d}{dx} f\Big|_{x=a} > 0$$

e $x > a$, então $f(x) > f(a)$. Se

$$\frac{d}{dx} f\Big|_{x=a} > 0$$

e $x < a$, então $f(x) < f(a)$. Em qualquer uma das situações f é localmente crescente no ponto a.

Teorema 6.16. *Uma função $f : d \subseteq \mathbb{R} \to \mathbb{R}$ é globalmente decrescente se $f'(x) < 0$ para todo x pertencente a d.*

Demonstração: Análoga à prova anterior.

Muitos outros teoremas envolvendo funções crescentes e decrescentes com derivadas podem ser enunciados e provados. Mas o que temos acima é suficiente para os nossos propósitos.

Definição 6.6. Logaritmo natural *é uma função*

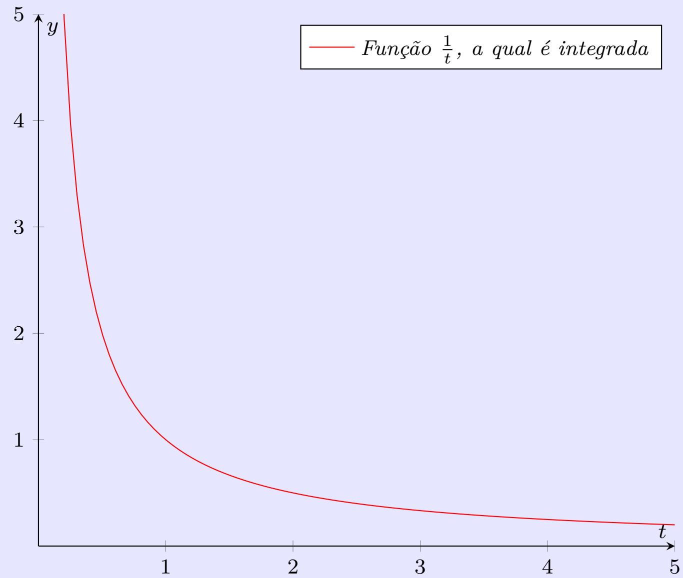
$$\ln : \{x \in \mathbb{R} \mid x > 0\} \to \mathbb{R}$$

tal que

$$\ln(x) = \int_1^x \frac{1}{t} dt.$$

Em outras palavras, logaritmo natural é uma função definida por uma integral.

Exemplo 6.20. *O logaritmo natural de* 3 *é a área da região de* $\mathbb{R}^2$ *abaixo de* $\frac{1}{t}$*, acima do eixo* t *e ladeada pelas retas* $t = 1$ *e* $t = 3$*.*



O próximo teorema mostra que nem sempre logaritmo natural assume valores reais positivos.

Teorema 6.17. *Se* $\ln$ *é a função logaritmo natural da Definição 6.6, então*

i: $\ln(1) = 0$*;*

ii: $\ln(x) > 0$ *se* $x > 1$*;*

iii: $\ln(x) < 0$ *se* $0 < x < 1$.

Demonstração: i: Definição 6.6 garante que

$$\ln(1) = \int_1^1 \frac{1}{t} dt.$$

A definição de integral de Riemann, para o caso em que o limite superior de integração é idêntico ao limite inferior, garante que

$$\int_1^1 \frac{1}{t} dt = 0.$$

ii: Se $x > 1$, então o domínio de integração de

$$\int_1^x \frac{1}{t} dt$$

é o intervalo não degenerado $[1, x]$. Mas, neste intervalo, a função integrada $f(t) = \frac{1}{t}$ assume somente imagens estritamente positivas. Logo,

$$\int_1^x \frac{1}{t} dt > 0,$$

como consequência imediata do Teorema 6.7.

iii: Se $0 < x < 1$, então o domínio de integração de

$$\int_x^1 \frac{1}{t} dt$$

é o intervalo não degenerado $[x, 1]$. Mas, neste intervalo, a função $\frac{1}{t}$ assume imagens estritamente positivas. Logo,

$$\int_x^1 \frac{1}{t} dt > 0.$$

No entanto,

$$\ln(x) = \int_1^x \frac{1}{t} dt.$$

Logo, a definição de integral de Riemann garante que

$$\int_1^x \frac{1}{t} dt < 0.$$

Isso conclui a prova dos três itens.

Teorema 6.18. *A função* ln *é injetora.*

Demonstração: Temos que

$$\frac{d}{dx}\ln(x) = \frac{d}{dx}\int_1^x \frac{1}{t}dt = \frac{1}{x},$$

por conta do item I do Teorema Fundamental do Cálculo. Logo,

$$\frac{d}{dx}\ln(x) > 0,$$

uma vez que o domínio de ln é o conjunto dos reais estritamente positivos. Isso implica que ln é crescente, de acordo com Teorema 6.15. Uma vez que qualquer função crescente é injetora, então ln é injetora.

Teorema 6.19. *A função* ln *é sobrejetora.*

Demonstração: Apresentamos apenas um esboço da prova.

Por um lado, se x é uma cópia real de um inteiro maior do que 1, então

$$\ln(x) > \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{x}.$$

Para perceber isso, basta verificar que o somatório do lado direito da desigualdade acima é uma soma de Riemann no intervalo $[1, x]$, onde cada elemento da partição do domínio de integração $[1, x]$ de $\ln(x)$ tem medida 1.

Por outro lado, a série harmônica $\sum \frac{1}{n}$, estendida para os reais, é divergente, de acordo com o Teorema 5.32. Logo,

$$\lim_{x\to\infty} \ln(x) = \infty,$$

uma vez que $\ln(x)$ é estritamente positivo para todo x à direita de 1.

Além disso, a função $\frac{1}{t}$ usada para definir logaritmo natural é simétrica em relação à reta $y = t$. Logo,

$$\lim_{x\to 0^+} \ln(x) = -\infty$$

(lembrar que $\ln(x) < 0$ para reais x no intervalo aberto $(0, 1)$).

Para finalizar, ln é diferenciável, o que implica que é contínua (Teorema 5.25). Logo, para qualquer y real existe x

tal que $\ln(x) = y$, uma vez que as imagens $\ln(x)$ da função ln percorrem todos os reais, de $-\infty$ a ∞ (i.e., todos os números reais, sejam negativos ou positivos).

Teorema 6.20. *A função* ln *admite inversa.*

Demonstração: Ver Teoremas 6.18 e 6.19, os quais implicam que ln é bijetora, bem como Teorema 4.21: toda função bijetora admite inversa.

Seção 62

A inversa de ln

Sumário
Índice
Rede

Na Seção 61 provamos que logaritmo natural admite inversa. Por outro lado, exponencial $\exp(x)$ é definida como solução de uma equação diferencial sob uma condição de contorno. No próximo teorema mostramos que logaritmo natural e exponencial são inversas uma da outra. Uma vez que o domínio de logaritmo natural é o conjunto de todos os reais estritamente positivos, o próximo resultado permite inferir que exponencial de qualquer real x jamais é negativo.

Teorema 6.21. *A inversa de* ln *é a função exponencial* exp.

Demonstração: Seja $\ln : \{x \in \mathbb{R} \mid x > 0\} \to \mathbb{R}$ tal que

$$\ln(x) = \int_1^x \frac{1}{t} dt.$$

Para fins de abreviação, chamemos $\ln(x)$ de $y(x)$. Seja

$$g : \mathbb{R} \to \{x \in \mathbb{R} \mid x > 0\}$$

a inversa de y, cuja existência é garantida pelo Teorema 6.20. Logo, para qualquer x real,

$$y(g(x)) = x,$$

de acordo com Definição 4.15.

Logo,

$$\frac{d}{dx}(y(g(x))) = \frac{d}{dx}x.$$

De acordo com o Teorema 6.3 (lembrar também que

$$\frac{d}{dx}\ln(x) = x^{-1},$$

como mostrado na prova do Teorema 6.18), temos

$$\frac{1}{g(x)}g'(x) = 1.$$

Logo, $g'(x) = g(x)$. Uma vez que $\ln(1) = 0$, logo, $g(0) = 1$. Mas esta é exatamente a definição de função exponencial dada na Seção 56. Logo, $g(x) = \exp(x)$.

Seção 63

Aplicação elementar

Sumário
Índice
Rede

Entre os elementos que ocorrem na natureza, Polonium (Polônio, em português) é o mais radioativo. Existem 42 isótopos conhecidos deste elemento descoberto em 1898 pelo casal Marie e Pierre Curie. Polonium-210 (abreviado como ^{210}Po), por exemplo, tem meia-vida de 138,376 dias (*meia-vida* de um isótopo é o tempo necessário para a sua massa reduzir à metade). Em contato com o ar, a radiação deste isótopo é visível a olho nu, emitindo uma luminosidade azulada. Quanto tempo demora para que um quilograma de ^{210}Po seja reduzido a um grama?

Neste caso, podemos modelar matematicamente o fenômeno de decaimento radioativo através do emprego de funções. Se mapearmos massa m através de uma função real $m(t)$ dependente de tempo t, o modelo usual assume que

$$\frac{d}{dt}m = km,$$

sendo k uma constante de proporcionalidade cujo valor depende do material em processo de decaimento. Essa proposta se traduz da seguinte maneira:

A taxa de variação de massa em relação ao tempo
é proporcional à massa.

Em outras palavras, quanto maior a massa, maior a taxa de variação da massa em relação ao tempo.

Logo,

$$m(t) = c\exp(kt),$$

onde c é uma constante real que define uma condição de contorno.

Com efeito, a função acima satisfaz a equação diferencial

$$\frac{d}{dt}m = km.$$

Para o caso em que $t = 0$, temos $m(0) = c$. Logo, c pode ser interpretado como massa inicial m_0. Logo,

$$m(t) = m_0 \exp(kt).$$

Uma vez que a meia-vida do isótopo em questão é de apenas 138,376 dias, logo

$$\frac{m_0}{2} = m_0 \exp(138, 376k).$$

Portanto, $\exp(138, 376k) = 0, 5$. Logo, $\ln(\exp(138, 376k)) = \ln(0, 5)$, o que implica que $138, 376k = -0, 693147$. Finalmente,

$$k = -0, 00500916d^{-1},$$

sendo que d denota 'dias' e d^{-1} denota 'por dia'.

O valor da constante de proporcionalidade k é negativo justamente porque, no problema em questão, a taxa de variação

$$\frac{dm}{dt}$$

é negativa, sendo $m(t)$ sempre positivo. Ou seja, está ocorrendo *perda* de massa ao longo do tempo.

Uma vez determinada a constante de proporcionalidade k do modelo usual, para descrever decaimento radioativo, podemos responder à questão proposta.

Temos que

$$1 = 1000 \exp(-0, 00500916t),$$

uma vez que queremos determinar o tempo t consumido (em dias) para transformar mil gramas de ^{210}Po em um grama.

Isso implica em

$$\exp(-0, 00500916t) = 0, 001.$$

Logo,

$$\ln(\exp(-0,00500916t)) = \ln(0,001) = -6,90775.$$

Isso implica que

$$-0,00500916t = -6,90775.$$

Logo,

$$t = 1379,02d,$$

o que corresponde a $3,77556$ anos (três anos, nove meses e dez dias).

Em menos de quatro anos um quilograma de Polonium-210 é reduzido a um grama.

Observar que um grama de Polonium-210 é suficiente para matar cinquenta milhões de pessoas, e adoecer outras cinquenta milhões, por envenenamento radioativo [36].

✍ Sabendo que a meia-vida de ^{14}C (isótopo Carbono 14) é de 5730 anos, qual é a massa final de dois gramas deste isótopo após oitenta milhões de anos? Para resolver este problema empregue o modelo usual de decaimento radioativo, o qual assume que a taxa de variação de massa em relação à passagem de tempo é proporcional à massa.

STOP O exercício acima é algo que pode ser divertido para reflexões. Por um lado, se o leitor encarar a questão de um ponto de vista puramente matemático, perceberá que será necessário calcular a exponencial de um valor real com ordem de grandeza 10^3. No entanto, calculadoras científicas usualmente não contam com capacidade de processamento para esse tipo de conta. Se o leitor tentar empregar uma calculadora científica típica, não será capaz de obter uma resposta para, digamos, exponencial de 9000. Esta, portanto, é uma ótima oportunidade para a natureza humana demonstrar sua capacidade criativa. Com efeito, 9000 é a adição de 90 com 90, com cem ocorrências da parcela 90. Calculadoras científicas conseguem processar a exponencial de 90. Uma vez que a exponencial de uma soma é o produto de exponenciais (Teorema 6.5), agora o problema passa a ser fácil de resolver, numa parceria entre tecnologia e espírito humano.

STOP Por outro lado, o exercício proposto é um problema físico. Problemas de caráter físico não podem ser resolvidos levando em conta apenas aspectos matemáticos. Com efeito, processos de datação por Carbono-14 não são confiáveis para períodos tão longos quanto os oitenta milhões de anos sugeridos. Logo, rogamos ao leitor que pense com bastante carinho sobre a questão levantada. Ciência não se sustenta por manuais técnicos que ditam normas a serem incondicionalmente cumpridas. Ciência é uma atividade de profunda responsabilidade intelectual.

Seção 64

Um olhar sobre o paraíso

Sumário
Índice
Rede

Seja p um número real maior do que zero. Logo,

$$\frac{d}{dx}\ln(px) = \frac{1}{px}p = \frac{1}{x} = \frac{d}{dx}\ln(x).$$

Usamos acima derivada de função composta (Teorema 6.3), além do fato de que

$$\frac{d}{dx}\ln(x) = x^{-1},$$

conforme demonstração do Teorema 6.18.

Isso significa que ambas as funções $\ln(px)$ e $\ln(x)$ têm a mesma derivada

$$\frac{1}{x}.$$

Logo,

$$\ln(px) = \ln(x) + C,$$

onde C é uma constante real.

Se $x = 1$, temos $\ln(p) = \ln(1) + C$. Logo, $C = \ln(p)$. Consequentemente,

$$\boxed{\ln(px) = \ln(x) + \ln(p).}$$

Ou seja, foi provado acima que o logaritmo natural de um produto px entre fatores reais estritamente positivos p e x é igual à adição do logaritmo natural de p com o logaritmo natural de x. Em jargão popular (mais semelhante a um bordão popular nos dias de hoje), *logaritmo natural do produto é a soma de logaritmos naturais.*

Por outro lado,

$$\frac{d}{dx}\ln\left(\frac{1}{x}\right) = \frac{1}{1/x}(-x^{-2}) = \frac{-x}{x^2} = \frac{-1}{x} = \frac{d}{dx}(-\ln(x)).$$

Em outras palavras, ambas as funções $\ln\left(\frac{1}{x}\right)$ e $-\ln(x)$ têm a mesma derivada

$$\frac{-1}{x}.$$

Logo,

$$\ln\left(\frac{1}{x}\right) = -\ln(x) + C.$$

Se $x = 1$, então $C = 0$. Portanto,

$$\ln\left(\frac{1}{x}\right) = -\ln(x).$$

Uma vez que

$$\ln\left(\frac{x}{p}\right) = \ln\left(x\frac{1}{p}\right),$$

então

$$\ln\left(\frac{x}{p}\right) = \ln(x) + \ln\left(\frac{1}{p}\right) = \ln(x) - \ln(p).$$

Ou seja, *logaritmo natural de uma razão é a diferença de logaritmos naturais*.

Nos exemplos que seguem o Teorema 6.5 mostramos que exponencial da soma é o produto de exponenciais das parcelas da soma. Aqui, por conta do fato de logaritmo natural ser a inversa da exponencial, mostramos que o logaritmo natural de um produto é a soma dos logaritmos naturais dos fatores desse produto. Além disso, logaritmo natural de uma razão é a diferença entre os logaritmos naturais dos termos da razão.

Considere agora a equação diferencial

$$\frac{dy}{dt} = ky,$$

a qual é exatamente a mesma que foi utilizada na solução do decaimento radioativo de ^{210}Po, na Seção 63.

Logo,

$$\frac{1}{y}\frac{dy}{dt} = k.$$

Para que possamos passar da *forma diferencial* acima para uma *forma integral* (isso por conta do Teorema Fundamental do Cálculo), basta percebermos que

$$\int_{y_0}^{y_F} \frac{1}{y} dy = \lim_{\|P_y\|\to 0} \sum_i \frac{1}{z_i}\Delta y_i = \lim_{\|P_t\|\to 0} \sum_i \frac{1}{z_i}\frac{\Delta y_i}{\Delta t_i}\Delta t_i =$$

$$\int_{t_0}^{t_F} \frac{1}{y}\frac{dy}{dt} dt = \int_{t_0}^{t_F} k dt,$$

sendo que P_y e P_t denotam partições nos eixos y e t, respectivamente.

Lembrar que estamos sempre assumindo que $y = y(t)$, ou seja, y é uma função de t (ou seja, os termos do domínio de y são chamados de t).

STOP Alguns autores justificam a passagem da forma diferencial para a integral de maneira muito mais breve, porém falaciosa:

$$\frac{dy}{dt} = ky$$

implica em

$$\frac{1}{y} dy = k dt$$

que, por sua vez, implica em

$$\int_{y_0}^{y_F} \frac{1}{y} dy = \int_{t_0}^{t_F} k dt.$$

A passagem da primeira para a segunda fórmula (antes de 'concluir' a forma integral) sugere que $\frac{dy}{dt}$ é uma razão entre reais dy e dt. No entanto,

$$\frac{dy}{dt}$$

não é uma razão entre números reais!

Logo, esta estratégia (comumente empregada em textos de física teórica e engenharia, na qual $\frac{dy}{dt}$ é tratada como uma razão), apesar de funcionar como regra mnemônica, não consiste em justificativa no contexto de cálculo diferencial e integral padrão.

Alguns autores chegam a se referir a dy e dt como *infinitesimais*, sendo que no cálculo padrão **não** há infinitesimais (ver Seção 52 sobre o tema). O conceito de infinitesimal é típico, e extremamente importante, em duas outras teorias de cálculo diferencial e integral que não necessitam de limites para qualificar derivadas e integrais. Essas formas diferentes de cálculo diferencial e integral são *análise não standard* [40] e *análise infinitesimal suave* [4].

Agora que sabemos que

$$\frac{dy}{dt} = ky$$

implica em

$$\int_{y_0}^{y_F} \frac{1}{y} dy = \int_{t_0}^{t_F} k dt,$$

temos que

$$\ln(y)\Big|_{y_0}^{y_F} = kt\Big|_{t_0}^{t_F},$$

por aplicação do Teorema Fundamental do Cálculo.

Logo, $\ln(y_F) - \ln(y_0) = k\Delta t$, sendo $\Delta t = t_F - t_0$. Logo, $\ln(y_F/y_0) = k\Delta t$. Portanto, $y_F/y_0 = \exp(k\Delta t)$, ou seja,

$$y_F = y_0 \exp(kt),$$

se assumirmos que $t_F = t$ e $t_0 = 0$. Observar que esta é exatamente a solução para o problema de decaimento radioativo de ^{210}Po.

Ou seja, a forma integral

$$\int_{y_0}^{y_F} \frac{1}{y} dy = \int_{t_0}^{t_F} k dt$$

(por *separação de variáveis*, i.e., todas as ocorrências de y estão do mesmo lado da igualdade e todas as ocorrências de t estão do outro lado) da equação diferencial

$$\frac{dy}{dt} = ky$$

(a qual define a exponencial de kt com condições de contorno $y(0) = y_0$) se mostra solúvel através da definição de logaritmo natural via integração de Riemann.

O problema de decaimento radioativo é revisitado sob um ponto de vista completamente diferente na Seção 95.

Essa discussão ilustra o papel de derivadas e integrais, anunciado no primeiro parágrafo da Seção 47. Derivadas permitem modelar (via equações diferenciais) fenômenos físicos (como, e.g., decaimento radioativo). Integrais, por outro lado, viabilizam previsões de longo termo, as quais são mapeadas por funções que são soluções de equações diferenciais. Este é um dos papeis principais do Teorema Fundamental do Cálculo: viabilizar soluções de equações diferenciais via processo de integração.

Tudo isso é formulado em uma teoria de conjuntos sustentada por apenas duas 'colunas': igualdade $=$ e pertinência $\in$. Portanto, aqui radica parte do valor estético de ZF: dois conceitos apenas, $=$ e $\in$, abrem portas para um vasto universo de possibilidades para estudos e aplicações.

Em 1926, oito anos após a morte de Cantor, David Hilbert afirmou:

> *Ninguém poderá nos expulsar do paraíso que Cantor criou para nós.*

De fato, o paraíso de Cantor ainda está sendo conhecido, lentamente, por milhares de matemáticos do mundo todo. Foi este paraíso que inspirou Ernst Zermelo, Abraham Fraenkel, John von Neumann, Kurt Gödel e muitos outros, até os dias de hoje. Mesmo sem sabermos ao certo o que é possível fazer em terras tão exóticas, até o presente momento já temos uma boa noção de sua extraordinária beleza.

No romance *Princess Napraxine*, a escritora britânica Ouida afirma que

> *familiaridade é um mágico cruel com a beleza, mas gentil com a feiura.*

Em outras palavras, o belo deve resistir à familiaridade.

Nesta acepção, ZF é bela. Ainda não há perspectivas de plena familiaridade com o seu poder de alcance.

Nenhum teorema sobre teoria de conjuntos é atribuído a Hilbert, o primeiro grande defensor da teoria de conjuntos. Mas este exerceu uma poderosa influência sobre muitos outros que decidiram conhecer o paraíso concebido por Cantor. Hilbert foi possivelmente o último matemático de visão universal sobre este ramo do conhecimento. O tom profético de sua visão sobre o que é importante em matemática repercute até os dias de hoje. Mas esta é outra longa história não cabível neste livro.

Seção 65

Pseudomatemática

Sumário
Índice
Rede

Na Seção anterior exibimos um exemplo de prática de má qualidade no contexto de cálculo diferencial e integral padrão, ao mostrarmos a não justificabilidade da simplificação

$$\frac{dy}{dt} = ky \Rightarrow \frac{1}{y}dy = kdt.$$

Este é um exemplo de *pseudomatemática*, uma atividade muito comum na qual há a tentativa de imitar procedimentos matemáticos sem qualquer atenção a rigor ou razão.

Até mesmo Thomas Hobbes (um dos pais da filosofia política) foi vítima de si mesmo, por conta de práticas matemáticas sem qualquer fundamentação racional. Hobbes acreditava ter resolvido o insolúvel problema da quadratura do círculo. Essa questão deu origem à famosa controvérsia entre Hobbes e John Wallis, a qual durou décadas, durante o século 17.

Frequentemente matemáticos do mundo todo são importunados por pessoas que negam o Argumento da Diagnonal de Cantor ou os Teoremas de Incompletude de Gödel, entre outras sandices.

O termo 'pseudomatemática' foi cunhado pelo lógico Augustus De Morgan, em 1915. Nas palavras de De Morgan:

> O pseudomatemático é uma pessoa que lida com a matemática como um macaco que brinca com uma navalha. A criatura tenta se barbear, imitando seu mestre; mas, sem qualquer noção sobre o ângulo em que a lâmina deve ser posicionada, acaba cortando a própria garganta.

Porém, não há qualquer procedimento efetivo que permita discernir matemática de pseudomatemática. Se houvesse, possivelmente o destino da matemática poderia ser entregue às máquinas. Não podemos esquecer as duras críticas que Cantor recebeu por sua teoria de conjuntos. Na visão de Kronecker, por exemplo, o que Cantor propôs não era algo digno de atenção.

Como dizia Cantor,

A essência da matemática radica em sua liberdade.

Ou seja, pode haver raras vantagens ao ouvirmos os ingênuos. Ainda assim, é recomendável que o leitor sempre tome muito cuidado.

Aos curiosos, há um excelente livro sobre pseudomatemática, de Underwood Dudley [14].

Seção 66

Quanto é a^x?

Sumário
Índice
Rede

Aqui respondemos a uma das questões da Introdução.

Definição 6.7. Número de Euler *é o número real e tal que*

$$\ln(e) = 1.$$

Logo, a definição acima garante que $e > 1$ (Teorema 6.17). Existem várias técnicas para calcular o número de Euler e em sua representação decimal usual. Uma delas faz uso do fato de que exponencial é a inversa de logaritmo natural. Logo, $\exp(1) = e$. Portanto,

$$e = 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \cdots .$$

O truncamento desta série, obtido pela soma das primeiras cem mil parcelas, i.e.,

$$\sum_{n=0}^{99999} \frac{1}{n!},$$

nos fornece um valor aproximado de

$$e \approx 2,71827.$$

Leonhard Euler provou a irracionalidade de e. O Teorema de Lindemann-Weierstraß prova que e é *transcendente*, ou seja, não existe equação polinomial com coeficientes reais racionais tal que e seja solução desta equação. Todo número real transcendente é irracional, apesar da recíproca desta última afirmação não ser teorema.

Exemplo 6.21. *Já provamos anteriormente que* $\sqrt{2}$ *é irracional. No entanto,* $\sqrt{2}$ *é solução da equação polinomial*

$$x^2 - 2 = 0,$$

cujos coeficientes são todos racionais. Logo, $\sqrt{2}$ *não é um número real transcendente.*

Números reais não transcendentes são chamados de *algébricos*.

Exemplo 6.22.

i: $\sqrt{2}$ *é um real algébrico, uma vez que é solução da equação*

$$x^2 - 2 = 0;$$

ii: 2 *é um real algébrico, uma vez que é solução da equação* $x - 2 = 0$.

Logo, reais algébricos podem ser racionais ou irracionais.

Não é uma tarefa fácil provar que um número real qualquer é irracional ou transcendente. Por exemplo, até hoje não se sabe se os reais $e\pi$, $e + \pi$ ou $\pi - e$ (entre muitos outros) são irracionais ou transcendentes.

Relembrando conceitos já vistos aqui, sejam a um número real e n um inteiro estritamente positivo. Logo, a^n é o produto de a por a com n ocorrências de a. Se n é um inteiro negativo, então a^n é o simétrico multiplicativo de a^{-n}, desde que a seja diferente de 0. Se a é um número real diferente de zero, então $a^0 = a^{n-n} = \frac{a^n}{a^n} = 1$. Se a é um número real qualquer e n é um inteiro não nulo, então $a^{\frac{1}{n}} = b$ sss $b^n = a$. Neste caso denotamos $a^{\frac{1}{n}}$ como $\sqrt[n]{a}$. Se p e q são inteiros tais que $q \neq 0$, então $a^{\frac{p}{q}} = \sqrt[q]{a^p}$.

A extensão de potências a^n, de inteiros n para racionais $n = \frac{p}{q}$, foi introduzida por John Wallis, em seu livro *Arithmetica Infinitorum*, de 1656.

A questão que devemos responder agora é o conceito de a^x, para x um real qualquer, de modo que a^x seja consistente com os casos já discutidos até aqui, nos quais x é um racional. Observar que nem sempre existe número real y tal que $y = a^{\frac{p}{q}}$ se $a < 0$.

Sejam x um número real estritamente positivo e a um número real racional. Logo,

$$\frac{d}{dx}\ln(x^a) = \frac{1}{x^a}ax^{a-1} = \frac{a}{x} = a\frac{d}{dx}\ln(x) = \frac{d}{dx}a\ln(x).$$

Isso demonstra que as funções reais $\ln(x^a)$ e $a\ln(x)$ têm a mesma derivada

$$\frac{a}{x},$$

se $x > 0$ e a é um real racional.

Logo,

$$\ln(x^a) = a\ln(x) + C,$$

onde C é uma constante real.

Se $x = 1$, então $C = 0$. Portanto,

$$\ln(x^a) = a\ln(x).$$

Uma vez que $x^a = \exp(\ln(x^a))$ (por conta do Teorema 6.21), então

$$x^a = \exp(a\ln(x)),$$

se $x > 0$ e a é um real racional.

Isso significa que uma definição para a^x, assumindo x um real qualquer, deve ser consistente com o teorema dado acima. Esta é a estratégia adotada na próxima definição.

Definição 6.8. *Sejam a um número real estritamente positivo e x um número real qualquer. Então*

$$a^x = \exp(x\ln(a)).$$

Uma vez que ln e exp são funções reais já definidas e a última definição é consistente com o teorema

$$(a > 0 \wedge x \in \mathbb{Q}) \Rightarrow a^x = \exp(x\ln(a)),$$

então fomos bem sucedidos na conceituação de a^x para $a > 0$ e x real. Observar que o símbolo $\mathbb{Q}$ foi usado aqui como notação abusiva, uma vez que estamos tratando com reais racionais.

Exemplo 6.23. i: $5^{\sqrt{2}} = \exp(\sqrt{2}\ln(5))$;

ii: $e^x = \exp(x\ln(e))$; *logo,* $e^x = \exp(x)$; *com efeito,* $\ln(e) = 1$.

O segundo item do Exemplo acima justifica a prática comum de escrever a exponencial de x simplesmente como e^x, sendo e o número de Euler. Além disso, é prática comum ler e^x como ‘exponencial de x’. Logo,

Teorema 6.22. *$a^x = e^{x\ln(a)}$, se $a > 0$ e x é um real qualquer.*

O próximo resultado, conhecido como *a identidade de Euler*, estabelece uma inesperada relação entre π (uma constante da geometria euclidiana plana), o número de Euler e a unidade imaginária i dos complexos.

Teorema 6.23 (Identidade de Euler).

$$e^{i\pi} = -1.$$

Demonstração: Basta usar o teorema $\exp(x) = e^x$, estendido para os complexos, em parceria com o Teorema de Euler, demonstrado na Seção 57.

Frequentemente a Identidade de Euler é mencionada como exemplo de profunda beleza matemática. Isso por conta de uma inesperada e elegante conexão entre um conceito geométrico (a razão π entre o perímetro de uma circunferência e seu diâmetro), um conceito analítico (o número de Euler) e um conceito algébrico (a unidade imaginária). Apesar deste teorema não estar enunciado em qualquer trabalho publicado de Leonhard Euler, parece evidente que ele conhecia o resultado [59].

Teorema 6.24. *Se $a > 0$ e x é um real qualquer, então*

$$\frac{d}{dx}a^x = \ln(a)a^x.$$

Demonstração:

$$\frac{d}{dx}a^x = \frac{d}{dx}e^{x\ln(a)} = \ln(a)e^{x\ln(a)} = \ln(a)a^x,$$

de acordo com Teoremas 6.22 e 6.3.

Exemplo 6.24. i:

$$\frac{d}{dx}5^x = \ln(5)5^x;$$

ii:

$$\int_a^b 5^x dx = \frac{5^x}{\ln(5)}\bigg|_a^b;$$

por conta do Teorema Fundamental do Cálculo e Teorema 6.24.

Lembrar que a equação diferencial

$$\frac{dy}{dx} = ky,$$

com a condição de contorno

$$y(0) = y_0,$$

admite como solução única a função

$$y(x) = y_0 e^{kx}.$$

No entanto, se $k = \ln(a)$ (i.e., $a = e^k$), a última afirmação é equivalente a dizer que

$$y(x) = y_0 a^x$$

é solução única da mesma equação diferencial. Em particular, o problema do decaimento do isótopo Polonium-210 (Seção 63) pode ser alternativamente modelado como

$$m(t) = m_0 0,995003^t,$$

uma vez que

$$k = -0,00500916 d^{-1}$$

e

$$e^k = 0,995003.$$

Teorema 6.25. *A combinação linear de soluções quaisquer da equação diferencial $y' = ky$ também é solução da mesma equação.*

Demonstração: Devemos provar que, se y_1 e y_2 são soluções de $y' = ky$, então qualquer função definida por

$$c_1 y_1 + c_2 y_2$$

também é solução da mesma equação, onde c_1 e c_2 são reais quaisquer.

Sejam y_1 e y_2 soluções da equação diferencial $y' = ky$. Logo

$$y_1' = ky_1 \quad \text{e} \quad y_2' = ky_2.$$

Logo,

$$c_1 y_1' = c_1 k y_1,$$

onde c_1 é um real qualquer.

Por conta do Teorema 5.20 sobre derivada de constante multiplicada por função, temos que

$$(c_1 y_1)' = k(c_1 y_1).$$

Isso prova que $c_1 y_1$ é solução da equação diferencial $y' = ky$. Analogamente, temos que

$$y_2' = k y_2$$

implica em

$$(c_2 y_2)' = k(c_2 y_2),$$

onde c_2 é um real qualquer. Isso prova que $c_2 y_2$ também é solução da equação diferencial $y' = ky$.

Se somarmos ambos os lados de $(c_1 y_1)' = k(c_1 y_1)$ por um mesmo termo, a nova igualdade se mantém como teorema. Logo,

$$(c_1 y_1)' + (c_2 y_2)' = k(c_1 y_1) + k(c_2 y_2).$$

Portanto, de acordo com Teorema 5.21 sobre derivada da soma de funções,

$$(c_1 y_1 + c_2 y_2)' = k(c_1 y_1 + c_2 y_2).$$

Isso prova que a combinação linear

$$c_1 y_1 + c_2 y_2$$

de y_1 com y_2 é solução de $y' = ky$.

O último teorema pode ser generalizado para uma vasta gama de equações diferenciais conhecidas na literatura como *equações diferenciais lineares homogêneas*. Esse resultado tem significativo impacto no estudo de equações diferenciais tanto lineares quanto não lineares, no sentido de que resultados de álgebra linear podem ser aproveitados no estudo de equações diferenciais. Essa questão é discutida na Parte 8. O que podemos adiantar é que o resultado acima significa que o conjunto de soluções da equação diferencial

$$y' = ky$$

define um *espaço vetorial real* de *uma dimensão*, onde os *vetores* são funções reais.

Seção 67

Logaritmo

Sumário
Índice
Rede

Sabemos que

$$1^x = e^{x \ln(1)} = e^0 = 1.$$

Isso implica que $f : \mathbb{R} \to \mathbb{R}$ dada por

$$f(x) = 1^x$$

é uma função constante e, portanto, não injetiva. Não obstante, funções definidas por a^x, com $a \neq 1$, contam com comportamento bem diferente.

Teorema 6.26. *Se $a > 0$ e $a \neq 1$, então $f : \mathbb{R} \to \mathbb{R}$, dada por*

$$f(x) = a^x,$$

é crescente para $a > 1$ e decrescente para $a < 1$.

Demonstração: Temos que $f(x) = a^x = e^{x \ln(a)}$. Uma vez que a exponencial de qualquer número real é estritamente positiva, então a^x é estritamente positiva. Além disso,

$$\frac{d}{dx} a^x = \frac{d}{dx} e^{x \ln(a)} = \ln(a) e^{x \ln(a)} = \ln(a) a^x.$$

Logo, se $a > 1$, então

$$\left.\frac{d}{dx} a^x\right|_{x=b} > 0,$$

para todo b real (função f é crescente, de acordo com Teorema 6.15).

Se $0 < a < 1$, então

$$\left.\frac{d}{dx} a^x\right|_{x=b} < 0,$$

para todo b real (função f é decrescente, de acordo com Teorema 6.16).

O último teorema deixa claro que $f(x) = a^x$ é injetiva se $a \neq 1$. Logo, se definirmos

$$f : \mathbb{R} \to \{x \in \mathbb{R} \mid x > 0\},$$

tal que

$$f(x) = a^x$$

e $a \neq 1$, então f é bijetora e, portanto, inversível. Este fato viabiliza a definição de logaritmos.

Definição 6.9.

$$\log_a(n) = x \ \vdots\ a^x = n,$$

se a e n são números reais estritamente positivos e $a \neq 1$.

Lemos $\log_a(n)$ como '*logaritmo* de n na base a'. Se não houver risco de confusão, podemos escrever $\log_a(n)$ como $\log_a n$. Logo,

$$\log_e n = \ln n,$$

se e é o número de Euler. Neste sentido, logaritmo natural passa a ser um caso particular de logaritmo.

Teorema 6.27. *Para todo real $a > 0$ tal que $a \neq 1$,*

$$\log_a a = 1.$$

Demonstração: $\log_a a = x$ sss $a^x = a$. Mas $a^x = e^{x\ln(a)}$. Uma vez que a^x é injetiva para $a \neq 1$ e $x = 1$ é solução da equação $e^{x\ln(a)} = a$, logo essa solução é única. Logo, $\log_a a = 1$.

Teorema 6.28. *Se $a > 0$, $a \neq 1$, $m > 0$ e $n > 0$, então*

$$\log_a(mn) = \log_a(m) + \log_a(n).$$

Demonstração: $\log_a(m) = x$ sss $a^x = m$. Logo,

$$e^{x\ln(a)} = m.$$

Logo, $x\ln(a) = \ln(m)$, o que implica em

$$x = \frac{\ln(m)}{\ln(a)}.$$

Ou seja,

$$\log_a(m) = \frac{\ln(m)}{\ln(a)} \quad \text{e} \quad \log_a(n) = \frac{\ln(n)}{\ln(a)}.$$

Logo,

$$\log_a(mn) = \frac{\ln(mn)}{\ln(a)} = \frac{\ln(m)+\ln(n)}{\ln(a)} =$$

$$\frac{\ln(m)}{\ln(a)} + \frac{\ln(n)}{\ln(a)} = \log_a(m) + \log_a(n).$$

Isso encerra a prova. Notar que usamos o teorema sobre logaritmo natural de um produto.

✍ Analogamente, é possível provar que

$$\log_a\left(\frac{m}{n}\right) = \log_a(m) - \log_a(n).$$

Teorema 6.29 (Mudança de Base). *Se $a > 0$, $b > 0$, $a \neq 1$, $b \neq 1$ e $m > 0$, então*

$$\log_b(m) = \log_a(m)\log_b(a).$$

Demonstração:

$$\frac{\log_b(m)}{\log_a(m)} = \frac{\ln(m)}{\ln(b)}\frac{\ln(a)}{\ln(m)} = \frac{\ln(a)}{\ln(b)} = \log_b(a).$$

Logo,

$$\log_b(m) = \log_a(m)\log_b(a).$$

Equivalentemente,

$$\log_a(m) = \log_b(m)/\log_b(a).$$

O último resultado acima é conhecido como *Teorema de Mudança de Base de Logaritmos*.

Observar que

$$\log_a(a^x) = \frac{\ln(a^x)}{\ln(a)} = \frac{\ln(e^{x\ln(a)})}{\ln(a)} = \frac{x\ln(a)}{\ln(a)} = x,$$

se $a \neq 1$.

Logo, $\log_a(x)$ é a inversa de a^x. Por conta do Teorema 6.26, isso implica que a função

$$\log_a : \{x \in \mathbb{R} \mid x > 0\} \to \mathbb{R}$$

dada por

$$\log_a(x)$$

é injetiva, se $a \neq 1$.

Seção 68

Logaritmo como isomorfismo entre grupos

Sumário

Índice

Rede

✂ O leitor pode ignorar esta Seção, sem prejuízo para o restante da leitura. A discussão aqui apenas coloca uma perspectiva puramente algébrica para logaritmos.

Na visão do policéfalo Nicolas Bourbaki, matemática é o estudo de três *estruturas-mãe*:

- algébricas,
- topológicas e
- de ordem.

Grosso modo, estruturas algébricas se referem a conjuntos munidos de operações, como adição e multiplicação entre reais. Estruturas topológicas são aquelas que tratam de noções sobre 'vizinhança', 'proximidade'. Tais conceitos podem ser formulados através de conjuntos munidos de *topologias*. Finalmente, estruturas de ordem são conjuntos munidos de relações de ordem (parcial, total, entre outras).

Apesar de, hoje em dia, esta ser uma visão *démodé* (até porque a teoria de conjuntos de Bourbaki conta com formulação diferente da teoria ZF), ela pode ser uma primeira aproximação interessante para uma ampla visão sobre matemática. Cálculo diferencial e integral padrão, por exemplo, pode ser percebido como uma combinação dessas três grandes estruturas. No entanto, resultados de cálculo diferencial e integral podem ser observados pelo prisma de uma única dessas estruturas. Daí o exemplo abaixo, o qual promove uma avaliação puramente algébrica sobre logaritmos.

Magmas, *monoides*, *grupos*, *anéis*, *corpos* e *espaços vetoriais*, entre outros exemplos, são estruturas algébricas bem conhecidas. Exploramos brevemente *grupos* e sua relação com logaritmos.

Um *grupo* $\mathfrak{G}$ é uma tripla ordenada $\mathfrak{G} = (g, \star, e)$ que satisfaz os seguintes axiomas:

G1: $g \neq \varnothing$;

G2: $\star : g \times g \to g$ é uma função; abreviamos $\star(a, b) = c$ como $a \star b = c$;

G3: $e \in g$;

G4: $\forall a(a \in g \Rightarrow (a \star e = e \star a = a))$;

G5: $\forall a \forall b \forall c((a \in g \wedge b \in g \wedge c \in g) \Rightarrow ((a \star b) \star c = a \star (b \star c)))$;

G6: $\forall a(a \in g \Rightarrow \exists b(b \in g \wedge a \star b = b \star a = e))$; abreviamos b como a^{-1}.

Por abuso de linguagem, é usual se referir ao conjunto g como grupo. Neste contexto, é comum autores afirmarem que um grupo é um conjunto g munido de uma operação binária $\star$ e de um elemento privilegiado e, que satisfaz os axiomas acima listados. Adotamos aqui o mesmo abuso de linguagem, de agora em diante.

Do ponto de vista intuitivo, os axiomas dizem o seguinte.

G1: todo grupo g é um conjunto não vazio;

G2: o grupo g é munido de uma operação binária $\star$ *fechada* em g, ou seja, para quaisquer elementos a e b de g, $a \star b$ é um elemento de g; do ponto de vista da linguagem de ZF, $\star$ é tão somente uma função com domínio $g \times g$ e co-domínio g;

G3: o elemento privilegiado e pertence ao grupo g;

G4: o elemento privilegiado e é neutro à direita e neutro à esquerda, relativamente à operação $\star$;

G5: a operação $\star$ é associativa;

G6: todo elemento a do grupo g admite um simétrico à direita e um simétrico à esquerda, relativamente à operação $\star$.

De maneira mais resumida, um grupo g é um conjunto não vazio, munido de uma operação binária fechada e associativa, com elemento neutro e elementos simétricos. A função $\star$ de um grupo é comumente chamada de *ação* do grupo.

Exemplo 6.25. i: *Seja*

$$R^+ = \{r \in \mathbb{R} \mid r > 0\}.$$

Logo, $(R^+, \cdot, 1)$ *é um grupo, se* $\cdot$ *é a multiplicação usual entre números reais; neste caso, estamos interpretando* g *como* R^+, $\star$ *como* $\cdot$, e *como* 1.

II: $(\mathbb{R}, +, 0)$ *é um grupo.*

III: $(\omega, +, 0)$ *não é um grupo, se* $+$ *for a adição usual entre naturais. Com efeito, axioma G6 não é satisfeito. Por exemplo, não há simétrico de 2 relativamente a* $+$.

DEFINIÇÃO 6.10. *Sejam*

$$\mathfrak{G} = (g, \star, e) \quad e \quad \mathfrak{G}' = (g', \star', e')$$

grupos. Dizemos que h *é um* isomorfismo *entre os grupos* $\mathfrak{G}$ *e* $\mathfrak{G}'$ *sss* $h : g \to g'$ *é uma bijeção tal que*

$$h(a \star b) = h(a) \star' h(b)$$

para todos a *e* b *pertencentes a* g.

Ou seja, isomorfismos entre grupos g e g' são bijeções $f : g \to g'$ que mantém *invariantes* as ações dos grupos envolvidos. Intuitivamente falando, tanto faz se operarmos $a \star b$ e então aplicarmos h para obter $h(a \star b)$, ou aplicarmos h sobre a e b para, somente então, operarmos $h(a) \star' h(b)$, obtemos sempre o mesmo resultado.

Como discutido na Seção 41, nenhum inteiro é racional, apesar de racionais copiarem os inteiros. No entanto, essa cópia dos inteiros entre racionais pode ser mapeada pelos inteiros (ou vice-versa), através de um isomorfismo entre dois grupos, conforme o próximo EXEMPLO. Isso porque a linguagem usada para definir inteiros e racionais é a mesma, a saber, a linguagem de ZF.

EXEMPLO 6.26. *Seja*

$$\mathbb{Q}_{\mathbb{Z}} = \{r \in \mathbb{Q} \mid r \textit{ copia um inteiro}\}.$$

Logo, $(\mathbb{Q}_{\mathbb{Z}}, +', 0')$ *é um grupo, onde* $+'$ *é a adição entre racionais e* $0'$ *é o neutro aditivo entre racionais. Ademais,* $(\mathbb{Z}, +, 0)$ *também é um grupo, sendo que* $+$ *é a adição entre inteiros e* 0 *é o neutro aditivo entre inteiros. Empregamos os símbolos* $+$ *e* $+'$*, bem como* 0 *e* $0'$*, para destacar que são conceitos distintos.*

✍ *Consideremos agora a função* $h : \mathbb{Q}_{\mathbb{Z}} \to \mathbb{Z}$ *dada por*

$$h(r) = s \Leftrightarrow r \textit{ copia o inteiro } s.$$

Logo, h *é um isomorfismo entre os grupos* $(\mathbb{Q}_{\mathbb{Z}}, +', 0')$ *e* $(\mathbb{Z}, +, 0)$. *Cabe ao leitor mostrar os detalhes, a partir do que foi discutido em Seções anteriores.*

Para efeitos práticos, isso corresponde a dizer que os inteiros, munidos de adição, são algebricamente indiscerníveis dos racionais que copiam os inteiros, munidos de adição entre eles. Em particular, do ponto de vista algébrico (onde a álgebra é examinada por propriedades de grupos), o neutro aditivo racional é indiscernível do neutro aditivo inteiro.

O próximo teorema oferece um exemplo de isomorfismo muito mais interessante, uma vez que revela logaritmo como um isomorfismo entre grupos.

Teorema 6.30. *Se $a > 0$ e $a \neq 1$, então a função $\log_a$ define um isomorfismo entre os grupos*

$$(R^+, \cdot, 1) \quad e \quad (\mathbb{R}, +, 0)$$

do Exemplo 6.25.

Demonstração: De acordo com o Teorema 6.28,

$$\log_a(m \cdot n) = \log_a(m) + \log_a(n).$$

Além disso, $\log_a$ é uma bijeção $\log_a : R^+ \to \mathbb{R}$.

Moral da história: Do ponto de vista de teoria de grupos, $(R^+, \cdot, 1)$ e $(\mathbb{R}, +, 0)$ são indiscerníveis, justamente por serem grupos isomorfos entre si. Em particular, 0 e 1 são algebricamente indiscerníveis do ponto de vista do isomorfismo do último teorema.

Uma vez que sabemos que, entre os reais, $0 \neq 1$, o último teorema mostra que os números reais são muito mais do que simples estruturas algébricas de grupo. 'Filtrar' logaritmos sob a ótica de operações algébricas como adição e multiplicação entre reais, pode nos tornar 'daltônicos' a respeito dos reais.

Seção 69

Resumo da ópera

Sumário
Índice
Rede

Esta sexta parte pode ser resumida como se segue.

- O objetivo do cálculo diferencial e integral é o estudo e a aplicação de equações diferenciais, as quais são fórmulas $u = v$ onde há pelo menos uma ocorrência de um operador diferencial.

- O estudo e a aplicação de equações diferenciais depende do Teorema Fundamental do Cálculo.
- As funções seno e co-seno são definidas como soluções de uma equação diferencial. O que diferencia seno de co-seno são as condições de contorno impostas à equação diferencial usada para defini-las.
- As interpretações geométricas usuais para seno e co-seno são teoremas de cálculo diferencial e integral padrão.
- Fórmulas usuais de Trigonometria (o estudo de funções circulares) dependem de cálculo diferencial e integral estendido para os complexos.
- Logaritmo natural, por definição, é uma integral de Riemann.
- Logaritmos são definidos a partir de conceitos de cálculo diferencial e integral.

Seção 70

Notas históricas

Sumário
Índice
Rede

Historicamente, trigonometria é estudada pelo menos desde o período helenístico há mais de dois milênios, com o objetivo de aplicações em astronomia e engenharia. No entanto, do ponto de vista da matemática hodierna, os antigos conceitos trigonométricos não eram formulados de maneira racional. Assumir, por exemplo, que seno de um ângulo interno de um triângulo retângulo é, *por definição*, a razão entre a medida do cateto oposto ao ângulo e a medida da hipotenusa, não permite calcular o seno, digamos, de um radiano. O que torna operacional o cálculo de seno da medida de um ângulo é, hoje em dia, sua definição como solução de um problema de contorno. Isso mostra que qualquer noção de racionalidade depende de contextos históricos.

Plimpton 322: Tabela trigonométrica babilônica que antecede o período helenístico em pelo menos mil anos

Fonte: Norman Wildberger.

Com relação a logaritmos, eles foram concebidos em 1614 por John Napier, ou seja, antes do advento do cálculo diferencial e integral. Isso ajuda a ilustrar o fato de que o desenvolvimento histórico da matemática é um processo de difícil compreensão.

PARTE 7

Geometria euclidiana

Antes de iniciarmos a geometria euclidiana, é conveniente falarmos um pouco sobre o *Programa de Suppes*. Os estudos de geometria sintética e até mesmo álgebra linear na Parte 8 podem ser percebidos como exemplos pontuais do programa mencionado.

Seção 71
Predicados conjuntistas

Sumário
Índice
Rede

Lecionando filosofia na universidade Stanford, durante os anos 1950, Patrick Suppes produziu algumas notas de aula sobre o papel de teoria de conjuntos para os fundamentos da ciência. Em 1962 ele distribuiu entre interessados uma extensa monografia com maior detalhamento sobre aquelas notas, sob o título provisório *Set-Theoretical Structures in Science*. Em 2002, décadas de investigações sobre o tema foram reunidas no livro *Representation and Invariance of Scientific Structures* [54]. Em 2014 Suppes faleceu.

Nesta Seção discutimos brevemente sobre o famoso *Programa de Suppes*, o qual é resumido pelo autor em um *slogan*:

Axiomatizar uma teoria é definir um predicado conjuntista.

A ideia é simples. Considerando que

I: uma teoria de conjuntos qualquer permite fundamentar vastas porções da matemática e

II: essas vastas porções da matemática são empregadas para mapear fenômenos do mundo real,

então é possível usar a linguagem e a lógica de uma teoria de conjuntos para formular teorias referentes às ciências reais, como física e linguística, entre outras.

Naturalmente, tal estratégia garante economia de pensamento, no sentido de tirar proveito de tudo aquilo que teorias de conjuntos têm a oferecer.

Inspirados no Programa de Suppes, Newton da Costa e Rolando Chuaqui desenvolveram o conceito de *Predicado de Suppes*, o qual é um assunto sofisticado demais para os propósitos desta obra [11]. No entanto, podemos qualificar a proposta de Suppes no contexto de ZF da seguinte maneira:

Definição 7.1. *Um* predicado conjuntista $\mathfrak{P}$ *para uma* 'teoria' $\mathcal{T}$ *é a seguinte abreviação metalinguística:*

$$\mathfrak{P}(\mathcal{T}) \vdots \exists x_1 \exists x_2 \cdots \exists x_n \exists r_1 \exists r_2 \cdots \exists r_m (\mathcal{T} = \langle x_1, x_2 \cdots, x_n, r_1, r_2 \cdots, r_m \rangle \wedge \textit{ axiomas de } \mathcal{T}),$$

onde

- *cada x_i é um conjunto,*
- *cada r_j é uma relação na qual há pelo menos uma ocorrência de algum conjunto x_i em seu domínio ou co-domínio e, finalmente,*
- *os axiomas de $\mathcal{T}$ são fórmulas nas quais ocorrem pelo menos um dos termos $x_1, \cdots, x_n$ ou um dos termos $r_1, \cdots, r_m$.*

Obviamente, a definição acima não é suficientemente clara, até porque Suppes jamais se preocupou em formular rigorosamente suas ideias. A estratégia dele, para desenvolver e veicular o *slogan* acima, era sustentada por exemplos pontuais. Alguns desses exemplos são discutidos aqui, como teoria de grupos, aritmética, espaços vetoriais, corpos, geometria euclidiana, espaços de probabilidades e mecânica clássica não relativística de partículas.

Os termos $x_1, \cdots, x_n$ e $r_1, \cdots, r_m$ são chamados de *conceitos primitivos* da 'teoria' $\mathcal{T}$.

A justificativa para o uso de aspas em '*teoria*' é a seguinte: de acordo com a visão acima, toda 'teoria' $\mathcal{T}$ é um conjunto, em particular, uma $(n+m)$-upla ordenada. No entanto, Suppes propõe seu programa de axiomatização para as ciências formais (matemática, lógica e ciência da computação), bem como as ciências reais (física, química, economia, psicologia, entre outras). Por um lado, é usual se referir a ZF como uma teoria, apesar de ZF certamente não ser um conjunto. Por outro, uma teoria física como a mecânica clássica, dificilmente é aceitável como um conjunto. Afinal, fazem parte da mecânica clássica certos conceitos que escapam do domínio de uma teoria de conjuntos como ZF. Exemplos bem conhecidos são experimentos e significados intuitivos de princípios físicos. Logo, a palavra 'teoria' assume múltiplas conotações na literatura especializada, a ponto de não haver um consenso sobre o que é de fato uma teoria.

No contexto aqui discutido, assumimos como teoria um sistema formal na acepção dada por Elliott Mendelson, em seu livro [38]. A teoria de Zermelo-Fraenkel é um caso particular de teoria, no sentido de que qualificamos linguagem formal e lógica subjacente. Logo, a proposta de Suppes não é cabível para qualificar teorias. Por conta disso, preferimos nos referir a $\mathcal{T}$ como uma 'teoria' (entre aspas).

Com relação aos símbolos metalinguísticos $\langle$ e $\rangle$, estes são chamados de 'abre parênteses' e 'fecha parênteses', respectivamente. Tais símbolos cumprem o mesmo papel de (e). Mas é uma prática comum o emprego de $\langle$ e $\rangle$ no contexto de predicados conjuntistas.

Como primeira ilustração de predicado conjuntista, consideremos a 'teoria' de grupos. Esta foi brevemente apresentada na Seção 68. Usando a Definição 7.1, podemos conceituar o predicado conjuntista $\mathfrak{G}$ 'ser um grupo' da seguinte maneira:

Definição 7.2.

$$\mathfrak{G}(\mathcal{G}) \vdots\ \exists g\, \exists \star\, (\mathcal{G} = \langle g, \star\rangle \wedge g \neq \varnothing \wedge \star \subset (g \times g) \times g \wedge$$
$$\forall a \forall b((a \in g \wedge b \in g) \Rightarrow \exists! c((a,b),c) \in \star) \wedge$$
$$\exists e(e \in g \wedge \forall a(a \in g \Rightarrow (((a,e),a) \in \star \wedge ((e,a),a) \in \star))) \wedge$$
$$\forall a \forall b \forall c(\star(\star(a,b),c) = \star(a, \star(b,c))) \wedge$$
$$\forall a \exists b(\star(a,b) = e \wedge \star(b,a) = e)).$$

Comparando Definições 7.1 e 7.2, temos o que se segue.

- O predicado $\mathfrak{P}$ da Definição 7.1 é interpretado como o predicado $\mathfrak{G}$ da Definição 7.2.
- A 'teoria' $\mathcal{T}$ da Definição 7.1 é interpretada como a 'teoria' $\mathcal{G}$ da Definição 7.2.
- Conjunto x_1 da Definição 7.1 é interpretado como o conjunto g da Definição 7.2; logo, o valor de n da Definição 7.1 é 1.
- Relação r_1 da Definição 7.1 é interpretada como a função $\star$ da Definição 7.2; logo, o valor de m da Definição 7.1 é 1.
- Os axiomas de $\mathcal{T}$ mencionados na Definição 7.1, para a interpretação $\mathcal{G}$, são as fórmulas

$$g \neq \varnothing,$$

$$\star \subset (g \times g) \times g,$$

$$\forall a \forall b((a \in g \wedge b \in g) \Rightarrow \exists! c((a,b),c) \in \star),$$

$$\exists e(e \in g \wedge \forall a(a \in g \Rightarrow (((a,e),a) \in \star \wedge ((e,a),a) \in \star))),$$

$$\forall a \forall b \forall c(\star(\star(a,b),c) = \star(a,\star(b,c)))$$

e

$$\forall a \exists b(\star(a,b) = e \wedge \star(b,a) = e)$$

da Definição 7.2.

$\mathfrak{G}(\mathcal{G})$ se lê '$\mathcal{G}$ é um grupo'. Logo, $\mathcal{G}$ é um grupo sss for um par ordenado $\langle g, \star \rangle$ em conjunção com a conjunção de seis fórmulas. Essas seis fórmulas são os axiomas de grupo, os quais são discutidos nos próximos parágrafos. Por abuso de linguagem, chamamos o conjunto g de grupo.

O axioma

$$g \neq \varnothing$$

afirma que todo grupo é um conjunto não vazio. Isso corresponde ao postulado G1 da Seção 68.

O axioma

$$\star \subset (g \times g) \times g$$

afirma que $\star$ é uma relação com domínio $g \times g$ e co-domínio g, exatamente como se exige na Definição 7.1.

O axioma

$$\forall a \forall b((a \in g \wedge b \in g) \Rightarrow \exists! c((a,b),c) \in \star),$$

em conjunção com a fórmula do parágrafo acima, estabelece que $\star$ é uma função com domínio $g \times g$ e co-domínio g. Isso corresponde ao postulado G2 da Seção 68.

O axioma

$$\exists e(e \in g \wedge \forall a(a \in g \Rightarrow (((a, e), a) \in \star \wedge ((e, a), a) \in \star)))$$

corresponde aos postulados G3 e G4 da Seção 68.

Observar que, na formulação anterior de grupo, sugerimos o termo privilegiado e como um dos conceitos primitivos de grupo. No entanto, Definição 7.2 deixa claro que essa manobra não é necessária. Basta um axioma que garanta a existência de e, como foi feito agora. Além disso, observar que, na definição de predicado conjuntista, exigem-se como conceitos primitivos conjuntos e relações entre esses conjuntos. No entanto, na definição de grupo não há qualquer relação na qual o termo e seja domínio ou co-domínio. Neste sentido, é uma questão de coerência com a Definição 7.1 que o termo e não seja listado entre os conceitos primitivos de grupo.

É claro que a prática matemática nem sempre funciona assim. Muitos autores listam o termo e como um dos conceitos primitivos de grupo, mesmo sabendo que não há necessidade disso. Essa prática ocorre simplesmente porque fica mais fácil escrever o que é um grupo quando se assume e como um conceito primitivo.

A fórmula

$$\forall a \forall b \forall c(\star(\star(a, b), c) = \star(a, \star(b, c)))$$

afirma que $\star$ é associativa. Isso corresponde ao postulado G5 da Seção 68.

Finalmente, a fórmula

$$\forall a \exists b(\star(a, b) = e \wedge \star(b, a) = e))$$

corresponde ao postulado G6 da Seção 68.

Resumidamente,

um grupo é um conjunto munido de uma operação binária fechada, associativa, com elemento neutro e elemento simétrico.

O texto em destaque acima é o discurso usual, o qual emprega rigor, mas não formalismo, como discutido na Seção 8. Declarar informalmente o que é um grupo não é erro matemático. Mas é incompetência matemática não saber traduzir formalmente a afirmação 'grupo é um conjunto munido de uma operação binária fechada, associativa, com elemento neutro e elemento simétrico'. Sim, a última afirmação foi um juízo de valor. Logo, o leitor pode contestar o que foi dito sem prejuízo ao restante da leitura.

Para efeitos práticos, comumente predicados conjuntistas, para uma teoria como a de grupos, são escritos da forma como se apresenta na Seção 68.

Quando alguém afirma que um grupo é um conjunto munido de uma operação binária fechada, associativa, com elemento neutro e elemento simétrico, está tacitamente usando um predicado conjuntista para definir o que é um grupo, mesmo que não saiba disso. É uma situação análoga àquela de *Monsieur* Jourdain, personagem principal da peça *Le Bourgeois Gentilhomme*, de Molière. *Monsieur* Jourdain fica encantado ao descobrir que passou a vida toda falando em prosa, sem saber disso.

Definição 7.3. *Seja $\mathfrak{P}$ um predicado conjuntista, onde*

$$\mathfrak{P}(\mathcal{T}) \vdots \exists x_1 \exists x_2 \cdots \exists x_n \exists r_1 \exists r_2 \cdots \exists r_m (\mathcal{T} = \langle x_1, x_2 \cdots, x_n, r_1, r_2 \cdots, r_m \rangle \wedge \text{ axiomas de } \mathcal{T}),$$

nos moldes da Definição 7.1. Uma interpretação *de $\mathfrak{P}$ é qualquer atribuição de valores para os termos $x_1, x_2 \cdots, x_n, r_1, r_2 \cdots, r_m$. Um* modelo *de $\mathfrak{P}$ é uma interpretação de $\mathfrak{P}$ na qual os axiomas de $\mathcal{T}$ são teoremas.*

Exemplo 7.1. *No* Exemplo 6.25*, os três itens* i, ii *e* iii *são interpretações de grupo. Mas apenas itens* i *e* ii *são modelos de grupo.*

Mantendo este espírito de rigor no lugar de formalimo, finalmente podemos iniciar nossos estudos sobre geometria na próxima Seção.

Neste livro utilizamos predicados conjuntistas para definir números naturais, inteiros, racionais, reais e complexos nas Seções anteriores, ainda que o leitor possa perceber isso somente agora (lembrar de *Monsieur* Jourdain!). Mas usamos a mesma técnica para formular *plano euclidiano* na Seção 76, *espaços vetoriais reais* na Seção 80, *corpos* e *espaços vetoriais* quaisquer na Seção 96, espaços métricos na Seção 86, espaços de probabilidades na Seção 102 e até mesmo mecânica clássica não relativística de partículas na Seção 110.

Seção 72

Plano de incidência

Sumário
Índice
Rede

Em 1899 David Hilbert publicou um texto revolucionário sobre geometria. Trata-se do histórico *Grundlagen der Geometrie*. Uma tradução para o inglês pode ser encontrada em [23]. Na tradução portuguesa [24] há um *blurb* de Maria do Pilar Ribeiro e José da Silva Paulo onde se lê o seguinte:

> *Não é este um livro de texto de geometria elementar, mas esta tradução é dedicada aos nossos professores da matéria e aos estudantes de matemática das nossas universidades. Os tradutores têm a esperança de que um cuidadoso estudo dos vários problemas deste livro contribuirá indirectamente para implantar neles a ideia de que, em geral, os males do ensino da geometria nas nossas escolas só superficialmente residem em deficiências de ordem pedagógica, mas antes se encontram na falta de contacto com os problemas vivos, actuais, da matéria que se ensina e do indispensável treino para a investigação desses problemas.*

O impacto da obra de Hilbert resultou, entre muitas outras coisas, no livro *Fundamentos da Geometria*, de Benedito Castrucci [10], o qual é uma leitura altamente recomendável, apesar do texto contar com muitos erros de digitação.

O foco das obras citadas é geometria euclidiana, pelo menos numa acepção mais moderna do que a obra *Elementos*, de Euclides de Alexandria.

Geometria euclidiana é um ramo da *geometria sintética*, a qual trata também de *geometria absoluta*, *geometria não-euclidiana*, *geometria afim*, *geometria não-Desarguesiana*, *geometria projetiva*, *geometria não-Paschiana*, entre outras. Geometria euclidiana deve servir de fundamentação para a *geometria analítica plana*, a qual é discutida na Seção 90.

Todas as Seções desta Parte, com exceção da primeira, são uma adaptação do livro de Castrucci [10]. Mas há diferenças entre nossa versão e aquela de Castrucci, no que se refere à formulação de certos postulados. Em contrapartida, no livro citado a abordagem é muito mais detalhada, apesar do autor não qualificar qual teoria de conjuntos é empregada. Como já sugerido, aqui geometria euclidiana é tratada como um predicado conjuntista (no contexto de ZF) cujos axiomas são divididos em cinco grupos:

I: Incidência;

II: Ordem;

III: Congruência;

IV: Continuidade;

V: Paralelismo.

Introduzimos nesta Seção o primeiro grupo de postulados.

Definição 7.4. *Um* plano de incidência *é um par ordenado*

$$\mathfrak{p} = \langle \pi, \rho \rangle$$

tal que as seguintes fórmulas são teoremas.

GE1: $\forall a \forall b((a \in \pi \wedge b \in \pi \wedge a \neq b) \Rightarrow \exists r(r \in \rho \wedge a \in r \wedge b \in r))$.

GE2: $\forall a \forall b \forall r \forall s((a \in \pi \wedge b \in \pi \wedge r \in \rho \wedge s \in \rho \wedge a \neq b \wedge a \in r \cap s \wedge b \in r \cap s) \Rightarrow r = s)$.

GE3: $\forall r(r \in \rho \Rightarrow \exists a \exists b(a \in \pi \wedge b \in \pi \wedge a \neq b \wedge a \in r \wedge b \in r))$.

GE4: $\exists a \exists b \exists c(a \in \pi \wedge b \in \pi \wedge c \in \pi \wedge \forall r(r \in \rho \Rightarrow a \notin r \vee b \notin r \vee c \notin r))$.

Em um plano de incidência $\mathfrak{p} = \langle \pi, \rho \rangle$ chamamos os elementos de π de *pontos* e os elementos de ρ de *retas*. Logo, π é o conjunto de pontos do plano de incidência e ρ é o conjunto de retas do mesmo plano de incidência. Uma vez que elementos de um conjunto não vazio são conjuntos, então pontos e retas são conjuntos.

Por abuso de linguagem é usual se referir a π como *plano*.

Definição 7.5. *Seja $\langle \pi, \rho \rangle$ um plano de incidência. Se a é um ponto (i.e., $a \in \pi$), r é uma reta (i.e., $r \in \rho$) e $a \in r$, dizemos que 'o ponto a* incide *sobre a reta r' ou, equivalentemente, 'a reta r* passa *pelo ponto a'.*

Pontos incidentes sobre uma mesma reta em um plano de incidência são chamados de colineares*.*

Axioma GE1 diz que, para quaisquer dois pontos distintos a e b, existe uma reta que passa por ambos. Ou seja, dois pontos distintos são sempre colineares.

Axioma GE2 afirma que, se dois pontos distintos a e b incidem sobre reta r e reta s, então $r = s$. Em outras palavras, pontos distintos não são apenas colineares, mas também definem uma única reta que passa por ambos.

Postulado GE3 diz que toda reta r admite pelo menos dois pontos distintos incidentes sobre ela.

Finalmente, GE4 garante que existem pontos a, b e c tais que nenhuma reta r incide sobre os três. Ou seja, em qualquer plano de incidência devem existir pelo menos três pontos não colineares.

Exemplo 7.2. *Sejam*

$$\pi = \{1, 2, 3\}$$

e

$$\rho = \{\{1, 2\}, \{1, 3\}, \{2, 3\}\}.$$

Neste caso, $\langle \pi, \rho \rangle$ é um plano de incidência. Com efeito, os quatro postulados são teoremas para esta interpretação. Por exemplo, existem apenas três possíveis escolhas de pares de pontos distintos:

I: *Para os pontos 1 e 2 existe a reta $\{1, 2\}$ incidente sobre ambos;*

II : *Para os pontos 1 e 3 existe a reta $\{1, 3\}$ incidente sobre ambos;*

III: *Para os pontos 2 e 3 existe a reta $\{2, 3\}$ incidente sobre ambos.*

Logo, axioma GE1 *é teorema.*

✍ *Os demais postulados podem ser verificados pelo leitor.*

✍ No Exemplo acima foi ilustrado um plano de incidência

$$\langle\{1,2,3\},\{\{1,2\},\{1,3\},\{2,3\}\}\rangle$$

com apenas três pontos e três retas. Recomendamos que o leitor prove que o número mínimo de pontos em um plano de incidência é 3. Além disso, é um exercício interessante a concepção de um plano de incidência com quatro pontos.

O leitor deve ter observado que a formulação dos postulados de incidência é desnecessariamente complicada. Afinal, os axiomas propõem ideias simples mas através de fórmulas muito longas. Podemos evitar essa inconveniência utilizando quantificadores relativizados, a exemplo do que foi feito na Seção 35.

Definição 7.6. *Sejam $\mathcal{P}$ uma fórmula e ϱ um conjunto. Logo:*

$$\forall_{\varrho} a(\mathcal{P}) \vdots \forall a(a \in \varrho \Rightarrow \mathcal{P}),$$

$$\exists_{\varrho} a(\mathcal{P}) \vdots \exists a(a \in \varrho \wedge \mathcal{P}).$$

Importante perceber que os quantificadores relativizados usados na Seção 35 são casos particulares da Definição acima. Com efeito, naquela Seção o conjunto ϱ é simplesmente o conjunto dos racionais estritamente positivos.

Graças às abreviações metalinguísticas acima, podemos reescrever os quatro postulados de incidência de maneira muito mais fácil de ler, como se segue.

ge1: $\forall_{\pi} a \forall_{\pi} b(a \neq b \Rightarrow \exists_{\rho} r(a \in r \wedge b \in r))$.

ge2: $\forall_{\pi} a \forall_{\pi} b \forall_{\rho} r \forall_{\rho} s((a \neq b \wedge a \in r \cap s \wedge b \in r \cap s) \Rightarrow r = s)$.

ge3: $\forall_{\rho} r \exists_{\pi} a \exists_{\pi} b(a \neq b \wedge a \in r \wedge b \in r)$.

ge4: $\exists_{\pi} a \exists_{\pi} b \exists_{\pi} c(\forall_{\rho} r(a \notin r \vee b \notin r \vee c \notin r))$.

Em outras palavras, relativizamos os quantificadores universal e existencial aos conjuntos π e ρ de pontos e retas.

Os demais postulados de geometria euclidiana a serem introduzidos nas próximas Seções são todos escritos empregando quantificadores relativizados, de acordo com a Definição 7.6

Seção 73

Axiomas de ordem

Sumário
Índice
Rede

Qualificamos aqui o que é um ponto entre dois pontos.

Definição 7.7. *Seja* $\mathfrak{p} = \langle \pi, \rho \rangle$ *um plano de incidência. Seja ainda*

$$\underline{\quad} : \pi \times \pi \to \pi$$

uma relação (ver Definição 3.11) tal que

$$((a,c),b) \in \underline{\quad}$$

é denotado abreviadamente por

$$\underline{abc}$$

e lido como 'o ponto b está entre os pontos a e c'.

Dizemos que

$$\mathfrak{o} = \langle \mathfrak{p}, \underline{\quad} \rangle$$

é um plano quase-ordenado *sss as seguintes fórmulas são teoremas.*

GE5: $\forall_\pi a \forall_\pi b \forall_\pi c(\underline{abc} \Rightarrow \exists_\rho r(a \in r \wedge b \in r \wedge c \in r))$;

GE6: $\forall_\pi a \forall_\pi b \forall_\pi c(\underline{abc} \Rightarrow (a \neq b \wedge a \neq c \wedge b \neq c))$;

GE7: $\forall_\pi a \forall_\pi b \forall_\pi c(\underline{abc} \Rightarrow \underline{cba})$;

GE8: $\forall_\pi a \forall_\pi b(a \neq b \Rightarrow \exists c(\underline{abc}))$;

GE9: $\forall_\pi a \forall_\pi b \forall_\pi c(\underline{abc} \Rightarrow (\neg\underline{acb} \wedge \neg\underline{bac}))$.

A sequência de símbolos $\underline{abc}$ é uma abreviação metalinguística para a fórmula

$$((a,c),b) \in \underline{\quad},$$

onde $\underline{\quad}$ é uma relação com domínio $\pi \times \pi$ e codomínio π, em um plano de incidência $\mathfrak{p} = \langle \pi, \rho \rangle$.

Notar que planos quase-ordenados são definidos a partir de planos de incidência.

Axioma GE5 diz que, se um dado ponto b está entre os pontos a e c, então existe reta que passa por a, b e c.

Postulado ge6 afirma que, se um dado ponto b está entre os pontos a e c, então os três pontos envolvidos são distintos entre si, tomados dois a dois.

Postulado ge7 estabelece que, se um dado ponto b está entre os pontos a e c, então este mesmo ponto b está entre c e a.

Axioma ge8 afirma que, dados dois pontos a e b distintos entre si, então existe ponto c tal que b está entre a e c.

Finalmente, axioma ge9 diz que, se um dado ponto b está entre os pontos a e c, então c não está entre a e b e, além disso, a não está entre b e c.

Precisamos agora conhecer um pouco melhor os inteiros, para que possamos exibir um modelo de plano quase-ordenado.

Definição 7.8. *Um inteiro x é* múltiplo *de um inteiro y sss existe inteiro z tal que $x = yz$.*

Exemplo 7.3. *Os inteiros* 5 *e* -10 *são múltiplos de* -5. *Com efeito,* $5 = (-1)(-5)$ *e* $-10 = 2(-5)$.

✍ Notar que 0 é múltiplo de qualquer inteiro. Além disso, nenhum inteiro diferente de 0 é múltiplo de 0. Outro teorema útil é o seguinte: se $x = yz$, onde x, y e z são inteiros, então x é múltiplo de ambos y e z.

Definição 7.9. *Um* divisor *de um inteiro x é qualquer inteiro y tal que x é múltiplo de y.*

Exemplo 7.4. i: *Os divisores de qualquer primo x (ver Definição 4.2) são apenas x e 1;*

ii: *Os divisores de* 12 *são* 1, 2, 3, 4, 6 *e* 12.

Definição 7.10. *Inteiros estritamente positivos m e n são* primos entre si *sss o único divisor em comum entre m e n é 1.*

Exemplo 7.5. i: 5 *e* 4 *são primos entre si. Com efeito, o único divisor em comum entre eles é* 1*;*

ii: 1 *e* 1 *são primos entre si;*

iii: *2 e 2 não são primos entre si, uma vez que compartilham o mesmo divisor* $2 \neq 1$*;*

iv: *6 e 4 não são primos entre si.*

Nos Exemplos 7.6, 7.7, 7.8 e 7.9 a seguir exibimos e examinamos um modelo de plano quase-ordenado.

Exemplo 7.6. *Seja* $\pi = \mathbb{Z}^2$*, ou seja, cada ponto pertencente a* π *é um par ordenado* (x, y) *de inteiros. Sejam ainda*

$$r^{m,n}_{p,q} = \Big\{(x,y) \in \mathbb{Z}^2 \mid (x = p \wedge y = q) \vee \Big((x \neq p \vee y \neq q) \Rightarrow \Big(x - p \text{ é múltiplo de } m \wedge \frac{x-p}{y-q} = \frac{m}{n}\Big)\Big)\Big\},$$

onde m*,* n*,* p *e* q *são inteiros,* $n \neq 0$ *e* $|m|$ *e* $|n|$ *são primos entre si; e*

$$r^{m,0}_{p,q} = \{(x,y) \in \mathbb{Z}^2 \mid y = q\}.$$

✍ *Se definirmos* ρ *como o conjunto de todos os* $r^{m,n}_{p,q}$*, onde* p*,* q*,* m *e* n *são inteiros, então* $\langle \pi, \rho \rangle$ *é um plano de incidência. Cabe ao leitor provar.*

No Exemplo acima cada reta pertencente a ρ é um conjunto não vazio de pares ordenados de inteiros, onde cada um desses pares ordenados é um ponto do plano de incidência $\langle \pi, \rho \rangle$.

✍ A condição '$|m|$ e $|n|$ são primos entre si' é desnecessária, bastando que n seja diferente de 0. Sugerimos que o leitor examine postulado ge2 no que se refere a tal exigência, para fins de avaliação, pelo menos de um ponto de vista intuitivo.

Exemplo 7.7. *No plano de incidência* $\langle \pi, \rho \rangle$ *do* Exemplo 7.6 *a reta* $r^{1,1}_{0,0}$ *é o conjunto*

$$\{(0,0), (1,1), (-1,-1), (2,2), (-2,-2), \cdots\}.$$

Observar que $r^{1,1}_{0,0} = r^{-1,-1}_{0,0} = r^{1,1}_{1,1} = r^{-1,-1}_{1,1} = r^{n,n}_{p,p}$ *para qualquer* n *inteiro diferente de* 0*.*

Outro exemplo de reta é o conjunto

$$r^{3,5}_{1,2} = \{(1,2), (4,7), (-2,-3), (7,12), (-5,-8) \cdots\},$$

o qual é igual a $r^{-3,-5}_{1,2}$*, que é igual a* $r^{3,5}_{4,7}$ *e assim por diante.*

Notar que

$$r^{1,1}_{0,0} \cap r^{3,5}_{1,2} = \varnothing.$$

Finalmente, para ilustrar uma reta para o caso em que $n = 0$, *temos*

$$r^{3,0}_{1,2} = \{(0,2),(1,2),(-1,2),(2,2),(-2,2),\cdots\}.$$

Neste caso,

$$r^{3,0}_{1,2} = r^{m,0}_{p,2}$$

para quaisquer p *e* m *inteiros.*

Observar que

$$r^{3,0}_{1,2} \cap r^{1,1}_{0,0} = \{(2,2)\}$$

e

$$r^{3,0}_{1,2} \cap r^{3,5}_{1,2} = \{(1,2)\}.$$

Exemplo 7.8. *No plano de incidência* $\langle \pi, \rho \rangle$ *do* Exemplo 7.6, *consideremos a seguinte relação* $\_\_ : \pi \times \pi \to \pi$ *dada por*

$$\underline{(x,y)(x',y')(x'',y'')}$$

sss

- I: (x,y), (x',y') *e* (x'',y'') *pertencem à mesma reta e*
- II: $(x < x' < x'' \vee x'' < x' < x) \vee ((x = x' = x'') \wedge (y < y' < y'' \vee y'' < y' < y))$.

Naturalmente,

$$\underline{(x,y)(x',y')(x'',y'')} \ \text{ sss } \ (((x,y),(x'',y'')),(x',y')) \in \_\_.$$

Neste caso, $\langle \langle \pi, \rho \rangle, \_\_ \rangle$ *é um plano quase-ordenado, onde lemos*

$$\underline{(x,y)(x',y')(x'',y'')}$$

como (x',y') *está entre* (x,y) *e* (x'',y'').

Exemplo 7.9. *Seguindo* Exemplos 7.6, 7.7 *e* 7.8, *o ponto* $(1,1)$ *está entre* $(2,2)$ *e* $(-1,-1)$. *Com efeito, os três pontos envolvidos pertencem à reta* $r^{1,1}_{0,0}$ *e, além disso,* $-1 < 1 < 2$.

Analogamente, $(1,2)$ *está entre os pontos* $(4,7)$ *e* $(-2,-3)$. *Neste caso, a reta que passa pelos três é* $r^{3,5}_{1,2}$. *O mesmo ponto* $(1,2)$ *está também entre os pontos* $(0,2)$ *e* $(2,2)$, *apesar da reta que passa pelos três ser* $r^{3,0}_{1,2}$, *a qual é diferente de* $r^{3,5}_{1,2}$.

Em compensação, o ponto $(1, 2)$ *não está entre*

$$(4, 7) \text{ e } (0, 2).$$

Com efeito, os pontos $(1, 2)$ *e* $(4, 7)$ *incidem sobre a reta* $r^{3,5}_{1,2}$*, enquanto os pontos* $(1, 2)$ *e* $(0, 2)$ *incidem sobre a reta* $r^{3,0}_{1,2}$*. Uma vez que* $r^{3,0}_{1,2} \cap r^{3,5}_{1,2} = \{(1, 2)\}$*, de acordo com* EXEMPLO 7.7*, então estamos falando de retas distintas. Logo,* $(1, 2)$*,* $(4, 7)$ *e* $(0, 2)$ *são pontos não colineares neste modelo.*

✍ A prova de que

$$\langle\langle \pi, \rho\rangle, \_\_\rangle,$$

do EXEMPLO 7.8, garante as fórmulas GE5∼GE9 como teoremas, é quase imediata. A demonstração de que GE1∼GE4 são teoremas nesta interpretação é um pouco mais ardilosa.

DEFINIÇÃO 7.11. *Sejam* a *e* b *pontos de um plano quase-ordenado* $\langle\langle \pi, \rho\rangle, \_\_\rangle$*. Um* segmento fechado de reta *(ou* segmento*)* $[a, b]$ *é o conjunto*

$$[a, b] = \{x \in \pi \mid x = a \vee x = b \vee \underline{axb}\}.$$

Além disso, um segmento aberto $]a, b[$ *é o conjunto*

$$]a, b[= \{x \in \pi \mid \underline{axb}\},$$

onde $a \neq b$*.*

EXEMPLO 7.10. *Seguindo os* EXEMPLOS 7.6, 7.7 *e* 7.8*, o conjunto*

$$x = \{(-2, -3), (1, 2), (4, 7), (7, 12)\}$$

é o segmento de reta

$$[(-2, -3), (7, 12)].$$

Com efeito, os pontos $(1, 2)$ *e* $(4, 7)$ *são os únicos entre* $(-2, -3)$ *e* $(7, 12)$*.*

Além disso, x *admite nove subconjuntos próprios que são segmentos. Dois deles são*

$$\{(-2, -3), (1, 2), (4, 7)\} \text{ e } \{(4, 7)\}.$$

Notar também que $\{(4, 7)\}$ *é um segmento fechado e um segmento aberto. Com efeito,* $\{(4, 7)\} = [(4, 7)] =](1, 2), (7, 12)[$*.*

No Exemplo acima são exibidos três segmentos de reta:

- um com quatro pontos,
- um com três pontos e
- outro com um único ponto.

Naturalmente, no espaço quase-ordenado do Exemplo 7.8 todo segmento de reta admite no mínimo um ponto. Além disso, todo segmento de reta no mesmo espaço é um conjunto finito.

Definição 7.12. *Se os pontos a, b e c são colineares (ver Definição 7.5) em um plano quase-ordenado, denotamos isso por $\overline{abc}$.*

Observar que $\overline{abc}$ é uma fórmula. Obviamente,

$$\underline{abc} \Rightarrow \overline{abc}$$

é teorema. No entanto, a recíproca não é. ✍ Com efeito, basta exibir pontos a, b e c colineares tais que o ponto a esteja entre b e c.

Notar também que, apesar da Definição 7.5 se referir a colinearidade de pontos em um plano de incidência, todo plano quase-ordenado é um caso particular de plano de incidência. Logo, é consistente usar a Definição 7.5 para tratar de colinearidade de pontos em um plano quase-ordenado.

Definição 7.13. *Um plano quase-ordenado $\langle\langle \pi, \rho\rangle, \_\_\rangle$ é um* plano ordenado *sss a fórmula abaixo é teorema.*

ge10: $\forall_\pi a \forall_\pi b \forall_\pi c \forall_\rho r((\neg\overline{abc} \wedge a \notin r \wedge b \notin r \wedge c \notin r \wedge \exists d(r \cap [a,b] = \{d\})) \Rightarrow (\exists e(r \cap [b,c] = \{e\}) \vee \exists f(r \cap [a,c] = \{f\})))$.

Postulado ge10 é o famoso *Axioma de Pasch*, em referência a Moritz Pasch (matemático alemão que exerceu forte influência na obra de David Hilbert). De um ponto de vista meramente intuitivo, os pontos a, b e c não colineares, assumidos no axioma, devem definir vértices de um triângulo cujos lados são os segmentos $[a,b]$, $[b,c]$ e $[c,a]$. Neste contexto, ge10 diz o seguinte:

Se uma reta, em um plano ordenado, intersecta um dos lados do triângulo dado por a, b e c, sem passar por qualquer vértice, então a mesma reta intersecta um dos outros lados.

Exemplo 7.11. *O plano quase-ordenado*

$$\langle\langle\pi,\rho\rangle,\_\_\rangle,$$

do Exemplo 7.8, *não é um plano ordenado. Para provar isso, basta exibir um contraexemplo para a fórmula* ge10.

Seja r *a reta* $r_{0,0}^{1,1}$. *Sejam* a, b *e* c *os pontos não colineares*

$$(1,2),\quad (3,2)\quad e\quad (4,7),$$

respectivamente. Notar que nenhum deles incide sobre $r_{0,0}^{1,1}$.

Os pontos $(1,2)$ *e* $(3,2)$ *definem a reta*

$$r_{1,2}^{3,0},$$

a qual intersecta $r_{0,0}^{1,1}$ *no ponto* $(2,2)$, *conforme* Exemplo 7.7. *Logo, todas as condições que antecedem a condicional de* ge10 *são satisfeitas, onde o ponto* d *é justamente* $(2,2)$.

No entanto, a reta definida por a *e* c *é*

$$r_{1,2}^{3,5},$$

cuja interseção com $r_{0,0}^{1,1}$ *é o conjunto vazio. Além disso, a reta definida por* b *e* c *é* $r_{3,2}^{1,5}$, *cuja interseção com* $r_{0,0}^{1,1}$ *também é o conjunto vazio.*

Logo, a reta $r_{0,0}^{1,1}$ *intersecta o lado* $[a,b]$ *do triângulo com vértices* a, b *e* c, *mas não intersecta o lado* $[b,c]$ *e nem o lado* $[a,c]$, *garantindo dessa maneira que* ge10 *não é teorema no plano quase-ordenado em questão. Temos, assim, um plano quase-ordenado que não é um plano ordenado.*

Para exibirmos um plano ordenado, pedimos ao leitor um pouco de paciência. Chegaremos lá.

Definição 7.14. *Sejam* r *uma reta de um plano ordenado e* o, b *e* c *pontos incidentes sobre* r, *de modo que* $\underline{boc}$, *ou seja,* o *está entre* b *e* c. *Então os conjuntos*

$$r_o^b = \{x \in r - \{o\} \mid \underline{box}\}$$

e

$$r_o^c = \{x \in r - \{o\} \mid \underline{cox}\}$$

são semirretas com origem o *ou, simplesmente,* semirretas, *se não houver risco de confusão.*

Em [10] é demonstrado que, para toda reta r de um plano ordenado, é possível definir uma relação de equivalência sobre $r - \{o\}$ que particiona este conjunto em exatamente dois subconjuntos próprios. Tal teorema, omitido aqui, justifica a última definição, no sentido de que os dois subconjuntos próprios de $r - \{o\}$ são exatamente as semirretas acima.

Lembrar que, de acordo com GE7,

$$\underline{box} \Leftrightarrow \underline{xob}$$

e

$$\underline{cox} \Leftrightarrow \underline{xoc}.$$

O que fica evidente a partir da Definição 7.14 são os seguintes teoremas.

I: $r_o^b \cap r_o^c = \varnothing$;

II: $r_o^b \cup r_o^c = r - \{o\}$;

III: $c \in r_o^b \wedge b \in r_o^c$;

IV: $b \notin r_o^b \wedge c \notin r_o^c$.

V: $o \notin r_o^b \wedge o \notin r_o^c$.

Recomendamos ao leitor que os demonstre.

Definição 7.15. *Seja a um ponto de um plano ordenado. Se r_a é uma semirreta, então*

$$r_a \cup \{a\}$$

é uma semirreta fechada.

Ou seja, semirretas fechadas são conjuntos cujos elementos são todos os pontos de uma semirreta, bem como a origem da semirreta.

Seção 74

Axiomas de congruência

Sumário
Índice
Rede

Etimologicamente falando, geometria era o estudo da 'medição da terra'. Entre civilizações egípcias e babilônicas de milhares de anos atrás, geometria era uma ciência física, cujos princípios eram determinados experimentalmente, a partir de modos de percepção humana sobre o mundo onde vivemos. Neste sentido, havia a preocupação com medições de comprimentos, áreas, volumes e ângulos, sem qualquer qualificação para tais conceitos (pelo menos nos moldes do que hoje se entende por qualificação).

Hoje em dia, porém, a abordagem sintética para a geometria dispensa quaisquer considerações sobre medidas, no sentido de atribuir a objetos geométricos números reais que digam quanto mede, por exemplo, um ângulo ou um segmento.

Essa mesma abordagem sintética foi posteriormente utilizada para a formulação de teorias físicas, onde a Gravitação Universal de Newton foi formulada por Hartry Field sem qualquer necessidade do emprego de números. Detalhes em [15].

Nesta Seção mostramos como é possível tratar de relações de congruência em geometria sem a invocação de medidas dadas por números reais. A ideia intuitiva é simples: medidas podem corresponder biunivocamente a elementos de uma partição. Por exemplo, um segmento de reta s tem medida m se m for o conjunto de todos os segmentos de reta com a mesma medida de s. Para evitar a óbvia circularidade da última frase, basta introduzirmos uma relação de equivalência que cumpra o papel de particionar o conjunto de todos os segmentos de reta, nos moldes dos Teoremas 3.10 e 3.11. Essa relação de equivalência se chama *congruência*.

Como destacado, a interpretação pretendida para congruência é a seguinte: dois segmentos são congruentes se compartilham a mesma medida.

No entanto, os axiomas GE11∼GE14 a seguir permitem provar que congruência é tão somente uma relação de equivalência, em situação análoga (mas não equivalente!) à equipotência entre conjuntos, conforme Definição 4.17 na Seção 33. Logo, não há a necessidade de qualquer referência a medidas dadas por números reais. Geometria sintética é essencialmente o estudo de geometria sem medidas (dadas por números reais) e sem cooordenadas.

Geometria sintética é geometria sem números.

Mais adiante mostramos como se relaciona a geometria sintética com a geometria analítica, a qual é uma geometria com números.

Definição 7.16. *Sejam* $\mathfrak{o} = \langle\langle \pi, \rho \rangle, \_\_ \rangle$ *um plano ordenado,*

$$s = \{x \in \wp(\pi) \mid \exists_\pi a \exists_\pi b (x = [a, b])\}$$

e

$$\downarrow_{\dot{a}} = \{r \in \wp(\pi) \mid r \text{ é uma semirreta fechada com origem } a\},$$

onde $a \in \pi$*, de acordo com Definições 7.14 e 7.15.*

Em outras palavras, s *é o conjunto de todos os segmentos de* $\mathfrak{o}$ *e* $\downarrow_{\dot{a}}$ *é o conjunto de todas as semirretas fechadas com origem* a*.*

Seja ainda $\approxeq$ *uma relação em* s*, i.e.,*

$$\approxeq \subseteq s \times s.$$

Dizemos que $\approxeq$ *é uma* relação de congruência *sss as seguintes fórmulas são teoremas.*

GE11: $\forall_\pi a \forall_\pi b \forall_\pi c \forall_{\downarrow_{\dot{c}}} r \exists_\pi d (d \in r \wedge [a, b] \approxeq [c, d])$.

GE12: $\forall_\pi a \forall_\pi b \forall_\pi c \forall_\pi a' \forall_\pi b' \forall_\pi c' ((\underline{abc} \wedge \underline{a'b'c'} \wedge [a, b] \approxeq [a', b'] \wedge [b, c] \approxeq [b', c']) \Rightarrow [a, c] \approxeq [a', c'])$.

GE13: $\forall_\pi a \forall_\pi b \forall_\pi c \forall_\pi d \forall_\pi e \forall_\pi f (([a, b] \approxeq [c, d] \wedge [e, f] \approxeq [c, d]) \Rightarrow [a, b] \approxeq [e, f])$.

GE14: $\forall_\pi a \forall_\pi b \forall_\pi c \forall_\pi a' \forall_\pi b' \forall_\pi c' ((\neg\overline{abc} \wedge \neg\overline{a'b'c'}) \Rightarrow \forall_\pi d \forall_\pi d' ((d \in]a, b[\wedge d' \in]a', b'[\wedge [a, b] \approxeq [a', b'] \wedge [b, c] \approxeq [b', c'] \wedge [a, c] \approxeq [a', c'] \wedge [a, d] \approxeq [a', d']) \Rightarrow [d, c] \approxeq [d', c']))$.

Postulado GE11 estabelece que, dados pontos a e b e uma semirreta fechada com origem c, então existe ponto d incidente sobre a semirreta fechada tal que os segmentos $[a, b]$ e $[c, d]$ são congruentes.

Axioma GE12 afirma que, se b está entre a e c e, além disso, b' está entre a' e c', então a congruência entre os segmentos $[a, b]$ e $[a', b']$, em conjunção com a congruência entre os segmentos $[b, c]$ e $[b', c']$, implica na congruência entre $[a, c]$ e $[a', c']$.

Fórmula GE13 sugere a transitividade da congruência. Obviamente é necessário provar que congruência é simétrica, para inferir essa transitividade. Ou seja, devemos garantir que

$$[e, f] \approxeq [c, d] \Leftrightarrow [c, d] \approxeq [e, f]$$

é teorema, para interpretarmos GE13 como transitividade. Mas isso é feito no Teorema 7.1, logo adiante.

Com relação ao postulado GE14, este trata de triângulos. Afinal, são assumidos pontos a, b e c não colineares, o que garante que eles são distintos dois a dois. Comentário análogo sobre os pontos a', b' e c'. Portanto, temos um triângulo com lados $[a, b]$, $[b, c]$ e $[a, c]$, bem como um triângulo de lados $[a', b']$, $[b', c']$ e $[a', c']$. Os vértices do primeiro são a, b e c, enquanto os vértices do segundo são a', b' e c'. É justamente GE14 que cria oportunidade para falarmos sobre ângulos congruentes, algo a ser discutido mais adiante.

No contexto do parágrafo acima, são também assumidos pontos d e d' incidentes sobre os lados $[a, b]$ e $[a', b']$, respectivamente, de modo que nenhum deles é qualquer vértice. Dadas todas essas condições (deste e do parágrafo anterior), GE14 diz o seguinte: a congruência entre $[a, b]$ e $[a', b']$, entre $[b, c]$ e $[b', c']$, entre $[a, c]$ e $[a', c']$ e entre $[a, d]$ e $[a', d']$, implica na congruência entre $[d, c]$ e $[d', c']$. Intuitivamente falando, os segmentos $[d, c]$ e $[d', c']$ 'atravessam o interior' de seus respectivos triângulos, definindo novos triângulos com vértices a, c e d e vértices b, c e d, e com lados respectivamente congruentes aos lados dos triângulos com vértices a', c' e d' e vértices b', c' e d'.

Teorema 7.1. *A relação de congruência $\approxeq$ da Definição 7.16 é de equivalência.*

Demonstração: De acordo com a Definição 3.14, devemos provar que $\approxeq$ é reflexiva, simétrica e transitiva. Por conta disso, dividimos essa demonstração em três partes.

Reflexividade: De acordo com GE11, se a e b definem um segmento $[a, b]$ e c é um ponto qualquer (origem de uma semirreta), então existe d tal que $[a, b] \approxeq [c, d]$. Mas GE13 garante que

$$([a, b] \approxeq [c, d] \wedge [a, b] \approxeq [c, d]) \Rightarrow [a, b] \approxeq [a, b],$$

onde substituímos $[e, f]$ por $[a, b]$.

Logo,

$$\forall_\pi a \forall_\pi b([a, b] \approxeq [a, b]).$$

Simetria: Foi provado no primeiro passo que $[a', b'] \approxeq [a', b']$. Supor que $[a, b] \approxeq [a', b']$. Logo, GE13 garante que

$$([a', b'] \approxeq [a', b'] \wedge [a, b] \approxeq [a', b']) \Rightarrow [a', b'] \approxeq [a, b].$$

Portanto, $[a,b] \approxeq [a',b'] \Rightarrow [a',b'] \approxeq [a,b]$. Naturalmente isso é equivalente à fórmula

$$[a,b] \approxeq [a',b'] \Leftrightarrow [a',b'] \approxeq [a,b].$$

TRANSITIVIDADE: GE13 afirma que

$$([a,b] \approxeq [c,d] \wedge [e,f] \approxeq [c,d]) \Rightarrow [a,b] \approxeq [e,f].$$

Mas foi provado acima que

$$[e,f] \approxeq [c,d] \Leftrightarrow [c,d] \approxeq [e,f].$$

Logo,

$$([a,b] \approxeq [c,d] \wedge [c,d] \approxeq [e,f]) \Rightarrow [a,b] \approxeq [e,f].$$

Isso encerra a prova.

Se s é o conjunto de todos os segmentos de reta em um plano ordenado (como apresentado na Definição 7.16), então a 'medida' de um segmento de reta $[a,b]$ pode ser dada simplesmente por

$$\{x \in s \mid x \approxeq [a,b]\}.$$

Portanto, Teoremas 3.10 e 3.11 garantem que $\approxeq$ particiona s em classes de equivalência, as quais são tão somente 'medidas', no sentido acima. Neste contexto, uma 'medida' não é qualquer número real.

Colocamos a palavra 'medida' entre aspas porque o conceito de *medida*, em *teoria da medida*, demanda o emprego de números reais, algo a ser discutido na Seção 103. Ou seja, usamos o termo 'medida' aqui apenas em um sentido meramente intuitivo.

DEFINIÇÃO 7.17. *Sejam r_o e s_o semirretas com a mesma origem o. Se não existe reta r tal que $r_o \subset r \wedge s_o \subset r$, dizemos que*

$$r_o \cup s_o$$

é o ângulo $\widehat{r_o s_o}$.

Também podemos denotar o ângulo $r_o \cup s_o$ por $\widehat{aob}$, onde a e b são pontos incidentes sobre r_o e s_o, respectivamente, e ambos diferentes de o.

✍ Ou seja, um ângulo é a união de semirretas não colineares que compartilham a mesma origem. A justificativa para podermos denotar o ângulo $\widehat{r_o s_o}$ por $\widehat{aob}$ reside no axioma GE2. Cabe ao leitor escrever os detalhes.

Definição 7.18. *Sejam r_o, s_o, m_u e n_u semirretas. Sejam ainda $\widehat{r_o s_o}$ e $\widehat{m_u n_u}$ ângulos. Dizemos que $\widehat{r_o s_o}$ e $\widehat{m_u n_u}$ são congruentes, e escrevemos isso como*

$$\widehat{r_o s_o} \approxeq \widehat{m_u n_u},$$

sss

$$\forall_\pi b \forall_\pi c \forall_\pi b' \forall_\pi c' ((b \in r_o \wedge c \in s_o \wedge b' \in m_u \wedge c' \in n_u \wedge [o, b] \approxeq [u, b'] \wedge [o, c] \approxeq [u, c']) \Rightarrow [b, c] \approxeq [b', c']).$$

Intuitivamente falando, o segmento $[b, c]$ nos dá a 'abertura' do ângulo

$$\widehat{r_o s_o},$$

sendo que essa 'abertura' é dada a partir dos 'parâmetros' b e c, uma vez que b incide sobre a semirreta r_o e c incide sobre a semirreta s_o.

Analogamente, o segmento $[b', c']$ nos dá a 'abertura' do ângulo

$$\widehat{m_u n_u},$$

sendo que essa 'abertura' é dada a partir dos 'parâmetros' b' e c', uma vez que b' incide sobre a semirreta m_u e c' incide sobre a semirreta n_u.

Neste contexto, a definição acima estabelece que ângulos congruentes contam com 'aberturas' congruentes, desde que os 'parâmetros de avaliação' definam segmentos congruentes. Esse mesmo critério de avaliação de 'abertura' é dado pelas condições $[o, b] \approxeq [u, b']$ e $[o, c] \approxeq [u, c']$.

Desnecessário enfatizar que nenhum ângulo é congruente a qualquer segmento. Afinal, congruência entre segmentos é uma relação sobre o conjunto de todos os segmentos, enquanto congruência entre ângulos é uma relação sobre o conjunto de todos os ângulos.

Definição 7.19. *Seja r uma reta de um plano ordenado*

$$\mathfrak{o} = \langle \langle \pi, \rho \rangle, \_\_ \rangle.$$

Definimos $\varkappa_r : \pi - r \to \pi - r$ como uma relação em $\pi - r$ dada por

$$(a, b) \in \varkappa_r \;\; sss \;\; \nexists p (p \in r \wedge \underline{apb}).$$

Na última definição $\pi - r$ é o conjunto dos pontos do plano π, exceto aqueles que pertencem à reta r. Observar que

$$\nexists p(p \in r \wedge \underline{apb}) \Leftrightarrow \forall p(p \in r \Rightarrow \neg \underline{apb})$$

é teorema. Logo, Definição 7.19 é equivalente a

$$(a, b) \in \varkappa_r \ \text{ sss } \ \forall p(p \in r \Rightarrow \neg \underline{apb}).$$

Teorema 7.2. *Para cada reta r de um plano ordenado*

$$\mathfrak{o} = \langle\langle \pi, \rho\rangle, \_\_\,\rangle,$$

$\varkappa_r$ é uma relação de equivalência.

Demonstração: Devemos provar que $\varkappa_r$ é reflexiva, simétrica e transitiva. Dividimos a demonstração em três partes.

Reflexividade: De acordo com GE6, $\forall_\pi a \forall_\pi p(\neg \underline{apa})$. Em particular, se p incide sobre r e a pertence a $\pi - r$, então $\neg \underline{apa}$. Logo, $(a, a) \in \varkappa_r$.

Simetria: Sabemos que

$$(a, b) \in \varkappa_r \Leftrightarrow \neg \underline{apb},$$

se p incide sobre r. Mas, GE7 implica que

$$\neg \underline{apb} \Leftrightarrow \neg \underline{bpa}.$$

Uma vez que

$$(b, a) \in \varkappa_r \Leftrightarrow \neg \underline{bpa},$$

então

$$(a, b) \in \varkappa_r \Leftrightarrow (b, a) \in \varkappa_r.$$

Transitividade: Devemos considerar três ocorrências de pontos a, b e c, uma vez que precisamos provar que

$$((a, b) \in \varkappa_r \wedge (b, c) \in \varkappa_r) \Rightarrow (a, c) \in \varkappa_r.$$

Para facilitar a demonstração, podemos dividi-la em etapas: o caso em que a, b e c são colineares e o caso em que não são. Na última etapa os pontos a, b e c são vértices de um triângulo e, por conta disso, o Axioma de Pasch (GE10) deve ser usado. Deixamos os detalhes para o leitor.

Ou seja, uma reta divide um plano ordenado em duas partes.

Em [10] (página 74) há uma demonstração de que, para toda reta r, o quociente

$$(\pi - r)/\varkappa_r$$

tem exatamente duas classes de equivalência. Se o leitor não recorda o que é o quociente de um conjunto por uma relação de equivalência, ver parágrafo que segue Teorema 3.11. Se denotarmos essas classes de equivalência por α_{1r} e α_{2r}, temos que

$$(\pi - r)/\varkappa_r = \{\alpha_{1r}, \alpha_{2r}\},$$

onde α_{1r} e α_{2r} são chamados de *semiplanos*.

Ou seja, qualquer reta r divide um plano ordenado em dois semiplanos, os quais não têm interseção entre si e tais que a união desses semiplanos produz o conjunto $\pi - r$.

Graças a esse último resultado, podemos encerrar esta Seção com a próxima definição.

Definição 7.20. *Seja α_r o conjunto de semiplanos definidos por uma reta r em um plano ordenado $\mathfrak{o} = \langle\langle\pi, \rho\rangle, \_\_\rangle$. Seja ainda*

$$\downarrow_o = \{r \in \wp(\pi) \mid r \textit{ é uma semirreta com origem } o\},$$

onde $o \in \pi$, de acordo com Definições 7.14 e 7.15. Um plano absoluto *é um par ordenado*

$$\langle\mathfrak{o}, \approxeq\rangle$$

onde GE11, GE12, GE13, GE14 *e a fórmula abaixo são teoremas.*

GE15: $\forall_\pi o \forall_\pi o' \forall_{\downarrow_o} a \forall_{\downarrow_o} b \forall_{\downarrow_{o'}} a' \forall_\rho r \forall_{\alpha_r} \beta (a' \subset r \Rightarrow \exists! b'(b' \in \downarrow_{o'} \wedge \widehat{ab} \approxeq \widehat{a'b'} \wedge b' \subset \beta))$.

Do ponto de vista intuitivo, axioma GE15 diz o que se segue. Dados

i: um ângulo $\widehat{ab}$ definido por semirretas a e b com a mesma origem o,

ii: uma reta r que divide o plano π em dois semiplanos e

iii: uma semirreta a', com origem o', que esteja contida na reta r, então, para cada semiplano β, existe uma única semirreta b' com origem o', contida em β, de modo que o ângulo $\widehat{a'b'}$ é congruente ao ângulo $\widehat{ab}$.

Ao leitor interessado, este último axioma corresponde ao postulado $\mathcal{C}_5$ em [10], páginas 124 e 125. Naquele texto há um erro no enunciado.

Seção 75

Axioma de continuidade

Sumário
Índice
Rede

Para que possamos enunciar o próximo postulado de geometria euclidiana, precisamos ser capazes de ordenar pontos de uma reta, em um plano ordenado. Em seguida seremos capazes de qualificar o que é um *eixo*, o qual é uma reta munida de orientação, também conhecida como *reta orientada*.

Definição 7.21. *Seja r uma reta em um plano ordenado*

$$\mathfrak{o} = \langle\langle \pi, \rho\rangle, \_\_\rangle.$$

Seja ainda

$$\downarrow_{(r)} = \{s \in \wp(\pi) \mid s \text{ é uma semirreta } \wedge s \subset r\}.$$

O conjunto $\rightleftharpoons_r$ é uma relação em $\downarrow_{(r)}$ dada como se segue:

$$a \rightleftharpoons_r b \quad \text{sss} \quad a \subseteq b \vee b \subseteq a.$$

Lemos $a \rightleftharpoons_r b$ como 'a semirreta a tem a mesma orientação *da semirreta b'.*

Usamos a notação $\downarrow_{(r)}$ (para designar o conjunto de todas as semirretas contidas na reta r) para não confundir com o conjunto $\downarrow_r$ das semirretas com origem r (onde r é um ponto).

Teorema 7.3. *A relação $\rightleftharpoons_r$ na Definição 7.21 é de equivalência.*

Demonstração: Devemos provar que $\rightleftharpoons_r$, no conjunto $\downarrow_{(r)}$ das semirretas contidas em r, é reflexiva, simétrica e transitiva. Por conta disso, dividimos a prova em três partes.

Reflexividade: Uma vez que todo conjunto é subconjunto de si mesmo, em particular, se a é uma semirreta contida em r, então $a \subseteq a$. Logo, $a \rightleftharpoons_r a$.

Simetria: Temos que $a \rightleftharpoons_r b$ sss $a \subseteq b \vee b \subseteq a$. No entanto, $(a \subseteq b \vee b \subseteq a) \Leftrightarrow (b \subseteq a \vee a \subseteq b)$ é teorema. Logo, $a \rightleftharpoons_r b$ sss $b \rightleftharpoons_r a$.

Transitividade: Temos que

$$a \rightleftharpoons_r b \ \text{ sss } \ a \subseteq b \vee b \subseteq a$$

e

$$b \rightleftharpoons_r c \ \text{ sss } \ b \subseteq c \vee c \subseteq c.$$

Dessa maneira, há quatro possibilidades que devem ser avaliadas.

(i) Se $a \subseteq b$ e $b \subseteq c$, então é imediato que $a \subseteq c$, o que implica em $a \rightleftharpoons_r c$ (i.e., transitividade para a primeira possibilidade).

(ii) Se $a \subseteq b$ e $c \subseteq b$ (ou seja, duas semirretas a e c estão contidas em uma mesma semirreta b), então

$$a \subseteq c \vee c \subseteq a$$

(✍ recomendamos que o leitor prove isso).

Seja qual for o caso, a última fórmula é equivalente a $a \rightleftharpoons_r c$ (i.e., transitividade para a segunda possibilidade).

(iii) Se $b \subseteq a$ e $b \subseteq c$ (ou seja, uma mesma semirreta b está contida em duas semirretas a e c), então

$$a \subseteq c \vee c \subseteq a$$

(✍ recomendamos que o leitor prove isso).

Seja qual for o caso, a última fórmula é equivalente a $a \rightleftharpoons_r c$ (i.e., transitividade para a terceira possibilidade).

(iv) Finalmente, se $b \subseteq a$ e $c \subseteq b$, temos situação análoga ao item (i). Logo, $c \subseteq a$, o que implica em $a \rightleftharpoons_r c$ (i.e., transitividade para a quarta e última possibilidade).

O próximo teorema garante que toda reta de um plano ordenado admite exatamente duas possíveis orientações.

Teorema 7.4. *Seja r uma reta em um plano ordenado*

$$\mathfrak{o} = \langle\langle\pi, \rho\rangle, \_\_\rangle.$$

Se

$$\downarrow_{(r)} = \{s \in \wp(\pi) \mid s \text{ é uma semirreta } \wedge s \subset r\}$$

e $\rightleftharpoons_r$ é a relação de equivalência em $\downarrow_{(r)}$ dada pela Definição 7.21, então o quociente

$$\downarrow_{(r)} / \rightleftharpoons_r$$

admite apenas dois elementos.

Demonstração: Seja a uma semirreta de $\downarrow_{(r)}$ com origem o. Seja ainda

$$a^{\star} = r - (a \cup \{o\}).$$

Logo, $a^{\star}$ é uma semirreta pertencente a $\downarrow_{(r)}$ (✍ cabe ao leitor provar). Além disso, $a \cap a^{\star} = \varnothing$, o que implica que $a \not\subseteq a^{\star}$ e $a^{\star} \not\subseteq a$. Portanto, a e $a^{\star}$ são representantes de classes de equivalência diferentes, ou seja, $\neg(a \rightleftharpoons_r a^{\star})$.

O próximo passo é considerar uma semirreta b pertencente a $\downarrow_{(r)}$ tal que $b \neq a$ e $b \neq a^{\star}$. Neste caso, b necessariamente pertence à classe de equivalência com representante a ou à classe de equivalência com representante $a^{\star}$. ✍ Cabe ao leitor concluir a demonstração.

As classes de equivalência do último teorema são as duas únicas orientações de qualquer reta de um plano ordenado, conforme a próxima definição.

Definição 7.22. *Seja r uma reta em um plano ordenado. Se*

$$\downarrow_{(r)} = \{s \in \wp(\pi) \mid s \text{ é uma semirreta } \wedge s \subset r\}$$

e $\rightleftharpoons_r$ é a relação de equivalência em $\downarrow_{(r)}$ dada pela Definição 7.21, então o quociente

$$\downarrow_{(r)} / \rightleftharpoons_r$$

é definido como

$$\{\rightharpoonup_r, \leftharpoondown_r\},$$

onde as classes de equivalência $\rightharpoonup_r$ *e* $\leftharpoondown_r$ *são chamadas de* orientações *da reta* r. *Além disso, cada* *par ordenado* $(r, \rightharpoonup_r)$ *e* $(r, \leftharpoondown_r)$ *é chamado de* reta orientada *ou* eixo.

Em outras palavras, uma reta orientada é uma reta r munida de uma orientação, a qual pode ser $\rightharpoonup_r$ ou $\leftharpoondown_r$. Logo, toda reta de um plano ordenado admite duas retas orientadas definíveis a partir dela.

Retas orientadas, em um plano ordenado, permitem ordenar pontos de uma mesma reta, como se percebe a seguir.

Definição 7.23. *Seja* x *uma orientação de uma reta* r *em um plano ordenado. Se* a *e* b *são pontos de* r *tais que*

I: a *define semirreta* r_a *pertencente a* x,

II: b *define semirreta* r_b *pertencente a* x *e*

III: $r_b \subseteq r_a$,

então a precede b *na orientação* x *e denotamos isso como*

$$a \preceq_x b.$$

Se $a \neq b$, *nas condições dadas acima, dizemos que* a precede estritamente b *e denotamos isso como*

$$a \prec_x b.$$

Teorema 7.5. *Sejam* a *e* b *pontos distintos de uma reta* r *em um plano ordenado. Logo*

$$a \prec_{\rightharpoonup_r} b \Leftrightarrow b \prec_{\leftharpoondown_r} a.$$

Teorema 7.6. *Seja* x *uma orientação de uma reta* r *em um plano ordenado. Logo,* $\preceq_x$ *é uma relação de ordem total no eixo* (r, x).

✍ As provas dos dois últimos teoremas ficam a cargo do leitor.

Definição 7.24. *Seja* (r, z) *uma reta orientada em um plano ordenado. Dizemos que o* *par ordenado* (x, y) *é um* corte de Dedekind *de* r *sss*

I: $x \neq \varnothing \wedge y \neq \varnothing$,

II: $x \subset r \wedge y \subset r$,

III: $x \cup y = r \wedge x \cap y = \varnothing$ e

IV: $\forall p \forall q((p \in x \wedge q \in y) \Rightarrow p \preceq_z q)$.

O fato de cortes de Dedekind dependerem da relação de ordem total $\preceq_z$, para uma dada orientação z, justifica o emprego de par ordenado para defini-los.

Naturalmente, todo corte de Dedekind particiona uma reta orientada em dois conjuntos x e y, de modo que os pontos pertencentes a x precedem os pontos pertencentes a y.

Definição 7.25. *Seja* $\mathfrak{c} = \langle \mathfrak{o}, \cong \rangle$ *um plano absoluto, onde*

$$\mathfrak{o} = \langle \langle \pi, \rho \rangle, \_\_ \rangle$$

é um plano ordenado.

Dizemos que $\mathfrak{c}$ *é um* plano absoluto contínuo *sss a fórmula abaixo for teorema.*

GE16:

$$\forall_\rho r \forall z \forall x \forall x(((x, y) \text{ é corte de Dedekind de } r \wedge$$
$$z \text{ é orientação de } r) \Rightarrow$$
$$\exists s(s \in r \wedge \forall_r p \forall_r q((p \prec s \Rightarrow p \in x) \wedge (s \prec q \Rightarrow q \in y)))).$$

✍ É um exercício interessante escrever formalmente o postulado GE16.

A fórmula GE16 é conhecida como *axioma de Dedekind*, apesar de Joseph Bertrand ter trabalhado com o mesmo assunto antes de Richard Dedekind.

Seguindo a orientação z de uma reta r e lembrando que (x, y) é um corte de Dedekind de r, o postulado acima garante o seguinte: o ponto s da reta r (a existência de s é garantida pelo postulado GE16) é o último de x ou o primeiro de y, relativamente à orientação z.

O propósito deste axioma é claro. Se, por exemplo, uma reta t define dois semiplanos α_{1t} e α_{2t} e, além disso, uma reta r passa por um ponto p no semiplano α_{1t} e por um ponto q no semiplano α_{2t}, então o postulado de continuidade dado acima garante que há interseção entre as retas t e r. Tal interseção é o ponto s. Notar que isso não ocorre no modelo de plano quase-ordenado do Exemplo 7.8.

Seção 76

Axioma de paralelismo

Sumário
Índice
Rede

Podemos finalmente conceituar geometria euclidiana plana.

Definição 7.26. *Um* plano euclidiano *é um plano absoluto contínuo*

$$\langle\langle\langle \pi, \rho\rangle, \underline{}\rangle, \approxeq\rangle$$

onde a fórmula abaixo é teorema.

GE17:

$\forall_\rho r \forall_\rho a \forall_\rho b \forall_\pi p((p \notin r \wedge p \in a \cap b \wedge a \cap r = \varnothing \wedge b \cap r = \varnothing) \Rightarrow a = b)$.

A fórmula GE17 é conhecida como *Postulado das Paralelas* ou *axioma de Playfair*, em referência a John Playfair. Sua leitura é muito intuitiva. São assumidas retas r, a e b, e um ponto p que não incide sobre r, mas incide sobre a e b. Se as retas a e b forem paralelas a r (ou seja, $a \cap r = \varnothing \wedge b \cap r = \varnothing$), então a e b são a mesma reta. Em outras palavras, dada uma reta r e um ponto p não incidente sobre r, existe uma única reta que passa por p e é paralela a r.

✍ Obviamente GE17 pode ficar com um enunciado mais curto se for reescrito usando o quantificador $\exists!$.

Definição 7.26 é equivalente à seguinte fórmula:

Um plano euclidiano é uma quádrupla ordenada

$$\langle \pi, \rho, \underline{}, \approxeq\rangle,$$

onde as fórmulas GE1∼GE17 *são teoremas.*

Neste contexto, as fórmulas

I: GE1, GE2, GE3, e GE4 são os axiomas de incidência,

II: GE5, GE6, GE7, GE8, GE9 e GE10 são os axiomas de ordem,

III: GE11, GE12, GE13, GE14 e GE15 são os axiomas de congruência,

IV: GE16 é o axioma de continuidade e

V: GE17 é o axioma de paralelismo.

Seção 77

Modelo de plano euclidiano

Sumário
Índice
Rede

Os únicos modelos que vimos até agora, relativos a geometria sintética, foram para plano de incidência (Exemplo 7.2) e plano quase-ordenado (Exemplo 7.8), o qual é também um plano de incidência. Nesta Seção, porém, exibimos um modelo muito conhecido para plano euclidiano, a saber, o *plano cartesiano*.

Plano cartesiano é a quádrupla ordenada

$$\langle \pi, \rho, \_\_, \approxeq \rangle,$$

na qual interpretamos π, ρ, $\_\_$ e $\approxeq$ como se segue nos próximos parágrafos.

Interpretação de π: O conjunto π é $\mathbb{R} \times \mathbb{R}$, onde $\mathbb{R}$ é o conjunto dos números reais, conforme Seção 39. Em outras palavras, $\pi = \mathbb{R}^2$.

Isso significa que interpretamos pontos do plano cartesiano como pares ordenados (x, y) de números reais.

Interpretação de ρ: Uma *reta* $r_{(a,b,c)}$ é o conjunto

$$r_{(a,b,c)} = \{(x, y) \in \mathbb{R}^2 \mid ax + by = c\},$$

onde a e b não são ambos nulos.

A tripla ordenada (a, b, c) é chamada de *parâmetros da reta* $r_{(a,b,c)}$. Ou seja, na tripla ordenada (a, b, c), o parâmetro real a é o termo que multiplica x na equação $ax+by = c$, enquanto b é aquele que multiplica y, e c é o parâmetro real que não multiplica nem por x e nem por y (comumente chamado de *termo independente*). Lembrar que x e y são usados para definir os pontos (x, y) que *incidem* sobre a reta $r_{(a,b,c)}$. Em outras palavras, (x, y) incide sobre $r_{(a,b,c)}$ sss $(x, y) \in r_{(a,b,c)}$.

✍ Notar que, se $r_{(a,b,c)}$ é uma reta, então

$$r_{(a,b,c)} = r_{(\alpha a,\alpha b,\alpha c)},$$

para qualquer α real não nulo.

Definimos ρ como

$$\rho = \{r_{(a,b,c)} \in \wp(\mathbb{R}^2) \mid a \in \mathbb{R} \wedge b \in \mathbb{R} \wedge c \in \mathbb{R} \wedge (a \neq 0 \vee b \neq 0)\}.$$

Ou seja, ρ é o conjunto de todas as retas do plano cartesiano.

Observar que

$$(x, y) \in r_{(a,b,c)} \Leftrightarrow ax + by = c.$$

Portanto, dados pontos distintos (x_1, y_1) e (x_2, y_2), temos como teorema a seguinte fórmula:

$$((x_1, y_1) \in r_{(a,b,c)} \wedge (x_2, y_2) \in r_{(a,b,c)}) \Leftrightarrow ax_1 + by_1 = ax_2 + by_2,$$

observando que tanto $ax_1 + by_1$ quanto $ax_2 + by_2$ são iguais ao parâmetro c. Ou seja, pontos (x_1, y_1) e (x_2, y_2), distintos entre si, são incidentes sobre a mesma reta r sss existem a e b tais que $ax_1 + by_1 = ax_2 + by_2$. Isso garante que axioma GE2 é teorema nesta interpretação.

Com relação a GE1, sejam (x_1, y_1) e (x_2, y_2) pontos de $\mathbb{R}^2$ tais que $(x_1, y_1) \neq (x_2, y_2)$. Logo, a equação $ax_1 + by_1 = ax_2 + by_2$ admite soluções reais onde a e b não são ambos nulos. Com efeito,

$$ax_1 + by_1 = ax_2 + by_2 \Leftrightarrow a(x_1 - x_2) = b(y_2 - y_1).$$

A última equação admite solução para ambos a e b nulos. Mas, lembrando que $(x_1, y_1) \neq (x_2, y_2)$, então $x_1 - x_2 \neq 0$ ou $y_2 - y_1 \neq 0$. Logo, também existem soluções para a e b, onde $a \neq 0$ ou $b \neq 0$. Isso garante que existe reta incidente sobre (x_1, y_1) e (x_2, y_2). Logo, GE1 é teorema.

✍ As demonstrações de GE3 e GE4 ficam a cargo do leitor.

Interpretação de __: Entre os reais há uma relação de ordem total $\leq$, como discutido na Seção 39. Também há uma relação de ordem parcial $<$ dada por $r < s$ sss $r \leq s \wedge r \neq s$. Logo, para quaisquer reais r e s distintos entre si temos que $r < s$ ou $s < r$. No primeiro caso, podemos definir o intervalo aberto (r, s), enquanto que, no segundo caso, podemos definir o intervalo aberto (s, r). Sobre o conceito de intervalo aberto no conjunto dos reais, ver Seção 39. Ou seja, dados reais r e s distintos entre si, sempre existe intervalo aberto definido por r e s, no sentido deste intervalo ser (r, s) ou (s, r).

STOP Não confundir um intervalo aberto (r, s) com um ponto (r, s) do plano cartesiano. É uma prática comum o emprego de uma mesma notação para conceitos não equivalentes entre si.

Neste contexto, se r e s são reais distintos, um real t está *entre* r e s sss t pertence ao intervalo aberto definido por r e s. Notar que tal definição implica que $r \neq t$ e $s \neq t$.

Graças à definição acima de real t entre dois reais r e s (a qual é possível por conta da relação de ordem parcial $<$ entre os reais), podemos agora interpretar a relação de ordem $\underline{\quad}$.

Dados os pontos (x_1, y_1) e (x_2, y_2) tais que $(x_1, y_1) \neq (x_2, y_2)$, dizemos que o ponto (x_e, y_e) *está entre* (x_1, y_1) e (x_2, y_2) sss os pontos (x_1, y_1), (x_2, y_2) e (x_e, y_e) incidem sobre a mesma reta e, além disso, uma e apenas uma das fórmulas abaixo é teorema:

- $x_e = x_1 = x_2$ e y_e está entre y_1 e y_2.
- $y_e = y_1 = y_2$ e x_e está entre x_1 e x_2.
- x_e está entre x_1 e x_2 e y_e está entre y_1 e y_2.

Escrevemos isso como $\underline{(x_1, y_1)(x_e, y_e)(x_2, y_2)}$.

✍ As provas de GE5∼GE9 são imediatas. A demonstração do Axioma de Pasch (postulado GE10), porém, consome considerável esforço. Mas é um típico exercício de geometria analítica plana. Deixamos essas provas como exercícios para o leitor.

✍ Para facilitar a demonstração de GE10, é interessante que o leitor prove o seguinte teorema.

Teorema 7.7. *Se*

$$r_{(a,b,c)} = \{(x, y) \in \mathbb{R}^2 \mid ax + by = c\}$$

e

$$r_{(d,e,f)} = \{(x, y) \in \mathbb{R}^2 \mid dx + ey = f\}$$

são retas do plano cartesiano, então

$$r_{(a,b,c)} \cap r_{(d,e,f)} = \varnothing \Leftrightarrow \exists \alpha (a = \alpha d \wedge b = \alpha e \wedge c \neq \alpha f).$$

Interpretação de $\approxeq$: A partir do momento em que definimos o que é um ponto entre dois pontos do plano cartesiano, basta empregar a Definição 7.11 (sobre segmentos fechados) para a interpretação deste conceito aqui.

Uma vez que sabemos o que é um segmento de reta

$$[(x_1, y_1), (x_2, y_2)],$$

definimos o *comprimento de* $[(x_1, y_1), (x_2, y_2)]$ como

$$\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}.$$

Feito isso, dizemos que os segmentos

$$[(x_1, y_1), (x_2, y_2)] \text{ e } [(x_3, y_3), (x_4, y_4)]$$

são *congruentes* sss

$$\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} = \sqrt{(x_3 - x_4)^2 + (y_3 - y_4)^2}.$$

Denotamos isso como $[(x_1, y_1), (x_2, y_2)] \approxeq [(x_3, y_3), (x_4, y_4)]$.

Essas informações facilitam a demonstração das fórmulas GE11, GE12, GE13, GE14 e GE15, as quais, novamente, são típicos exercícios de geometria analítica plana.

A título de curiosidade, no Exemplo 8.42 da Seção 88 provamos que o comprimento de um segmento $[(x_1, y_1), (x_2, y_2)]$ é também uma *distância* entre os pontos (x_1, y_1) e (x_2, y_2). Este fato garante que

$$[(x_1, y_1), (x_2, y_2)] \approxeq [(x_2, y_2), (x_1, y_1)].$$

Afinal

$$\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}.$$

✍ A demonstração de GE17 (Axioma de Playfair) é outro exercício de geometria analítica que pode ser resolvido usando o Teorema 7.7. Basta manter em mente que retas paralelas, no plano cartesiano, são retas com interseção vazia.

✍ Com relação a cortes de Dedekind (postulado GE16), recordar que, em $\mathbb{R}$, toda sequência de Cauchy é convergente. Isso garante que GE16 é teorema nesta interpretação de plano euclidiano. Se tivéssemos interpretado π como $\mathbb{Q}^2$, GE16 não seria teorema. Cabe ao leitor provar isso.

Geometria euclidiana plana é o estudo do conjunto

$$\langle\langle\langle \pi, \rho \rangle, \_\_ \rangle, \approxeq \rangle,$$

desde que os postulados GE1$\sim$GE17 sejam teoremas de ZF.

Seção 78

Resumo da ópera

Sumário
Índice
Rede

Apesar de geometria ter nascido há milhares de anos, a partir de modos de percepção humana sobre espaço, obviamente o assunto evoluiu para um elevado nível de abstração. Quando Euclides de Alexandria escreveu sua grande obra *Elementos*, ele apenas assumiu implicitamente certos postulados, sem efetivamente enunciá-los. Exemplo bem conhecido é o Axioma de Pasch, o qual só foi explicitado dois milênios após a obra de Euclides.

STOP O texto original de Euclides foi perdido, por conta da destruição da Biblioteca de Alexandria. Nas traduções que sobreviveram há algumas 'definições'. Uma delas diz que 'um ponto é aquilo que não tem partes'. Em outro momento, Euclides escreve que 'uma reta tem comprimento, mas não largura'. Claramente essas afirmações não são definições, uma vez que naquele texto não há qualificação para os conceitos de 'aquilo', 'partes', 'comprimento' ou 'largura'.

Por conta disso, a grande revolução promovida por Hilbert, em seu *Grundlagen der Geometrie*, foi a proposta de que conceitos como pontos e retas não são definíveis. A geometria euclidiana, nos moldes da proposta de Hilbert, é uma teoria de caráter meramente sintático, desprovido de significado.

O que fizemos aqui foi adaptar as ideias de Hilbert para a linguagem de ZF, a qual é uma teoria formal cuja linguagem não conta com qualquer contraparte semântica.

Como reza a lenda (sobre uma suposta conversa entre Hilbert e Otto Blumenthal, em uma estação de trem em Berlim), se trocarmos as palavras 'ponto', 'reta' e 'plano' por 'mesa', 'cadeira' e 'copo de cerveja', tudo o que interessa é que os axiomas da geometria euclidiana sejam teoremas em uma dada interpretação. Dessa maneira podemos garantir que tal interpretação é um modelo de geometria euclidiana.

Seção 79

Notas históricas

Sumário
Índice
Rede

O registro mais antigo de estudo sistemático da geometria é a obra *Elementos*, de Euclides de Alexandria (*Εὐκλείδηζ*). Fragmentos de cópias do texto original podem ser encontrados na Biblioteca do Vaticano.

Durante mais de dois mil anos matemáticos tentaram provar o postulado das paralelas a partir dos demais postulados originais de Euclides, sem sucesso. Isso por conta de uma visão intuitiva de que o postulado das paralelas não era 'autoevidente'.

Foi somente no século 19 que Nikolai Lobachevsky publicou uma prova de que tal postulado era independente dos demais. Por conta disso, o método axiomático utilizado por Euclides deixou de ser uma ferramenta meramente didática para o ensino de geometria e passou a ser alvo de interesse matemático. A obra de Lobachevsky também abriu portas para a percepção de geometria como tema de estudos independentes dos modos de percepção humana sobre espaço. Por fim, a contribuição deste matemático russo serviu de inspiração para muitos pensadores questionarem conhecimentos tradicionais não apenas da matemática, mas também de outras áreas do saber.

Por conta do impacto acima mencionado, William Kingdon Clifford chegou a escrever que Lobachevsky foi o 'Copérnico da geometria', enquanto Eric Temple Bell foi além, afirmando que Lobachevsky foi o 'Copérnico de todo o pensamento humano' [3].

O asteroide 1858 Lobachevsky, a cratera lunar Lobachevsky, a universidade russa Lobatchevsky e a canção Lobachevsky, de Tom Lehrer, são algumas das homenagens póstumas a este grande nome da ciência.

Retrato de Lobachevsky, feito por Lev Kryukov.

Fonte: Wikipedia.

PARTE 8

Álgebra linear

ℵ

Nesta oitava parte discutimos a respeito de noções indispensáveis para o estudo de álgebra linear, bem como suas relações com geometria euclidiana e cálculo diferencial e integral.

Seção 80

Espaços vetoriais reais

Sumário
Índice
Rede

O estudo de *espaços vetoriais* e *transformações lineares* entre espaços vetoriais é o que se chama de *álgebra linear*. Para que essa afirmação seja compreensível, é necessário qualificar o que são espaços vetoriais e transformações lineares. Começamos com um caso muito particular, conhecido como *espaço vetorial real de dimensão finita*. O conceito de transformação linear é examinado mais adiante. Espaços vetoriais diferentes de espaços vetoriais reais são brevemente discutidos na Seção 96. Espaços vetoriais de dimensão infinita são examinados na Seção 97.

Antes, porém, é interessante motivarmos o assunto. Fazemos isso com um exemplo histórico.

Em 1873 James Clerk Maxwell publicou seu famoso livro *A Treatise on Electricity and Magnetism* (Um Tratado sobre Eletricidade e Magnetismo). A partir de um sistema de vinte e quatro equações, Maxwell formulou uma visão unificada para campos elétricos e campos magnéticos, dando início àquilo hoje conhecido como eletromagnetismo. Posteriormente Oliver Heaviside reescreveu as equações originais de Maxwell empregando funções definidas sobre espaços vetoriais. Essa estratégia reduziu as vinte e quatro equações originais de Maxwell a apenas quatro.

O exemplo acima é apenas um caso muito simples e bem conhecido para ilustrar o poder de síntese de espaços vetoriais. Outro exemplo bem mais radical é o caso da mecânica quântica. Sem espaços vetoriais, não existiria hoje qualquer formulação teórica minimamente sensata para descrever os fenômenos do mundo quântico. Consequentemente, não existiria este arquivo PDF que o leitor está contemplando, como podemos verificar ao final desta Parte.

Definição 8.1. *Um* espaço vetorial real $\mathcal{V}$ *é uma quíntupla ordenada*

$$\langle V, \mathbb{R}, +, \cdot, \overline{0} \rangle$$

tal que as seguintes fórmulas são teoremas.

V1: $V \neq \varnothing$;

V2: $+ : V \times V \to V$ *é uma função, onde abreviamos* $+(u, v)$ *como* $u + v$, *sendo* u *e* v *elementos de* V;

V3: $\cdot : \mathbb{R} \times V \to V$ *é uma função, onde abreviamos* $\cdot(\alpha, u)$ *como* $\alpha \cdot u$ *ou, simplesmente,* αu, *sendo* α *um elemento de* $\mathbb{R}$ *e* u *um elemento de* V;

V4: $\overline{0} \in V$;

V5: *Se* u *pertence a* V, *então*

$$u + \overline{0} = u;$$

V6: *Se* u *e* v *são elementos de* V, *então*

$$u + v = v + u;$$

V7: *Se* u, v *e* w *pertencem a* V, *então*

$$(u + v) + w = u + (v + w);$$

V8: *Se* u *pertence a* V, *então existe* v *pertencente a* V *tal que*

$$u + v = \overline{0};$$

V9: *Se α pertence a $\mathbb{R}$ e u e v pertencem a V, então*

$$\alpha(u + v) = \alpha u + \alpha v;$$

V10: *Se α e β pertencem a $\mathbb{R}$ e u pertence a V, então*

$$(\alpha + \beta)u = \alpha u + \beta u;$$

V11: *Se α e β pertencem a $\mathbb{R}$ e u pertence a V, então*

$$(\alpha\beta)u = \alpha(\beta u);$$

V12: *Se 1 é o neutro multiplicativo de $\mathbb{R}$ e u pertence a V, então*

$$1u = u.$$

Se

$$\mathcal{V} = \langle V, \mathbb{R}, +, \cdot, \overline{0}\,\rangle$$

é um espaço vetorial real, chamamos V de *conjunto de vetores* ou *espaço de vetores*.

Alguns autores, por abuso de linguagem, se referem a V como espaço vetorial. Não adotamos essa convenção aqui.

Os elementos de V são chamados de *vetores*. Com relação a $\mathbb{R}$, este é o corpo dos reais, já definido na Seção 39. Seus elementos são chamados de *escalares*.

A função $+$ é chamada de *adição de vetores*. A função $\cdot$ é chamada de *multiplicação de escalar por vetor*. Apesar da notação ser a mesma, essas operações não podem ser confundidas com adição e multiplicação entre reais.

O termo $\overline{0}$ é chamado de *vetor nulo*.

Para evitarmos sobrecarga de notação, adotamos a seguinte convenção: todos os vetores diferentes do vetor nulo são denotados por letras latinas minúsculas, enquanto os escalares são sempre denotados por letras gregas minúsculas. Logo, $\alpha + \beta$ é uma adição entre reais, enquanto $u + v$ é uma adição entre vetores. Analogamente, $\alpha \cdot \beta$ é uma multiplicação entre escalares, enquanto $\alpha \cdot u$ é uma multiplicação entre um real α e um vetor u.

Postulado V1 diz que todo espaço de vetores tem pelo menos um vetor. Uma vez que axioma V4 afirma que o vetor nulo é um vetor, todo espaço de vetores conta com pelo menos o vetor nulo entre seus elementos. Obviamente V1 é desnecessário, uma vez que V4 implica em V1.

V2 estabelece que a adição entre vetores é uma operação binária, ou seja, é aplicável a duas ocorrências de vetores. Além disso, o mesmo postulado afirma que adição entre vetores é uma operação fechada, i.e., vetor mais vetor é vetor.

V3 diz que escalar *vezes* vetor é vetor.

V5 afirma que o vetor nulo é nulo relativamente à adição entre vetores. Isso justifica seu nome.

V6 e V7 estabelecem, respectivamente, a comutatividade e a associatividade da adição entre vetores.

V8 diz que todo vetor admite simétrico relativamente à adição de vetores (chamado de simétrico aditivo). Se $u + v = \overline{0}$, denotamos v por $-u$. Equivalentemente, $u = -v$.

V9 é a distributividade da adição de vetores relativamente à multiplicação de escalar por vetor, enquanto V10 exige a distributividade da adição de escalares relativamente à multiplicação de escalar por vetor.

STOP Observar atentamente a fórmula V11:

$$(\alpha\beta)u = \alpha(\beta u).$$

Neste caso, $\alpha\beta$ é a multiplicação do real α pelo real β. Essa multiplicação resulta em um real γ. Neste sentido

$$(\alpha\beta)u = \gamma u.$$

Do lado direito da igualdade, não obstante, temos apenas duas ocorrências da multiplicação de escalar por vetor e nenhuma ocorrência da multiplicação de real por real. Com efeito, βu é um vetor, de acordo com postulado V3. Logo, $\alpha(\beta u)$ também é um vetor, novamente usando postulado V3. Logo, V11 impõe uma estreita conexão entre multiplicação de real por real, algo definido na Seção 39, e multiplicação de real por vetor.

Finalmente, V12 afirma que a multiplicação do real 1 por qualquer vetor u resulta no próprio vetor u. Mas de forma alguma esse postulado deve ser interpretado como a garantia de existência de um neutro multiplicativo. Afinal, 1 denota um escalar, o qual é neutro multiplicativo entre reais. O termo u, por sua vez, é um vetor.

SEÇÃO 81

Modelos de espaços vetoriais reais

SUMÁRIO
ÍNDICE
REDE

O que Definição 8.1 coloca é o seguinte: qualquer quíntupla ordenada

$$\langle V, \mathbb{R}, +, \cdot, \overline{0}\, \rangle$$

na qual as fórmulas V1∼V12 são teoremas, é um espaço vetorial real.

Uma vez escolhidos os conjuntos V, $+$, $\cdot$ e $\overline{0}$, a quíntupla ordenada

$$\langle V, \mathbb{R}, +, \cdot, \overline{0}\rangle$$

é uma *interpretação* de espaço vetorial. Se, dada a interpretação $\langle V, \mathbb{R}, +, \cdot, \overline{0}\, \rangle$, os axiomas V1∼V12 são teoremas de ZF, então essa interpretação é um *modelo* de espaço vetorial real. Logo, 'modelo de espaço vetorial real' e 'espaço vetorial real' são sinônimos.

Colocamos a seguir algumas possíveis interpretações de espaço vetorial real, avaliando se elas são modelos de espaços vetoriais reais.

ESPAÇO $\mathbb{R}^2$ USUAL

Seja $\langle \mathbb{R}^2, \mathbb{R}, +, \cdot, (0,0)\rangle$ uma quíntupla ordenada tal que

$$+ : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}^2$$

é uma função na qual se abrevia $+((a,b),(c,d))$ como $(a,b)+(c,d)$ e

$$\cdot : \mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}^2$$

é uma função na qual se abrevia $\cdot(\alpha,(a,b))$ como $\alpha \cdot (a,b)$ ou simplesmente $\alpha(a,b)$. Além disso, essas funções são definidas da seguinte maneira:

$$(a,b) + (c,d) = (a+c, b+d)$$

e

$$\alpha \cdot (a,b) = (\alpha a, \alpha b).$$

Uma primeira crítica que o leitor poderia fazer é a seguinte: se $(a,b) \in \mathbb{R}^2$, por que não denotar o par ordenado (a,b) por (α,β), uma vez que a e b são números reais?

Pois bem, o que está em jogo aqui são apenas *duas* operações: adição entre pares ordenados de reais e multiplicação de real por par ordenado de reais. Neste contexto, usamos letras gregas minúsculas apenas para denotar os reais que multiplicam por um par ordenado de reais. Usamos letras latinas minúsculas, nesta interpretação particular de espaço vetorial real, para denotar os reais a e b que definem o par ordenado (a, b).

A quíntupla ordenada $\langle \mathbb{R}^2, \mathbb{R}, +, \cdot, (0,0)\rangle$ é de fato uma interpretação de espaço vetorial real $\langle V, \mathbb{R}, +, \cdot, \overline{0}\,\rangle$, no sentido de que interpretamos

- o espaço de vetores V como $\mathbb{R}^2$,
- a adição $+$ de vetores como adição entre pares ordenados de reais,
- a multiplicação $\cdot$ de escalar por vetor como a multiplicação de real por par ordenado de reais e
- o vetor nulo $\overline{0}$ como o par ordenado $(0, 0)$, onde 0 é o real nulo aditivo.

Porém, a questão importante é se essa primeira interpretação de espaço vetorial real é um *modelo* de espaço vetorial real. Para que seja o caso, é necessário que todos os axiomas de espaço vetorial real sejam teoremas de ZF nesta interpretação. Ou seja, os doze axiomas dados na Seção 80 devem ser examinados um a um.

Axioma V1 exige que V seja não vazio. Essa fórmula é teorema na interpretação dada. Com efeito, existe pelo menos um par ordenado de reais, por exemplo, $(\sqrt{2} - 4\sqrt{7}, \pi)$. Logo, $\mathbb{R}^2$ é um conjunto não vazio.

Postulado V2 exige que adição de vetores seja dada por uma função fechada no espaço de vetores. Logo, tal fórmula é teorema no contexto da interpretação dada. Com efeito, adição de pares ordenados de reais foi definida como uma função

$$+ : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}^2.$$

Além disso, se (a, b) e (c, d) são pares ordenados de reais, então $(a, b) + (c, d)$, dado por $(a + c, b + d)$, é um par ordenado de reais simplesmente porque a adição de reais é fechada nos reais (real mais real é um real).

V3 demanda que multiplicação de real por vetor seja definida por uma função fechada no espaço de vetores. Logo, V3 é teorema no contexto da interpretação aqui sugerida. Com efeito, a multiplicação de real por par ordenado de reais foi definida por uma função

$$\cdot : \mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}^2$$

dada por $\alpha(a, b) = (\alpha a, \alpha b)$.

Além disso, $(\alpha a, \alpha b)$ pertence a $\mathbb{R}^2$, uma vez que multiplicação de real por real é fechada nos reais.

Axioma V4 exige que o vetor nulo seja vetor. Novamente temos um teorema, uma vez que $(0, 0)$ é um par ordenado de reais e, portanto, $(0, 0) \in \mathbb{R}^2$.

V5 afirma que qualquer vetor u somado ao vetor nulo $\overline{0}$ é o próprio u. Novamente temos um teorema. Com efeito, para qualquer par ordenado (a, b) de $\mathbb{R}^2$ temos

$$(a, b) + (0, 0) = (a + 0, b + 0) = (a, b).$$

Ou seja, o fato de zero real ser nulo aditivo garante que o par ordenado $(0, 0)$ é nulo relativamente à operação $+$ definida para pares ordenados de reais.

V6 exige que adição de vetores seja comutativa. Tal fórmula também é teorema de ZF para a interpretação dada. Afinal,

$$(a, b) + (c, d) = (a + c, b + d),$$

de acordo com a definição de adição de pares ordenados de reais. No entanto,

$$(a + c, b + d) = (c + a, d + b),$$

uma vez que adição de reais é comutativa. Finalmente,

$$(c + a, d + b) = (c, d) + (a, b),$$

por conta da definição de adição de pares ordenados de reais. A transitividade da igualdade garante, portanto, que

$$(a, b) + (c, d) = (c, d) + (a, b).$$

Postulado V7 demanda a associatividade da adição de vetores. Mas,

$$((a, b)+(c, d))+(e, f) = (a+c, b+d)+(e, f) = ((a+c)+e, (b+d)+f) =$$

$$(a+(c+e), b+(d+f)) = (a, b)+(c+e, d+f) = (a, b)+((c, d)+(e, f)).$$

Observar que, na sequência de igualdades acima, foi usada a definição de adição de pares ordenados de reais, bem como o teorema que garante a associatividade da adição de reais. Logo, V7 é teorema para esta interpretação.

Axioma V8 exige que todo vetor admita um simétrico relativamente à operação de adição de vetores. Ora, se (a, b) é um termo de $\mathbb{R}^2$, temos que $(-a, -b)$ também pertence a $\mathbb{R}^2$. Afinal, todo real admite simétrico aditivo. Além disso,

$$(a, b) + (-a, -b) = (a + (-a), b + (-b)) = (0, 0),$$

sendo que $(0, 0)$ é interpretado aqui como o vetor nulo. Portanto, V8 é teorema nesta interpretação.

V9 estabelece a distributividade da adição de vetores relativamente à multiplicação de escalar por vetor. Logo, V9 é teorema. Com efeito,

$$\alpha((a, b) + (c, d)) = \alpha(a + c, b + d) = (\alpha(a + c), \alpha(b + d)) =$$
$$(\alpha a + \alpha c, \alpha b + \alpha d) = (\alpha a, \alpha b) + (\alpha c, \alpha d) = \alpha(a, b) + \alpha(c, d).$$

✍ Levando em conta que os parênteses definem quais são as primeiras operações a serem efetuadas, cabe ao leitor justificar passo a passo essa última demonstração.

Postulado V10 demanda a distributividade da adição de escalares relativamente à multiplicação de escalar por vetor. Notar que

$$(\alpha + \beta)(a, b) = ((\alpha + \beta)a, (\alpha + \beta)b) = (\alpha a + \beta a, \alpha b + \beta b) =$$
$$(\alpha a, \alpha b) + (\beta a, \beta b) = \alpha(a, b) + \beta(a, b).$$

Logo, V10 é teorema.

V11 afirma que $(\alpha\beta)u = \alpha(\beta u)$ deve ser teorema. Mas

$$(\alpha\beta)(a, b) = ((\alpha\beta)a, (\alpha\beta)b) = (\alpha(\beta a), \alpha(\beta b)) =$$
$$\alpha(\beta a, \beta b) = \alpha(\beta(a, b)).$$

Ou seja, a associatividade da multiplicação de reais garante que, de fato, V11 é teorema de ZF nesta interpretação de espaço vetorial real.

Finalmente, V12 exige que $1u = u$ seja teorema, para qualquer vetor u do espaço vetorial. Observar que

$$1(a, b) = (1(a), 1(b)) = (a, b).$$

Isso encerra a prova de que a interpretação dada para espaço vetorial real é de fato um modelo de espaço vetorial real. Este primeiro exemplo aqui discutido é conhecido como *espaço vetorial real* $\mathbb{R}^2$ *usual* ou, simplesmente, *espaço* $\mathbb{R}^2$ *usual*.

Interpretação que não é modelo

Seja $\langle \mathbb{R}^2, \mathbb{R}, +, \cdot, (0, -1) \rangle$ uma quíntupla ordenada tal que

$$+ : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}^2$$

é uma função na qual se abrevia $+((a, b), (c, d))$ como $(a, b) + (c, d)$ e

$$\cdot : \mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}^2$$

é uma função na qual se abrevia $\cdot(\alpha, (a, b))$ como $\alpha \cdot (a, b)$. Além disso, essas funções são definidas da seguinte maneira:

$$(a, b) + (c, d) = (a + c, b + d + 1)$$

e

$$\alpha \cdot (a, b) = (\alpha a, \alpha b).$$

✍ Neste caso, axioma V9 não é teorema para esta interpretação. Basta assumir, por exemplo, $\alpha = 3$. Além disso, V10 também não é teorema. Os demais postulados são teoremas. É obviamente recomendável que o leitor verifique isso. No entanto, basta um axioma não ser teorema para termos um exemplo de interpretação de espaço vetorial real que não é modelo de espaço vetorial real.

Espaço $M_{m\times n}$ usual

Nesta Subseção mostramos que matrizes reais também podem ser vetores. Antes, é necessário qualificar o que é uma matriz real.

Definição 8.2. *Sejam*

$$l_m = \{1, 2, 3, \cdots, m\}$$

e

$$c_n = \{1, 2, 3, \cdots, n\}.$$

Ou seja, l_m é o conjunto dos naturais i tais que $1 \leq i \leq m$, enquanto c_n é o conjunto dos naturais j tais que $1 \leq j \leq n$. Uma matriz real de m linhas e n colunas *é uma função*

$$a : l_m \times c_n \to \mathbb{R}.$$

Chamamos l_m *de conjunto* linhas *e* c_n *de conjunto* colunas*.*

✍ Ou seja, toda matriz real é uma restrição finita de uma função real com domínio $\omega \times \omega$. A recíproca desta última afirmação não é teorema.

O fato de produto cartesiano não ser comutativo é o que permite discernir linhas de colunas em uma matriz real.

Exemplo 8.1. *Se* $m = 3$ *e* $n = 2$*, então*

$$a = \{((1,1), a(1,1)), ((1,2), a(1,2)), ((2,1), a(2,1)),$$
$$((2,2), a(2,2)), ((3,1), a(3,1)), ((3,2), a(3,2))\}$$

é uma matriz real de três linhas e duas colunas, desde que as imagens $a(1,1)$*,* $a(1,2)$*,* $a(2,1)$*,* $a(2,2)$*,* $a(3,1)$ *e* $a(3,2)$ *sejam números reais.*

Neste exemplo a *é uma função com domínio* $\{1,2,3\} \times \{1,2\}$*.*

A notação usual para uma matriz real a é uma disposição retangular em linhas e colunas envolvidas por um par de parênteses, de modo que cada imagem $a(i,j)$ é denotada por a_{ij}. No caso do último exemplo dado acima, a matriz a é escrita simplesmente como

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix},$$

sendo $a_{ij} = a(i,j)$ para $1 \leq i \leq 3$ e $1 \leq j \leq 2$.

Teorema 8.1. *Sejam* $a : l_m \times c_n \to \mathbb{R}$ *e* $b : l_p \times c_q \to \mathbb{R}$ *matrizes reais. Logo,* $a = b$ *sss* $m = p$*,* $n = q$ *e* $a_{ij} = b_{ij}$ *para quaisquer* i *e* j *tais que* $1 \leq i \leq m$ *e* $1 \leq j \leq n$*.*

Demonstração: Matrizes reais, de acordo com Definição 8.2, são funções. Ademais, funções são casos particulares de conjuntos. Portanto, basta empregar o Axioma da Extensionalidade e Teorema 3.1.

Uma vez que produto cartesiano não é comutativo, se $m \neq n$, então a matriz $a : l_m \times c_n \to \mathbb{R}$ é diferente da matriz $b : c_n \times l_m \to \mathbb{R}$. Com efeito, a matriz b tem n linhas e m colunas.

Agora consideremos a seguinte interpretação para espaço vetorial real.

$$\mathcal{M}_{m\times n} = \langle M_{m\times n}, \mathbb{R}, +, \cdot, \bigcirc \rangle,$$

onde

- $M_{m\times n}$ é o conjunto de todas as matrizes reais com m linhas e n colunas;
- $+ : M_{m\times n} \times M_{m\times n} \to M_{m\times n}$ é uma função dada por
$$+(a, b)_{ij} = a_{ij} + b_{ij};$$
- $\cdot : \mathbb{R} \times M_{m\times n} \to M_{m\times n}$ é uma função dada por
$$\cdot(\alpha, a)_{ij} = \alpha a_{ij};$$
- $\bigcirc$ é a matriz real de m linhas e n colunas tal que $\bigcirc_{ij} = 0$ para todo i e para todo j tais que $1 \leq i \leq m$ e $1 \leq j \leq n$.

✍ O conjunto $M_{m\times n}$ pode ser facilmente definido usando o Esquema de Separação de ZF. Recomendamos que o leitor faça isso.

A função $+$ é conhecida como a *adição usual de matrizes reais*. A função $\cdot$ é conhecida como a *multiplicação usual de um real α por uma matriz real a de m linhas e n colunas*. Finalmente, a matriz $\bigcirc$ é chamada de *matriz nula*, aquela cujas imagens são todas iguais ao zero real.

EXEMPLO 8.2. *Se $m = 3$ e $n = 2$, temos que*

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{pmatrix} = \begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} \\ a_{21}+b_{21} & a_{22}+b_{22} \\ a_{31}+b_{31} & a_{32}+b_{32} \end{pmatrix},$$

$$\alpha \cdot \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix} = \begin{pmatrix} \alpha a_{11} & \alpha a_{12} \\ \alpha a_{21} & \alpha a_{22} \\ \alpha a_{31} & \alpha a_{32} \end{pmatrix}$$

e

$$\bigcirc = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

✍ Para cada par de naturais m e n, tais que $m \geq 1$ e $n \geq 1$, a quíntupla ordenada

$$\mathcal{M}_{m\times n} = \langle M_{m\times n}, \mathbb{R}, +, \cdot, \bigcirc \rangle$$

é um espaço vetorial real conhecido como $M_{m\times n}$ *usual*. Para provar isso, basta demonstrar que cada axioma de espaço vetorial real é teorema para cada interpretação $\mathcal{M}_{m\times n}$. Por exemplo, axiomas V6 e V7 são consequências imediatas do fato de que a adição entre reais é comutativa e associativa. Com efeito, adição entre matrizes é definida a partir da adição de reais.

Logo, nesta Subseção exibimos uma infinidade de espaços vetoriais reais: um para cada par de valores m e n.

Espaços vetoriais reais de funções

Espaços vetorias reais de matrizes não são os únicos casos de espaços vetoriais reais de funções. Considere, por exemplo, a seguinte interpretação para espaço vetorial real.

$$\mathfrak{F} = \langle \mathcal{C}^0, \mathbb{R}, +, \cdot, \bigcirc \rangle,$$

onde

- $\mathcal{C}^0$ é o conjunto de todas as funções reais contínuas com domínio $\mathbb{R}$;
- $+ : \mathcal{C}^0 \times \mathcal{C}^0 \to \mathcal{C}^0$ é uma função dada por
$$+(f, g)(x) = f(x) + g(x);$$
- $\cdot : \mathbb{R} \times \mathcal{C}^0 \to \mathcal{C}^0$ é uma função dada por
$$\cdot(\alpha, f)(x) = \alpha f(x);$$
- $\bigcirc$ é a função real $\bigcirc : \mathbb{R} \to \mathbb{R}$ tal que
$$\bigcirc(x) = 0$$
para todo real x.

Uma vez que adição de funções reais contínuas é uma função real contínua e multiplicação de um real α por uma função real contínua é uma função contínua, então não ocorre qualquer inconsistência na definição da interpretação $\mathfrak{F}$ dada acima.

✍ Novamente temos um exemplo de espaço vetorial real. Basta verificar os axiomas, um a um. Neste caso os vetores são funções reais contínuas com domínio $\mathbb{R}$.

$\langle \mathcal{C}^0, \mathbb{R}, +, \cdot, \bigcirc \rangle$ é chamado de *espaço* $\mathcal{C}^0$ *usual*. Se trocarmos $\mathcal{C}^0$ por $\mathcal{C}^k$, temos o *espaço* $\mathcal{C}^k$ *usual*, onde $\mathcal{C}^k$ é o conjunto das funções reais k vezes diferenciáveis e com derivada de ordem k contínua.

Espaço $\mathbb{R}^n$ usual

Seja $\langle \mathbb{R}^n, \mathbb{R}, +, \cdot, \bigcirc \rangle$ uma quíntupla ordenada tal que

$$+ : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$$

é uma função na qual se abrevia $+((a_1, a_2, \cdots, a_n), (b_1, b_2, \cdots, b_n))$ como

$$(a_1, a_2, \cdots, a_n) + (b_1, b_2, \cdots, b_n)$$

e

$$\cdot : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$$

é uma função na qual se abrevia $\cdot(\alpha, (a_1, a_2, \cdots, a_n))$ como

$$\alpha \cdot (a_1, a_2, \cdots, a_n).$$

Além disso, essas funções são definidas da seguinte maneira:

$$(a_1, a_2, \cdots, a_n) + (b_1, b_2, \cdots, b_n) = (a_1 + b_1, a_2 + b_2, \cdots, a_n + b_n)$$

e

$$\alpha \cdot (a_1, a_2, \cdots, a_n) = (\alpha a_1, \alpha a_2, \cdots, \alpha a_n).$$

✍ Se assumirmos também que $\bigcirc$ é a n-upla ordenada cujas entradas são todas iguais a 0, temos que

$$\langle \mathbb{R}^n, \mathbb{R}, +, \cdot, \bigcirc \rangle$$

é mais um exemplo de espaço vetorial real, conhecido como $\mathbb{R}^n$ *usual*. A prova desse resultado é análoga àquela feita para $\mathbb{R}^2$ usual.

Observar que o caso particular em que $n = 1$ implica que

$$\langle \mathbb{R}, \mathbb{R}, +, \cdot, 0 \rangle$$

é um espaço vetorial real, onde $+$ e $\cdot$ são as operações de adição e multiplicação entre reais, respectivamente.

Naturalmente, o caso em que $n = 2$ corresponde a $\mathbb{R}^2$ usual.

O caso em que $n = 3$ é usado em uma aplicação de espaços vetoriais reais em mecânica de partículas, na Seção 110.

Espaço $\mathcal{C}^\infty$ usual

Seja

$$\mathcal{R} = \langle \mathcal{C}^\infty, \mathbb{R}, +, \cdot, \bigcirc \rangle$$

uma interpretação de espaço vetorial real, onde

I: $\mathcal{C}^\infty$ é o conjunto de todas as funções reais com domínio $\mathbb{R}$ que admitem derivada de qualquer ordem;

II: $+ : \mathcal{C}^\infty \times \mathcal{C}^\infty \to \mathcal{C}^\infty$ é uma função tal que

$$+(f, g)(x) = f(x) + g(x);$$

III: $\cdot : \mathbb{R} \times \mathcal{C}^\infty \to \mathcal{C}^\infty$ é uma função tal que

$$\cdot(\alpha, f)(x) = \alpha f(x);$$

IV: $\bigcirc : \mathbb{R} \to \mathbb{R}$ é a função identicamente nula, ou seja, $\bigcirc(x) = 0$ para todo x real.

✍ Neste caso, $\mathcal{R}$ é um espaço vetorial real, conhecido como o *espaço $\mathcal{C}^\infty$ usual*.

Exemplo 8.3. ✍ *Se $p : \mathbb{R} \to \mathbb{R}$ é uma função polinomial, então p é um vetor de $\mathcal{C}^\infty$ usual.*

✍ *As funções seno, co-seno e exponencial também são vetores de $\mathcal{C}^\infty$ usual.*

A função $f : \mathbb{R} \to \mathbb{R}$ dada por

$$f(x) = |x|$$

não é um vetor de $\mathcal{C}^\infty$ usual. Com efeito, basta ver o Exemplo 5.26 *que segue a demonstração do Teorema 5.25.*

Resumindo

Dessa maneira fica claro que existem espaços vetoriais reais tais que seus vetores podem ser números reais, n-uplas ordenadas de números reais, matrizes reais, funções reais contínuas e funções reais diferenciáveis um número arbitrário de vezes. De maneira análoga é possível definir espaços vetoriais reais de funções reais quaisquer (desde que compartilhem o mesmo domínio), bem como funções reais que admitem derivada de ordem n.

Seção 82

Teoremas básicos sobre espaços vetoriais reais

Sumário
Índice
Rede

Todos os espaços vetoriais reais – sejam aqueles cujos vetores são matrizes, números reais, n-uplas ordenadas de reais ou funções reais – compartilham certas propriedades algébricas em comum.

Teorema 8.2. *Se* $\langle V, \mathbb{R}, +, \cdot, \overline{0}\,\rangle$ *é um espaço vetorial real, então o vetor nulo* $\overline{0}$ *é o único nulo aditivo em* V*.*

Demonstração: Supor que existe vetor nulo $\overline{0}'$ diferente de $\overline{0}$ em V. Logo,

$$\overline{0} + \overline{0}' = \overline{0}$$

e

$$\overline{0} + \overline{0}' = \overline{0}'$$

(lembrar que adição de vetores é comutativa). Portanto, a transitividade da igualdade implica que $\overline{0} = \overline{0}'$. $\perp$

Exemplo 8.4. *Em* $\mathbb{R}^2$ *usual,* $(0, 0)$ *é o único vetor nulo aditivo.*

Teorema 8.3. *Para qualquer vetor* v *de um espaço vetorial real*

$$\langle V, \mathbb{R}, +, \cdot, \overline{0}\,\rangle$$

temos que

$$v + v = v \Rightarrow v = \overline{0}.$$

Demonstração: Consequência do teorema anterior.

Teorema 8.4. *Para qualquer vetor* v *de um espaço vetorial real*

$$\langle V, \mathbb{R}, +, \cdot, \overline{0}\,\rangle$$

temos que

$$0 \cdot v = \overline{0}.$$

Demonstração: Observar que

$$0 \cdot v = (0+0) \cdot v.$$

Mas, de acordo com axioma V10,

$$(0+0) \cdot v = 0 \cdot v + 0 \cdot v.$$

A transitividade da igualdade garante, portanto, que

$$0 \cdot v = 0 \cdot v + 0 \cdot v.$$

Logo, Teorema 8.3 implica que $0 \cdot v = \overline{0}$.

Exemplo 8.5. *Num espaço vetorial real qualquer de matrizes reais de duas linhas e duas colunas, se*

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

é um vetor, então

$$0 \cdot \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Teorema 8.5. *Se v é vetor de $\langle V, \mathbb{R}, +, \cdot, \overline{0} \rangle$, temos que*

$$(-1) \cdot v = -v.$$

Demonstração: Sabemos que

$$0 \cdot v = (1 + (-1)) \cdot v.$$

Logo, Teorema 8.4 implica que

$$(1 + (-1)) \cdot v = \overline{0}.$$

Portanto, de acordo com axioma V10,

$$1 \cdot v + (-1) \cdot v = \overline{0}.$$

Logo, de acordo com axioma V12,

$$v + (-1) \cdot v = \overline{0}.$$

Finalmente, postulado V8 garante que

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

sendo $-v$ o simétrico aditivo (relativamente à adição de vetores) de v.

Exemplo 8.6. *Num espaço vetorial real*

$$\mathfrak{F} = \langle \mathcal{C}^0, \mathbb{R}, +, \cdot, \bigcirc \rangle$$

de funções reais contínuas, se $f : \mathbb{R} \to \mathbb{R}$ *é um vetor dado por* $f(x) = -\cos(x)$*, então* $(-1) \cdot f$ *é um vetor* $g : \mathbb{R} \to \mathbb{R}$ *tal que* $g(x) = \cos(x)$.

Teorema 8.6. *Para qualquer escalar* α *de um espaço vetorial real*

$$\langle V, \mathbb{R}, +, \cdot, \overline{0}\, \rangle$$

temos que $\alpha \cdot \overline{0} = \overline{0}$.

Demonstração: Temos que

$$\alpha \cdot \overline{0} = \alpha \cdot (\overline{0} + \overline{0}) = \alpha \cdot \overline{0} + \alpha \cdot \overline{0}.$$

Logo, $\alpha \cdot \overline{0} = \overline{0}$.

Exemplo 8.7. *Num espaço vetorial real qualquer de matrizes reais de duas linhas e duas colunas, se* α *é um escalar real e*

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

é o vetor nulo, então

$$\alpha \cdot \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Seção 83

Subespaços

Sumário
Índice
Rede

O conceito de *subespaço* é útil para criar exemplos de espaços vetoriais reais sem avaliar todos os postulados V1∼V12 da Definição 8.1, Seção 80.

Definição 8.3. *Seja* $\mathcal{V} = \langle V, \mathbb{R}, +, \cdot, \overline{0}\, \rangle$ *um espaço vetorial real. Dizemos que* $\mathcal{W} = \langle W, \mathbb{R}, \oplus, \odot, \overline{0}\, \rangle$ *é um* subespaço *de* $\mathcal{V}$ *sss*

i: $W \subseteq V$*;*

ii: $\overline{0} \in W$;

iii: $(u \in W \wedge v \in W) \Rightarrow (u \oplus v \in W \wedge u \oplus v = u + v)$;

iv: $(u \in W \wedge \alpha \in \mathbb{R}) \Rightarrow (\alpha \odot u \in W \wedge \alpha \odot u = \alpha \cdot u)$.

Um subespaço de um espaço vetorial real $\langle V, \mathbb{R}, +, \cdot, \overline{0}\,\rangle$ é definido por um subconjunto W do espaço de vetores V e restrições $\oplus$ e $\odot$ de $+$ e $\cdot$, respectivamente aos domínios $W \times W$ e $\mathbb{R} \times W$ (sobre restrição de função, ver Definição 4.12). Além disso, $\oplus$ e $\odot$ são fechadas em W (vetor de W $\oplus$ vetor de W é um vetor de W, e $\alpha \odot v$ é um vetor de W se v é vetor de W) e o vetor nulo $\overline{0}$ pertence a W.

O motivo para a mudança de notação de $+$ para $\oplus$, e $\cdot$ para $\odot$, é claro: restrições de uma dada função podem ter domínios diferentes; logo, se for o caso, são funções diferentes de acordo com o Axioma da Extensionalidade de ZF.

Teorema 8.7. *Se $\mathcal{V} = \langle V, \mathbb{R}, +, \cdot, \overline{0}\,\rangle$ é um espaço vetorial real, então $\mathcal{V}$ é subespaço de si mesmo.*

Demonstração: Lembrando que todo conjunto é subconjunto dele mesmo e, consequentemente, toda função é restrição dela mesma, a prova é imediata.

Teorema 8.8. *$\mathcal{W} = \langle \{\overline{0}\}, \mathbb{R}, \oplus, \odot, \overline{0}\,\rangle$ é subespaço do espaço vetorial real $\mathcal{V} = \langle V, \mathbb{R}, +, \cdot, \overline{0}\,\rangle$, se $\oplus$ é restrição de $+$ ao conjunto $\{\overline{0}\} \times \{\overline{0}\}$, e $\odot$ é restrição de $\cdot$ ao conjunto $\mathbb{R} \times \{\overline{0}\}$.*

Demonstração: Com efeito,

i: $\{\overline{0}\} \subset V$,

ii: $\overline{0} \in \{\overline{0}\}$,

iii: $\overline{0} \oplus \overline{0} = \overline{0}$ e,

iv: para qualquer real α, $\alpha \odot \overline{0} = \overline{0}$ (Teorema 8.4).

Os subespaços $\mathcal{V}$ e $\mathcal{W}$, mencionados nos dois últimos teoremas, são chamados de *subespaços triviais* de $\mathcal{V}$. Neste sentido, todo espaço vetorial real admite pelo menos um subespaço.

Os espaços vetoriais reais que admitem um único subespaço são aqueles nos quais o conjunto de vetores conta com um único elemento, a saber, o vetor nulo. Portanto, qualquer espaço vetorial real com pelo menos um vetor a mais, além do vetor nulo, admite pelo menos dois subespaços. Mas os subespaços relevantes são os não triviais, como ilustrado a seguir.

Exemplo 8.8. *Seja*

$$\mathcal{V} = \langle \mathbb{R}^2, \mathbb{R}, +, \cdot, (0,0) \rangle$$

o espaço vetorial real $\mathbb{R}^2$ *usual (conforme Seção 81).*

Seja

$$\mathcal{R} = \langle r, \mathbb{R}, \oplus, \odot, (0,0) \rangle$$

definido como se segue.

i: $r = \{(x,y) \in \mathbb{R}^2 \mid ax + by = 0\}$*, onde* a *e* b *são reais não simultaneamente nulos;*

ii: $\oplus : r \times r \to \mathbb{R}^2$ *e* $\odot : \mathbb{R} \times r \to \mathbb{R}^2$ *são restrições de* $+$ *e* $\cdot$*, respectivamente.*

Logo, $\mathcal{R}$ *é subespaço não trivial de* $\mathcal{V}$*.*

Com efeito, sejam (m,n) *e* (p,q) *elementos de* r*. Logo,*

$$am + bn = 0 \ \text{ e } \ ap + bq = 0,$$

o que implica em

$$am + ap + bn + bq = 0.$$

Logo,

$$a(m+p) + b(n+q) = 0,$$

o que implica que $(m+p, n+q)$ *também pertence a* r*.*

Mas $(m+p, n+q)$ *é igual a* $(m,n) + (p,q)$ *que, por sua vez, é igual a* $(m,n) \oplus (p,q)$ *(afinal,* $\oplus$ *é uma restrição de* $+$*). Isso significa que* $\oplus$ *é fechada em* r*.*

Analogamente, se (m,n) *pertence a* r*, então* $\alpha \odot (m,n)$ *também pertence a* r*. Com efeito,*

$$\alpha \odot (m,n) = \alpha \cdot (m,n) = (\alpha m, \alpha n).$$

Mas, se (m,n) *pertence a* r*, então* $am + bn = 0$*, o que implica que* $\alpha(am + bn) = 0$*; isso, por sua vez, implica que*

$$a\alpha m + b\alpha n = 0.$$

Portanto, $(\alpha m, \alpha n) \in r$*. Além disso,* $(0,0) \in r$*, pois* $a(0) + b(0) = 0$*, para quaisquer reais* a *e* b*.*

✍ Cada conjunto r do último Exemplo (definido pela escolha dos valores reais a e b) corresponde a uma *reta que passa pela origem* $(0,0)$, conforme Seção 77. Logo, o que provamos acima é que, qualquer reta que passa pela origem de $\mathbb{R}^2$ define um subconjunto de vetores de $\mathbb{R}^2$ que, por sua vez, define um subespaço do espaço vetorial real $\mathbb{R}^2$ usual. A versão resumida e informal da demonstração feita no último Exemplo segue no próximo parágrafo.

Por um lado, o vetor nulo $(0,0)$ pertence a qualquer reta que passa pela origem. Por outro, se r é uma reta que passa pela origem, a adição de pontos quaisquer da reta r resulta em um ponto na reta r. Além disso, a multiplicação de qualquer real α por um ponto de r resulta num ponto de r.

Em outras palavras, $\mathbb{R}^2$ usual admite uma infinidade de subespaços não triviais, um para cada escolha de reais a e b, desde que não sejam ambos nulos.

Os conjuntos s do próximo Exemplo são *retas que não passam pela origem* $(0,0)$. Logo, o que provamos a seguir é que nenhuma reta que não passa pela origem de $\mathbb{R}^2$ define subespaço deste espaço vetorial real.

Exemplo 8.9. *Seja*

$$\mathcal{V} = \langle \mathbb{R}^2, \mathbb{R}, +, \cdot, (0,0) \rangle$$

o espaço vetorial real $\mathbb{R}^2$ *usual (conforme Seção 81).*

Seja

$$\mathcal{R} = \langle s, \mathbb{R}, \oplus, \odot, (0,0) \rangle$$

definido como se segue.

i: $s = \{(x,y) \in \mathbb{R}^2 \mid ax + by = c\}$, *sendo* $c \neq 0$ *e* $a \neq 0 \vee b \neq 0$*;*

ii: $\oplus : s \times s \to \mathbb{R}^2$ *e* $\odot : \mathbb{R} \times s \to \mathbb{R}^2$ *são restrições de* $+$ *e* $\cdot$*, respectivamente.*

Logo, $\mathcal{R}$ *não é subespaço de* $\mathcal{V}$*. Com efeito, basta verificar que* $(0,0)$ *não pertence a* s*. Afinal,*

$$a(0) + b(0) = 0,$$

para quaisquer reais a *e* b*.*

O próximo Exemplo ilustra como certos espaços vetoriais reais de funções admitem infinitos subespaços.

EXEMPLO 8.10. *Seja*

$$\mathfrak{F} = \langle \mathcal{C}^0, \mathbb{R}, +, \cdot, \bigcirc \rangle$$

o espaço vetorial real de funções contínuas, como discutido na Seção 81.

Seja $\mathcal{C}^1$ *o conjunto de funções reais diferenciáveis com domínio* $\mathbb{R}$ *e cujas derivadas são contínuas.*

Teorema 5.25 garante que

$$\mathcal{C}^1 \subset \mathcal{C}^0,$$

i.e.,toda função diferenciável é contínua.

Teorema 5.21 garante que, se $f \in \mathcal{C}^1$ *e* $g \in \mathcal{C}^1$*, então*

$$f + g \in \mathcal{C}^1$$

(derivada da soma é a soma das derivadas).

Teorema 5.20 garante que, se $f \in \mathcal{C}^1$*, então*

$$\alpha \cdot f \in \mathcal{C}^1$$

(derivada de constante vezes função é constante vezes a derivada da função).

Logo, se

$$\mathfrak{G} = \langle \mathcal{C}^1, \mathbb{R}, \oplus, \odot, \bigcirc \rangle,$$

onde $\oplus$ *é uma restrição de* $+$ *ao domínio* $\mathcal{C}^1 \times \mathcal{C}^1$*, e* $\odot$ *é uma restrição de* $\cdot$ *ao domínio* $\mathbb{R} \times \mathcal{C}^1$*, então* $\mathfrak{G}$ *é subespaço não trivial de* $\mathfrak{F}$*.*

Ademais, se substituirmos, na discussão acima, $\mathcal{C}^1$ *por* $\mathcal{C}^2$ *(conjunto de funções reais diferenciáveis duas vezes, com domínio* $\mathbb{R}$*, cujas derivadas segundas são contínuas), então temos novo subespaço não trivial de* $\mathfrak{F}$*.*

Discussão análoga vale para $\mathcal{C}^k$*, o conjunto de funções reais* k *vezes diferenciáveis, com domínio* $\mathbb{R}$*, cujas derivadas de ordem* k *são contínuas.*

No último EXEMPLO mostramos que funções reais continuamente diferenciáveis (aquelas cujas derivadas são contínuas) definem um subespaço do espaço vetorial real usual de funções reais contínuas com domínio $\mathbb{R}$. Também indicamos que funções continuamente diferenciáveis k vezes igualmente definem subespaços do mesmo espaço vetorial real.

✍ Notar que

$$\mathcal{C}^\infty \subset \mathcal{C}^{k+1} \subset \mathcal{C}^k \subset \mathcal{C}^1 \subset \mathcal{C}^0,$$

para qualquer k natural maior do que 1, sendo que $f \in \mathcal{C}^\infty$ sss f admite derivada de qualquer ordem.

Em outras palavras, existem, por exemplo, funções reais três vezes diferenciáveis, mas não quatro.

EXEMPLO 8.11. *Seja $f : \mathbb{R} \to \mathbb{R}$ uma função dada por*

$$f(x) = \sqrt[3]{x^{11}}.$$

Por conta da Definição de Wallis, $\sqrt[3]{x^{11}} = x^{\frac{11}{3}}$. Por conta do Teorema 5.31,

$$f'(x) = \frac{11}{3} x^{\frac{8}{3}}.$$

Aplicando novamente Teorema 5.31, temos

$$f''(x) = \frac{88}{9} x^{\frac{5}{3}} \quad e \quad f'''(x) = \frac{440}{27} x^{\frac{2}{3}}.$$

Se o domínio de f fosse $\mathbb{R} - \{0\}$, a derivada quarta seria

$$f^{(4)}(x) = \frac{880}{81} x^{\frac{-1}{3}},$$

que é simplesmente

$$f^{(4)}(x) = \frac{880}{81} \frac{1}{x^{\frac{1}{3}}} = \frac{880}{81} \frac{1}{\sqrt[3]{x}}.$$

No entanto, f admite o zero real como um dos elementos de seu domínio. Lembrando que zero é o único real sem simétrico multiplicativo, obviamente não pode existir derivada quarta de f no ponto zero.

Resumidamente, f é uma função que admite derivada terceira mas não derivada quarta. Neste caso,

$$f \in \mathcal{C}^3$$

mas

$$f \notin \mathcal{C}^4,$$

o que implica que

$$\mathcal{C}^4 \subset \mathcal{C}^3,$$

ou seja, $\mathcal{C}^4$ é subconjunto próprio de $\mathcal{C}^3$.

Teorema 8.9. *Todo subespaço de um espaço vetorial real é um espaço vetorial real.*

Demonstração: Seja

$$\mathcal{W} = \langle W, \mathbb{R}, \oplus, \odot, \overline{0} \rangle$$

subespaço do espaço vetorial real

$$\mathcal{V} = \langle V, \mathbb{R}, +, \cdot, \overline{0} \rangle.$$

Como $\overline{0} \in W$, então axioma V1 é teorema na interpretação $\mathcal{W}$.

Itens III e IV da Definição 8.3 garantem que axiomas V2 e V3 são teoremas na mesma interpretação.

Axioma V4 é trivialmente verificado.

Finalmente, axiomas V5∼V12 são teoremas na interpretação $\mathcal{W}$ por conta do fato de que $\oplus$ e $\odot$ são restrições de $+$ e $\cdot$, respectivamente.

Uma vez conhecido um espaço vetorial real

$$\mathcal{V} = \langle V, \mathbb{R}, +, \cdot, \overline{0} \rangle,$$

fica mais fácil definir se

$$\mathcal{W} = \langle W, \mathbb{R}, \oplus, \odot, \overline{0} \rangle$$

é espaço vetorial real. Basta que W seja subconjunto de V, o vetor nulo pertença a W e as restrições $\oplus$ e $\odot$ sejam fechadas em W.

Exemplo 8.12. *Seja*

$$\mathcal{V} = \langle \mathbb{R}^2, \mathbb{R}, +, \cdot, (0,0) \rangle$$

o espaço vetorial real $\mathbb{R}^2$ *usual (conforme Seção 81).*

Seja

$$\mathcal{R} = \langle r, \mathbb{R}, \oplus, \odot, (0,0) \rangle$$

definido como se segue.

I: $r = \{(x,y) \in \mathbb{R}^2 \mid ax + by = 0\}$, *onde* $a \neq 0 \vee b \neq 0$*;*

II: $\oplus : r \times r \to \mathbb{R}^2$ *e* $\odot : \mathbb{R} \times r \to \mathbb{R}^2$ *são restrições de* $+$ *e* $\cdot$*, respectivamente.*

Logo, $\mathcal{R}$ *é um espaço vetorial real. Ver* Exemplo 8.8.

Exemplo 8.13. *Seja*

$$\mathcal{M}_{2\times 2} = \langle M_{2\times 2}, \mathbb{R}, +, \cdot, \bigcirc \rangle$$

o espaço vetorial real das matrizes reais de duas linhas e duas colunas.

Seja

$$\mathcal{N}_{2\times 2} = \langle N_{2\times 2}, \mathbb{R}, \oplus, \odot, \bigcirc \rangle$$

definido como se segue:

I: $\mathcal{N}_{2\times 2}$ *é subconjunto de* $\mathcal{M}_{2\times 2}$ *tal que, uma matriz* b *pertence a* $\mathcal{N}_{2\times 2}$ *sss* $b_{11} = 0$ *e* $b_{12} = 0$. *Ou seja,*

$$b = \begin{pmatrix} 0 & 0 \\ b_{21} & b_{22} \end{pmatrix},$$

onde b_{21} *e* b_{22} *são números reais quaisquer.*

II: $\oplus$ *é restrição de* $+$ *ao domínio* $\mathcal{N}_{2\times 2} \times \mathcal{N}_{2\times 2}$; *e* $\odot$ *é restrição de* $\cdot$ *ao domínio* $\mathbb{R} \times \mathcal{N}_{2\times 2}$.

✍ *Logo,* $\mathcal{N}_{2\times 2}$ *é um espaço vetorial real, uma vez que se trata de subespaço de* $\mathcal{M}_{2\times 2}$. *Cabe ao leitor fazer a demonstração.*

Seção 84

Dependência e independência linear

Sumário
Índice
Rede

Certos espaços vetoriais reais podem ser univocamente determinados por uma quantia finita de vetores, ainda que exista uma quantia não finita de vetores no mesmo espaço. Estudamos isso nesta e na próxima Seção.

Definição 8.4. *Sejam* $\mathcal{V} = \langle V, \mathbb{R}, +, \cdot, \overline{0} \rangle$ *um espaço vetorial real e*

$$\{v_1, v_2, \cdots, v_n\}$$

um conjunto de n *vetores de* V. *Dizemos que um vetor* v *pertencente a* V *é uma* combinação linear *dos vetores de* $\{v_1, v_2, \cdots, v_n\}$ *sss existem escalares reais* $\alpha_1, \alpha_2, \cdots, \alpha_n$ *tais que*

$$v = \sum_{i=1}^{n} \alpha_i v_i.$$

Cabe ao leitor recordar que já lidamos com combinações lineares de funções anteriormente, no Teorema 6.25, na discussão sobre operadores diferenciais (Seção 53) e na discussão sobre soluções da equação diferencial $y'' + y = 0$ (final da Seção 54). Ou seja, o que fizemos anteriormente foi uma preparação para o estudo de álgebra linear.

Exemplo 8.14. *Seja*

$$x = \{(1,0,0),(0,1,0),(0,0,1),(1,1,1)\}$$

um conjunto de vetores de $\mathbb{R}^3$ *usual (caso particular de* $\mathbb{R}^n$ *usual, conforme Seção 81). Neste caso, o vetor* $(\pi, \sqrt{2}, -7)$ *é uma combinação linear dos vetores de* x*. Com efeito,*

$$(\pi, \sqrt{2}, -7) = \pi(1,0,0) + \sqrt{2}(0,1,0) + (-7)(0,0,1) + 0(1,1,1).$$

Observar, porém, que esta não é a única possível combinação linear dos vetores de x *para obter* $(\pi, \sqrt{2}, -7)$*. Afinal, podemos ter também a seguinte combinação linear:*

$$(\pi, \sqrt{2}, -7) =$$

$$0(1,0,0) + (\sqrt{2} - \pi)(0,1,0) + (-7-\pi)(0,0,1) + \pi(1,1,1).$$

Além disso, qualquer vetor (a,b,c) *de* $\mathbb{R}^3$ *usual é combinação linear dos vetores de* x*:*

$$(a,b,c) = a(1,0,0) + b(0,1,0) + c(0,0,1) + 0(1,1,1).$$

Exemplo 8.15. *Seja* $y = \{(1,0,0),(0,1,0)\}$ *um conjunto de vetores de* $\mathbb{R}^3$ *usual. Logo, o vetor* $(2,2,2)$ *não é uma combinação linear dos vetores de* y*. Com efeito, se*

$$(2,2,2) = \alpha(1,0,0) + \beta(0,1,0),$$

então

$$(2,2,2) = (\alpha, \beta, 0),$$

uma contradição.

No entanto, qualquer vetor $(a,b,0)$ *deste espaço é combinação linear dos vetores de* y*. Com efeito,*

$$(a,b,0) = a(1,0,0) + b(0,1,0).$$

✍ No Exemplo 8.14 $(1,1,1)$ é combinação linear dos demais vetores de x. Cada vetor de x é combinação linear dos demais.

Definição 8.5. *Sejam* $\mathcal{V} = \langle V, \mathbb{R}, +, \cdot, \overline{0}\,\rangle$ *um espaço vetorial real e*

$$\{v_1, v_2, \cdots, v_n\}$$

um conjunto de n *vetores pertencentes a* V*.*

Dizemos que x *é* linearmente independente *sss nenhum vetor de* x *é uma combinação linear dos demais. Podemos escrever isso como '*x *é L.I.', onde L.I. abrevia 'linearmente independente'. Caso contrário,* x *é* linearmente dependente *e escrevemos isso como '*x *é L.D.', onde L.D. abrevia 'linearmente dependente'.*

Exemplo 8.16. *Seja* $y = \{(1,0,0), (0,1,0)\}$ *um conjunto de vetores de* $\mathbb{R}^3$ *usual. Logo,* y *é L.I. Com efeito, se*

$$(1,0,0) = \alpha(0,1,0),$$

então

$$(1,0,0) = (0,\alpha,0),$$

uma contradição; além disso, se

$$(0,1,0) = \beta(1,0,0),$$

então

$$(0,1,0) = (\beta,0,0),$$

uma contradição. Portanto, nenhum dos dois vetores de y *é combinação linear do único vetor que resta em* y*.*

Exemplo 8.17. *Seja* $\mathcal{V} = \langle S, \mathbb{R}, \oplus, \odot, \bigcirc\rangle$*, onde*

- I: $S = \{y \in \mathcal{C}^\infty \mid y'' + y = 0\}$,
- II: $\oplus$ *é restrição de* $+$ *(no espaço* $\mathcal{C}^\infty$ *usual) a* S*,*
- III: $\odot$ *é restrição de* $\cdot$ *(no espaço* $\mathcal{C}^\infty$ *usual) a* S*, e*
- IV: $\bigcirc$ *é a função real identicamente nula, com domínio* $\mathbb{R}$*.*

Se y_1 *e* y_2 *pertencem a* S*, então a combinação linear* $\alpha y_1 + \beta y_2$ *também pertence a* S *(✍ demonstração análoga à do Teorema 6.25). Ademais,* $\bigcirc \in S$*. Logo* $\mathcal{V}$ *é subespaço de* $\mathcal{C}^\infty$ *usual, o que implica que* $\mathcal{V}$ *é espaço vetorial real.*

Lembrando que as funções seno e co-seno pertencem a S*, temos que* $\{\text{sen}, \cos\}$ *é L.I. neste espaço. Com efeito, não existe* α *tal que* $\text{sen} = \alpha \cos$*, nem* β *tal que* $\cos = \beta\text{sen}$*. Ver Seção 54.*

✍ É altamente recomendável que o leitor prove circunstancialmente todas as afirmações feitas no último EXEMPLO. As ideias usadas na demonstração do Teorema 6.25 podem ser facilmente adaptadas para provar que o EXEMPLO acima de fato descreve um subespaço não trivial de $\mathcal{C}^{\infty}$ usual.

Com relação à tese de que $\{\text{sen}, \cos\}$ é linearmente independente, basta observar o que se segue. Uma vez que $\text{sen}(0) = 0$ e $\cos(0) = 1$, não há real β tal que $\cos(0) = \beta\text{sen}(0)$. Logo, não há β tal que

$$\cos(x) = \beta\text{sen}(x)$$

para todo x real.

A prova de que não existe α de modo que $\text{sen}(x) = \alpha\cos(x)$ é análoga.

Definição 8.5 (sobre conjuntos de vetores linearmente independentes) não é uma ferramenta muito prática para efeitos de cálculos. Se um conjunto de vetores conta com n elementos, precisamos testar cada um deles para determinar se o conjunto é linearmente independente. O próximo teorema, porém, oferece um critério mais econômico para determinar se um conjunto finito de vetores é L.I.

TEOREMA 8.10. *Um conjunto de vetores*

$$x = \{v_1, v_2, \cdots, v_n\}$$

de um espaço vetorial real

$$\langle V, \mathbb{R}, +, \cdot, \overline{0}\,\rangle$$

é linearmente independente sss a equação

$$\sum_{i=1}^{n} \alpha_i \cdot v_i = \overline{0}$$

admite uma única solução, onde cada α_i é um escalar.

DEMONSTRAÇÃO: Obviamente a equação

$$\sum_{i=1}^{n} \alpha_i \cdot v_i = \overline{0}$$

sempre admite pelo menos uma solução. Basta fazer $\alpha_i = 0$ para todo i tal que $1 \leq i \leq n$. Teorema 8.4 e axioma V5 garantem isso. Esta é chamada de *solução trivial* da equação acima.

Supor que a mesma equação admite outra solução, além da trivial. Isso implica que existe pelo menos um α_j, diferente de 0, tal que a igualdade acima é teorema. Por conta da comutatividade e da associatividade da adição de vetores, podemos assumir que esse α_j é α_1, sem perda de generalidade. Logo, multiplicando ambos os lados da igualdade acima pelo simétrico multiplicativo de α_1, podemos reescrever a mesma equação como

$$v_1 + \sum_{i=2}^{n} (\alpha_i \alpha_1^{-1}) \cdot v_i = \overline{0},$$

uma vez que

$$\alpha_1 \alpha_1^{-1} = 1 \text{ e } 1 \cdot v_1 = v_1,$$

de acordo com axioma V12. Logo,

$$v_1 = \sum_{i=2}^{n} (-\alpha_i \alpha_1^{-1}) \cdot v_i,$$

por conta do Teorema 8.5. Isso prova que v_1 é combinação linear dos demais vetores de x. Portanto, acabamos de provar que, se a equação em questão admite outra solução além da trivial, então x é L.D.

Por outro lado, supor agora que x é L.D. Uma vez que o Axioma da Extensionalidade garante que a ordem dos vetores listados em x é irrelevante, podemos assumir, sem perda de generalidade, que v_1 é combinação linear dos demais vetores em x. Ou seja, existem $\beta_2, \beta_3, \cdots, \beta_n$ tais que

$$v_1 = \sum_{i=2}^{n} \beta_i \cdot v_i.$$

Mas isso implica que

$$v_1 + \sum_{i=2}^{n} (-\beta_i) \cdot v_i = \overline{0},$$

a qual é a mesma equação do enunciado do Teorema, onde $\alpha_1 = 1$ e os demais α_i são iguais a $-\beta_i$. Uma vez que $\alpha_1 \neq 0$, então a equação admite outra solução além da trivial.

Com este último passo, provamos que a equação em questão admite mais de uma solução sss x é L.D. Fórmulas 7 e 12 da lista de dezessete teoremas da Seção 10 concluem a prova.

Exemplo 8.18. *Em* $\mathcal{C}^\infty$ *usual o conjunto*

$$\{\text{sen}, \cos\}$$

é L.I., como já adiantado no Exemplo 8.17. *Mas agora usamos Teorema 8.10 para provar o mesmo resultado. Se*

$$\alpha\, sen(x) + \beta \cos(x) = 0,$$

para todo x *real (observar que o símbolo* 0 *à direita da igualdade corresponde à função identicamente nula), então* $\alpha = 0$ *e* $\beta = 0$*, de acordo com Seção 54.*

Exemplo 8.19. *Em* $\mathbb{R}^2$ *usual o conjunto*

$$\{(1,1),(2,2)\}$$

é L.D. Com efeito, se

$$\alpha \cdot (1,1) + \beta \cdot (2,2) = (0,0),$$

então existem outras soluções para essa equação, além da trivial $(\alpha = \beta = 0)$*. Basta assumir, por exemplo,* $\alpha = -2$ *e* $\beta = 1$*.*

Teorema 8.11. *Seja*

$$x = \{v_1, v_2, \cdots, v_n\}$$

um conjunto de vetores de um espaço vetorial real

$$\langle V, \mathbb{R}, +, \cdot, \overline{0}\,\rangle.$$

Se algum dos vetores de x *for* $\overline{0}$*, então* x *é L.D.*

Demonstração: Considere a equação

$$\sum_{i=1}^{n} \alpha_i \cdot v_i = \overline{0}.$$

Se existe j tal que $v_j = \overline{0}$, então, independentemente de qualquer valor real α_j, temos

$$\alpha_j \cdot v_j = \overline{0},$$

de acordo com Teorema 8.4. Logo, a equação acima admite infinitas soluções, além da trivial. Basta assumir qualquer $\alpha_j \neq 0$. Logo, Teorema 8.10 garante que x é L.D.

Naturalmente, a recíproca do último teorema não é teorema (Exemplo 8.19).

Seção 85

Espaços vetoriais reais de dimensão finita

Sumário
Índice
Rede

Qualificamos aqui o conceito de *base* para certos espaços vetoriais reais.

Definição 8.6. *Seja*

$$\mathcal{V} = \langle V, \mathbb{R}, +, \cdot, \overline{0} \rangle$$

um espaço vetorial real.

Um conjunto de vetores

$$x = \{v_1, v_2, \cdots, v_n\}$$

gera $\mathcal{V}$ *sss todo vetor* v *pertencente a* V *pode ser obtido por combinação linear (Definição 8.4) dos vetores de* x.

Exemplo 8.20. *Seja* $\mathcal{V} = \langle S, \mathbb{R}, \oplus, \odot, \bigcirc \rangle$, *onde*

I: $S = \{y \in \mathcal{C}^\infty \mid y'' + y = 0\}$,

II: $\oplus$ *é restrição de* $+$ *(no espaço* $\mathcal{C}^\infty$ *usual) a* S,

III: $\odot$ *é restrição de* $\cdot$ *(no espaço* $\mathcal{C}^\infty$ *usual) a* S, *e*

IV: $\bigcirc$ *é a função real identicamente nula, com domínio* $\mathbb{R}$.

O conjunto $\{\text{sen}, \cos\}$ *gera os vetores de* $\mathcal{V}$, *conforme final da Seção 54. Com efeito, qualquer função* y, *onde* $y'' + y = 0$, *é tal que*

$$y(x) = \alpha \cos(x) + \beta \,\text{sen}(x).$$

Exemplo 8.21. *Seja*

$$x = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}.$$

Logo, x *não gera* $M_{3\times 2}$ *usual. Com efeito, se*

$$\begin{pmatrix} a & b \\ c & d \\ e & f \end{pmatrix} = \alpha \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} + \beta \begin{pmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} + \gamma \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \end{pmatrix} + \delta \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix},$$

então
$$\begin{pmatrix} a & b \\ c & d \\ e & f \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ \gamma & 0 \\ 0 & \delta \end{pmatrix}.$$

Logo, nenhum vetor
$$\begin{pmatrix} a & b \\ c & d \\ e & f \end{pmatrix},$$
onde d ou e são diferentes de 0, é combinação linear dos vetores de x.

Exemplo 8.22. *O conjunto x do último* Exemplo *gera o subespaço de $M_{3\times 2}$ usual cujos vetores são*
$$\begin{pmatrix} a & b \\ c & 0 \\ 0 & f \end{pmatrix}.$$
Com efeito, basta fazer $\alpha = a$, $\beta = b$, $\gamma = c$ e $\delta = f$.

Exemplo 8.23. *O conjunto*
$$y = \left\{ \begin{pmatrix} 0 & 0 \\ 2 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ -\pi & 0 \\ 0 & 0 \end{pmatrix} \right\}$$
gera o subespaço de $M_{3\times 2}$ usual cujos vetores são
$$\begin{pmatrix} 0 & 0 \\ c & 0 \\ 0 & 0 \end{pmatrix}.$$
Com efeito, se
$$\begin{pmatrix} 0 & 0 \\ c & 0 \\ 0 & 0 \end{pmatrix} = \alpha \begin{pmatrix} 0 & 0 \\ 2 & 0 \\ 0 & 0 \end{pmatrix} + \beta \begin{pmatrix} 0 & 0 \\ -\pi & 0 \\ 0 & 0 \end{pmatrix},$$
basta fazer $\alpha = \frac{c}{2}$ e $\beta = 0$.

Outra opção é fazer $\alpha = 0$ e $\beta = \frac{-c}{\pi}$.

Definição 8.7. *Seja $\mathcal{V} = \langle V, \mathbb{R}, +, \cdot, \overline{0} \rangle$ um espaço vetorial real. O conjunto de vetores $\{v_1, v_2, \cdots, v_n\}$ é uma* base finita *(ou* base*) de $\mathcal{V}$ sss x gera $\mathcal{V}$ e x é linearmente independente.*

Exemplo 8.24. *O conjunto $d = \{(2,2),(2,1)\}$ é uma base de $\mathbb{R}^2$ usual. Provamos isso a seguir.*

Por um lado, se

$$\alpha(2,2) + \beta(2,1) = (0,0),$$

então

$$(2\alpha + 2\beta, 2\alpha + \beta) = (0,0),$$

o que implica em

$$2\alpha + 2\beta = 0 \;\; e \;\; 2\alpha + \beta = 0.$$

Portanto, $2\beta = \beta$, fórmula esta que é teorema apenas para $\beta = 0$. Mas, se $\beta = 0$, então $\alpha = 0$. Logo, a primeira equação acima admite apenas a solução trivial para garanti-la como teorema. Portanto, Teorema 8.10 garante que d é L.I.

Por outro lado, se (a,b) é um vetor qualquer de $\mathbb{R}^2$ usual e

$$(a,b) = \alpha(2,2) + \beta(2,1),$$

então

$$(a,b) = (2\alpha + 2\beta, 2\alpha + \beta).$$

Portanto, $2\alpha + 2\beta = a$ e $2\alpha + \beta = b$. Mas essas duas últimas fórmulas são teoremas somente se

$$\alpha = b - \frac{a}{2} \quad e \quad \beta = a - b.$$

Afinal, para quaisquer a e b reais, temos

$$(a,b) = \left(b - \frac{a}{2}\right)(2,2) + (b-a)(2,1).$$

Logo, d gera $\mathbb{R}^2$ usual.

Portanto, Definição 8.7 garante que d é uma base de $\mathbb{R}^2$ usual.

Exemplo 8.25. *O conjunto $c = \{(2,2),(2,1),(1,1)\}$ não é uma base de $\mathbb{R}^2$ usual. Por um lado, conforme último* Exemplo, *c gera $\mathbb{R}^2$, uma vez que*

$$(a,b) = \alpha(2,2) + \beta(2,1) + \gamma(1,1)$$

se

$$\alpha = b - \frac{a}{2}, \quad \beta = a - b \quad e \quad \gamma = 0.$$

Porém, c é L.D. Com efeito, $(2,2) = 0(2,1) + 2(1,1)$.

EXEMPLO 8.26. *O conjunto*

$$y = \{(1,0,0), (0,1,0)\}$$

do EXEMPLO 8.16 *não é uma base de* $\mathbb{R}^3$ *usual.*

Apesar de y *ser L.I., não gera* $\mathbb{R}^3$ *usual. Com efeito, se* $(a,b,c) \in \mathbb{R}^3$ *e*

$$(a,b,c) = \alpha(1,0,0) + \beta(0,1,0),$$

então

$$(a,b,c) = (\alpha, \beta, 0).$$

Logo, nenhum vetor (a,b,c) *deste espaço vetorial real, tal que* $c \neq 0$*, pode ser obtido por combinação linear dos vetores de* y*.*

Porém, y *é uma base do supespaço de* $\mathbb{R}^3$ *usual cujos vetores são da forma* $(a,b,0)$*.*

EXEMPLO 8.27. *O conjunto* x *do* EXEMPLO 8.21 *não é uma base de* $M_{3\times 2}$ *usual. Isso porque* x *não gera o espaço vetorial real* $M_{3\times 2}$ *usual.*

EXEMPLO 8.25 ilustra uma situação na qual um conjunto de vetores gera um espaço vetorial real, mas não é linearmente independente. EXEMPLO 8.26, não obstante, exibe uma situação na qual um conjunto de vetores é linearmente independente, mas não gera o espaço vetorial real dado. Portanto, as duas condições exigidas na Definição 8.7 (sobre base) são independentes entre si.

Observar também que um mesmo espaço vetorial real pode admitir mais de uma base.

EXEMPLO 8.28. *Ambos os conjuntos*

$$\{(2,2), (2,1)\} \ \ e \ \ \{(1,0), (0,1)\}$$

são bases de $\mathbb{R}^2$ *usual.*

Em contrapartida, qualquer espaço vetorial real cujo único vetor é o vetor nulo $\overline{0}$ não admite base alguma, por conta do Teorema 8.11.

Graças aos conceitos de base e bijeção (ver Definição 4.16), podemos agora introduzir *bases ordenadas*, bem como *coordenadas* de um vetor em certos espaços vetoriais reais. Como vemos a seguir, uma base ordenada é uma base finita munida de uma bijeção com um ordinal finito.

Definição 8.8. *Sejam* $\mathcal{V}$ *um espaço vetorial real e*

$$b = \{v_1, v_2, \cdots, v_n\}$$

uma base de $\mathcal{V}$. *Seja ainda* $f : n \to b$ *uma bijeção, onde* n *é um* *ordinal finito* *e* $f(i) = v_{i+1}$ *para cada* i *pertencente a* n. *Se, para qualquer vetor* v *de* $\mathcal{V}$ *tivermos*

$$v = \sum_{i=1}^{n} \alpha_i v_i,$$

os escalares α_i *definem as* coordenadas

$$(\alpha_1, \alpha_2, \cdots, \alpha_n)$$

de v *relativamente ao par ordenado* (b, f). *Chamamos* (b, f) *de* base ordenada *de* $\mathcal{V}$.

É uma prática comum escrever as coordenadas de um vetor, relativamente a uma base ordenada, na forma de n-uplas ordenadas

$$(\alpha_1, \alpha_2, \cdots, \alpha_n).$$

Isso porque os vetores de qualquer base são obviamente distintos entre si. Apesar disso, eventualmente as coordenadas de um vetor, relativamente a uma base ordenada, podem ser iguais entre si. Ou seja, a coordenada α_1 é o escalar que multiplica por v_1, a coordenada α_2 é o escalar que multiplica por v_2, e assim por diante. O que define a ordem na n-upla ordenada é a bijeção f da base ordenada (b, f).

Exemplo 8.29. *No* Exemplo 8.24 *foi provado que*

$$d = \{(2, 2), (2, 1)\}$$

é uma base de $\mathbb{R}^2$ *usual*.

Logo, as coordenadas do vetor $(\sqrt{2}, 4)$ *relativamente à base* d *são*

$$\left(\frac{8 - \sqrt{2}}{2}, \sqrt{2} - 4\right).$$

Com efeito,

$$(\sqrt{2}, 4) = \frac{8 - \sqrt{2}}{2}(2, 2) + (\sqrt{2} - 4)(2, 1).$$

Estamos assumindo que a bijeção f *da base ordenada* (d, f) *é dada por* $f : 2 \to d$, *onde* $f(0) = (2, 2)$ *e* $f(1) = (2, 1)$.

Se assumirmos uma base ordenada (d, g) *onde* $g : 2 \to d$ *é dada por* $g(0) = (2, 1)$ *e* $g(1) = (2, 2)$*, então as coordenadas do mesmo vetor são*

$$\left(\sqrt{2} - 4, \frac{8 - \sqrt{2}}{2}\right),$$

relativamente à nova base ordenada.

Em contrapartida, as coordenadas do mesmo vetor relativamente à base

$$\{(1, 0), (0, 1)\}$$

ordenada por $h : 2 \to \{(1, 0), (0, 1)\}$*, onde* $h(0) = (1, 0)$ *e* $h(1) = (0, 1)$*, são* $(\sqrt{2}, 4)$*. Afinal,*

$$(\sqrt{2}, 4) = \sqrt{2}(1, 0) + 4(0, 1).$$

Comumente a bijeção f de uma base ordenada (b, f) não é explicitada, uma vez que a própria ordem $(\alpha_1, \alpha_2, \cdots, \alpha_n)$ já deixa clara a bijeção f. Neste contexto, é usual a expressão 'coordenadas de um vetor relativamente a uma base finita' no lugar de 'coordenadas de um vetor relativamente a uma base finita ordenada'. Neste livro adotamos essa convenção, como ocorre no próximo teorema.

Teorema 8.12. *Se um espaço vetorial real* $\mathcal{V}$ *admite base*

$$b = \{v_1, v_2, \cdots, v_n\}$$

com n *elementos (onde* n *é um natural), então as coordenadas de qualquer vetor* v *de* $\mathcal{V}$ *são únicas.*

Demonstração: Sejam $(\alpha_1, \alpha_2, \cdots, \alpha_n)$ as coordenadas de um vetor v qualquer de $\mathcal{V}$ relativamente à base b. Logo,

$$v = \sum_{i=1}^{n} \alpha_i v_i.$$

Supor que v admite coordenadas $(\beta_1, \beta_2, \cdots, \beta_n)$ relativamente à mesma base b. Logo,

$$v = \sum_{i=1}^{n} \beta_i v_i.$$

Portanto,

$$\sum_{i=1}^{n} \alpha_i v_i - \sum_{i=1}^{n} \beta_i v_i = \overline{0}.$$

Mas isso implica que

$$\sum_{i=1}^{n}(\alpha_i - \beta_i)v_i = \overline{0}.$$

Lembrando que b é linearmente independente (de acordo com Definição 8.7), Teorema 8.10 garante que, para todo i tal que $1 \leq i \leq n$, temos $\alpha_i - \beta_i = 0$. Logo, $\alpha_i = \beta_i$. Portanto, as coordenadas são únicas.

Naturalmente, como já alertado, na prova do último teorema assumimos implicitamente uma base ordenada (b, f), onde $f : n \to b$ é uma bijeção dada por $f(i) = v_{i+1}$, para todo i pertencente a n.

Teorema 8.13. *Se um espaço vetorial real $\mathcal{V}$ admite base*

$$b = \{v_1, v_2, \cdots, v_n\}$$

com n elementos, então qualquer conjunto

$$c = \{w_1, w_2, \cdots, w_m\}$$

de m vetores de $\mathcal{V}$, onde $m > n$, é linearmente dependente.

Demonstração: Cada w_j de c é uma combinação linear de vetores de b, uma vez que b gera $\mathcal{V}$. Logo,

$$w_j = \sum_{i=1}^{n} \alpha_{ji} v_i,$$

para cada j tal que $1 \leq j \leq m$. Considere agora

$$\sum_{j=1}^{m} \beta_j w_j = \overline{0}.$$

Se provarmos que esta última admite mais de uma solução, além da trivial (i.e., existe pelo menos um β_j diferente de 0, tal que a equação acima é teorema), encerramos a prova.

De acordo com axioma V11 da Seção 80, temos

$$\left(\sum_{j=1}^{m} \beta_j \alpha_{j1}\right) v_1 + \left(\sum_{j=1}^{m} \beta_j \alpha_{j2}\right) v_2 + \cdots + \left(\sum_{j=1}^{m} \beta_j \alpha_{jn}\right) v_n = \overline{0}.$$

Mas esta última equação é teorema se cada somatório

$$\sum_{j=1}^{m} \beta_j \alpha_{ji}$$

que multiplica cada v_i for 0.

Logo, temos um sistema de n equações com m valores β_j a serem definidos para garantir que todas as n equações são teoremas. Lembrando que $m > n$, isso implica que existe pelo menos um β_j diferente de 0. Esta última afirmação pode ser demonstrada por indução ao longo de todos os i, de 1 a n. ✍ Recomendamos que o leitor conclua a prova.

Teorema 8.14. *Se o conjunto de vetores $\{v_1, v_2, \cdots, v_n\}$ gera um espaço vetorial real e o conjunto de vetores*

$$\{w_1, w_2, \cdots, w_m\}$$

do mesmo espaço é L.I., então $m \leq n$.

Demonstração: Notar que, na prova do Teorema 8.13, não foi necessário usar o fato de que $\{v_1, v_2, \cdots, v_n\}$ é L.I. Foi suficiente assumir que tal conjunto gera o espaço. Logo, o enunciado aqui colocado é consequência imediata da prova do Teorema 8.13.

Teorema 8.15. *Se um espaço vetorial real $\mathcal{V}$ admite base*

$$b = \{v_1, v_2, \cdots, v_n\}$$

com n elementos, então qualquer outra base de $\mathcal{V}$ tem n elementos.

Demonstração: Seja

$$c = \{w_1, w_2, \cdots, w_m\}$$

outra base de $\mathcal{V}$. Como c gera o espaço e b é L.I., então Teorema 8.14 garante que $n \leq m$.

Analogamente, como b gera o espaço e c é L.I., então $m \leq n$. Mas $n \leq m$ e $m \leq n$ sss $m = n$.

Agora que sabemos que o número n de elementos de uma base de um espaço vetorial real é invariante, caso exista base

$$b = \{v_1, v_2, \cdots, v_n\},$$

podemos finalmente introduzir nova definição, suportada pelo último teorema acima.

Definição 8.9. *Se um espaço vetorial real* $\mathcal{V}$ *admite base com* n *elementos, dizemos que* $\mathcal{V}$ *tem* n dimensões. *Escrevemos*

$$\dim(\mathcal{V}) = n.$$

Exemplo 8.30. Exemplo 8.24 *prova que* $\mathbb{R}^2$ *usual tem duas dimensões.*

Exemplo 8.31. Exemplo 8.26 *prova que*

$$y = \{(1,0,0),(0,1,0)\}$$

é uma base do supespaço de $\mathbb{R}^3$ *usual cujos vetores são da forma* $(a,b,0)$.

Logo, tal subespaço é um espaço vetorial real com duas dimensões.

Exemplo 8.32. *Seja* $\mathcal{V} = \langle S, \mathbb{R}, \oplus, \odot, \bigcirc \rangle$, *onde*

i: $S = \{y \in \mathcal{C}^\infty \mid y'' + y = 0\}$,

ii: $\oplus$ *é restrição de* $+$ *(no espaço* $\mathcal{C}^\infty$ *usual) a* S,

iii: $\odot$ *é restrição de* $\cdot$ *(no espaço* $\mathcal{C}^\infty$ *usual) a* S, *e*

iv: $\bigcirc$ *é a função real identicamente nula, com domínio* $\mathbb{R}$.

Neste caso, $\{\operatorname{sen}, \cos\}$ *é uma base de* $\mathcal{V}$, *de acordo com* Exemplos 8.17 *e* 8.20. *Logo* $\dim(\mathcal{V}) = 2$.

Exemplo 8.33. *O espaço vetorial real* $\langle \mathbb{R}, \mathbb{R}, +, \cdot, 0 \rangle$ *tem uma dimensão. Com efeito,* $\{\sqrt{3}\}$ *é uma base de tal espaço. Afinal, a única solução da equação* $\alpha\sqrt{3} = 0$ *é a trivial, o que prova que o conjunto* $\{\sqrt{3}\}$ *é L.I. Além disso, qualquer vetor* r *de* $\mathbb{R}$ *pode ser obtido por combinação linear dos vetores de* $\{\sqrt{3}\}$. *Com efeito, se* $r = \alpha\sqrt{3}$, *basta fazer*

$$\alpha = \frac{r}{\sqrt{3}}.$$

Neste caso, as coordenadas de qualquer vetor r *de* $\mathbb{R}$ *relativamente à base* $\{\sqrt{3}\}$ *são simplesmente* $\frac{r}{\sqrt{3}}$.

Perceber que qualquer base de um espaço vetorial real de uma dimensão, como no Exemplo acima, admite uma única base ordenada (Definição 8.8) correspondente a ela.

EXEMPLO 8.34. *Qualquer espaço vetorial real* $\langle\{\bar{0}\}, \mathbb{R}, +, \cdot, \bar{0}\,\rangle$ *tem zero dimensões, de acordo com Teorema 8.11.*

TEOREMA 8.16. *Seja* $\mathcal{W} = \langle W, \mathbb{R}, \oplus, \odot, \bar{0}\,\rangle$ *um* subespaço *de um espaço vetorial real* $\mathcal{V} = \langle V, \mathbb{R}, +, \cdot, \bar{0}\,\rangle$*. Se* $\dim(\mathcal{V}) = n$*, então* $\dim(\mathcal{W}) \leq n$*.*

✍ A prova fica a cargo do leitor.

Espaços vetoriais reais com n dimensões, onde n é um natural, são chamados de *espaços vetoriais reais de dimensão finita*. No EXEMPLO 8.39 e na Seção 97 discutimos sobre espaços vetoriais reais que não têm dimensão finita. Logo, nem todo espaço vetorial real admite base finita. Mas, antes de examinarmos essa questão, é relevante discutirmos outros assuntos.

SEÇÃO 86

Espaços métricos

SUMÁRIO
ÍNDICE
REDE

Apesar do estudo de *espaços métricos* usualmente não ser mencionado em livros de álgebra linear, seu impacto sobre espaços vetoriais é muito marcante. Por conta disso, dedicamos esta Seção a um breve estudo sobre o tema.

DEFINIÇÃO 8.10. *Um par ordenado* $\mathfrak{m} = \langle m, d\rangle$ *é um* espaço métrico *sss*

M1: $m \neq \varnothing$*;*

M2 - DISTÂNCIA: $d : m \times m \to \mathbb{R}$ *é uma função cujas imagens são denotadas por* $d(a, b)$*;*

M3 - IDENTIDADE DOS INDISCERNÍVEIS:

$$d(a, b) = 0 \Leftrightarrow a = b;$$

M4 - SIMETRIA:

$$d(a, b) = d(b, a);$$

M5 - DESIGUALDADE TRIANGULAR:

$$d(a, c) \leq d(a, b) + d(b, c).$$

Obviamente espaços métricos são definidos por meio de um predicado conjuntista, nos moldes da Seção 71. Logo, no presente contexto, todo espaço métrico é um caso muito particular de conjunto, em ZF.

A função d é chamada de *métrica* ou *função-distância*. Para cada par ordenado (a, b) pertencente a $m \times m$, dizemos que $d(a, b)$ é a *distância entre a e b*. Os elementos a e b de m são chamados de *pontos* do espaço métrico. Em outras palavras, uma métrica é uma função real, enquanto uma distância é um número real.

Pontos de um espaço métrico são *indiscerníveis* sss a distância entre eles é zero. Axioma M3 assume que pontos de um espaço métrico são indiscerníveis se, e somente se, forem idênticos.

M4 estabelece que a distância entre a e b é também a distância entre b e a.

Finalmente, M5 garante a ideia intuitiva de que 'desvios' em um espaço métrico não são 'atalhos'. A distância entre um ponto 'de partida' a e um ponto de 'chegada' c é menor ou igual a quaisquer desvios que passem por um ponto b qualquer, antes de chegar de a até c.

Exemplo 8.35. *Seja $\langle \mathbb{R}, d \rangle$ uma interpretação de espaço métrico, onde $d : \mathbb{R}^2 \to \mathbb{R}$ é uma função dada por*

$$d(r, s) = |r - s|.$$

Fórmulas M1 e M2 da Definição 8.10 são teoremas triviais. Além disso,

$$|r - s| = 0 \Leftrightarrow r = s$$

é teorema, o que implica que M3 é teorema.

Fórmula M4 é teorema por conta de $|r - s| = |s - r|$ para quaisquer reais r e s.

✍ *Finalmente, cabe ao leitor provar que*

$$|r - t| \leq |r - s| + |s - t|.$$

A última implica que M5 é teorema para a interpretação $\langle \mathbb{R}, d \rangle$. Logo, temos aqui um primeiro modelo de espaço métrico.

O leitor deve ter percebido que o espaço métrico acima foi amplamente usado nas discussões da Seção 44 dedicada a limites de funções reais. Analogamente, $\langle \mathbb{Q}, d \rangle$ é um espaço métrico se $d : \mathbb{R}^2 \to \mathbb{R}$ é uma função definida como $d(r,s) = |r - s|$. Este segundo exemplo foi utilizado nas discussões sobre sequências convergentes da Seção 35.

Exemplo 8.36. ✍ *Seja $\langle m, d \rangle$ um par ordenado, onde $m \neq \varnothing$ e $d : m \times m \to \mathbb{R}$ é a função dada por*

$$d(a,b) = \begin{cases} 1 & \text{se } a \neq b \\ 0 & \text{se } a = b. \end{cases}$$

Neste caso $\langle m, d \rangle$ é um espaço métrico. Cabe ao leitor o ônus da prova.

O Exemplo acima é conhecido como *espaço métrico discreto*. Ele também demonstra que qualquer conjunto m não vazio pode ser munido de métrica.

Teorema 8.17. *Em todo conjunto não vazio é possível definir uma métrica.*

Demonstração: Exemplo 8.36 prova isso.

Teorema 8.18. *Em um espaço métrico nenhuma distância é um real negativo.*

Demonstração: Seja $\langle m, d \rangle$ um espaço métrico. Sejam ainda a e b pontos quaisquer do espaço. Logo, de acordo com a Desigualdade Triangular (axioma M5), temos

$$d(a,b) + d(b,a) \geq d(a,a).$$

Logo, o postulado de Simetria M4 implica que

$$d(a,b) + d(a,b) \geq d(a,a).$$

Logo, a Identidade dos Indiscerníveis M3 garante que

$$2d(a,b) \geq 0.$$

Portanto,

$$d(a,b) \geq 0.$$

O estudo de espaços métricos conta com muitos outros resultados e aplicações. Exemplos já explorados aqui são os conceitos de limites de funções reais e de sequências racionais. No entanto, na próxima Seção exibimos outras aplicações no contexto de espaços vetoriais reais. Mostramos, entre outras coisas, que certos espaços vetoriais reais podem ser munidos de métricas não triviais (ou seja, diferentes da métrica discreta do Exemplo 8.36).

Seção 87

Produto interno

Sumário
Índice
Rede

Qualificamos aqui *produto interno*, uma operação binária entre vetores que produz escalares.

Definição 8.11. *Seja* $\mathcal{V} = \langle V, \mathbb{R}, +, \cdot, \overline{0}\, \rangle$ *um espaço vetorial real. Uma função*

$$\langle\,,\rangle : V \times V \to \mathbb{R},$$

cujas imagens $\langle\,,\rangle(u,v)$ *são denotadas abreviadamente por* $\langle u, v\rangle$*, é um* produto interno *em* $\mathcal{V}$ *se as seguintes fórmulas são teoremas.*

pi1: $\langle u, v\rangle = \langle v, u\rangle$;

pi2: $\langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$;

pi3: $\langle \alpha u, v\rangle = \alpha\langle u, v\rangle$, *onde* α *é um escalar;*

pi4: $\langle u, u\rangle > 0$ *se* $u \neq \overline{0}$.

Referimo-nos a $(\mathcal{V}, \langle\,,\rangle)$ *como um* espaço vetorial real $\mathcal{V}$ munido de produto interno $\langle\,,\rangle$.

STOP É importante que o leitor não confunda pares ordenados $\langle a, b\rangle$ com produto interno $\langle a, b\rangle$. Para que não exista risco de confusão, sempre qualificamos quando é um caso ou o outro. No caso de produto interno, lemos $\langle a, b\rangle$ como 'produto interno entre a e b'.

Axioma pi1 é chamado de *simetria*. O nome 'simetria' é preferido no lugar de 'comutatividade' por conta de produto interno não ser uma operação fechada no espaço de vetores, apesar de ser definida sobre pares ordenados de vetores.

Postulados pi2 e pi3 são conhecidos como *bilinearidade*.

Finalmente, pi4 é chamado de *positividade*.

Produtos internos são casos particulares de *funcionais*. Funcionais são funções cujos domínios são espaços vetoriais ou produtos cartesianos de um espaço vetorial por ele mesmo, tais que suas imagens são escalares. Neste contexto,

um produto interno em um espaço vetorial real é um funcional simétrico, bilinear e positivo.

Teorema 8.19. *Sejam* $\mathcal{V} = \langle V, \mathbb{R}, +, \cdot, \bar{0}\,\rangle$ *um espaço vetorial real e*

$$b = \{v_1, v_2, \cdots, v_n\}$$

uma base de $\mathcal{V}$. *Se*

$$(\alpha_1, \alpha_2, \cdots, \alpha_n) + (\beta_1, \beta_2, \cdots, \beta_n) =$$
$$(\alpha_1 + \beta_1, \alpha_2 + \beta_2, \cdots, \alpha_n + \beta_n)$$

e

$$\gamma(\alpha_1, \alpha_2, \cdots, \alpha_n) = (\gamma\alpha_1, \gamma\alpha_2, \cdots, \gamma\alpha_n),$$

onde $\gamma \in \mathbb{R}$ *e* $(\alpha_1, \alpha_2, \cdots, \alpha_n)$ *e* $(\beta_1, \beta_2, \cdots, \beta_n)$ *são coordenadas de vetores quaisquer de* V, *relativamente a* b, *então* $\mathcal{V}$ *pode ser munido de produto interno.*

Demonstração: Para qualquer vetor v existem n escalares α_i tais que

$$v = \sum_{i=1}^{n} \alpha_i v_i.$$

Ou seja, as coordenadas de v relativamente à base b são

$$(\alpha_1, \alpha_2, \cdots, \alpha_n).$$

Consideremos a função $\langle\, , \rangle : V \times V \to \mathbb{R}$ tal que

$$\langle(\alpha_1, \alpha_2, \cdots, \alpha_n), (\beta_1, \beta_2, \cdots, \beta_n)\rangle = \sum_{i=1}^{n} \alpha_i\beta_i.$$

Provamos a seguir que essa função define um produto interno em $\mathcal{V}$.

Em primeiro lugar, $\langle\, , \rangle$ é um funcional. Além disso, temos

$$\langle(\alpha_1, \alpha_2, \cdots, \alpha_n), (\beta_1, \beta_2, \cdots, \beta_n)\rangle = \sum_{i=1}^{n} \alpha_i\beta_i = \sum_{i=1}^{n} \beta_i\alpha_i =$$
$$\langle(\beta_1, \beta_2, \cdots, \beta_n), (\alpha_1, \alpha_2, \cdots, \alpha_n)\rangle,$$

por conta da comutatividade da multiplicação entre reais.

Portanto, PI1 é teorema.

Em segundo lugar,

$$\langle(\alpha_1, \alpha_2, \cdots, \alpha_n) + (\beta_1, \beta_2, \cdots, \beta_n), (\gamma_1, \gamma_2, \cdots, \gamma_n)\rangle = \langle(\alpha_1 + \beta_1, \alpha_2 + \beta_2, \cdots, \alpha_n + \beta_n), (\gamma_1, \gamma_2, \cdots, \gamma_n)\rangle.$$

Mas este último termo é idêntico a

$$\sum_{i=1}^{n}(\alpha_i + \beta_i)\gamma_i,$$

o qual é igual a

$$\sum_{i=1}^{n}(\alpha_i\gamma_i + \beta_i\gamma_i) = \sum_{i=1}^{n}\alpha_i\gamma_i + \sum_{i=1}^{n}\beta_i\gamma_i = \langle(\alpha_1, \alpha_2, \cdots, \alpha_n), (\gamma_1, \gamma_2, \cdots, \gamma_n)\rangle + \langle(\beta_1, \beta_2, \cdots, \beta_n), (\gamma_1, \gamma_2, \cdots, \gamma_n)\rangle.$$

Isso prova que PI2 é teorema.

Em terceiro lugar,

$$\langle\alpha(\alpha_1, \alpha_2, \cdots, \alpha_n), (\beta_1, \beta_2, \cdots, \beta_n)\rangle = \langle(\alpha\alpha_1, \alpha\alpha_2, \cdots, \alpha\alpha_n), (\beta_1, \beta_2, \cdots, \beta_n)\rangle.$$

Observar que α não é necessariamente idêntico a qualquer α_i. Mas o último termo é igual a

$$\sum_{i=1}^{n}\alpha\alpha_i\beta_i = \alpha\sum_{i=1}^{n}\alpha_i\beta_i = \alpha\langle(\alpha_1, \alpha_2, \cdots, \alpha_n), (\beta_1, \beta_2, \cdots, \beta_n)\rangle.$$

Isso prova que PI3 é teorema.

Finalmente,

$$\langle(\alpha_1, \alpha_2, \cdots, \alpha_n), (\alpha_1, \alpha_2, \cdots, \alpha_n)\rangle = \sum_{i=1}^{n}\alpha_i\alpha_i = \sum_{i=1}^{n}\alpha_i^2.$$

Os axiomas da Definição 8.1 garantem que

$$(\alpha_1, \alpha_2, \cdots, \alpha_n) = \overline{0}$$

sss cada α_i é 0.

Logo, se algum α_i for diferente de 0, então

$$\langle(\alpha_1, \alpha_2, \cdots, \alpha_n), (\alpha_1, \alpha_2, \cdots, \alpha_n)\rangle > 0,$$

o que prova que PI4 é teorema.

✍ Notar que o produto interno sugerido na última demonstração não é o único possível. A função $\langle\,,\rangle : V \times V \to \mathbb{R}$ tal que

$$\langle(\alpha_1, \alpha_2, \cdots, \alpha_n), (\beta_1, \beta_2, \cdots, \beta_n)\rangle = \sum_{i=1}^{n} c_i \alpha_i \beta_i$$

também define um produto interno em $\mathcal{V}$, se cada c_i for um real estritamente positivo. Cabe ao leitor demonstrar isso. No caso em que cada c_i é 1, temos o *produto interno canônico* (o mesmo da demonstração acima) em uma vasta gama de espaços vetoriais reais de dimensão finita (incluindo qualquer $\mathbb{R}^n$ usual). Portanto, um mesmo espaço vetorial real de dimensão finita admite uma infinidade de possíveis produtos internos.

Exemplo 8.37. *A função $\langle\,,\rangle : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ dada por*

$$\langle(a, b), (c, d)\rangle = ac + bd$$

é um produto interno em $\mathbb{R}^2$ usual.

Exemplo 8.38. *Seja $\mathcal{V} = \langle S, \mathbb{R}, \oplus, \odot, \bigcirc\rangle$, onde*

I: $S = \{y \in \mathcal{C}^\infty \mid y'' + y = 0\}$,

II: *$\oplus$ é restrição de $+$ (no espaço $\mathcal{C}^\infty$ usual) a S,*

III: *$\odot$ é restrição de $\cdot$ (no espaço $\mathcal{C}^\infty$ usual) a S, e*

IV: *$\bigcirc$ é a função real identicamente nula, com domínio $\mathbb{R}$.*

O conjunto $\{\text{sen}, \cos\}$ é uma base de $\mathcal{V}$ (conforme Exemplo *8.32), o que implica que qualquer função f de S é da forma $f = \alpha\,\text{sen} + \beta\cos$. Logo, se*

$$f = \alpha\,\text{sen} + \beta\cos$$

e

$$g = \gamma\,\text{sen} + \delta\cos$$

são vetores de S, então $\langle\,,\rangle : S^2 \to \mathbb{R}$, dada por

$$\langle f, g\rangle = \pi(\alpha\gamma + \beta\delta),$$

é um produto interno em $\mathcal{V}$. Lembrar que π é uma constante real estritamente positiva.

Nem todo espaço vetorial real é de dimensão finita, como se verifica no próximo Exemplo.

EXEMPLO 8.39. *Seja* $\mathcal{V} = \langle S, \mathbb{R}, \oplus, \odot, \bigcirc \rangle$, *onde*

I: $S = \{y \in \mathcal{C}^\infty \mid y$ *é polinomial com domínio* $[-a, a]\}$, *sendo* $a \neq 0$,

II: $\oplus$ *é restrição de* $+$ *(no espaço* $\mathcal{C}^\infty$ *usual) a* $S \times S$,

III: $\odot$ *é restrição de* $\cdot$ *(no espaço* $\mathcal{C}^\infty$ *usual) a* $\mathbb{R} \times S$, *e*

IV: $\bigcirc$ *é a função real identicamente nula, com domínio* $\mathbb{R}$.

Por um lado, adição de funções polinomiais é uma função polinomial, e multiplicação entre um real e uma função polinomial é uma função polinomial. Isso ocorre mesmo em um domínio definido por um intervalo fechado não degenerado $[-a, a]$. *Por outro, a função identicamente nula* $\bigcirc$ *é polinomial (de grau 0). Logo, de acordo com Definição 8.3,* $\mathcal{V}$ *é subespaço de* $\mathcal{C}^\infty$ *usual. Portanto, de acordo com Teorema 8.9,* $\mathcal{V}$ *é espaço vetorial real.*

Supor que $\mathcal{V}$ *admite base finita*

$$b = \{p_1, p_2, \cdots, p_n\},$$

onde cada p_i *pertencente a* b *é polinomial. Logo, existe natural* m *tal que* m *é o grau do polinômio de maior grau entre todas as funções de* b. *Portanto, se* p *for uma função polinomial de grau maior do que* m, p *não é combinação linear dos vetores de* b, *por conta do Teorema Fundamental da Álgebra (ver* OBSERVAÇÃO FINAL *da Seção 43).*

O fato de um espaço vetorial real não ter base finita não impede necessariamente a definição de um produto interno:

EXEMPLO 8.40. *No espaço vetorial real* $\mathcal{V}$ *do* EXEMPLO 8.39, *considere a seguinte função* $\langle\, , \rangle : S^2 \to \mathbb{R}$, *definida por*

$$\langle p, q \rangle = \int_{-a}^{a} p(x)q(x)dx.$$

O produto entre funções polinomiais é polinomial. Além disso, toda polinomial é integrável. Logo, $\langle\, , \rangle$ *é um funcional.*

Além disso,

$$\int_{-a}^{a} p(x)q(x)dx = \int_{-a}^{a} q(x)p(x)dx,$$

o que garante que PI1 *é teorema.*

Também temos que

$$\langle p+q, r\rangle = \int_{-a}^{a}((p(x)+q(x))r(x))dx =$$

$$\int_{-a}^{a}(p(x)r(x)+q(x)r(x))dx =$$

$$\int_{-a}^{a}p(x)r(x) + \int_{-a}^{a}q(x)r(x)dx = \langle p, r\rangle + \langle q, r\rangle.$$

Logo, pi2 *é teorema. Foi usado acima o Teorema 6.10.*

Ademais,

$$\langle \alpha p, q\rangle = \int_{-a}^{a}(\alpha p(x))q(x)dx = \int_{-a}^{a}\alpha p(x)q(x)dx =$$

$$\alpha\int_{-a}^{a}p(x)q(x)dx = \alpha\langle p, q\rangle,$$

o que garante que pi3 *é teorema. Novamente foi empregado o Teorema 6.10.*

Finalmente,

$$\langle p, p\rangle = \int_{-a}^{a}p(x)^2dx \geq 0,$$

o que prova que pi4 *também é teorema. Foi usado o Teorema 6.7.*

Logo, $\langle\,,\rangle$ *é um produto interno.*

Exemplo 8.41. ✍ *No mesmo espaço vetorial real* $\mathcal{V}$ *do* Exemplo 8.39*, a função* $\langle\,,\rangle : S^2 \to \mathbb{R}$*, definida por*

$$\langle p, q\rangle = c\int_{-a}^{a}p(x)q(x)dx$$

é um produto interno em $\mathcal{V}$*, se* $c > 0$*. Cabe ao leitor a prova.*

✍ Se, em um dado espaço vetorial real não trivial (de dimensão finita ou não), for possível definir um produto interno $\langle\,,\rangle$, então é possível definir uma infinidade de outros produtos internos.

Teorema 8.20. *Seja*

$$\mathcal{V} = \langle V, \mathbb{R}, +, \cdot, \overline{0}\,\rangle$$

um espaço vetorial real munido de produto interno $\langle\,,\rangle$*. Então, para qualquer vetor* u *do espaço, temos*

$$\langle u, \overline{0}\rangle = 0.$$

Demonstração: De acordo com pi1, $\langle u, \overline{0}\rangle = \langle \overline{0}, u\rangle$. De acordo com pi2 e axioma V5 da Seção 80,

$$\langle \overline{0}, u\rangle = \langle \overline{0}+\overline{0}, u\rangle = \langle \overline{0}, u\rangle + \langle \overline{0}, u\rangle.$$

Mas a equação

$$\langle \overline{0}, u\rangle = \langle \overline{0}, u\rangle + \langle \overline{0}, u\rangle$$

somente é teorema se $\langle \overline{0}, u\rangle = 0$.

Seção 88

Norma de um vetor

Sumário

Índice

Rede

Todo espaço vetorial real munido de produto interno $\langle\,,\rangle$ é um espaço métrico cuja função-distância é induzida por $\langle\,,\rangle$. Este é o tema principal desta Seção.

Definição 8.12. *Seja $\mathcal{V} = \langle V, \mathbb{R}, +, \cdot, \overline{0}\,\rangle$ um espaço vetorial real munido de produto interno $\langle\,,\rangle$. A* norma *de um vetor v é*

$$\parallel v \parallel = \sqrt{\langle v, v\rangle}.$$

Teorema 8.21. *Seja $\mathcal{V} = \langle V, \mathbb{R}, +, \cdot, \overline{0}\,\rangle$ um espaço vetorial real munido de produto interno $\langle\,,\rangle$. Então $\parallel \overline{0} \parallel = 0$.*

Demonstração: Consequência imediata de Teorema 8.20 e da definição de norma.

Teorema 8.22. *Seja $\mathcal{V} = \langle V, \mathbb{R}, +, \cdot, \overline{0}\rangle$ um espaço vetorial real munido de produto interno $\langle\,,\rangle$. Então*

$$\parallel \lambda v \parallel = |\lambda| \cdot \parallel v \parallel,$$

onde λ é um escalar.

Demonstração: De acordo com Definição 8.12 e axiomas pi1 e pi3 da Definição 8.11,

$$\parallel \lambda v \parallel = \sqrt{\langle \lambda v, \lambda v\rangle} = \sqrt{\lambda\langle v, \lambda v\rangle} =$$

$$\sqrt{\lambda^2\langle v, v\rangle} = \sqrt{\lambda^2}\sqrt{\langle v, v\rangle} = |\lambda| \cdot \parallel v \parallel .$$

Teorema 8.23 (Desigualdade de Cauchy-Schwarz). *Se*

$$\mathcal{V} = \langle V, \mathbb{R}, +, \cdot, \overline{0} \rangle$$

é um espaço vetorial real, munido de produto interno

$$\langle\,,\rangle : V \times V \to \mathbb{R},$$

e u e v são vetores de V, então

$$|\langle u, v\rangle| \leq \parallel u \parallel \cdot \parallel v \parallel .$$

Demonstração: Se $v = \overline{0}$, a prova é imediata, graças ao Teorema 8.20. Situação análoga para $u = \overline{0}$, uma vez que produto interno é simétrico, de acordo com pi1.

Agora consideremos a situação na qual $v \neq \overline{0}$. Neste caso, Definição 8.12 de norma, bem como os axiomas da Definição 8.11, garantem que

$$0 \leq \parallel u + \lambda v \parallel^2 = \langle u + \lambda v, u + \lambda v\rangle =$$
$$\langle u, u + \lambda v\rangle + \langle \lambda v, u + \lambda v\rangle =$$
$$\langle u, u\rangle + \langle u, \lambda v\rangle + \langle \lambda v, u\rangle + \langle \lambda v, \lambda v\rangle =$$
$$\langle u, u\rangle + 2\lambda\langle u, v\rangle + \lambda^2\langle v, v\rangle.$$

Levando em conta que essa desigualdade vale para qualquer λ real, façamos

$$\lambda = -\frac{\langle u, v\rangle}{\langle v, v\rangle},$$

uma vez que $v \neq \overline{0}$ e, portanto, $\langle v, v\rangle \neq 0$.

Logo, de acordo com a desigualdade acima,

$$0 \leq \langle u, u\rangle - 2\frac{\langle u, v\rangle}{\langle v, v\rangle}\langle u, v\rangle + \frac{\langle u, v\rangle^2}{\langle v, v\rangle^2}\langle v, v\rangle =$$

$$\langle u, u\rangle - \frac{\langle u, v\rangle^2}{\langle v, v\rangle}.$$

Portanto,

$$\langle u, u\rangle \geq \frac{\langle u, v\rangle^2}{\langle v, v\rangle},$$

o que implica em $\langle u, u\rangle\langle v, v\rangle \geq \langle u, v\rangle^2$. Logo, Definição 8.12 garante que

$$|\langle u, v\rangle| \leq \parallel u \parallel \cdot \parallel v \parallel .$$

Uma das consequências da desigualdade de Cauchy-Schwarz é a famosa desigualdade triangular, resultado estratégico para a prova de que normas induzidas por produtos internos também induzem métricas.

Teorema 8.24 (Desigualdade Triangular). *Seja*

$$\mathcal{V} = \langle V, \mathbb{R}, +, \cdot, \overline{0} \rangle$$

um espaço vetorial real munido de produto interno $\langle \, , \rangle$.

Se a *e* b *são vetores de* V*, então*

$$\| \, a + b \, \| \leq \| \, a \, \| + \| \, b \, \| \, .$$

Demonstração: De acordo com Definição 8.12 e axiomas pi1 e pi2 da Definição 8.11,

$$\| \, a + b \, \|^2 = \langle a + b, a + b \rangle = \langle a, a + b \rangle + \langle b, a + b \rangle =$$
$$\langle a, a \rangle + \langle a, b \rangle + \langle b, a \rangle + \langle b, b \rangle =$$
$$\| \, a \, \|^2 + 2\langle a, b \rangle + \| \, b \, \|^2 \leq$$
$$\| \, a \, \|^2 + 2|\langle a, b \rangle| + \| \, b \, \|^2 \, .$$

Mas Teorema 8.23 sobre a desigualdade de Cauchy-Schwarz garante que este último termo é menor ou igual a

$$\| \, a \, \|^2 + 2 \, \| \, a \, \| \cdot \| \, b \, \| + \| \, b \, \|^2,$$

sendo este último idêntico a $(\| \, a \, \| + \| \, b \, \|)^2$. Uma vez que toda norma de qualquer vetor é maior ou igual a zero, então

$$\| \, a + b \, \| \leq \| \, a \, \| + \| \, b \, \| \, .$$

O próximo teorema mostra que a norma em um espaço vetorial real munido de produto interno pode ser usada para calcular distâncias (ver Definição 8.10) entre vetores.

Teorema 8.25. *Seja* $\mathcal{V} = \langle V, \mathbb{R}, +, \cdot, \overline{0} \rangle$ *um espaço vetorial real munido de produto interno* $\langle \, , \rangle$. *Seja ainda* $d : V \times V \to \mathbb{R}$ *uma função dada por*

$$d(u, v) = \| \, u - v \, \| \, .$$

Logo, $\langle V, d \, \rangle$ *é um espaço métrico.*

Demonstração: Basta provarmos que a função d define uma métrica no espaço V de vetores. Mas d é uma função cujas imagens são reais e cujo domínio é não vazio. Logo, axiomas M1 e M2 de espaço métrico (Definição 8.10) são teoremas.

Sobre M3 (Definição 8.10), observar o seguinte.

$$d(u,u) = \| \, u - u \, \| = \| \, \overline{0} \, \| \; .$$

Logo, Teorema 8.21 garante que $d(u,u) = 0$.

Por outro lado, axioma PI4, em parceria com Definição 8.12, garante que o vetor nulo é o único cuja norma é zero. Logo, se $d(u,v) = 0$, então $\| \, u - v \, \| = 0$, o que implica em $u - v = \overline{0}$. Mas isso somente ocorre se $u = v$. Portanto, M3 é teorema.

Sobre M4, observar que

$$d(u,v) = \| \, u - v \, \| = \| \, (-1)(v - u) \, \|,$$

de acordo com Teorema 8.5. No entanto, o termo à direita da última igualdade é $|-1| \cdot \| \, (v-u) \, \|$, de acordo com Teorema 8.22. Logo, $d(u,v) = \| \, v - u \, \|$, sendo que $\| \, v - u \, \|$ é $d(v,u)$. Portanto, M4 é teorema.

O último postulado da Definição 8.10 a ser avaliado é M5. Ou seja, precisamos provar que

$$\| \, u - w \, \| \leq \| \, u - v \, \| + \| \, v - w \, \| \; .$$

Para isso, basta substituir a por $u - v$ e b por $v - w$ no Teorema 8.24.

Exemplo 8.42. *Seja $\mathbb{R}^2$ usual munido do produto interno canônico*

$$\langle \; , \; \rangle : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$$

dado por $\langle (a,b), (c,d) \rangle = ac + bd$.

Neste caso, a norma induzida pelo produto interno $\langle \; , \; \rangle$ induz também a seguinte métrica $d : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$:

$$d((a,b),(c,d)) = \sqrt{\langle (a,b) - (c,d), (a,b) - (c,d) \rangle} =$$

$$\sqrt{\langle (a-c, b-d), (a-c, b-d) \rangle} = \sqrt{(a-c)^2 + (b-d)^2}.$$

Esta é a métrica euclidiana *em $\mathbb{R}^2$ usual. É também o comprimento do segmento de reta $[(a,b),(c,d)]$, de acordo com Seção 77. Logo, $\mathbb{R}^2$ usual, munido do produto interno canônico, é o plano cartesiano.*

Exemplo 8.43. *Em* $\mathbb{R}^2$ *usual munido do produto interno canônico (como no* Exemplo *anterior), uma circunferência* c *com centro* (a, b) *é o conjunto dos pontos* (x, y) *equidistantes de* (a, b)*. Logo,*

$$c = \{(x, y) \in \mathbb{R}^2 \mid \sqrt{(x-a)^2 + (y-b)^2} = r\},$$

sendo $r > 0$ *a distância que define o raio da circunferência* c*.*

Notar que a definição da circunferência c *em* $\mathbb{R}^2$ *usual foi feita por aplicação do Esquema de Separação.*

Exemplo 8.44. *Considere o espaço vetorial real* $\langle \mathbb{R}, \mathbb{R}, +, \cdot, 0 \rangle$ *usual. Neste caso, a multiplicação entre reais (os quais são vetores, além de escalares) é o produto interno canônico de tal espaço. Simetria, por exemplo, é consequência da comutatividade da multiplicação entre reais.*

A norma induzida pelo produto interno canônico neste espaço induz também a seguinte métrica

$$d : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$$

dada por

$$d(r, s) = \sqrt{\langle r-s, r-s \rangle} = \sqrt{(r-s)^2} = |r-s|.$$

Desnecessário dizer que esta é exatamente a mesma métrica usada na definição de limite de função real na Seção 44. Observar que $d(r, s) = |r - s|$ *é a* métrica euclidiana na reta dos reais.

Exemplo 8.45. *No espaço vetorial real* $\mathcal{V}$ *do* Exemplo 8.39*, considere a seguinte função* $\langle \, , \rangle : S^2 \to \mathbb{R}$*, definida por*

$$\langle p, q \rangle = \int_{-a}^{a} p(x)q(x)dx.$$

Foi demonstrado no Exemplo 8.40 *que* $\langle \, , \rangle$ *é um produto interno. Logo, é possível calcular a distância, induzida por* $\langle \, , \rangle$*, entre as seguintes funções:*

$$p : [-a, a] \to \mathbb{R} \quad \textit{dada por} \quad p(x) = x^2$$

e

$$q : [-a, a] \to \mathbb{R} \quad \textit{dada por} \quad q(x) = 2x.$$

✍ *Recomendamos que o leitor faça as contas.*

✍ Outro exercício interessante é definir um espaço vetorial real de matrizes, bem como um produto interno neste espaço. Em seguida, mostrar como calcular a distância entre matrizes deste espaço a partir da métrica induzida pelo produto interno escolhido.

Uma vez que um mesmo espaço vetorial real pode ser munido de uma infinidade de produtos internos, cada um deles induz uma nova métrica no mesmo espaço.

A seguir começamos a nos aproximar de *geometria analítica*.

SEÇÃO 89

Ortogonalidade

SUMÁRIO
ÍNDICE
REDE

Ortogonalidade é um conceito algébrico, como se percebe na próxima definição. Perpendicularismo, porém, é um conceito geométrico a ser discutido na próxima Seção.

DEFINIÇÃO 8.13. *Vetores u e v de um espaço vetorial real munido de produto interno $\langle\, , \rangle$ são* ortogonais entre si *sss*

$$\langle u, v\rangle = 0.$$

Um conjunto de vetores $\{v_1, v_2, \cdots, v_n\}$ é ortogonal *sss v_i é ortogonal a v_j para quaisquer i e j tais que $i \neq j$, $1 \leq i \leq n$ e $1 \leq j \leq n$.*

Uma vez que produto interno é simétrico, se u e v são vetores ortogonais entre si, então v e u são ortogonais entre si. Logo, é usual dizer que u é ortogonal a v, se $\langle u, v\rangle = 0$.

TEOREMA 8.26. *Seja*

$$\mathcal{V} = \langle V, \mathbb{R}, +, \cdot, \overline{0}\,\rangle$$

um espaço vetorial real munido de produto interno $\langle\ ,\ \rangle$. Logo, o vetor nulo $\overline{0}$ é ortogonal a qualquer vetor de V.

DEMONSTRAÇÃO: Consequência imediata do Teorema 8.20 sobre produto interno entre vetor nulo e qualquer outro vetor de um mesmo espaço vetorial real.

TEOREMA 8.27. *Seja* $\mathcal{V} = \langle V, \mathbb{R}, +, \cdot, \bar{0} \rangle$ *um espaço vetorial real munido de produto interno* $\langle\ ,\rangle$*. Logo, qualquer conjunto* $\{v_1, v_2, \cdots, v_n\}$ *de vetores ortogonais não nulos é linearmente independente.*

DEMONSTRAÇÃO: Se $x = \{v_1, v_2, \cdots, v_n\}$ é ortogonal então, para quaisquer v_i e v_j pertencentes a x, temos $\langle v_i, v_j \rangle = 0$ se $j \neq i$. Considere agora a fórmula

$$\sum_{j=1}^{n} \alpha_j v_j = \bar{0}.$$

Se provarmos que os únicos valores α_j a garantirem que a fórmula acima é teorema são aqueles que correspondem à solução trivial, demonstramos que x é L.I., de acordo com Teorema 8.10. A partir desta equação temos

$$\left\langle \sum_{j=1}^{n} \alpha_j v_j, v_i \right\rangle = \langle \bar{0}, v_i \rangle,$$

para cada v_i pertencente a x. Logo, axiomas PI2 e PI3 da Definição 8.11 e Teorema 8.20 garantem que

$$\sum_{j=1}^{n} \alpha_j \langle v_j, v_i \rangle = 0.$$

Uma vez que x é ortogonal, então a igualdade acima implica em

$$\alpha_i \langle v_i, v_i \rangle = 0.$$

No entanto, lembrando que os vetores pertencentes a x são não nulos, então axioma PI4 da Definição 8.11 garante que $\langle v_i, v_i \rangle \neq 0$. Logo, α_i deve ser zero para cada i tal que $1 \leq i \leq n$. Portanto, Teorema 8.10 implica que x é linearmente independente.

A recíproca do último teorema não é teorema, como se ilustra no próximo EXEMPLO.

EXEMPLO 8.46. *O conjunto* $\{(\pi, \sqrt{3}), (\sqrt{3}, \pi)\}$ *é linearmente independente em* $\mathbb{R}^2$ *usual munido do produto interno canônico* $\langle\ ,\rangle$*. Cabe ao leitor provar isso. No entanto,*

$$\langle (\pi, \sqrt{3}), (\sqrt{3}, \pi) \rangle = 2\pi\sqrt{3} \neq 0.$$

A contrapositiva do Teorema 8.27 garante que, em um conjunto

$$x = \{v_1, v_2, \cdots, v_n\}$$

de vetores não nulos, linearmente dependente, existe pelo menos um par $\{v_i, v_j\}$ de vetores distintos não ortogonais entre si.

Exemplo 8.47. *Seja $\mathcal{V} = \langle V, \mathbb{R}, +, \cdot, \overline{0}\,\rangle$ um espaço vetorial real munido de produto interno $\langle\,,\rangle$. Se $\{u, v\}$ é um conjunto de vetores não nulos de V, linearmente dependente, então existe α real tal que $u = \alpha v$. Logo,*

$$\parallel u \parallel = \sqrt{\langle u, u\rangle} = \sqrt{\langle \alpha v, \alpha v\rangle} =$$

$$\sqrt{\alpha^2 \langle v, v\rangle} = |\alpha| \sqrt{\langle v, v\rangle} = |\alpha| \cdot \parallel v \parallel .$$

Este último Exemplo inspira a próxima definição.

Definição 8.14. *Seja $\mathcal{V} = \langle V, \mathbb{R}, +, \cdot, \overline{0}\rangle$ um espaço vetorial real munido de produto interno. Dizemos que um vetor não nulo v de V é* unitário *sss*

$$\parallel v \parallel = 1.$$

Vetores unitários nada têm a ver com conjuntos unitários.

Teorema 8.28. *Seja $\mathcal{V} = \langle V, \mathbb{R}, +, \cdot, \overline{0}\rangle$ um espaço vetorial real munido de produto interno $\langle\,,\rangle$. Logo, para qualquer vetor não nulo v de V existe vetor w tal que*

$$v = \alpha \cdot w$$

e

$$\parallel w \parallel = 1.$$

Demonstração: Basta fazer $\alpha = \parallel v \parallel$. Com efeito, se $v = \parallel v \parallel \cdot w$, então $w = \parallel v \parallel^{-1} \cdot v$. Logo,

$$\parallel w \parallel = \sqrt{\langle w, w\rangle} = \sqrt{\langle \parallel v \parallel^{-1} \cdot v, \parallel v \parallel^{-1} \cdot v\rangle} =$$

$$\sqrt{\parallel v \parallel^{-2} \langle v, v\rangle} = \parallel v \parallel^{-1} \sqrt{\langle v, v\rangle} = \parallel v \parallel^{-1} \cdot \parallel v \parallel = 1.$$

O último teorema justifica a prática de *normalização* de vetores em espaços vetoriais reais munidos de produto interno. A normalização de um vetor v não nulo é feita definindo um vetor w tal que $\{v, w\}$

é L.D. e

$$w = \frac{1}{\| v \|} \cdot v.$$

À luz dos Teoremas 8.27 e 8.28 introduzimos o seguinte.

Definição 8.15. *Seja $\mathcal{V} = \langle V, \mathbb{R}, +, \cdot, \overline{0} \rangle$ um espaço vetorial real munido de produto interno $\langle\, , \rangle$. Um conjunto*

$$x = \{v_1, v_2, \cdots, v_n\}$$

define uma base ortonormal *de $\mathcal{V}$ sss*

I: *x é uma base de $\mathcal{V}$,*

II: *x é ortogonal e*

III: *cada vetor de x é unitário.*

Exemplo 8.48. ✍ *Considere $\mathbb{R}^3$ usual munido do produto interno canônico*

$$\langle\, , \rangle : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$$

dado por

$$\langle (a, b, c), (d, e, f) \rangle = ad + be + cf.$$

Neste caso,

$$\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$$

define uma base ortonormal de $\mathbb{R}^3$ usual, enquanto

$$\{(1, 0, 0), (0, 1, 0), (0, 0, \pi)\}$$

é uma base para o mesmo espaço, mas não ortonormal.

Exemplo 8.49. ✍ *Seja $\mathcal{P}_2 = \langle P_2, \mathbb{R}, +, \cdot, \bigcirc \rangle$ o espaço vetorial real tal que*

I: *P_2 é o conjunto das funções reais polinomiais de grau menor ou igual a 2 e domínio $\mathbb{R}$,*

II: *$+$ é a adição usual de funções polinomiais de grau menor ou igual a 2, ou seja, se p e q são funções de P_2, então $p+q$ é uma função tal que $(p+q)(x) = p(x) + q(x)$,*

III: *$\cdot$ é a multiplicação usual de real por uma função polinomial de grau menor ou igual a 2, ou seja, se $p \in P_2$ e α é um real, então αp é uma função tal que $(\alpha p)(x) = \alpha p(x)$, e* iv: $\bigcirc$ *é a função identicamente nula com domínio* $\mathbb{R}$.

Cabe ao leitor provar que este é um espaço vetorial real.

Logo, se p, q *e* r *são funções reais com domínio* $\mathbb{R}$ *tais que* $p(x) = 1$, $q(x) = x$ *e* $r(x) = x^2$, *então o conjunto*

$$m = \{p, q, r\}$$

define uma base para $\mathcal{P}_2$ *(o que implica que* $\mathcal{P}_2$ *tem três dimensões).*

Neste contexto, qualquer função $v : \mathbb{R} \to \mathbb{R}$ *de* P_2 *tal que*

$$v(x) = ax^2 + bx + c,$$

tem coordenadas

$$(a, b, c)$$

relativamente à base m.

Se definirmos o produto interno $\langle\, ,\rangle : P_2 \times P_2 \to \mathbb{R}$ *como*

$$\langle ax^2 + bx + c, dx^2 + ex + f\rangle = adx^2 + bex + cf,$$

então m *é uma base ortonormal de* $\mathcal{P}_2$ *relativamente a este produto interno. Recomendamos ao leitor que prove isso.*

Seção 90

Noções elementares sobre geometria analítica

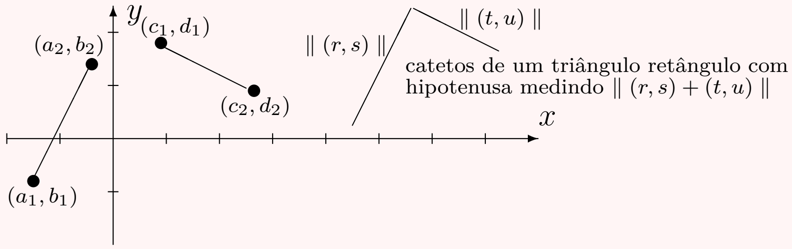
Sumário
Índice
Rede

Geometria analítica plana, grosso modo, é o estudo de uma interpretação do plano euclidiano (como apresentado na Parte 7) dada por $\mathbb{R}^2$ usual, o qual é munido do produto interno canônico. Detalhes no Exemplo 8.42.

Geometria analítica espacial, por sua vez, é uma extensão da geometria analítica plana para $\mathbb{R}^3$ usual, também munido de produto interno canônico.

Geometria analítica trata de geometria analítica plana e geometria analítica espacial.

Espaços $\mathbb{R}^n$, para $n > 3$, podem ser usados para generalizar os resultados de geometria analítica. Discutimos aqui apenas sobre o plano cartesiano, o qual mostramos a seguir que pode ser identificado com $\mathbb{R}^2$ usual munido do produto interno canônico.



Na imagem acima ilustramos ideias intuitivas para o próximo teorema.

TEOREMA 8.29. *Ortogonalidade entre vetores p e q do plano cartesiano é equivalente a perpendicularismo entre segmentos de reta definidos por p e q.*

DEMONSTRAÇÃO: Um segmento de reta do plano cartesiano corresponde a um conjunto definido por dois pontos (a_1, b_1) e (a_2, b_2), onde

$$(a_1, b_1) \neq (a_2, b_2),$$

conforme imagem acima.

O *comprimento* deste segmento é

$$d((a_1, b_1), (a_2, b_2)) = \sqrt{(a_1 - a_2)^2 + (b_1 - b_2)^2},$$

conforme EXEMPLO 8.42.

Lembrando que

$$d((a_1, b_1), (a_2, b_2)) = \| (a_1, a_2) - (b_1, b_2) \|,$$

temos que

$$d((a_1, b_1), (a_2, b_2)) = \| (r, s) \|,$$

onde $r = a_1 - b_1$ e $s = a_2 - b_2$.

Ou seja, (r, s) é um vetor que corresponde a segmentos de reta dados por pontos (a_1, b_1) e (a_2, b_2), distintos entre si, com comprimento $\| (r, s) \|$ e tais que

$$(a_1 - a_2, b_1 - b_2) = (r, s).$$

Analogamente, pontos (c_1, d_1) e (c_2, d_2), distintos entre si, definem segmentos de reta de comprimento $\| (t, u) \|$, onde $t = c_1 - c_2$ e $u = d_1 - d_2$.

No entanto,

$$\| (r,s)+(t,u) \|^2 =$$
$$\| (r,s) \|^2 + \| (t,u) \|^2 + 2\langle (r,s),(t,u)\rangle .$$

Se interpretarmos $\| (r,s) \|$ e $\| (t,u) \|$ como comprimentos de catetos de um triângulo retângulo, a fórmula acima é o *Teorema de Pitágoras* se $\langle (r,s),(t,u)\rangle = 0$, onde a hipotenusa tem comprimento $\| (r,s)+(t,u) \|$.

Lembrando que catetos de um triângulo retângulo são perpendiculares entre si, a igualdade acima é o Teorema de Pitágoras se os vetores (r,s) e (t,u) forem ortogonais entre si, ou seja, $\langle (r,s),(t,u)\rangle = 0$.

Naturalmente, os vetores p e q do enunciado do teorema são (r,s) e (t,u), respectivamente.

Apesar da demonstração acima não ser rigorosa (uma vez que não enunciamos o Teorema de Pitágoras na Parte 7), esperamos que o leitor perceba a relação entre ortogonalidade e perpendicularismo, pelo menos no contexto do plano cartesiano. Mesmo assim, tudo o que é desenvolvido nesta Seção pode ser transposto para uma qualificação rigorosa dos conceitos envolvidos.

Na Seção 77 introduzimos uma definição para reta no plano cartesiano. No entanto, no contexto de álgebra linear, é possível expressar o mesmo conceito como um teorema.

Teorema 8.30. *Uma reta em $\mathbb{R}^2$ usual, munido do produto interno canônico, é o conjunto*

$$\{(x,y) \in \mathbb{R}^2 \mid ax+by=c\},$$

onde a, b e c são números reais tais que a e b não são simultaneamente nulos.

Demonstração: De acordo com os axiomas de incidência, uma reta pode ser definida em $\mathbb{R}^2$ usual por dois pontos. Isso equivale a afirmar que uma reta pode ser definida por um ponto e uma direção, no seguinte sentido.

Seja (a,b) um vetor não nulo de $\mathbb{R}^2$. Quaisquer pontos distintos (x_1,y_1) e (x_2,y_2) de uma reta perpendicular a (a,b) são tais que $\langle (x_2,y_2)-(x_1,y_1),(a,b)\rangle = 0$, de acordo com

Teorema 8.29. Neste sentido, a reta é definida, por exemplo, pelo ponto (x_1, y_1) e pela direção

$$(x_2, y_2) - (x_1, y_1)$$

perpendicular a (a, b).

Logo,

$$\langle (x_2 - x_1, y_2 - y_1), (a, b)\rangle = 0, \quad \text{ou seja,}$$

$$a(x_2 - x_1) + b(y_2 - y_1) = 0, \quad \text{que é equivalente a}$$

$$ax_2 + by_2 = ax_1 + by_1.$$

Portanto, dado um ponto (x_1, y_1), todos os pontos (x, y) tais que

$$(x, y) - (x_1, y_1)$$

é perpendicular a

$$(a, b),$$

são os pontos de uma mesma reta (a qual é perpendicular a (a, b)).

Em outras palavras, para quaisquer (x, y) desta reta, temos que

$$ax + by$$

é o mesmo valor real constante.

Se chamarmos tal constante de c, temos que

$$ax + by = c.$$

A razão para exigirmos que a e b não podem ser simultaneamente nulos é o Teorema 8.26: o vetor nulo é ortogonal a todo e qualquer vetor do espaço. Logo, $(0, 0)$ não pode definir uma reta no plano cartesiano.

Foi provado, portanto, que plano cartesiano e $\mathbb{R}^2$ usual munido de produto interno canônico são conceitos equivalentes.

Observar que os parâmetros a e b, na equação

$$ax + by = c,$$

definem a direção da reta, a qual deve ser perpendicular a (a, b). Já o parâmetro c permite localizar 'onde está' a reta, conforme discutido no próximo Exemplo.

Exemplo 8.50. *Se* $r = \{(x, y) \in \mathbb{R}^2 \mid ax + by = c\}$ *é uma reta, então*

I: $c = 0$ *implica que* r *passa pela origem* $(0, 0)$, *pois*
$$a(0) + b(0) = 0.$$

II: $a = 0$ *implica que* r *é uma reta* horizontal, *ou seja, paralela ao eixo* x *ou coincidente com ele; com efeito,*
$$y = \frac{c}{b},$$
onde $b \neq 0$, *uma vez que* $(a, b) \neq (0, 0)$, *como exige Teorema 8.30; logo, neste caso, a reta está a um distância*
$$\left|\frac{c}{b}\right|$$
do eixo x;

III: $b = 0$ *implica que* r *é uma reta* vertical, *ou seja, paralela ao eixo* y *ou coincidente com ele; com efeito,*
$$x = \frac{c}{a},$$
onde $a \neq 0$, *uma vez que* $(a, b) \neq (0, 0)$, *como exige Teorema 8.30; logo, neste caso, a reta está a um distância*
$$\left|\frac{c}{a}\right|$$
do eixo y;

IV: $b \neq 0$ *implica que* r *é uma reta não vertical; com efeito,*
$$y = -\frac{ax}{b} + \frac{c}{b},$$
onde $-\frac{a}{b}$ *é chamado de* coeficiente angular *de* r *e* $\frac{c}{b}$ *é conhecido como* coeficiente linear *de* r; *observar que neste caso a reta passa pelo ponto* $\left(0, \frac{c}{b}\right)$, *entre muitos outros.*

No contexto acima, uma reta r dada por
$$r = \{(x, y) \in \mathbb{R}^2 \mid ax + by = c\}$$
é paralela a uma reta
$$s = \{(x, y) \in \mathbb{R}^2 \mid a'x + b'y = c'\}$$
se, e somente se, existe λ real tal que
$$(a, b) = \lambda(a', b') \text{ e } c \neq \lambda c'.$$

Uma reta

$$t = \{(x, y) \in \mathbb{R}^2 \mid a''x + b''y = c''\}$$

é coincidente com r se, e somente se, existe λ real tal que

$$(a, b) = \lambda(a'', b'') \text{ e } c = \lambda c''.$$

Uma reta

$$u = \{(x, y) \in \mathbb{R}^2 \mid a'''x + b'''y = c'''\}$$

é perpendicular a r se, e somente se,

$$\langle (a, b), (a''', b''') \rangle = 0.$$

✍ Cabe ao leitor provar as três últimas afirmações, as quais são teoremas.

Seção 91

Transformações lineares

Sumário
Índice
Rede

Transformações lineares são funções, cujos domínios e codomínios são espaços vetoriais, que *preservam* a estrutura algébrica de tais espaços. Este conceito é tornado preciso a seguir.

Definição 8.16. *Sejam*

$$\mathcal{V} = \langle V, \mathbb{R}, +, \cdot, \overline{0} \rangle \quad e \quad \mathcal{W} = \langle W, \mathbb{R}, \oplus, \odot, \overline{\overline{0}} \rangle$$

espaços vetoriais reais. $\mathcal{T}$ *é uma* transformação linear *de* $\mathcal{V}$ *em* $\mathcal{W}$ *sss*

I: $\mathcal{T} : V \to W$ *é uma função;*

II: $\mathcal{T}(u + v) = \mathcal{T}(u) \oplus \mathcal{T}(v)$*;*

III: $\mathcal{T}(\alpha \cdot u) = \alpha \odot \mathcal{T}(u)$*.*

Ou seja, transformações lineares $\mathcal{T} : V \to W$ entre espaços vetoriais reais são funções que fazem o seguinte:

I: A cada vetor u de V, $\mathcal{T}(u)$ é um vetor de W.

II: Se $+$ é a adição de vetores em V e $\oplus$ é a adição de vetores em W, então os dois processos a seguir produzem o mesmo vetor de W:

(i) somar vetores u e v de V e, em seguida, calcular a imagem $\mathcal{T}(u+v)$ de $u+v$ relativamente à função $\mathcal{T}$ e

(ii) calcular as imagens $\mathcal{T}(u)$ e $\mathcal{T}(v)$ de u e v pela função $\mathcal{T}$ e somar tais imagens no espaço W, ou seja, $\mathcal{T}(u) \oplus \mathcal{T}(v)$.

iii: Se $\cdot$ é a multiplicação entre escalares reais α e vetores u de V, então os dois processos a seguir produzem o mesmo vetor de W:

(i) multiplicar α por u, obtendo $\alpha \cdot u$, e então calcular a imagem $\mathcal{T}(\alpha \cdot u)$ e

(ii) calcular a imagem $\mathcal{T}(u)$ de u e multiplicar α por $\mathcal{T}(u)$ no espaço W, ou seja, $\alpha \odot \mathcal{T}(u)$.

Em particular, se $\langle V, \mathbb{R}, +, \cdot, \overline{0} \rangle$ é um espaço vetorial real, então $f : V \to V$, tal que $f(u) = u$, é uma transformação linear do espaço vetorial nele mesmo.

Exemplo 8.51. *Considere o espaço vetorial real*

$$\mathcal{P}_2 = \langle P_2, \mathbb{R}, +, \cdot, \bigcirc \rangle$$

das funções polinomiais de grau menor ou igual a 2, conforme Exemplo 8.49. *Considere agora o espaço vetorial real* $\mathbb{R}^3$ *usual, ou seja,*

$$\mathcal{R}^3 = \langle \mathbb{R}^3, \mathbb{R}, +, \cdot, (0,0,0) \rangle.$$

Podemos definir a seguinte transformação linear $\mathcal{T} : P_2 \to \mathbb{R}^3$*: se*

$$p(x) = ax^2 + bx + c$$

é um vetor de P_2*, então*

$$\mathcal{T}(p) = (a, b, c).$$

Considerando o último Exemplo, item i da Definição 8.16 é trivialmente um teorema. Examinemos agora as demais exigências para transformações lineares.

Se $p : \mathbb{R} \to \mathbb{R}$ e $q : \mathbb{R} \to \mathbb{R}$ são funções tais que

$$p(x) = ax^2 + bx + c \quad \text{e} \quad q(x) = a'x^2 + b'x + c',$$

então $(p+q) : \mathbb{R} \to \mathbb{R}$ é uma função polinomial de grau menor ou igual a 2 tal que

$$(p+q)(x) = (ax^2 + bx + x) + (a'x^2 + b'x + c') =$$

$$(a+a')x^2 + (b+b')x + (c+c').$$

Neste caso,

$$\mathcal{T}(p+q) = (a+a', b+b', c+c'),$$

de acordo com a definição de $\mathcal{T}$ no EXEMPLO 8.51.

No entanto, $\mathcal{T}(p) = (a, b, c)$ e $\mathcal{T}(q) = (a', b', c')$. Logo,

$$\mathcal{T}(p) + \mathcal{T}(q) = (a+a', b+b', c+c').$$

Isso garante que item II da Definição 8.16 é teorema, uma vez que a transitividade da igualdade implica que $\mathcal{T}(p+q) = \mathcal{T}(p) + \mathcal{T}(q)$.

Finalmente, $\alpha \cdot p : \mathbb{R} \to \mathbb{R}$ é uma função polinomial de grau menor ou igual a 2 tal que

$$(\alpha \cdot p)(x) = \alpha(ax^2 + bx + c) = \alpha a x^2 + \alpha b x + \alpha c,$$

o que implica que

$$\mathcal{T}(\alpha \cdot p) = (\alpha a, \alpha b, \alpha c).$$

Porém, uma vez que $\mathcal{T}(p) = (a, b, c)$, então

$$\alpha \cdot \mathcal{T}(p) = \alpha \cdot (a, b, c) = (\alpha a, \alpha b, \alpha c).$$

Logo, item III da Definição 8.16 também é teorema.

Isso conclui a demonstração de que EXEMPLO 8.51 de fato ilustra uma transformação linear do espaço das funções polinomiais de grau menor ou igual a 2 no espaço vetorial real $\mathbb{R}^3$ usual.

Notar também que as coordenadas de qualquer vetor (a, b, c) de $\mathbb{R}^3$ usual, relativamente à *base canônica* $\{(1,0,0), (0,1,0), (0,0,1)\}$, são (a, b, c). Além disso, as coordenadas de qualquer vetor

$$p(x) = ax^2 + bx + c$$

do espaço P_2, relativamente à base ortonormal do EXEMPLO 8.51, são igualmente (a, b, c).

EXEMPLO 8.52. *Seguindo* EXEMPLO 8.51, *seja agora*

$$\mathcal{U} : P_2 \to \mathbb{R}^3$$

uma função tal que, para cada $p \in P_2$ *onde*

$$p(x) = ax^2 + bx + c,$$

temos

$$\mathcal{U}(p) = (2a - 3c, 0, 2b).$$

Neste caso, $\mathcal{U}$ *também é uma transformação linear.*

Considerando este novo Exemplo, item i da Definição 8.16 é trivialmente um teorema. Examinemos agora as demais exigências para transformações lineares.

Se $p : \mathbb{R} \to \mathbb{R}$ e $q : \mathbb{R} \to \mathbb{R}$ são funções tais que

$$p(x) = ax^2 + bx + c$$

e

$$q(x) = a'x^2 + b'x + c',$$

então $(p + q) : \mathbb{R} \to \mathbb{R}$ é uma função polinomial de grau menor ou igual a 2 tal que

$$(p + q)(x) =$$
$$(ax^2 + bx + x) + (a'x^2 + b'x + c') = (a + a')x^2 + (b + b')x + (c + c').$$

Neste caso,

$$\mathcal{T}(p + q) = (2(a + a') - 3(c + c'), 0, 2(b + b')),$$

de acordo com a definição de $\mathcal{T}$ no Exemplo 8.51.

No entanto, $\mathcal{T}(p) = (2a - 3c, 0, 2b)$ e $\mathcal{T}(q) = (2a' - 3c', 0, 2b')$. Logo,

$$\mathcal{T}(p) + \mathcal{T}(q) = (2a - 3c, 0, 2b) + (2a' - 3c', 0, 2b') =$$
$$(2a - 3c + 2a' - 3c', 0 + 0, 2b + 2b') =$$
$$(2(a + a') - 3(c + c'), 0, 2(b + b')).$$

Isso garante que item ii da Definição 8.16 é teorema, uma vez que a transitividade da igualdade implica que $\mathcal{T}(p + q) = \mathcal{T}(p) + \mathcal{T}(q)$.

Finalmente, $\alpha \cdot p : \mathbb{R} \to \mathbb{R}$ é uma função polinomial de grau menor ou igual a 2 tal que

$$(\alpha \cdot p)(x) = \alpha(ax^2 + bx + c) = \alpha ax^2 + \alpha bx + \alpha c,$$

o que implica que

$$\mathcal{T}(\alpha \cdot p) = (2\alpha a - 3\alpha c, 0, 2\alpha b).$$

Porém, uma vez que $\mathcal{T}(p) = (2a - 3c, 0, 2b)$, então

$$\alpha \cdot \mathcal{T}(p) = \alpha \cdot (2a - 3c, 0, 2b) = (2\alpha a - 3\alpha c, 0, 2\alpha b).$$

Logo, item iii da Definição 8.16 também é teorema.

Isso conclui a demonstração de que Exemplo 8.52 também ilustra uma transformação linear do espaço das funções polinomiais de grau menor ou igual a 2 no espaço vetorial real $\mathbb{R}^3$ usual.

Exemplo 8.53. *Seguindo* Exemplos 8.51 *e* 8.52, *seja agora*

$$\mathcal{R} : P_2 \to \mathbb{R}^3$$

uma função tal que, para cada $p \in P_2$ *onde*

$$p(x) = ax^2 + bx + c,$$

temos

$$\mathcal{R}(p) = (2a - 3c, 7, 2b).$$

✍ *Neste caso,* $\mathcal{R}$ *não é uma transformação linear. Com efeito, itens* II *e* III *da Definição 8.16 não são teoremas. Cabe ao leitor justificar.*

Exemplo 8.54. ✍ *Em* $\mathbb{R}^3$ *usual, seja* $f : \mathbb{R}^3 \to \mathbb{R}^3$ *tal que*

$$f(x, y, z) = (\alpha x + \beta y + \gamma z, \delta x + \varepsilon y + \zeta z, \eta x + \theta y + \kappa z).$$

Logo, f *é uma transformação linear de* $\mathbb{R}^3$ *usual em* $\mathbb{R}^3$ *usual, para quaisquer reais* α, β, γ, δ, ε, ζ, η, θ *e* κ. *Recomendamos que o leitor prove isso.*

Teorema 8.31. *Sejam*

$$\mathcal{V} = \langle V, \mathbb{R}, +, \cdot, \overline{0} \rangle \ \text{ e } \ \mathcal{W} = \langle W, \mathbb{R}, \oplus, \odot, \overline{\overline{0}} \rangle$$

espaços vetoriais reais. Se $\mathcal{T} : V \to W$ *é uma transformação linear, então*

$$\mathcal{T}(\overline{0}) = \overline{\overline{0}}.$$

Demonstração: Teorema 8.4 garante que, para qualquer vetor $v \in V$,

$$\mathcal{T}(\overline{0}) = \mathcal{T}(0 \cdot v).$$

Mas item III da Definição 8.16 exige que a fórmula

$$\mathcal{T}(\alpha \cdot u) = \alpha \odot \mathcal{T}(u)$$

seja teorema. Portanto,

$$\mathcal{T}(0 \cdot v) = 0 \odot \mathcal{T}(v).$$

Uma vez que $\mathcal{T}(v)$ é um vetor de um espaço vetorial real, então Teorema 8.4 garante que

$$0 \odot \mathcal{T}(v) = \overline{\overline{0}}.$$

Exemplo 8.55. *Teorema 8.31 garante que a função* $\mathcal{R}$ *do* Exemplo 8.53 *não é uma transformação linear. Com efeito, se* $\bigcirc : \mathbb{R} \to \mathbb{R}$*, dada por*

$$\bigcirc(x) = 0,$$

é o vetor nulo de P_2*, então* $\mathcal{R}(\bigcirc) = (0, 7, 0)$ *é diferente do vetor nulo* $(0, 0, 0)$ *de* $\mathbb{R}^3$ *usual.*

Se um espaço vetorial real $\langle V, \mathbb{R}, +, \cdot, \bar{0}\,\rangle$ admite base finita b, qualquer vetor $v \in V$ é uma combinação linear única dos elementos de b (Teorema 8.12). Graças a isso, transformações lineares $\mathcal{T}$ podem ser univocamente determinadas a partir de imagens dos elementos de b relativamente a $\mathcal{T}$, como se percebe no próximo teorema.

Teorema 8.32. *Sejam*

$$\mathcal{V} = \langle V, \mathbb{R}, +, \cdot, \bar{0}\,\rangle \;\; e \;\; \mathcal{W} = \langle W, \mathbb{R}, \oplus, \odot, \bar{\bar{0}}\,\rangle$$

espaços vetoriais reais e

$$b = \{v_1, v_2, \cdots, v_n\}$$

uma base de $\mathcal{V}$*.*

Se $\mathcal{R} : b \to W$ *é uma função, então existe uma única transformação linear*

$$\mathcal{T} : V \to W$$

tal que $\mathcal{R}$ *é restrição de* $\mathcal{T}$*.*

Demonstração: Se $\mathcal{R} : b \to W$ é uma função, então

$$\mathcal{R}(v_i) = w_i,$$

para todo i tal que $1 \leq i \leq n$, sendo cada w_i pertencente a W.

Se $b = \{v_1, v_2, \cdots, v_n\}$ é uma base de $\mathcal{V}$, então cada vetor v de V é dado por

$$v = \sum_{i=1}^{n} \alpha_i v_i.$$

Se $\mathcal{R}$ é uma restrição de $\mathcal{T}$, então

$$\mathcal{T}(v_i) = w_i,$$

para todo i tal que $1 \leq i \leq n$.

Se $\mathcal{T}$ é uma transformação linear de $\mathcal{V}$ em $\mathcal{W}$, então

$$\mathcal{T}(v) = \mathcal{T}\left(\sum_{i=1}^{n} \alpha_i v_i\right) = \sum_{i=1}^{n} \mathcal{T}(\alpha_i v_i) = \sum_{i=1}^{n} \alpha_i \mathcal{T}(v_i),$$

graças aos itens II e III da Definição 8.16. Logo, basta conhecer as imagens $\mathcal{T}(v_i)$ para definir as imagens de uma transformação linear

$$\mathcal{T} : \mathcal{V} \to \mathcal{W}.$$

Ou seja, apesar de poder existir uma infinidade de funções

$$\mathcal{T} : V \to W$$

tais que $\mathcal{R}$ é restrição de $\mathcal{T}$, apenas uma delas é transformação linear.

Cada função $\mathcal{R} : b \to W$ define uma e apenas uma transformação linear de $\mathcal{V}$ em $\mathcal{W}$.

Exemplo 8.56. *Seja $\mathcal{R} : \{(1,0),(0,1)\} \to \mathbb{R}^3$ uma função tal que*

$$\mathcal{R}(1,0) = \mathcal{R}(0,1) = (5,7,9),$$

onde $\{(1,0),(0,1)\}$ é a base canônica de $\mathbb{R}^2$ usual e $\mathbb{R}^3$ é o espaço de vetores de $\mathbb{R}^3$ usual.

Logo, existe uma única transformação linear $\mathcal{T} : \mathbb{R}^2 \to \mathbb{R}^3$ tal que $\mathcal{R}$ é restrição de $\mathcal{T}$. Com efeito,

$$\mathcal{T}(x,y) = \mathcal{T}(x(1,0) + y(0,1)) = \mathcal{T}(x(1,0)) + \mathcal{T}(y(0,1)) =$$
$$x\mathcal{T}(1,0) + y\mathcal{T}(0,1) = x(5,7,9) + y(5,7,9) =$$
$$(5x,7x,9x) + (5y,7y,9y) = (5x+5y, 7x+7y, 9x+9y).$$

✍ Recomendamos que o leitor faça uma versão do Exemplo acima na qual seja trocada apenas a função $\mathcal{R}$, substituindo-a por uma função injetiva.

Os próximos teoremas mostram que transformações lineares também podem ser interpretadas como vetores.

Teorema 8.33. *Sejam*

$$\mathcal{V} = \langle V, \mathbb{R}, +, \cdot, \bar{0} \rangle \;\; e \;\; \mathcal{W} = \langle W, \mathbb{R}, \oplus, \odot, \bar{\bar{0}} \rangle$$

espaços vetoriais reais.

A função $\mathcal{T} : V \to W$ dada por $\mathcal{T}(u) = \bar{\bar{0}}$ é uma transformação linear de $\mathcal{V}$ em $\mathcal{W}$.

Demonstração: Item i da Definição 8.16 é imediato.

Sobre item ii, notar que

$$\mathcal{T}(u+v)=\bar{\bar{0}},\ \ \mathcal{T}(u)=\bar{\bar{0}}\ \text{ e }\ \mathcal{T}(v)=\bar{\bar{0}}.$$

Logo, de acordo com axioma V5 da Definição 8.1,

$$\mathcal{T}(u+v)=\mathcal{T}(u)\oplus\mathcal{T}(v).$$

Finalmente, sobre item iii, Teorema 8.6 garante que

$$\alpha\odot\mathcal{T}(u)=\bar{\bar{0}},$$

uma vez que $\mathcal{T}(u)=\bar{\bar{0}}$.

Logo, lembrando que $\mathcal{T}(\alpha\cdot u)=\bar{\bar{0}}$, temos que

$$\mathcal{T}(\alpha\cdot u)=\alpha\odot\mathcal{T}(u)$$

também é teorema.

Provamos acima que uma função constante (cuja constante é um vetor nulo), com domínio em um espaço de vetores, é uma transformação linear.

A seguir mostramos que escalar vezes transformação linear é uma transformação linear, desde que essa operação seja definida a partir da estrutura algébrica do co-domínio, o qual é um espaço de vetores.

Teorema 8.34. *Sejam*

$$\mathcal{V}=\langle V,\mathbb{R},+,\cdot,\bar{0}\rangle\ \ e\ \ \mathcal{W}=\langle W,\mathbb{R},\oplus,\odot,\bar{\bar{0}}\rangle$$

espaços vetoriais reais, α um escalar e $\mathcal{T}:V\to W$ uma transformação linear. Então $\alpha\boxdot\mathcal{T}:V\to W$, dada por

$$\alpha\boxdot\mathcal{T}(u)=\alpha\odot\mathcal{T}(u),$$

é uma transformação linear.

Demonstração: Item i da Definição 8.16 é trivial. Para item ii, notar que

$$\alpha\boxdot\mathcal{T}(u+v)=\alpha\odot\mathcal{T}(u+v)=\alpha\odot(\mathcal{T}(u)\oplus\mathcal{T}(v))=$$
$$\alpha\odot\mathcal{T}(u)\oplus\alpha\odot\mathcal{T}(v)=\alpha\boxdot\mathcal{T}(u)\oplus\alpha\boxdot\mathcal{T}(v).$$

Na sequência acima de quatro ocorrências da igualdade, a primeira é justificada pela definição de multiplicação de escalar por transformação linear, assumida no enunciado do teorema. A segunda é decorrente da hipótese de que $\mathcal{T}$ é uma transformação linear. A terceira é consequência do axioma V9 da Definição 8.1. Na quarta e última novamente é usada a definição de multiplicação de escalar por transformação linear.

Sobre item III da Definição 8.16, temos que

$$\alpha \boxdot \mathcal{T}(\beta \cdot u) = \alpha \odot \mathcal{T}(\beta \cdot u) =$$
$$\alpha \odot (\beta \odot \mathcal{T}(u)) = (\alpha\beta) \odot \mathcal{T}(u) =$$
$$(\beta\alpha) \odot \mathcal{T}(u) = \beta \odot (\alpha \odot \mathcal{T}(u)) = \beta \odot (\alpha \boxdot \mathcal{T}(u)),$$

onde β é um escalar.

✍ Cabe ao leitor justificar cada uma das seis últimas ocorrências da igualdade.

No teorema a seguir é provado que a soma de transformações lineares, definidas sobre um mesmo domínio, é também uma transformação linear, desde que essa operação seja definida a partir da estrutura algébrica do co-domínio, o qual é um espaço de vetores.

Teorema 8.35. *Sejam*

$$\mathcal{V} = \langle V, \mathbb{R}, +, \cdot, \overline{0}\,\rangle \ \ e \ \ \mathcal{W} = \langle W, \mathbb{R}, \oplus, \odot, \overline{\overline{0}}\,\rangle$$

espaços vetoriais reais.

Sejam

$$\mathcal{R} : V \to W \ \ e \ \ \mathcal{T} : V \to W$$

transformações lineares.

Logo, $\mathcal{R} \boxplus \mathcal{T} : V \to W$, definida como

$$\mathcal{R} \boxplus \mathcal{T}(u) = \mathcal{R}(u) \oplus \mathcal{T}(u),$$

é uma transformação linear.

Demonstração: Item I da Definição 8.16 é trivial. Para item II, temos que

$$\mathcal{R} \boxplus \mathcal{T}(u+v) = \mathcal{R}(u+v) \oplus \mathcal{T}(u+v) =$$
$$(\mathcal{R}(u) \oplus \mathcal{R}(v)) \oplus (\mathcal{T}(u) \oplus \mathcal{T}(v)) =$$
$$(\mathcal{R}(u) \oplus \mathcal{T}(u)) \oplus (\mathcal{R}(v) \oplus \mathcal{T}(v)) =$$
$$\mathcal{R} \boxplus \mathcal{T}(u) \oplus \mathcal{R} \boxplus \mathcal{T}(v).$$

Com relação a item III da Definição 8.16, observar que

$$\mathcal{R} \boxplus \mathcal{T}(\alpha \cdot u) =$$
$$\mathcal{R}(\alpha \cdot u) \oplus \mathcal{T}(\alpha \cdot u) = \alpha \odot \mathcal{R}(u) \oplus \alpha \odot \mathcal{T}(u) =$$
$$\alpha \odot (\mathcal{R}(u) \oplus \mathcal{T}(u)) = \alpha \odot \mathcal{R} \boxplus \mathcal{T}(u),$$

onde α é um escalar.

Logo, $\mathcal{R} \boxplus \mathcal{T}$ é uma transformação linear.

✍ Cabe ao leitor justificar cada um dos passos da prova.

Teorema 8.36. *Sejam*

$$\mathcal{V} = \langle V, \mathbb{R}, +, \cdot, \overline{0} \rangle \ \ e \ \ \mathcal{W} = \langle W, \mathbb{R}, \oplus, \odot, \overline{\overline{0}} \rangle$$

espaços vetoriais reais.

Seja $\mathfrak{T}$ o conjunto de todas as transformações lineares de $\mathcal{V}$ em $\mathcal{W}$. Logo, $\langle \mathfrak{T}, \mathbb{R}, \boxplus, \boxdot, \bigcirc \rangle$ é um espaço vetorial real, onde

I: *$\boxplus$ é a adição de transformações lineares usada no Teorema 8.35,*

II: *$\boxdot$ é a multiplicação entre escalar e transformação linear usada no Teorema 8.34 e*

III: *$\bigcirc : V \to W$ é a função dada por $\bigcirc(u) = \overline{\overline{0}}$.*

Demonstração: ✍ A prova de que axiomas V1~V12 da Definição 8.1 são teoremas nesta interpretação fica muito facilitada, graças aos teoremas 8.33, 8.34 e 8.35.

Detalhes ficam para o leitor.

Exemplo 8.57. *Continuando* Exemplo 8.54, *toda transformação linear de $\mathbb{R}^3$ usual em $\mathbb{R}^3$ usual é uma função $f : \mathbb{R}^3 \to \mathbb{R}^3$ tal que*

$$f(x, y, z) = (\alpha x + \beta y + \gamma z, \delta x + \varepsilon y + \zeta z, \eta x + \theta y + \kappa z),$$

onde α, β, γ, δ, ε, ζ, η, θ e κ são reais.

Podemos reescrever isso na forma matricial, como se segue.

$$\begin{pmatrix} \alpha & \beta & \gamma \\ \delta & \varepsilon & \zeta \\ \eta & \theta & \kappa \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \alpha x + \beta y + \gamma z \\ \delta x + \varepsilon y + \zeta z \\ \eta x + \theta y + \kappa z \end{pmatrix},$$

sendo que a operação acima é a multiplicação usual entre matrizes reais (ver Definição 8.17 imediatamente abaixo, para recordar). Cada matriz

$$\begin{pmatrix} \alpha & \beta & \gamma \\ \delta & \varepsilon & \zeta \\ \eta & \theta & \kappa \end{pmatrix}$$

corresponde a uma transformação linear. Se todas as entradas forem nulas, temos a transformação linear $\bigcirc$ do último teorema. Logo, o espaço vetorial real de todas as transformações lineares de $\mathbb{R}^3$ usual em $\mathbb{R}^3$ usual tem nove dimensões.

Se o leitor não recorda o conceito de multiplicação matricial usual, aqui vai.

Definição 8.17. *Sejam*

$$a : l_m \times c_n \to \mathbb{R}$$

e

$$b : l'_n \times c'_p \to \mathbb{R}$$

matrizes reais, de acordo com a Definição 8.2.

Ou seja, o número de colunas da matriz a coincide com o número de linhas da matriz b. O produto de a por b é uma matriz

$$c : l_m \times c'_p \to \mathbb{R}$$

tal que cada entrada c_{ij} da matriz c é dada por

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}.$$

Exemplo 8.58. *Ver multiplicação matricial do* Exemplo 8.57.

O espaço vetorial real das transformações lineares de $\mathbb{R}^3$ usual em $\mathbb{R}^3$ usual do Exemplo 8.57 é simplesmente o espaço $M_{3\times 3}$ usual.

A representação matricial de transformações lineares entre espaços vetoriais reais de dimensão finita depende das bases ordenadas (Definição 8.8) utilizadas para os espaços vetoriais reais envolvidos. No Exemplo 8.57 utilizamos base canônica para $\mathbb{R}^3$ usual. Não avançamos sobre este importante tópico aqui.

Seção 92

Imagem de uma transformação linear

Sumário
Índice
Rede

Provamos aqui que transformações lineares definem subespaços de seus contradomínios.

Definição 8.18. *Sejam*

$$\mathcal{V} = \langle V, \mathbb{R}, +, \cdot, \overline{0}\,\rangle$$

e

$$\mathcal{W} = \langle W, \mathbb{R}, \oplus, \odot, \overline{\overline{0}}\,\rangle$$

espaços vetoriais reais. Se $\mathcal{T} : V \to W$ *é uma transformação linear, dizemos que*

$$\mathfrak{Im}(\mathcal{T}) = \{w \in W \mid \exists v(v \in V \wedge \mathcal{T}(v) = w)\}$$

é a imagem *de* $\mathcal{T}$.

Ou seja, a imagem de uma transformação linear $\mathcal{T}$ é o conjunto dos vetores w tais que $w = \mathcal{T}(v)$, para algum v do domínio de $\mathcal{T}$.

Exemplo 8.59. *Considere o espaço vetorial real*

$$\mathcal{P}_2 = \langle P_2, \mathbb{R}, +, \cdot, \bigcirc\rangle$$

das funções polinomiais de grau menor ou igual a 2, conforme Exemplo 8.49. *Seja* $\mathcal{D} : \mathcal{P}_2 \to \mathcal{P}_2$ *a função dada por*

$$\mathcal{D}(p(x)) = \frac{d}{dx}p(x),$$

onde $p(x)$ *é um vetor de* $\mathcal{P}_2$.

Observar que $\mathcal{D}$ *é uma transformação linear. Com efeito, item* I *da Definição 8.16 é imediato; item* II *decorre do fato de que derivada da soma é a soma de derivadas (Teorema 5.21); item* III *é consequência do fato de que derivada de constante vezes função é constante vezes derivada da função (Teorema 5.20).*

Se (a, b, c) *são as coordenadas de* $p(x)$ *relativamente à base ordenada canônica de* $\mathcal{P}_2$, *então* $\mathcal{D}(p(x))$ *é um vetor com coordenadas* $(0, 2a, b)$. *Com efeito,*

$$\frac{d}{dx}(ax^2 + bx + c) = 0 + 2ax + b.$$

Neste caso, $\mathfrak{Im}(\mathcal{D})$ é o conjunto das funções polinomiais de grau menor ou igual a 1.

No Exemplo acima a imagem da transformação linear $\mathcal{D}$ define um subespaço do contradomínio de $\mathcal{D}$. Isso não é coincidência, como se percebe no próximo teorema.

Teorema 8.37. *Sejam*

$$\mathcal{V} = \langle V, \mathbb{R}, +, \cdot, \overline{0}\,\rangle$$

e

$$\mathcal{W} = \langle W, \mathbb{R}, \oplus, \odot, \overline{\overline{0}}\,\rangle$$

espaços vetoriais reais. Se

$$\mathcal{T} : V \to W$$

é uma transformação linear, então

$$\langle \mathfrak{Im}(\mathcal{T}), \mathbb{R}, \oplus', \odot', \overline{\overline{0}}\rangle$$

é subespaço de $\mathcal{W}$, onde $\oplus'$ e $\odot'$ são restrições de $\oplus$ e $\odot$, respectivamente.

Demonstração: Item I da Definição 8.3 é imediato.

Item II é consequência do Teorema 8.31.

Sobre item III, se w_1 e w_2 pertencem a $\mathfrak{Im}(\mathcal{T})$, então existem v_1 e v_2 tais que $\mathcal{T}(v_1) = w_1$ e $\mathcal{T}(v_2) = w_2$. Mas

$$\mathcal{T}(v_1 + v_2) = \mathcal{T}(v_1) \oplus \mathcal{T}(v_2) = w_1 \oplus w_2.$$

Portanto, $w_1 \oplus w_2$ pertence a $\mathfrak{Im}(\mathcal{T})$.

Para finalizar, se $w \in \mathfrak{Im}(\mathcal{T})$, então existe $v \in V$ tal que $\mathcal{T}(v) = w$. Mas

$$\mathcal{T}(\alpha \cdot v) = \alpha \odot \mathcal{T}(v) = \alpha \odot w,$$

se α é um escalar. Logo, $\alpha \odot w$ pertence a $\mathfrak{Im}(\mathcal{T})$.

Seção 93

Núcleo de uma transformação linear

Sumário
Índice
Rede

Provamos aqui que transformações lineares definem subespaços de seus domínios.

Definição 8.19. *Sejam*

$$\mathcal{V} = \langle V, \mathbb{R}, +, \cdot, \overline{0}\,\rangle \;\; e \;\; \mathcal{W} = \langle W, \mathbb{R}, \oplus, \odot, \overline{\overline{0}}\,\rangle$$

espaços vetoriais reais. Se

$$\mathcal{T} : V \to W$$

é uma transformação linear, dizemos que

$$\mathfrak{N}(\mathcal{T}) = \{v \in V \mid \mathcal{T}(v) = \overline{\overline{0}}\}$$

é o núcleo *de* $\mathcal{T}$*.*

Ou seja, o núcleo de uma transformação linear $\mathcal{T}$ é o conjunto de todos os vetores v do domínio de $\mathcal{T}$ tais que $\mathcal{T}(v)$ é o vetor nulo do co-domínio de $\mathcal{T}$.

Exemplo 8.60. *Considere o espaço vetorial real*

$$\mathcal{P}_2 = \langle P_2, \mathbb{R}, +, \cdot, \bigcirc\rangle$$

das funções polinomiais de grau menor ou igual a 2, conforme Exemplo 8.49*. Seja* $\mathcal{D} : \mathcal{P}_2 \to \mathcal{P}_2$ *a função dada por*

$$\mathcal{D}(p(x)) = \frac{d}{dx}p(x),$$

onde $p(x)$ *é um vetor de* $\mathcal{P}_2$*. Como visto no* Exemplo 8.59*,* $\mathcal{D}$ *é uma transformação linear.*

O núcleo de $\mathcal{D}$ *é o conjunto das funções reais constantes. Com efeito, se* $p(x) = c$*, então*

$$\frac{d}{dx}p(x) = 0,$$

ou seja, a derivada de qualquer função constante é a função identicamente nula $\bigcirc$*.*

No Exemplo acima o núcleo da transformação linear $\mathcal{D}$ define um subespaço do domínio de $\mathcal{D}$. Isso não é coincidência:

Teorema 8.38. *Sejam*

$$\mathcal{V} = \langle V, \mathbb{R}, +, \cdot, \overline{0}\,\rangle$$

e

$$\mathcal{W} = \langle W, \mathbb{R}, \oplus, \odot, \overline{\overline{0}}\,\rangle$$

espaços vetoriais reais.

Se $\mathcal{T} : V \to W$ *é uma transformação linear, então*

$$\langle \mathfrak{N}(\mathcal{T}), \mathbb{R}, \oplus', \odot', \bar{\bar{0}} \rangle$$

é subespaço de $\mathcal{W}$*, onde* $\oplus'$ *e* $\odot'$ *são restrições de* $\oplus$ *e* $\odot$*, respectivamente.*

Demonstração: Item i da Definição 8.3 é imediato.

Item ii é consequência do Teorema 8.31.

Sobre item iii, se v_1 e v_2 pertencem a $\mathfrak{N}(\mathcal{T})$, então $\mathcal{T}(v_1) = \mathcal{T}(v_2) = \bar{\bar{0}}$. Mas

$$\mathcal{T}(v_1 + v_2) = \mathcal{T}(v_1) \oplus \mathcal{T}(v_2) = \bar{\bar{0}} \oplus \bar{\bar{0}} = \bar{\bar{0}}.$$

Logo, $v_1 + v_2$ pertence ao núcleo de $\mathcal{T}$.

Finalmente, com relação ao item iv da Definição 8.3, se $v \in \mathfrak{N}(\mathcal{T})$, então $\mathcal{T}(v) = \bar{\bar{0}}$. Mas

$$\mathcal{T}(\alpha \cdot v) = \alpha \odot \mathcal{T}(v) = \alpha \odot \bar{\bar{0}} = \bar{\bar{0}}.$$

Logo, $\alpha \cdot v$ pertence ao núcleo de $\mathcal{T}$.

Importante notar os seguintes fatos sobre os Exemplos 8.59 e 8.60.

- O espaço vetorial real das funções polinomiais de grau menor ou igual a 2 tem três dimensões;
- O subespaço $\mathfrak{Im}(\mathcal{D})$ tem duas dimensões;
- O subespaço $\mathfrak{N}(\mathcal{D})$ tem uma dimensão;
- $2+1 = 3$, onde $2 = \dim(\mathfrak{Im}(\mathcal{D}))$, $1 = \dim(\mathfrak{N}(\mathcal{D}))$ e $3 = \dim(\mathcal{P}_2)$.

Novamente isso não é mera coincidência, como se percebe no próximo resultado.

Teorema 8.39 (Núcleo e Imagem). *Sejam*

$$\mathcal{V} = \langle V, \mathbb{R}, +, \cdot, \bar{0} \rangle$$

e

$$\mathcal{W} = \langle W, \mathbb{R}, \oplus, \odot, \bar{\bar{0}} \rangle$$

espaços vetoriais reais, onde $\mathcal{V}$ *tem* n *dimensões. Se* $\mathcal{T} : V \to W$ *é uma transformação linear, então*

$$\dim(\mathfrak{N}(\mathcal{T})) + \dim(\mathfrak{Im}(\mathcal{T})) = n.$$

Demonstração: ✍ Basta provar que, se

$$\{v_1, v_2, \cdots, v_p\}$$

é uma base para $\mathfrak{N}(\mathcal{T})$ e

$$\{\mathcal{T}(u_1), \mathcal{T}(u_2), \cdots, \mathcal{T}(u_q)\}$$

é uma base para $\mathfrak{Im}(\mathcal{T})$, então

$$\{v_1, v_2, \cdots, v_p, u_1, u_2, \cdots, u_q\}$$

é uma base para $\mathcal{V}$.

Uma vez que $\dim(\mathcal{V}) = n$, segue que $p+q=n$. Deixamos a prova para o leitor.

O teorema acima é o célebre *Teorema do Núcleo e Imagem*.

Uma generalização considerável deste resultado, pelo menos para certos operadores lineares, é o Teorema de Atiyah–Singer [45], o qual se refere ao *índice analítico* de *operadores diferenciais elípticos*. Neste contexto, índices analíticos estão intimamente relacionados com a dimensão de um espaço vetorial real cujos vetores são funções que são soluções de uma dada equação diferencial.

Levando em conta as discussões nesta Seção e na anterior, se uma transformação linear $\mathcal{T}$ é injetiva, então seu núcleo é o subespaço trivial do domínio de $\mathcal{T}$ formado apenas pelo vetor nulo do domínio. Se $\mathcal{T}$ for sobrejetiva, sua imagem é o subespaço trivial do co-domínio de $\mathcal{T}$ formado por todos os vetores deste co-domínio.

Seção 94

Operadores lineares

Sumário
Índice
Rede

O*peradores lineares* são transformações lineares nas quais domínio e contradomínio são o mesmo espaço de vetores.

Definição 8.20. *Seja $\mathcal{V} = \langle V, \mathbb{R}, +, \cdot, \overline{0}\rangle$ um espaço vetorial real. Uma transformação linear*

$$\mathcal{T} : V \to V$$

é um operador linear.

Exemplo 8.61. *A transformação linear* $\mathcal{D}$ *dos* Exemplos 8.59 *e* 8.60 *é um operador linear.*

Exemplo 8.62. *Seja* $\mathfrak{F} = \langle \mathcal{C}^\infty, \mathbb{R}, +, \cdot, \bigcirc \rangle$ *o espaço vetorial real onde*

- $\mathcal{C}^\infty$ *é o conjunto de todas as funções reais diferenciáveis um número arbitrário de vezes;*
- $+ : \mathcal{C}^\infty \times \mathcal{C}^\infty \to \mathcal{C}^\infty$ *é uma função dada por*
$$+(f, g)(x) = f(x) + g(x);$$
- $\cdot : \mathbb{R} \times \mathcal{C}^\infty \to \mathcal{C}^\infty$ *é uma função dada por*
$$\cdot(\alpha, f)(x) = \alpha f(x);$$
- $\bigcirc$ *é a função real* $\bigcirc : \mathbb{R} \to \mathbb{R}$ *tal que* $\bigcirc(x) = 0$ *para todo real* x.

✍ *Logo, qualquer operador diferencial (ver Seção 53) definido sobre* $\mathcal{C}^\infty$ *é um operador linear sobre* $\mathfrak{F}$.

Para efeitos práticos, isso significa que teoremas de álgebra linear sobre operadores lineares e espaços vetoriais reais sobre os quais eles atuam encontram repercussão no estudo de equações diferenciais.

Seção 95

Autovalores e autovetores

Sumário
Índice
Rede

Se $\mathcal{V} = \langle V, \mathbb{R}, +, \cdot, \overline{0} \rangle$ é um espaço vetorial real e
$$\mathcal{T} : V \to V$$
é um operador linear sobre V, um subespaço
$$\mathcal{W} = \langle W, \mathbb{R}, \oplus, \odot, \overline{\overline{0}} \rangle$$
de $\mathcal{V}$ é *invariante* sob a ação de $\mathcal{T}$ se, e somente se, para qualquer vetor w pertencente a W, temos $\mathcal{T}(w)$ pertencente a W.

Isso equivale a afirmar que a imagem da restrição de $\mathcal{T}$ a W está contida em W.

Naturalmente, todo espaço vetorial real é invariante sob a ação de qualquer operador linear definido sobre ele. Por conta disso, estamos interessado apenas nos casos não triviais. A próxima definição sugere a investigação de subespaços de uma dimensão que sejam invariantes sob a ação de operadores lineares.

Definição 8.21. *Sejam*

$$\mathcal{V} = \langle V, \mathbb{R}, +, \cdot, \bar{0} \rangle$$

um espaço vetorial real e $\mathcal{T} : V \to V$ *um operador linear sobre* V*. Um vetor não nulo de* V *é um* autovetor *de* $\mathcal{T}$ *sss existe* λ *real tal que*

$$\mathcal{T}(v) = \lambda \cdot v.$$

Referimo-nos ao real λ *como* autovalor *do operador linear* $\mathcal{T}$*.*

A equação acima é conhecida como equação de autovalores*.*

Mais adiante vemos alguns exemplos de operadores lineares que admitem autovetores (e, consequentemente, autovalores), bem como exemplos que não admitem. No entanto, antes disso, é útil compreender o próximo teorema.

Teorema 8.40. *Se um operador linear* $\mathcal{T}$ *admite autovetor* v*, então qualquer combinação linear não nula de* v *também é um autovetor de* $\mathcal{T}$*.*

Demonstração: Por hipótese, sabemos que existe autovalor λ de $\mathcal{T}$ tal que

$$\mathcal{T}(v) = \lambda \cdot v,$$

onde $\cdot$ é a multiplicação de escalar por vetor no espaço vetorial real em questão.

Seja w uma combinação linear não nula de v, ou seja, $w = \alpha \cdot v$, onde $\alpha \neq 0$ (lembrar que o vetor nulo jamais é autovetor de operador linear algum). Logo,

$$\mathcal{T}(w) = \mathcal{T}(\alpha \cdot v) = \alpha \cdot \mathcal{T}(v) = \alpha \cdot (\lambda \cdot v) =$$
$$(\alpha\lambda) \cdot v = (\lambda\alpha) \cdot v = \lambda \cdot (\alpha \cdot v).$$

Em outras palavras,

$$\mathcal{T}(\alpha \cdot v) = \lambda \cdot (\alpha \cdot v),$$

o que faz de $\alpha \cdot v$ um autovetor de $\mathcal{T}$ com o mesmo autovalor.

Ou seja, se um operador linear admite um autovalor λ, existe uma infinidade de autovetores correspondentes a λ.

Exemplo 8.63. *Se $\mathcal{T}$ é um operador linear identidade (i.e., para qualquer vetor v temos $\mathcal{T}(v) = v$), em um espaço vetorial real, então $\mathcal{T}$ admite um único autovalor λ, a saber, $\lambda = 1$. Isso significa que qualquer vetor não nulo deste espaço vetorial real é um autovetor de $\mathcal{T}$.*

Exemplo 8.64. *Seja $\mathcal{T}$ o operador linear em $\mathbb{R}^2$ usual, na base canônica $\{(1,0),(0,1)\}$, dado por*

$$\mathcal{T}(x,y) = (3x + 2y, 2x).$$

Se $\mathcal{T}$ admite autovalor λ, então $\mathcal{T}(x,y) = \lambda \cdot (x,y)$, onde $\cdot$ é a multiplicação de escalar por vetor em $\mathbb{R}^2$ usual.

Logo, devemos ter

$$(3x + 2y, 2x) = (\lambda x, \lambda y)$$

que, por sua vez, é equivalente a

$$3x + 2y = \lambda x \text{ e } 2x = \lambda y.$$

Logo,

$$y = \frac{2x}{\lambda} \quad e \quad y = \frac{x(\lambda - 3)}{2}.$$

As duas equações acima garantem que $x \neq 0 \Leftrightarrow y \neq 0$. Logo, para que ambas sejam teoremas, basta que

$$\frac{\lambda - 3}{2} = \frac{2}{\lambda},$$

o que implica que λ pode ser -1 ou 4, de acordo com Seção 43.

Os autovetores correspondentes ao autovalor -1 são

$$(x, -2x),$$

onde $x \neq 0$. Os autovetores correspondentes ao autovalor 4 são

$$(2y, y),$$

onde $y \neq 0$.

Observar que esses autovetores são linearmente independentes. Além disso, qualquer um deles (e.g., $(1,-2)$ e $(-2,-1)$) é base de um subespaço de uma dimensão, invariante sob a ação de $\mathcal{T}$.

Para que o leitor desenvolva uma visão intuitiva sobre o que está acontecendo no último Exemplo, consideremos o seguinte.

Aprendemos acima que qualquer vetor $(2c, c)$ é autovetor do operador linear

$$\mathcal{T}(x, y) = (3x + 2y, 2x),$$

desde que c seja diferente de 0.

Em particular, $(2, 1)$ é um autovetor com autovalor 4. Neste caso,

$$\mathcal{T}(2, 1) = (3(2) + 2(1), 2(2)) = (6 + 2, 4) = (8, 4) = 4 \cdot (2, 1).$$

Ou seja, $\mathcal{T}(2, 1)$ é uma combinação linear $\lambda(2, 1)$ de $(2, 1)$, onde λ é o autovalor 4. Além disso, $(2c, c)$ define um subespaço de $\mathbb{R}^2$ invariante sob a ação de $\mathcal{T}$.

O mesmo fenômeno não ocorre com vetores que não são autovetores de $\mathcal{T}(x, y)$. Por exemplo, $\mathcal{T}(3, 1) = (11, 6)$, sendo que $(11, 6)$ não é combinação linear de $(3, 1)$.

Comentários análogos valem para os autovetores $(c, -2c)$ com autovalor -1, onde $c \neq 0$.

Exemplo 8.65. ✍ *Seja $\mathcal{R}$ o operador linear em $\mathbb{R}^2$ usual, na base canônica, dado por*

$$\mathcal{R}(x, y) = (y, -x).$$

Logo, $\mathcal{R}$ não admite qualquer autovalor.

Exemplo 8.66. *Seja $\mathfrak{F} = \langle \mathcal{C}^\infty, \mathbb{R}, +, \cdot, \bigcirc \rangle$ o espaço vetorial real onde*

- *$\mathcal{C}^\infty$ é o conjunto de todas as funções reais diferenciáveis um número arbitrário de vezes;*
- *$+ : \mathcal{C}^\infty \times \mathcal{C}^\infty \to \mathcal{C}^\infty$ é a adição usual entre funções reais;*
- *$\cdot : \mathbb{R} \times \mathcal{C}^\infty \to \mathcal{C}^\infty$ é a multiplicação usual entre real e função real;*
- *$\bigcirc$ é a função real $\bigcirc : \mathbb{R} \to \mathbb{R}$ tal que $\bigcirc(x) = 0$ para todo real x.*

✍ *Logo, o operador linear $\mathcal{D}$ sobre $\mathcal{C}^\infty$ dado por*

$$\mathcal{D}(f(x)) = \frac{d}{dx} f(x)$$

admite infinitos autovalores.

Com efeito, de acordo com Seção 63, se

$$\frac{d}{dx}f(x) = \lambda f(x),$$

então $f(x) = f(0)\exp^{\lambda x}$*. Ou seja, cada real* λ *é um autovalor do operador linear de derivação. Os autovetores correspondentes são as funções* $f(x)$*. Notar que autovalores diferentes correspondem a autovetores linearmente independentes.*

No caso do problema do decaimento radioativo de Polonium-210, discutido na Seção 63, o modelo proposto

$$\frac{dm}{dt} = km$$

é uma equação de autovalores.

Resolver a equação diferencial acima é equivalente a determinar os autovetores do operador de derivação de primeira ordem

$$\frac{d}{dt}$$

correspondentes ao autovalor k. Para cada valor $m(0)$ real positivo, $m(t) = m(0)\exp(kt)$ é autovetor correspondente a k. Neste mesmo caso o próprio autovetor é uma função dependente do autovalor.

Cada elemento, substância ou isótopo conta com sua própria meia-vida, a qual define univocamente o valor de k. Cada meia-vida pode ser matematicamente mapeada por um autovalor k do operador de derivação sobre um espaço vetorial real de funções reais.

Exemplo 8.67. *Seja* $\bigcirc$ *o operador nulo sobre um espaço vetorial real qualquer, ou seja,*

$$\bigcirc(v) = \overline{0}$$

para todo vetor v*, onde* $\overline{0}$ *é o vetor nulo do espaço dado.*

Neste caso, $\bigcirc$ *admite apenas o real* 0 *como autovalor. Com efeito,*

$$\bigcirc(v) = 0 \cdot v$$

para qualquer vetor v*, por conta do Teorema 8.4.*

Logo, qualquer vetor não nulo v *do espaço é um autovetor correspondente ao autovalor* 0*.*

Nos Exemplos 8.66 e 8.67 mostramos que o zero real pode ser autovalor de um operador linear, apesar de nenhum autovetor ser nulo. Também mostramos que nem todo operador linear admite autovalor, como ocorre no Exemplo 8.65. Mas um fenômeno comum aos últimos exemplos é o fato de que autovalores diferentes correspondem a autovetores linearmente independentes. Isso não é mera coincidência, porém um teorema.

Teorema 8.41. *Autovalores diferentes do mesmo operador linear correspondem a autovetores linearmente independentes.*

Demonstração: A prova é feita por indução, para um conjunto qualquer de n autovetores correspondentes a n autovalores distintos dois a dois.

O caso em que $n = 1$ é imediato, por conta do Teorema 8.11 e do fato de que nenhum autovetor é nulo.

✍ A prova de que o caso para $n-1$ autovetores implica no caso para n autovetores fica a cargo do leitor. Sugestão: usar Teorema 8.10.

Detalhes sobre o papel de autovalores de operadores lineares podem ser encontrados em [35]. Por exemplo, existem métodos muito mais econômicos (do ponto de vista computacional) para determinar autovalores de operadores lineares sobre espaços vetoriais reais de dimensão finita do que aquele que foi empregado no Exemplo 8.64.

Seção 96

Outros espaços vetoriais

Sumário
Índice
Rede

Até este momento investigamos brevemente espaços vetoriais reais. Mas existem outros espaços vetoriais. Para que possamos qualificar isso, precisamos saber o que é um *corpo*, o qual é definido através de um predicado conjuntista, nos moldes da Seção 71.

Definição 8.22. *Uma quíntupla ordenada*

$$\mathfrak{K} = \langle K, +, \cdot, 0, 1\rangle$$

é um corpo *se as seguintes fórmulas são teoremas.*

K1: $K \neq \varnothing$ *(seus elementos diferentes de* 0 *e* 1 *são denotados por letras gregas minúsculas);*

K2: $+ : K \times K \to K$ *é uma função tal que* $+(\alpha, \beta) = \alpha + \beta$*;*

K3: $\cdot : K \times K \to K$ *é uma função tal que* $\cdot(\alpha, \beta) = \alpha \cdot \beta$ *(podemos também escrever simplesmente* $\alpha\beta$ *no lugar de* $\alpha \cdot \beta$*);*

K4: $0 \in K \wedge 1 \in K$*;*

K5: $\alpha + \beta = \beta + \alpha$

K6: $\alpha\beta = \beta\alpha$*;*

K7: $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$*;*

K8: $(\alpha\beta)\gamma = \alpha(\beta\gamma)$*;*

K9: $\alpha + 0 = \alpha$*;*

K10: $\alpha \cdot 1 = \alpha$*;*

K11: $\forall_K \alpha \exists_K \beta(\alpha + \beta = 0)$*;* β *é o* simétrico aditivo *de* α*, denotado por* $-\alpha$*;*

K12: $\forall_K \alpha(\alpha \neq 0 \Rightarrow \exists_K \beta(\alpha\beta = 1)$*;* β *é o* simétrico multiplicativo *de* α*, denotado por* α^{-1}*;*

K13: $\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$*.*

Usamos quantificadores relativizados (Definição 7.6) em K11 e K12. Por exemplo, K11 se lê ‘para todo α pertencente a K existe β pertencente a K tal que $\alpha + \beta$ é o neutro aditivo 0 (ver K9).

Exemplo 8.68. *De acordo com Seção 31,*

$$\langle \mathbb{Q}, +, \cdot, 0, 1 \rangle$$

é um corpo, onde + é a adição entre racionais, · é a multiplicação entre racionais, 0 é o neutro aditivo entre os racionais e 1 é o neutro multiplicativo entre os racionais.

Exemplo 8.69. *De acordo com Seção 39,*

$$\langle \mathbb{R}, +, \cdot, 0, 1 \rangle$$

é um corpo, onde + é a adição entre reais, · é a multiplicação entre reais, 0 é o neutro aditivo entre os reais e 1 é o neutro multiplicativo.

Exemplo 8.70. ✍ *De acordo com Seção 40,*

$$\langle \mathbb{C}, +, \cdot, 0, 1\rangle$$

é um corpo, onde + é a adição entre complexos, · é a multiplicação entre complexos, 0 é o neutro aditivo entre os complexos e 1 é o neutro multiplicativo entre os complexos.

Exemplo 8.71. *De acordo com Seção 29,*

$$\langle \omega, +, \cdot, 0, 1\rangle$$

não é um corpo, sendo + a adição entre naturais, · a multiplicação entre naturais, 0 o neutro aditivo entre os naturais e 1 o neutro multiplicativo entre os naturais.

Com efeito, axiomas K11 *e* K12 *não são teoremas nesta interpretação.*

Exemplo 8.72. *De acordo com Seção 30,*

$$\langle \mathbb{Z}, +, \cdot, 0, 1\rangle$$

não é um corpo, sendo + a adição entre inteiros, · a multiplicação entre inteiros, 0 o neutro aditivo entre os inteiros e 1 o neutro multiplicativo entre os inteiros.

Com efeito, axioma K12 *não é teorema nesta interpretação.*

Exemplo 8.73. *De acordo com Seção 39,*

$$\langle \mathbb{I}, +, \cdot, 0, 1\rangle$$

não é um corpo, se $\mathbb{I}$ *é o conjunto dos reais irracionais, + é a adição entre reais irracionais, · é a multiplicação entre reais irracionais, 0 é o neutro aditivo entre os reais e 1 é o neutro multiplicativo entre os reais.*

Com efeito, axiomas K2, K3 *e* K4 *não são teoremas nesta interpretação. Por exemplo, além de 0 e 1 não serem irracionais, o irracional* $\sqrt{2}$ *somado do irracional* $-\sqrt{2}$ *não é um irracional.*

Vimos acima três Exemplos de interpretações de corpo que são modelos de corpo e três que não são. As três primeiras são os exemplos mais comuns de corpos, no estudo de espaços vetoriais. A ideia é admitir escalares que não sejam necessariamente números reais, como vemos adiante. Mas, antes, é interessante mais um Exemplo.

Exemplo 8.74. *Seja*

$$\mathbf{F}_4 = \langle \{0, 1, 2, 3\}, +, \cdot, 0, 1\rangle$$

uma interpretação *de corpo onde as operações* $+$ *e* $\cdot$ *são definidas pelas tabelas abaixo.*

$+$ em $\mathbf{F}_4$				
$+$	0	1	2	3
0	0	1	2	3
1	1	0	3	2
2	2	3	0	1
3	3	2	1	0

$\cdot$ em $\mathbf{F}_4$				
$\cdot$	0	1	2	3
0	0	0	0	0
1	0	1	2	3
2	0	2	3	1
3	0	3	1	2

✍ *Para fins de ilustração,* $3 + 2 = 1$ *e* $2 \cdot 2 = 3$.

Neste caso $\mathbf{F}_4$ *é um corpo* finito*, no sentido de que o conjunto* $\{0, 1, 2, 3\}$ *conta com apenas quatro elementos.*

✍ *Cabe ao leitor verificar que os postulados da Definição* 8.22 *são teoremas.*

Por exemplo, levando em conta que ambas as tabelas dadas são simétricas em relação à diagonal principal, comutatividade da adição K5 *e comutatividade da multiplicação* K6 *são imediatos.*

Observar que, no Exemplo acima,

- 0 é neutro aditivo e 1 é neutro multiplicativo;
- o simétrico aditivo de n é o próprio n, se $n \in \{0, 1, 2, 3\}$;
- 0 não admite simétrico multiplicativo;
- o simétrico multiplicativo de 1 é o próprio 1;
- o simétrico multiplicativo de 2 é 3;
- o simétrico multiplicativo de 3 é 2.

ℹ ✍ Muitos outros exemplos de corpos finitos podem ser apresentados, os quais são também conhecidos como *corpos de Galois*. A *ordem* de um corpo finito $\langle K, +, \cdot, 0, 1\rangle$ é o número de elementos de K. Se o corpo tem ordem n, então n é um primo ou uma potência de um primo (Definição 4.2). No Exemplo acima a ordem de $\mathbf{F}_4$ é 2^2.

Finalmente podemos definir espaços vetoriais quaisquer.

Definição 8.23. *Um conjunto*

$$\mathcal{V} = \langle V, \mathfrak{K}, \oplus, \odot, \overline{0} \rangle$$

é um espaço vetorial *se as seguintes fórmulas são teoremas (os comentários entre parênteses não fazem parte dos postulados).*

V1: $V \neq \varnothing$ *(os elementos de* V *são chamados de* vetores*);*

V1': $\mathfrak{K}$ *é um corpo* $\langle K, +, \cdot, 0, 1 \rangle$ *(os elementos de* K *são chamados de* escalares*);*

V2: $\oplus : V \times V \to V$ *é uma função, onde abreviamos* $\oplus(u, v)$ *como* $u \oplus v$*, sendo* u *e* v *elementos de* V *(chamamos* $\oplus$ *de* adição de vetores*);*

V3: $\odot : K \times V \to V$ *é uma função, onde abreviamos* $\odot(\alpha, u)$ *como* $\alpha \odot u$ *ou, simplesmente,* αu*, sendo* α *um elemento de* K *e* u *um elemento de* V *(chamamos* $\odot$ *de* multiplicação de escalar por vetor*);*

V4: $\overline{0} \in V$ *(*$\overline{0}$ *é o* vetor nulo*);*

V5: *Se* u *pertence a* V*, então* $u \oplus \overline{0} = u$;

V6: *Se* u *e* v *são elementos de* V*, então* $u \oplus v = v \oplus u$;

V7: *Se* u*,* v *e* w *pertencem a* V*, então* $(u \oplus v) \oplus w = u \oplus (v \oplus w)$;

V8: *Se* u *pertence a* V*, então existe* v *pertencente a* V *tal que* $u \oplus v = \overline{0}$ *(*v *é chamado de* simétrico aditivo de u *e denotado por* $-u$*);*

V9: *Se* α *pertence a* K *e* u *e* v *pertencem a* V*, então*

$$\alpha \odot (u \oplus v) = (\alpha \odot u) \oplus (\alpha \odot v);$$

V10: *Se* α *e* β *pertencem a* K *e* u *pertence a* V*, então*

$$(\alpha + \beta) \odot u = (\alpha \odot u) \oplus (\beta \odot u);$$

V11: *Se* α *e* β *são escalares e* u *é um vetor, então*

$$(\alpha \cdot \beta) \odot u = \alpha \odot (\beta \odot u);$$

V12: *Se* 1 *é o neutro multiplicativo de* $\mathfrak{K}$ *e* u *pertence a* V*, então* $1 \odot u = u$.

Novamente adotamos a seguinte convenção: todos os vetores diferentes do vetor nulo são denotados por letras latinas minúsculas, enquanto os escalares são sempre denotados por letras gregas minúsculas, desde que não sejam 0 ou 1.

Se um espaço vetorial admite como único vetor o nulo, ele é dito um *espaço vetorial trivial*.

Se o corpo $\mathfrak{K}$ é o corpo $\mathbb{R}$ dos reais (ver Exemplo 8.69), então o espaço vetorial é chamado de *espaço vetorial real*, conforme Seção 80. Se o corpo $\mathfrak{K}$ é o corpo $\mathbb{Q}$ dos reais racionais (ver Exemplo 8.68), então o espaço vetorial é chamado de *espaço vetorial racional*. Se o corpo $\mathfrak{K}$ é o corpo $\mathbb{C}$ dos complexos (ver Exemplo 8.70), então o espaço vetorial é chamado de *espaço vetorial complexo*.

Exemplo 8.75. *De acordo com* Exemplo 8.33, *o espaço vetorial real*

$$\langle \mathbb{R}, \mathbb{R}, +, \cdot, 0 \rangle$$

tem uma dimensão. Isso porque qualquer base dele admite um único elemento. A base canônica é $\{1\}$.

Porém, o espaço vetorial racional

$$\mathcal{V} = \langle \mathbb{R}, \mathbb{Q}, +, \cdot, 0 \rangle$$

não admite base com apenas um vetor. Com efeito, supor que existe base

$$b = \{r\}$$

para $\mathcal{V}$. *Neste caso,* r *é real racional ou real irracional. Afinal, os vetores de* $\mathcal{V}$ *são números reais quaisquer, enquanto os escalares são apenas reais racionais.*

Se r *for real racional, nenhum vetor* s *de* $\mathbb{R}$ *que seja real irracional pode ser obtido por combinação linear dos elementos de* $\{r\}$. *Com efeito, qualquer escalar* α *do corpo* $\mathbb{Q}$ *é um real racional e, por isso,* $\alpha \cdot r$ *é real racional e, portanto, diferente de* s.

Por outro lado, se r *for real irracional, nenhum vetor* s *de* $\mathbb{R}$ *que seja real racional pode ser obtido por combinação linear dos elementos de* $\{r\}$. *Afinal, qualquer escalar* α *do corpo* $\mathbb{Q}$ *é um real racional e, por isso,* $\alpha \cdot r$ *é real irracional e, portanto, diferente de* s.

No Exemplo acima foi provado que a reta dos reais sobre o corpo dos racionais não tem uma dimensão, apesar da mesma reta dos reais — mas desta vez sobre o corpo dos reais — ter uma dimensão.

É possível provar que o espaço vetorial racional acima não admite qualquer base finita e, portanto, não tem dimensão finita.

Para lidarmos com o conceito de *dimensão infinita* em um espaço vetorial real, ou qualquer outro, precisamos estender o conceito de base, como fazemos adiante, na Seção 97.

Um conceito de importância estratégica é o de *espaço de Hilbert*. Um espaço de Hilbert é um espaço vetorial complexo munido de produto interno (devidamente definido de maneira a generalizar a Definição 8.11) de modo que a norma induzida por este produto interno defina um *espaço métrico completo*. Por sua vez, um espaço métrico (Definição 8.10) é completo se toda sequência de Cauchy definida nele (relativamente à métrica) for convergente.

Espaços de Hilbert são a principal base matemática para o estudo de mecânica quântica não relativística. Física quântica, por sua vez, é responsável por grandes fatias do PIB de países desenvolvidos, servindo de suporte teórico de tecnologias para a concepção e fabricação de computadores, *smartphones*, televisores, instrumentos de telecomunicações, *lasers*, aparelhos de GPS, relógios atômicos, máquinas de ressonância magnética, entre muitos outros.

STOP Portanto, espaços vetoriais impregnam o cotidiano do leitor, independentemente de seu interesse sobre o assunto.

✍ Um exercício interessante é investigar a possibilidade de definir espaços vetoriais sobre corpos finitos, como $\mathbf{F}_4$, do Exemplo 8.74.

Seção 97

Espaços vetoriais de dimensão infinita

Sumário
Índice
Rede

Como sabemos, nem todo espaço vetorial tem base finita.

Definição 8.24. *Seja* $\mathcal{V} = \langle V, \mathfrak{K}, \oplus, \odot, \overline{0}\rangle$ *um espaço vetorial sobre um corpo* $\mathfrak{K} = \langle K, +, \cdot, 0, 1\rangle$. *Um vetor* v *de* V *é* combinação linear *de vetores* v_1, v_2, $\cdots$, v_n *sss existem escalares* α_1, α_2, $\cdots$, α_n *pertencentes a* K *tais que*

$$v = \sum_{i=1}^{n} \alpha_i v_i.$$

Em outras palavras, a definição acima generaliza Definição 8.4, no sentido de permitir escalares de um corpo qualquer.

Definição 8.25. *Seja $\mathcal{V} = \langle V, \mathfrak{K}, \oplus, \odot, \overline{0}\rangle$ um espaço vetorial sobre um corpo $\mathfrak{K}$. Um conjunto finito*

$$x = \{v_1, v_2, \cdots, v_n\}$$

de vetores de V é linearmente independente *sss nenhum dos vetores de x é combinação linear dos demais elementos de x. Caso contrário, dizemos que x é* linearmente dependente.

Definição 8.26. *Seja $\mathcal{V} = \langle V, \mathfrak{K}, \oplus, \odot, \overline{0}\rangle$ um espaço vetorial sobre um corpo $\mathfrak{K}$. Um conjunto x de vetores de V é* linearmente independente *sss qualquer subconjunto finito de x é linearmente independente, de acordo com Definição 8.25. Caso contrário, dizemos que x é* linearmente dependente.

Exemplo 8.76. *Seja $\mathcal{V} = \langle V, \mathbb{R}, \oplus, \odot, \bigcirc\rangle$, onde*

i: *$V = \{y \in \mathcal{C}^\infty \mid y$ é polinomial com domínio $\mathbb{R}\}$,*

ii: *$\oplus$ é restrição de $+$ (no espaço $\mathcal{C}^\infty$ usual) a $V \times V$,*

iii: *$\odot$ é restrição de $\cdot$ (no espaço $\mathcal{C}^\infty$ usual) a $\mathbb{R} \times V$, e*

iv: *$\bigcirc$ é a função real identicamente nula, com domínio $\mathbb{R}$.*

Este é um espaço vetorial real similar àquele do Exemplo 8.39. *A diferença reside apenas nos domínios das funções polinomiais pertencentes a $\mathcal{C}^\infty$. Seja*

$$b = \{p_i \in V \mid p_i : \mathbb{R} \to \mathbb{R} \text{ é uma função dada por } p_i(x) = x^i, \text{ onde } i \in \omega\}.$$

Em outras palavras, b é o conjunto das funções monomiais com coeficientes iguais a 1 e grau i, onde i é um natural.

Escrevendo de outra maneira, $b = \{1, x, x^2, x^3, x^4, \cdots\}$, onde cada x^i é uma abreviação para uma função monomial com coeficiente 1. Uma vez que ω é infinito e b é equipotente a ω (ver Seção 33), então b é infinito.

Teorema Fundamental da Álgebra (Seção 43) garante que qualquer subconjunto finito de b é linearmente independente. Logo, b é linearmente independente, conforme Definição 8.26.

O último Exemplo ilustra o grande alcance da Definição 8.26, no sentido de conceituar independência linear (e, consequentemente, dependência linear) para conjuntos finitos e conjuntos infinitos de vetores de um dado espaço vetorial sobre um corpo qualquer.

Definição 8.27. *Sejam $\mathcal{V} = \langle V, \mathfrak{K}, \oplus, \odot, \overline{0}\rangle$ um espaço vetorial, sobre um corpo $\mathfrak{K}$, e $s \subseteq V$ um conjunto não vazio. Seja ainda*

$$c = \{x \in \wp(s) \mid x \text{ é linearmente independente}\}.$$

Se c for não vazio, dizemos que um elemento maximal de c (ver Definição 4.20) é um subconjunto de s maximal linearmente independente.

Em outras palavras, dado um espaço de vetores V, um subconjunto s de V é maximal linearmente independente se, e somente se, para qualquer vetor v de V diferente de todos os demais pertencentes a s temos que $s \cup \{v\}$ é linearmente dependente.

No caso particular de espaços vetoriais reais $\mathcal{V}$ de dimensão finita, uma base b de $\mathcal{V}$ é um conjunto L.I. que gera o espaço de vetores de $\mathcal{V}$ (Definição 8.7). Nem todo conjunto L.I. gera um espaço, como já vimos (Exemplo 8.26). Mas toda base b de $\mathcal{V}$ é maximal linearmente independente. Qualquer outro vetor adicionado a b implica que o novo conjunto é L.D. Essa simples ideia permite generalizar o conceito de base de um espaço vetorial qualquer, seja real ou não.

Definição 8.28. *Seja $\mathcal{V} = \langle V, \mathfrak{K}, \oplus, \odot, \overline{0}\rangle$ um espaço vetorial sobre um corpo $\mathfrak{K}$. Dizemos que b é uma* base *de $\mathcal{V}$ sss b é um* subconjunto de V maximal linearmente independente.

Exemplo 8.77. *O conjunto b do* Exemplo 8.76 *é uma base do espaço vetorial do mesmo* Exemplo. *Afinal, qualquer outra função polinomial acrescentada a*

$$b = \{1, x, x^2, x^3, x^4, \cdots\}$$

produz um conjunto linearmente dependente.

✍ *Outra possível base para aquele espaço é o conjunto*

$$d = \{p_i \in V \mid p_i : \mathbb{R} \to \mathbb{R} \text{ é uma função dada por}$$
$$p_i(x) = -38x^i, \text{ onde } i \in \omega\}.$$

Uma infinidade de outras bases podem ser exibidas para o espaço vetorial real das funções polinomiais com domínio $\mathbb{R}$. Porém, é teorema que todas as possíveis bases são equipotentes entre si.

✍ O leitor deve observar que Definição 8.7 é um caso particular da Definição 8.28. É um exercício edificante provar isso.

Seção 98

Resumo da ópera

Sumário
Índice
Rede

Ao contrário do que uns e outros dizem por aí, vetores não são segmentos de reta orientados ou 'entes' com 'módulo', 'direção' e 'sentido', o que quer que signique essa nomenclatura maluca. Um vetor é tão somente um elemento de um espaço de vetores. Para saber o que é um espaço de vetores é necessário qualificar o conceito de espaço vetorial (Definição 8.23). Fizemos isso usando um predicado conjuntista (Definição 7.1) formulado na linguagem de ZF.

Neste contexto, vetores podem ser pares ordenados de números reais, os quais são consistentes com conceitos de geometria euclidiana, conforme Seção 90. Mas podem ser também n-uplas ordenadas de reais ou de complexos, funções reais ou funções complexas, bem como matrizes e até mesmo transformações lineares entre espaços vetoriais. Se não for definida uma norma em um espaço vetorial, não pode haver o tal do 'módulo' propagado por certos professores. Uma vez que cada possível produto interno induz uma norma diferente, é preciso muito cuidado para qualificar sobre o que se está falando.

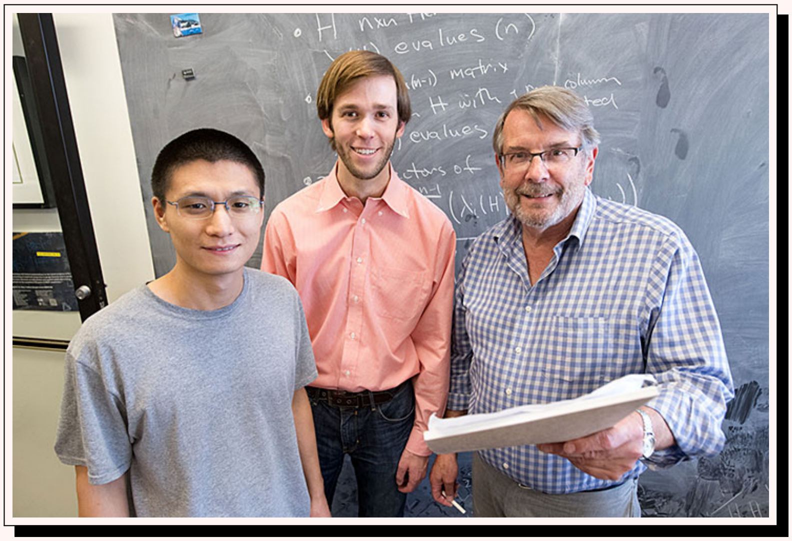
Um dos possíveis modelos de espaço vetorial real é o plano cartesiano, desde que munido de produto interno canônico. Uma vez que o plano cartesiano é modelo de plano euclidiano, esse fato justifica interpretar tais casos muito particulares de vetores como segmentos orientados de reta. Mas, se um vetor é uma matriz real, tal matriz não é um segmento de reta. Se um vetor é a função exponencial, esta também não é um segmento de reta.

Seção 99

Notas históricas

Sumário
Índice
Rede

Muito difícil pontuar em qual momento surgiram as primeiras ideias que conduziram à atual visão sobre espaços vetoriais. A partir da geometria analítica de René Descartes e Pierre de Fermat, Giuso Bellavitis concebeu a noção de segmentos orientados. Em meados do século 19 Arthur Cayley introduziu a notação matricial. Mas foi Hermann Grassmann o primeiro a perceber a necessidade de tratar de estruturas algébricas com objetos abstratos e não necessariamente matrizes ou pares ordenados de reais. Os conceitos de independência linear, dimensão e produto interno surgiram com Grassmann. Giuseppe Peano delineou a atual definição de espaço vetorial em 1888.



Da esquerda para a direita, Xining Zhang, Peter Denton e Stephen Parke, os principais descobridores da IAA

Fonte: Brookhaven National Laboratory.

Apesar do estudo de espaços vetoriais ser antigo e bem estabelecido, em 2021 foi publicado um resultado básico até então desconhecido sobre o tema. Trata-se da *Identidade entre Autovetores e Autovalores* (IAA), a qual é dada por uma fórmula com grande impacto sobre aplicações. Autovalores de um operador linear são fáceis de calcular, enquanto autovetores consomem um esforço computacional muito maior. No entanto, graças à IAA, esse quadro mudou radicalmente, pelo menos para operadores lineares hermitianos que atuam sobre espaços vetoriais de dimensão finita. Detalhes em [13].

PARTE 9

Probabilidades

Nesta parte discutimos sobre um dos assuntos mais importantes da atualidade: probabilidades.

Seção 100

Motivação

Sumário
Índice
Rede

Em 1898 Morgan Robertson publicou *Futility*, história ficcional sobre o navio de passageiros *Titan* que, em uma noite de abril, colide contra um *iceberg* e naufraga. Em 1912, em uma noite abril, *Titanic* colide contra um *iceberg* e afunda.

Numa sexta-feira de 1865, Abraham Lincoln foi assassinado no Teatro Ford, com um tiro na cabeça, em frente à esposa. Numa sexta-feira de 1963, John Kennedy foi assassinado em um carro Ford (uma *limousine* Lincoln), com um tiro na cabeça, em frente à esposa. Ambos tiveram sucessores com o sobrenome Johnson.

Em setembro de 2009 a loteria da Bulgária sorteou exatamente os mesmos seis números em dois jogos consecutivos. Uma comissão foi designada pelo Ministro dos Esportes para investigar o caso, mas nenhuma fraude foi detectada.

Uma questão natural sobre os três exemplos históricos dados acima é a seguinte: quais são as chances de ocorrências de tais coincidências?

Como bem argumenta David J. Hand em seu famoso livro [21], 'devemos esperar o inesperado'.

Enquanto alguns percebem o caso *Titan-Titanic* como evidência de uma profecia, o caso Lincoln-Kennedy como uma macabra conexão entre dois presidentes distantes por um século, e o caso da loteria búlgara como evidência de crime, o que realmente está em jogo aqui é um conflito entre realidade e modos de percepção da realidade.

Matemática não existe como proposta para compreender a realidade. Mas matemática é uma ferramenta muito útil para mapear fenômenos do mundo real. É claro que mapas podem informar direções equivocadas. Porém, neste caso, o problema não reside no emprego de mapas, mas em quem criou o mapa.

Uma das ferramentas matemáticas mais utilizadas para mapear chances de ocorrências de eventos é o conceito de *probabilidade*. A partir da próxima Seção fazemos isso, novamente utilizando predicados conjuntistas no contexto de ZF.

Na Seção 105 discutimos sobre os casos Titan-Titanic, Lincoln-Kennedy e loteria da Bulgária.

Seção 101

σ-álgebra

Sumário
Índice
Rede

Probabilidade é uma função. Logo, demanda um domínio. Tal domínio é uma álgebra de eventos, também conhecida como σ-álgebra. Portanto, precisamos conhecer este conceito antes de qualificarmos o que é uma função de probabilidades. Mas, antes de conceituarmos σ-álgebras, precisamos de um conceito preliminar.

Definição 9.1. *Um conjunto x é* enumerável *sss x é equipotente a algum subconjunto de ω.*

Sobre o conceito de equipotência (denotada por $\sim$), ver Definição 4.17. Sobre ω, este é o conjunto dos números naturais, conforme Definição 3.5.

Em outras palavras, x é enumerável se, e somente se, existe bijeção

$$f : x \to r,$$

onde $r \subseteq \omega$.

Exemplo 9.1. i: *O conjunto ω dos números naturais é enumerável. Com efeito, $\omega \sim \omega$.*

ii: *O conjunto $\{\omega, S(\omega)\}$ é enumerável, onde $S(\omega)$ é o sucessor de ω (Definição 3.4). Com efeito,*

$$\{\omega, S(\omega)\} \sim 2,$$

lembrando que $2 \subseteq \omega$, onde

$$2 = \{\varnothing, \{\varnothing\}\},$$

conforme Seção 23. Aliás, o leitor deve perceber que todo ordinal finito é elemento e subconjunto de ω.

iii: *O conjunto $\mathbb{R}$ dos números reais não é enumerável. Afinal, $\mathbb{R}$ não é equipotente a qualquer subconjunto de ω, conforme Seção 34.*

Definição 9.2. *Seja x um conjunto. Dizemos que Σ é uma σ-*álgebra *de x (ou, simplesmente, uma σ-*álgebra*) sss*

i: *$\Sigma \subseteq \wp(x)$, onde $\wp(x)$ é a* potência *de x;*

ii: *se s pertence a Σ, então $x - s$ pertence a Σ;*

iii: *$x \in \Sigma$;*

iv: *se y é um conjunto enumerável de elementos de Σ, então*

$$\bigcup_{s \in y} s$$

é um elemento de Σ.

Em outras palavras, os elementos de uma σ-álgebra (lê-se 'sigma-álgebra') de um dado conjunto x são subconjuntos de x, de modo que o próprio x é um deles. Além disso, se um dado s pertence à σ-álgebra, seu *complementar relativamente a x* também pertence, sendo que o complementar de s relativamente a x é o conjunto de todos os termos pertencentes a x, exceto aqueles que pertencem a s. Por último, qualquer união finitária de n elementos da σ-álgebra é um elemento da σ-álgebra; e qualquer união arbitrária de uma quantia enumerável de elementos da σ-álgebra é um elemento da σ-álgebra.

Exemplo 9.2. ✍

I: *Sejam* $\{m\}$ *e* $\{n\}$ *conjuntos, onde chamamos* $\{m\}$ *de* cara *e* $\{n\}$ *de* coroa. *Logo,* $\wp(\{m, n\})$ *é uma* σ*-álgebra do par* $\{m, n\}$. *Observar que* cara *e* coroa *são subconjuntos de* $\{m, n\}$.

II: *Seja* x *um conjunto qualquer. Logo,* $s = \{\varnothing, x\}$ *é uma* σ*-álgebra de* x. *Observar que, no caso particular em que* $x = \varnothing$, *temos* $s = \wp(x)$. *Para os demais casos,* $s \neq \wp(x)$. *Isso prova que nem toda* σ*-álgebra de um conjunto* x *é a potência de* x.

III: *Seja* $y = \{a, b, c\}$ *um conjunto com três elementos. Logo, o conjunto* $t = \{\{a\}, \{b, c\}, \varnothing, y\}$ *é uma* σ*-álgebra de* y. *Observar que* $t \neq \wp(y)$.

Teorema 9.1. *A potência de* x *é uma* σ*-álgebra de* x.

Demonstração: Basta provarmos que todos os quatro itens da Definição 9.2 são teoremas para o caso do enunciado.

Sobre item I, observar que $\wp(x) \subseteq \wp(x)$ (Teorema 3.5).

Sobre item II, se s é subconjunto de x (i.e., pertence à potência de x), então $x - s$ é subconjunto de x e, portanto, pertence à potência de x.

Sobre item III, $x \in \wp(x)$, uma vez que todo conjunto é subconjunto dele mesmo (Teorema 3.5).

Finalmente, sobre item IV, observar que, se y é um conjunto de subconjuntos de x (seja enumerável ou não), então a união arbitrária de todos os elementos de y é um subconjunto de x e, portanto, um elemento da potência de x.

ℹ A recíproca do último teorema não é um teorema. O conjunto b de todos os *conjuntos de Borel* de $\mathbb{R}$ é uma σ-álgebra de $\mathbb{R}$, apesar de $b \neq \wp(\mathbb{R})$. No entanto, foge aos nossos propósitos o estudo de conjuntos de Borel. Queremos aqui apenas introduzir e discutir noções elementares sobre probabilidades. Neste contexto, o que interessa saber, pelo menos por enquanto, é que existem σ-álgebras diferentes das potências de conjuntos, além dos casos envolvendo conjuntos finitos dos itens II e III do Exemplo 9.2.

Teorema 9.2. *O conjunto vazio pertence à σ-álgebra de qualquer conjunto.*

Demonstração: Seja x um conjunto qualquer. Se Σ é sua σ-álgebra, então item III da Definição 9.2 garante que $x \in \Sigma$. Mas item II garante que $x - x$ também pertence a Σ. Uma vez que

$$x - x = \varnothing,$$

então a substitutividade da igualdade implica que $\varnothing \in \Sigma$.

Teorema 9.3. *A interseção entre duas σ-álgebras quaisquer de um conjunto x é uma σ-álgebra de x.*

✍ Deixamos a demonstração deste último como um divertido exercício para o leitor.

Exemplo 9.3. *Seja $x = \{a, b, c\}$ um conjunto com três elementos. Logo, os conjuntos*

$$s = \{\{a\}, \{b, c\}, \varnothing, x\}$$

e

$$t = \{\{b\}, \{a, c\}, \varnothing, x\}$$

são σ-álgebras de x. Além disso, $s \cap t = \{x, \varnothing\}$ é uma σ-álgebra de x, conforme item II *do* Exemplo 9.2.

No entanto, observar que

$$r = s \cup t = \{\{a\}, \{b\}, \{a, c\}, \{b, c\}, \varnothing, x\}$$

não é uma σ-álgebra de x. Com efeito, $\{a\} \cup \{b\} = \{a, b\}$ é uma união enumerável de elementos de r, apesar de $\{a, b\}$ não pertencer a r. Logo, item IV *da Definição 9.2 não é teorema para este caso.*

O último Exemplo deixa claro que união arbitrária de σ-álgebras não é necessariamente uma σ-álgebra. Obviamente, no caso particular da união de $\wp(x)$ com qualquer outra σ-álgebra de x é uma σ-álgebra de x.

✍ O último teorema pode ser generalizado para *interseções arbitrárias* de σ-álgebras. Interseção arbitrária de termos de x pode ser definida a partir do Axioma da União e do Esquema de Separação, como se segue.

DEFINIÇÃO 9.3.

$$\bigcap_{t\in x} t = \left\{ r \in \bigcup_{t\in x} t \mid \forall z(z \in x \Rightarrow r \in z) \right\}$$

✍ Neste caso, interseção finitária (Definição 3.10) é um caso particular de interseção arbitrária

A ideia intuitiva de interseção arbitrária é simples. A interseção arbitrária de todos os elementos de um dado x é um subconjunto da união arbitrária de x cujos elementos são apenas aqueles que são comuns a todos os elementos de x.

O lado esquerdo da igualdade acima se lê 'interseção de todos os conjuntos t pertencentes a x'.

SEÇÃO 102

Espaço de probabilidades

SUMÁRIO
ÍNDICE
REDE

Finalmente podemos conceituar probabilidades.

DEFINIÇÃO 9.4. *Um* espaço de probabilidades $\mathfrak{p}$ *é uma tripla ordenada*

$$\mathfrak{p} = \langle \Omega, \Sigma, p \rangle$$

tal que as seguintes fórmulas são teoremas.

P1: Σ *é uma* σ*-álgebra de* Ω*;*

P2: $p : \Sigma \to \mathbb{R}$ *é uma função, onde* $p(e) \geq 0$*, para todo* $e \in \Sigma$*;*

P3: $p(\Omega) = 1$*;*

P4: *se* y *é um conjunto enumerável de elementos de* Σ*, dois a dois disjuntos (ou seja, para quaisquer* r *e* s *pertencentes a* y*, temos que* $r \neq s \Rightarrow r \cap s = \varnothing$*), então*

$$p\left(\bigcup_{s\in y} s\right) = \sum_{s\in y} p(s).$$

Se $\mathfrak{p} = \langle \Omega, \Sigma, p \rangle$ é um espaço de probabilidades, chamamos o conjunto Ω de *espaço amostral*.

Os elementos da σ-álgebra Σ de um espaço de probabilidades são chamados de *eventos*.

Logo, todo evento é um subconjunto do espaço amostral. Em particular, o espaço amostral é um evento, de acordo com item iii da Definição 9.2.

A função p, em um espaço de probabilidades $\langle \Omega, \Sigma, p \rangle$, é chamada de *função de probabilidade* ou, simplesmente, *probabilidade*. Se s é um evento, lemos $p(s)$ como 'probabilidade de ocorrer o evento s' ou 'probabilidade do evento s'.

Neste contexto, axioma p1 diz que eventos contam com uma estrutura algébrica definida pelo fechamento de uniões arbitrárias enumeráveis. Ou seja, qualquer união enumerável de eventos é um evento.

Axioma p2 diz que a probabilidade de qualquer evento é um número real maior ou igual a zero.

Postulado p3 afirma que a probabilidade do evento Ω é 1.

Observar que Ω é a união arbitrária (enumerável ou não) de todos os elementos de qualquer σ-álgebra de Ω. Neste contexto é importante o leitor perceber que uma σ-álgebra não precisa necessariamente ser um conjunto enumerável.

Segue um resultado surpreendente. Se uma σ-álgebra é enumerável, ela é obrigatoriamente finita.

Em outras palavras, se Σ é uma σ-álgebra infinita, então qualquer função injetora $f : \omega \to \Sigma$ é não sobrejetora. Logo, nenhuma σ-álgebra é equipotente a ω.

Dois eventos r e s distintos entre si e tais que $r \cap s = \varnothing$, são ditos *mutuamente excludentes* ou, simplesmente, *disjuntos*.

Axioma p4 afirma que a probabilidade da união arbitrária de eventos mutuamente excludentes é a soma das probabilidades individuais de tais eventos.

Observar que o termo

$$\sum_{s \in y} p(s)$$

é um somatório se y for um conjunto finito. Além disso, o mesmo termo é uma série, se y for enumerável e infinito. No último caso, P4 garante que a série é necessariamente convergente.

Graças ao Teorema 9.2, o conjunto vazio é um evento em qualquer espaço amostral e, consequentemente, em qualquer espaço de probabilidades. Por conta disso, ele merece um nome especial. O conjunto vazio é chamado de *evento impossível*. Ademais, se $\langle \Omega, \Sigma, p \rangle$ é um espaço de probabilidades, dizemos que Ω é o *evento inevitável*

Teorema 9.4. *A probabilidade do evento impossível, em qualquer espaço de probabilidades, é zero.*

Demonstração: Seja $\mathfrak{p} = \langle \Omega, \Sigma, p \rangle$ um espaço de probabilidades. Teorema 9.2 e item III da Definição 9.2 garantem que ambos Ω e $\varnothing$ são eventos. Postulado P4 da Definição 9.4 garante que

$$p(\Omega \cup \varnothing) = p(\Omega) + p(\varnothing),$$

uma vez que $\Omega \cap \varnothing = \varnothing$. No entanto, postulado P3 garante que $p(\Omega) = 1$. Logo,

$$1 = 1 + p(\varnothing).$$

Isso implica que

$$p(\varnothing) = 0.$$

Uma consequência imediata do último teorema e de P3 é que o espaço amostral jamais é vazio, em um espaço de probabilidades.

Teorema 9.5. *Qualquer conjunto não vazio pode ser o espaço amostral de um espaço de probabilidades.*

Demonstração: Seja x um conjunto não vazio. Como visto no item II do Exemplo 9.2, $s = \{\varnothing, x\}$ é uma σ-álgebra de x. Logo, basta definir a função $p : s \to [0, 1]$ dada por $p(x) = 1$ e $p(\varnothing) = 0$.

Teorema 9.6 (Monotonicidade). *Se r e s são eventos em um espaço de probabilidades $\langle \Omega, \Sigma, p \rangle$, então*

$$r \subseteq s \Rightarrow p(r) \leq p(s).$$

Demonstração: Se $r \subseteq s$, seja $t = s - r$, ou seja, o conjunto dos termos pertencentes a s que não pertencem a r. Logo, $r \cap t = \varnothing$. Além disso, $r \cup t = s$. Portanto, axioma P4 garante que

$$p(r \cup t) = p(r) + p(t).$$

Mas $p(r \cup t) = p(s)$. Logo,

$$p(s) = p(r) + p(t).$$

Lembrando que todas as probabilidades são maiores ou iguais a zero, então $p(s) \geq p(r)$.

O último teorema é conhecido como *monotonicidade da probabilidade*. Como vemos adiante, sua recíproca não é teorema.

Teorema 9.7. *Se e é um evento em um espaço de probabilidades com probabilidade p, então*

$$0 \leq p(e) \leq 1.$$

Demonstração: Todo evento e é elemento de uma σ-álgebra e, portanto, um subconjunto do espaço amostral Ω. Logo, a monotonicidade da probabilidade, garantida pelo teorema anterior, implica que $p(e) \leq p(\Omega)$.

Uma vez que axiomas P2 e P3 da Definição 9.4 garantem que $p(e) \geq 0$ e $p(\Omega) = 1$, respectivamente, então $p(e) \in [0, 1]$, onde $[0, 1]$ é o intervalo fechado dos números reais entre 0 e 1, incluindo 0 e 1.

Teorema 9.8. *Se r e s são eventos em um espaço de probabilidades $\langle \Omega, \Sigma, p \rangle$, então*

$$p(r \cup s) = p(r) + p(s) - p(r \cap s).$$

Demonstração: Sabemos que $r \cap (s - r) = \varnothing$ e $r \cup (s - r) = r \cup s$. Logo, P4 implica que

$$p(r \cup s) = p(r) + p(s - r).$$

Mas o evento $s - r$ é também idêntico a $s - (r \cap s)$. Logo, a última equação em destaque pode ser reescrita como

$$p(r \cup s) = p(r) + p(s - (r \cap s)).$$

No entanto, $s = (s - (r \cap s)) \cup (r \cap s)$, sendo que

$$(s - (r \cap s)) \cap (r \cap s) = \varnothing.$$

Logo, p4 implica que

$$p(s) = p(s - (r \cap s)) + p(r \cap s).$$

Uma vez que o termo $p(s - (r \cap s))$ é comum à última e à antepenúltima equação em destaque, então

$$p(r \cup s) = p(r) + p(s) - p(r \cap s).$$

Levando em conta que união finitária é definível a partir da disjunção, e interseção finitária é definível a partir de conjunção, o teorema acima estabelece o seguinte:

A probabilidade de ocorrer o evento r ou *o evento s é a*
probabilidade de ocorrer o evento r mais
a probabilidade de ocorrer o evento s menos
a probabilidade de ocorrer o evento r e *o evento s.*

Exemplo 9.4. *No* Exemplo 9.2 *vimos que, se $\{m\}$ e $\{n\}$ são conjuntos (chamados, respectivamente, de* cara *e* coroa*), então*

$$\wp(\{m, n\}) = \{\varnothing, \{m\}, \{n\}, \{m, n\}\}$$

é uma σ-álgebra do par $\{m, n\}$.

Podemos, portanto, definir um espaço de probabilidades

$$\langle \{m, n\}, \{\varnothing, \{m\}, \{n\}, \{m, n\}\}, p\rangle,$$

onde p é uma função com domínio $\{\varnothing, \{m\}, \{n\}, \{m, n\}\}$ dada, por exemplo, por

$$p(\varnothing) = 0,$$

$$p(\{m\}) = \frac{1}{2},$$

$$p(\{n\}) = \frac{1}{2}$$

e

$$p(\{m, n\}) = 1.$$

Neste caso, a probabilidade de ocorrer o evento cara *é $\frac{1}{2}$, a probabilidade de ocorrer o evento* coroa *é $\frac{1}{2}$, a probabilidade de ocorrer o evento* cara e coroa *é 0 e a probabilidade de ocorrer o evento* cara ou coroa *é 1.*

Exemplo 9.5. ✍ *Seguindo o último* Exemplo, *se*

$$\langle \{m, n\}, \{\varnothing, \{m\}, \{n\}, \{m, n\}\}, q\rangle$$

é uma tripla ordenada, onde q é uma função com domínio

$$\{\varnothing, \{m\}, \{n\}, \{m, n\}\}$$

dada por

$$q(\varnothing) = 0,$$
$$q(\{m\}) = \frac{1}{3},$$
$$q(\{n\}) = \frac{2}{3}$$

e

$$q(\{m, n\}) = 1,$$

então tal tripla ordenada também é um espaço de probabilidades.

Exemplo 9.6. ✍ *Seguindo os dois últimos* Exemplos, *se*

$$\langle \{m, n\}, \{\varnothing, \{m\}, \{n\}, \{m, n\}\}, r\rangle$$

é uma tripla ordenada, onde r é uma função com domínio

$$\{\varnothing, \{m\}, \{n\}, \{m, n\}\}$$

dada por

$$r(\varnothing) = 0,$$
$$r(\{m\}) = 1,$$
$$r(\{n\}) = 0$$

e

$$r(\{m, n\}) = 1,$$

então tal tripla ordenada também é um espaço de probabilidades.

Neste caso, a probabilidade de ocorrer coroa *é zero, enquanto a probabilidade de ocorrer o evento* cara *é um.*

Os três últimos Exemplos são contraexemplos para a recíproca do Teorema 9.6 da Monotonicidade de Probabilidades.

O evento impossível $\varnothing$ é a interseção entre eventos mutuamente excludentes. Sua probabilidade, como já mostrado, é sempre zero. No entanto, no último Exemplo temos um evento não vazio com probabilidade zero. Se definirmos uma restrição r' da função r sobre o domínio $\{\varnothing, \{m\}, \{m, n\}\}$, com o propósito de 'nos livrarmos' do evento não vazio com probabilidade zero, perceberemos que o novo conjunto $\{\varnothing, \{m\}, \{m, n\}\}$ não é uma σ-álgebra. Com efeito, o postulado II da Definição 9.2 não é teorema nesta interpretação. Logo, r' não é uma probabilidade.

Se ainda insistirmos na ideia de nos livrarmos do evento não vazio com probabilidade zero, a melhor estratégia é redefinir o espaço amostral como, no caso do Exemplo acima,

$$\{m, n\} - \{n\} = \{m\}.$$

Em seguida definimos uma restrição da probabilidade r para o novo espaço amostral.

Porém, precisamos também aprender a conviver com eventos não vazios de probabilidade zero. O problema, não obstante, é o ímpeto natural de pessoas interpretarem probabilidade nula como sinônimo de impossibilidade de ocorrência do evento (em um sentido intuitivo). Vemos no próximo Exemplo que a interpretação intuitiva de probabilidades é um pouco mais ardilosa do que isso.

Exemplo 9.7. *Seja* $\mathfrak{b} = \langle [0,1], \wp([0,1]), p \rangle$, *onde*

$$p : \wp([0,1]) \to [0,1]$$

é uma função definida da seguinte maneira.

$$p(x) = \begin{cases} b - a & \textit{se } x \textit{ é o intervalo real } [a,b] \textit{ ou} \\ & (a,b) \textit{ ou } [a,b) \textit{ ou } (a,b], \\ & \textit{onde } a \in [0,1] \wedge b \in [0,1] \wedge a < b \\ 0 & \textit{se } x \textit{ é o intervalo degenerado } [a,a], \\ & \textit{onde } a \in [0,1] \\ \sum_{s \in y} p(s) & \textit{se existe conjunto enumerável } y \textit{ tal que} \\ & \textit{seus elementos são intervalos dois a dois} \\ & \textit{disjuntos contidos em } [0,1] \textit{ e } x \subseteq \bigcup_{s \in y} s \\ \text{☕} & \textit{para os demais casos.} \end{cases}$$

O símbolo ☕ designa o fato de que o leitor precisará de muitas xícaras de café para uma plena compreensão deste Exemplo, *uma vez que não apresentamos aqui todos os detalhes. Como* *dizia Paul Erdös, 'matemático é uma máquina que transforma café em teoremas'. Mas não se desespere! O que realmente interessa aos nossos propósitos elementares é detalhado aqui.*

Existe uma conhecida σ-álgebra Σ definida sobre $[0,1]$, de modo que $\Sigma \subset \wp([0,1])$. Trata-se do conjunto de todos os conjuntos de Borel *contidos em $[0,1]$. Para efeitos ilustrativos deste* Exemplo, *basta que o leitor saiba as seguintes informações:*

- I: *qualquer intervalo, degenerado ou não, contido em $[0,1]$, é um conjunto de Borel;*
- II: *qualquer união enumerável de conjuntos de Borel é um conjunto de Borel.*

Uma vez definido o domínio da probabilidade p como $\wp([0,1])$, item ☕ nos diz que, se x é elemento de $\wp([0,1])$ mas não é um conjunto de Borel, então $p(x) = 0$. No entanto, há conjuntos de Borel com probabilidade zero também. Qualquer intervalo fechado degenerado é um conjunto de Borel com probabilidade nula. Até mesmo o conjunto de todos os reais racionais pertencentes a $[0,1]$ é um conjunto de Borel. Isso porque esse conjunto é enumerável (conforme Seção 34) e, além disso, é a união arbitrária de todos os intervalos degenerados $[r,r]$, onde r é um real racional. Por conta disso, tal conjunto tem probabilidade zero (o que implica que a probabilidade do evento x definido por todos os reais irracionais pertencentes a $[0,1]$ é 1).

Neste caso, $\mathfrak{b}$ é um espaço de probabilidades. Em particular,

$$p([0,1]) = 1 - 0 = 1, \ \ p\left(\left[\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right]\right) = 0 \ \ e$$

$$p\left(\left(\frac{1}{3}, \frac{1}{2}\right] \cup \left(\frac{2}{3}, \frac{3}{4}\right)\right) = \frac{1}{2} - \frac{1}{3} + \frac{3}{4} - \frac{2}{3} = \frac{3}{12}.$$

De acordo com a Seção 24, é possível definir uma função

$$c : \{[0,1]\} \to [0,1]$$

tal que $c([0,1]) = r$, onde r é um elemento escolhido pelo Axioma da Escolha de ZFC, sendo r um real do intervalo $[0,1]$. Tal função c é conhecida como função escolha. *Neste caso, o real r permite definir o singleton $\{r\}$ que, por sua vez, é o intervalo fechado degenerado $[r,r]$. O Axioma da Escolha escolheu um evento do espaço de probabilidades acima com probabilidade nula.*

Se pensarmos numa analogia meramente intuitiva, podemos interpretar o espaço amostral $[0,1]$ do último Exemplo como um alvo. Neste contexto, o Axioma da Escolha pode ser interpretado como um dardo lançado em direção ao alvo. A probabilidade do dardo acertar o alvo é um. Isso porque $p([0,1]) = 1$. A probabilidade do dardo acertar a região $(\frac{1}{3}, \frac{1}{2}] \cup (\frac{2}{3}, \frac{3}{4})$ do alvo é $\frac{3}{12}$. Em contrapartida, a probabilidade do dardo acertar um ponto $\{s\}$ específico (onde s é um real pertencente ao alvo e, portanto, $\{s\}$ é um evento da σ-álgebra) é zero. Apesar disso, o dardo acerta de fato um ponto $\{r\}$, cuja probabilidade é nula.

Definição 9.5. *Seja $\mathfrak{p} = \langle \Omega, \Sigma, p \rangle$ um espaço de probabilidades no qual Ω é equipotente a um ordinal finito n (ou seja, Ω tem n elementos). Dizemos que $\mathfrak{p}$ é um espaço* equiprovável *sss para todo evento unitário $\{x\}$ de Σ tivermos*

$$p(\{x\}) = \frac{1}{n}.$$

Exemplo 9.8. Exemplo 9.4 *se refere a um espaço equiprovável.* Exemplos 9.5 e 9.6 *se referem a espaços não equiprováveis.*

Seção 103

Probabilidade condicional

Sumário
Índice
Rede

O conceito de espaço de probabilidade apresentado na última Seção encontra ampla aplicabilidade. Do ponto de vista matemático, uma função de probabilidade é um caso muito particular de *medida*, tema de expressivo interesse em vários ramos da matemática e da matemática aplicada.

Se trocarmos o axioma p3 da Definição 9.4 pela fórmula

$$p(\varnothing) = 0,$$

a tripla ordenada $\langle \Omega, \Sigma, p \rangle$ daquela definição passa a ser um *espaço de medidas* e a função p é uma *medida* definida sobre Σ.

Ou seja, σ-álgebras não são usadas apenas no estudo de probabilidades, mas também na investigação de medidas. No último caso, medidas servem de fundamentação para o conceito de *integral de Lebesgue* [29]. Outra aplicação de σ-álgebras é em estatística, na qualificação de *estatística suficiente*.

No entanto, historicamente, probabilidades nasceram como propostas para teorias físicas. Tal tradição histórica obviamente persiste até os dias de hoje, justamente por conta do grande sucesso dessa ferramenta.

Tal contexto histórico é responsável pela confusão muito comum entre os conceitos de *medida* e *medição*. Afinal, toda probabilidade é uma medida, enquanto probabilidades são também fortemente motivadas por problemas do mundo real.

Medidas são funções p em um espaço de medidas, conforme a definição acima. Medições, por sua vez, são processos físicos de comparação entre objetos ou eventos do mundo real com outros objetos e eventos do mundo real. Em inúmeros contextos sócio-linguísticos as palavras 'medida' e 'medição' são sinônimos. Mas, no que se refere a um discernimento entre matemática e física, é preciso ter muito cuidado.

No sentido acima exposto, a definição de espaço de probabilidades funciona muito bem para aquilo que chamamos de *uma tentativa*. Jogando uma moeda, é possível modelar matematicamente a probabilidade de ocorrer o evento *cara* ou o evento *coroa* em uma única tentativa. Se a medida de probabilidade para o evento cara é $\frac{1}{2}$, experimentos no mundo real devem apontar que a medição de ocorrências de cara, para uma moeda não viciada, é aproximadamente 50% (para uma quantia 'suficientemente grande' de tentativas).

Em um baralho de cartas misturadas, em princípio, é possível determinar a probabilidade de alguém escolher um dois de copas em uma única tentativa. Basta usarmos a Definição 9.4, no sentido de definir uma σ-álgebra que mapeie todos os possíveis eventos e uma função de probabilidades.

No entanto, aplicações de probabilidades exigem mais. É de interesse respondermos qual é a probabilidade de um evento, se um evento anterior já ocorreu. O que está implícito neste discurso é que, em uma primeira tentativa, há a probabilidade de ocorrência de um evento. Em uma próxima tentativa, a probabilidade de um evento pode depender do que já ocorreu. O que já ocorreu é uma tentativa. O que ocorre a seguir é uma nova tentativa.

No caso de um jogo de cara-ou-coroa, a probabilidade de uma moeda não viciada resultar em cara, após uma primeira tentativa, é $\frac{1}{2}$. Em uma segunda tentativa, ao se jogar a mesma moeda, a probabilidade de ocorrer o evento cara continua sendo $\frac{1}{2}$. Este é um exemplo bem conhecido de eventos que são independentes entre si. Porém, no caso do baralho de cartas misturadas mencionado acima, se as cartas escolhidas não forem devolvidas ao baralho, as probabilidades de cada evento dependem da ocorrência de eventos em tentativas anteriores. Temos aqui uma situação de eventos dependentes daquilo que já ocorreu ou daquilo que poderia ter ocorrido, dependendo da interpretação que se promova para fins de mapeamento do mundo real.

Logo, a pergunta natural é como mapear matematicamente o conceito de *uma tentativa*. Mais importante, como mapear a ideia de que tentativas ocorrem sequencialmente, no sentido de que devemos discernir tentativas anteriores de tentativas posteriores?

Pois bem. Mais uma vez ZF conta com aparato suficiente para lidar com essa situação: trata-se do conceito de produto cartesiano (Definição 3.7) e, consequentemente, de par ordenado. Com efeito, um par ordenado (m, n) é igual a um par ordenado (p, q) sss $m = p$ e $n = q$ (Teorema 3.4). Isso significa, para efeitos práticos, que a ordem em que os termos ocorrem em um par ordenado é relevante para fins de identificação do mesmo. Neste sentido, podemos interpretar a primeira entrada m de um par ordenado (m, n) como primeira tentativa, e a segunda entrada n como segunda tentativa. Uma vez que pares ordenados permitem definir r-uplas ordenadas, onde r é um ordinal finito, isso significa que podemos lidar com quantas tentativas quisermos, desde que (por enquanto) seja uma quantia finita de tentativas.

Outra questão importante é a seguinte: qualquer que seja a proposta para a definição de uma probabilidade condicional a eventos de tentativas anteriores, ela deve ser consistente com a existência de eventos independentes entre si, bem como com a existência de eventos dependentes. Obviamente os conceitos de eventos dependentes e eventos independentes são meras arbitrariedades matemáticas. Mas, uma vez que probabilidades foram concebidas para mapear fenômenos do mundo real, é altamente desejável que elas façam isso de forma bem sucedida do ponto de vista experimental. A discussão sobre o sucesso de probabilidades é feita na Seção 105.

De agora em diante, por uma questão de mera conveniência, eventualmente podemos nos referir a uma σ-álgebra como uma *álgebra de eventos*.

Definição 9.6. *Sejam x e y conjuntos não vazios e Σ uma σ-álgebra de $x \times y$. Se $r_1 \subseteq x$ e $r_2 \subseteq y$, definimos e_{r_1} e e_{r_2} como se segue.*

$$\forall t \left(t \in r_1 \Leftrightarrow \bigcup_{\{(t,k)\} \in \Sigma} \{(t,k)\} \in e_{r_1} \right)$$

e

$$\forall t \left(t \in r_2 \Leftrightarrow \bigcup_{\{(k,t)\} \in \Sigma} \{(k,t)\} \in e_{r_2} \right).$$

Dizemos que r_1 é 'evento' de x correspondente *a e_{r_1} de Σ, e r_2 é 'evento' de y* correspondente *a e_{r_2} de Σ. Equivalentemente, e_{r_1} é* correspondente *a r_1 e e_{r_2} é* correspondente *a r_2.*

Os subscritos 1 e 2 em e_{r_1} e e_{r_2} são usados para discernir primeira entrada de segunda entrada em um par ordenado pertencente a um evento de Σ.

Importantíssimo notar o emprego de aspas nos conceitos de 'evento' de x e 'evento' de y. Os motivos para isso são os seguintes:

i: Estamos falando de um conjunto $x \times y$ munido de uma σ-álgebra. Logo, não estamos tratando (ainda) de um espaço de probabilidades, uma vez que sequer mencionamos qualquer função de probabilidades. Obviamente nossa meta é empregar esses conceitos para tratar de certos tipos especiais de espaços de probabilidades. Mas eventos se referem apenas a espaços de probabilidades, e não necessariamente a elementos de uma σ-álgebra. Com efeito, σ-álgebras são empregadas em muitas outras situações, além de probabilidades.

ii: Muito mais importante é o fato de estarmos falando de subconjuntos de x e de y. Portanto, não estamos tratando apenas de subconjuntos de $x \times y$ na última definição. Se usarmos $x \times y$ como espaço amostral e Σ como álgebra de eventos, para definirmos um espaço de probabilidades (acrescentando uma função de probabilidades com domínio Σ), apenas os elementos de Σ são eventos deste espaço, os quais são subconjuntos de $x \times y$. Nenhum elemento de uma σ-álgebra de $x \times y$ é elemento ou subconjunto de x ou sequer elemento de uma σ-álgebra de x. Além disso, observar que o produto cartesiano de σ-álgebras não é uma σ-álgebra de conjunto algum. Daí a necessidade de cuidados especiais!

Notar também que o emprego de bicondicionais na definição acima. Todo 'evento' r_1 de x corresponde a um e apenas um elemento e_{r_1} de Σ, e todo elemento e_{r_1} de Σ corresponde a um e apenas um subconjunto r_1 de x. Comentário análogo vale para 'eventos' de y.

O que a última definição estabelece é que um evento e_{r_1} da álgebra de eventos de $x \times y$ é o conjunto de todos os pares ordenados (t, k) de Σ tais que t pertence a r_1 (lembrando que r_1 é um subconjunto qualquer de x). Analogamente, e_{r_2} da álgebra de eventos de $x \times y$ é o conjunto de todos os pares ordenados (k, t) de Σ tais que t pertence a r_2 (lembrando que r_2 é um subconjunto qualquer de y).

A definição acima não é usual em textos sobre sobre probabilidades. Em geral, a literatura não é clara sobre esses conceitos, apelando muitas vezes a uma visão meramente intuitiva sobre probabilidades condicionais. Neste livro, porém, nos comprometemos ao esclarecimento sobre como ZF e ZFC permitem fundamentar grande parte da prática matemática. Daí a necessidade da última definição, antes que possamos qualificar o que é uma probabilidade condicional.

Antes de irmos ao ponto principal desta Seção, se a σ-álgebra Σ de $x \times y$ for $\wp(x \times y)$, naturalmente cada 'evento' de x ou de y é elemento das σ-álgebras $\wp(x)$ e $\wp(y)$, respectivamente.

Definição 9.7. *Seja $\langle x \times y, \Sigma, p\rangle$ um espaço de probabilidades. Seja*

$$\mathfrak{p}_c : \Sigma \times \Sigma \to [0, 1]$$

uma função definida como

$$\mathfrak{p}_c(e_1, e_2) = \frac{p(e_1 \cap e_2)}{p(e_2)},$$

onde $p(e_2) \neq 0$ e $\mathfrak{p}_c(e_1, e_2)$ é denotada abreviadamente como $\mathfrak{p}_c(e_1 \mid e_2)$.

Se r_1 é um 'evento' de x correspondente ao evento e_{r_1} de Σ, e r_2 é um 'evento' de y correspondente ao evento e_{r_2} de Σ, dizemos que a probabilidade de r_1 condicionada a r_2 *é $\mathfrak{p}_c(e_{r_1} \mid e_{r_2})$. Além disso, a* probabilidade de r_2 condicionada a r_1 *é $\mathfrak{p}_c(e_{r_2} \mid e_{r_1})$.*

A função $\mathfrak{p}_c$ *é chamada de* probabilidade condicional. *Evento* e_{r_2} *é chamado de* condicionante *em* $\mathfrak{p}_c(e_{r_1} \mid e_{r_2})$, *enquanto* e_{r_1} *é o* condicionante *em* $\mathfrak{p}_c(e_{r_2} \mid e_{r_1})$.

Por abuso de linguagem, podemos nos referir a $\mathfrak{p}_c(e_{r_1} \mid e_{r_2})$ como a *probabilidade de* e_{r_1} *condicionada a* e_{r_2}, e $\mathfrak{p}_c(e_{r_2} \mid e_{r_1})$ como a *probabilidade de* e_{r_2} *condicionada a* e_{r_1}.

Por conta da monotonicidade da probabilidade (Teorema 9.6),

$$p(e_1 \cap e_2) \leq p(e_2).$$

Logo,

$$\mathfrak{p}_c(e_1 \mid e_2) \leq 1,$$

para quaisquer pares ordenados (e_1, e_2) de eventos de Σ.

No entanto, notar que a probabilidade condicional não é uma função de probabilidade, apesar de ser definida a partir de uma. Com efeito, seu domínio não é uma álgebra de eventos, mas um produto cartesiano de uma álgebra de eventos por ela mesma. Como já foi destacado, o produto cartesiano entre uma álgebra de eventos e ela mesma não é uma álgebra de eventos.

Apesar disso, para efeitos práticos, a probabilidade condicional opera como uma probabilidade de ocorrência de um evento, desde que condicionada à ocorrência de outro evento. Justificamos essa última afirmação no próximo parágrafo.

Dado um evento e_{r_2}, as imagens $\mathfrak{p}_c(e_{r_1} \mid e_{r_2})$ da probabilidade condicional $\mathfrak{p}_c$ são imagens de uma função de probabilidade $\mathfrak{p}_{e_{r_2}}$ com domínio Σ. Logo, podemos definir o conjunto $\mathfrak{p}_{c_{e_{r_2}}}$ como a união arbitrária de todas as probabilidades $\mathfrak{p}_{e_{r_2}}$, para todos os possíveis e_{r_2}. Comentário análogo vale para a probabilidade de e_{r_2} condicionada a e_{r_1}. Isso faz da função de probabilidade condicional $\mathfrak{p}_c$ a união do conjunto $\mathfrak{p}_{c_{e_{r_2}}}$ de todas as probabilidades $\mathfrak{p}_{e_{r_2}}$ com o conjunto $\mathfrak{p}_{c_{e_{r_1}}}$ de todas as probabilidades $\mathfrak{p}_{e_{r_1}}$.

Ou seja, uma probabilidade condicional $\mathfrak{p}_c(e \mid e')$ tem as mesmas imagens de uma função de probabilidades aplicada sobre o evento e, desde que o evento e' de Σ seja fixado.

Outra observação relevante é o fato de que a probabilidade de um evento condicionada a um condicionante pode ser interpretada como um fator de correção aplicado à probabilidade de ocorrência do condicionante. Com efeito, basta observar que

$$p(e_1 \cap e_2) = p(e_2)\mathfrak{p}_c(e_1 \mid e_2) = p(e_1)\mathfrak{p}_c(e_2 \mid e_1).$$

Logo,

$$\frac{p(e_2)}{p(e_1)} = \frac{\mathfrak{p}_c(e_2 \mid e_1)}{\mathfrak{p}_c(e_1 \mid e_2)}.$$

A igualdade acima é o célebre Teorema de Bayes, assunto a ser discutido na próxima Seção.

Exemplo 9.9. *Em um jogo com uma moeda não viciada, qual é a probabilidade de ocorrer cara após ocorrer coroa?*

Podemos definir o conjunto $x = \{m, n\}$, onde $\{m\}$ é o 'evento' cara, e $\{n\}$ é o 'evento' coroa. Uma vez que a segunda tentativa usa a mesma moeda empregada na primeira tentativa, devemos definir uma álgebra de eventos para o espaço amostral

$$x \times x = \{(m, m), (m, n), (n, m), (n, n)\},$$

a qual é $\wp(x \times x)$. Notar que $\wp(x \times x)$ tem 16 elementos. Cada um desses elementos é um evento.

Assumir que a moeda é não viciada equivale a afirmar que a probabilidade com domínio $\wp(x \times x)$ é dada da seguinte maneira:

$$p(e) = \frac{\textit{número de elementos de } e}{\textit{número de elemento de } x \times x},$$

i.e., $\langle x \times x, \wp(x \times x), p\rangle$ é um espaço equiprovável (Definição 9.5).

O evento condicionante e_2, correspondente ao 'evento' coroa, é $\{(m, n), (n, n)\}$.

O evento e_1 sobre o qual queremos determinar a probabilidade condicional é $\{(m, n), (m, m)\}$.

Logo, $e_1 \cap e_2 = \{(m, n)\}$. Isso implica que

$$p(e_1 \cap e_2) = \frac{1}{4},$$

enquanto

$$p(e_2) = \frac{2}{4}.$$

Logo,

$$\mathfrak{p}_c(e_1 \mid e_2) = \frac{\frac{1}{4}}{\frac{2}{4}} = \frac{1}{2}.$$

Definição 9.8. *Sejam* $\langle x \times y, \Sigma, p\rangle$ *um espaço de probabilidades e* $\mathfrak{p}_c$ *uma probabilidade condicional com domínio* $\Sigma \times \Sigma$. *Dizemos que o evento* e_1 é independente do evento e_2 *sss*

$$\mathfrak{p}_c(e_1 \mid e_2) = p(e_1).$$

Caso contrário, dizemos que e_1 *e* e_2 *são eventos* dependentes.

Exemplo 9.10. Exemplo 9.9 *ilustra evento* e_1 *independente do evento* e_2. *Uma vez que o espaço de probabilidades é equiprovável,*

$$p(e_1) = \frac{2}{4} = \frac{1}{2}.$$

Mas este é exatamente o mesmo valor de $\mathfrak{p}_c(e_1 \mid e_2)$.

Exemplo 9.11. *Consideremos dois dados de seis faces, numeradas de 1 a 6, não viciados. Digamos que um dos dados é azul e o outro é vermelho. Qual é a probabilidade de ocorrer uma face numerada 2 no dado azul, se a soma das faces numeradas de ambos os dados for menor do que 6?*

Neste caso, podemos definir o conjunto $x = \{1, 2, 3, 4, 5, 6\}$, *onde cada evento unitário é uma ocorrência de uma face numerada de um dado qualquer. Definimos uma álgebra de eventos para o espaço amostral*

$$x \times x,$$

a qual é

$$\wp(x \times x).$$

Notar que $x \times x$ *tem* 36 *elementos e* $\wp(x \times x)$ *tem* 2^{36} *elementos (i.e., mais de* 68 *bilhões de elementos). Cada um desses elementos é um evento.*

O que permite discernir as tentativas, neste caso, é a discernibilidade dos dados, uma vez que um é azul e o outro é vermelho.

Assumir que os dados são não viciados equivale a afirmar que a probabilidade com domínio $\wp(x \times x)$ *é dada da seguinte maneira:*

$$p(e) = \frac{\text{número de elementos de } e}{\text{número de elemento de } x \times x},$$

ou seja,

$$\langle x \times x, \wp(x \times x), p\rangle$$

é um espaço equiprovável.

O evento condicionante e_2, *correspondente ao 'evento' 'a soma das faces é menor do que* 6*', é*

$$\{(1,1),(1,2),(1,3),(1,4),(2,1),(2,2),(2,3),(3,1),(3,2),(4,1)\}.$$

O evento e_1 *sobre o qual queremos determinar a probabilidade condicional é*

$$\{(2,1),(2,2),(2,3),(2,4),(2,5),(2,6)\},$$

onde o 'evento' $\{2\}$ *se refere ao dado azul, aqui retratado na primeira entrada dos pares ordenados acima.*

Logo, $e_1 \cap e_2 = \{(2,1),(2,2),(2,3)\}$. *Isso implica que*

$$p(e_1 \cap e_2) = \frac{3}{36},$$

enquanto

$$p(e_2) = \frac{10}{36}.$$

Logo,

$$\mathfrak{p}_c(e_1 \mid e_2) = \frac{\frac{3}{36}}{\frac{10}{36}} = \frac{3}{10}.$$

Notar que

$$p(e_1) = \frac{6}{36} = \frac{1}{6} \neq \frac{3}{10}.$$

Logo, e_1 *e* e_2 *são eventos dependentes.*

Teorema 9.9. *Dois eventos* e_1 *e* e_2 *em um espaço de probabilidades com probabilidade* p *e probabilidade condicional* $\mathfrak{p}_c$ *são independentes sss* $p(e_1 \cap e_2) = p(e_1)p(e_2)$.

Demonstração: ✍ Basta usar a definição de eventos independentes dada a partir de probabilidade condicional (Definição 9.8), bem como o próprio conceito de probabilidade condicional (Definição 9.7).

STOP Aqui é cabível um alerta. É uma prática comum na literatura a 'definição' de probabilidade condicional da seguinte maneira:

$$p(e_1 \mid e_2) = \frac{p(e_1 \cap e_2)}{p(e_2)},$$

sem deixar claro qual é o domínio da nova função introduzida. O próprio emprego da mesma letra p usada para duas funções diferentes (uma probabilidade e uma probabilidade condicional) também pode colaborar para dificuldades nos estudos. Com efeito, apesar da probabilidade condicional depender de uma álgebra de eventos, seu domínio não é uma álgebra de eventos.

Seção 104

Teorema de Bayes

Sumário
Índice
Rede

Em uma noite de nevoeiro, um táxi é envolvido em um grave acidente, com vítima fatal. Mas o motorista fugiu no mesmo veículo, sem prestar socorro.

Uma única testemunha prestou relatório às autoridades. Em seu depoimento, ela afirmou que o táxi era azul. Como parte da investigação, a polícia testou a confiabilidade da testemunha, submetendo-a a condições semelhantes de visibilidade. Afinal, naquela cidade havia apenas duas operadoras de táxis. Em uma delas os veículos eram todos azuis. Em outra, eram todos verdes. Dependendo das condições de luminosidade, não é de espantar que alguém confunda uma cor com a outra.

Após vários testes, foi finalmente avaliado que a testemunha era capaz de dizer corretamente a cor do táxi em oitenta por cento das simulações. Logo, temos um depoimento com elevado grau de confiabilidade, apesar das condições adversas. A questão agora é a seguinte: o que mais provavelmente aconteceu naquela dramática noite? A vítima foi atropelada por um táxi azul ou por um táxi verde?

O problema acima foi formulado pela primeira vez por Amos Tversky e Daniel Kahneman [20]. O objetivo desses psicólogos israelenses era avaliar uma tendência natural de pessoas julgarem fatos com rapidez e eficiência, mas sem dados completos que exijam tempo e energia para pensar racionalmente. O testemunho confiável certamente não é o bastante para lidar com o problema proposto. É neste momento que entra em cena um exemplo muito simples de aplicação do Teorema de Bayes, tema desta Seção.

A ideia é a seguinte. Uma informação é a probabilidade do táxi ser azul, se a testemunha relata a cor azul. Outra, é a probabilidade da testemunha relatar a cor azul, se o táxi é azul. Estes não são necessariamente o mesmo número.

Teorema 9.10 (Teorema de Bayes). *Sejam*

$$\langle x \times y, \Sigma, p \rangle$$

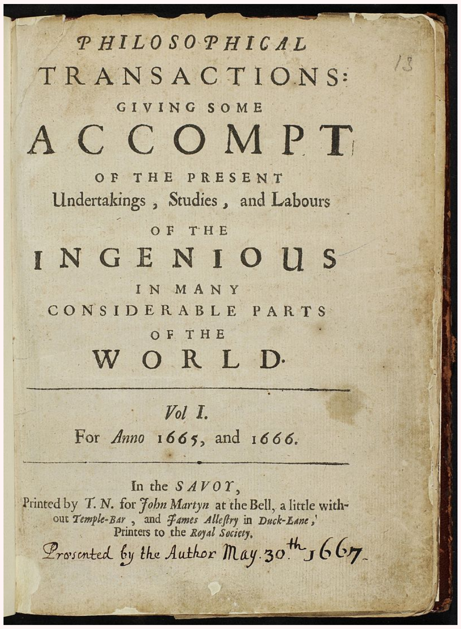
um espaço de probabilidades e $\mathfrak{p}_c$ *uma probabilidade condicional com domínio* $\Sigma \times \Sigma$. *Logo,*

$$\mathfrak{p}_c(e_1 \mid e_2) = \frac{\mathfrak{p}_c(e_2 \mid e_1) p(e_1)}{p(e_2)},$$

se $e_1 \neq e_2$ *e* $p(e_2) \neq 0$.

Demonstração: A demonstração já foi feita no último parágrafo que antecede o Exemplo 9.9.

O resultado acima é devido ao reverendo presbiteriano Thomas Bayes, apesar deste jamais ter publicado seu mais famoso trabalho. Foi graças a Richard Price que hoje conhecemos o autor do célebre teorema. Price editou o texto original de Bayes em *Philosophical Transactions*, o primeiro periódico da história a ser dedicado exclusivamente à ciência.


PHILOSOPHICAL
TRANSACTIONS:
GIVING SOME
ACCOMPT
OF THE PRESENT
Undertakings, Studies, and Labours
OF THE
INGENIOUS
IN MANY
CONSIDERABLE PARTS
OF THE
WORLD.

Vol I.
For Anno 1665, and 1666.

In the SAVOY,
Printed by T. N. for John Martyn at the Bell, a little without Temple-Bar, and James Allestry in Duck-Lane, Printers to the Royal Society.
Presented by the Author May 30th 1667

Capa do primeiro volume de Philosophical Transactions

Fonte: Wikipedia.

Exemplo 9.12. *No problema proposto por Tversky e Kahneman, colocado no início desta Seção, a probabilidade da testemunha relatar a cor corretamente é importante, sem sombra de dúvida. Mas devemos considerar também a probabilidade de um táxi ser de fato azul.*

Chamemos de $\{\mathcal{A}\}$ *e* $\{\mathcal{V}\}$ *os 'eventos' 'táxi azul' e 'táxi verde', respectivamente. Também chamemos de* $\{\mathcal{TA}\}$ *e* $\{\mathcal{TV}\}$ *os 'eventos' 'testemunha afima que táxi era azul' e 'testemunha afima que táxi era verde', respectivamente.*

Consideremos dois possíveis cenários.

Cenário i: *A probabilidade* $p(\{\mathcal{V}\})$ *de um táxi ser verde é maior do que a probabilidade da testemunha relatar corretamente a cor do veículo. Para fins de ilustração, digamos que* $p(\{\mathcal{V}\}) = 0,85$.

Cenário ii: *A probabilidade* $p(\{\mathcal{V}\})$ *de um táxi ser verde é menor do que a probabilidade da testemunha relatar corretamente a cor do veículo. Para fins de ilustração, digamos que* $p(\{\mathcal{V}\}) = 0,5$.

Começamos com o Cenário i.

De acordo com o problema proposto, a probabilidade da testemunha relatar cor azul, no caso em que o táxi tem cor azul, é $0,80$. *Logo,*

$$\mathfrak{p}_c(\{\mathcal{TA}\} \mid \{\mathcal{A}\}) = 0,80.$$

O que queremos responder é a probabilidade do táxi ser azul, diante do fato da testemunha afirmar que o veículo era azul. Ou seja, precisamos calcular

$$\mathfrak{p}_c(\{\mathcal{A}\} \mid \{\mathcal{TA}\}).$$

Teorema de Bayes estabelece que

$$\mathfrak{p}_c(\{\mathcal{A}\} \mid \{\mathcal{TA}\}) = \frac{\mathfrak{p}_c(\{\mathcal{TA}\} \mid \{\mathcal{A}\})p(\{\mathcal{A}\})}{p(\{\mathcal{TA}\})}.$$

Uma vez que táxis verdes e táxis azuis são eventos mutuamente excludentes, sabemos que $p(\{\mathcal{A}\}) = 1 - p(\{\mathcal{V}\})$.

Isso, no Cenário i, *implica que* $p(\{\mathcal{A}\}) = 0,15$. *Mas como conhecer* $p(\{\mathcal{TA}\})$*?*

$p(\{\mathcal{TA}\})$ é a probabilidade da testemunha relatar que o táxi era azul. Mas só existem duas possibilidades: (i) a testemunha diz que é azul, sendo o táxi azul; (ii) a testemunha diz que é azul, sendo o táxi verde. Lembrando que táxis azuis e táxis verdes são mutuamente excludentes, logo,

$$\{\mathcal{A}\} \cap \{\mathcal{TA}\} \ \text{ e } \ \{\mathcal{V}\} \cap \{\mathcal{TA}\}$$

são mutuamente excludentes.

Portanto,

$$p(\{\mathcal{TA}\}) = p(\{\mathcal{A}\} \cap \{\mathcal{TA}\}) + p(\{\mathcal{V}\} \cap \{\mathcal{TA}\}).$$

Aplicando o Teorema de Bayes sobre as duas parcelas à direita da igualdade na última equação, temos

$$p(\{\mathcal{TA}\}) = p(\{\mathcal{A}\})\mathfrak{p}_c(\{\mathcal{TA}\} \mid \{\mathcal{A}\}) + p(\{\mathcal{V}\})\mathfrak{p}_c(\{\mathcal{TA}\} \mid \{\mathcal{V}\}),$$

lembrando que $\mathfrak{p}_c(\{\mathcal{TA}\} \mid \{\mathcal{V}\}) = 0,2$, *uma vez que os eventos 'dizer a cor corretamente' e 'dizer a cor errada' são mutuamente excludentes e* $\mathfrak{p}_c(\{\mathcal{TA}\} \mid \{\mathcal{A}\}) = 0,8$.

Podemos agora finalmente determinar $\mathfrak{p}_c(\{\mathcal{A}\} \mid \{\mathcal{TA}\})$:

$$\mathfrak{p}_c(\{\mathcal{A}\} \mid \{\mathcal{TA}\}) = \frac{0,15(0,80)}{0,15(0,80) + 0,85(0,2)}.$$

Isso nos dá um valor aproximado de $0,41$.

Ou seja, no Cenário i, *há 41% de chances do táxi ter sido azul, apesar do relato afirmar que era azul. Isso implica que mais provavelmente a vítima foi atropelada por um táxi verde, com 59% de chances de ser o caso.*

✍ Cenário ii *fica como exercício para o leitor.*

O Exemplo acima dá suporte a uma famosa declaração de Carl Sagan:

Alegações extraordinárias demandam evidências extraordinárias.

Se uma pessoa afirma ter visto uma nave extraterrestre, existe uma probabilidade de seu depoimento estar certo. No que se refere a tal probabilidade, dificilmente é aceitável que seja 1. Com efeito, pessoas podem interpretar erroneamente o que veem. Se a probabilidade de não ocorrência de naves extraterrestres em nossos quintais for maior do que a probabilidade de acerto em um depoimento dessa natureza, o que mais provavelmente ocorreu é que nenhuma nave extraterrestre visitou o quintal da testemunha.

O que foi dito no último parágrafo não é uma tentativa de desacreditar testemunhos, espalhados pelo mundo, de eventos extraordinários, como avistamentos de lobisomens ou fantasmas. Mas é um alerta das severas limitações para lidarmos com depoimentos de tais eventos.

Seção 105

Mapeamento com probabilidades

Sumário
Índice
Rede

Como já comentado, o conceito de probabilidade nasceu da necessidade humana de identificar padrões matemáticos específicos que parecem ocorrer em certos fenômenos do mundo real. Se admitirmos que é possível mapear tais fenômenos através de predicados conjuntistas, a questão a ser respondida é a seguinte: quais fenômenos podem ser mapeados por probabilidades?

Neste sentido, estamos falando de possíveis interpretações pretendidas para probabilidades, probabilidades condicionais e outros conceitos relacionados, como *distribuição de probabilidades*, *variáveis aleatórias*, *probabilidades fuzzy*, *probabilidades em espaços topológicos*, *geometria estocástica*, *processos estocásticos*, *processos de Markov*, *teoria das decisões*, *mecânica estatística*, *lógica indutiva*, *inferências Bayesianas* etc.

Se nos concentrarmos apenas em probabilidades e probabilidades condicionais, as interpretações pretendidas mais comuns para esses conceitos podem ser divididas em três grandes grupos.

Probabilidade como suporte epistemológico: Neste caso, probabilidades servem ao propósito de fornecer medidas objetivas de evidências de relações entre fenômenos. Por exemplo, levando em conta as atuais taxas de consumo de recursos naturais em parceria com avanços tecnológicos, a probabilidade de colapso irreversível da civilização nas próximas décadas é acima de 90% [7].

Probabilidade como grau de crença: Por exemplo, 'não tenho certeza que o mundo como o conhecemos deixará de existir em 2040, mas provavelmente será o caso'.

Probabilidade como conceito físico: Neste caso, probabilidades são aplicáveis a vários sistemas físicos, independentemente de avaliações subjetivas. Não importa o que uma pessoa pensa sobre uma moeda não viciada, se ela for jogada dez mil vezes, em aproximadamente cinco mil vezes resultará em coroa.

Os poucos exemplos aqui explorados se referem ao terceiro grupo. A literatura sobre probabilidades é suficiente para consumir o tempo de uma vida e ainda assim ficar muito longe de esgotar o tema.

Para finalizar esta brevíssima introdução ao assunto, retomemos a provocação colocada na Seção 100.

Sobre o caso Titan-Titanic, o mero acaso explica as coincidências entre o romance *Futility* e o naufrágio do Titanic? Afinal, assim como podemos obter duas coroas seguidas por mero acaso em um jogo de cara-ou-coroa, podemos atribuir o mesmo acaso ao romance? Ou estamos diante de uma intrigante evidência de profética previsão do futuro?

Pois bem. Infelizmente, teoria de probabilidades, por si só, não permite responder a essa questão. É necessário um espaço de probabilidades que mapeie o mundo real. Logo, como propor um espaço amostral para fins de avaliação? Esse espaço amostral deve incluir todos os romances já escritos e todos os eventos reais já ocorridos? Se a meta é comparar ficção com realidade, qual é o período de tempo a ser considerado? Décadas? Séculos? Milênios? Essa questão é relevante o bastante para ser digna de pesquisa?

Situação análoga ocorre com o caso Lincoln-Kennedy. Por exemplo, Lincoln teve quatro filhos. Kennedy teve três. Logo, não há coincidência entre ambos, no que se refere a número de filhos. Quais outros aspectos devem ser levados em consideração para que seja definido um espaço amostral? Devemos considerar apenas os presidentes dos Estados Unidos ou todos os estadistas que já existiram?

Com relação à loteria da Bulgária, esta já existia há mais de cinquenta anos, antes do inusitado evento das mesmas dezenas em dois sorteios consecutivos. Além disso, não é a única loteria do mundo. Se escolhermos como espaço amostral o conjunto de todas as loterias espalhadas pelo planeta, é tão inusitado assim que, em uma delas, ocorra tal coincidência? Afinal, os registros mais antigos de loterias datam da época da Dinastia Han, na China, há mais de dois mil anos.

A mesma mente humana, capaz de reconhecer padrões na natureza e transformá-los em matemática, também é capaz de interpretar coincidências como atos divinos ou até mesmo milagres. Não é por acaso que Georg Cantor percebia na teoria de conjuntos uma forma de conhecer a mente de Deus [57].

Seção 106

Resumo da ópera

Sumário
Índice
Rede

Qualificamos probabilidades e probabilidades condicionais na linguagem de ZFC. Neste contexto, mostramos que

- probabilidades são casos muito particulares de medidas;
- qualquer conjunto não vazio pode ser espaço amostral em um espaço de probabilidades;
- probabilidades foram concebidas para resolver problemas do mundo real;
- a interpretação intuitiva de probabilidades é um problema extraordinariamente difícil;
- probabilidades condicionais não são probabilidades, mas uma concatenação de probabilidades.

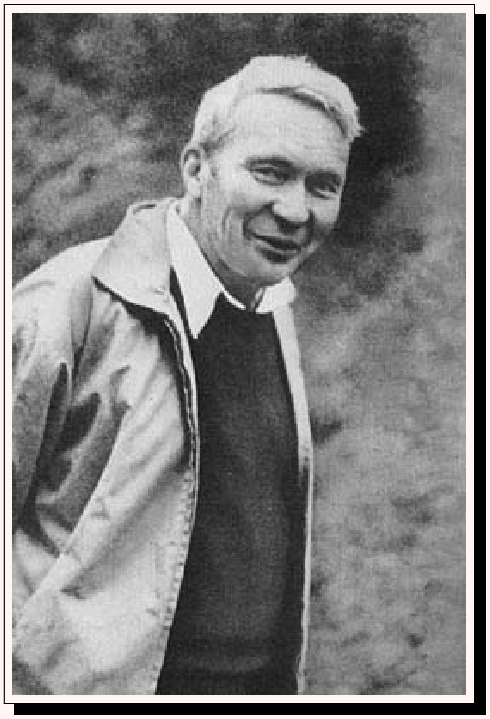
Seção 107

Notas históricas

Sumário
Índice
Rede

Em 1900 David Hilbert apresentou no Congresso Internacional de Matemáticos, de Paris, uma histórica palestra que, posteriormente, resultou em uma lista de 23 problemas que, em sua opinião, eram as principais questões legadas pelos matemáticos do século 19 aos do século 20. Os problemas por ele enunciados serviram, e ainda servem, para direcionar relevantes trabalhos nos campos da matemática pura e da matemática aplicada. Em sua lista, é de especial interesse aqui, o sexto problema:

Investigações sobre os fundamentos da geometria sugerem o problema: tratar do mesmo modo, por meio de axiomas, as ciências físicas nas quais a matemática desempenha papel importante: são prioritárias a teoria de probabilidades e a mecânica.



Andrey Kolmogorov

Fonte: https://www.kolmogorov.com/.

Claramente o texto acima mostra que teoria de probabilidades era um ramo das ciências físicas. Foi somente em 1933 que Andrey Kolmogorov apresentou uma solução parcial ao Sexto Problema de Hilbert, ao publicar um livro onde são introduzidos essencialmente os mesmos axiomas que o leitor encontra na Definição 9.4.

Mais do que um teórico, Kolmogorov colaborou com as forças armadas russas durante a Segunda Guerra Mundial, para proteger Moscou dos bombardeiros alemães. Sua arma: estatística. O matemático russo desenvolveu uma engenhosa distribuição estocástica de balões barragem.

Por conta de sua vasta obra, hoje o nome Kolmogorov está entre os mais citados em matemática, incluindo *homologia de Kolmogorov*, *espaços de Kolmogorov*, *paradoxo de Borel-Kolmogorov*, *critério de Kolmogorov*, *equação de Fisher-Kolmogorov* e *complexidade de Kolmogorov*, entre algumas dezenas de outras referências, incluindo, naturalmente, os axiomas de Kolmogorov para probabilidades.

PARTE 10

Informações complementares

Nesta parte discutimos sobre alguns assuntos complementares.

Seção 108

Newton-Raphson

Sumário
Índice
Rede

Seja $f : \mathbb{R} \to \mathbb{R}$ uma função diferenciável tal que existem valores em um intervalo $[a, b]$ onde $f(x)$ muda de sinal. A existência de pelo menos um $c \in [a, b]$, tal que $f(c) = 0$, é garantida pelo fato de f ser contínua (Teorema 5.25). Isso é consequência do Teorema do Valor Intermediário, mencionado muito brevemente na Seção 59.

Supor que existe apenas um $c \in [a, b]$ tal que $f(c) = 0$. Logo, podemos introduzir o método de Newton-Raphson, se certas condições forem atendidas. Detalhes em um bom livro sobre análise numérica.

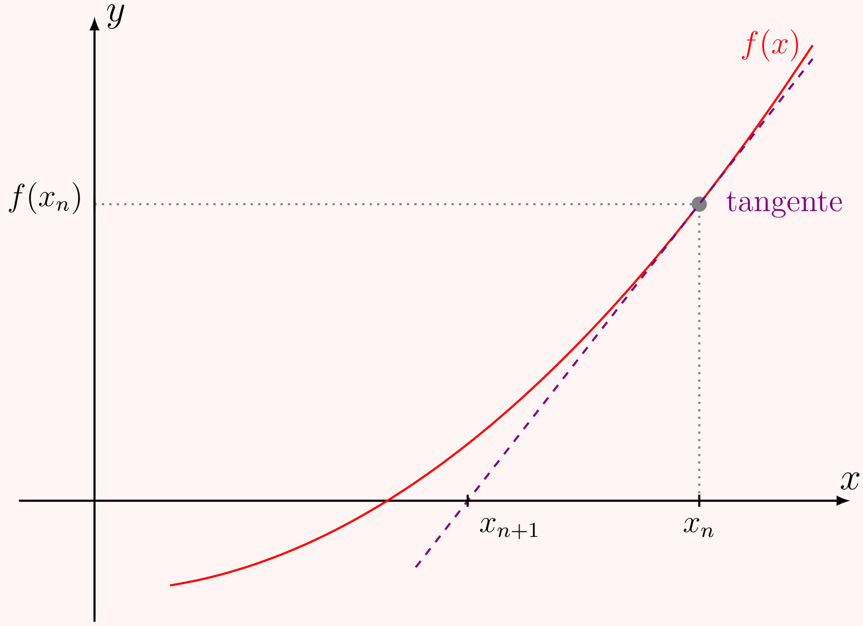
Se x_n é uma entrada de $f(x_n)$, x_{n+1} é obtido da seguinte maneira:

$$f'(x_n) = \frac{f(x_n) - 0}{x_n - x_{n+1}}.$$

A derivada $f'(x_n)$ é o coeficiente angular de uma reta definida por

$$(x_n, f(x_n)) \quad \text{e} \quad (x_{n+1}, 0),$$

conforme a imagem a seguir.



Logo,

$$x_n - x_{n+1} = \frac{f(x_n)}{f'(x_n)}.$$

Finalmente,

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

A fórmula acima descreve o *método de Newton-Raphson*. Observar que, se $f(x_n) = 0$, então $x_{n+1} = x_n$, para todo n natural. Além disso, se $f'(x_n) = 0$, o método de Newton-Raphson não é aplicável.

Exemplo 10.1. *Como calcular a raiz quadrada x de um número real $a \geq 0$?*

Por um lado temos $x = \sqrt{a}$. Logo, $x^2 = a$, o que implica que

$$x^2 - a = 0.$$

Logo, a raiz quadrada x de a é um zero de $f : \mathbb{R} \to \mathbb{R}$ dada por

$$f(x) = x^2 - a.$$

O método de Newton-Raphson permite obter aproximações com precisão arbitrária para $x = \sqrt{a}$. A função recursiva é dada por

$$x_{n+1} = x_n - \frac{x_n^2 - a}{2x_n},$$

sendo $x_0 \neq 0$ uma estimativa inicial para a raiz quadrada de a.

A função recursiva pode ser reescrita como

$$x_{n+1} = \frac{2x_n^2}{2x_n} - \frac{x_n^2 - a}{2x_n}.$$

Logo,

$$x_{n+1} = \frac{x_n + \frac{a}{x_n}}{2}.$$

Observar que esta é a mesma sequência introduzida na Seção 39, no caso em que x_0 é um número racional e $a = 2$. Naquela Seção há exemplos ilustrativos da sequência recursiva acima para valores de $x_0 = 2$ e $x_0 = 5$ com o propósito de obter aproximações racionais para $x = \sqrt{2}$.

Uma curiosidade histórica é que a equação

$$x_{n+1} = \frac{x_n + \frac{a}{x_n}}{2}$$

reproduz o *Método Babilônico para o Cálculo de Raiz Quadrada*, concebido há mais de quatro mil anos. Os matemáticos babilônicos não eram capazes de justificar o método acima por meio de cálculo diferencial, que só foi concebido no século 17. Mas empregavam a mesma ideia: aproximações de raízes quadradas por médias aritméticas.

O Exemplo acima pode ser estendido para aproximações de $x = \sqrt[m]{a}$, onde m é um inteiro positivo maior do que 1:

$$x_{n+1} = \frac{(m-1)x_n + \frac{a}{x_n^{m-1}}}{m}.$$

Ou seja, médias aritméticas de m termos (notar que todos os m termos ocorrem no numerador) podem ser usadas para obtermos aproximações de $\sqrt[m]{a}$, ideia essa que, aparentemente, jamais passou pela cabeça dos pensadores do Império Babilônico.

A representação decimal dessas aproximações permite estabelecer um critério de parada em função do número de casas de precisão desejada. No entanto, este método não exige que x_0 seja racional ou que f seja polinomial. Logo, pode ser aplicado a funções f envolvendo exponenciais, logaritmos, senos, co-senos, entre outras (incluindo composições não triviais), desde que (entre outras condições) sejam localmente diferenciáveis em um intervalo aberto onde há um zero de f.

Exemplo 10.2. *Um dos zeros de seno é π, conforme Seção 57. Portanto, para obter aproximações de π, basta aplicar o Método de Newton-Raphson, uma vez que seno é diferenciável.*

Se $f : \mathbb{R} \to \mathbb{R}$ é uma função dada por $f(x) = \operatorname{sen}(x)$, então

$$x_{n+1} = x_n - \frac{\operatorname{sen}(x_n)}{\cos(x_n)}.$$

Falta agora uma semente *x_0, para obter x_1, x_2, x_3 etc.*

Podemos usar, como inspiração, alguma referência histórica. Afinal, é divertido.

De acordo com Arquimedes de Siracusa, π está entre

$$\frac{223}{71} \quad e \quad \frac{22}{7},$$

ou seja, entre

$$3,140845070422535 \quad e \quad 3,142857142857142,$$

com precisão de 15 dígitos após a vírgula.

Outra referência histórica é a Bíblia Sagrada. No Primeiro Livro de Reis, Capítulo 7, Versículo 23, lê-se o seguinte:

> Fez o tanque de metal fundido, redondo, medindo quatro metros e meio de diâmetro e dois metros e vinte e cinco centímetros de altura. Era preciso um fio de treze metros e meio para medir a sua circunferência.

A versão acima, entre dezenas de traduções para o português brasileiro feitas desde o final do século 19, é de https://www.bibliaon.com/. Logo,

$$\pi = \frac{13 + \frac{1}{2}}{4 + \frac{1}{2}} = 3.$$

Mas o metro não era padrão adotado há 2700 anos. No original em hebraico clássico o tanque tem dez cúbitos curtos *de diâmetro e trinta* cúbitos curtos *de circunferência, o que corresponde a*

$$\pi = \frac{30}{10} = 3.$$

Na literatura especializada em história bíblica há extensas discussões sobre o cúbito curto e o cúbito longo, entre outras dezenas de unidades de medição empregadas no Antigo Testamento. Mas, *o que nos interessa é a proporção de valores. Logo, afirmar que o Primeiro Livro de Reis estabelece que* $\pi = 3$, *é uma tese segura, com pouco espaço para debate.*

Observar que $\operatorname{sen}(3)$ *é um real estritamente positivo, o qual pode ser calculado com precisão arbitrária a partir do truncamento da série de potências discutida na Seção 54. Além disso,* $\operatorname{sen}(\frac{22}{7})$ *é um real negativo, que pode ser calculado da mesma maneira. Neste caso,* π *está em algum lugar do corpo totalmente ordenado dos números reais, entre a estimativa bíblica* 3 *e uma das aproximações de Arquimedes,* $\frac{22}{7}$.

Apesar da proposta sagrada não ser tão precisa quanto a estimativa de Arquimedes, escolhemos ela como semente x_0, *para fins de uma avaliação superficial do desempenho do Método de Newton-Raphson.*

Logo,

$$x_0 = 3,000000000000000 \qquad x_1 = 3,142546543074278$$

$$x_2 = 3,141592653300477 \qquad x_3 = 3,141592653589793$$

$$x_4 = 3,141592653589793$$

Com apenas quatro iterações, conseguimos uma aproximação de π *com quinze dígitos após a vírgula. O critério de parada é o fato de que, com uma precisão de quinze dígitos, existe n natural tal que*

$$x_{n+1} - x_n = 0,000000000000000.$$

Esse valor n é 3. Logo, o processo recursivo é interrompido em x_4. *Esse exemplo ajuda a ilustrar a eficiência do Método de Newton-Raphson.*

Seção 109

Método de Euler

Sumário
Índice
Rede

Métodos numéricos implementáveis em máquinas não servem apenas ao propósito de determinar zeros de certas funções reais diferenciáveis. É possível também estimar funções que sejam soluções aproximadas de equações diferenciais.

Seja $x : \mathbb{R} \to \mathbb{R}$ uma função $x(t)$ real tal que

$$\frac{d^2x}{dt^2} = \alpha,$$

sendo α um número real qualquer. Logo,

$$\frac{dx}{dt} = \alpha t + \beta,$$

sendo $\beta \in \mathbb{R}$. Portanto,

$$x(t) = \alpha\frac{t^2}{2} + \beta t + \gamma,$$

onde $\gamma \in \mathbb{R}$.

Se usarmos t para mapear tempo em segundos e $x(t)$ para mapear posição em metros de uma partícula em $\mathbb{R}$, dependente de tempo, temos que

$$\frac{d^2x}{dt^2} = \alpha$$

é a *aceleração constante* da partícula em metros por segundo por segundo, e

$$\frac{dx}{dt}$$

é a sua *velocidade instantânea* em metros por segundo. Observar que $x(0) = \gamma$. Logo, γ pode ser usada para mapear posição x_0 metros no instante 0 segundos. Além disso,

$$\left.\frac{dx}{dt}\right|_{t=0} = \beta,$$

o que significa que podemos usar β para mapear velocidade instantânea v_0 no instante $t = 0$. Logo, $\beta = v_0$ e $\gamma = x_0$ são condições de contorno da equação diferencial

$$\frac{d^2x}{dt^2} = \alpha.$$

Outra maneira de examinarmos o problema acima é através da substituição de $\frac{dx}{dt}$ por $v(t)$, uma função de velocidade instantânea. Logo, $\frac{d^2x}{dt^2} = \alpha$ é equivalente a

$$\frac{dv}{dt} = \alpha.$$

A forma integral dessa última equação diferencial é

$$\int_{v_0}^{v_F} dv = \int_{t_0}^{t_F} \alpha dt.$$

Logo, $v\Big|_{v_0}^{v_F} = \alpha t\Big|_{t_0}^{t_F}$. Logo, $v_F - v_0 = \alpha(t_F - t_0)$, o que implica em

$$v_F = v_0 + \alpha(t_F - t_0).$$

Se substituirmos t_0 por 0 e t_F por t, temos a mesma igualdade

$$\frac{dx}{dt} = \alpha t + \beta,$$

onde $\beta = v_0$ e $v_F = \frac{dx}{dt}$.

A forma integral da última equação diferencial é

$$\int_{x_0}^{x_F} dx = \int_{t_0}^{t_F} (\alpha t + \beta) dt.$$

Logo, $x\Big|_{x_0}^{x_F} = (\alpha \frac{t^2}{2} + \beta t)\Big|_{t_0}^{t_F}$. Isso implica em

$$x_F - x_0 = \alpha \frac{{t_F}^2}{2} + \beta t_F - \frac{{t_0}^2}{2} - \beta t_0.$$

Logo,

$$x(t) = \alpha \frac{t^2}{2} + \beta t + x_0,$$

se substituirmos t_0 por 0 e t_F por t.

No contexto de mecânica clássica [17], o *estado* de uma partícula sujeita às leis de Newton pode ser representado por um par ordenado

$$(x(t), v(t))$$

(posição e velocidade em cada instante de tempo t), se a partícula tiver uma massa constante em relação a tempo. A ideia intuitiva é simples: o estado da partícula deve descrever onde a partícula está e em qual velocidade se encontra a cada instante de tempo.

Suponha que uma partícula com massa m constante (em relação ao tempo) mude seu estado $(x(t), v(t))$ de acordo com a Segunda Lei de Newton:

$$F(t) = m\frac{dv}{dt},$$

onde $F(t)$ é uma função que descreve força resultante sobre a partícula.

A forma integral desta equação diferencial é

$$\int_{v_0}^{v_F} mdv = \int_{t_0}^{t_F} F(t)dt.$$

Sendo m uma constante em relação a t, a integral do lado esquerdo pode ser facilmente calculada a partir do Teorema Fundamental do Cálculo. Mas a integral do lado direito da igualdade é algo bem diferente, pelo fato de depender da função-força $F(t)$.

Se $F(t)$ for uma função constante, a aplicação do Teorema Fundamental do Cálculo é viável, uma vez que funções constantes admitem primitivas dadas por funções elementares.

Existe vasta variedade de funções elementares $F(t)$ que admitem primitivas dadas por funções elementares, como polinomiais, trigonométricas, exponenciais e logarítmicas. No entanto, nem sempre isso ocorre. Por exemplo,

$$F(t) = e^{-t^2},$$

$$F(t) = \frac{1}{\ln t}$$

e

$$F(t) = \frac{e^{-t}}{t}$$

são funções elementares que não admitem primitivas elementares. Para detalhes, consultar o Algoritmo de Risch [39].

Para casos envolvendo funções elementares que não admitem primitivas elementares, uma opção é o emprego de métodos numéricos a partir de funções recursivas. O *Método de Euler* é o mais simples para esse propósito.

O Método de Euler foi concebido por Leonhard Euler no século 18 (ou seja, muito antes do advento de tecnologias digitais) para oferecer soluções aproximadas de equações diferenciais nas quais ocorrem uma derivada de primeira ordem mas nenhuma derivada de ordem superior, desde que uma condição de contorno seja dada.

No caso ilustrado acima, podemos reescrever a forma diferencial

$$F(t) = m\frac{dv}{dt}$$

como uma equação a *diferenças finitas* na qual uma derivada é tratada de fato como uma razão entre números reais, como mostrado abaixo. Por conta disso, o Método de Euler fornece apenas uma solução aproximada da equação diferencial. Neste caso, temos:

$$m\frac{v_{n+1} - v_n}{\tau} = F(t_n),$$

sendo que

$$t_{n+1} = t_n + \tau,$$

onde τ é o *passo de integração numérica*. Logo,

$$v_{n+1} = v_n + \frac{\tau F(t_n)}{m}$$

e

$$t_{n+1} = t_n + \tau.$$

A partir de uma condição de contorno $F(t_0)$ e um passo de integração τ adequadamente escolhido, em princípio é possível obter pontos (t_n, v_n) que correspondem a uma restrição de uma aproximação da primitiva de $F(t)$, uma vez que qualquer máquina somente é capaz de processar uma quantia finita de informações .

Há generalizações do Método de Euler, como os *Métodos de Runge-Kutta*. Recomendamos ao leitor que procure informações sobre o tema.

Seção 110

Predicados conjuntistas para teorias físicas

Sumário
Índice
Rede

Como ressaltado na Seção 107, o Sexto Problema de Hilbert coloca a seguinte questão:

> *Investigações sobre os fundamentos da geometria sugerem o problema: tratar do mesmo modo, por meio de axiomas, as ciências físicas nas quais a matemática desempenha papel importante: são prioritárias a teoria de probabilidades e a mecânica.*

Kolmogorov resolveu esse problema para a teoria de probabilidades (Parte 9). No entanto, mecânica se revelou assunto um pouco mais complicado.

A primeira metade do século 20 testemunhou o nascimento da física moderna, a qual foi definida pela mecânica quântica e pelas teorias da relatividade (restrita e geral) de Einstein. Posteriormente surgiram as teorias quânticas de campos. Logo, aquilo que Hilbert chamava de mecânica é algo que hoje pode ser chamado de mecânica clássica não relativística.

Apresentamos a seguir a pioneira formulação que John Charles Chenoweth McKinsey, Alvin Sugar e Patrick Colonel Suppes introduziram para a mecânica clássica não-relativística de partículas, em 1953. O sistema proposto reflete aspectos essenciais da mecânica newtoniana. Além disso, oferece uma axiomatização, via predicado conjuntista, suficientemente rica para uma discussão filosófica a respeito de outros tópicos, como as mecânicas de Hertz [46] e de Mach [47].

Nesta Seção, $\mathbb{R}^3$ denota o espaço vetorial real usual de triplas ordenadas de reais, munido das operações usuais de adição de triplas ordenadas de reais e de multiplicação de real por tripla ordenada de reais. Também assumimos a base canônica para $\mathbb{R}^3$ e que este mesmo espaço é munido do produto interno canônico. Intuitivamente falando, identificamos $\mathbb{R}^3$ com o espaço físico. Espaço físico, por sua vez, se refere a uma potencial totalidade de possíveis posições e direções relativas de objetos do mundo real.

A seguir, alguns conceitos preliminares.

Seja $f : T \to \mathbb{R}^3$ uma função, onde T é um intervalo de números reais. Dizemos, neste caso, que f é uma *função vetorial* em $\mathbb{R}^3$.

Se $t \in T$, temos que

$$f(t) = (x(t), y(t), z(t)),$$

onde x, y e z são funções reais com domínio T.

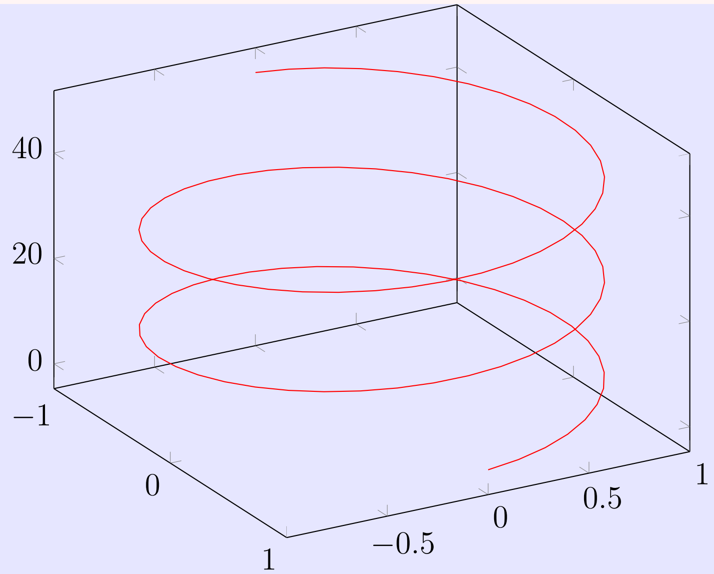
Neste contexto,

$$\frac{d^n}{dt^n} f(t) = \left(\frac{d^n}{dt^n} x(t), \frac{d^n}{dt^n} y(y), \frac{d^n}{dt^n} z(t) \right).$$

Exemplo 10.3. *A função $f : \mathbb{R} \to \mathbb{R}^3$ dada por*

$$f(t) = (\cos t, \operatorname{sen} t, 3t)$$

descreve uma helicoidal *em $\mathbb{R}^3$.*



Neste caso particular,

$$\frac{d}{dt}f(t) = (-\operatorname{sen} t, \cos t, 3) \quad e \quad \frac{d^2}{dt^2}f(t) = (-\cos t, -\operatorname{sen} t, 0).$$

Rigorosamente falando, a última igualdade na última definição não é a definição de derivada de função vetorial em $\mathbb{R}^3$, mas um teorema que pode ser demonstrado a partir da definição de derivada de funções vetoriais. Não apresentamos os detalhes porque, para os atuais propósitos, são desnecessários.

A seguir uma adaptação do trabalho de McKinsey, Sugar e Suppes.

Definição 10.1. $\mathcal{P} = \langle P, T, \mathbf{s}, m, \mathbf{f}, \mathbf{g} \rangle$ *é um* sistema não relativístico de partículas *sss:*

mcp1: *P é um conjunto finito não vazio;*

mcp2: *T é um intervalo de números reais;*

mcp3: *$\mathbf{s} : P \times T \to \mathbb{R}^3$ é uma função cujas imagens $\mathbf{s}(p,t)$ são denotadas por $\mathbf{s}_p(t)$;*

mcp4: *Para todo $t \in T$ existe*

$$\frac{d^2\mathbf{s}_p(t)}{dt^2};$$

mcp5: *$m : P \to \mathbb{R}$ é uma função tal que $m(p) > 0$ para todo $p \in P$;*

MCP6: $\mathbf{f} : P \times P \times T \to \mathbb{R}^3$ *é uma função tal que*

$$\mathbf{f}(p, q, t) = -\mathbf{f}(q, p, t);$$

MCP7: $\mathbf{g} : P \times T \to \mathbb{R}^3$ *é uma função, cujas imagens são denotadas por* $\mathbf{g}(p, t)$*;*

MCP8: *Se* $p \in P$ *e* $q \in P$*, então, para todo* t *pertencente a* T*,* $\mathbf{f}(p, q, t)$ *é uma combinação linear de* $\mathbf{s}_p(t) - \mathbf{s}_q(t)$*;*

MCP9: *Se* $p \in P$*,* $q \in P$ *e* $t \in T$*, então*

$$m(p)\frac{d^2\mathbf{s}_p(t)}{dt^2} = \sum_{q\in P} \mathbf{f}(p, q, t) + \mathbf{g}(p, t).$$

- Os elementos de P são chamados de *partículas*.
- Os elementos de T são *instantes de tempo* ou, simplesmente, *instantes*.
- Cada imagem $\mathbf{s}_p(t)$ da função $\mathbf{s}$ é a *posição da partícula* p *no instante* t.
- Cada imagem $m(p)$ da função m é a *massa* da partícula p.
- Cada imagem $\mathbf{f}(p, q, t)$ da função $\mathbf{f}$ é a *força da partícula* q *sobre a partícula* p *no instante* t.
- Cada imagem $\mathbf{g}(p, t)$ da função $\mathbf{g}$ é a *força perturbativa sobre a partícula* p *no instante* t.

Axioma MCP1 diz que todo sistema não relativístico de partículas tem pelo menos uma partícula. Além disso, o conjunto de partículas não pode ser infinito.

Postulado MCP2 apenas estabelece um conjunto T como parâmetro para definir funções de posição, força e força perturbativa. A interpretação pretendida é que T seja um intervalo de tempo.

Fórmula MCP3 garante que cada partícula p, em cada instante de tempo t, pode ser localizada por uma posição $\mathbf{s}_p(t)$ no espaço $\mathbb{R}^3$. A helicoidal no Exemplo 10.3 descreve uma possível *trajetória* de uma dada partícula em um dado intervalo de tempo.

Axioma MCP4 impõe que as posições de partículas são dadas por funções duas vezes diferenciáveis em relação a tempo. Naturalmente, a interpretação pretendida é que a derivada primeira descreva velocidade de cada partícula em cada instante de tempo, e a derivada segunda descreva aceleração.

MCP5 diz que toda partícula tem massa real estritamente positiva.

Uma vez que $\mathbb{R}^3$ é munido de produto interno canônico, isso garante que podemos definir a norma induzida pelo produto interno. Neste sentido, postulado MCP6 estabelece que a força da partícula q sobre a partícula p tem a mesma intensidade (dada pela norma) da força da partícula p sobre a partícula q, em cada instante t. Além disso, ambos os vetores têm a mesma direção, mas 'sentidos opostos' (um é simétrico aditivo do outro). A interpretação pretendida é que este postulado mapeie uma versão fraca da *Terceira Lei de Newton*.

Para efeitos práticos, é desejável a existência de forças perturbativas. Esta é a razão do axioma MCP7.

Postulado MCP8, em parceria com MCP6, deve mapear a *Terceira Lei de Newton*: a força que a partícula q exerce sobre p é um vetor com a direção da posição de p relativamente à posição de q. Esse axioma previne ambiguidades na determinação da direção de forças.

Finalmente, MCP9 deve mapear a *Segunda Lei de Newton* para partículas com massa constante relativamente ao parâmetro tempo. O lado direito da igualdade é o que se chama de *força resultante* sobre a partícula p.

Naturalmente, postulado MCP9 é uma equação diferencial. Resolver tal equação diferencial significa determinar o *estado* de cada partícula em cada instante de tempo, sendo que o estado de uma partícula p no instante t é o par ordenado

$$\left(\mathbf{s}_p(t), \frac{d}{dt}\mathbf{s}_p(t)\right).$$

Teorema 10.1. *A força que uma partícula exerce sobre ela mesma, em um sistema não relativístico de partículas, é sempre nula.*

Demonstração: De acordo com MCP6,

$$\mathbf{f}(p,p,t) = -\mathbf{f}(p,p,t).$$

Logo,

$$\mathbf{f}(p,p,t) + \mathbf{f}(p,p,t) = (0,0,0).$$

Logo,

$$\mathbf{f}(p,p,t) = (0,0,0).$$

Teorema 10.2. Primeira Lei de Newton.

Demonstração: De acordo com mcp9, se a força resultante sobre uma partícula p é nula, então

$$\frac{d^2\mathbf{s}_p(t)}{dt^2} = (0, 0, 0),$$

uma vez que $m(p) > 0$.

Logo,

$$\frac{d\mathbf{s}_p(t)}{dt} = (\alpha, \beta, \gamma),$$

onde α, β e γ são constantes reais.

Logo,

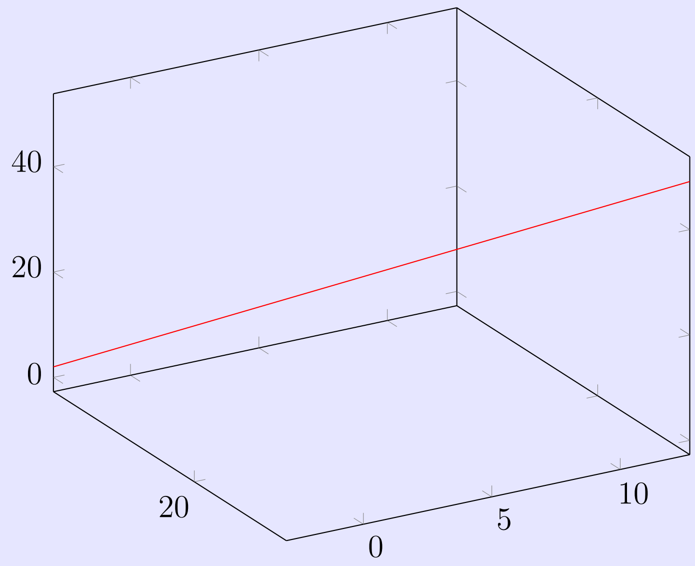
$$\mathbf{s}_p(t) = (\alpha t + a, \beta t + b, \gamma t + c).$$

Mas esta é exatamente uma trajetória retilínea, se α, β ou γ forem diferentes de 0, ou um estado de repouso, se α, β e γ forem todos nulos.

Exemplo 10.4. ✍ *A função $f : \mathbb{R} \to \mathbb{R}^3$ dada por*

$$f(t) = (2t + 1, t - 3, 3t + 2)$$

descreve uma trajetória retilínea em $\mathbb{R}^3$.



Neste caso particular,

$$\frac{d}{dt}f(t) = (2, 1, 3)$$

e

$$\frac{d^2}{dt^2}f(t) = (0, 0, 0).$$

Uma limitação óbvia do sistema axiomático acima é o fato das partículas terem massas invariantes em relação a tempo. Mas esta é uma limitação facilmente contornável se modificarmos adequadamente os axiomas MCP5 e MCP9.

No entanto, há uma limitação muito mais severa: o fato de que essa formulação assume explicitamente um único papel para forças: mudar o estado de uma partícula ao longo do tempo. É bem sabido que forças, no contexto de mecânica clássica, contam com outro papel relevante: deformar corpos.

Apesar disso, o sistema concebido por McKinsey, Sugar e Suppes foi pioneiro no emprego de técnicas modernas de axiomatização para teorias físicas. Desde então, muitas outras propostas surgiram na literatura. Algumas delas se refeream a sistemas axiomáticos para a mecânica dos meios contínuos, para teorias de campos (clássicas ou quânticas) e para a termodinâmica, entre muitos outros exemplos de teorias físicas.

Sobre o impacto dessas ideias, ver [12]. O que mostramos aqui é uma porção insignificante sobre o que o método axiomático pode fazer pela física teórica.

Seção 111

Modelos de ZF

Sumário
Índice
Rede

Como enfatizado na Seção 1, a linguagem $\mathfrak{S}$ de ZF é desprovida de semântica. Apesar das vantagens já discutidas sobre essa característica, há dificuldades inerentes a ZF que exigem algum tipo de consideração de caráter semântico.

Como já foi dito, matemáticos são caçadores de teoremas não triviais. Neste contexto, considere o seguinte fato:

$$\vdash_{\text{ZF}-\{\text{Axioma do Par}\}} \text{Axioma do Par}.$$

O que está escrito acima é o seguinte: em ZF − {Axioma do Par} (a qual é uma teoria formal com todos os postulados de ZF, exceto o Axioma do Par) o Axioma do Par é teorema, como se verifica a seguir.

Teorema 10.3. $\vdash_{ZF-\{Axioma\ do\ Par\}}$ *Axioma do Par.*

Demonstração: Axioma do Vazio e Teorema 3.2 garantem a existência e unicidade do conjunto vazio $\varnothing$. Axioma da Potência garante a existência do conjunto

$$t = \wp(\wp(\varnothing)) = \{\varnothing, \{\varnothing\}\}.$$

Sejam r e s conjuntos quaisquer e $\mathcal{F}(x, y)$ a fórmula

$$(x = \varnothing \Rightarrow y = r) \wedge (x \neq \varnothing \Rightarrow y = s).$$

Logo, aplicando o Esquema de Substituição sobre t, usando a fórmula $\mathcal{F}(x, y)$, obtemos o par $\{r, s\}$.

Resumidamente, o que foi provado acima é que o Axioma do Par é desnecessário, desde que tenhamos os postulados Vazio, Potência e Substituição. A teoria formal axiomática ZF − {Axioma do Par} é equivalente a ZF, no sentido de que todos os teoremas de uma são teoremas da outra. Por conta disso, alguns autores simplesmente omitem esse postulado em certas formulações de ZF e ZFC.

Se $\mathcal{F}$ é um postulado de ZFC, dizemos que $\mathcal{F}$ é *independente* dos demais axiomas de ZFC sss

$$\nvdash_{ZFC-\{\mathcal{F}\}} \mathcal{F}.$$

Caso contrário, dizemos que $\mathcal{F}$ é *dependente*.

Logo, o último teorema é equivalente à seguinte proposição.

Proposição 10.1. *O Axioma do Par é dependente dos demais postulados de ZF.*

Consequentemente, é natural questionar se fenômeno análogo ocorre com outros axiomas de ZF e ZFC. Afinal, matemáticos querem genuinamente conhecer essa teoria formal.

No entanto, não é tão fácil assim responder se outros postulados podem ser simplesmente omitidos, mantendo todos os teoremas. Provar a dependência de um postulado é, em princípio, fácil. Basta exibir uma demonstração. Provar a independência, porém, é algo que exige ferramentas metamatemáticas. Afinal, se um matemático não consegue exibir uma demonstração, isso não implica que tal demonstração não exista. O Axioma da Escolha é um exemplo histórico bem conhecido.

A primeira pessoa a trazer alguma luz sobre o tema foi Kurt Gödel, apesar de Abraham Fraenkel e Andrzej Mostowski terem apresentado resultados relacionados ao Axioma da Escolha para uma variação de ZF conhecida como ZFU [22].

Gödel criou um *modelo* de ZF, hoje conhecido como L, ou, *universo construtível de Gödel*. Para isso ele precisou qualificar qual é um possível *universo de discurso* L de ZF.

Em outras palavras, uma vez que os axiomas de ZF empregam quantificadores lógicos, o que *podem* significar fórmulas como

'para todo x, isso ou aquilo acontece'?

O que é 'para todo'? Obviamente, esse 'para todo' não pode incluir objetos como elefantes, calças desbotadas ou molas de grampolas. É neste momento que um modelo de ZF cumpre o papel de estabelecer um possível universo de discurso para uma teoria como ZF.

O que Gödel propôs foi um universo de discurso L minimamente necessário para satisfazer todos os axiomas de ZF. A ideia, intuitivamente, é a seguinte.

Um conjunto y é *definível* a partir de um conjunto x se existe uma fórmula Φ tal que t pertence a y sss t pertence a x e t satisfaz a fórmula Φ. Uma vez estabelecido o que é um conjunto definível a partir de outro, é possível introduzir um universo de conjuntos hierarquicamente definíveis a partir do conjunto vazio como se segue.

i: Admite-se a existência de um conjunto vazio chamado de L_0;

ii: Se existir fórmula Φ que permita definir um novo conjunto a partir de L_0, este novo conjunto é elemento de um conjunto L_1;

iii: De fato, é possível definir um conjunto x a partir de L_0 da seguinte forma: t pertence a x sss t pertence a L_0 e Φ (seja qual for a fórmula Φ, neste caso muito particular); com efeito, nenhum t pertence a vazio; logo, neste caso muito especial, qualquer fórmula funciona; logo, x é novamente o vazio; a partir disso é definido o conjunto L_1 cujos elementos são todos os conjuntos definíveis a partir de L_0, ou seja, o próprio L_0; em outras palavras, L_1 conta com um único elemento, a saber, L_0; escrevemos $L_1 = \{L_0\}$;

iv: Se existir fórmula Φ que permita definir um novo conjunto x a partir de L_1, este novo x é elemento de um conjunto L_2; logo, os elementos de L_2 são apenas L_0 e L_1, definíveis pela escolha apropriada de fórmulas; ou seja, $L_2 = \{L_0, \{L_0\}\}$;

v: Se existir fórmula Φ que permita definir um novo conjunto x a partir de L_2, este novo x é elemento de um conjunto L_3; logo, os elementos de L_3 são apenas quatro, definíveis a partir de fórmulas apropriadas para este fim: L_0, L_1, L_2 e um conjunto que tem como único elemento L_1; escrevemos $L_3 = \{L_0, \{L_0\}, \{L_0, \{L_0\}\}, \{\{L_0\}\}\}$;

vi: Repetimos o processo acima para L_4, L_5, L_6 e assim por diante, até cobrir todos os ordinais finitos;

vii: Chamamos de L_ω o conjunto que satisfaz a seguinte condição: um conjunto x pertence a L_ω sss x pertence a algum L_n, onde n é um ordinal finito;

viii: Chamamos de $L_{\omega+1}$ o conjunto dos conjuntos definíveis a partir de L_ω; $L_{\omega+2}$ o conjunto dos conjuntos definíveis a partir de $L_{\omega+1}$, e assim por diante, até novamente cobrir todos os ordinais finitos;

ix: Chamamos de $L_{2\omega}$ o conjunto que satisfaz a seguinte condição: um conjunto x pertence a $L_{2\omega}$ sss x pertence a algum $L_{\omega+n}$, onde n é um ordinal finito;

x: Chamamos de $L_{2\omega+1}$ o conjunto dos conjuntos definíveis a partir de $L_{2\omega}$; $L_{2\omega+2}$ o conjunto dos conjuntos definíveis a partir de $L_{2\omega+1}$, e assim por diante, até novamente cobrir todos os ordinais finitos;

xi: Chamamos de $L_{3\omega}$ o conjunto que satisfaz a seguinte condição: um conjunto x pertence a $L_{3\omega}$ sss x pertence a algum $L_{2\omega+n}$, onde n é um ordinal finito;

xii: Repetimos o procedimento acima para $L_{4\omega}$, $L_{5\omega}$, e assim por diante.

xiii: Finalmente, dizemos que um conjunto x pertence a L sss x pertence a algum L_p, onde p é um ordinal finito n ou p é $m\omega+n$.

No contexto acima, um conjunto é qualquer x que pertence a L.

Pois bem. O que Gödel provou é que todos os axiomas de ZFC são satisfeitos em L.

Por exemplo, o Axioma do Vazio é satisfeito graças à existência de L_0, o qual é um elemento de L. O Axioma do Par é satisfeito graças ao seguinte fato: dados x e y pertencentes a L, existem L_p e L_q tais que x pertence a L_p e y pertence a L_q, para algum p e algum q da construção acima; logo, existe algum L_r tal que r é maior do que ambos p e q e tal que o conjunto $\{x, y\}$ pertence a L_r; logo, o par $\{x, y\}$ pertence a L, ou seja, é um conjunto.

Uma vez que todos os axiomas de ZFC são satisfeitos em L, Gödel provou com isso que os axiomas de ZF não permitem inferir a negação do Axioma da Escolha como teorema. Ou seja, apesar de até hoje não existir prova de que ZF é consistente, pelo menos Gödel provou que, se ZF for consistente, então ZFC também é.

Observar que o universo construtível L de Gödel oferece uma possível interpretação para o conceito de conjunto. Neste sentido, uma fórmula qualquer de ZFC é *verdadeira* em L sss essa fórmula for satisfeita em L. Caso contrário, a fórmula é *falsa* em L. Logo, todos os axiomas de ZFC são verdadeiros em L, enquanto a negação do Axioma da Escolha é falsa em L.

No entanto, todo esse esforço de Gödel não foi suficiente para responder se o próprio Axioma da Escolha é teorema ou não de ZF. Quem respondeu a essa questão foi Paul Cohen, na segunda metade do século 20.

Cohen criou uma técnica hoje conhecida como *forcing*, a qual permite criar outros modelos de ZF a partir, por exemplo, do universo L de Gödel. Graças a forcing é possível criar um modelo M de ZF no qual L está contido mas tal que $L \neq M$. Em um dos modelos M Cohen provou que todos os axiomas de ZF são verdadeiros, mas o Axioma da Escolha é falso. Logo, ZF é consistente tanto com o Axioma da Escolha quanto com a negação do Axioma da Escolha. Portanto, o Axioma da Escolha não pode ser demonstrado a partir dos demais postulados de ZF. Com efeito, se pudesse, qualquer modelo de ZF seria também modelo de ZFC.

Qualquer modelo de ZF ou ZFC oferece possíveis interpretações para conjuntos e para a pertinência $\in$. Neste sentido, modelos de ZF e de ZFC qualificam o que é 'para todo'. Com efeito, ao enunciar 'para todo conjunto x' estamos falando apenas dos x pertencentes a um universo de discurso particular que opera como modelo de ZF ou ZFC. Teorias como ZF e ZFC podem ter muitos universos de discurso. Esta é a ambiguidade inerente aos quantificadores lógicos $\forall$ e $\exists$. Não existe um único possível universo de discurso para ZF ou ZFC, ou seja, não existe uma única possível interpretação para a totalidade de conjuntos de ZF ou ZFC.

Além disso, os conceitos de *verdade* e *falsidade* são relativos a modelos. Ou seja, dada uma fórmula Φ de ZF, como saber se essa fórmula é verdadeira ou falsa? Para responder a essa questão é necessário qualificar o modelo que está sendo usado. No modelo L de Gödel o Axioma da Escolha é verdadeiro. No modelo M de Cohen, a mesma fórmula é falsa.

Logo, os conceitos semânticos de verdade e falsidade em uma linguagem formal como a de ZF têm uma conotação muito diferente dos conceitos de verdade e falsidade em uma linguagem natural como o português.

Uma questão natural é a seguinte: se existe modelo para ZF e ZFC, por que esse modelo não permite provar a consistência dessas teorias? A resposta é simples. Qualquer modelo, por exemplo, de ZFC qualifica apenas um único possível universo de discurso para a teoria. Mas ZF e ZFC são muito mais do que qualquer modelo possa revelar. Uma vez que a consistência de ZF ou ZFC não pode ser teorema da própria teoria [28], modelos não são capazes de responder sobre consistência.

Seção 112

Princípio de Partição

Sumário
Índice
Rede

O *Princípio de Partição* (PP) é a seguinte fórmula:

Para toda função sobrejetora $f : x \to y$ existe uma função injetora $g : y \to x$.

Exemplo 10.5. *Seja*

$$f : \{1, 2, 3, 4, 5, 6\} \to \{1, 2, 3\}$$

dada por $f(1) = f(2) = 1$, $f(3) = f(4) = 2$, $f(5) = f(6) = 3$.

Cada elemento do contradomínio de f é imagem de algum elemento de seu domínio. Isso garante que f é sobrejetora. Neste caso particular certamente existe

$$g : \{1, 2, 3\} \to \{1, 2, 3, 4, 5, 6\}$$

injetora. Por exemplo, g pode ser dada por

$$g(1) = 1, \ \ g(2) = 2 \ \ e \ \ g(3) = 3.$$

No EXEMPLO acima PP é teorema. Naturalmente, se f for finita, PP sempre é teorema. O problema, no entanto, é provar tal fórmula para toda e qualquer função sobrejetora.

TEOREMA 10.4. $\vdash_{ZFC}$ *Princípio de Partição*

DEMONSTRAÇÃO: Seja $f : x \to y$ sobrejetora. Para cada elemento s de y existe subconjunto u_s de x tal que

$$\forall r(r \in u_s \Rightarrow f(r) = s).$$

Uma vez que f é função, então

$$s \neq s' \Rightarrow u_s \cap u_{s'} = \varnothing.$$

O conjunto de todos os u_s, para todos os s pertencentes a y, define uma partição p_x de x (Definição 3.16). Logo, podemos aplicar o Axioma da Escolha sobre p_x. O Axioma da Escolha 'escolhe' um e apenas um elemento c_s pertencente a cada u_s da partição p_x. Portanto, podemos definir uma função injetora $g : y \to x$ dada por

$$g(s) = c_s.$$

EXEMPLO 10.6. *Seja $f : \{1, 2, 3, 4, 5, 6\} \to \{1, 2, 3\}$ dada por*

$$f(1) = f(2) = 1, \ \ f(3) = f(4) = 2, \ \ f(5) = f(6) = 3,$$

exatamente como no EXEMPLO 10.5.

Neste caso, a partição de $\{1, 2, 3, 4, 5, 6\}$ induzida por f é

$$p = \{\{1, 2\}, \{3, 4\}, \{5, 6\}\}.$$

Se aplicarmos o Axioma da Escolha sobre p podemos obter, por exemplo, o conjunto escolha

$$\{1, 3, 5\},$$

entre outras possibilidades.

Se for o caso, a injetora $g : y \to x$ é dada por

$$g(1) = 1, \ \ g(2) = 3 \ \ e \ \ g(3) = 5.$$

Obviamente a função injetora do EXEMPLO *anterior não pode ser obtida por emprego do Axioma da Escolha, pelo menos nos moldes da prova da último teorema. Mas, seja como for, a estratégia adotada na demonstração do Teorema 10.4 funciona para toda e qualquer sobrejeção $f : x \to y$.*

Porém, a questão agora é a seguinte: sabendo que o Axioma da Escolha implica em PP, podemos garantir a recíproca de tal teorema? Formalmente,

$$\vdash_{ZF+\{PP\}} \text{Axioma da Escolha?}$$

Quando Ernst Zermelo propôs o Axioma da Escolha há mais de um século, ele o fez para garantir PP no contexto da teoria de conjuntos proposta pelo próprio Zermelo. No entanto, nunca foi capaz de responder se PP implica em AE (Axioma da Escolha). Até os dias de hoje esta é uma questão em aberto.

Todos os modelos de ZF, até hoje concebidos, nos quais o Princípio de Partição é verdadeiro, o Axioma da Escolha também é verdadeiro. Este fato parece sugerir que PP implica em AE. No entanto, não existe qualquer classificação de todos os possíveis modelos de ZF. Consequentemente, a atual teoria de modelos tem se mostrado insuficiente para resolver esse problema.

Em [44] há uma prova de que o Axioma da Escolha é independente do Princípio de Partição em uma variação de ZF conhecida como ZFU (Zermelo-Fraenkel com átomos).

SEÇÃO 113

O que omitimos

SUMÁRIO
ÍNDICE
REDE

Muitos assuntos relevantes, relacionados aos temas abordados, foram omitidos neste livro. Fazemos abaixo uma breve lista de apenas alguns deles. Paralelamente, recomendamos leituras complementares.

TEORIAS FORMAIS: A linguagem de ZF (a qual é a mesma de ZFC) é um caso particular de *linguagem de primeira ordem*.

Linguagens de primeira ordem, por sua vez, são casos particulares de linguagens formais. Ambos os assuntos podem ser estudados em [38], livro extraordinariamente didático. As últimas edições são as melhores. Na mesma obra há uma detalhada discussão sobre a teoria de conjuntos NBG (von Neumann, Bernays, Gödel). Esta é uma teoria de conjuntos que garante a existência de classes próprias, termos que admitem elementos mas que não são conjuntos. Em [8] há uma boa discussão sobre a teoria de Zermelo-Fraenkel em uma linguagem de segunda ordem, também conhecida como ZF_2. Em [56] há um exemplo de teoria de conjuntos cuja linguagem formal prescinde de variáveis, quantificadores lógicos e até mesmo conectivos lógicos.

Tópicos de teoria de conjuntos: Se a pessoa não está interessada em aspectos formais de teoria de conjuntos, uma opção é o estudo de *teoria ingênua* (também conhecida como *teoria intuitiva* de conjuntos). Neste sentido recomendamos [50]. Para um primeiro estudo detalhado sobre teoria de Zermelo-Fraenkel ver [28]. No entanto, não é recomendável a leitura deste último antes de um seguro conhecimento sobre os conteúdos iniciais de [38]. Os modelos para ZF em [28] podem ser um tanto difíceis de compreender para o iniciante. É recomendável uma cuidadosa leitura de [1] para entender os propósitos de teoria de modelos. Sobre o Axioma da Escolha, ver [27].

Cardinalidades: Ordinais finitos podem ser estendidos para outra classe de objetos, a saber, os ordinais. Todo ordinal finito é um ordinal, mas a recíproca não é teorema. Além disso, a pertinência define uma boa ordem em qualquer ordinal. Logo, qualquer ordinal admite um menor elemento relativamente à pertinência. Isso permite definir cardinalidade de um conjunto (intuitivamente falando, a quantia de elementos do conjunto). A cardinalidade de um conjunto x é o menor ordinal (relativamente à pertinência) equipotente a x. O estudo de cardinalidades encontra enorme impacto sobre assuntos como teoria da medida, entre outros [28] [30].

Análise matemática: Assim como a Seção 39, sobre números reais, apresenta um modelo de corpo (no sentido das discussões nas Seções 71 e 96), também é possível provar que a mesma interpretação para números reais é um modelo de *corpo ordenado completo* [34]. De maneira análoga, as discussões sobre complexos na Seção 40 apenas exibem um modelo para *corpo topológico algebricamente fechado* [58]. Corpos ordenados completos e corpos topológicos algebricamente fechados são casos particulares de corpo.

Geometria sintética: Avaliar a independência dos postulados da geometria euclidiana é essencial para a qualificação de outras geometrias, como a absoluta, as não euclidianas e as não Paschianas. Detalhes na obra de Castrucci [10].

Álgebra matricial: Operações definidas sobre espaços de matrizes (incluindo *posto*, *determinante*, *escalonamento*, entre outras) são essenciais para o estudo de representação matricial de operadores lineares definidos sobre espaços vetoriais de dimensão finita. Detalhes em [35].

Cálculo padrão: Muitas aplicações de derivadas foram omitidas. Técnicas de integração (como *substituição de variáveis*, *substituições trigonométricas*, *integração por partes*, entre outras) são essenciais para a aplicabilidade de cálculo diferencial e integral padrão. Para uma abordagem intuitiva, [55]. Para um tratamento mais próximo da análise matemática, [19].

Mecânica clássica: Entre as aplicações mais usuais de cálculo padrão estão as teorias físicas, como mecânica clássica. Texto padrão sobre mecânica clássica: [17]. Tratamento para a mecânica clássica como uma teoria de campos: [2]. Para uma visão histórica e filosófica sobre o conceito de massa em diferentes teorias da física: [26].

Leituras complementares: Em [54] há uma extensa análise do emprego de linguagens de teoria de conjuntos no estudo de probabilidades, mecânica clássica, linguística e outras áreas. Em [57] há uma fascinante discussão sobre a abordagem de Bolzano para o conceito de infinito, ressaltando até mesmo o papel da religião sobre a concepção da teoria de conjuntos. O artigo de divulgação científica [52] oferece uma extraordinária visão sobre parte do impacto de teoria de conjuntos em análise matemática. Em [31] há uma proposta para privilegiar o emprego de funções no estudo de matemática.

PARTE 11

Por que tantos nomes em matemática?

Nesta última Parte discutimos sobre as motivações para o emprego de nomes em matemática, bem como seu impacto.

Seção 114
Nomes como arbitrariedades

Sumário
Índice
Rede

A matemática carrega algo de hermético em sua prática. Segundo Joannes Philoponus, filósofo neoplatônico cristão do século 6, na porta de entrada da Academia de Platão lia-se a frase

Ninguém deve entrar se for ignorante em geometria.

Não se sabe se a lenda procede. Mas essa famosa frase não deixa de encerrar em si uma visão comumente nutrida há milênios. Como já mencionado, Poincaré afirmou que o matemático nasce, não se cria. Independentemente de quaisquer considerações sobre a visão do grande matemático francês, um exemplo que ajuda a ilustrar a separação entre matemáticos e não matemáticos é exatamente ZF, como se discute a seguir.

ZF é conhecida como uma teoria de conjuntos. No entanto, seus conceitos primitivos são dois predicados binários usualmente denotados pelos símbolos $=$ e $\in$. Em ZF não se define o que é um conjunto. Sequer existe a necessidade de qualificar o que são conjuntos. Logo, ZF não é uma teoria sobre objetos chamados de conjuntos.

Se um leigo espera aprender algo sobre conjuntos, ao estudar ZF, apenas aprenderá sobre pertinência e igualdade. Isso porque ZF é uma teoria sobre as relações entre pertinência e igualdade. Neste sentido, uma pessoa com pouca familiaridade com matemática pode se sentir desorientada. Afinal, por que matemáticos se referem a ZF como uma teoria de conjuntos? Por que não chamar ZF de 'Teoria da Pertinência Extensional'? O primeiro postulado próprio de ZF é o Axioma da Extensionalidade, o qual estabelece de imediato o propósito de identificar um conjunto x a partir de todos os y que pertencem a x. Graças ao Axioma da Extensionalidade é possível garantir a unicidade de qualquer união arbitrária, de qualquer potência, de qualquer par. Todos os demais axiomas próprios 'confiam' na rigidez do Axioma da Extensionalidade. Qualquer objeto x que possamos conceber precisa da *Sagrada Lei da Extensionalidade* para garantir que x é um conjunto de ZF.

Para que o leitor possa compreender melhor essa questão, promovemos uma pequena mudança de nomenclatura, apenas para fins de breve ilustração.

Em primeiro lugar, chamemos os termos de ZF de 'homens'.

Em seguida, chamemos a pertinência de *contemplação*. No lugar de dizermos que 'o conjunto x pertence ao conjunto y' diremos que 'o homem x contempla o homem y'. Além disso, chamemos a igualdade de *fidelidade*. No lugar de dizermos que 'o conjunto x é igual ao conjunto y', dizemos que 'o homem x é fiel ao homem y'.

Como próximo passo, impomos na forma de postulado a *Sagrada Lei da Fidelidade* (nome que substitui a expressão 'Substitutividade da Igualdade'): todo homem x é fiel apenas a si mesmo. Logo, ao afirmarmos que x é fiel a y, estamos apenas dizendo que x pode ser chamado de y, assim como y pode ser chamado de x. Mas, no final das contas, x e y são apenas nomes de um único homem, uma vez que todo homem é fiel apenas a si mesmo.

Neste contexto, a *Sagrada Lei da Extensionalidade* (nome que substitui a expressão 'Axioma da Extensionalidade') afirma o seguinte.

Se todo homem t que contempla o homem x é também um homem que contempla o homem y e, além disso, todo homem t que contempla o homem y é também um homem que contempla o homem x, então o homem x é fiel ao homem y.

Se um leitor qualquer contestar a Sagrada Lei da Fidelidade, sugerindo que ele mesmo consegue ser fiel a outras pessoas, tal leitor não estará entendendo o ponto principal: o nome 'fidelidade' é apenas um nome, assim como o nome 'igualdade' é apenas um nome, nada além disso.

Se chamamos $\in$ de pertinência ou contemplação, isso é irrelevante do ponto de vista matemático. Relevante é apenas aquilo que os postulados demandam sobre os predicados $\in$ e $=$. Neste sentido, os nomes 'pertinência' e 'igualdade' são meras arbitrariedades, não identificáveis com situações ordinárias do cotidiano de não-matemáticos.

Porém, a prática mostra que nomes dados a predicados, termos ou fórmulas devem, pelo menos em princípio, refletir uma interpretação pretendida a tais conceitos. Ao chamarmos os termos de ZF de conjuntos, isso deve refletir a intuição de que um termo qualquer deve, de alguma forma, refletir a interpretação pretendida de uma coleção de objetos. Por conta disso o nome dado ao predicado $\in$ deve estar de acordo com tal interpretação pretendida. Um nome, em princípio, adequado é 'pertinência'.

Logo, os nomes adotados para conceitos acabam impactando sobre a própria prática matemática, como se percebe na próxima Seção.

Seção 115

Nomes como arbitrariedades impactantes

Sumário
Índice
Rede

Nomes desempenham dois papeis relevantes na matemática.

i: Tornam a prática matemática viável. Por exemplo, digamos que não fosse usado o nome 'número natural' ou qualquer outro nome para nos referirmos aos ordinais finitos. Além disso, digamos que não exista uma única definição explícita abreviativa que permita conferir nomes a termos, predicados ou fórmulas. Em outras palavras, digamos que não sejam definidos os seguintes conceitos:

- conjunto vazio,
- conjunto indutivo,
- sucessor de um conjunto,
- união finitária e, claro,
- conjunto ω dos números naturais.

Como poderíamos nos referir a um termo n qualquer como número natural? Neste caso, seria necessário dizer

$$\forall w((\forall m((\forall p(\neg(p \in m))) \Rightarrow m \in w) \land \forall t(t \in w \Rightarrow$$

$$\forall r((\forall s(s \in r \Leftrightarrow (s \in t \lor s = t))) \Rightarrow r \in w))) \Rightarrow n \in w).$$

O que está escrito acima é que o termo n (o único de ocorrência livre na fórmula) é elemento comum a todo e qualquer conjunto indutivo w.

Levando em conta que números inteiros são definidos a partir de classes de equivalência de pares ordenados de naturais (Seção 30), racionais são definidos a partir de classes de equivalência de pares ordenados de inteiros (Seção 31), e reais são definidos a partir de classes de equivalência de sequências de Cauchy de racionais (Seção 39), o leitor já pode imaginar a grande dificuldade para escrever algo como 'seja 1 o neutro multiplicativo dos reais', caso não adotássemos qualquer nome para o neutro multiplicativo dos reais ou para os demais conceitos usados para defini-lo.

II: CONFEREM PODER DE CONTROLE COGNITIVO SOBRE CONCEITOS. Exemplo bem conhecido foi a extraordinária habilidade de Alexander Grothendieck para atribuir nomes provocativos a conceitos, antes mesmo de uma plena compreensão sobre os mesmos. *Esquemas*, os quais generalizam *variedades algébricas*, são um caso bastante famoso. Outro exemplo histórico marcante foi a atitude de Cantor, ao afirmar que em matemática existem diversos tipos de *infinitos*. Um dos nomes escolhidos por Cantor foi 'infinito enumerável'. Por consequência, isso abriu espaço para o 'infinito não enumerável'. Tal controle cognitivo exerce

> um poder de criação: coisas passam a existir quando recebem nomes, como bem insistiu Henri Poincaré. No excelente artigo [18], Loren Graham compara esse ato de criação com o Gênesis do Antigo Testamento: 'Faça-se a luz! E a luz foi feita'. O nome 'luz' surgiu antes da própria luz.

No entanto, nas palavras de Nikolai Luzin [18], nomes fazem com que se perca 'as partes nebulosas e escuras que nossa intuição sussurra para nós'. Neste sentido, nomes podem restringir nossos modos de percepção sobre conceitos matemáticos.

Por exemplo, ao nos referirmos aos termos de ZF como conjuntos, podemos deixar de perceber que conjuntos podem ser interpretados como objetos que nada têm a ver com coleções de outros objetos, como ocorre nos modelos parciais de Abian e LaMacchia [1].

Principalmente para aqueles que iniciam seus estudos em matemática, nomes usualmente encontrados na literatura especializada podem facilmente confundir. Citamos a seguir alguns exemplos pontuais.

Seção 116

Inércia histórica

Sumário
Índice
Rede

Leopold Kronecker chegou a dizer que

> Deus criou os inteiros, o resto é obra dos homens.

Independentemente do significado desta famosa frase, números naturais, inteiros, racionais, irracionais, transcendentes, algébricos, reais e complexos são apenas casos especiais de termos, se formulados na linguagem de ZF. Termos, por sua vez, são conceitos abstratos desprovidos de significado. O número natural 2015, portanto, não existe no mundo real. Ainda que uma pessoa creia ser capaz de contar 2015 canivetes de bolso, apenas os canivetes de bolso são objetos reais. O número natural 2015 não é qualquer objeto do mundo real que esteja intrinsecamente associado a 2015 canivetes de bolso. Uma pessoa pode errar a contagem de canivetes de bolso sem que os canivetes em si informem de imediato esse erro de contagem.

A palavra 'natural', em contrapartida, se refere à natureza, podendo até mesmo ser sinônimo de banal, comum. Na qualidade de desesperados advogados do diabo, assumamos temporariamente que números naturais foram concebidos com o propósito de contar objetos do mundo real, algo que parece perfeitamente natural, banal, comum. Quantos números naturais podemos realmente usar para este propósito?

Estima-se haver algo entre 10^{78} e 10^{82} átomos no universo. Em 1920, um garoto de nove anos de idade nomeou o número natural 10^{100} como '*googol*', a pedido de seu tio, o matemático Edward Kasner. Logo, um googol é maior do que o número de átomos em todo o universo observável.

É claro que podemos imaginar situações potencialmente reais envolvendo números naturais bem maiores. Por exemplo, segundo Claude Shannon, o número total de possíveis variações no jogo de xadrez é algo entre 10^{111} e 10^{123}, uma quantia muito maior do que um googol. Porém, um *googolplex* é obviamente muito maior. Com efeito, um googolplex é $10^{10^{100}}$, ou seja, o número 1 seguido de um googol de zeros.

Não importa quantos objetos desejemos contar − entre canivetes de bolso, átomos no universo observável ou variações no jogo de xadrez − sempre há números naturais muito maiores. Isso porque o conjunto dos números naturais é infinito, algo que não ocorre na natureza. Portanto, números naturais não são naturais, nas acepções usuais do termo.

Um aspecto mais alarmante sobre a não banalidade do conjunto ω dos números naturais se refere ao conjunto de todos os subconjuntos de ω, i.e., sua potência. No universo construtível L de Gödel os subconjuntos de ω são todos definíveis. Porém, em qualquer extensão não trivial de L, via forcing, há subconjuntos de ω que não são definíveis. Se não há banais coleções infinitas na natureza, o que poderia haver de banal em uma coleção infinita não definível?

Para finalizar, o que há de banal ao afirmar que n é número natural se, e somente se,

$$\forall w((\forall m((\forall p(\neg(p \in m))) \Rightarrow m \in w) \wedge \forall t(t \in w \Rightarrow$$

$$\forall r((\forall s(s \in r \Leftrightarrow (s \in t \vee s = t))) \Rightarrow r \in w))) \Rightarrow n \in w)?$$

Um neófito que esteja iniciando seus estudos de matemática pode desfrutar de equivocada tranquilidade ao saber que está estudando números naturais. Afinal, é natural contar. Mas contar frutas em uma cornucópia pouco tem a ver com o conceito de número natural, a não ser pelas origens históricas do termo. Há muito mais entre números naturais, além de um, dois e muitos.

Comentários análogos podem ser feitos sobre os números inteiros, racionais, irracionais, reais, transcendentes e complexos. Números reais, por exemplo, não são objetos reais. Afinal, são conjuntos. O que seria transcendido por números reais transcendentes? O que há de complexo em números complexos, uma vez que todo complexo não passa de um par ordenado de reais? O que há de tão especial na unidade imaginária, uma vez que ela não é o único conjunto definível em ZF? Apenas a unidade imaginária apela à imaginação? Números irracionais são números incapazes de raciocinar? Os racionais raciocinam?

As origens históricas desses nomes não podem ser confundidas com as concepções atuais sobre tais conceitos. As funções circulares seno e co-seno ilustram este ponto muito bem. Historicamente, o termo 'seno' deriva do latim *sinus*, que pode ser traduzido como 'baía', 'seio' ou 'dobra', dependendo do contexto. Seja como for, as origens históricas do termo seno derivam de um apelo visual a formato. Porém, nos dias de hoje a função seno é definida como solução de uma equação diferencial, dadas condições de contorno. Essa definição pode ser estendida para incluir o seno de números complexos e até mesmo matrizes. Logo, a atual visão matemática da função seno pouco tem a ver com qualquer intuição de apelo visual.

Seção 117

Nomes que confundem

Sumário
Índice
Rede

Alguns nomes empregados na matemática denunciam questionável escolha de vocabulário entre certos profissionais. Um exemplo marcante é o conceito de *símbolo de um operador diferencial linear* D. O símbolo de D é tão somente um polinômio p obtido pela substituição de derivadas parciais (um caso especial de operador diferencial) por termos que ocorrem em p. Uma vez que toda a matemática emprega símbolos, na acepção da discussão na Seção 7, obviamente o símbolo de um operador diferencial é um símbolo, apesar de nem todo símbolo ser o símbolo de um operador diferencial. A palavra 'símbolo' aqui admite duas acepções não equivalentes entre si. Uma é o símbolo D usado para se referir a um dado operador diferencial linear; outra é o polinômio p correspondente ao operador linear D.

Outro exemplo digno de nota é o Princípio de Partição (PP), discutido na Seção 112. De acordo com PP, se $f : x \to y$ é uma função sobrejetora, então existe função injetora $g : y \to x$. Toda função sobrejetora $f : x \to y$ define uma partição p do domínio x de f (ver Teorema 10.4), de modo que, se z é elemento de p, então f restrita a z é uma função constante. Se estivermos tratando de ZFC, podemos usar o Axioma da Escolha para garantir a existência da função injetora $g : y \to x$ de maneira que g é a inversa de uma função escolha bijetora c cujo domínio é subconjunto de x e cujo co-domínio é y (Teorema 10.4). Exemplos são discutidos na Seção 112. Porém, PP não exige que a função injetora $g : y \to x$ seja a inversa de qualquer função escolha c. Portanto, em princípio, PP é indiferente à partição p de x induzida pela função sobrejetora f. Isso mostra que o nome 'Princípio de Partição' é infeliz. É um nome que sugere atenção especial a uma informação irrelevante.

Muitos outros exemplos podem ser mencionados sobre potenciais confusões de leitura. Mas, para finalizar, comentamos sobre um dos mais graves: *demonstração por indução*.

Demonstrações por indução − como aquelas empregadas nos Teoremas 4.2 e 4.6, entre outros − são feitas por infinitas aplicações de Modus Ponens, o qual é um argumento dedutivo, não indutivo. Argumentos dedutivos, como discutido na Seção 9, são relações entre fórmulas, de modo que uma única fórmula é consequência imediata de outras. Argumentos indutivos, por sua vez, são relações entre fórmulas nas quais se atribui um *grau de suporte* para inferir uma fórmula a partir de outras. Neste sentido, a inferência indutiva não é necessariamente única. A abordagem mais usual para expressar graus de suporte é através do emprego de funções de probabilidade condicional, conforme Seção 103. Detalhes podem ser vistos na Seção 105, bem como no livro de Ian Hacking [20]. No Exemplo 9.12 foi ilustrado que mais provavelmente a vítima foi atropelada por um táxi verde. O grau de suporte dessa inferência indutiva em especial é de 59%.

Portanto, a palavra *indução* é um exemplo de polissemia em matemática. Mais de um sentido é atribuído à mesma terminologia.

Seção 118

Neologismos e polissemia

Sumário
Índice
Rede

Aparentemente o matemático persa Sharaf al-Dīn al-Tūsī foi o primeiro a sugerir a ideia intuitiva de termos dependentes de outros, no século 12. Isso sugere que al-Tūsī foi o primeiro a sugerir algo semelhante ao atual conceito de função. Há quem defenda que o pensador persa chegou a introduzir rudimentos de cálculo diferencial e integral [25]. Mas, se insistirmos que somente coisas com nomes podem existir, então fica bem mais fácil determinar quando nasceu o importantíssimo conceito de função em matemática: foi no século 17.

Gottfried Leibniz foi o primeiro a usar o termo 'função', para se referir à dependência de uma variável relativamente a outras. Mas, obviamente, o que Leibniz entendia por funções não coincide necessariamente com os atuais conceitos para tal termo. Afinal, em ZF, funções são casos particulares de conjuntos. A Teoria ZF, por sua vez, nasceu apenas no século 20.

Isaac Newton, contemporâneo de Leibniz, não usava o termo 'função'. No lugar disso, ele se referia a variáveis independentes como *fluentes* e variáveis dependentes como *relata quantitas*. Apesar da grande influência da obra de Newton, foi a terminologia de Leibniz que se estabeleceu, o qual introduziu também as palavras 'constante' e 'parâmetro' ao vocabulário matemático.

Mas alguns matemáticos claramente demonstram grande preocupação com a introdução de neologismos nesta área do saber. Um deles foi Ralph Philip Boas Jr. Em seu artigo de 1981 [6], Boas faz vários alertas. Listamos apenas alguns.

i: Não tente desfazer erros do passado. Se alguém acredita ser capaz de criar um novo nome mais adequado para um conceito matemático usual, possivelmente este alguém está certo. No entanto, matemática é uma atividade social que envolve um esforço coletivo de milhares de pessoas. Convencê-las a mudar terminologia é possivelmente uma perda de tempo. Com efeito, matemáticos preferem se preocupar com matemática, não com nomes.

ii: Examine a literatura, antes de introduzir nova terminologia. Palavras como *distribuição*, *função característica*, *norma* e *módulo*, assumem diferentes significados, dependendo do ramo matemático em contexto.

iii: Não crie novos nomes para conceitos que são usados uma única vez. Apesar de nomes serem úteis para fins de concisão de afirmações, os conceitos matemáticos são mais importantes do que seus nomes.

Uma excelente discussão sobre a história dos nomes em matemática pode ser encontrada no link http://www.economics.soton.ac.uk/staff/aldrich/Mathematical%20Words.htm.

Item i acima aponta para uma contradição inerente entre matemáticos: matemática é mais urgente do que nomes, apesar de nomes exercerem impacto sobre a matemática. Mas essa atitude é compreensível. Com efeito, matemáticos são seres humanos e, portanto, seres inconsistentes.

Referências

[1] A. Abian and S. LaMacchia, *On the consistency and independence of some set-theoretical axioms*, ***Notre Dame Journal of Formal Logic***, vol. XIX (1978), pp. 155–158.

[2] V. I. Arnol'd, ***Mathematical Methods of Classical Mechanics***, Springer, 1997.

[3] E. T. Bell, ***Men of Mathematics***, Touchstone Books, 1986.

[4] J. Bell, ***A Primer of Infinitesimal Analysis***, Cambridge University Press, Cambridge, 2008.

[5] ———, ***Intuitionistic Set Theory***, https://publish.uwo.ca/~jbell/INTBOOK(Repaired).pdf, 2018.

[6] R. P. Boas, *Can we make mathematics intelligible?*, ***The American Mathematical Monthly***, vol. 88 (1981), pp. 727–731.

[7] M. Bologna and G. Aquino, *Deforestation and world population sustainability: a quantitative analysis*, ***Nature Scientific Reports***, vol. 10 (2020), p. 7631.

[8] T. Button and S. Walsh, ***Philosophy and Model Theory***, Oxford University Press, Oxford, 2018.

[9] W. Carnielli and M. E. Coniglio, ***Paraconsistent Logic: Consistency, Contradiction and Negation***, Springer, 2018.

[10] B. Castrucci, ***Fundamentos da Geometria***, FFCL, Santo André, 1973.

[11] N. C. A. da Costa and R. Chuaqui, *On Suppes set theoretical predicates*, ***Erkenntnis***, vol. 29 (1988), pp. 95–112.

[12] N. C. A. da Costa and F. A. Doria, ***On Hilbert's Sixth Problem***, Springer, 2021.

[13] P.B. Denton, S. J. Parke, T. Tao, and X. Zhang, *Eigenvectors from eigenvalues: a survey of a basic identity in linear algebra*, ***Bulletin of the American Mathematical Society***, vol. DOI: https://doi.org/10.1090/bull/1722 (2021).

[14] U. Dudley, ***Mathematical Cranks***, MAA Press, Rhode Island, 2019.

[15] H. Field, ***Science Without Numbers***, Princeton University Press, New Jersey, 1980.

[16] A. A. Fraenkel, *The notion "definite" and the independence of the axiom of choice*, ***From Frege to Gödel*** (J. van Heijenoort, editor), Harvard University Press, Cambridge, 1967, pp. 284–289.

[17] H. Goldstein, C. Poole, and J. Safko, ***Classical Mechanics***, Pearson, 2001.

[18] L. Graham, *The power of names*, ***Theology and Science***, vol. 9 (2011), pp. 157–164.

[19] H. Guidorizzi, ***Um Curso de Cálculo, Volume 1***, LTC, 2018.

[20] I. Hacking, ***An Introduction to Probability and Inductive Logic***, Cambridge University Press, Cambridge, 2001.

[21] D. J. Hand, ***The Improbability Principle***, Scientific American, 2014.

[22] J. Heijenoort, ***From Frege to Gödel: A Source Book in Mathematical Logic***, Harvard University Press, 2002.

[23] D. Hilbert, ***Foundations of Geometry***, Open Court, Chicago, 1980.

[24] ———, ***Fundamentos da Geometria***, Gradiva, Lisboa, 2003.

[25] J. P. Hogendijk, *Sharaf al-Dīn al-Ṭūsī on the number of positive roots of cubic equations*, ***Historia Mathematica***, vol. 16 (1989), pp. 69–85.

[26] M. Jammer, ***Concepts of Mass in Classical and Modern Physics***, Dover, New York, 2010.

[27] T. Jech, ***The Axiom of Choice***, Dover, New York, 1973.

[28] ———, ***Set Theory***, Springer, Berlin, 2003.

[29] W. Johnston, ***The Lebesgue Integral: An Elementary Approach***, Mathematical Association of America, 2015.

[30] A. Kanamori, ***The Higher Infinite***, Springer, Berlin, 2003.

[31] F. Klein, ***Elementary Mathematics from a Higher Standpoint: Volume 1***, Springer, Berlin, 2016.

[32] N. Koblitz, ***A Course in Number Theory and Cryptography***, Springer, Berlin, 1994.

[33] S. Kuhn, *The derivative á la caratheodory*, ***The American Mathematical Monthly***, vol. 98 (1991), pp. 40–44.

[34] S. Lang, ***Algebra***, Springer, New York, 2002.

[35] E. L. Lima, ***Álgebra Linear***, IMPA, Rio de Janeiro, 2004.

[36] R. B. McFee and J. B. Leikin, *Death by Polonium-210: lessons learned from the murder of former soviet spy Alexander Litvinenko*, ***Seminars in Diagnostic Pathology***, vol. 26 (2009), pp. 61–67.

[37] R. McKee, ***Story: substância, estrutura, estilo e os princípios da escrita de roteiro***, Arte & Letra, 2017.

[38] E. Mendelson, ***Introduction to Mathematical Logic***, Chapman & Hall, London, 1997.

[39] R. H. Risch, *The problem of integration in finite terms*, ***Transactions of the American Mathematical Society***, vol. 139 (1969), pp. 167–189.

[40] A. M. Robert, ***Nonstandard analysis***, Dover, New York, 2011.

[41] D. (ed.) Runes, ***Dictionary of Philosophy***, Littlefield, Adams & Co., Inc., New Jersey, 1971.

[42] B. Russell, *Recent work on the principles of mathematics*, ***International Monthly***, vol. 4 (1901), pp. 83–101.

[43] A. S. Sant'Anna, ***O que é uma Definição***, Manole, Barueri, SP, 2005.

[44] A. S. Sant'Anna, R. Brodzinski, M. P. P. França, and O. Bueno, *Flow: the Axiom of Choice is independent from the Partition Principle in ZFU*, ***submitted for publication***.

[45] A. S. Sant'Anna, N. C. A. da Costa, and F. A. Doria, *The Atiyah-Singer index theorem and the gauge field copy problem*, ***Journal of Physics A: Mathematical and General***, vol. 30 (1997), pp. 5511–5516.

[46] A. S. Sant'Anna and C. Garcia, *Gravitation in Hertz's mechanics*, ***Foundations of Physics Letters***, vol. 16 (2003), pp. 559–572.

[47] A. S. Sant'Anna and C. A. S. Maia, *Axioms for Mach's mechanics*, ***Foundations of Physics Letters***, vol. 14 (2001), pp. 247–262.

[48] D. Scott and D. McCarty, *Reconsidering ordered pairs*, ***The Bulletin of Symbolic Logic***, vol. 14 (2008), pp. 379–397.

[49] C. E. Shannon, *A mathematical theory of communication*, ***Bell System Technical Journal***, vol. 27 (1948), pp. 379–423.

[50] A. Shen and N. K. Vereshchagin, ***Basic Set Theory***, AMS, 2002.

[51] R. P. Stanley, ***Catalan Numbers***, Cambridge University Press, Cambridge, 2015.

[52] J. R. Steel, *What is a Woodin cardinal?*, ***Notices of the American Mathematical Society***, vol. 54 (2007), pp. 1146–1147.

[53] P. Suppes, ***Introduction to Logic***, van Nostrand, Princeton, 1957.

[54] ———, ***Representation and Invariance of Scientific Structures***, CSLI, Stanford, 2002.

[55] E. W. Swokowski, ***Cálculo com Geometria Analítica, Volume 1***, Makron, 1994.

[56] A. Tarski and S. Givant, ***A Formalization of Set Theory Without Variables***, AMS, 1987.

[57] K. Trlifajova, *Bolzano's infinite quantities*, ***Foundations of Science***, vol. 23 (2018), pp. 681–704.

[58] W. Wieslaw, ***Topological Fields***, Marcel Dekker, 1988.

[59] R. Wilson, ***Euler's Pioneering Equation : the Most Beautiful Theorem in Mathematics***, Oxford University Press, Oxford, 2018.

99

Índice Remissivo

Símbolos

Pessoas

Demais itens

Este documento foi editado em LaTeX e compilado por PDFTeXify.