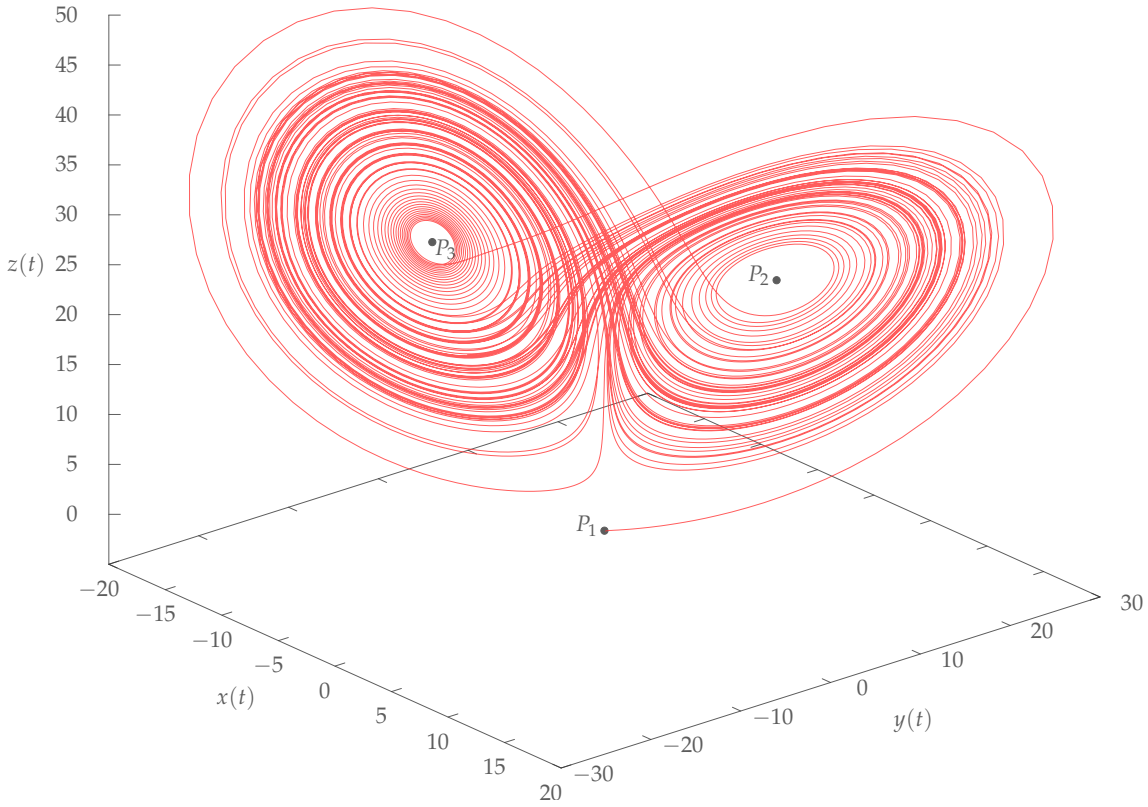
Métodos Computacionais da Física

Notas de aula das disciplinas LOM3260 e LOM3227

Luiz T. F. Eleno



Departamento de Engenharia de Materiais
Escola de Engenharia de Lorena
Universidade de São Paulo

versão 0.3

2021

Métodos Computacionais da Física

Notas de aula das disciplinas LOM3260 e LOM3227

Luiz T. F. Eleno

Departamento de Engenharia de Materiais
Escola de Engenharia de Lorena
Universidade de São Paulo

versão 0.3

2021

© 2021, Luiz T. F. Eleno.

Licença *Creative Commons CC BY-NC-ND 4.0:* Você tem o direito de compartilhar — copiar e redistribuir o material em qualquer suporte ou formato (o licenciante não pode revogar estes direitos desde que você respeite os termos da licença), de acordo com os termos seguintes: (1) Atribuição — você deve dar o crédito apropriado, prover um link para a licença e indicar se mudanças foram feitas. Você deve fazê-lo em qualquer circunstância razoável, mas de nenhuma maneira que sugira que o licenciante apoia você ou o seu uso. (2) Não Comercial — Você não pode usar o material para fins comerciais. (3) Sem Derivações — Se você remixar, transformar ou criar a partir do material, você não pode distribuir o material modificado. (4) Sem restrições adicionais — Você não pode aplicar termos jurídicos ou medidas de caráter tecnológico que restrinjam legalmente outros de fazerem algo que a licença permita. https://creativecommons.org/licenses/by-nc-nd/4.0/deed.pt_BR.

Editoração eletrônica: Luiz T. F. Eleno

Universidade de São Paulo (USP)
Escola de Engenharia de Lorena (EEL)
Departamento de Engenharia de Materiais (Demar)

Polo Urbo-Industrial Gleba AI-6 Caixa Postal 116
12600-970 Lorena – SP

Contato
+55 12 3159 9810 luizeleno@usp.br http://www.demar.eel.usp.br/docentes/luiz-tadeu-fernandes-eleno.html

versão eletrônica e código-fonte: https://github.com/luizeleno/LOM3227

Nota

Muito zelo e técnica foram empregados na edição desta obra. No entanto, podem ocorrer erros de digitação, impressão ou dúvida conceitual. Em qualquer das hipóteses, solicitamos a comunicação direta com o autor, para que possamos esclarecer a questão.

DADOS BIBLIOGRÁFICOS

Luiz T. F. Eleno
Métodos Computacionais da Física
2021, versão 0.3
Lorena: Demar–EEL–USP, 2021.
157 p. : il.

Inclui Bibliografia / Índice de assuntos / Índice de autores

20 de janeiro de 2021

Aprenda a resolver todo problema já resolvido.

Richard Feynman

Prefácio

Não é só deus que sabe: eu sei, e, ao final do semestre, vocês saberão.

Sidney Coleman

Métodos computacionais não significam apenas usar um computador para resolver problemas. Realmente, à Física Computacional moderna é fundamental o acesso a algum *hardware* de média a alta capacidade e o domínio de alguma linguagem de programação. No entanto, essa é uma das últimas etapas de um processo que, por falta de melhor nome, chamamos de **simulação** computacional: a etapa em que, efetivamente, usamos um computador para testar hipóteses e obter resultados a partir da **implementação** de algum **modelo teórico** de determinado problema do mundo real. Antes disso, porém, é preciso entender o fenômeno em questão, analisá-lo, *sintetizá-lo* (ou seja, simplificá-lo) e propor um modelo que, ainda que inevitavelmente mais simples que a realidade, consiga captar sua essência, ou, pelo menos, parte dela. É como criar um vinho de buquê específico, único e particular, que contenha as nuances e matizes das uvas, do terreno, da região e do próprio viticultor... Obviamente, nesse curso não pretendemos necessariamente formar enólogos peritos em diferenciar e criar diferentes safras de *grands crus*. Mas é sempre bom saber distinguir entre um bom vinho e um daqueles de garrafão!

Mais cedo ou mais tarde, no entanto, inevitavelmente acabamos chegando aos computadores e à implementação computacional. Veremos que, já nos modelos mais simples, as matrizes — e, portanto, a Álgebra Linear — começarão a nuclear por toda a parte, como gotas de orvalho em mudas recém-germinadas. A solução numérica de sistemas lineares, por exemplo, é geralmente vista em cursos introdutórios de Cálculo Numérico.[1] Uma aplicação mais avançada, a determinação numérica de autovalores e autovetores, é geralmente relegada para mais tarde, e, mais frequentemente, não é sequer apresentada. No entanto, encontrar autovalores e autovetores é o arroz-com-feijão (ou o spaghetti que acompanha o vinho) da Física Computacional. Isso sem falar nos já mencionados e onipresentes sistemas lineares. Por isso aqui, já nos primeiros capítulos, matrizes serão usadas sem pudor, seja para a obtenção de belas imagens fractais, seja para a solução de equações diferenciais, seja para analisar as notas musicais que formam um acorde, os harmônicos envolvidos e a forma de onda advinda do timbre e do ataque ao instrumento.

A disciplina "Métodos Computacionais da Física" (LOM3227) é ministrada aos alunos do 5º semestre do curso de Engenharia Física da Escola de Engenharia de Lorena de Universidade de São Paulo (EEL-USP). A essa altura do curso, os alunos (idealmente) já concluíram o ciclo fundamental de disciplinas de Física básica, Cálculo diferencial e integral, Álgebra Linear, Programação e Cálculo Numérico, além de uma introdução a técnicas matemáticas um pouco mais avançadas da Física Matemática. Esse é um caudaloso rio de conteúdo em apenas dois anos! É mais que provável que boa parte desse enorme volume de saber não tenha sido inteiramente assimilado, e uma parte tenha ido ralo abaixo; na melhor das hipóteses, encontra-se represado nos livros, apostilas, apresentações de *slides* e anotações de aula colecionados pelos estudantes em seus primeiros quatro semestres de curso. E isso não é um problema exclusivamente nacional, como mostra o meme ao lado ("quem somos?", "os estudantes!", "o que fazemos?", "estudamos!", "e depois?!?!", "esquecemos tudo!").

[1] Para os alunos de Engenharia Física da EEL-USP, os primeiros conceitos de cálculo numérico são apresentados na disciplina LOM3260 — Computação Científica em Python, ministrada no 2º semestre do curso.

Há um fator ainda mais crítico e que, de certa maneira, ajuda a turvar e dispersar tanto conhecimento supostamente, mas não realmente adquirido: muitas vezes, o conteúdo é ministrado sem bula — sem indicação e contra-indicação: os alunos, ou boa parte deles, não têm a menor ideia da utilidade de muitos dos conceitos vistos no começo do curso! Esse isolamento entre conceito e aplicação é, na práxis da universidade, quase sempre inevitável, pois as disciplinas introdutórias são frequentemente ministradas por docentes de áreas desconexas da prática da engenharia ou física aplicada, muitas vezes mais interessados em discursos outros que a aplicação (como o rigor matemático, com exaustiva demonstração de teoremas), a discentes também de diversas áreas, com diferentes motivações. É difícil de entender o porquê de ser, intrinsecamente, inevitável que boa parte dos alunos sinta-se alienada, sempre com pouco interesse nos conceitos que não entende o motivo de ter que aprender?

As presentes notas de aula são uma tentativa, talvez ambiciosa e possivelmente ingênua, de usar todo aquele volume fluido de conhecimentos do ciclo básico para irrigar e, quiçá, ver nascer alguns frutos, ao trazê-los para a terra firme do mundo da engenharia e da física aplicada modernas. Foi-se o tempo em que engenharia era simplesmente aplicar com bom senso alguns conceitos de física. É provável que um moderno profissional da área de Exatas vá passar boa parte do tempo contatando fornecedores e clientes, preenchendo formulários e relatórios, espremendo números e tabelas em figuras impactantes para apresentar seus resultados de forma otimista, e outras tarefas que exigirão, realmente, muito pouco da abordagem teórica e mecanicista dos primeiros semestres (a até dos demais) de qualquer curso de engenharia. No entanto, fica sempre a proverbial pulga atrás da orelha: você, estudante de engenharia ou física, e ainda mais de áreas na fronteira tecnológica com as quais lida um engenheiro físico (eletrônica, fotônica, materiais especiais, computação, automação), apreciaria a oportunidade de entender e aprender a aplicar na prática os conceitos lógicos e altamente úteis da Matemática e da Física? Se a sua resposta é negativa, abandone agora a carreira que escolheu: seu mundo não é o das Exatas e você está perdendo tempo. Caso sua resposta seja um Sim, veemente e seguro de si, pode parar de ler este prefácio: não preciso lhe convencer. Se, por fim, sua resposta também é um sim, mas acanhado e indeciso, coçando a cabeça, a única resposta que posso lhe dar nesse momento é: aprenda com o curso, leia estas notas de aula, resolva os exercícios, acompanhe as aulas. Talvez, ao final do semestre, seu tímido sim se torne um estrondoso Sim — ainda que a modéstia e a prudência me obriguem a dizer que já ficarei contente se não transformar muitos sim's em não's...

Luiz T. F. Eleno
Lorena, janeiro de 2019

Sobre o autor

Sou de São Paulo, capital, nascido em 1976. Concluí minha graduação em Engenharia Metalúrgica pela Escola Politécnica (EP) da Universidade de São Paulo (USP), em 2001, e logo engatilhei o mestrado na mesma instituição, com a dissertação "Incorporação do Volume ao Método Variacional de Clusters," defendida em 2003. Depois disso, embarquei para a Alemanha, onde passei quatro anos e meio ensopado em cerveja e trabalhando no *Max-Planck-Institut für Eisenforschung* (Instituto Max Planck de Pesquisa em Ligas Ferrosas), na cidade de Düsseldorf, de 2003 a 2008. Lá caí (literalmente) do trem, quebrei a perna direita, onde tenho até hoje três pinos e uma placa metálica — selando minha simbiose com a Metalurgia. Logo em seguida, e ainda de muletas, conheci o amor da minha vida, também paulistana, coincidências. Fomos juntos para Toulouse, França, por mais seis meses, onde fui pesquisador convidado do *Centre Interuniversitaire de Recherche et d'Ingénierie des Matériaux* (CIRIMAT) da *Université de Toulouse*. Voltamos, finalmente, ao Brasil, onde defendi minha tese de doutorado com o título "Dados empíricos e ab initio no método CALPHAD: os sistemas Fe–Cr–Mo–C e Nb–Ni–Si" em 2012, na EP-USP. Logo em seguida, dediquei-me ao Pós-Doutorado em Física pelo Instituto de Física (IF) da USP até 2014. Concluída mais essa etapa, começaram minhas atividades como Professor Doutor junto ao Departamento de Engenharia de Materiais (Demar) da Escola de Engenharia de Lorena (EEL) da USP, na cidade de Lorena, interior do estado de São Paulo, onde amarrei finalmente minha mula, que desamarro às vezes para saborear cerveja mundo afora, como na foto ao lado, em Oradea, Romênia. Quando não estou me dedicando a essas atividades, faço pesquisas em cálculo de propriedades eletrônicas e térmicas de materiais metálicos e ministro aulas às turmas de Engenharia de Materiais e Engenharia Física da EEL, nas disciplinas de Termodinâmica, Física Estatística, Mecânica Quântica, Computação Científica e Métodos Computacionais.

Sumário

Lista de Símbolos

Obs.: os números indicam a página em que os símbolos, constantes e unidades aparecem pela primeira vez. Estão listadas apenas as grandezas mais importantes, ou que recorrem ao longo do livro.

Abreviações

FFT *Fast Fourier Transform*, ou Transformada Rápida de Fourier. 89, 92

PRNG *Pseudorandom number generator*, ou Gerador de números pseudo-aleatórios. 7

Símbolos no alfabeto latino

m massa. 15

 © Luiz T. F. Eleno, v. 0.3, 2021

Parte I

Introdução

Capítulo 1

Prelúdio: calculando a natureza

> Física significa questionar, estudar, sondar a natureza. Você investiga e, se tiver sorte, encontra estranhas pistas.
>
> Lene Hau

1.1 Introdução: estranhos formatos

Uma bela árvore frondosa, como a da Figura 1.1, se ramifica em muitos galhos, às vezes entrelaçados, outras vezes mais livres, crescendo em direção ao sol. Esses ramos principais, por sua vez, também se dividem em vias mais estreitas, em muitos pequenos galhinhos, que também se subdividem. Vêm então as folhas, que também, quando chega a vez delas, e se olharmos bem de perto, têm uma estrutura ramificada, com uma nervura central que se reparte em vários outros canais menores, ordens de grandeza menores que o robusto tronco que sustenta toda a complexa estrutura natural, oscilando ao ritmo do vento. Isso sem falar das raízes...

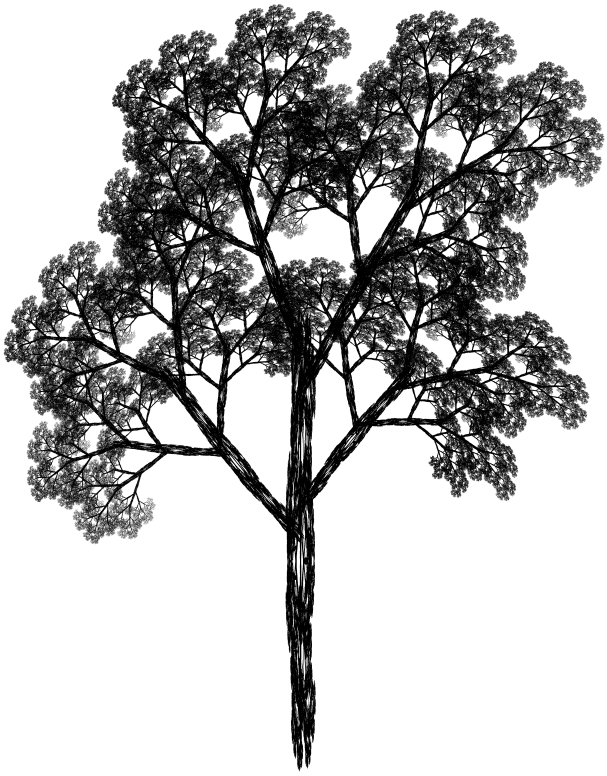


FIGURA 1.1: A árvore da figura não foi desenhada à mão: foi gerada no computador. É uma árvore fractal, criada por fórmulas incrivelmente simples que levam a estruturas fantasticamente complexas.

A natureza está repleta de tais padrões: um mesmo motivo que se repete em muitas escalas, muitas vezes tão similares que perdemos a noção do tamanho do objeto. Para usarmos o mesmo exemplo botânico do parágrafo anterior: um graveto, visto bem de perto, tem às vezes a aparência de uma árvore completa. Essa *autossimilaridade* da natureza é um dos aspectos que definem os objetos que físicos, matemáticos e cientistas da computação decidiram chamar de *fractais*. Um fractal é exatamente isso: algo que, visto como um todo ou em detalhe, de longe ou de perto, tem sempre a mesma aparência. Não precisamos ir muito longe: um segmento de reta é um fractal (ainda que sem muita graça!), pois continua reto visto tanto de longe quanto de perto. Verdade? No mundo idealizado, euclidiano, sim. Mas no mundo real é um pouco diferente: não existem linhas retas. De perto, com o auxílio de uma lupa, um segmento de reta, desenhado com lápis e régua no papel, apresentará inumeras irregularidades, devido à rugosidade da régua e do próprio papel, além da variação de espessura da grafita depositada. Mas nossa linha real continua tendo um caráter fractal: ampliações maiores levarão novamente a uma natureza fractal: uma região que, sob uma lupa fraca, parecia lisa, na verdade também será estranhamente rugosa sob outra mais potente.

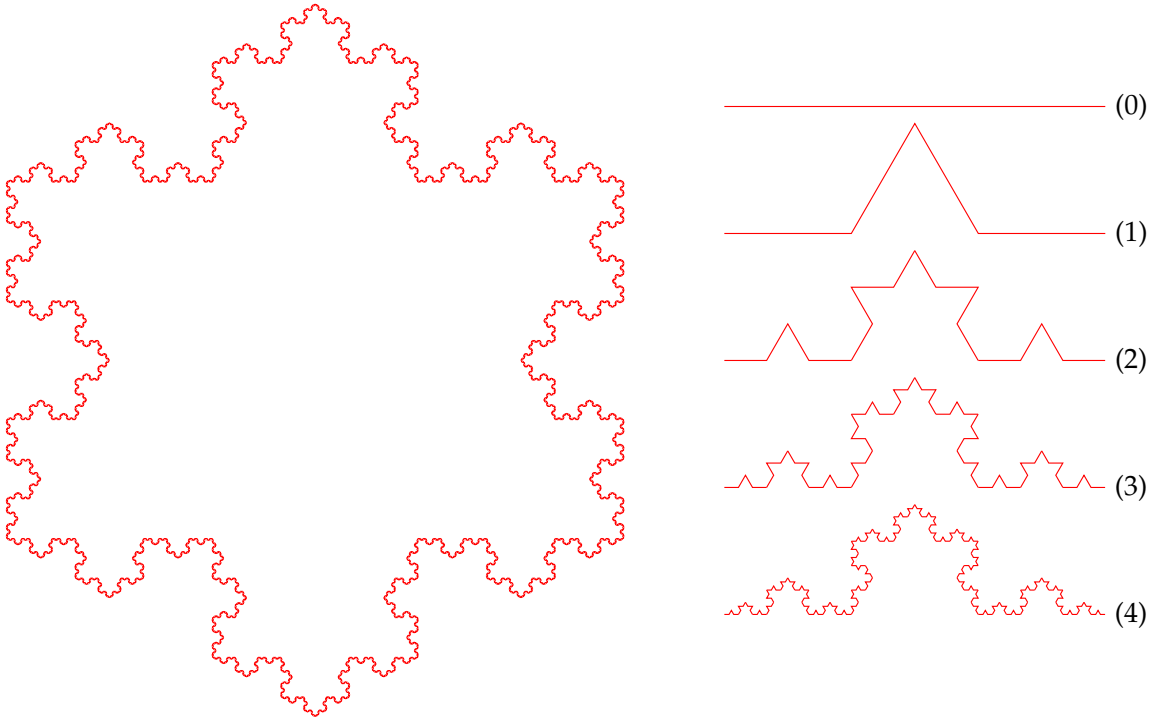


FIGURA 1.2: O fractal conhecido como Floco de Koch. Para construí-lo, comece com um triângulo equilátero e, em cada segmento de reta, troque o terço central por outros dois de mesmo comprimento. As quatro primeiras etapas estão ilustradas acima à direita.

Há um fractal que ilustra muito bem esse aspecto rugoso da natureza: a curva de Koch,[1] mostrada em toda sua glória na Figura 1.2. Ainda que altamente regular, esse fractal parece o litoral extremamente entrecortado de uma ilha vulcânica perdida no oceano, repleta de pequenas enseadas, costões, baías e penínsulas. Parece também uma das inúmeras configurações adotadas por um cristal (ou floco) de neve [1]. As primeiras etapas da construção da curva de Koch estão também indicadas na Figura 1.2. Comece com um triângulo equilátero. No passo 1, para cada um dos lados do triângulo (no passo 0, apenas um dos lados está indicado), remova o terço central (ou seja, divida o segmento em três partes iguais e remova a parte do meio), substituindo-o por outros dois segmentos de mesma extensão, formando um outro triângulo equilátero (do qual fica faltando a base). A partir daí, itere o processo para todos os segmentos de reta. Por exemplo, o passo 1 resulta em quatro segmentos de reta por lado do triângulo, que deverão passar pelo processo de cortar e colar que resulta no passo 2, que tem agora 16 segmentos de reta. O floco de Koch é a figura (o fractal) resultante de continuar o processo indefinidamente, ou seja, infinitamente. É natural esperar que um objeto desses não exista na prática: se não por outra coisa, uma hora o nível atômico é atingido e não mais temos a condição de continuidade dos segmentos. Mas isso não deve nos espantar, assim como não nos espanta o fato de que não existem linhas retas na prática, como já discutimos. Fractais são objetos tão idealizados quanto linhas retas. Mas ambos, fractais e linhas retas, são bem aproximados na natureza: pense nos litorais entrecortados e na linha do horizonte em alto mar, por exemplo.

Uma vez aceita a existência idealizada de um fractal como a curva de Koch, uma pergunta que podemos nos fazer é: qual é o seu comprimento? Vamos imaginar que os lados do triângulo original têm comprimento igual a uma unidade, e vamos focar em apenas um dos lados. Vejamos: no passo 1 da Figura 1.2, aumentamos o comprimento em $1/3$ de unidade, ou seja, temos agora um comprimento de $4/3$ unidade. Do passo 1 ao 2, aumentamos cada segmento em $1/9$ e, como tínhamos 4 segmentos no passo 1, ficamos com um comprimento de $4 \times (1/3 + 1/9) = 16/9 = (4/3)^2$ unidade no passo 2. É fácil agora generalizar: o comprimento L_n do objeto no passo n é dado por

$$L_n = \left(\frac{4}{3}\right)^n \tag{1.1.1}$$

Não é dificil de ver, então, que a curva de Koch (que equivale a fazer $n \to \infty$ na Eq. 1.1.1) tem um comprimento infinito. O interessante é que a área cercada pelo fractal não é infinita (se fosse, não poderíamos representá-la num pedaço de papel de área finita!). Ou seja, a curva de Koch tem um perímetro infinito delimitando uma área finita.

[1] Niels Fabian Helge von Koch (1870—1924), matemático sueco.

Vejamos agora um outro exemplo de fractal famoso, o Triângulo de Sierpiński, [2] mostrado na Figura 1.3, em que também algo inusitado acontece. Assim como o Floco de Koch, o Triângulo de Sierpiński também é o resultado de um processo iterativo. Na primeira etapa, removemos o triângulo cujos vértices são os pontos médios dos lados do triângulo original, o que gera outros três triângulos com metade do lado. Repetimos o processo em cada um desses novos triângulos, e ficamos agora com nove triângulos com lado igual a 1/4 do original; e assim por diante, indefinidamente. A Figura 1.3, na verdade, corresponde a apenas oito iterações, como você pode verificar (na versão digital) ampliando a figura.

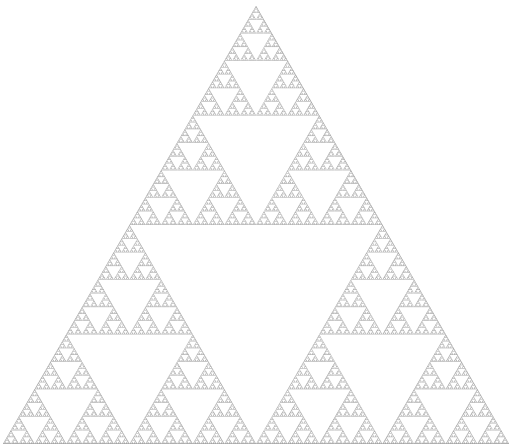


FIGURA 1.3: O fractal conhecido como Triângulo de Sierpiński. Apenas os oito primeiros passos estão representados.

A pergunta agora é: qual é a área do triângulo de Sierpinski? Seguindo o algoritmo descrito no parágrafo anterior, se a área do triângulo original é de 1 unidade, a cada passo n, a área A_n é dada por

$$A_n = \left(\frac{3}{4}\right)^n \tag{1.1.2}$$

e o fractal, que corresponde a $n \to \infty$, tem área nula! Ou seja, o Triângulo de Sierpiński, cujos pontos estão todos contidos num plano, que tem dimensão 2, é uma figura geométrica sem área. Por isso o aspecto enevoado da Figura 1.3, pois o Triângulo de Sierpinski é um objeto quase etéreo... Mas ele não é apenas um aglomerado desconexo de pontos (de dimensão zero), pois é possível demonstrar que, partindo de algum ponto do Triângulo, é possível chegar a qualquer outro ponto sem nunca sair do Triângulo — ou seja o Triângulo de Sierpiński é um objeto contínuo. Linhas (de dimensão 1) também não têm área (apesar de, no mundo real, terem espessura!), também podem ser representadas num plano e também são contínuas — mas o Triângulo de Sierpiński não é uma linha! Ele é um fractal e, como tal, talvez tenha uma *dimensão fractal*, ou seja, uma dimensão quebrada, entre 1 (como uma linha poligonal ou um círculo) e 2 (como um triângulo ou um disco), o que faz com que ele seja mais que um aglomerado de pontos e menos do que um objeto bidimensional.

Você pode se perguntar: e existe alguma aplicação prática para tudo isso, além de conceitos esquisitos e belas imagens? Sim, existem muitas: trocadores de calor, misturadores, antenas, além do estudo do câncer, de modelos de percolação, de trânsito e urbanismo, de solidificação, ... Veja mais detalhes em http://fractalfoundation.org. Uma aplicação pode, inclusive, estar nesse momento na sua mão (ou no seu bolso): seu celular pode ter (e, se não tem, o próximo quase certamente terá) uma antena fractal, no formato de um fractal como o de Koch ou de Sierpiński [2]. Obviamente, a antena não é um verdadeiro fractal, mas um objeto como os que ilustram o presente capítulo, com apenas alguns dos primeiros passos da sua construção. A ideia de fractal, no entanto, é a motivação por trás do projeto de tais antenas: compactar ao máximo o comprimento de antena no menor espaço possível, ou trabalhar o efeito amplificador de superfícies fractais, conseguindo uma boa captação num espectro amplo de frequências.

1.2 Dimensão fractal

Nada do que estamos falando aqui é muito novo. Na verdade, boa parte do que já vimos até agora, como os dois fractais apresentados (e também a árvore do começo do capítulo), surgiu no começo do século XX. Mas só começou a ser estudado por mais gente, e só começou a atrair a atenção dos físicos, a partir dos anos de 1980, quando Benoît Mandelbrot[3] publicou o divisor de águas *"The Fractal Geometry of Nature"* [3]. Mandelbrot cunhou o termo *fractal* para caracterizar objetos como os exemplos que vimos, buscando também inúmeros outros exemplos na natureza. Seu livro virou um clássico moderno da literatura científica.

Antes disso, porém, já no século XIX, diversos matemáticos começaram a investigar uma área que recebeu o nome de Topologia, e que, recentemente, ganhou ímpeto com o desenvolvimento de materiais com propriedades topológicas. No início, porém, a Topologia era pouco mais que uma curiosidade matemática, ainda que bastante abstrata, rica e interessante. Felix Hausdorff[4] foi um desses matemáticos, dentre alguns outros. Hausdorff decidiu ampliar e generalizar o conceito de dimensão de um objeto através de conceitos topológicos. Aqui, neste capítulo introdutório que deve servir para atrair e não repelir o leitor, não precisamos entrar em todos os detalhes técnicos (e não seria eu o mais indicado a fornecê-los!). Basta tentar

[2] Wacław Franciszek Sierpiński (1882–1969), matemático polonês (sim, tem um acento agudo no n e um corte no l!).
[3] Benoit B. Mandelbrot (1924–2010) polímata francês nascido na Polônia.
[4] Felix Hausdorff (1868–1942), matemático alemão.

entender intuitivamente a ideia da dimensão de Hausdorff, ou dimensão fractal, e aplicá-la aos dois fractais que vimos até aqui.

Pois bem. Sem maiores delongas, Hausdorff decidiu calcular a dimensão de qualquer objeto geométrico usando a seguinte expressão:[5]

$$D = \lim_{\epsilon \to 0} \frac{\log N(\epsilon)}{\log(1/\epsilon)} \tag{1.2.1}$$

que ficou conhecida posteriormente como dimensão de Hausdorff ou dimensão fractal [4]. Na Eq. (1.2.1), $N(\epsilon)$ é o número mínimo de caixas de mesmo tamanho que precisamos para recobrir todo o objeto, sendo ϵ o tamanho dessas caixas. O significado de "caixa" ficará mais claro com os exemplos a seguir.

Para tentar entender o que diz a Eq. (1.2.1), vejamos os exemplos da Figura 1.4. Na parte (a), temos um segmento de reta de comprimento L, do qual queremos determinar a dimensão — que sabemos ser igual a um. Então subdividimos o segmento em pedacinhos menores, de tamanho ϵ, que são as nossas caixas nesse caso. O número de caixas é $N(\epsilon) = L/\epsilon$, de forma que, segundo Hausdorff,

$$D = \lim_{\epsilon \to 0} \frac{\log(L/\epsilon)}{\log(1/\epsilon)} = \lim_{\epsilon \to 0} \left(-\frac{\log L}{\log \epsilon} + 1 \right) = 1\,, \tag{1.2.2}$$

como já sabíamos de antemão. Já no caso da Figura 1.4b, as caixas são quadradinhos de lado ϵ que recobrem todo o quadrado de lado L, do qual queremos saber a dimensão. Temos agora $(L/\epsilon)^2$ caixas de lado ϵ e portanto, usando a Eq. (1.2.1):

$$D = \lim_{\epsilon \to 0} \frac{\log(L/\epsilon)^2}{\log(1/\epsilon)} = \lim_{\epsilon \to 0} \left(-\frac{\log L^2}{\log \epsilon^2} + 2 \right) = 2\,, \tag{1.2.3}$$

o que não é de se espantar: um quadrado tem dimensão dois.

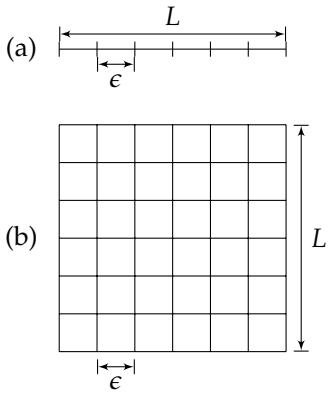


FIGURA 1.4: Caixas de dimensão ϵ usadas para calcular a dimensão de Hausdorf de (a) um segmento de reta e (b) de um quadrado.

Poderíamos continuar com outros exemplos, como uma linha curva qualquer (dimensão 1), ou um objeto tridimensional como um cubo ou uma pirâmide. A fórmula de Hausdorff forneceria sempre a dimensão correta. Mas vamos passar direto ao caso dos fractais que já vimos. Vamos começar com a curva de Koch. Voltemos então aos quatro primeiros passos da Figura 1.2. Basta focarmos num dos lados do triângulo original, que vamos imaginar de comprimento unitário. No passo 1, vamos adotar ϵ como um dos quatro segmentos de reta, portanto $\epsilon = 1/3$ e $N(\epsilon) = 4$. Continuando com o mesmo esquema, no passo 2, temos $\epsilon = 1/9 = 1/3^2$ e $N(\epsilon) = 16 = 4^2$; Generalizando: para o passo n, $\epsilon = 1/3^n$ e $N(\epsilon) = 4^n$. Na fórmula de Hausdorff, isso fica

$$D = \lim_{\epsilon \to 0} \frac{\log N(\epsilon)}{\log(1/\epsilon)} = \lim_{n \to \infty} \frac{\log 4^n}{\log 3^n} = \frac{\log 4}{\log 3} \approx 1.262 \tag{1.2.4}$$

Portanto, a curva de Koch tem uma dimensão quebrada, fracionária, o que o torna realmente um fractal! Ele tem dimensão um pouco maior que a de uma linha. Cuidado: é preciso fazer o limite $\epsilon \to 0$ (ou $n \to \infty$) para atingir a curva de Koch verdadeira. Qualquer passo intermediário é apenas uma coleção de segmentos de reta que, sabemos, são unidimensionais. É apenas o fato da curva de Koch continuar indefinidamente a repetir os passos da Figura 1.2 — ou seja, é apenas a sua autossimilaridade — que o torna um fractal e ter dimensão fracionária.

Fica como exercício mostrar que o Triângulo de Sierpiński também é um fractal, com dimensão dada por $D = \log 3/\log 2 \approx 1.585$. Com uma dimensão fracionária (ou fractal), ele é bem mais que um conjunto de linhas, mas fica aquém de um objeto bidimensional, cuja dimensão é exatamente 2.

1.3 Uma estratégia diferente

Nas seções anteriores, vimos maneiras de construir passo a passo um fractal, num processo de cortar (remover) e colar (acrescentar) partes à exaustão, seguindo uma receita muito bem determinada. Essa é uma metodologia que podemos chamar de *método determinístico*: sabemos exatamente o que acontecerá no próximo passo. Mas vamos agora tentar enxergar objetos fractais de outro ângulo, usando uma estratégia que chamaremos de *método estocástico* — do grego στόχος (stókhos), mira, palpite —, em que o acaso, na forma de números aleatórios (do latim *alea*, dado e, por extensão, sorte), desempenha um papel fundamental.

[5] qualquer base serve para os logaritmos, desde que seja a mesma tanto no numerador quanto no denominador.

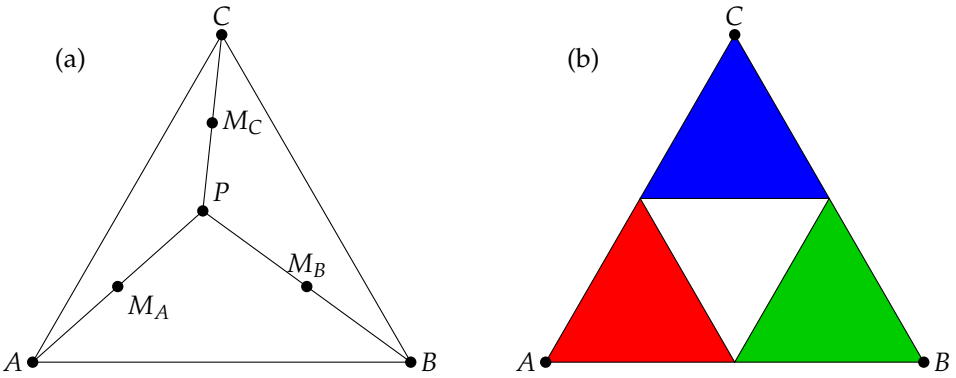


FIGURA 1.5: (a) Regras do jogo: qualquer ponto P no interior do triângulo é substituído pelo ponto médio (M_A, M_B ou M_C) do segmento entre ele e um dos três vértices do triângulo (A, B ou C), escolhidos aleatoriamente. O processo é repetido iterativamente com o novo ponto obtido. (b) A cada iteração, o ponto P é levado ao triângulo colorido mais próximo do vértice escolhido.

Vamos então estabelecer a seguinte receita — ou algoritmo, para usar um nome mais bonito. Considere primeiramente o triângulo equilátero da Figura 1.5a, de vértices $A = (0,0)$, $B = (1,0)$ e $C = (1/2,\sqrt{3}/2)$. O ponto $P = (x,y)$ é qualquer ponto do interior do triângulo. Faça então o seguinte: escolha aleatoriamente um dos três vértices (A, B ou C), determine o ponto médio do segmento de reta que o une a P (M_A, M_B ou M_C) e substitua P por esse novo ponto. Repita esse processo com as novas coordenadas, indefinidamente, escolhendo novamente um dós três vértices e movendo P, durante um número bastante grande de iterações. Faça um gráfico de todo o histórico da movimentação de P — que chamamos de *órbita* de P — ao longo do processo. O que você espera encontrar?

Vejamos, acompanhando a Figura 1.5b: se escolho, digamos, o vértice A, o ponto P vai terminar no triângulo colorido mais próximo desse vértice. Coisa parecida acontece se escolho um dos outros vértices, B ou C. Com mais e mais iterações, o ponto P ficará saltando aleatoriamente entre as três regiões coloridas.

Mas não apenas isso: você pode tentar se convencer de que haverá regiões no interior do triângulo nunca alcançadas pela órbita de P. Por exemplo, após a primeira iteração, P nunca estará no triângulo central (de ponta-cabeça). Da mesma maneira, P nunca estará, nas iterações seguintes, nos triângulos centrais das regiões coloridas, nem em diversas outras regiões que você talvez consiga descobrir quais são. Mas, se tem dificuldades em visualizar, faça o teste, programando um computador para fazer as contas pra você. Você vai precisar de uma biblioteca gráfica e de um gerador de números aleatórios. Se quiser, pode até usar uma planilha eletrônica para isso, que geralmente conta com tais recursos.

Para criar o algoritmo, você precisará calcular as coordenadas de M_A, M_B e M_C que, não é difícil de enxergar, são dadas por

$$M_i = \frac{1}{2}\left(x_P + x_i, y_P + y_i\right), \qquad (i = A, B, C) \tag{1.3.1}$$

e as coordenadas de A, B e C já foram mencionadas um pouco acima. Além disso, é preciso escolher, a cada iteração, um desses três pontos. A maneira mais fácil de se conseguir isso (e menos custosa, computacionalmente falando) é usar um *gerador de números pseudo-aleatórios* (PRNG, *Pseudorandom number generator*). Quando falamos em gerar números aleatórios, pensamos em produzir uma sequência tal que os valores e a ordem em que aparecem seja realmente *randômica*, ou seja, completamente ao acaso. Mas um computador é em grande parte uma máquina determinística e, portanto, usá-lo para isso é uma tarefa mais árdua do que parece. Para isso foram criados os PRNGs que, na verdade, são equações bastante complicadas que fornecem sequências periódicas de números, mas cuja periodicidade é tão alta que, para todos os efeitos, ela parece aleatória. Além disso, tais geradores têm uma propriedade importante do ponto de vista científico: a *reprodutibilidade*. Para usá-las, você precisa fornecer uma *semente* (*seed*) para iniciar a sequência. A semente é fornecida ao gerador, que cospe então o primeiro número aleatório A partir daí, o número apenas gerado, ao invés da semente, é usado para obter o seguinte com a fórmula do gerador. Podemos usar variáveis com certo grau de aleatoriedade como semente (o horário em que o cálculo foi iniciado, por exemplo, ou a umidade relativa do ar), o que garante ainda mais aleatoriedade. Mas, se quisermos reproduzir a mesma sequência de números aleatórios, basta usar a mesma semente, o que pode ser útil em algumas situações.

Além disso, a sequência de números pseudo-aleatórios precisa obedecer alguma *distribuição de probabilidades*, ou seja, deve responder à seguinte pergunta: qual a probabilidade de qualquer dos números da sequência ser encontrado em um dado intervalo? A quase totalidade dos PRNGs gera números uniformemente distribuídos entre dois limites, usualmente no intervalo $[0,1)$. Isso significa que a probabilidade de qualquer número gerado estar, por exemplo, no intervalo $[0,0.1)$ é de 10%, a mesma probabilidade que para o intervalo $[0.54,0.64)$. No Capítulo 12 veremos como obter outras distribuiçoes, como as distribui-

ções Normal, de Poisson, log-normal, de Pareto, ... Felizmente, por enquanto, só precisamos da distribuição uniforme.

Mas voltemos à nossa receita dos triângulos. Precisamos escolher entre três pontos, o que significa que nosso PRNG precisa fornecer, a cada iteração, uma escolha dentre três possibilidades. Por exemplo, ele poderia escolher aleatoriamente qualquer número inteiro do conjunto $\{1, 2, 3\}$, com a mesma probabilidade $1/3$. Ainda que existam PRNGs que fazem exatamente isso, há uma maneira mais fácil usando o gerador mais básico, que fornece números `float` no intervalo $[0, 1)$. Na Figura 1.6, vemos que podemos dividir o intervalo $[0, 1)$ em três partes congruentes. Quando o número aleatório r é gerado, escolhemos um dos valores $\{1, 2, \text{ou } 3\}$ com base no segmento em que r se encontra. Ou seja, nossa escolha será

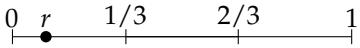


FIGURA 1.6: O número aleatório r foi gerado no intervalo $[0, 1/3)$. Como o PRNG usado fornece números distribuídos uniformemente no intervalo $[0, 1)$, a probabilidade disso ocorrer é de $1/3$. O mesmo vale para os outros dois intervalos mostrados, $[1/3, 2/3)$ e $[2/3, 1)$.

$$\text{escolha}(r) = \begin{cases} 1, & \text{se } r < 1/3 \\ 2, & \text{se } 1/3 \leqslant r < 2/3 \\ 3, & \text{se } r \geqslant 2/3 \end{cases} \tag{1.3.2}$$

De posse de um PRNG, não é muito difícil ensinar ao computador nossa receita do triângulo da Figura 1.5. A aplicação desse algoritmo, com 100 000 iterações, é mostrada na Figura 1.11 na página 13, colocada propositalmente ao final do capítulo. Surpresa! O que obtivemos? O Triângulo de Sierpiński! Mas, se na versão impressa, ou com uma pequena ampliação da versão digital, as Figuras 1.3 e 1.11 são muito parecidas, tente olhar mais de perto: os pontos do Triângulo de Sierpiński da Figura 1.11 tem um certo grau de aleatoriedade, diferente da construção regular da Figura 1.3. Então, volte agora à Figura 1.5 e tente entender por que o algoritmo descreve o Triângulo de Sierpiński. Não por acaso, a figura obtida é um fractal, pois o padrão da Figura 1.5b é repetido em todas as escalas (pelo menos dentro da resolução da Figura 1.11). Mas ela é chamada também de *atrator*, pois a órbita de P, que poderia ter até mesmo começado em algum lugar fora do triângulo ABC, acabará sendo atraída para ele e nele permanecerá em definitivo. A constante repetição do algoritmo garante a autossimilaridade do atrator, o que o torna um fractal.

Muitos outros fractais podem ser criados por esse algoritmo. Você encontra alguns deles na Figura 1.7. Por exemplo, o primeiro fractal mostrado na Figura 1.7, conhecido como *Dragão de Levy* [5], é gerado do seguinte modo: escolha um ponto inicial $P = (x, y)$ qualquer e determine sua órbita $P' = (x', y')$ de acordo com a seguinte receita:

$$T_1: \quad \begin{cases} x' = (x - y)/2 \\ y' = (x + y)/2 \end{cases}, \quad \text{prob.} = 50\% \tag{1.3.3a}$$

$$T_2: \quad \begin{cases} x' = (x + y + 1)/2 \\ y' = (-x + y + 1)/2 \end{cases}, \quad \text{prob.} = 50\% \tag{1.3.3b}$$

O material de apoio às presentes notas de aula contém as fórmulas para gerar todos os fractais da Figura 1.7. Fique à vontade para alterar os valores e criar seu próprio fractal! Mas, se quiser entender um pouco mais a fundo como essas fórmulas funcionam, precisamos falar de mais alguns conceitos.

1.3.1 Transformações afins no plano

Considere a Eq. (1.3.3)b, que é uma das fórmulas para se obter o dragão de Levy, como acabamos de ver. Veja que essa expressão também pode ser colocada em forma matricial:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 1/2 & 1/2 \\ -1/2 & 1/2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix} \tag{1.3.4}$$

Relações como essa, do tipo

$$X' = T_l X + T\,, \tag{1.3.5}$$

em que

$$X' = \begin{pmatrix} x' \\ y' \end{pmatrix}, \qquad X = \begin{pmatrix} x \\ y \end{pmatrix}, \qquad T_l = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \qquad T = \begin{pmatrix} t_x \\ t_y \end{pmatrix}, \tag{1.3.6}$$

são chamadas de *transformações afins no plano* (a, b, c, d, t_x, t_y são números reais quaisquer). O que uma transformação dessas faz é aplicar primeiramente sobre X uma *transformação linear* T_l, resultando em novas coor-

FIGURA 1.7: Exemplos de fractais gerados acompanhando a órbita de um ponto inicial qualquer por diversos algoritmos aleatórios. Um pouco de arte foi usada para colorir algumas das imagens.

denadas X_l dadas por

$$X_l = T_l X\,, \tag{1.3.7}$$

seguida de uma translação T das coordenadas X_l, resultando nas coordenadas X' da transformação afim:

$$X' = X_l + T \tag{1.3.8}$$

Uma maneira diferente de se representar transformações afins é usar uma *matriz ampliada* 3x3 da seguinte maneira:

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} a & b & t_x \\ c & d & t_y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \tag{1.3.9}$$

ou, mais compactamente,

$$X'_a = T_a X_a \tag{1.3.10}$$

onde está claro o significado dos vetores 3x1 X'_a e X_a, e da matriz ampliada 3x3 T_a, que representa a transformação afim. Ampliar as matrizes será útil para entender a aplicação de múltiplas transformações, como veremos mais abaixo.

Vamos agora tentar entender os efeitos de uma transformação afim sobre as coordenadas de um ponto a partir de algumas matrizes T_a simples, chamadas às vezes de transformações afins básicas. Para fins de

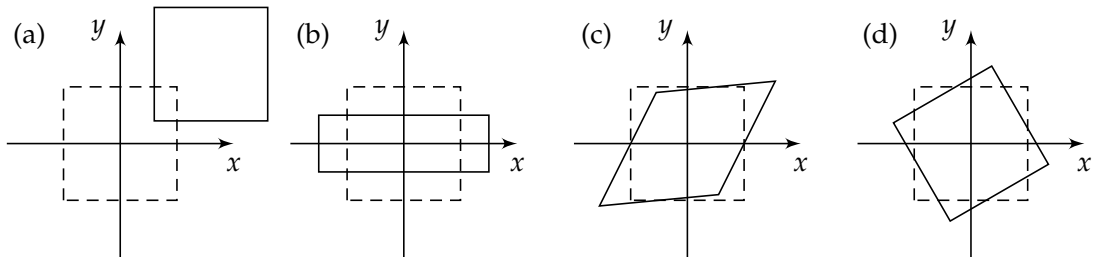


FIGURA 1.8: Ilustrando o efeito das transformações afins básicas de (a) translação T_t, (b) expansão/contração T_e, (c) cizalhamento T_c e (d) rotação T_r sobre o quadrado representado por linhas tracejadas.

visualização, é mais intrutivo aplicar a transformação a todo um conjunto de pontos, ao invés de a apenas um ponto isolado. Vamos então descrever os efeitos das transformações básicas sobre um quadrado de lado unitário. Acompanhe visualmente, na Figura 1.8, a descrição algébrica feita abaixo.

Translação: Uma translação pura significa apenas movimentar os pontos na direção do vetor de translação (t_x, t_y) a uma distância dada pelo módulo desse vetor. Na notação de matrizes ampliadas, a translação pura é dada por

$$T_t = \begin{pmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{pmatrix} \tag{1.3.11}$$

Expansão/contração: Nesse caso, a separação entre dois pontos originais é aumentada/diminuída de fatores s_x e s_y nas direções x e y, respectivamente. Não há outra forma de deformação além dessa ampliação. A matriz da transformação é dada por

$$T_e = \begin{pmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{1.3.12}$$

Repare que, se $0 < s_x < 1$, há uma contração na direção x, ao invés de uma expansão. Por outro lado, caso $s_x < 0$, ocorrerá, além da expansão/contração, também uma *reflexão* em relação ao eixo x. As mesmas observações valem para s_y em relação à direção y.

Cizalhamento: Diferentemente da expansão, uma transformação de cizalhamento deforma ângulos retos cujos lados sejam paralelos às direções dos eixos cartesianos. A matriz ampliada nesse caso é dada por

$$T_c = \begin{pmatrix} 1 & c_x & 0 \\ c_y & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{1.3.13}$$

Veja que os sinais de c_x e c_y apenas alteram o sentido do cizalhamento. Uma transformação cizalhante, apesar de alterar ângulos, tem o efeito de preservar paralelas linhas que já o eram no objeto original.

Rotação: Uma transformação de rotação não deforma o objeto original. Ele é apenas reorientado em relação aos eixos cartesianos, mas mantendo inalterados distâncias e ângulos relativos após a transformação. Uma rotação de um ângulo θ no sentido anti-horário é dada por

$$T_r = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta & 0 \\ \operatorname{sen}\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{1.3.14}$$

É importante observar que tal rotação é sempre feita em torno da origem do sistema de coordenadas.

1.3.2 Sequência de transformações afins

Considere a transformação afim T_a mostrada na Figura 1.9, que leva os pontos do quadrado C ao quadrilátero C'. Como qualquer transformação afim, ela é representada por uma matriz como na Eq. (1.3.9). No caso particular da Figura 1.9, a matriz da transformação T_a é

$$T_a = \begin{pmatrix} 0.91923 & -0.168135 & 1.4 \\ 0.807846 & 0.711218 & 1 \\ 0 & 0 & 1 \end{pmatrix} \tag{1.3.15}$$

Mas na prática, a maneira como criei C' a partir de C foi um pouco diferente. Na verdade, apliquei a seguinte sequência de transformações básicas:

1. ampliação de 20% em x e redução em 30% em y:

$$T_e = \begin{pmatrix} 1.2 & 1 & 0 \\ 0 & 0.7 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{1.3.16}$$

2. cizalhamento de 0.3 em x e 0.2 em y:

$$T_c = \begin{pmatrix} 1 & 0.3 & 0 \\ 0.2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{1.3.17}$$

3. rotação de 30° em torno da origem, no sentido anti-horário:

$$T_r = \begin{pmatrix} 0.5 & -0.86602 & 0 \\ 0.86602 & 0.5 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{1.3.18}$$

4. translação de um vetor $\vec{t} = (1.4, 1)$:

$$T_t = \begin{pmatrix} 1 & 0 & 1.4 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \tag{1.3.19}$$

É importante não esquecer que as transformações foram realizadas exatamente na ordem mostrada acima. Inverter a ordem do cizalhamento e da rotação, por exemplo, gera uma transformação completamente diferente! Você pode se convencer disso com exemplos simples, que você mesmo pode criar; mas também pode usar o material de apoio para testar no presente caso. Pois bem: a primeira transformação, T_e (ampliação/redução), aplicada a qualquer ponto P, gera um ponto P' dado por

$$P' = T_e P\,. \tag{1.3.20}$$

A seguir, a segunda transformação (T_c) opera sobre P', levando a P'' dado por

$$P'' = T_c P' = T_c T_e P\,. \tag{1.3.21}$$

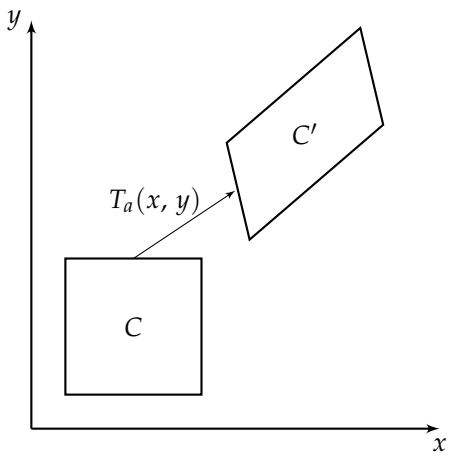


FIGURA 1.9: A transformação afim T_a leva os pontos do quadrado C para o quadrilátero C'.

E agora espero que você tenha entendido o espírito de como criar uma sequência de transformações: aplique-as consecutivamente a P, multiplicando-as pela esquerda às matrizes anteriores. O ponto final de C', após as quatro transformações listadas acima, será um ponto P^{iv} dado por

$$P^{iv} = T_a P = T_t\, T_r\, T_c\, T_e P\,, \tag{1.3.22}$$

o que implica em

$$T_a = T_t\, T_r\, T_c\, T_e\,. \tag{1.3.23}$$

Você pode verificar que a matriz T_a na Eq. (1.3.15) realmente é dada pelo produto das quatro matrizes das transformações básicas listadas acima, exatamente na ordem dada pela Eq. (1.3.23).

Fica agora claro, também algebricamente, por que diferentes sequências de transformações afins geram resultados distintos: T_a é uma sequência de matrizes multiplicadas entre si — mas a multiplicação de matrizes não tem a propriedade comutativa! Ou seja, dadas duas matrizes A e B quaisquer, em geral $A\,B \neq B\,A$, a não ser em casos particulares. Assim, inverter a ordem de duas transformações significa alterar a ordem em que as respectivas matrizes são multiplicadas, o que, nos diz a Álgebra Linear, fornecerá, via de regra, um resultado diferente.

1.4 O segredo da árvore fractal

Um bom mágico não revela o segredo dos seus truques — a não ser que os lance em livro! Este é um livro de métodos computacionais, e não de ilusionismo ou prestidigitação, mas mesmo assim vou revelar o segredo da árvore da Figura 1.1, já que agora, com a discussão das seções anteriores, temos amplas condições de entender como um fractal complexo como a árvore da Figura 1.1 é projetado. Ela é gerada pela combinação aleatória das seis transformações afins, com as respectivas probabilidades, mostradas no Quadro 1.1.

O efeito das transformações mostradas no Quadro 1.1 é visto na Figura 1.10. Cada uma das seis transformações age (por exemplo) sobre o retângulo em linhas tracejadas na Figura 1.10, levando-o a um dos seis quadriláteros coloridos, onde já podemos perceber o "esqueleto" da árvore fractal. Usando o conceito de sequência de transformação que vimos na última seção, é relativamente fácil determinar as matrizes de cada uma das seis transformações usadas — mas há um grau de subjetividade na criação da árvore: um quesito estético também faz parte do projeto da árvore. As probabilidades, também, são determinadas de maneira subjetiva, de forma a escurecer ou clarear partes do fractal (ou seja, aumentar ou diminuir a densidade de pontos).

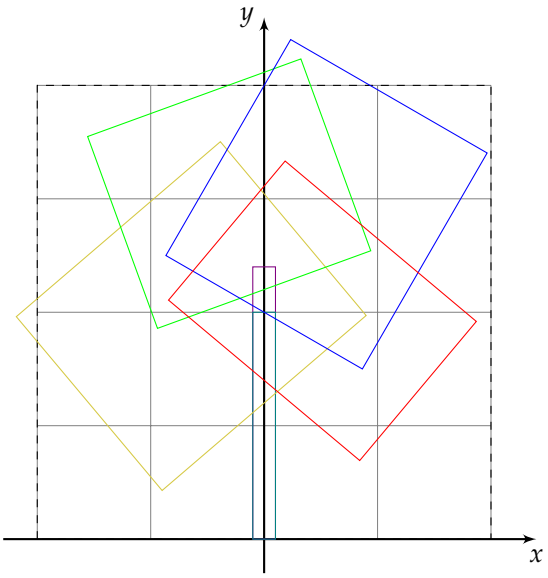


FIGURA 1.10: "Esqueleto" da árvore fractal da Figura 1.1, com as transformações afins do Quadro 1.1 aplicadas ao quadrado mostrado em linhas tracejadas.

A constante iteração faz com que, dentro de cada um dos quadriláteros da Figura 1.10, haja uma cópia em miniatura desse esqueleto, criando a autossimilaridade do atrator em formato de árvore. Vale lembrar que todos os fractais mostrados na Figura 1.7 são gerados pelo mesmo princípio. Como já mencionamos, você encontra as transformações (e suas respectivas probabilidades) no material de apoio.

Para chegar ao nível de detalhamento visto na Figura 1.1, foi utilizado um quadro de imagem de 4000×4800 pixels (da Wikipédia: *"um pixel é o menor ponto que forma uma imagem digital, sendo que um conjunto de pixels com várias cores formam a imagem inteira."*) e quarenta milhões de iterações. Usando um antiga (e eficiente) biblioteca em C++ dos primórdios do Linux (`libplot`, parte do pacote `plotutils`, https://www.gnu.org/software/plotutils) num *laptop* moderno (em 2018), a implementação leva em torno de 20 s para gerar um arquivo `png` com a imagem final. O código está no material de apoio.

$$T_1 = \begin{pmatrix} 0.05 & 0 & 0 \\ 0 & 0.6 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad p_1 = 10\%\,, \qquad T_2 = \begin{pmatrix} 0.05 & 0 & 0 \\ 0 & -0.5 & 1 \\ 0 & 0 & 1 \end{pmatrix} \quad p_2 = 10\%\,,$$

$$T_3 = \begin{pmatrix} 0.46 & -0.32 & 0 \\ 0.39 & 0.38 & 0.6 \\ 0 & 0 & 1 \end{pmatrix} \quad p_3 = 20\%\,, \quad T_4 = \begin{pmatrix} 0.47 & -0.154 & 0 \\ 0.171 & 0.423 & 1.1 \\ 0 & 0 & 1 \end{pmatrix} \quad p_4 = 20\%\,,$$

$$T_5 = \begin{pmatrix} 0.433 & 0.275 & 0 \\ -0.25 & 0.476 & 1 \\ 0 & 0 & 1 \end{pmatrix} \quad p_5 = 20\%\,, \quad T_6 = \begin{pmatrix} 0.421 & 0.257 & 0 \\ -0.353 & 0.306 & 0.7 \\ 0 & 0 & 1 \end{pmatrix} \quad p_6 = 20\%\,.$$

QUADRO 1.1: Transformações e probabilidades de escolha para gerar a árvore fractal da Figura 1.1.

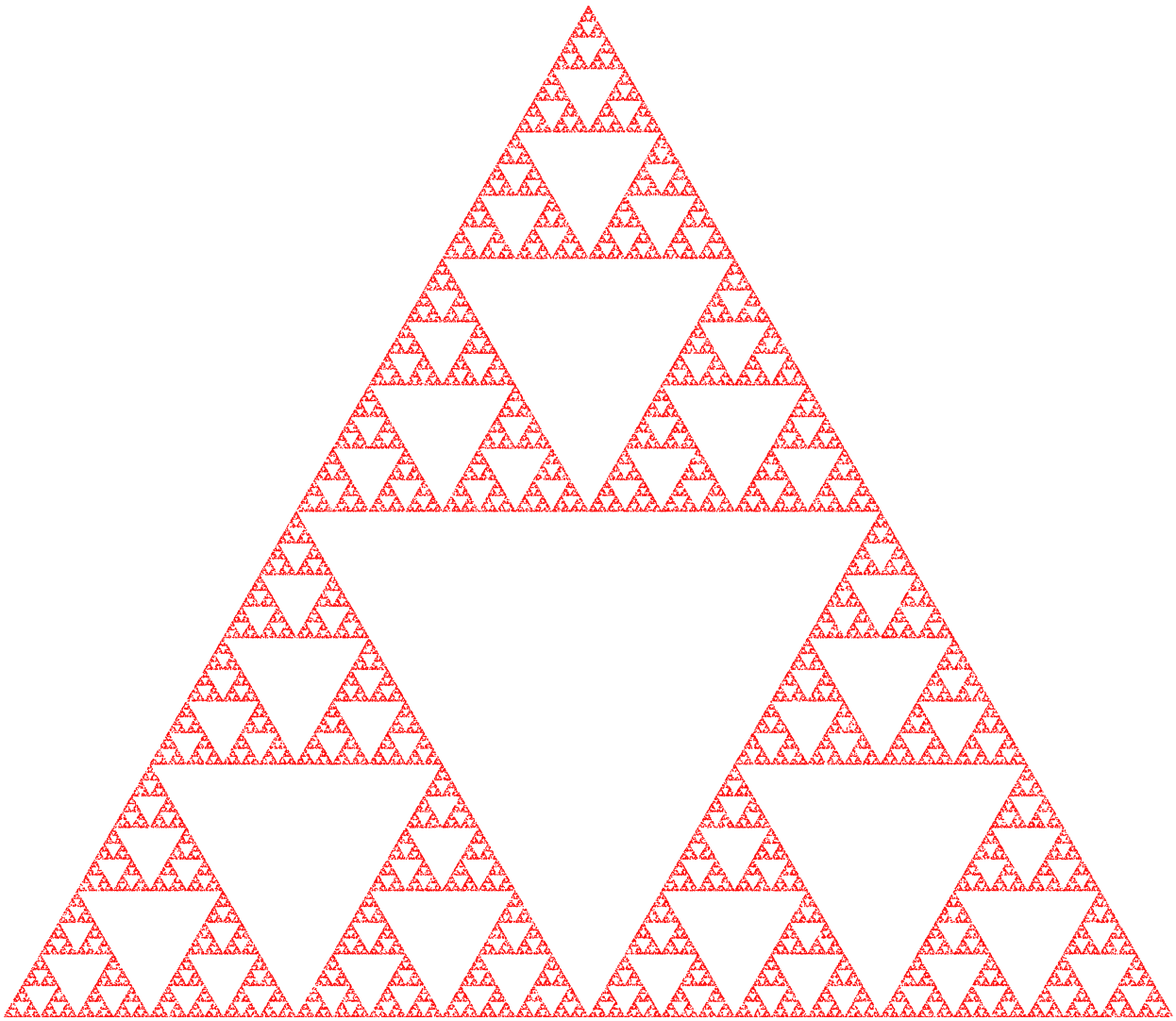


FIGURA 1.11: Resultado do algoritmo descrito no começo da seção 1.3. Surpresa: é o Triângulo de Sierpiński!

Referências

[1] K. Libbrecht e P. Rasmussen. *The Snowflake — Winter's secret beauty*. Stillwater (MN), EUA: Voyageur Press, 2004.

[2] R. G. Hohlfeld e N. Cohen. "Self-similarity and the geometric requirements for frequency independence in antennae". *Fractals* 7 (1999), 79–84.

[3] B. B. Mandelbrot. *The Fractal Geometry of Nature*. San Francisco: W. H. Freeman, 1983.

[4] N. Fiedler-Ferrara e C. P. C. do Prado. *Caos – uma introdução*. São Paulo: Edgard Blücher, 1994.

[5] S. Bailey, T. Kim e R. S. Strichartz. "Inside the Lévy Dragon". *The American Mathematical Monthly* 109 (2002), 689–703.

 © Luiz T. F. Eleno, v. 0.3, 2021

Capítulo 2

Modos normais

A carreira de um jovem físico teórico consiste em tratar o oscilador harmônico em níveis crescentes de abstração.

Sidney Coleman

2.1 O oscilador harmônico simples por outro ângulo

Você conhece o oscilador harmônico desde o ensino médio, pelo menos. Mesmo sabendo que esse é um problema antigo, vamos descrevê-lo tomando como base a Figura 2.1, apenas para introduzir a notação do restante do capítulo. Vemos, então, uma partícula pontual de massa m presa a uma mola de constante elástica k e comprimento natural (ou de equilíbrio) a. Quando a massa se desloca de x em relação à posição de equilíbrio, a mola impõe sobre a partícula uma força restauradora proporcional ao deslocamento, ou seja, $F = -kx$. A 2ª Lei de Newton nos diz que a dinâmica da partícula é regida pela seguinte equação diferencial ordinária (EDO):

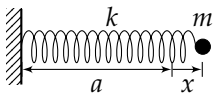


FIGURA 2.1: Modelo esquemático de um oscilador harmônico simples.

$$m\ddot{x} = -kx\,, \tag{2.1.1}$$

onde $\ddot{x}(t) = d^2x/dt^2$ é a aceleração da partícula. A Eq. (2.1.1) pode ser levemente reescrita como

$$\ddot{x} = -\omega_0^2 x\,, \tag{2.1.2}$$

onde introduzimos a chamada *frequência natural* ω_0 do oscilador,

$$\omega_0^2 = \frac{k}{m}\,. \tag{2.1.3}$$

Talvez o que veremos partir desse ponto seja novo para você. O que faremos é basicamente transformar a equação Eq. (2.1.2), que é uma EDO de segunda ordem, num sistema de EDOs de primeira ordem, que sabemos resolver facilmente. Vamos então escrever as duas equações a seguir:

$$\dot{x} = v \tag{2.1.4a}$$

$$\dot{v} = -\omega_0^2 x \tag{2.1.4b}$$

Mesmo aqui ainda não há nada de novo: $v(t) = \dot{x}(t) = dx/dt$ é a velocidade da partícula, cuja derivada temporal $\dot{v}(t)$ é a aceleração. A novidade vem agora: vamos escrever as Eqs. (2.1.4) em forma matricial,

$$\begin{pmatrix} \dot{x} \\ \dot{v} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -\omega_0^2 & 0 \end{pmatrix} \begin{pmatrix} x \\ v \end{pmatrix} \tag{2.1.5}$$

Veja que podemos identificar os vetores coluna que aparecem na Eq. (2.1.5) como um vetor $Y(t)$:

$$Y(t) = \begin{pmatrix} x \\ v \end{pmatrix} \quad \Rightarrow \quad \dot{Y}(t) = \begin{pmatrix} \dot{x} \\ \dot{v} \end{pmatrix} \tag{2.1.6}$$

Com isso, a Eq. (2.1.5) é escrita compactamente como

$$\dot{Y}(t) = WY(t)\,, \tag{2.1.7}$$

em que W é a matriz

$$W = \begin{pmatrix} 0 & 1 \\ -\omega_0^2 & 0 \end{pmatrix}. \tag{2.1.8}$$

Vamos agora recuar um instante na solução do problema original e relembrar a solução de uma EDO mais simples:

$$\dot{y}(t) = \omega y(t)\,, \tag{2.1.9}$$

em que ω é uma constante e $y(t)$ é uma função real de t. Estamos carecas de saber que essa equação é facilmente resolvida:

$$\dot{y}(t) = \omega y(t) \quad \Rightarrow \quad y(t) = Ae^{\omega t}\,. \tag{2.1.10}$$

Mas, comparando as Eqs. (2.1.7) e (2.1.9), vemos que elas têm a mesma cara. Vamos então buscar por uma solução para $Y(t)$ na forma

$$Y(t) = Ae^{\omega t} \tag{2.1.11}$$

em que ω é uma constante a ser determinada posteriormente e A agora é um vetor constante:

$$A = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \tag{2.1.12}$$

também a ser determinado posteriormente. Derivando a Eq. (2.1.11) em relação ao tempo t teremos $\dot{Y}(t) = \omega Ae^{\omega t}$. Igualando essa derivada à Eq. (2.1.7), podemos escrever

$$WAe^{\omega t} = \omega Ae^{\omega t} \tag{2.1.13}$$

o que nos leva a concluir que

$$WA = \omega A\,. \tag{2.1.14}$$

Cuidado! Não podemos cancelar A, que aparece nos dois lados da equação acima, porque, na verdade, trata-se de uma equação matricial. De fato, sabemos das Eqs. (2.1.8) e (2.1.12) que W e A são uma matriz 2×2 e um vetor coluna 2×1, respectivamente, e portanto a Eq. (2.1.14) é, na verdade, a seguinte equação matricial:

$$\begin{pmatrix} 0 & 1 \\ -\omega_0^2 & 0 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \omega \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \tag{2.1.15}$$

O formato da equação Eq. (2.1.14), no entanto, é um velho conhecido das aulas de Álgebra Linear (e que você sempre se perguntou para que serviria): é um problema de autovalores e autovetores, em que, dada uma matriz quadrada W, procuramos seus autovalores ω e os autovetores A correspondentes. Reescrevemos facilmente a Eq. (2.1.15) como

$$\begin{pmatrix} -\omega & 1 \\ -\omega_0^2 & -\omega \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{2.1.16}$$

que é um *sistema linear homogêneo* em a_1 e a_2 que só admite solução se

$$\det \begin{pmatrix} -\omega & 1 \\ -\omega_0^2 & -\omega \end{pmatrix} = 0 \tag{2.1.17}$$

O determinante nulo acima fornece a chamada *equação característica* da EDO do oscilador harmônico,

$$\omega^2 + \omega_0^2 = 0 \tag{2.1.18}$$

de onde extraímos os valores possíveis de ω (ou seja, os autovalores):

$$\omega = \pm i\omega_0 \tag{2.1.19}$$

Os autovetores correspondentes aos dois autovalores acima são também determinados facilmente:

$$\omega_\alpha = +i\omega_0 \quad \Rightarrow \quad A_\alpha = c_1 \begin{pmatrix} 1 \\ i\omega_0 \end{pmatrix} \quad \Rightarrow \quad Y_\alpha(t) = A_\alpha e^{i\omega_0 t} \tag{2.1.20a}$$

$$\omega_\beta = +i\omega_0 \quad \Rightarrow \quad A_\beta = c_2 \begin{pmatrix} 1 \\ -i\omega_0 \end{pmatrix} \quad \Rightarrow \quad Y_\beta(t) = A_\beta e^{-i\omega_0 t} \tag{2.1.20b}$$

em que c_1 e c_2 são constantes (que podem, sem perda de generalidade, ser igualadas a um) e usamos α e β apenas como identificadores dos dois valores possíveis para ω, A e também para Y, usando a Eq. (2.1.11).

Lembrando ainda das aulas de Álgebra Linear: autovalores distintos geram autovetores linearmente independentes e, se o número de autovalores distintos é igual à dimensão do espaço (ou seja, da matriz), os autovetores formam uma base completa. Em outras palavras, a solução geral da Eq. (2.1.14) é qualquer combinação linear das soluções obtidas com os autovalores e autovetores listados nas Eqs. (2.1.20):

$$Y(t) = c_\alpha Y_\alpha(t) + c_\beta Y_\beta(t) \tag{2.1.21}$$

ou, de forma explícita:

$$Y(t) = \begin{pmatrix} x(t) \\ \dot{x}(t) \end{pmatrix} = c_\alpha \begin{pmatrix} 1 \\ i\omega_0 \end{pmatrix} e^{i\omega_0 t} + c_\beta \begin{pmatrix} 1 \\ -i\omega_0 \end{pmatrix} e^{-i\omega_0 t} \tag{2.1.22}$$

Veja que obtemos de uma só tacada tanto a posição quanto a velocidade da partícula, o que configura uma das vantagens do método matricial. Na Eq. (2.1.22), c_α e c_β são constantes a serem determinadas a partir das condições iniciais

$$Y(0) = \begin{pmatrix} x(0) \\ v(0) \end{pmatrix} = \begin{pmatrix} x_0 \\ v_0 \end{pmatrix} . \tag{2.1.23}$$

Talvez você ainda se sinta um pouco desconfortável com tantas exponenciais complexas aparecendo num problema tão simples (e real!) quanto um oscilador harmônico. Mas a solução é facilmente reescrita em termos de senos e cossenos usando a identidade de Euler, encontrada no Apêndice B,

$$e^{\pm i\theta} = \cos\theta \pm i \operatorname{sen}\theta \tag{2.1.24}$$

que transforma, por exemplo, a posição

$$x(t) = c_\alpha e^{i\omega_0 t} + c_\beta e^{-i\omega_0 t} \tag{2.1.25}$$

numa forma mais amigável para quem não gosta de números complexos aparecendo nas equações,

$$x(t) = \alpha_1 \cos\omega_0 t + \beta_1 \operatorname{sen}\omega_0 t\,, \tag{2.1.26}$$

ainda que eles continuem escondidos nas novas constantes α_1 e β_1:

$$\alpha_1 = c_\alpha + c_\beta\,; \qquad \beta_1 = i(c_\alpha - c_\beta)\,. \tag{2.1.27}$$

Podemos transformar ainda mais a equação usando as identidades trigonométricas de prostaferese, também extraídas do Apêndice B:

$$\cos(a \pm b) = \cos a \cos b \mp \operatorname{sen} a \operatorname{sen} b \tag{2.1.28}$$

$$\operatorname{sen}(a \pm b) = \operatorname{sen} a \cos b \pm \cos a \operatorname{sen} b \tag{2.1.29}$$

o que nos leva a um novo formato equivalente para $x(t)$:

$$x(t) = r\cos(\omega_0 t + \phi) \tag{2.1.30}$$

sendo que r e ϕ, chamadas, respectivamente, de amplitude e fase do movimento periódico, são dadas por

$$r = \sqrt{\alpha_1^2 + \beta_1^2}\,; \qquad \phi = -\operatorname{arctg}(\beta_1/\alpha_1)\,. \tag{2.1.31}$$

e podem também ser determinadas diretamente das condições iniciais.

Vamos recapitular: o que fizemos até aqui foi, essencialmente, converter uma EDO de segunda ordem, a Eq. (2.1.2), numa EDO de primeira ordem, a Eq. (2.1.7), ainda que, para isso, tenhamos apelado para matrizes. É importante dizer que esse método, apesar de parecer, para o oscilador harmônico, um tiro de canhão para matar uma mosca, é fundamental para a solução computacional (ou seja, numérica) de EDOs

de ordem superior, já que os computadores são excelentes para lidar com matrizes!

Veremos a seguir uma série de casos em que o uso de matrizes nos ajuda a encontrar a solução de problemas muito mais cabeludos que o oscilador harmônico simples. Vamos então por a mão na massa e começar com o desenvolvimento algébrico dos modelos de osciladores para, ao final, conseguirmos entender a motivação por detrás do método numérico universal do capítulo 8, capaz de resolver sistemas de EDOs de qualquer ordem.

Na próxima seção, veremos a primeira generalização do oscilador harmônico, comumente encontrada nos cursos de Física Básica, o oscilador harmônico amortecido, e como podemos reformular matematicamente o problema usando matrizes. A não ser por essa pequena variação, não há muita coisa nova nessa abordagem alternativa, mas ela talvez sirva como uma pequena revisão sobre o assunto.

2.2 O oscilador harmônico amortecido

Vamos continuar considerando uma partícula de massa m presa a uma mola de constante k, como na Figura 2.2. No entanto, o oscilador está também vinculado a um sistema de amortecimento, que impõe uma força dissipativa proporcional à velocidade da partícula, $F = -\epsilon\dot{x}$, sendo ϵ uma medida da viscosidade do meio em que a partícula está inserida. Aplicando novamente a 2ª Lei de Newton, escrevemos a EDO para o oscilador harmônico amortecido:

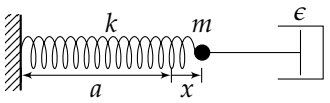


FIGURA 2.2: Um oscilador harmônico amortecido por uma força viscosa.

$$m\ddot{x} = -kx - \epsilon\dot{x} \tag{2.2.1}$$

que pode ser um pouco mais simplificada para

$$\ddot{x} = -\omega_0^2 x - 2\eta\dot{x} \tag{2.2.2}$$

se definirmos

$$2\eta = \frac{\epsilon}{m}, \tag{2.2.3}$$

com o mesmo ω_0 definido na Eq. (2.1.3). O fator 2 nas Eqs. (2.2.2) e (2.2.3) foi utilizado apenas por conveniência, para simplificar as próximas equações. Vamos considerar que k e ϵ, e portanto ω_0 e η, são constantes positivas.

Introduzindo agora a velocidade $v(t) = \dot{x}(t)$, como fizemos anteriormente, convertemos a EDO de segunda ordem num sistema de EDOs de primeira ordem:

$$\begin{pmatrix} \dot{x} \\ \dot{v} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -\omega_0^2 & -2\eta \end{pmatrix} \begin{pmatrix} x \\ v \end{pmatrix} \tag{2.2.4}$$

ou, mais compactamente, como

$$\dot{Y} = WY \tag{2.2.5}$$

que é muito parecida com a Eq. (2.1.7), mas com uma matriz W diferente:

$$W = \begin{pmatrix} 0 & 1 \\ -\omega_0^2 & -2\eta \end{pmatrix} \tag{2.2.6}$$

Já que os problemas são tão parecidos matematicamente, podemos adotar a mesma solução teste,

$$Y(t) = Ae^{\omega t}, \tag{2.2.7}$$

e procurar por ω e A que, sabemos, são os autovalores e autovetores, respectivamente, da matriz W, oriundos da equação abaixo:

$$WA = \omega A \tag{2.2.8}$$

A equação característica nesse caso é dada por

$$\det \begin{pmatrix} -\omega & 1 \\ -\omega_0^2 & -2\eta - \omega \end{pmatrix} = 0 \tag{2.2.9}$$

ou seja, calculando o determinante e simplificando:

$$\omega^2 + 2\eta\omega + \omega_0^2 = 0 \tag{2.2.10}$$

que fornece os seguintes autovalores ω:

$$\omega = -\eta \pm \sqrt{\eta^2 - \omega_0^2} \tag{2.2.11}$$

Vamos investigar caso a caso, em função do sinal da expressão debaixo do radical na Eq. (2.2.11). Vamos inicialmente criar uma variável auxiliar, δ, definida através da expressão

$$\delta^2 = \left|\eta^2 - \omega_0^2\right| . \tag{2.2.12}$$

Vejamos agora as três situações que podem ocorrer.

1. *amortecimento supercrítico:* $\eta^2 - \omega_0^2 > 0$

   Nesse caso, as duas soluções são reais:

   $$\omega = -\eta \pm \delta \tag{2.2.13}$$

   e, portanto, podemos escrever $x(t)$, por exemplo, como uma combinação linear das duas soluções possíveis:

   $$x(t) = c_\alpha e^{(-\eta+\delta)t} + c_\beta e^{(-\eta-\delta)t} \tag{2.2.14}$$

   ou, simplificando,

   $$x(t) = e^{-\eta t}\left(c_\alpha e^{\delta t} + c_\beta e^{-\delta t}\right) . \tag{2.2.15}$$

   onde as constantes c_α e c_β vêm das condições iniciais. Veja que queimamos a etapa de encontrar os autovetores, uma vez que eles apenas forneceriam automaticamente a velocidade além da posição. Mas podemos obter facilmente a velocidade derivando a Eq. (2.2.15) em relação ao tempo. Por outro lado, nos exemplos das próximas seções, será imprescindível determinar explicitamente os autovetores, ainda que não mais forneçam a velocidade, como veremos.

2. *amortecimento subcrítico:* $\eta^2 - \omega_0^2 < 0$

   Nessa situação, os autovalores são números complexos:

   $$\omega = -\eta \pm i\delta \tag{2.2.16}$$

   e temos portanto uma combinação linear de exponenciais complexas:

   $$x(t) = e^{-\eta t}\left(c_\alpha e^{i\delta t} + c_\beta e^{-i\delta t}\right) \tag{2.2.17}$$

   Usando a relação de Euler, já encontrada (Eq. 2.1.24), essa equação pode ser reescrita como

   $$x(t) = e^{-\eta t}\left(c_\alpha \cos\delta t + c_\beta \operatorname{sen}\delta t\right) , \tag{2.2.18}$$

   em que c_α e c_β não são, naturalmente, os mesmos encontrados na Eq. (2.2.17), ou ainda na forma mais compacta

   $$x(t) = a_0 e^{-\eta t}\cos(\delta t + \phi) . \tag{2.2.19}$$

   Os valores das constantes a_0 e ϕ (ou c_α e c_β) são obtidos a partir das condições iniciais.

3. *amortecimento crítico:* $\eta^2 - \omega_0^2 = 0$

   Esse é um caso peculiar, uma vez que as duas soluções para ω, encontradas nos casos anteriores, se tornam uma única solução degenerada:

   $$\omega = -\eta \tag{2.2.20}$$

   Portanto, a Eq. (2.2.7) não fornece mais duas soluções linearmente independentes, e sim uma única solução,

   $$x(t) = c_\alpha e^{-\eta t} . \tag{2.2.21}$$

   Podemos adotar o método de variação das constantes para tentar encontrar uma outra solução independente para o problema original — ou podemos consultar algum livro de Física Básica, que nos dirá que uma outra solução possível é

   $$Y(t) = c_\beta t e^{-\eta t} \tag{2.2.22}$$

   Com isso, a solução geral nesse caso é

   $$x(t) = c_\alpha e^{-\eta t} + c_\beta t e^{-\eta t} = e^{-\eta t}(c_\alpha + c_\beta t) . \tag{2.2.23}$$

© Luiz T. F. Eleno, v. 0.3, 2021

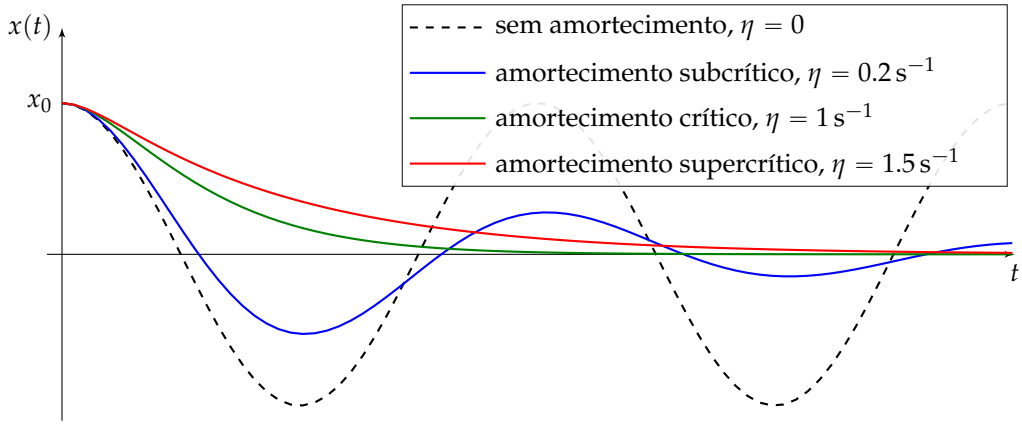
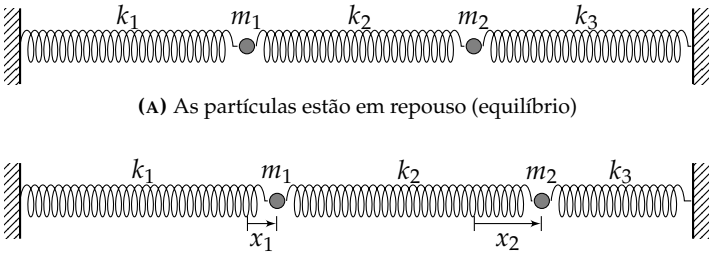


FIGURA 2.3: Exemplos de osciladores harmônicos de mesma frequência natural $\omega_0 = 1\,\mathrm{s}^{-1}$ e diferentes condições de amortecimento para as mesmas condições iniciais $x(0) = x_0$ e $\dot{x}(0) = 0$.

A Figura 2.3 mostra exemplos de osciladores harmônicos, de mesma frequência natural $\omega_0 = 1\,\mathrm{s}^{-1}$ e mesmas condições iniciais $x(0) = x_0$ e $\dot{x}(0) = 0$, mas com diferentes condições de amortecimento, cada uma delas correspondendo a um dos casos listados acima. Vemos que, nos três casos em que $\eta \neq 0$, o oscilador perde progressivamente energia até atingir o repouso (para $t \to \infty$). No entanto, o oscilador com amortecimento subcrítico continua oscilando ao redor da posição de equilíbrio ($x = 0$), ainda que com amplitude cada vez menor. Os outros dois osciladores, por outro lado, caminham monotonicamente para a posição de equilíbrio. O oscilador com amortecimento crítico é o limite entre os outros dois casos, além de ser o oscilador que tende de maneira mais rápida à posição de equilíbrio sem oscilar ao seu redor.

2.3 Dois osciladores

Agora vamos complicar um pouco o problema e acoplar dois osciladores harmônicos simples, como na Figura 2.4. Continuaremos a lidar com uma situação unidimensional se pensarmos apenas na vibração longitudinal das duas partículas de massa m_1 e m_2. Quando, como na Figura 2.4b, as partículas se deslocam das posições de equilíbrio de x_1 e x_2, respectivamente, elas sentem forças das molas de constantes k_1, k_2 e k_3. Por suas vez, as molas das extremidades estão presas a apoios fixos.



(A) As partículas estão em repouso (equilíbrio)

(B) As partículas se deslocaram de x_1 e x_2

FIGURA 2.4: Dois osciladores acoplados. (a), as partículas de massas m_1 e m_2 estão em repouso (equilíbrio); em (b), as partículas se deslocaram de x_1 e x_2 em relação à posição de equilíbrio, o que faz com que sintam forças restauradoras das molas de constantes elásticas k_1, k_2 e k_3.

A maneira mais fácil de calcular as forças é isolar as partículas e determinar seus deslocamentos relativos à posição original da Figura 2.4a. Por exemplo, a mola de constante k_1 sofre uma extensão x_1 pelo deslocamento da primeira partícula de x_1. Como a partícula 2 se desloca de x_2, a mola k_2 sofre uma extensão dada por $x_2 - x_1$, ou seja, a partícula 2 "estica" a mola 2, ao passo que a primeira partícula a faz "encolher". Coisa parecida se passa com a terceira e última mola. Não é assim tão difícil de enxergar que as leis de Newton para as duas partículas fornecem o seguinte sistema de EDOs:

$$m_1\ddot{x}_1 = -k_1x_1 + k_2(x_2 - x_1) \tag{2.3.1a}$$
$$m_2\ddot{x}_2 = -k_2(x_2 - x_1) - k_3x_2 \tag{2.3.1b}$$

que reescrevemos de forma mais simpática como

$$\ddot{x}_1 = -\Omega_{11}x_1 + \Omega_{12}x_2 \tag{2.3.2a}$$
$$\ddot{x}_2 = \Omega_{21}x_1 - \Omega_{22}x_2 \tag{2.3.2b}$$

se definirmos as seguintes constantes Ω_{ij}:

$$\Omega_{11} = \frac{k_1 + k_2}{m_1} \qquad \Omega_{12} = \frac{k_2}{m_1} \tag{2.3.3a}$$

$$\Omega_{21} = \frac{k_2}{m_2} \qquad \Omega_{22} = \frac{k_2 + k_3}{m_2} \tag{2.3.3b}$$

Em forma matricial, a Eq. (2.3.2) fica

$$\begin{pmatrix} \ddot{x}_1 \\ \ddot{x}_2 \end{pmatrix} = - \begin{pmatrix} \Omega_{11} & -\Omega_{12} \\ -\Omega_{21} & \Omega_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \tag{2.3.4}$$

ou, de maneira ainda mais reduzida, simplesmente como

$$\ddot{X} = -\Omega X \,. \tag{2.3.5}$$

onde Ω é a matriz quadrada cujos elementos são os Ω_{ij} da Eq. (2.3.4). Mas a Eq. (2.3.5) é em tudo semelhante à Eq. (2.1.2), da qual já conhecemos a solução! Podemos então queimar etapas e procurar, diretamente, por soluções na forma

$$X(t) = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} e^{i\omega t} = A e^{i\omega t} \tag{2.3.6}$$

sendo que as constantes A e ω precisam ainda ser determinadas. Mas já sabemos o que fazer: derivando duas vezes a função teste, obtemos

$$\ddot{X}(t) = -A\omega^2 e^{i\omega t} \tag{2.3.7}$$

e assim recairemos, novamente, no problema de encontrar os autovalores (ω^2) e autovetores (A) da matriz de coeficientes Ω:

$$\Omega A = \omega^2 A \,, \tag{2.3.8}$$

ou, explicitamente:

$$\begin{pmatrix} \Omega_{11} & -\Omega_{12} \\ -\Omega_{21} & \Omega_{22} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \omega^2 \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \tag{2.3.9}$$

2.3.1 Um caso particular: batimentos

Ao invés de resolver o caso geral da Eq. (2.3.9), que não traria muitas novidades, mas só complicações algébricas desnecessárias, vamos focar no problema particular quando as partículas têm a mesma massa m e as molas das extremidades têm a mesma constante k, diferentemente da mola central, cuja constante elástica é dada por K:

$$m_1 = m_2 = m \,, \qquad k_1 = k_3 = k \,, \quad k_2 = K \tag{2.3.10}$$

Vamos também mudar um pouco a notação e reescrever a Eq. (2.3.9) na seguinte forma:

$$\begin{pmatrix} \omega_1^2 + \omega_2^2 & -\omega_2^2 \\ -\omega_2^2 & \omega_1^2 + \omega_2^2 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \omega^2 \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \tag{2.3.11}$$

sendo que as novas constantes ω_1 e ω_2 são dadas, usando as Eqs. (2.3.3), por

$$\omega_1^2 = \frac{k}{m} \,, \qquad \omega_2^2 = \frac{K}{m} \tag{2.3.12}$$

Buscando pelos autovalores da matriz Ω da Eq. (2.3.11), encontramos dois valores distintos: $\omega_\alpha^2 = \omega_1^2$ e $\omega_\beta^2 = \omega_1^2 + 2\omega_2^2$. Vamos investigar os dois casos separadamente:

1. $\omega_\alpha^2 = \omega_1^2$:

   Obviamente, esse autovalor nos leva a $\omega_\alpha = \omega_1$. Os valores possíveis para ω na nossa solução teste são então $\pm\omega_\alpha$. Nos dois casos, os autovetores são da forma (verifique!)

$$A_\alpha = c_\alpha \begin{pmatrix} 1 \\ 1 \end{pmatrix} \tag{2.3.13}$$

© Luiz T. F. Eleno, v. 0.3, 2021

com c_α uma constante, o que nos fornece as seguintes soluções independentes:

$$X_{\alpha_1}(t) = A_\alpha e^{i\omega_\alpha t}\,, \qquad X_{\alpha_2}(t) = A_\alpha e^{-i\omega_\alpha t} \tag{2.3.14}$$

2. $\omega_\beta^2 = \omega_1^2 + 2\omega_2^2$:

Procedendo como no caso anterior, temos que $\omega_\beta = \sqrt{\omega_1^2 + 2\omega_2^2}$, que fornece, como você pode verificar, autovetores dados por

$$A_\beta = c_\beta \begin{pmatrix} 1 \\ -1 \end{pmatrix} \tag{2.3.15}$$

(c_β constante) e as seguintes soluções independentes:

$$X_{\beta_1}(t) = A_\beta e^{i\omega_\beta t}\,, \qquad X_{\beta_2}(t) = A_\beta e^{-i\omega_\beta t} \tag{2.3.16}$$

Juntando as quatro soluções encontradas acima, montamos a solução geral do problema:

$$X(t) = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \left(\alpha_1 e^{i\omega_\alpha t} + \alpha_2 e^{-i\omega_\alpha t}\right) + \begin{pmatrix} 1 \\ -1 \end{pmatrix} \left(\beta_1 e^{i\omega_\beta t} + \beta_2 e^{-i\omega_\beta t}\right) \tag{2.3.17}$$

na qual incorporamos as constantes c_α e c_β nas constantes α_1, α_2, β_1 e β_2.

Já sabemos que podemos reescrever exponenciais complexas como senos e cossenos ou, se preferirmos, como um cosseno e uma fase, como na expressão abaixo:

$$X(t) = \begin{pmatrix} 1 \\ 1 \end{pmatrix} r_\alpha \cos(\omega_\alpha t + \phi_\alpha) + \begin{pmatrix} 1 \\ -1 \end{pmatrix} r_\beta \cos(\omega_\beta t + \phi_\beta) \tag{2.3.18}$$

Repare que, seja de que maneira escolhermos representar a solução geral, temos sempre quatro constantes, o que é esperado, pois o sistema de EDOs do qual buscamos a solução, a Eq. (2.3.2), é formado por duas EDOs de segunda ordem, que dependem de quatro condições iniciais: as posições e velocidades iniciais dos dois osciladores. Vamos, para fixar as ideias, considerar a seguinte configuração inicial para o sistema:

$$x_1(0) = x_0\,, \qquad \dot{x}_1(0) = 0 \tag{2.3.19a}$$
$$x_2(0) = 0\,, \qquad \dot{x}_2(0) = 0 \tag{2.3.19b}$$

Ou seja, no instante inicial, ambas as partículas estão em repouso, mas o primeiro oscilador foi deslocado de x_0 em relação à sua posição de equilíbrio, mas mantendo o segundo oscilador na posição inicial. Na notação matricial, a condição inicial é representada na forma

$$X(0) = \begin{pmatrix} x_0 \\ 0 \end{pmatrix}\,, \qquad \dot{X}(0) = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{2.3.20}$$

o que fornece o sistema de quatro equações e quatro incógnitas abaixo:

$$r_\alpha \cos\phi_\alpha + r_\beta \cos\phi_\beta = x_0 \tag{2.3.21a}$$
$$r_\alpha \cos\phi_\alpha - r_\beta \cos\phi_\beta = 0 \tag{2.3.21b}$$
$$-\omega_\alpha r_\alpha \operatorname{sen}\phi_\alpha - \omega_\beta r_\beta \operatorname{sen}\phi_\beta = 0 \tag{2.3.21c}$$
$$-\omega_\alpha r_\alpha \operatorname{sen}\phi_\alpha + \omega_\beta r_\beta \operatorname{sen}\phi_\beta = 0 \tag{2.3.21d}$$

Esse sistema é resolvido quase imediatamente (se você for esperto e somar e subtrair equações convenientemente!), fornecendo para as constantes ϕ_α, ϕ_β, r_α e r_β os seguintes valores:

$$\phi_\alpha = \phi_\beta = 0\,, \qquad r_\alpha = r_\beta = \frac{x_0}{2} \tag{2.3.22}$$

o que nos leva à solução final do problema:

$$x_1(t) = \frac{x_0}{2}\left(\cos\omega_\alpha t + \cos\omega_\beta t\right) \tag{2.3.23a}$$
$$x_2(t) = \frac{x_0}{2}\left(\cos\omega_\alpha t - \cos\omega_\beta t\right) \tag{2.3.23b}$$

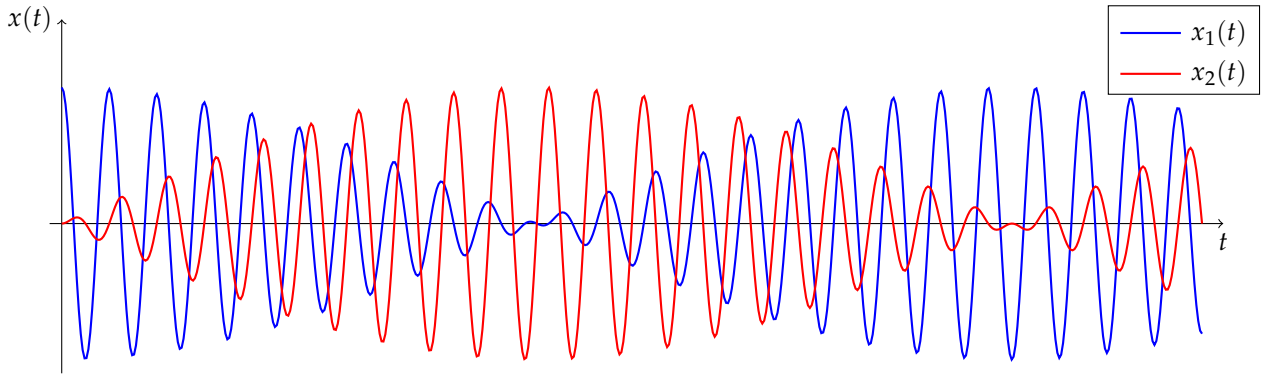


FIGURA 2.5: Batimentos em um sistema com duas frequências parecidas tal que $\overline{\omega} = 8$ e $\delta = 0.2$ (s^{-1}).

Vamos agora fazer uma transformação que pode parecer absurda: veja que, dados quaisquer dois números, a e b, é sempre verdade que

$$a = \frac{a+b}{2} + \frac{a-b}{2}, \qquad b = \frac{a+b}{2} - \frac{a-b}{2} \tag{2.3.24}$$

Com essas estranhas equações, mais as relações de prostaferese que encontramos no Apêndice B, podemos reescrever a Eq. (2.3.23) como

$$x_1(t) = x_0 \cos\overline{\omega}t \cos\delta t \tag{2.3.25a}$$
$$x_2(t) = x_0 \operatorname{sen}\overline{\omega}t \operatorname{sen}\delta t \tag{2.3.25b}$$

sendo $\overline{\omega}$ e δ definidos como

$$\overline{\omega} = \frac{\omega_\alpha + \omega_\beta}{2}, \qquad \delta = \frac{\omega_\beta - \omega_\alpha}{2} \tag{2.3.26}$$

Repare que $\overline{\omega}$ é a média dos dois valores de ω do sistema de osciladores, ao passo que δ é, fundamentalmente, a diferença entre eles. O que acontece se ω_α e ω_β são muito parecidos entre si? Ora, nesse caso, δ será um número muito pequeno. Com isso, $\cos\delta t$ decai muito lentamente à medida em que o tempo passa. Já $\operatorname{sen}\delta t$ cresce de forma bastante lenta. Isso dá origem ao fenômeno conhecido como *batimento*, ilustrado na Figura 2.5. As funções $\cos\overline{\omega}t$ e $\operatorname{sen}\overline{\omega}t$ são as responsáveis pela oscilação de menor período, enquanto $\cos\delta t$ e $\operatorname{sen}\delta t$ levam a uma *modulação* da amplitude dos dois osciladores.

Em termos de energia mecânica, o fenômeno de batimento pode ser entendido como uma transferência da energia cinética e potencial do sistema para cada um dos osciladores alternadamente. Por exemplo, quando a partícula da esquerda está oscilando com amplitude próxima de seu máximo valor ($x_0/2$), ou seja, com praticamente toda a energia do sistema, o segundo oscilador está muito próximo do repouso na posição de equilíbrio. Lenta e periodicamente a situação se reverte, com um período inversamente proporcional a δ.

2.4 Por que modos normais?

Antes de continuar com mais alguns exemplos, vamos tentar entender uma coisa: por que usamos a expressão *modos normais* no título do capítulo? A explicação passa pelo conceito de coordenadas generalizadas, visto em Mecânica Clássica. Aqui, tentaremos evitar complicar ainda mais o texto, e vamos, portanto, evitar introduzir ainda mais conceitos. Vamos nos limitar, então, a analisar o que se passa com a energia mecânica, que já começamos a abordamos no final da seção anterior. Uma introdução bastante elementar sobre o assunto, evitando conceitos de Física Clássica, é feita por Nussenszveig [1]. Por outro lado, se quiser ir mais a fundo no assunto, respire fundo mergulhe na Física Matemática e na Mecânica Clássica [2–4]. É um assunto fascinante!

Para seguir em frente, vamos escrever as energias cinética e potencial do sistema de dois osciladores da Figura 2.4, com $m_1 = m_2 = m$, $k_1 = k_3 = k$ e $k_2 = K$, que tratamos na seção anterior. Essas duas energias são dadas, respectivamente, por

$$T = \frac{1}{2}m\dot{x}_1^2 + \frac{1}{2}m\dot{x}_2^2 \tag{2.4.1a}$$
$$V = \frac{1}{2}kx_1^2 + \frac{1}{2}K(x_2 - x_1)^2 + \frac{1}{2}kx_2^2 \tag{2.4.1b}$$

Vamos focar, inicialmente, a energia potencial. No plano (x_1, x_2), e para V fixado em algum valor constante, a Eq. (2.4.1b) é a equação de uma elipse, tal como mostra a Figura 2.6. No entanto, os semi-eixos principais da elipse não apontam nas direções x_1 e x_2, e, sim, ao longo das direções y_1 e y_2, mostradas na Figura 2.6. O semi-eixo maior está na direção y_1, ao passo que, na direção y_2, encontramos o semi-eixo menor. Repare que podemos fazer uma mudança de coordenadas (uma rotação) e reescrever a Eq. (2.4.1b) em função das novas variáveis y_1 e y_2. A pergunta então é: qual a maneira mais fácil de realizar essa rotação?

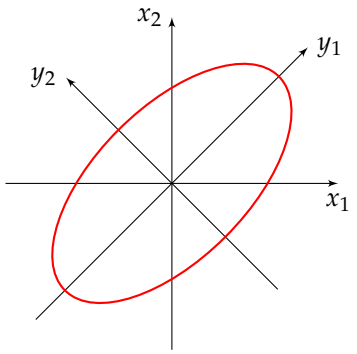


FIGURA 2.6: A energia potencial de um sistema de dois osciladores de mesma massa, acoplados por molas, forma uma elipse quando projetada no plano das posições x_1 e x_2 dos osciladores. Os eixos y_1 e y_2 apontam nas direções dos semi-eixos principais da elipse.

A resposta pode não agradar a alguns leitores: diagonalizar a matriz Ω... Veja no apêndice D que qualquer matriz diagonalizável Ω pode ser escrita como

$$\Omega = V\Lambda V^{-1} \tag{2.4.2}$$

onde Λ é uma matriz diagonal cujos elementos não-nulos são os autovalores de Ω, V é uma matriz cujas colunas são autovetores correspondentes aos autovalores em Λ e V^{-1} é a inversa de V. No caso em questão, sabemos da Eq. (2.3.11) que

$$\Omega = \begin{pmatrix} \omega_1^2 + \omega_2^2 & -\omega_2^2 \\ -\omega_2^2 & \omega_1^2 + \omega_2^2 \end{pmatrix} \tag{2.4.3}$$

cujos autovalores, como também vimos, são $\omega_\alpha^2 = \omega_1^2$ e $\omega_\beta^2 = \omega_1^2 + 2\omega_2^2$, com autovetores dados pelas Eqs. (2.3.13) e (2.3.15), respectivamente. Sabendo tudo isso, podemos escrever Λ e V (e, portanto, V^{-1}) da seguinte forma:

$$\Lambda = \begin{pmatrix} \omega_1^2 & 0 \\ 0 & \omega_1^2 + 2\omega_2^2 \end{pmatrix}, \qquad V = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \qquad V^{-1} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \tag{2.4.4}$$

Repare que adotamos $c_\alpha = c_\beta = 1/\sqrt{2}$, de forma a *normalizar* os autovetores A_α e A_β (normalizar um vetor v é, nesse caso, impor a condição $v^\dagger v = 1$, onde $v^\dagger$ é a transposta conjugada de v). Você pode verificar que a Eq. (2.4.2) realmente é satisfeita por essas matrizes! Veja, agora, que a equação original de autovalores e autovetores do problema, ou seja, a Eq. (2.3.8), pode ser colocada na seguinte forma:

$$V\Lambda V^{-1}A = \omega^2 A \tag{2.4.5}$$

Se multiplicarmos pela esquerda os dois lados da equação acima por V^{-1}, obteremos

$$V^{-1}V\Lambda V^{-1}A = \omega^2 V^{-1}A \tag{2.4.6}$$

Mas repare que $V^{-1}V = I$, sendo I a matriz identidade. Além disso, se fizermos a mudança de variável $V^{-1}A = A'$ (que é a rotação de que falamos anteriormente!) obteremos

$$\Lambda A' = \omega^2 A' \tag{2.4.7}$$

A Eq. (2.4.7) é uma equação para os autovalores (ω^2) e autovetores (A') da matriz Λ com os mesmos autovalores que Ω (já conhecidos!), e cujos autovetores são dados simplesmente por

$$A'_\alpha = V^{-1}A_\alpha = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad A'_\beta = V^{-1}A_\beta = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{2.4.8}$$

onde usamos as Eqs. (2.3.13) e (2.3.15) para A_α e A_β com, novamente, $c_\alpha = c_\beta = 1/\sqrt{2}$. Veja que A'_α e A'_β são (a menos da normalização) os *vetores canônicos* do espaço de matrizes 2×2. Isso faz todo sentido, já que Λ é uma matriz diagonal! Por isso, dizemos que Eq. (2.4.7) é a *forma diagonalizada* do problema.

Mas agora observe: a solução do problema, como vimos, é dada pela Eq. (2.3.17), que podemos escrever abreviadamente como

$$X(t) = A_\alpha\left(\alpha_1 e^{i\omega_\alpha t} + \alpha_2 e^{-i\omega_\alpha t}\right) + A_\beta\left(\beta_1 e^{i\omega_\beta t} + \beta_2 e^{-i\omega_\beta t}\right). \tag{2.4.9}$$

O que acontece se multiplicarmos ambos os lados da equação acima por V^{-1}? Acontece o seguinte:

$$Y(t) = V^{-1}X(t) = A'_\alpha \left(\alpha_1 e^{i\omega_\alpha t} + \alpha_2 e^{-i\omega_\alpha t}\right) + A'_\beta \left(\beta_1 e^{i\omega_\beta t} + \beta_2 e^{-i\omega_\beta t}\right) \tag{2.4.10}$$

onde introduzimos o vetor $Y(t)$, dado explicitamente por

$$Y(t) = \begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix} = V^{-1} \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = \begin{pmatrix} x_1(t) + x_2(t) \\ x_1(t) - x_2(t) \end{pmatrix} \tag{2.4.11}$$

Substituindo os valores já encontrados para A'_α e A'_β, $Y(t)$ é escrito como

$$Y(t) = \begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix} = \begin{pmatrix} \alpha_1 e^{i\omega_\alpha t} + \alpha_2 e^{-i\omega_\alpha t} \\ \beta_1 e^{i\omega_\beta t} + \beta_2 e^{-i\omega_\beta t} \end{pmatrix} \tag{2.4.12}$$

Calma que estamos terminando! Veja que, como $Y(t) = V^{-1}X(t)$, podemos inverter essa equação e escrever

$$X(t) = VY(t) \tag{2.4.13}$$

o que nos leva ao seguinte resultado:

$$x_1(t) = \frac{y_1(t) + y_2(t)}{2} \tag{2.4.14a}$$

$$x_2(t) = \frac{y_1(t) - y_2(t)}{2} \tag{2.4.14b}$$

E o que acontece se substituímos as Eqs. (2.4.14) nas Eqs. (2.4.1) para as energias cinética e potencial? Fazendo a substituição e simplificando, chegamos a

$$T = \frac{1}{4}m\dot{y}_1^2 + \frac{1}{4}m\dot{y}_2^2 \tag{2.4.15a}$$

$$V = \frac{1}{4}ky_1^2 + \frac{1}{4}(k+2K)y_2^2 \tag{2.4.15b}$$

Veja que a Eq. (2.4.15b) é exatamente a equação que buscávamos para a elipse da Figura 2.6 em função das variáveis y_1 e y_2! É interessante notar também que a energia cinética manteve a mesma forma de soma de quadrados de velocidades, mas, agora, de *variáveis generalizadas*, no sentido adotado e empregado em Mecânica Clássica. As Eqs. (2.4.15) são a maneira mais simples de escrever as expressões para as energias cinética e potencial, e são as formas diagonalizadas das respectivas *formas quadráticas* em duas variáveis, como são conhecidas expressões desse tipo. No fundo, na verdade diagonalizamos a *Lagrangiana* do sistema de osciladores. É nesse sentido que usamos o termo *modos normais de vibração*, o que responde a pergunta feita no título da presente seção.

Uma outra maneira de interpretar as Eqs. (2.4.15) é como dois osciladores harmônicos desacoplados, de massa $m/2$ e constantes de mola $k/2$ e $(k+2K)/2$, com deslocamentos dados por $y_1(t)$ e $y_2(t)$. Nesse sentido, vemos que os modos normais são uma maneira de desacoplar o sistema original de EDOs!

É interessante também investigar o significado físico dos modos normais de vibração, $y_1(t)$ e $y_2(t)$. Digamos que, ajustando as condições iniciais, as partículas se encontram no *autoestado* $y_1(t)$, ou seja, tal que $\beta_1 = \beta_2 = 0$ e, portanto, $y_2(t) = 0$. A Eq. (2.4.14)b nos diz que, nesse caso, $x_1(t) = x_2(t)$. O que significa isso? Que as partículas oscilam em sincronia, como na Figura 2.7a, tal que o movimento oscilatório das duas partículas acontece sempre na mesma direção e sentido com a mesma amplitude, ou seja, em fase, mantendo a mola central sem deformação. Por outro lado, se as condições iniciais são tais que $\alpha_1 = \alpha_2 = 0$, e consequentemente $y_1(t) = 0$, teremos, pela Eq. (2.4.14)a, que $x_1(t) = -x_2(t)$. Ou seja, agora as partículas oscilam em oposição de fase: enquanto uma se move para a direita, a outra vai pra esquerda, como na Figura 2.7b. A deformação da mola central é, portanto, o dobro da deformação das molas das extremidades.

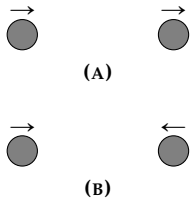


FIGURA 2.7: Modos normais de vibração do sistema de dois osciladores harmônicos acoplados de mesma massa e mesmas constantes elásticas: (a) em fase e (b) em oposição de fase.

Todas as outras soluções do problema são combinações lineares dos dois modos (normais) de vibração descritos no parágrafo anterior. É o que nos diz a Eq. (2.4.9).

2.5 Modelo da molécula de CO_2

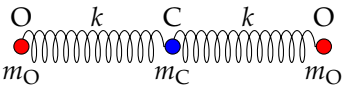


FIGURA 2.8: A molécula de CO_2 como três osciladores acoplados por duas molas.

Vamos agora estudar um caso diferente. Ao lidar realisticamente com objetos tão minúsculos quanto moléculas, a Mecânica Quântica é fundamental. Ainda assim, a Mecânica Clássica fornece visões qualitativas sobre o que podemos esperar antes de quantizar o problema. Nesse espírito, vamos estudar classicamente as vibrações longitudinais da molécula de CO_2. Para isso, vamos modelar a molécula como um sistema de três massas acopladas por duas molas, como mostra a Figura 2.8. As partículas pontuais, de massas m_O nas extremidades e m_C no centro, representam os átomos de oxigênio e carbono, respectivamente. Molas de constantes k ligam os dois átomos de O ao de C, e o sistema é livre de vínculos externos, a não ser pelo fato de que vamos impor a movimentação apenas no eixo x. Obviamente, o modelo de molécula, por exemplo, poderia se movimentar de tal maneira a deixar de ser linear. Mas vamos deixar isso de lado e focar apenas nas vibrações longitudinais.

Nesse ponto, provavelmente não deve ser um problema escrever rapidamente as equações do movimento a partir da 2ª Lei de Newton:

$$m_O\ddot{x}_1 = k(x_2 - x_1) \tag{2.5.1a}$$
$$m_C\ddot{x}_2 = -k(x_2 - x_1) + k(x_3 - x_2) \tag{2.5.1b}$$
$$m_O\ddot{x}_3 = -k(x_3 - x_2) \tag{2.5.1c}$$

sendo que x_1 e x_3 representam o deslocamento dos átomos de oxigênio das extremidades esquerda e direita, respectivamente, enquanto x_2 é o deslocamento do átomo de carbono, em relação às posições de equilíbrio, quando as molas estão em seus comprimentos de repouso. Se definirmos as duas grandezas abaixo,

$$\omega_0^2 = \frac{k}{m_O}, \qquad \mu = \frac{m_O}{m_C}, \tag{2.5.2}$$

podemos modificar um pouco o sistema para o seguinte formato:

$$\ddot{x}_1 = -\omega_0^2(x_1 - x_2) \tag{2.5.3a}$$
$$\ddot{x}_2 = -\mu\omega_0^2(-x_1 + 2x_2 - x_3) \tag{2.5.3b}$$
$$\ddot{x}_3 = -\omega_0^2(x_3 - x_2) \tag{2.5.3c}$$

ou, seguindo a estratégia até aqui, também em forma matricial:

$$\begin{pmatrix} \ddot{x}_1 \\ \ddot{x}_2 \\ \ddot{x}_3 \end{pmatrix} = -\omega_0^2 \begin{pmatrix} 1 & -1 & 0 \\ -\mu & 2\mu & -\mu \\ 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \tag{2.5.4}$$

Mais compactamente, essa última equação é simplesmente reescrita como

$$\ddot{X} = -\omega_0^2 M X\,. \tag{2.5.5}$$

Em relação ao problema da seção 2.3, há pouca diferença, a não ser pelo fato de termos uma matriz diferente, que agora tem três linhas e três colunas, ao invés de duas de cada. Nada nos impede, portanto, de buscar por soluções na forma

$$X = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} e^{i\omega t} = A e^{i\omega t} \tag{2.5.6}$$

o que nos leva, novamente, a uma equação de autovalores e autovetores:

$$\omega_0^2 M A = \omega^2 A \tag{2.5.7}$$

Apenas para simplificar a notação, vamos fazer

$$\lambda = \left(\frac{\omega}{\omega_0}\right)^2 \tag{2.5.8}$$

e reescrever a Eq. (2.5.7) como

$$MA = \lambda A \tag{2.5.9}$$

cuja equação característica vem do determinante nulo abaixo:

$$\det \begin{pmatrix} 1-\lambda & -1 & 0 \\ -\mu & 2\mu-\lambda & -\mu \\ 0 & -1 & 1-\lambda \end{pmatrix} = 0 \tag{2.5.10}$$

ou, calculando o determinante e simplificando,

$$\lambda^3 - 2(\mu+1)\lambda^2 + (2\mu+1)\lambda = 0 \tag{2.5.11}$$

que fornece três valores distintos para λ: $\lambda_\alpha = 0$, $\lambda_\beta = 1$ e $\lambda_\gamma = 2\mu+1$. Vamos investigar caso a caso o que acontece com a solução-teste da Eq. (2.5.6).

1. $\lambda = 0$:

   O autovetor correspondente a esse autovalor é

   $$A_\alpha = c_\alpha \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \tag{2.5.12}$$

   sendo c_α uma constante. Esse é um caso particular, pois fornece, pela Eq. (2.5.8), $\omega = 0$ e, portanto, a solução teste é degenerada: $X(t) = A =$ constante. Mas, como a EDO é de segunda ordem, esperamos que dependa de duas constantes, como vimos na seção 2.3. Por outro lado, não é difícil perceber que, nesse caso, Bt (B constante) também é uma solução do problema, o que nos leva à solução

   $$X_\alpha(t) = A_\alpha(c_1 + c_2 t) \tag{2.5.13}$$

   com c_1 e c_2 constantes a serem determinadas.

2. $\lambda = 1$:

   Esse caso é simples. Os autovetores são da forma

   $$A_\beta = c_\beta \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \tag{2.5.14}$$

   com c_β qualquer constante. Nesse caso, pela Eq. (2.5.8),

   $$\omega_\beta = \omega_0\,, \tag{2.5.15}$$

   o que nos leva ao autoestado

   $$X_\beta(t) = A_\beta \left(c_3 e^{i\omega_\beta t} + c_4 e^{-i\omega_\beta t} \right) \tag{2.5.16}$$

   com c_3 e c_4 constantes de integração.

3. $\lambda = 2\mu + 1$:

   Aqui também nada de novo: os autovetores são

   $$A_\gamma = c_\gamma \begin{pmatrix} 1 \\ -2\mu \\ 1 \end{pmatrix} \tag{2.5.17}$$

   com c_γ constante e uma frequência, pela Eq. (2.5.8), dada por

   $$\omega_\gamma = \omega_0\sqrt{2\mu+1} \tag{2.5.18}$$

   ao passo que o autoestado é

   $$X_\gamma(t) = A_\gamma \left(c_5 e^{i\omega_\gamma t} + c_6 e^{-i\omega_\gamma t} \right) \tag{2.5.19}$$

   com c_5 e c_6 constantes.

Com as três soluções acima, que já sabemos serem os modos normais do problema, escrevemos a solução

geral como

$$X(t) = \begin{pmatrix}1\\1\\1\end{pmatrix}(c_1 + c_2 t) + \begin{pmatrix}1\\0\\-1\end{pmatrix}\left(c_3 e^{i\omega_\beta t} + c_4 e^{-i\omega_\beta t}\right) + \begin{pmatrix}1\\-2\mu\\1\end{pmatrix}\left(c_5 e^{i\omega_\gamma t} + c_6 e^{-i\omega_\gamma t}\right) \tag{2.5.20}$$

ou, se preferirmos nos livrar das exponenciais complexas, como

$$X(t) = \begin{pmatrix}1\\1\\1\end{pmatrix}(c_1 + c_2 t) + \begin{pmatrix}1\\0\\-1\end{pmatrix} c_3 \cos(\omega_\beta t + \phi_\beta) + \begin{pmatrix}1\\-2\mu\\1\end{pmatrix} c_4 \cos\left(\omega_\gamma t + \phi_\gamma\right) \tag{2.5.21}$$

Em todo caso, temos seis constantes a determinar a partir das condições iniciais que fornecem as posições e velocidades das três partículas formadoras da molécula.

Poderíamos agora proceder como na seção 2.4 e diagonalizar o problema. No entanto, já sabemos o que encontraremos apenas analisando os autovetores que acabamos de encontrar. Por exemplo, o primeiro modo normal é apenas uma translação da molécula como um todo. Os três átomos apenas se deslocam em movimento retilíneo uniforme, carregando consigo as molas, que permanecem em seus comprimentos de repouso. Portanto, a molécula não vibra nesse primeiro modo normal, que está representado na Figura 2.9a.

Muito bem. Vamos então investigar o que se passa com o segundo modo normal da molécula. Nesse modo, basta fazer $c_1 = c_2 = c_4 = 0$ na Eq. (2.5.21) e perceber que temos necessariamente $x_1(t) = -x_3(t)$ e $x_2(t) = 0$. Ou seja, o átomo de carbono do centro permanece em repouso e os átomos de oxigênio se movem em oposição de fase, como mostrado na Figura 2.9b.

Por fim, no terceiro modo basta fazer $c_1 = c_2 = c_3 = 0$ na Eq. (2.5.21). Percebemos então que o movimento nesse modo é tal que

$$x_1(t) = -\frac{1}{2\mu} x_2(t) = x_3(t) \tag{2.5.22}$$

Qual a interpretação? Vemos que, como $x_1(t) = x_3(t)$, os átomos de O se movem em fase, como na Figura 2.9c. Por outro lado, o átomo de C se move no sentido contrário aos de O, mas com uma amplitude diferente. Se a amplitude dos átomos de O é c_4, como nos diz a Eq. (2.5.21), a do C é $-2\mu c_4$ (o sinal negativo indica que acontece em oposição de fase). E o que acontece com o centro de massa da molécula? Lembrando da Eq. (2.5.2) que $\mu = m_O/m_C$, temos que o centro de massa se desloca de

$$2m_O c_4 - m_C 2\mu c_4 = 0 \tag{2.5.23}$$

ou seja, ele permanece em repouso. Assim, o terceiro modo normal de vibração acontece com os oxigênios em fase, compensados por um movimento contrário do carbono de forma a manter o centro de massa da molécula em repouso.

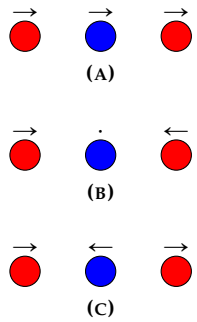


FIGURA 2.9: Modos normais de vibração da molécula de CO_2: (a) movimento retilíneo uniforme, sem vibração; (b) os átomos de O em oposição de fase, com o C em repouso; e (c) oxigênios em fase, com movimento do C de forma a manter o centro de massa da molécula em repouso.

Como na seção 2.3, qualquer movimento longitudinal da molécula de CO_2 é uma *superposição*, ou seja, uma combinação linear, dos três modos normais descritos acima. Você provavelmente sabe o que deve acontecer, mas é um bom exercício tentar encontrar as expressões das energias cinética e potencial em relação às variáveis transformadas $y_1(t)$, $y_2(t)$ e $y_3(t)$ dos modos normais, seguindo o receita da seção 2.4.

2.6 N osciladores harmônicos acoplados

Veremos agora um exemplo importante para o estudo de algumas propriedades vibracionais de sólidos cristalinos. Trata-se do problema de encontrar os modos normais de vibração de uma cadeia de N osciladores harmônicos acoplados em série, como na Figura 2.10. Os osciladores têm massa m_1, m_2, ..., m_N, ao passo que as molas têm constantes elásticas k_1, k_2, ..., k_{N+1}. Os dois osciladores das extremidades estão presos a paredes rígidas e fixas. O problema é, naturalmente, a extensão lógica do caso de dois osciladores acoplados que vimos na seção 2.3. Vamos partir diretamente para o caso de N osciladores, ao invés de apresentar primeiramente o caso de três deles, pois a dificuldade conceitual nos dois casos é a mesma. Será importante ter claro o conceito de solução do problema usando matrizes, já que, como tentamos mostrar até aqui, elas são essenciais para a solução.

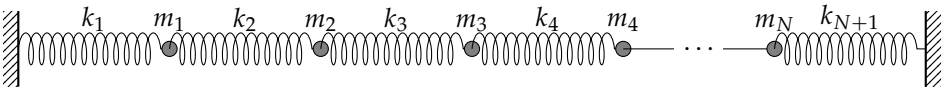


FIGURA 2.10: N osciladores harmônicos de massas m_1, m_2, ..., m_N acoplados em série por $N+1$ molas de constantes elásticas $k_1, k_2, \ldots, k_{N+1}$.

Com toda a experiência adquirida nas seções anteriores, deve ser quase imediato aplicar a Segunda Lei de Newton para escrever o sistema de N EDOs do sistema:

$$m_1\ddot{x}_1 = -k_1x_1 + k_2(x_2 - x_1) \tag{2.6.1a}$$
$$m_2\ddot{x}_2 = -k_2(x_2 - x_1) + k_3(x_3 - x_2) \tag{2.6.1b}$$
$$m_3\ddot{x}_3 = -k_3(x_3 - x_2) + k_4(x_4 - x_3) \tag{2.6.1c}$$
$$\vdots$$
$$m_N\ddot{x}_N = -k_N(x_N - x_{N-1}) + k_{N+1}x_N \tag{2.6.1d}$$

Para tentar escrever uma equação geral para a j-ésima equação nas Eqs. (2.6.1), vamos introduzir as duas *condições de contorno* abaixo:

$$x_0 = x_{N+1} = 0 \tag{2.6.2}$$

que equivalem ao deslocamento nulo das paredes fixas à que estão presas as molas das extremidades. Vamos então reescrever as Eqs. (2.6.1) da seguinte maneira

$$m_1\ddot{x}_1 = -k_1(x_1 - x_0) - k_2(x_1 - x_2) \tag{2.6.3a}$$
$$m_2\ddot{x}_2 = -k_2(x_2 - x_1) - k_3(x_2 - x_3) \tag{2.6.3b}$$
$$m_3\ddot{x}_3 = -k_3(x_3 - x_2) - k_4(x_3 - x_4) \tag{2.6.3c}$$
$$\vdots$$
$$m_N\ddot{x}_N = -k_N(x_N - x_{N-1}) - k_{N+1}(x_N - x_{N+1}) \tag{2.6.3d}$$

que naturalmente são válidas desde que as condições de contorno sejam satisfeitas. Com isso, todas as N EDOs das Eqs. (2.6.3) tem o mesmo formato, ou seja:

$$m_j\ddot{x}_j = -k_j(x_j - x_{j-1}) - k_{j+1}(x_j - x_{j+1}) \qquad (j = 1, 2, \ldots, N)\,. \tag{2.6.4}$$

Vamos trabalhar um pouco na simplificação da Eq. (2.6.4). Podemos escrevê-las no formato

$$\ddot{x}_j = \Omega_{j,j-1}x_{j-1} - \Omega_{jj}x_j + \Omega_{j,j+1}x_{j+1} \tag{2.6.5}$$

se definirmos as grandezas Ω_{ij} como abaixo:

$$\Omega_{j,j-1} = \frac{k_j}{m_j}\,, \qquad \Omega_{j,j+1} = \frac{k_{j+1}}{m_j}\,, \qquad \Omega_{jj} = \Omega_{j,j-1} + \Omega_{j,j+1} \tag{2.6.6}$$

Repare que a Eq. (2.6.5) é uma equação matricial:

$$\ddot{X} = \begin{pmatrix} \ddot{x}_1 \\ \ddot{x}_2 \\ \ddot{x}_3 \\ \vdots \\ \ddot{x}_N \end{pmatrix} = -\begin{pmatrix} \Omega_{11} & -\Omega_{12} & 0 & 0 & \ldots & 0 & 0 \\ -\Omega_{21} & \Omega_{22} & -\Omega_{23} & 0 & \ldots & 0 & 0 \\ 0 & -\Omega_{32} & \Omega_{33} & -\Omega_{34} & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \ldots & -\Omega_{N,N-1} & \Omega_{NN} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_N \end{pmatrix} \tag{2.6.7}$$

que, de maneira quase irritante, pode ser escrita compactamente como

$$\ddot{X} = -\Omega X \tag{2.6.8}$$

A matriz Ω na Eq. (2.6.7) é uma *matriz tridiagonal*, pois apenas os elementos da diagonal e das duas subdiagonais acima e abaixo dela podem ser diferentes de zero. Matrizes tridiagonais aparecem frequentemente em Física, sempre que um sistema apresente interações entre vizinhos, como no presente caso.

Já sabemos como encontrar a solução: buscamos funções-teste no formato

$$x_j(t) = a_j e^{i\omega t} \qquad (j = 1, 2, \ldots, N) \tag{2.6.9}$$

buscando a seguir pelos a_j e por ω após substituição na Eq. (2.6.5), que leva a

$$-\Omega_{j,j-1}a_{j-1} + (\Omega_{jj} - \omega^2)a_j - \Omega_{j,j+1}a_{j+1} = 0 \tag{2.6.10}$$

Em notação matricial, isso é equivalente, como já sabemos dos casos anteriores, a

$$\Omega A = \omega^2 A \tag{2.6.11}$$

e, novamente, ω^2 são os autovalores, com autovetores $A = (a_1 \quad a_2 \quad \ldots \quad a_N)^t$. Além disso, por conta das condições de contorno, dadas pela Eq. (2.6.2), precisamos impor que

$$a_0 = a_{N+1} = 0 \tag{2.6.12}$$

Vamos agora focar num caso particular em que todas as massas são idênticas a m e todas as molas têm a mesma constante k. Nesse caso, teremos

$$\Omega_{j,j-1} = \frac{k}{m} = \omega_0^2\,, \qquad \Omega_{j,j+1} = \frac{k}{m} = \omega_0^2\,, \qquad \Omega_{jj} = \Omega_{j,j-1} + \Omega_{j,j+1} = 2\omega_0^2 \tag{2.6.13}$$

onde voltamos a usar a boa e velha frequência natural ω_0^2. Nosso problema então adquire o seguinte formato matricial, após alguma simplificação:

$$\begin{pmatrix} 2 & -1 & 0 & 0 & \ldots & 0 & 0 \\ -1 & 2 & -1 & 0 & \ldots & 0 & 0 \\ 0 & -1 & 2 & -1 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \ldots & -1 & 2 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_N \end{pmatrix} = \lambda \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_N \end{pmatrix} \tag{2.6.14}$$

na qual usamos novamente a notação

$$\lambda = \left(\frac{\omega}{\omega_0}\right)^2 \tag{2.6.15}$$

Cada linha da Eq. (2.6.14) fornece uma equação, dada abreviadamente por

$$-a_{j-1} + 2a_j - a_{j+1} = \lambda a_j \qquad (j = 1, 2, \ldots, N) \tag{2.6.16}$$

e estamos interessados em encontrar valores de a_j para que seja satisfeita. Repare que os a_j são elementos dos autovetores do problema! Vamos buscar por soluções na forma

$$a_j = c \operatorname{sen} jk \tag{2.6.17}$$

sendo c e k constantes a determinar.[1] Por que esse formato para os a_j? Porque sabemos que existem relações de prostaferese que podem nos ajudar. Além disso, para $j = 0$, a Eq. (2.6.17) fornece $a_0 = 0$, satisfazendo a primeira condição de contorno. Por outro lado, precisamos impor

$$a_{N+1} = c \operatorname{sen}(N+1)k = 0 \tag{2.6.18}$$

para satisfazer a segunda condição de contorno, o que nos leva a

$$k = \frac{\pi n}{N+1}\,, \quad n = 1, 2, 3, \ldots, N \tag{2.6.19}$$

Quaisquer outros valores inteiros não-nulos de n também são aceitáveis, porém apenas repetem as soluções para n no intervalo referido na Eq. (2.6.19), como você pode verificar. Com isso, obtivemos (a menos da constante c) os N autovetores para o problema! Obviamente, precisamos conferir se eles são realmente

[1] É comum, em textos sobre Física do Estado Sólido, escrever $a_j = c \operatorname{sen} ajk$ na Eq. (2.6.17), sendo a é a distância de equilíbrio entre as partículas vizinhas) pois isso leva diretamente aos conceitos de espaço recíproco (ou espaço-k) e zonas de Brillouin. Aqui, no entanto, isso é supérfluo.

autovetores substituindo a Eq. (2.6.17) na Eq. (2.6.16):

$$-\operatorname{sen}(j-1)k + 2\operatorname{sen} jk - \operatorname{sen}(j+1)k = \lambda \operatorname{sen} jk \tag{2.6.20}$$

Essa equação é satisfeita desde que

$$\lambda = 2 - 2\cos k \tag{2.6.21}$$

ou seja,

$$\lambda = 2 - 2\cos\frac{\pi n}{N+1} \qquad (n = 1, 2, \ldots, N) \tag{2.6.22}$$

que são os N autovalores do sistema de osciladores!

Se quisermos determinar a constante c que aparece na Eq. (2.6.17), podemos impor a condição de normalização dos autovetores,

$$A^\dagger A = \sum_{j=1}^{N} a_j^2 = 1 \tag{2.6.23}$$

o que, após algumas manipulações algébricas (confira o Apêndice B), nos leva a

$$c = \sqrt{\frac{2}{N+1}} \tag{2.6.24}$$

e nos permite escrever

$$a_{jn} = \sqrt{\frac{2}{N+1}} \operatorname{sen}\frac{n\pi j}{N+1} \tag{2.6.25}$$

e

$$\lambda_n = 2 - 2\cos\frac{n\pi}{N+1} \tag{2.6.26}$$

Repare que indexamos os autovalores λ e autovetores A com um segundo índice (n) que, como vimos, pode adotar valores inteiros de 1 a N. Você pode verificar, para os casos $N = 2$ e $N = 3$, cujas matrizes Ω são, respectivamente

$$\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \quad \text{e} \quad \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}, \tag{2.6.27}$$

que as equações (2.6.25) e (2.6.26) realmente fornecem os autovetores e autovalores, calculando-os também pelo método tradicional da Álgebra Linear.

Como, agora, escrever a solução geral? Parece, mas não é assim tão complicado de entender:

$$x_j(t) = \sum_{n=1}^{N} a_{jn}\left(\alpha_n e^{i\omega_n t} + \beta_n e^{-i\omega_n t}\right) = \sum_{n=1}^{N} \operatorname{sen}\frac{n\pi j}{N+1}\left(\alpha_n e^{i\omega_n t} + \beta_n e^{-i\omega_n t}\right) \tag{2.6.28}$$

sendo que as $2N$ constantes α_n e β_n (que absorveram a constante c) vêm das condições iniciais de posição e velocidade no instante $t = 0$, ao passo que as frequências ω_n, pela Eq. (2.6.15), são dadas por

$$\omega_n = \omega_0\sqrt{\lambda_n} \tag{2.6.29}$$

Podemos, por fim, livrar-nos das exponenciais complexas e escrever a solução geral como

$$x_j(t) = \sum_{n=1}^{N} \alpha_n \operatorname{sen}\frac{n\pi j}{N+1}\cos\left(\omega_n t + \phi_n\right) \tag{2.6.30}$$

Podemos tentar verificar todo esse equacionamente checando se a Eq. (2.6.30) realmente funciona para o caso $N = 2$, que vimos na seção 2.3. De fato, nesse caso a Eq. (2.6.30) fornece

$$x_j(t) = \sum_{n=1}^{2} \alpha_n \operatorname{sen}\frac{n\pi j}{3}\cos\left(\omega_n t + \phi_n\right) = \alpha_1 \operatorname{sen}\frac{\pi j}{3}\cos\left(\omega_1 t + \phi_1\right) + \alpha_2 \operatorname{sen}\frac{2\pi j}{3}\cos\left(\omega_2 t + \phi_2\right) \tag{2.6.31}$$

Que fornece os deslocamentos para ambas as partículas:

$$x_1(t) = \alpha_1\frac{\sqrt{3}}{2}\cos\left(\omega_1 t + \phi_1\right) + \alpha_2\frac{\sqrt{3}}{2}\cos\left(\omega_2 t + \phi_2\right) \tag{2.6.32a}$$

© Luiz T. F. Eleno, v. 0.3, 2021

e

$$x_2(t) = \alpha_1 \frac{\sqrt{3}}{2} \cos(\omega_1 t + \phi_1) - \alpha_2 \frac{\sqrt{3}}{2} \cos(\omega_2 t + \phi_2) \tag{2.6.32b}$$

As Eqs. (2.6.32) fornecem fundamentalmente (com apenas algumas diferenças de notação) a mesma resposta que a Eq. (2.3.18). Você pode aproveitar esse momento para trabalhar o caso $N = 3$, buscando, por exemplo, a interpretação física dos modos normais, como fizemos nas seções anteriores.

Antes de terminar e partir para soluções numéricas, vamos atacar agora o problema das condições iniciais. Digamos que $x_j(0) = x_{j0}$ e $\dot{x}_j(0) = v_{j0}$ são os deslocamentos e as velocidades iniciais de cada um dos N osciladores, respectivamente. Usando a Eq. (2.6.30), podemos escrever

$$x_j(0) = \sum_{n=1}^{N} \alpha_n \operatorname{sen} \frac{n\pi j}{N+1} \cos\phi_n = x_{j0} \tag{2.6.33a}$$

$$\dot{x}_j(0) = \sum_{n=1}^{N} -\alpha_n \omega_n \operatorname{sen} \frac{n\pi j}{N+1} \operatorname{sen}\phi_n = v_{j0} \tag{2.6.33b}$$

Precisamos achar agora as $2N$ constantes do movimento, α_n e ϕ_n. Vamos fazer um teste e ver qual é o resultado da seguinte soma:

$$X_m = \sum_{j=1}^{N} x_j(0) \operatorname{sen} \frac{m\pi j}{N+1} = \sum_{j=1}^{N} \sum_{n=1}^{N} \alpha_n \operatorname{sen} \frac{n\pi j}{N+1} \cos\phi_n \operatorname{sen} \frac{m\pi j}{N+1} = \sum_{j=1}^{N} x_{j0} \operatorname{sen} \frac{m\pi j}{N+1} \tag{2.6.34}$$

onde m é qualquer um dentre os possíveis valores de n, ou seja, $m = 1, 2, \ldots, N$. A grande dificuldade está na dupla soma no termo central da Eq. (2.6.34). Vamos tentar rearranjar os termos invertendo a ordem das somas[2] em j e em n e colocando tudo que não depende de j em evidência:

$$X_m = \sum_{n=1}^{N} \alpha_n \cos\phi_n \sum_{j=1}^{N} \operatorname{sen} \frac{n\pi j}{N+1} \operatorname{sen} \frac{m\pi j}{N+1} \tag{2.6.35}$$

Se você tentou (e conseguiu!) obter o valor de c dado pela Eq. (2.6.24), é também capaz de verificar que

$$\sum_{j=1}^{N} \operatorname{sen} \frac{n\pi j}{N+1} \operatorname{sen} \frac{m\pi j}{N+1} = \frac{N+1}{2} \delta_{nm} \tag{2.6.36}$$

onde δ_{nm} é chamado de delta de Kronecker, um nome bonito para uma coisa muito simples:

$$\delta_{nm} = \begin{cases} 1, & \text{se } n = m \\ 0, & \text{se } n \neq m \end{cases} \tag{2.6.37}$$

o que faz com que a soma X_m tenha, na verdade, um único termo diferente de zero, ou seja,

$$X_m = \frac{N+1}{2} \alpha_m \cos\phi_m \tag{2.6.38}$$

Pensando em termos de Álgebra Linear, qual é o significado da Eq. (2.6.36)? Simplesmente, que os autovetores A_n do problema são ortogonais entre si. Considerando que a matriz que aparece na Eq. (2.6.14) é simétrica (em Quântica, diríamos que é também hermitiana), isso não é de se espantar.

Vamos agora juntar as Eqs. (2.6.34) e (2.6.38), trocando m por n, para voltar ao índice original:

$$X_n = \frac{N+1}{2} \alpha_n \cos\phi_n = \sum_{j=1}^{N} x_{j0} \operatorname{sen} \frac{n\pi j}{N+1} \tag{2.6.39}$$

Você pode verificar que, fazendo uma outra soma, mas usando a Eq. (2.6.33b) para as velocidades iniciais, descobrimos também que

$$V_n = -\frac{N+1}{2} \alpha_n \omega_n \operatorname{sen}\phi_n = \sum_{j=1}^{N} v_{j0} \operatorname{sen} \frac{n\pi j}{N+1} \tag{2.6.40}$$

[2] nem sempre isso é tão fácil de ser feito. Nesse caso, não há restrições pois os limites dos somatórios são constantes.

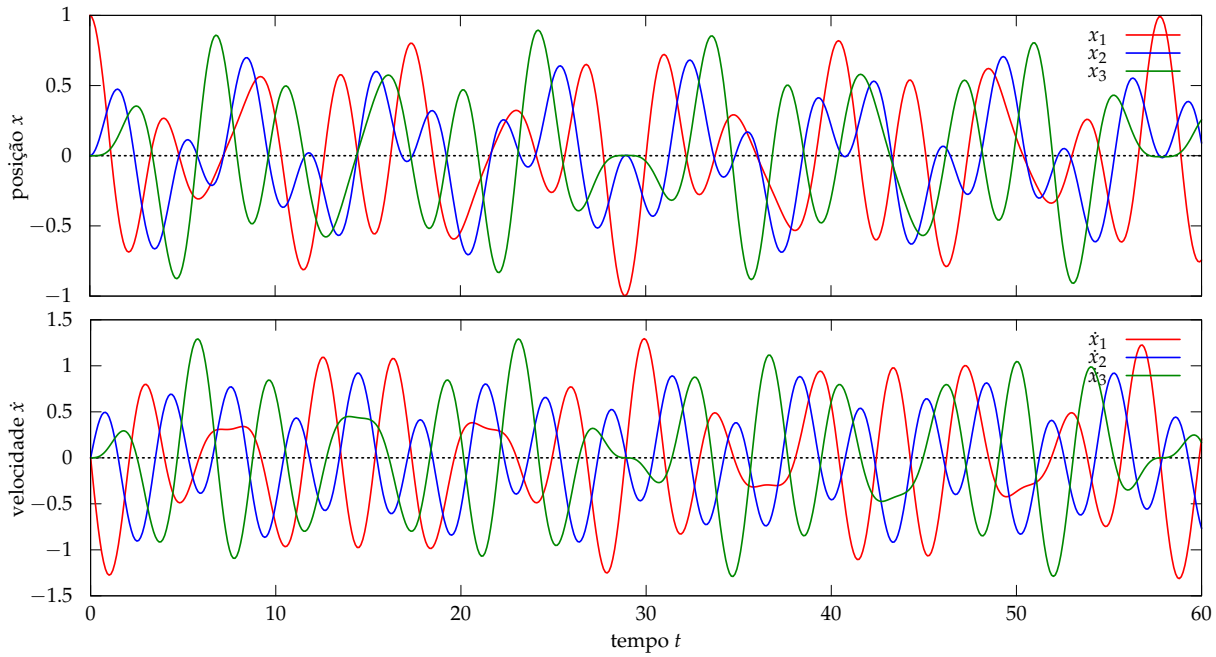


FIGURA 2.11: Posições e velocidades ao longo do tempo de três osciladores idênticos acoplados e presos a paredes fixas. Para o cálculo, adotamos $k = 1\,\mathrm{N/m}$ e $m = 1\,\mathrm{kg}$.

onde apelidamos a soma de V_n por simplicidade. Agora, manipulando as Eqs. (2.6.39) e (2.6.40), não é difícil descobrir que

$$\alpha_n = \frac{2}{N+1}\sqrt{X_n^2 + \left(\frac{V_m}{\omega_n}\right)^2} \tag{2.6.41a}$$

$$\phi_n = -\operatorname{arctg}\left(\frac{V_n}{\omega_n X_n}\right) \tag{2.6.41b}$$

e temos finalmente a solução completa do problema de N osciladores acoplados.[3] Ufa!

Como um teste final, verifique que as Eqs. (2.6.41) realmente acabam fornecendo o mesmo resultado que as Eqs. (2.3.23) para o caso $N = 2$, com as condições iniciais da Eq. (2.3.20).

Para ilustrar a dinâmica de um sistema de N osciladores, vamos resolver o problema para $N = 3$ com a seguinte condição inicial:

$$x(0) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad \dot{x}(0) = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \tag{2.6.42}$$

para o caso em que $k = 1\,\mathrm{N/m}$ e $m = 1\,\mathrm{kg}$. A solução é vista na Figura 2.11, tanto para a posição quanto para a velocidade das partículas. Tente observar o intrincado jogo de transferência de energia cinética e potencial entre as partículas!

Vamos parar por aqui, mas ecoando o início da seção, onde dissemos que o problema em questão é importante para sólidos cristalinos (mas não apenas!). De fato, os fônons, responsáveis, entre outras coisas, pela capacidade térmica vibracional dos sólidos, nada mais são do que os modos normais de vibração de um agregado de átomos, levando em conta todos os modos de oscilação, longitudinais e transversais.

Referências

[1] H. M. Nussenszveig. *Curso de Física Básica*. 2ª ed. Vol. 2. Rio de Janeiro: Editora Edgard Blücher, 2011.

[2] H. Goldstein, C. P. P. Jr. e J. L. Safko. *Classical Mechanics*. 3ª ed. New York, NY, USA: Pearson, 2001.

[3] M. L. Boas. *Mathematical Methods in the Physical Sciences*. 3rd. Hoboken (NJ): Wiley, 2006.

[4] H. Gould, J. Tobochnik e W. Christian. *An introduction to Computer Simulation Methods*. 3rd. San Francisco (CA): Pearson, 2007.

[3] É preciso ter cuidado com o sinal ao calcular ϕ_n usando a Eq. (2.6.41b). Boa parte das linguagens de programação tem uma função `atan2` ou `arctan2` que fornece a resposta correta com base nos sinais do seno e do cosseno fornecidos como argumentos.

Exercícios

2. 1 — Duas partículas de mesma massa m estão presas na vertical por uma mola de constante k e comprimento natural ℓ_0, como mostra a Figura 2.12a. A partícula de cima está ainda presa a uma parede, fixa e rígida, por uma mola idêntica à anterior. Considere apenas vibrações longitudinais na vertical.

(a) Escreva as equações diferenciais que regem o movimento das duas partículas.

(b) Determine as possíveis frequências de vibração das partículas.

(c) Desacople o sistema encontrando os modos normais de vibração.

2. 2 — A Figura 2.12b mostra dois pêndulos acoplados por uma mola de constante k. Cada pêndulo é formado por uma partícula de massa m presa por uma barra rígida e inextesível de massa desprezível e comprimento l. A aceleração da gravidade no local é g.

(a) Escreva as equações diferenciais que regem o movimento das partículas.

(b) Simplifique as equações do item (a) considerando apenas o regime de pequenas oscilações, no qual vale a aproximação $\operatorname{sen}\theta \approx \theta$.

(c) Encontre as possíveis frequências de vibração do sistema.

(d) Determine os modos normais de vibração.

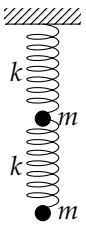
2. 3 — Ao longo de todo o capítulo, estudamos sempre o movimento de oscilação longitudinal de uma sequência de osciladores harmônicos acoplados. Na Figura 2.12c, por outro lado, vemos um exemplo de oscilações transversais de quatro partículas de massa m acopladadas entre si por molas de constante k, num plano horizontal, com o conjunto preso a duas paredes fixas por duas outras molas de constante k. Nesse caso, você pode considerar que as partículas movimentam-se apenas na direção perpendicular ao eixo que passa pelas extremidades das molas ligadas às paredes. Considere que as forças restaurativas são (em módulo) proporcionais ao deslocamento transversal, ou seja, são forças que seguem a Lei de Hooke. **Dica:** se quiser, use o código fornecido no material de apoio, ao invés de fazer as contas na mão.

(a) Escreve as equações do movimento.

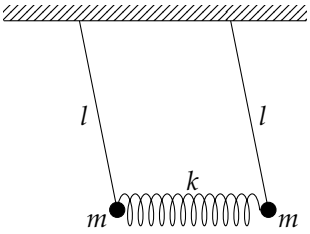
(b) Determine as possíveis frequências de vibração.

(c) Determine os modos normais de vibração e interprete-os fisicamente, esquematizando graficamente o resultado.

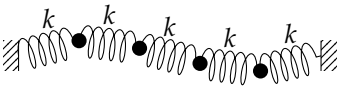
(d) Resolva as equações do movimento para a condição inicial em que todos os osciladores estão em repouso na posição de equilíbrio, exceto o segundo oscilador a partir da esquerda, que está deslocado para uma posição a_0 além da posição de equilíbrio.



(A) Ex. 2.1



(B) Ex. 2.2



(C) Ex. 2.3

FIGURA 2.12

Parte II

Noções de Cálculo Numérico

Capítulo 3

Zero de funções

 © Luiz T. F. Eleno, v. 0.3, 2021

Capítulo 4

Sistemas lineares

 © Luiz T. F. Eleno, v. 0.3, 2021

Capítulo 5

Ajuste de curvas por mínimos quadrados

> É a marca de uma mente culta ficar satisfeita com o grau de precisão que a natureza do assunto admite e não buscar exatidão onde apenas uma aproximação é possível.
>
> Aristóteles, Ética a Nicômaco

5.1 Introdução

Para dar início ao capítulo, vou seguir os passos de Helene [1], uma excelente introdução ao método dos mínimos quadrados, e usar o mesmo exemplo que ele utilizou, ainda que, aqui, eu siga uma abordagem menos formal. Imagine então que, num laboratório, um pesquisador realizou medidas que levaram ao seguinte sistema de equações para a determinação das grandezas a, b e c:

$$\begin{cases} a & = -0.5 \\ a + b & = 4.5 \\ b + c & = 4.5 \\ c & = 4.0 \end{cases} \tag{5.1.1}$$

É imediatamente claro que não existem valores de a, b e c que satisfaçam todas igualdades da Eq. (5.1.1) simultaneamente. Por exemplo, se acreditarmos na primeira delas, então $a = -0.5$ e, da segunda, $b = 5$. Com isso, a terceira fornece $c = -0.5$, o que conflita com a quarta, que nos diz que $c = 4$. Por outro lado, é possível que exista uma maneira de encontrar os melhores valores aproximados para as grandezas a, b e c, de acordo com algum critério pré-estabelecido (as "regras do jogo"). Uma das possibilidades é usar o chamado desvio quadrático (ou erro quadrático) que, nesse caso, é dado por

$$D_q(a, b, c) = (a + 0.5)^2 + (a + b - 4.5)^2 + (b + c - 4.5)^2 + (c - 4.0)^2 \tag{5.1.2}$$

Repare que, se todas as equações fossem satisfeitas exatamente, o desvio quadrático seria zero, e teríamos uma solução exata para as quatro equações. Como isso não é possível nesse caso, elevar cada termo ao quadrado antes de somá-los garante que todos os termos são positivos e que um termo não cancela o outro por terem sinais diferentes. Com isso, como $D_q(a, b, c) \geqslant 0$, deve existir um conjunto de valores para a, b e c que minimiza o desvio quadrático, bastando, para isso, encontrar a, b e c que anulem as derivadas parciais:

$$\frac{\partial D_q}{\partial a} = 0, \qquad \frac{\partial D_q}{\partial b} = 0, \qquad \frac{\partial D_q}{\partial c} = 0 \tag{5.1.3}$$

Isso nos leva ao seguinte sistema de equações:

$$\begin{cases} 2(a + 0.5) + 2(a + b - 4.5) = 0 \\ 2(a + b - 4.5) + 2(b + c - 4.5) = 0 \\ 2(b + c - 4.5) + 2(c - 4.0) = 0 \end{cases} \tag{5.1.4}$$

que, simplificado, fica

$$\begin{cases} 2a + b = 4.0 \\ a + 2b + c = 9.0 \\ b + 2c = 8.5 \end{cases} \tag{5.1.5}$$

Apenas para referência posterior, vamos reescrever o sistema Eq. (5.1.5) em forma matricial:

$$\begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 4.0 \\ 9.0 \\ 8.5 \end{pmatrix} \tag{5.1.6}$$

Esse sistema de equações tem uma única solução, dada por

$$a = 0.625, \qquad b = 2.75, \qquad c = 2.875 \tag{5.1.7}$$

Essa é a solução para a, b e c que minimiza o desvio quadrático. Não acredita? Então faça

$$a = 0.625 + \delta_a, \qquad b = 2.75 + \delta_b, \qquad c = 2.875 + \delta_c \tag{5.1.8}$$

de forma que a solução que supostamente minimiza o desvio quadrático é obtida quando $\delta_a = \delta_b = \delta_c = 0$. Substitua a Eq. (5.1.8) na Eq. (5.1.2) e simplifique. Você vai obter

$$D_q = 5.0625 + \delta_a^2 + (\delta_a + \delta_b)^2 + (\delta_b + \delta_c)^2 + \delta_c^2 \tag{5.1.9}$$

e fica evidente que o mínimo desvio quadrático é 5,0625, quando $\delta_a = \delta_b = \delta_c = 0$, e só pode aumentar quando $\delta_a, \delta_b, \delta_c \neq 0$. Portanto, os valores de a, b e c na Eq. (5.1.7) são os melhores valores para tais grandezas, no sentido de que minimizam o desvio quadrático definido na Eq. (5.1.2). Tal estratégia de minimizar uma função que é uma soma de quadrados (isto é, o desvio quadrático) dá nome ao método: Mínimos Quadrados.

Você poderia agora se perguntar: o experimento parece apontar para a, b e c dados pela Eq. (5.1.7), mas qual é o grau de confiabilidade desse resultado? Em outras palavras, qual é o desvio padrão nos valores obtidos para a, b e c? É uma excelente pergunta que precisa ser (e será) respondida mais adiante. Veremos que, na verdade, mais correto seria dizer que[1]

$$a = 0.625 \pm 1.949, \qquad b = 2.75 \pm 2.25, \qquad c = 2.875 \pm 1.949 \tag{5.1.10}$$

Veremos depois como obter os desvios padrão mostrados na Eq. (5.1.10) (você vai calculá-los no Ex. 1). Além disso, há uma forte correlação entre as grandezas a, b e c, medida pelas covariâncias entre elas (veremos o que isso significa para você determiná-las resolvendo o Ex. 1).

Por fim, perceba que os verdadeiros valores de a, b e c continuam desconhecidos. Os valores dados pela Eq. (5.1.10), ainda que sejam a melhor aproximação de acordo com o critério adotado (ou seja, menor desvio quadrático), ainda assim são apenas isso: aproximações.

5.2 Ajuste de curvas a pontos experimentais

A aplicação mais conhecida do método dos mínimos quadrados é o ajuste de curvas, como ilustra a Figura 5.1. Temos um conjunto de pontos (x_j, y_j) $(j = 1, \ldots, N)$, resultados de alguma medida experimental, por exemplo. Os dados tem bastante dispersão, ou seja, não se encaixam facilmente a uma função suave. No entanto, há uma clara tendência crescente: quando x aumenta, y também parece crescer, descontadas as variações de "curto alcance," ou seja, exceto pelo zigue-zague de baixa amplitude entre pontos consecutivos. Assim, poderíamos tentar descrever os dados ajustando uma curva, dada (por exemplo) por uma função $f_a(x)$, em que $a = (a_1, a_2, \ldots, a_M)$ é um conjunto de parâmetros desconhecidos *a priori*. Por exemplo, poderíamos tentar ajustar a parábola

$$f_a(x) = a_1 + a_2 x + a_3 x^2 \tag{5.2.1}$$

aos dados experimentais, tentando encontrar os melhores valores de $a = (a_1, a_2, a_3)$ — melhores de acordo com algum critério, como a minimização do desvio quadrático. É o que vamos fazer no restante do capítulo.

Para começar, vamos montar uma expressão para o erro quadrático. A diferença entre $f_a(x_j)$, o valor da função em $x = x_j$, e o valor medido de y no mesmo x, ou seja, y_j, é o que chamamos de resíduo em x_j:

$$r_j = f_a(x_j) - y_j \tag{5.2.2}$$

[1] repare que, nesse caso, não é um problema o desvio em a ser maior que o seu valor, pois é plenamente aceitável ser $a < 0$.

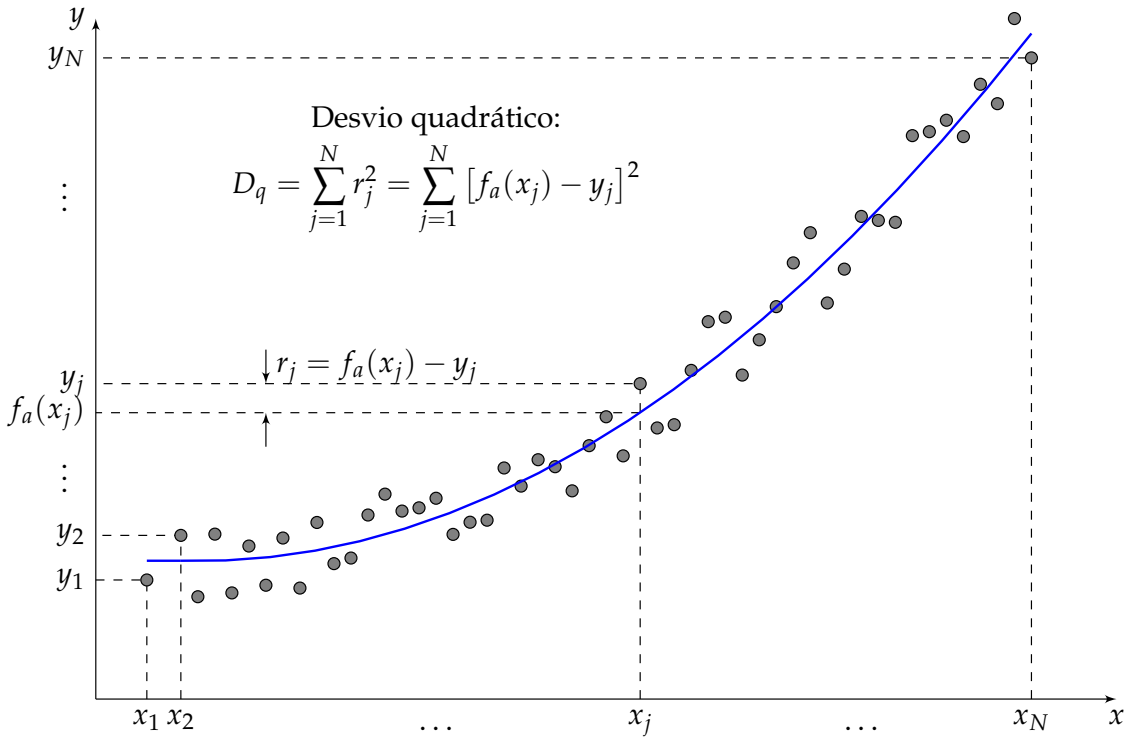


FIGURA 5.1: Exemplo de ajuste de uma função $f_a(x)$ (linha contínua) a N dados experimentais (círculos). O desvio quadrático é a soma dos quadrados dos resíduos $r_j = f_a(x_j) - y_j$. A função $f_a(x)$ depende de uma série de parâmetros $a = (a_1, a_2, \ldots, a_M)$ a serem determinados minimizando o desvio quadrático em relação a a.

O desvio quadrático é então a soma dos quadrados dos resíduos:

$$D_q = \sum_{j=1}^{N} r_j^2 = \sum_{j=1}^{N} \left[f_a(x_j) - y_j\right]^2 \tag{5.2.3}$$

Obtemos então os coeficientes a_i ($i = 1, \ldots, M$) minimizando D_q, ou seja, encontrando os a_j que forneçam um mínimo local:

$$\frac{\partial D_q}{\partial a_i} = 0\,, \qquad (i = 1, \ldots, M) \tag{5.2.4}$$

Vamos começar com um exemplo simples, o ajuste de uma linha reta, antes de descrever outras estratégias e generalizações para funções lineares nos parâmetros a e, mais para o final do capítulo, ajuste de funções não-lineares.

5.3 Regressão linear

O caso mais simples e, até certo ponto, mais importante de ajuste de curvas é a chamada regressão linear, no qual tentamos ajustar uma linha reta aos pontos experimentais. Ou seja, procuramos ajustar a função

$$f_a(x) = a_1 + a_2 x\,, \tag{5.3.1}$$

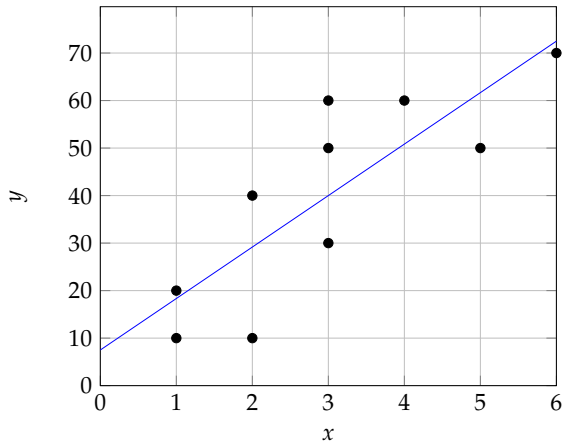
que depende de apenas dois parâmetros (a_1, a_2) a um conjunto de N pontos experimentais (x_j, y_j), com $j = 1, \ldots, N$. Usando a Eq. (5.2.3) o desvio quadrático será, nesse caso:

$$D_q = \sum_{j=1}^{N} \left[f_a(x_j) - y_j\right]^2 = \sum_{j=1}^{N} \left(a_1 + a_2 x_j - y_j\right)^2 \tag{5.3.2}$$

Os melhores valores para a_1 e a_2 são aqueles que minimizam D_q, e podemos encontrá-los fazendo

$$\frac{\partial D_q}{\partial a_1} = 0\,, \qquad \frac{\partial D_q}{\partial a_2} = 0 \tag{5.3.3}$$

© Luiz T. F. Eleno, v. 0.3, 2021



j	x_j	y_j
1	1	10
2	1	20
3	2	10
4	2	40
5	3	30
6	3	50
7	3	60
8	4	60
9	5	50
10	6	70

FIGURA 5.2: Ajuste de uma reta (linha contínua) a dados experimentais (círculos) pelo método dos mínimos quadrados. A equação da reta é $y(x) = 7.5 + 10.8333x$.

Usando a Eq. (5.3.2), teremos então:

$$\frac{\partial D_q}{\partial a_1} = \sum_{j=1}^{N} 2(a_1 + a_2 x_j - y_j) = 0\,, \qquad \frac{\partial D_q}{\partial a_2} = \sum_{j=1}^{N} 2(a_1 + a_2 x_j - y_j)x_j = 0 \tag{5.3.4}$$

que nos fornece um sistema de duas equações em a_1 e a_2 que pode ser simplificado para

$$\left(\sum_{j=1}^{N} 1\right) a_1 + \left(\sum_{j=1}^{N} x_j\right) a_2 = \sum_{j=1}^{N} y_j \tag{5.3.5a}$$

$$\left(\sum_{j=1}^{N} x_j\right) a_1 + \left(\sum_{j=1}^{N} x_j^2\right) a_2 = \sum_{j=1}^{N} x_j y_j \tag{5.3.5b}$$

(5.3.5c)

Para simplificar a notação, vamos fazer

$$\sum_{j=1}^{N} 1 = N\,, \qquad \sum_{j=1}^{N} x_j = S_x\,, \qquad \sum_{j=1}^{N} y_j = S_y\,, \qquad \sum_{j=1}^{N} x_j^2 = S_{xx}\,, \qquad \sum_{j=1}^{N} x_j y_j = S_{xy} \tag{5.3.6}$$

e reescrever a Eq. (5.3.5c) como

$$\begin{cases} N a_1 + S_x a_2 = S_y \\ S_x a_1 + S_{xx} a_2 = S_{xy} \end{cases} \tag{5.3.7}$$

Veja que a Eq. (5.3.7) é um sistema linear com duas equações e duas incógnitas, que pode ser colocado em forma matricial

$$\begin{pmatrix} N & S_x \\ S_x & S_{xx} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} S_y \\ S_{xy} \end{pmatrix} \tag{5.3.8}$$

e imediatamente resolvido:

$$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} N & S_x \\ S_x & S_{xx} \end{pmatrix}^{-1} \begin{pmatrix} S_y \\ S_{xy} \end{pmatrix} = \frac{1}{N S_{xx} - S_x^2} \begin{pmatrix} S_{xx} & -S_x \\ -S_x & N \end{pmatrix} \begin{pmatrix} S_y \\ S_{xy} \end{pmatrix} \tag{5.3.9}$$

ou seja,

$$a_1 = \frac{S_{xx} S_y - S_x S_{xy}}{N S_{xx} - S_x^2}\,, \qquad a_2 = \frac{N S_{xy} S_y - S_x S_y}{N S_{xx} - S_x^2} \tag{5.3.10}$$

Vamos aplicar o método de regressão linear que acabamos de deduzir aos dados mostrados na Figura 5.2. Os valores de N, S_x, S_y, S_{xx} e S_{xy} são facilmente obtidos da Figura 5.2:

$$N = 10\,, \qquad S_x = 30\,, \qquad S_y = 400\,, \qquad S_{xx} = 114\,, \qquad S_{xy} = 1460 \tag{5.3.11}$$

Se quisermos, podemos usar esses valores para montar o sistema na Eq. (5.3.8):

$$\begin{pmatrix} 10 & 30 \\ 30 & 114 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 400 \\ 1460 \end{pmatrix} \tag{5.3.12}$$

e resolver o sistema, mas a Eq. (5.3.10) fornece diretamente os seguintes valores para a_1 e a_2:

$$a_1 = 7.5\,, \qquad a_2 = 10.8333 \tag{5.3.13}$$

Portanto, a reta que melhor se ajusta aos dados, pelo método dos mínimos quadrados, é dada pela equação

$$y(x) = 7.5 + 10.8333x \tag{5.3.14}$$

que também está mostrada na Figura 5.2.

5.4 Formalismo matricial do método de mínimos quadrados

Poderíamos continuar com a mesma estratégia para o ajuste de funções mais complicadas — por exemplo, uma parábola, ou um polinômio de grau maior do que 2. Por tal estratégia, partimos do desvio quadrático D_q dado pela Eq. (5.2.3), escrita em termos de M parâmetros $a_1, a_2, \ldots, a_M$, encontramos as derivadas parciais de D_q em relação aos a_j, os igualamos a zero e resolvemos o sistema resultante. É exatamente o que faremos nesta seção, mas apenas para o caso em que a função a ser ajustada é uma combinação linear de funções conhecidas, ou seja, $f_a(x)$ na Eq. (5.2.3) é escrita como

$$f_a(x) = a_1 f_1(x) + a_2 f_2(x) + \ldots + a_M f_M(x) = \sum_{k=1}^{M} a_k f_k(x) \tag{5.4.1}$$

onde as funções $f_k(x)$ ($k = 1, \ldots, M$) são conhecidas (na verdade, você as escolhe!) e os coeficientes a_k são os parâmetros a serem ajustados pelo método dos mínimos quadrados. Repare que a regressão linear da seção anterior, a Eq. (5.3.1), é um caso particular da Eq. (5.4.1), com apenas duas funções ($M = 2$), sendo que $f_1(x) = 1$ e $f_2(x) = x$.

O ajuste de funções da forma da Eq. (5.4.1) é chamado de ajuste linear, mesmo que, tecnicamente, não estejamos ajustando uma reta, mas sim uma combinação linear de funções conhecidas—a função $f_a(x)$ é linear nos coeficientes a_j. Algumas funções, apesar de não serem da forma da Eq. (5.4.1), podem ser colocadas nessa forma depois de algumas manipulações, como veremos na seção 5.6. Outras equações mais complicadas, por outro lado, não podem ser colocadas nesse formato e precisam de métodos e estratégias de ajuste não-linear, alguns dos quais a serem discutidos na seção 5.7.

Pois então muito bem, vamos começar. Inicialmente, juntando as Eqs. (5.2.3) e (5.4.1), escrevemos o desvio quadrático como

$$D_q = \sum_{j=1}^{N} \left[\sum_{k=1}^{M} a_k f_k(x_j) - y_j \right]^2 \tag{5.4.2}$$

e precisamos calcular as derivadas parciais de D_q em relação aos a_i e igualá-las a zero:

$$\frac{\partial D_q}{\partial a_i} = 0 \qquad (i = 1, \ldots, M) \tag{5.4.3}$$

Respire fundo e analise bem a próxima equação, que contém a derivada parcial de D_q em relação a qualquer um dos parâmetros a_i que aparecem na Eq. (5.4.3):

$$\frac{\partial D_q}{\partial a_i} = \sum_{j=1}^{N} 2 \left[\sum_{k=1}^{M} a_k f_k(x_j) - y_j \right] f_i(x_j) = 0 \qquad (i = 1, \ldots, M) \tag{5.4.4}$$

Simplificando a Eq. (5.4.4), podemos reescrevê-la como

$$\sum_{j=1}^{N} \sum_{k=1}^{M} a_k f_k(x_j) f_i(x_j) = \sum_{j=1}^{N} y_j f_i(x_j) \qquad (i = 1, \ldots, M) \tag{5.4.5}$$

Chegou a hora das matrizes entrarem em ação! Você está mais do que convidado a refrescar sua memória lendo o Apêndice C ou seu livro de Álgebra Linear favorito antes de prosseguir. De qualquer maneira,

vamos começar definindo uma matriz F, de dimensões $N \times M$, cujos elementos F_{ij} são dados por

$$F_{ij} = f_j(x_i) \tag{5.4.6}$$

Explicitamente, a matriz F é dada por

$$F = \begin{pmatrix} f_1(x_1) & f_2(x_1) & \dots & f_M(x_1) \\ f_1(x_2) & f_2(x_2) & \dots & f_M(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ f_1(x_N) & f_2(x_N) & \dots & f_M(x_N) \end{pmatrix} \tag{5.4.7}$$

No caso da regressão linear, por exemplo, em que $M = 2$, $f_1(x) = 1$ e $f_2(x) = x$, a matriz F seria dada simplesmente por

$$F = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_N \end{pmatrix} \tag{5.4.8}$$

Com a definição dos elementos de F segundo a Eq. (5.4.6), podemos reescrever a Eq. (5.4.5) como

$$\sum_{j=1}^{N} \sum_{k=1}^{M} a_k F_{jk} F_{ji} = \sum_{j=1}^{N} y_j F_{ji} \qquad (i = 1, \dots, M) \tag{5.4.9}$$

que vamos escrever num formato ligeiramente diferente:

$$\sum_{j=1}^{N} \sum_{k=1}^{M} F_{ji} F_{jk} a_k = \sum_{j=1}^{N} F_{ji} y_j \qquad (i = 1, \dots, M) \tag{5.4.10}$$

Lembre-se agora da definição de matriz transposta para escrever $(F^T)_{ij} = F_{ji}$. Substitua isso nos termos F_{ji} que aparecem na Eq. (5.4.10):

$$\sum_{j=1}^{N} \sum_{k=1}^{M} (F^T)_{ij} F_{jk} a_k = \sum_{j=1}^{N} (F^T)_{ij} y_j \qquad (i = 1, \dots, M) \tag{5.4.11}$$

Agora repare que, pela Eq. (C.4.4), o lado direito da Eq. (5.4.11) é a i-ésima linha do produto $F^T y$, sendo y o vetor-coluna com os valores y_j:

$$y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{pmatrix} \tag{5.4.12}$$

Do lado esquerdo da Eq. (5.4.11), segundo a Eq. (C.7.6), temos a i-ésima linha do produto $F^T F a$, sendo a o vetor-coluna com os valores dos parâmetros a_j:

$$a = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_M \end{pmatrix} \tag{5.4.13}$$

Portanto, a Eq. (5.4.11) é apenas o formato explícito, linha por linha, da seguinte equação matricial:

$$F^T F a = F^T y \tag{5.4.14}$$

que é a equação que nos trará a resposta para os coeficientes a:

$$a = \left(F^T F\right)^{-1} F^T y \tag{5.4.15}$$

5.4.1 Exemplos de aplicação

Tenho quase certeza de que você está um pouco perdido neste ponto, principalmente se você tem dificuldade em entender as equações envolvendo índices de matrizes em transpostas e produtos. Realmente, é preciso uma enorme concentração e dedicação para acompanhar o desenvolvimento de todas as equações, saindo de somas em vários índices até a abstração das somas pela notação matricial como nas Eqs. (5.4.14) e (5.4.15). Por isso, talvez o uso do método em alguns exemplos torne a sua aplicação e utilidade, se não a dedução, mais evidente.

Vamos começar retomando o exemplo dado pela regressão linear dos dados da Figura 5.2 na página 44. Para este problema, a matriz F pode ser escrita como (vou escrever F^T para economizar espaço)

$$F^T = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 2 & 2 & 3 & 3 & 3 & 4 & 5 & 6 \end{pmatrix} \tag{5.4.16}$$

e o vetor coluna y é dado por (vou escrever y^T para novamente economizar espaço)

$$y^T = \begin{pmatrix} 10 & 20 & 10 & 40 & 30 & 50 & 60 & 60 & 50 & 70 \end{pmatrix} \tag{5.4.17}$$

Vamos agora calcular $F^T F$ (e aqui fica difícil economizar espaço!)

$$F^T F = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 2 & 2 & 3 & 3 & 3 & 4 & 5 & 6 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 2 \\ 1 & 2 \\ 1 & 3 \\ 1 & 3 \\ 1 & 3 \\ 1 & 4 \\ 1 & 5 \\ 1 & 6 \end{pmatrix} = \begin{pmatrix} 10 & 30 \\ 30 & 114 \end{pmatrix} \tag{5.4.18}$$

e também $F^T y$:

$$F^T y = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 2 & 2 & 3 & 3 & 3 & 4 & 5 & 6 \end{pmatrix} \begin{pmatrix} 10 \\ 20 \\ 10 \\ 40 \\ 30 \\ 50 \\ 60 \\ 60 \\ 50 \\ 70 \end{pmatrix} = \begin{pmatrix} 400 \\ 1460 \end{pmatrix} \tag{5.4.19}$$

A Eq. (5.4.14), portanto, se torna

$$\begin{pmatrix} 10 & 30 \\ 30 & 114 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 400 \\ 1460 \end{pmatrix} \tag{5.4.20}$$

que é idêntica à Eq. (5.3.12) que obtivemos anteriormente! A solução, portanto, é a mesma que a vista antes:

$$\begin{pmatrix} a_1 & a_2 \end{pmatrix}^T = \begin{pmatrix} 7.5 & 10.8333 \end{pmatrix} \tag{5.4.21}$$

Como esperado, a regressão linear que vimos na seção 5.3 é apenas um caso particular do método mais geral dado pela equação Eq. (5.4.14).

Mais um exemplo: e se tentássemos ajustar uma parábola aos mesmos dados da Figura 5.2? Ou seja, agora quero ajustar a função

$$f_a(x) = a_1 + a_2 x + a_3 x^2 \tag{5.4.22}$$

aos dados experimentais, e portanto $M = 3$, com $f_1(x)$=1, $f_2(x) = x$ e $f_3(x) = x^2$. Usando os dados disponíveis na tabela que acompanha a Figura 5.2, montamos a matriz F (indicarei aqui a transposta, novamente por questões de espaço):

$$F^T = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 2 & 2 & 3 & 3 & 3 & 4 & 5 & 6 \\ 1 & 1 & 4 & 4 & 9 & 9 & 9 & 16 & 25 & 36 \end{pmatrix} \tag{5.4.23}$$

 © Luiz T. F. Eleno, v. 0.3, 2021

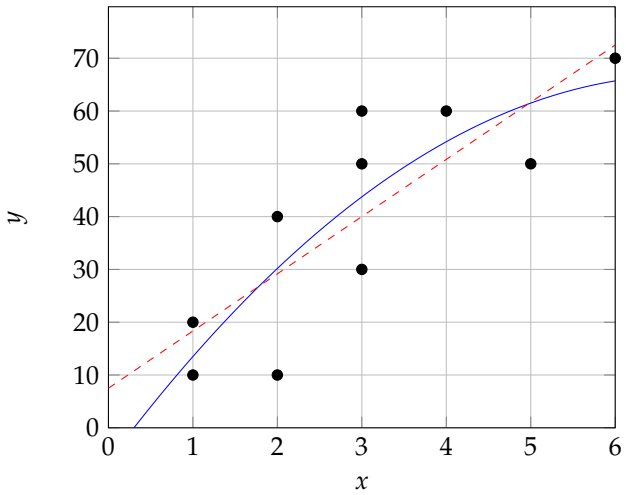


FIGURA 5.3: Ajuste de uma parábola (linha contínua) a dados experimentais (círculos) pelo método dos mínimos quadrados. A reta de melhor ajuste (linha tracejada) também está indicada. Os dados são os mesmos da Figura 5.2.

Com isso, as matrizes F^TF e F^Ty, como você pode verificar, serão dadas por

$$F^TF = \begin{pmatrix} 10 & 30 & 114 \\ 30 & 114 & 504 \\ 114 & 504 & 2454 \end{pmatrix}, \qquad F^Ty = \begin{pmatrix} 400 \\ 1460 \\ 6220 \end{pmatrix} \tag{5.4.24}$$

A partir da Eq. (5.4.14), ou seja, $F^TFa = F^Ty$, encontramos então o vetor de coeficientes resolvendo o sistema linear resultante:

$$a^T = \begin{pmatrix} a_1 & a_2 & a_3 \end{pmatrix} = \begin{pmatrix} -6.305 & 21.363 & -1.560 \end{pmatrix} \tag{5.4.25}$$

A equação da parábola que melhor se ajusta aos dados experimentais é portanto

$$f(x) = -6.305 + 21.363x - 1.560x^2 \tag{5.4.26}$$

Você pode ver o resultado do ajuste na Figura 5.3, onde também está representada a regressão linear, feita anteriormente, para fins de comparação.

Para encerrar esta seção, vamos retornar ao exemplo usado na introdução a este capítulo. Lá vimos como encontrar uma solução aproximada para um sistema linear com mais equações (4) do que incógnitas (3). Vou repetir aqui o sistema da Eq. (5.1.1), mas agora em forma matricial:

$$\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} -0.5 \\ 4.5 \\ 4.5 \\ 4.0 \end{pmatrix} \tag{5.4.27}$$

Já sabemos que não existem a, b e c que satisfaçam a Eq. (5.4.27), por isso buscamos a solução aproximada, já encontrada na introdução. Por outro lado, o ajuste de uma função dada por combinação linear de funções, com o qual trabalhamos na presente seção, é exatamente o mesmo problema, apenas disfarçado com uma roupa diferente. Veja: nesta seção, descobrimos como ajustar uma função dada pela Eq. (5.4.1) a um conjunto de dados (x_j, y_j). Com isso, o que fizemos foi tentar resolver um sistema de equações na forma

$$\sum_{k=1}^{M} a_k f_k(x_j) = y_j \qquad (j = 1, \ldots, N) \tag{5.4.28}$$

ou, usando matrizes,

$$Fa = y \tag{5.4.29}$$

na qual F, y e a são as matrizes definidas anteriormente nas Eqs. (5.4.6), (5.4.12) e (5.4.13), respectivamente. A Eq. (5.4.29) também não tem solução, ou seja, não existem coeficientes a_k que o satisfaça. A Eq. (5.4.27), portanto, é parte do mesmo conjunto de problemas, sintetizado pela Eq. (5.4.29), que resolvemos nesta seção e cuja solução é a Eq. (5.4.14). Assim, à partir da Eq. (5.4.27), extraímos as matrizes F e y, com as quais

podemos calcular $F^T F$ e $F^T y$:

$$F^T F = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}, \qquad F^T y = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} -0.5 \\ 4.5 \\ 4.5 \\ 4.0 \end{pmatrix} = \begin{pmatrix} 4.0 \\ 9.0 \\ 8.5 \end{pmatrix} \tag{5.4.30}$$

Repare que as matrizes $F^T F$ e $F^T y$ na Eq. (5.4.30) são as mesmas que aparecem na Eq. (5.1.6). A solução, portanto, é a mesma! Isso demonstra que o campo de aplicação da Eq. (5.4.14) é maior do que imaginamos a princípio. De fato, o método dos mínimos quadrados pode ser aplicado a inúmeros problemas, e não apenas ao ajuste de curvas [2, 3].

5.5 Estimativa de erros: a matriz de covariância

Lembra-se das aulas de Estatística? Vamos precisar de uma coisinha ou outra que você aprendeu por lá. Se você tem um conjunto de N medidas de uma certa grandeza y (por exemplo, os valores y_j que você usou para ajustar alguma curva), o valor médio (ou simplesmente média, ou ainda, valor esperado) de y, indicado por $\langle y \rangle$ ou $\overline{y}$, é dado por

$$\langle y \rangle = \frac{1}{N} \sum_{j=1}^{N} y_j \tag{5.5.1}$$

O símbolo $\langle\rangle$ é usado para representar o valor esperado do que está dentro dele. Definimos assim a variância do conjunto de dados y_j por

$$v_y = \sigma_y^2 = \left\langle (y - \overline{y})^2 \right\rangle \tag{5.5.2}$$

Já o desvio padrão é definido como a raiz quadrada da variância:

$$\sigma_y = \sqrt{v_y} \tag{5.5.3}$$

Considere agora duas variáveis, y_1 e y_2, cada uma com sua média $\overline{y}_1$ e $\overline{y}_2$, respectivamente. A covariância entre as duas variáveis é definida como

$$v_{y_1 y_2} = \left\langle (y_1 - \overline{y}_1)(y_2 - \overline{y}_2) \right\rangle \tag{5.5.4}$$

E, de maneira mais geral, para um conjunto de N variáveis $y_1, y_2, \ldots, y_N$, definimos uma matriz de covariância, $V^{(y)}$, com cada elemento $V_{ij}^{(y)}$ dado por

$$V_{ij}^{(y)} = \left\langle (y_i - \overline{y}_i)(y_j - \overline{y}_j) \right\rangle \tag{5.5.5}$$

Repare que os elementos da diagonal principal de $V_{ij}^{(y)}$ são as variâncias das grandezas y_i, ou seja,

$$V_{ii}^{(y)} = v_{y_i} \tag{5.5.6}$$

Além disso, $V_{ij}^{(y)}$ é uma matriz simétrica pois, pela Eq. (5.5.5), temos que

$$V_{ji}^{(y)} = V_{ij}^{(y)} \tag{5.5.7}$$

Uma última definição antes de continuar: o coeficiente de correlação entre duas variáveis y_i e y_j é dado por

$$\rho_{ij}^{(y)} = \frac{V_{ij}^{(y)}}{\sigma_{y_i} \sigma_{y_j}} \tag{5.5.8}$$

Você é capaz de perceber que sempre $|\rho_{ij}^{(y)}| \leqslant 1$? É uma aplicação da desigualdade de Schwarz, demonstrada em quase todo livro de Álgebra Linear [4, por exemplo]. O coeficiente de correlação tem uma interpretação bastante útil: se $\rho_{ij}^{(y)} > 0$, subestimar (superestimar) y_1 significa que y_2 provavelmente estará também subestimado (superestimado). Já quando $\rho_{ij}^{(y)} < 0$, se uma variável está subestimada, a outra provavelmente estará

superestimada. Quando mais próximo $|\rho_{ij}^{(y)}|$ estiver de 1, mais a palavra *provavelmente* pode ser substituída por *certamente* [1].

De posse dos conceitos de Estatística discutidos brevemente acima, podemos tentar fazer uma propagação de erros para encontrar as variâncias nos parâmetros obtidos por um ajuste de mínimos quadrados. Ou seja, digamos que, por exemplo, determinamos os parâmetros a e b de uma reta $y(x) = a + bx$ que melhor se ajustam a um conjunto de dados y_j ($1 \leqslant y \leqslant N$). Qual será o grau de confiança, medido por suas variâncias, no valor desses parâmetros?

Para responder a essa pergunta, vamos retornar à Eq. (5.4.15) que nos fornece os parâmetros do ajuste e repetida aqui para maior facilidade:

$$a = (F^T F)^{-1} F^T y \tag{5.4.15}$$

Vou chamar de M toda a combinação de matrizes que aparecem antes de y, ou seja,

$$M = (F^T F)^{-1} F^T \tag{5.5.9}$$

de modo que posso escrever o vetor-coluna a como

$$a = My \tag{5.5.10}$$

Repare que M é uma matriz $M \times N$. Pela definição de produto de matrizes (ver o Apêndice C), cada um dos coeficientes a_i é escrito como uma combinação linear dos valores de y:

$$a_i = \sum_{k=1}^{N} M_{ik} y_k \qquad (1 \leqslant i \leqslant M) \tag{5.5.11}$$

Podemos então buscar pela matriz de covariância dos coeficientes a, $V^{(a)}$, cujos elementos são dados por

$$V_{ij}^{(a)} = \big\langle (a_i - \overline{a}_i)(a_j - \overline{a}_j) \big\rangle \tag{5.5.12}$$

Substituindo a Eq. (5.5.11) na (5.5.12):

$$V_{ij}^{(a)} = \left\langle \left(\sum_{k=1}^{N} M_{ik} y_k - \overline{a}_k \right) \left(\sum_{l=1}^{N} M_{jl} y_l - \overline{a}_l \right) \right\rangle \tag{5.5.13}$$

onde usei um índice diferente (l) dentro dos segundos parênteses para não confundi-lo com o índice k usado nos primeiros. Veja agora que, pela linearidade do valor médio, temos

$$\overline{a}_k = \left\langle \sum_{k=1}^{N} M_{ik} y_k \right\rangle = \sum_{k=1}^{N} M_{ik} \overline{y}_k \tag{5.5.14}$$

e reescrevemos a Eq. (5.5.13) como

$$V_{ij}^{(a)} = \left\langle \left(\sum_{k=1}^{N} M_{ik} (y_k - \overline{y}_k) \right) \left(\sum_{l=1}^{N} M_{jl} (y_l - \overline{y}_l) \right) \right\rangle \tag{5.5.15}$$

ou ainda, usando as propriedades de linearidade do valor médio:

$$V_{ij}^{(a)} = \sum_{k=1}^{N} \sum_{l=1}^{N} M_{ik} M_{jl} \big\langle (y_k - \overline{y}_k)(y_l - \overline{y}_l) \big\rangle \tag{5.5.16}$$

Use agora a definição dos elementos da matriz de covariância, ou seja, a Eq. (5.5.5), para escrever

$$V_{ij}^{(a)} = \sum_{k=1}^{N} \sum_{l=1}^{N} M_{ik} M_{jl} V_{kl}^{(y)} \tag{5.5.17}$$

Basta agora um pequeno rearranjo, introduzindo M^T, a transposta de M, para chegar a

$$V_{ij}^{(a)} = \sum_{k=1}^{N} \sum_{l=1}^{N} M_{ik} V_{kl}^{(y)} (M^T)_{lj} \tag{5.5.18}$$

e, usando a Eq. (C.7.5), descobrimos que a matriz de covariância dos parâmetros a, $V^{(a)}$, é o produto de três matrizes:

$$V^{(a)} = MV^{(y)}M^T \tag{5.5.19}$$

Portanto, se conhecemos a matriz de covariância dos dados, ou seja, $V^{(y)}$, e calculamos M usando a Eq. (5.5.9), conheceremos também a matriz de covariância dos parâmetros ajustados e, por conseguinte, também as covariâncias e desvios padrão de cada um deles. Neste texto, no entanto, estamos trabalhando com dados (x_j, y_j) para os quais não definimos barras de erros (esse caso é tratado em Helene [1], por exemplo). Qual matriz $V^{(y)}$ devemos usar nesse caso? É possível demonstrar, usando técnicas avançadas de Estatística, como o teste de χ^2 (Helene [1] vem novamente em nossa auxílio!), que, quando não conhecemos as covariâncias dos dados, podemos supô-los seguindo aproximadamente uma distribuição gaussiana e, nesse caso, teremos a seguinte expressão para a matriz de covariância em y:

$$V^{(y)} = \frac{D_q}{N-M} I \tag{5.5.20}$$

sendo D_q o nosso já conhecido desvio quadrático dado pela Eq. (5.4.2) e I é a matriz identidade que, nesse caso, deve ter ordem M. Lembre que N é o número de pontos experimentais e M o número de parâmetros no modelo que escolhemos, de modo que $N - M$ é o chamado número de graus de liberdade do problema. Substituindo agora ese último resultado na Eq. (5.5.19) e simplificando, teremos

$$V^{(a)} = \frac{D_q}{N-M} MM^T \tag{5.5.21}$$

No entanto, a matriz M é dada pela Eq. (5.5.9), o que nos leva a

$$V^{(a)} = \frac{D_q}{N-M} (F^TF)^{-1}F^T \left((F^TF)^{-1}F^T\right)^T \tag{5.5.22}$$

e aqui começa um verdadeiro *tour de force* para simplificar a última expressão, usando as propriedades listadas na seção C.6:

$$\begin{aligned}
V^{(a)} &= \frac{D_q}{N-M} (F^TF)^{-1}F^TF \left((F^TF)^{-1}\right)^T \\
V^{(a)} &= \frac{D_q}{N-M} I \left((F^TF)^{-1}\right)^T \\
V^{(a)} &= \frac{D_q}{N-M} \left((F^TF)^T\right)^{-1} \\
V^{(a)} &= \frac{D_q}{N-M} \left(F^TF\right)^{-1}
\end{aligned} \tag{5.5.23}$$

A última equação é a expressão procurada para a matriz de covariância dos parâmetros do ajuste linear por mínimos quadrados. Ela nos fornecerá as barras de erros para o resultado do ajuste.

Por fim, é interessante encontrar uma maneira de obter o desvio quadrático também em notação matricial, o que facilitará nossa vida na hora de calculá-lo. Com a definição das matrizes F, y e a dadas pelas Eqs. (5.4.6), (5.4.12) e (5.4.12), reescrevemos a Eq. (5.4.2) como

$$D_q = \sum_{j=1}^{N} (F_{jk}a_k - y_j)^2 = (Fa - y)^T(Fa - y) \tag{5.5.24}$$

Repare que

$$R = Fa - y \tag{5.5.25}$$

é a matriz-coluna com os resíduos, definidos bem no começo do capítulo pela Eq. (5.2.2).

Vejamos agora todo o trabalho de estimativa de erros em ação num dos exemplos trabalhados anteriormente: a regressão linear da seção 5.3, revista usando a abordagem matricial, o que resultou nas Eqs. (5.4.16)–(5.4.21). Para calcular a matriz de resíduos, dada pela Eq. (5.5.25), vamos usar F, y e a dados pelas Eqs. (5.4.16), (5.4.17) e (5.4.21), respectivamente. Obtemos, pela Eq. (5.5.25):

$$R^T = \begin{pmatrix} 8.3333 & -1.6667 & 19.1667 & -10.8333 & 10 & -10 & -20 & -9.1667 & 11.6667 & 2.5 \end{pmatrix} \tag{5.5.26}$$

e com isso, usando a Eq. (5.5.24):

$$D_q = 1383.33 \tag{5.5.27}$$

 © Luiz T. F. Eleno, v. 0.3, 2021

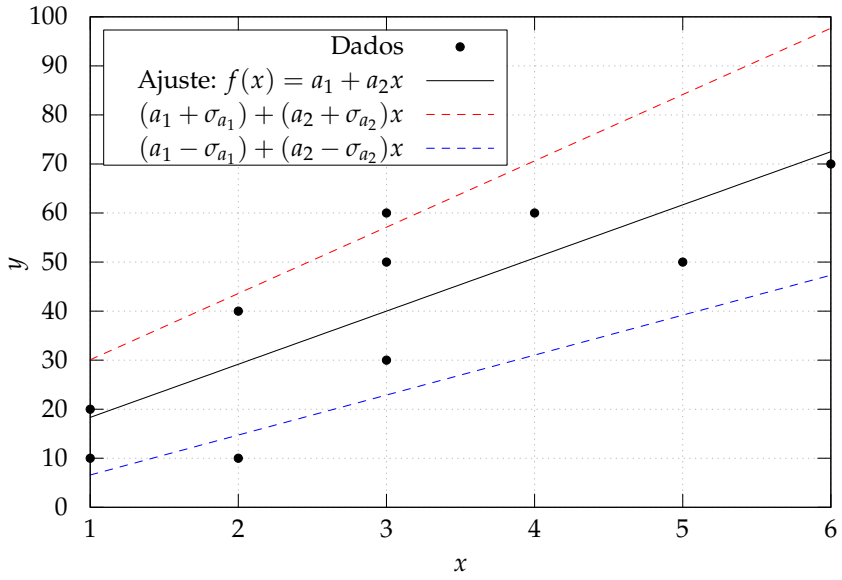


FIGURA 5.4: Novamente a regressão linear (linha contínua) aos dados da 5.2 (círculos), agora também mostrando o intervalo de confiança definido pelas retas com as estimativas dos desvios padrão dos coeficientes a_1 e a_2. (linhas tracejadas)

Usando agora $F^T F$ dado pela Eq. (5.4.18), além de $N - M = 10 - 2 = 8$, na Eq. (5.5.23), obtemos

$$V^{(a)} = \begin{pmatrix} 82.1354 & -21.6146 \\ -21.6146 & 7.2049 \end{pmatrix} \tag{5.5.28}$$

As variâncias v_{a_1} e v_{a_2} nos parâmetros do ajuste $y_a(x) = a_1 + a_2 x$ estão na diagonal principal da matriz da Eq. (5.5.28). Os desvios padrão são dados pala sua raiz quadrada, ou seja:

$$\sigma_{a_1} = 9.0629\,, \qquad \sigma_{a_2} = 2.6842 \tag{5.5.29}$$

De posse coeficientes e de seus respectivos desvios padrão, podemos montar duas outras retas,

$$y_p(x) = (a_1 + \sigma_{a_1}) + (a_2 + \sigma_{a_2})x \tag{5.5.30a}$$
$$y_n(x) = (a_1 - \sigma_{a_1}) + (a_2 - \sigma_{a_2})x \tag{5.5.30b}$$

que nos fornecem uma espécie de intervalo de confiança. Se os dados (x, y) realmente são descritos por uma linha reta, ela pode estar em qualquer lugar dentro de tal intervalo (poderia também estar fora do intervalo, com menor probabilidade). A reta dos mínimos quadrados é apenas a melhor hipótese para a reta real, que continuamos sem conhecer de fato.

Podemos também montar a matriz com os coeficientes de correlação, dados de forma análoga à Eq. (5.5.8):

$$\rho^{(a)} = \begin{pmatrix} 1 & -0.8885 \\ -0.8885 & 1 \end{pmatrix} \tag{5.5.31}$$

Como vimos, o valor negativo do coeficiente de correlação entre a_1 e a_2 indica que, se a_1 está superestimado, provavelmente a_2 estará subestimado, e vice-versa. O módulo do coeficiente (0.885) é relativamente próximo de 1, o que nos dá uma indicação da qualidade do ajuste (o valor seria unitário para ajuste perfeito).

5.5.1 Coeficiente de correlação numa regressão linear

Numa regressão linear, existe um teste que nos dá de forma relativamente fácil a adequação do ajuste de uma reta aos dados. Ele é chamado de coeficiente de correlação e é definido por[2]

$$\mathrm{cor}(x, y) = \frac{v_{xy}}{\sigma_x \sigma_y} \tag{5.5.32}$$

[2]Apesar de terem o mesmo nome, não confundir os coeficientes de correlação definidos na Eq. (5.5.8) para o ajuste de funções lineares nos parâmetros. São coisas diferentes!

em que v_{xy} é a covariância entre os conjuntos de dados (x, y) e σ_x, σ_y são, respectivamente, os desvios padrão dos dados para x e y.

O coeficiente de correlação está sempre no intervalo $-1 \leqslant \text{cor}(x, y) \leqslant 1$. Quanto mais próximo de 1 estiver $|\text{cor}(x, y)|$, melhor é o ajuste a uma linha reta. Um valor positivo de $\text{cor}(x, y)$ indica que a reta tem inclinação positiva, e inclinação negativa se $\text{cor}(x, y) > 0$.

No exemplo que apresentamos anteriormente, com os dados da Figura 5.2 o coeficiente de correlação é dado por

$$\text{cor}(z, y) = \frac{26}{1.549 \times 20.494} = 0.819 \tag{5.5.33}$$

O que indica que os dados seguem uma tendência linear com razoável aproximação (0,819 está razoavelmente próximo de 1), com inclinação positiva , como de fato mostrado na Figura 5.2.

5.6 Problemas linearizáveis

Às vezes, estamos interessados em ajustar funções que não são combinações lineares nos parâmetros, ou seja, que não seguem o formato da Eq. (5.4.1), o que invalidaria o formalismo matricial usado na seção 5.4. Muitas vezes, no entanto, é possível, através de algumas manipulações algébricas, transformar o problema num formato compatível com o da Eq. (5.4.1). Esse processo é chamado de linearização da função ajustável. Isso nem sempre é possível e, às vezes, é difícil enxergar logo de cara se certo problema é ou não linearizável. Nesta seção, vou mostrar apenas alguns exemplos de casos em que isso é possível. Alguns casos são mais "clássicos" e são encontrados comumente em livros de cálculo numérico, ao passo que outros são menos tradicionais e mais específicos a algumas áreas da Física ou da Química.

O primeiro exemplo é a função exponencial

$$f(x) = ae^{bx} \tag{5.6.1}$$

onde a e b são os parâmetros ajustáveis. Podemos linearizar tal função calculando o logaritmo de ambos os lados, obtendo

$$g(x) = \ln f(x) = \ln a + bx \tag{5.6.2}$$

Chamando $\ln a = c$, podemos reescrever $g(x)$ como

$$g(x) = c + bx$$

que é uma função linear em relação aos novos parâmetros b e c. Uma vez obtido seus valores por mínimos quadrados, basta fazer $a = e^c$ para encontrar o parâmetro da função original.

Um exemplo importante e parecido com o anterior aparece frequentemente em Física Estatística, Química e áreas correlatas, como Engenharia de Materiais. Trata-se de ajustar uma função do tipo Arrhenius:

$$f(T) = Ae^{-Q/RT} \tag{5.6.3}$$

em que A e Q são os parâmetros ajustáveis e R é uma constante (geralmente a constante universal dos gases). Extraindo o logaritmo de ambos os lados leva a

$$g(T) = \ln f(T) = \ln A - \frac{Q}{R}\frac{1}{T} \tag{5.6.4}$$

Fazendo uma mudança de variáveis $z = 1/T$, chamando $a = \ln A$, $b = Q/R$ e reescrevendo:

$$g(z) = a + bz \tag{5.6.5}$$

Em que agora buscamos ajustar $g(z)$ aos dados transformados z_j (veja que é necessário calcular $z_j = 1/T_j$ para todos os dados T_j, $1 \leqslant j \leqslant N$). Uma vez obtidos a e b, obtemos os parâmetros originais fazendo $A = e^a$ e $Q = bR$.

O próximo exemplo é o de um monômio em x:

$$f(x) = kx^n \tag{5.6.6}$$

onde k e n são os parâmetros buscados. Novamente, podemos calcular os logaritmos e escrever

$$g(x) = \ln f(x) = \ln k + n \ln x \tag{5.6.7}$$

Fazendo a mudança de variável independente $z = \ln x$ e chamando $a = \ln k$:

$$g(z) = a + nz \tag{5.6.8}$$

E agora podemos ajustar g(z) aos dados transformados, obtendo assim a e n. O parâmetro original k será dado por $k = e^a$.

Mais um exemplo em que tomar os logaritmos consegue linearizar o problema, agora de forma mais impressionante, é quando queremos ajustar a seguinte função:

$$f(t) = 1 - e^{-kt^n} \tag{5.6.9}$$

em que k e n são os parâmetros ajustáveis (ver o Ex. 7). A seguinte sequência de manipulações algébricas dá conta do recado:

$$1 - f(t) = e^{-kt^n} \tag{5.6.10a}$$

$$\frac{1}{1 - f(t)} = e^{kt^n} \tag{5.6.10b}$$

$$\ln \frac{1}{1 - f(t)} = kt^n \tag{5.6.10c}$$

$$\ln \left[\ln \frac{1}{1 - f(t)} \right] = \ln k + n \ln t \tag{5.6.10d}$$

Faça então a mudança de variáveis $z = \ln t$, chame todo o lado esquerdo de $g(z)$ e $a = \ln k$ para reescrever

$$g(z) = a + nz \tag{5.6.11}$$

que é linear nos parâmetros a e n. Uma vez ajustados, $k = e^a$ fornece o parâmetro original.

Um último exemplo, um pouco mais complicado do que os anteriores, é o de uma função racional como a seguinte [5]:

$$f(x) = \frac{c_1 + c_2 x + c_3 x^2}{1 + c_4 x + c_5 x^2} \tag{5.6.12}$$

A ideia aqui é multiplicar ambos os lados da equação acima pelo denominador para escrever

$$f(x) + c_4 x f(x) + c_5 x^2 f(x) = c_1 + c_2 x + c_3 x^2 \tag{5.6.13}$$

e isolar o primeiro $f(x)$ do lado esquerdo:

$$f(x) = c_1 + c_2 x + c_3 x^2 - c_4 x f(x) - c_5 x^2 f(x) \tag{5.6.14}$$

Se rebatizarmos essa função de $g(x)$ (apenas no lado esquerdo!), podemos reescrevê-la como

$$g(x) = c_1 + c_2 x + c_3 x^2 - c_4 x f(x) - c_5 x^2 f(x) \tag{5.6.15}$$

que é linear nos parâmetros c_k ($1 \leqslant k \leqslant 5$). Repare que, neste exemplo, estamos usando um truque bem mais sutil do que os usados nos exemplos anteriores. Sequer foi necessário rebatizar os parâmetros ou transformar a variável independente.

Acredito que os exemplos ilustrados acima cobrem boa parte dos casos práticos mais importantes de funções em que podemos linearizar o problema de ajuste por mínimos quadrados. Alguns outros casos são encontrados nos exercícios ao final do capítulo.

5.7 Ajuste não-linear

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim

rutrum.

Referências

[1] O. Helene. *Método dos Mínimos Quadrados com formalismo matricial*. São Paulo: Editora Livraria da Física, 2006.

[2] W. H. Press, S. A. Teukolsky, W. T. Vetterling e B. P. Flannery. *Numerical Recipes 3rd Edition: The Art of Scientific Computing*. 3ª ed. New York, NY, USA: Cambridge University Press, 2007.

[3] J. Nocedal e S. J. Wright. *Numerical Optimization*. 2nd. New York: Springer, 2006.

[4] J. L. Boldrini, S. I. R. Costa, V. L. Figueiredo e H. G. Wetzler. *Álgebra Linear*. 3ª ed. São Paulo: Harbra, 1980.

[5] A. F. P. de Castro Humes, I. S. H. de Melo, L. K. Yoshida e W. T. Martins. *Noções de Cálculo Numérico*. São Paulo: McGraw-Hill, 1984.

Exercícios

5. 1 — Obtenha a matriz de covariância para o problema visto na seção 5.1, asim como os desvios padrão fornecidos na Eq. (5.1.10).

5. 2 — A tabela abaixo contém notas de listas de exercícios (L) e projeto final (P) de trinta alunos de computação científica:

L	P	L	P	L	P
302	45	323	83	343	83
325	72	337	99	290	74
285	54	337	70	326	76
339	54	304	62	233	57
334	79	319	66	254	45
322	65	234	51	314	42
331	99	337	53	344	79
279	63	351	100	185	59
316	65	339	67	340	75
347	99	343	83	316	45

Faça uma regressão linear e determine as notas L que preveem

(a) $P \geqslant 90$;

(b) $P \geqslant 60$.

5. 3 — Medidas de entalpia molar de mistura ($\Delta\overline{H}_M$) em ligas líquidas In–Sn (índio-estanho) a 773 K, em função da concentraçao molar de In (x_{In}), são fornecidas abaixo:

x_{In}	$\Delta\overline{H}_M$ (J/mol)	x_{In}	$\Delta\overline{H}_M$ (J/mol)
0.10	-183	0.50	-316
0.20	-374	0.79	-165
0.30	-437	0.69	-241
0.40	-383	0.60	-301
0.50	-331	0.90	-87

Usualmente é feito um ajuste por polinômios especiais, chamados polinômios de Redlich-Kister:

$$\Delta H_M = x_{\text{In}} x_{\text{Sn}} \sum_{n=0}^{n_{\max}} L_n (x_{\text{In}} - x_{\text{Sn}})^n \tag{5.7.1}$$

sendo L_n ($n = 0, 1, \ldots, n_{\max}$) parâmetros a serem determinados e $x_{\text{Sn}} = 1 - x_{\text{In}}$. Ajuste os dados para os casos

(a) $n_{\max} = 0$;

(b) $n_{\max} = 1$;

(c) $n_{\max} = 2$.

© Luiz T. F. Eleno, v. 0.3, 2021

(d) $n_{\max} = 3$.

5. 4 — A tabela abaixo exibe dados de atividade solar média anual (x) e precipitação pluvial anual média (y) no observatório de Uccle, Bélgica. Calcule o coeficiente de correlação entre os dados e mostre que não há correlação entre os dados.

Ano	x	y	Ano	x	y	Ano	x	y	Ano	x	y
1901	2.7	700	1925	44.3	1075	1949	134.7	521	1973	38.0	690
1902	5.0	762	1926	63.9	896	1950	83.9	951	1974	34.5	1039
1903	24.4	854	1927	60.0	837	1951	69.4	878	1975	15.5	734
1904	42.0	663	1928	77.8	882	1952	31.5	926	1976	12.6	541
1905	63.5	912	1929	64.9	688	1953	13.9	557	1977	27.5	855
1906	53.8	821	1930	35.7	953	1954	4.4	741	1978	92.5	777
1907	62.0	622	1931	21.2	858	1955	38.0	616	1979	155.4	839
1908	48.5	678	1932	11.1	858	1956	141.7	795	1980	154.6	913
1909	43.9	842	1933	5.7	738	1957	190.2	801	1981	140.5	1016
1910	18.6	990	1934	8.7	707	1958	184.8	834	1982	115.9	800
1911	5.7	741	1935	36.1	916	1959	159.0	560	1983	66.6	689
1912	3.6	941	1936	79.7	763	1960	112.3	962	1984	45.9	931
1913	1.4	801	1937	114.4	900	1961	53.9	903	1985	17.9	758
1914	9.6	877	1938	109.6	711	1962	37.5	862	1986	13.4	946
1915	47.4	910	1939	88.8	928	1963	27.9	713	1987	29.2	908
1916	57.1	1054	1940	67.8	837	1964	10.2	785	1988	100.2	1005
1917	103.9	851	1941	47.5	744	1965	15.1	1073	1989	157.6	639
1918	80.6	848	1942	30.6	841	1966	47.0	1054	1990	142.6	759
1919	63.6	980	1943	16.3	738	1967	93.8	707	1991	145.7	794
1920	37.6	760	1944	9.6	776	1968	105.9	776	1992	94.3	916
1921	26.1	417	1945	33.2	745	1969	105.5	776	1993	54.6	857
1922	14.2	938	1946	92.6	861	1970	104.5	727	1994	29.9	894
1923	5.8	917	1947	151.6	640	1971	66.6	691	1995	17.5	763
1924	16.7	849	1948	136.3	792	1972	68.9	710	1996	8.6	745

Fonte: J. Meeus, *Astronomical Algorithms*, 2nd ed., 1998.

5. 5 — Deseja-se ajustar a função

$$f(\theta) = a \operatorname{sen}(\theta - \theta_0) \tag{5.7.2}$$

a um conjunto de dados experimentais. Linearize o problema colocando-o na forma

$$f(\theta) = A \operatorname{sen}(\theta) + B \cos(\theta)\,, \tag{5.7.3}$$

determinando a relação entre os parâmetros originais (a e θ_0)e os novos parâmetros A e B.

5. 6 — Linearizando o problema, ajuste a função

$$f(x) = \frac{a}{1 + bx^2} \tag{5.7.4}$$

ao seguinte conjunto de dados experimentais:

x	-4	-3	-2	-1	0	1	2	3	4
y	0.158	0.207	0.468	1.097	1.333	0.924	0.510	0.271	0.139

Determine também os desvios padrão estimados para os valores de a e b.

5. 7 — A equação de Johnson-Mehl-Avrami-Kolmogorov (JMAK) descreve a cinética de transformações de fases no estado sólido que ocorrem por um processo conhecido por *nucleação e crescimento*:

$$f(t) = 1 - e^{-kt^n} \tag{5.7.5}$$

onde k e n são constantes e $f(t)$ é a fração do material que já sofreu a transformação de fases no instante $t \geqslant 0$. Durante a recristalização de uma liga de cobre, a fração recristalizada foi medida por metalografia quantitativa, resultando na seguinte tabela:

t (min)	f	t (min)	f
1	0.00081	11	0.50448
2	0.00271	12	0.63679
3	0.00362	13	0.70907
4	0.01411	14	0.86573
5	0.04115	15	0.87770
6	0.05788	16	0.91696
7	0.10016	17	0.93913
8	0.17230	18	0.95511
9	0.24418	19	0.98450
10	0.42060	20	0.99342

Encontre os valores de n e k que melhor descrevem os dados experimentais.

5. 8 — A tabela abaixo mostra a evolução da malha ferroviária mundial durante a Revolução Industrial (fonte: E. Hobsbawm, *The Age of Capital*):

Ano	malha ferroviária (km)
1831	332
1841	8 591
1846	17 424
1851	38 022
1856	68 148
1861	106 886
1866	145 114
1871	235 375
1876	309 641

(a) Ajuste a equação

$$y(x) = ae^{b(x-1831)} \tag{5.7.6}$$

aos dados, sendo $y(x)$ a extensão da malha ferroviária no ano x, e veja que esse ajuste puramente exponencial superestima em muito o valor alcançado em 1900, que foi de 790 294 km de linhas férreas no mundo.

(b) Um ajuste melhor é o chamado *modelo de crescimento logístico*, dado por

$$y(x) = \frac{a}{1 + be^{-k(t-t_0)}} \tag{5.7.7}$$

Esse é um ajuste não-linear e não-linearizável. Use a função `curve_fit` da biblioteca `scipy.optimize` para refazer os cálculos do item (a) usando $t_0 = 1831$.

(c) Acrescente o valor para 1900 ao conjunto de dados experimentais e refaça o ajuste do modelo logístico. Veja que agora a concordância é excelente!

 © Luiz T. F. Eleno, v. 0.3, 2021

Capítulo 6

Interpolação

Capítulo 7

Integração numérica

Capítulo 8

Equações diferenciais ordinárias

Thou still unravish'd bride of quietness,
Thou foster-child of silence and slow time,
Sylvan historian, who canst thus express
A flowery tale more sweetly than our rhyme?

John Keats, *Ode on a Grecian urn*

8.1 Introdução

No capítulo 2, vimos uma série de problemas descritos (ou modelados) por equações diferenciais ordinárias (EDOs). Em todas as situações analisadas, as EDOs vinham das leis de Newton da Mecânica Clássica, acopladas a uma relação constitutiva (a lei de Hooke para uma mola ideal sem atrito). Boa parte da Física é isso: descrever a natureza de um fenômeno observando a taxa de variação (ou seja, as derivadas) das quantidades envolvidas.

Apesar do imenso trabalho envolvido em resolvê-las, ainda assim conseguimos obter naquele capítulo soluções analíticas para todos os problemas apresentados — o que não causa muito espanto: não seriam apresentados se não fosse possível! No entanto, os problemas do Capítulo 2, apesar de extremamente úteis pelos *insights* que fornecem, são altamente idealizados: molas ideais que obedecem a lei de Hooke, sem atrito, etc. Um outro fator complicador: todas as matrizes que vimos naquela ocasião eram constantes, ou seja, não dependiam nem do tempo, nem da posição. Todo o edifício construído sobre a análise por autovalores e autovetores é arruinado caso qualquer constante elástica não seja constante!

Mais uma maneira de complicador nossa vida é o seguinte: considere o problema de EDO com condição inicial abaixo:

$$\frac{dy}{dt} = y'(t) = y(t) + 2t\,, \qquad y(0) = 1 \tag{8.1.1}$$

Em alguma disciplina de cálculo, você aprendeu a resolver esse tipo de problema, cuja solução exata, facilmente encontrada, é

$$y(t) = 3e^t - 2(t+1)\,, \tag{8.1.2}$$

como você pode verificar. Mas basta complicar um pouquinho a EDO,

$$y'(t) = y^2(t) + 2t\,, \qquad y(0) = 1\,. \tag{8.1.3}$$

para que a solução fique extremamente complicada[1], em termos de funções especiais da Física Matemática, como a função Gama e as funções de Bessel — que foram inventadas exatamente para resolver EDOs que não tem solução sem elas! Porém, para EDOs apenas um pouco mais inusitadas, como a seguinte,

$$y'(t) = \operatorname{sen} y(t) + 2t\,, \qquad y(0) = 1\,, \tag{8.1.4}$$

ninguém inventou funções especiais,[2] e é preciso, por exemplo, recorrer a expansões em série de Frobenius, que você encontrou em Física Matemática, para buscar por uma solução e criar, efetivamente, uma nova

[1] não vou mostrá-la aqui para não assustar ninguém, mas você pode visualizá-la digitando `y'[t] == y[t]^2 + 2t, y[0] == 1` no site www.wolframalpha.com.

[2] o www.wolframalpha.com não consegue achar a solução para essa EDO!

função especial. Muitas vezes, essa é uma abordagem desnecessária — ou, pelo menos, trabalhosa demais. Há uma estratégia mais amena e direta: métodos numéricos. É o que veremos nesse capítulo.

Todos os métodos numéricos aproximativos para a solução de EDOS com condição inicial resolvem apenas o seguinte problema:

$$y'(t) = f(t, y(t)), \qquad y(0) = y_0 \tag{8.1.5}$$

onde y_0, o valor da solução $y(t)$ no instante[3] inicial t_0 (usualmente $t_0 = 0$), é uma constante conhecida e $f(t, y)$ é uma função qualquer (mas geralmente contínua) de duas variáveis, também conhecida. Além disso, os métodos não fornecem a função $y(t)$ completa, mas apenas valores em alguns pontos determinados por parâmetros do método. Ou seja, os métodos fornecem a função $y(t)$ para valores discretizados de t.

À primeira vista, resolver a Eq. (8.1.5) pode parecer pouco, já que muitas das EDOs de interesse não são de primeira ordem, e não é sempre que conseguimos isolar a derivada, escrevendo-a em função de t e y. Mas, como veremos, praticamente qualquer outra EDO de interesse prático pode ser colocada numa forma equivalente à da Eq. (8.1.5), bastando generalizá-la usando... (adivinhou!) vetores e matrizes.

8.2 Método de Euler

Voltamos a encontrar o suiço Leonhard Euler, que, até onde se sabe, foi o primeiro a propor um método numérico para resolver EDOs com condição inicial. A ideia de Euler é usar a série de Taylor de $y(t)$,

$$y(t+h) = y(t) + y'(t)h + \frac{1}{2!}y''(t)h^2 + \frac{1}{3!}y'''(t)h^3 + \mathcal{O}(h^4) \tag{8.2.1}$$

Na Eq. (8.2.1), o último termo $\mathcal{O}(h^4)$ usa a chamada notação Grande-O (*Big-O notation*), que significa que a equação tem mais termos, mas o de maior importância depende da quarta potência de h. Mantendo apenas termos até a primeira ordem em h na Eq. (8.2.1), teremos

$$y(t+h) \approx y(t) + y'(t)h \tag{8.2.2}$$

Substituindo agora a Eq. (8.1.5) na Eq. (8.2.1), podemos reescrever a proposta de Euler como

$$y(t+h) \approx y(t) + f(t, y(t))h \tag{8.2.3}$$

A partir daqui, podemos criar um processo iterativo: estabelecemos um passo h e calculamos a função $y(t)$ apenas nos valores t_n de t dados por

$$t_{n+1} = t_n + h = t_0 + nh, \qquad n = 0, 1, 2, \ldots \tag{8.2.4}$$

(em que $t_0 = 0$, geralmente), o que nos leva aos seguintes valores y_n, usando a Eq. (8.2.3):

$$y_{n+1} = y(t_{n+1}) = y(t_n) + hf(t_n, y(t_n)) \tag{8.2.5}$$

ou, mais compactamente,

$$y_{n+1} = y_n + hf_n \tag{8.2.6}$$

onde introduzimos também a notação $f(t_n, y(t_n)) = f_n$.

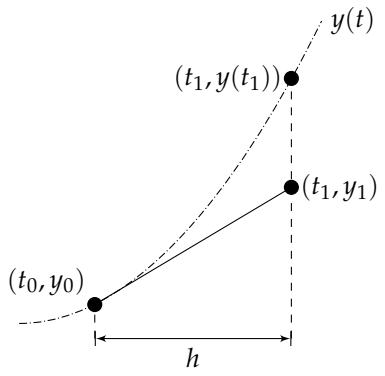


FIGURA 8.1: Uma representação visual do primeiro passo do método de Euler: $y(t_1)$ é aproximado por y_1 pela extrapolação linear seguindo a reta tangente a $y(t)$ em $t = t_0$.

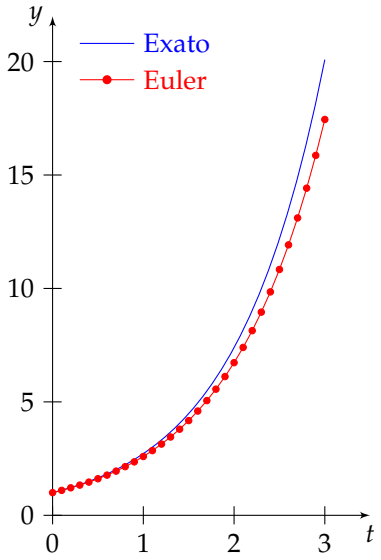
A interpretação geométrica do método de Euler é mostrada na Figure 8.1. Fundamentalmente, o método consiste em aproximar o valor da função $y(t)$ no instante $t_1 = t_0 + h$ pelo valor dado pela extrapolação linear a partir da reta tangente à curva em t_0. O ponto (t_0, y_0) é conhecido da condição inicial, mas esse é o único valor de $y(t)$ conhecido exatamente.

É fácil ver que o erro do método de Euler tende a aumentar à medida em que nos afastamos da condição inicial. Vamos observar isso usando um exemplo: busquemos a solução da EDO da Eq. (8.1.5) quando $f(x, y) = y$, ou seja,

$$y'(t) = y(t), \tag{8.2.7}$$

com a condição inicial $y(0) = 1$. A solução exata é facilmente encontrada: $y(t) = e^t$. Vamos aplicar o método de Euler com um passo $h = 0.1$ e calcular a solução para o intervalo $0 \leqslant t \leqslant 3$. Com a Eq. (8.2.6) montamos

[3] a variável independente t não é necessariamente o tempo, mas vamos usar a expressão *instante* simplesmente por comodidade.



n	t_n	y_n	$y(t_n)$
0	0.0	1.000	1.000
1	0.1	1.100	1.105
2	0.2	1.210	1.221
3	0.3	1.331	1.350
4	0.4	1.464	1.492
5	0.5	1.611	1.649
6	0.6	1.772	1.822
7	0.7	1.949	2.014
8	0.8	2.144	2.226
⋮	⋮	⋮	⋮
27	2.7	13.110	14.880
28	2.8	14.421	16.445
29	2.9	15.863	18.174
30	3.0	17.449	20.086

FIGURA 8.2: Aplicação do método de Euler para a EDO $y'(t) = y(t)$ com $y(0) = 1$, usando $h = 0.1$ no intervalo $0 \leqslant t \leqslant 3$. A última coluna da tabela à direita é a solução exata, $y(t) = e^t$.

a Tabela apresentada na Figura 8.2, representada também graficamente na mesma Figura. A propagação contínua dos erros, à medida em que t aumenta, é mais que evidente nesse exemplo.

Obviamente, podemos tentar diminuir os erros envolvidos no método numérico diminuindo o valor do passo h. A desvantagem disso é que demoramos mais para chegar ao valor $t = 3$ que, talvez, seja o que importa para determinada aplicação. O método de Euler, como vemos, é uma maneira progressiva de se chegar ao resultado, construindo passo a passo a curva $y(t)$. Se quisermos velocidade de cálculo, aumentamos o passo h, mas perdemos precisão ao final; se queremos maior precisão, diminuímos o passo e, com isso, aumentamos o custo (ou seja, o tempo) do cálculo.

Assim, a grande vantagem do método de Euler, sua simplicidade, torna-se também, simultaneamente, seu grande problema: como y_n na Eq. (8.2.6) depende linearmente de h, cada passo introduz um erro proporcional a h. Se o erro dependesse de h^2, por exemplo, teríamos automaticamente uma maior precisão sem alterar h. Na prática, então, ao invés de simplesmente reduzir h, uma segunda opção é a geralmente adotada: utilizar um método mais sofisticado, que garanta maior precisão sem aumentar desnecessariamente o custo computacional. Veremos algumas estratégias a seguir.

8.3 Métodos de Runge-Kutta

Vimos, na seção anterior, que o método de Euler usa a série de Taylor truncada na primeira derivada para obter uma aproximação para a solução $y(t)$ da Eq. (8.1.5). Por outro lado, se o erro de nosso método numérico crescesse com h^2 (e não com h, como no método de Euler), esse por si só seria um grande incremento na precisão: se $h = 0.1$, o erro aumentaria proporcionalmente a $h^2 = 0.01$, dez vezes menos que no método de Euler. Poderíamos então tentar buscar outras aproximações adicionando termos à série de Taylor, termos que dependeriam de h^2, h^3, etc. Foi o que Runge e Kutta fizeram em dois trabalhos no final do século XIX e início do século XX,[4] criando uma hierarquia de métodos cada vez mais precisos que receberam o nome de métodos de Runge-Kutta.

Vamos começar com apenas mais um termo, truncando a série de Taylor no termo quadrático em h^2:

$$y(t+h) = y(t) + y'(t)h + \frac{1}{2}y''(t)h^2 + \mathcal{O}(h^3) \tag{8.3.1}$$

Como fizemos anteriormente, podemos substituir a primeira derivada na Eq. (8.3.1) usando a Eq. (8.1.5): $y' = f(t,y)$. Mas o que fazer com $y''(t)$, a segunda derivada? Ora, podemos calculá-la! Usando o que sabemos da regra da cadeia para derivadas de funções de duas variáveis, podemos, partindo da Eq. (8.1.5), escrever formalmente y'' como

$$y''(t) = \frac{\partial f}{\partial t} + \frac{\partial f}{\partial y}y'(t) = \frac{\partial f}{\partial t} + \frac{\partial f}{\partial y}f(t,y) \tag{8.3.2}$$

[4] Carl David Tolmé Runge (1856–1927) e Martin Wilhelm Kutta (1867–1944), matemáticos alemães.

onde usamos novamente a Eq. (8.1.5) para substituir $y'(t)$ por $f(t,y)$. Com isso, usando as Eqs. (8.1.5) e (8.3.2) reescrevemos a Eq. (8.3.1) como

$$y(t+h) = y(t) + hf(t,y) + \frac{h^2}{2}\left[\frac{\partial f}{\partial t} + \frac{\partial f}{\partial y}f(t,y)\right] \tag{8.3.3}$$

onde deixamos de lado a indicação do erro, $\mathcal{O}(h^3)$. A Eq. (8.3.3) pode ser — e é realmente — usada como um método numérico para a solução da Eq. (8.1.5). No entanto, sua implementação computacional exige que as derivadas parciais da função $f(t,y)$ sejam calculadas. Para isso, ou você as determina manualmente e as fornece ao código, ou você precisa que alguma biblioteca de manipulação algébrica simbólica (o que se chama de *Computer Algebra System*) faça isso por você. Veja que esse passo intermediário está ausente do método de Euler, em que precisamos simplesmente da função $f(t,y)$ e não de suas derivadas. Seria interessante se conseguíssemos essa simplificação também aqui. Runge e Kutta propuseram fazer isso usando a seguinte expressão em substituição à Eq. (8.3.3):

$$y(t+h) = y(t) + h\left[\alpha f(t,y) + \beta f(t+\gamma h, y+\delta h)\right] \tag{8.3.4}$$

onde α, β, γ e δ são constantes a serem determinadas de forma que as Eqs. (8.3.3) e (8.3.4) forneçam o mesmo resultado. Veja que, na Eq. (8.3.4), precisamos, ao invés das duas derivadas parciais de f, de apenas dois de seus valores, calculados em (t,y) e $(t+\gamma h, y+\delta h)$. O computador agradece.

Para encontrar as constantes α, β, γ e δ, vamos primeiramente expandir $f(t+\gamma h, y+\delta h)$ em série de Taylor, truncando-a no primeiro termo:

$$f(t+\gamma h, y+\delta h) \approx f(t,y) + \gamma h\frac{\partial f}{\partial t} + \delta h\frac{\partial f}{\partial y} \tag{8.3.5}$$

o que nos permite reescrever a Eq. (8.3.4) como

$$y(t+h) = y(t) + (\alpha+\beta)hf(t,y) + \beta h^2\left[\gamma\frac{\partial f}{\partial t} + \delta\frac{\partial f}{\partial y}\right] \tag{8.3.6}$$

Comparando agora termo a termo as Eqs. (8.3.3) e (8.3.6), vemos que a escolha mais óbvia que podemos fazer (apesar de não ser a única) é adotar os seguintes valores

$$\alpha = \beta = \frac{1}{2}, \qquad \gamma = 1, \qquad \delta = f(t,y). \tag{8.3.7}$$

Substituindo esses valores de volta na Eq. (8.3.4), temos finalmente a expressão para o método de Runge-Kutta de ordem 2:

$$y(t+h) = y(t) + \frac{h}{2}\left[f(t,y) + f(t+h, y+hf)\right] \tag{8.3.8}$$

onde usamos a abreviação $f = f(t,y)$ para simplificar um pouco a expressão.

Assim, podemos usar a Eq. (8.3.8) como um método interativo, como no método de Euler. Calculamos os valores de $y(t)$ apenas em alguns valores de t igualmente espaçados,

$$t_{n+1} = t_0 + nh, \qquad n = 0, 1, 2, \ldots \tag{8.3.9}$$

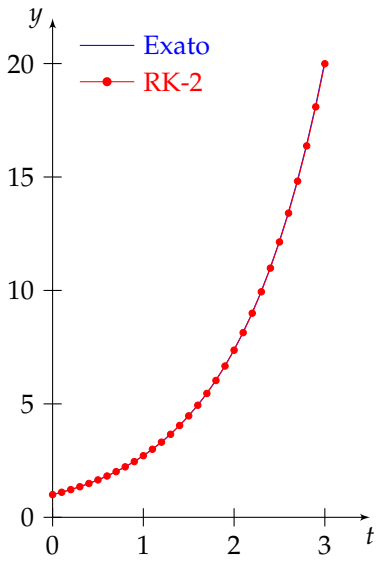
que nos fornecem os valores

$$y_{n+1} = y(t_{n+1}) = y(t_n) + \frac{h}{2}\left(K_1 + K_2\right) \tag{8.3.10}$$

onde K_1 e K_2 são dados por

$$K_1 = f(t_n, y_n) = f_n, \qquad K_2 = f(t_n + h, y_n + hK_1). \tag{8.3.11}$$

Vamos testar o método usando o mesmo exemplo que aquele que aquele visto quando discutimos o método de Euler: $y'(t) = y(t)$, com $y(0) = 1$, cuja solução exata é $y(t) = e^t$. A Figura 8.3 ilustra a aplicação do método de Runge-Kutta de ordem 2 com um passo $h = 0.1$ no intervalo $0 \leqslant t \leqslant 3$. Na escala da Figura 8.3, as duas curvas são indistinguíveis, mas, da tabela que acompanha a acompanha, vemos que o método consegue uma precisão de uma casa decimal até o passo $n = 29$. Compare os resultados da Figura 8.3 com aqueles da Figura 8.2 e veja que conseguimos uma melhoria impressionante com o custo de apenas uma chamada adicional à função $f(t,y)$ por iteração (e mais algumas somas e multiplicações).



n	t_n	K_1	K_2	y_n	$y(t_n)$
0	0.0	—	—	1.000	1.000
1	0.1	1.000	1.100	1.105	1.105
2	0.2	1.105	1.216	1.221	1.221
3	0.3	1.221	1.343	1.349	1.350
4	0.4	1.349	1.484	1.491	1.492
5	0.5	1.491	1.640	1.647	1.649
6	0.6	1.647	1.812	1.820	1.822
7	0.7	1.820	2.002	2.012	2.014
8	0.8	2.012	2.213	2.223	2.226
9	0.9	2.223	2.445	2.456	2.460
⋮	⋮	⋮	⋮	⋮	⋮
27	2.7	13.410	14.751	14.818	14.880
28	2.8	14.818	16.299	16.374	16.445
29	2.9	16.374	18.011	18.093	18.174
30	3.0	18.093	19.902	19.993	20.086

FIGURA 8.3: Aplicação do método de Runge-Kutta de ordem 2 (RK-2) para a EDO $y'(t) = y(t)$ com $y(0) = 1$, usando $h = 0.1$ no intervalo $0 \leqslant t \leqslant 3$. A última coluna da tabela à direita é a solução exata, $y(t) = e^t$.

8.3.1 Métodos de Runge-Kutta de ordem superior

Para melhorar ainda mais a precisão, poderíamos continuar a acrescentar termos à Eq. (8.3.4):

$$\begin{aligned} y(t+h) = y(t) + h\,[&\beta_0 f(t,y) + \beta_1 f(t+\gamma_1 h, y+\delta_1 h) + \\ &\beta_2 f(t+\gamma_2 h, y+\delta_2 h) + \beta_3 f(t+\gamma_3 h, y+\delta_3 h) + \ldots]\,, \end{aligned} \tag{8.3.12}$$

e igualar essa expressão à série de Taylor de ordem correspondente ao número de termos na expressão acima. Lembre que, no método de Runge-Kutta de ordem 2, que acabamos de estudar, igualamos a Eq. (8.3.4), com $f(x,y)$ calculada em dois pontos, à série de Taylor de ordem 2. Poderíamos então continuar na mesma toada e obter uma hierarquia de aproximações cada vez mais precisas. Estes são os chamados Métodos de Runge-Kutta.

Na prática, o o método de Runge-Kutta de ordem 4 é o mais utilizado, por aliar boa precisão à facilidade de implementação e baixo custo computacional. Esse método corresponde a truncar a Eq. (8.3.12) para quatro chamadas à função $f(x,y)$, ou seja, até β_3. Pouparemos o leitor do trabalho de deduzir as equações do método, um processo mecânico e extremamente tedioso. Forneceremos apenas o resultado final, usando a mesma notação que aquela usada nos métodos anteriores. Assim, novamente adotamos um passo h e calculamos a função $y(t)$ apenas nos valores $t_n = t_0 + nh$, obtendo com isso uma sequência de pontos y_n, calculados pela relação recursiva abaixo:

$$y_{n+1} = y_n + \frac{h}{6}\,(K_1 + 2K_2 + 2K_3 + K_4) \tag{8.3.13}$$

em que as grandezas K_1, K_2, K_3 e K_4 são valores da função $f(t,y)$ calculados nos seguintes pontos:

$$K_1 = f(x_n, y_n) \qquad K_2 = f(x_n + h/2, y_n + hK_1/2)$$
$$K_3 = f(x_n + h/2, y_n + hK_2/2) \qquad K_4 = f(x_{n+1}, y_n + hK_3)$$

Para o exemplo que estamos acompanhando até aqui, não há diferença visual entre os métodos de Runge-Kutta de ordem 2 e 4, razão pela qual não apresentaremos o resultado em forma gráfica. No entanto, percebemos pela Tabela 8.1 que o método de ordem 4 garante uma precisão de pelo menos quatro casas decimais até o último ponto! Em relação ao método de ordem 2, o custo aumentou em apenas duas chamadas extras à função $f(x,y)$, além de mais algumas poucas somas e multiplicações, o que continua plenamente aceitável, computacionalmente falando.

8.4 Métodos adaptativos

Os métodos vistos acima, em geral, fazem um bom trabalho mesmo se usados sem outros aprimoramentos. No entanto, eles infelizmente não tem um desempenho tão bom em alguns casos um pouco mais compli-

TABELA 8.1: Aplicação do método de Runge-Kutta de ordem 4 para a EDO $y'(t) = y(t)$ com $y(0) = 1$, usando $h = 0.1$ no intervalo $0 \leqslant t \leqslant 3$. A última coluna da tabela à direita é a solução exata, $y(t) = e^t$.

n	t_n	K_1	K_2	K_3	K_4	y_n	exato
0	0.0					1.00000	1.00000
1	0.1	1.000	1.050	1.053	1.105	1.10517	1.10517
2	0.2	1.105	1.160	1.163	1.221	1.22140	1.22140
3	0.3	1.221	1.282	1.286	1.350	1.34986	1.34986
4	0.4	1.350	1.417	1.421	1.492	1.49182	1.49182
5	0.5	1.492	1.566	1.570	1.649	1.64872	1.64872
6	0.6	1.649	1.731	1.735	1.822	1.82212	1.82212
7	0.7	1.822	1.913	1.918	2.014	2.01375	2.01375
8	0.8	2.014	2.114	2.119	2.226	2.22554	2.22554
9	0.9	2.226	2.337	2.342	2.460	2.45960	2.45960
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
27	2.7	13.464	14.137	14.171	14.881	14.87970	14.87973
28	2.8	14.880	15.624	15.661	16.446	16.44461	16.44465
29	2.9	16.445	17.267	17.308	18.175	18.17410	18.17415
30	3.0	18.174	19.083	19.128	20.087	20.08549	20.08554

cados, mas importantes na prática. Por exemplo, considere a função $y(t)$, mostrada esquematicamente na Figura 8.4, que é a solução de uma EDO do tipo que estamos considerando, dada pela Eq. (8.1.5). Adotando qualquer um dos métodos numéricos acima, com um passo h sempre constante, como o mostrado na Figura 8.4, fazemos — adotando por um momento a linguagem da Estatística — uma amostragem não representativa do comportamento da função no intervalo considerado. Veja que, entre t_1 e t_2, a curva muda rapidamente de inclinação (passa, na verdade, por um máximo), o que significa que a derivada $y'(t)$ varia fortemente para instantes nessa vizinhança. Entre t_3 e t_4 há outra região de mudança súbita de inclinação, ainda que com menor intensidade. Por outro lado, acima de t_4, a curva segue suave, sem grandes alterações no valor da derivada. É provável que nossos métodos numéricos anteriores falhem miseravelmente na região $t_0 < t < t_4$, com o passo h como na Figura 8.4, pois não consegue dar conta de toda essa variação na derivada. Além disso, como vimos, o erro tende a crescer à medida em que avançamos para t_n maiores. Como já começamos errando feio no começo, o erro nos instantes posteriores, ainda que a curva seja suave nessas regiões, só vai aumentar. Assim, mesmo com uma estratégia numérica de alta precisão como o método de Runge-Kutta, há alguns casos em que ele pode falhar.

Em virtude do que foi discutido no parágrafo anterior, poderíamos mitigar o problema da precisão diminuindo o passo h, fazendo assim uma melhor amostragem dos pontos da função $y(t)$. De fato, essa é uma solução possível, que certamente trará uma diminuição do erro do método numérico. No entanto, esse solução não é ótima: veja que, na Figura 8.4, as regiões acima de t_3 não têm assim tanta necessidade de um passo menor, como a região $t < t_3$. Idealmente, temos apenas que diminuir o passo na região $t_0 \leqslant t \leqslant t_4$, para dar conta da grande variação da derivada. No entanto, isso não é assim tão fácil: não conhecemos de antemão a função $y(t)$ para saber onde escolher um passo menor! Sabemos apena $y'(t)$ e, em princípio, poderíamos calcular a derivada segunda $y''(t)$ para saber onde a derivada primeira é mais problemática. Mas isso é muito custoso computacionalmente — é exatamente para evitar isso que existem os métodos de Runge-Kutta, afinal de contas! A solução para o problema é então um compromisso entre a precisão e o custo computacional, que é feito nos chamados *métodos adaptativos*, às vezes também ditos *métodos de malha adaptativa*.

No método adaptativo de Runge-Kutta-Fehlberg (RKF), o passo h é inicialmente adotado como fixo, como na Figura 8.4. Estabelecemos também uma *tolerância* η. Digamos que já usamos o método para calcular o

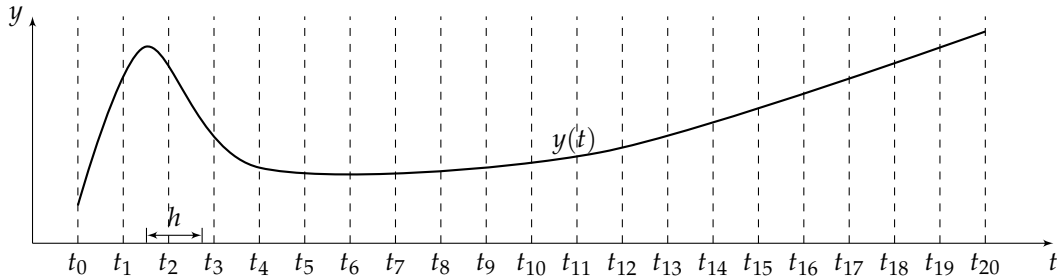


FIGURA 8.4: Uma função $y(t)$ com algumas regiões mais suaves e outras mais sinuosas.

valor y_n em $t = t_n$. Para calcular o valor y_{n+1}, usamos dois métodos padrão, digamos, os métodos de Runge-Kutta de ordem 4 e 5 (RK-4 e RK-5). Caso a diferença entre os dois valores encontrados seja menor que η, aceitamos o passo e vamos para o próximo valor de t. Mas, caso a diferença seja maior que η, isso é uma indicação de que a derivada varia muito entre t_n e t_{n+1}. Então, recuamos em t e recalculamos a função y num valor intermediário, usando provisoriamente um passo $h/2$, por exemplo, o que nos dará dois outros valores, de acordo com os métodos RK-4 e RK-5. Se novamente a diferença entre eles for maior que η, recuamos mais um pouco, com um passo $h/4$, continuando a fazer isso até que a diferença entre os valores seja menor que a tolerância η. E continuamos assim até alcançar o instante t_{n+1}, quando começamos novamente todo o processo até atingir t_{n+2}, etc., até o último instante.

O método de Runge-Kutta-Felhberg é um dos mais utilizados na prática. Implementações desse método podem ser encontradas na maior parte das bibliotecas numéricas, seja qual for a linguagem implementada, como na biblioteca `scipy.integrate`. Geralmente, os dois métodos convencionais usados para as comparações são os de Runge-Kutta de ordens 4 e 5, razão pela qual geralmente adota-se o acrônimo RKF45 (ou variações sobre o tema) para mencioná-lo.

Debaixo do guarda-chuva que é a denominação Métodos Adaptativos, existem muitas outras estratégias, com diferentes graus de complexidade e precisão. Não vemos motivos para descrevê-los em detalhes nestas notas de aula. Detalhes podem ser encontrados, por exemplo, em Press et al. [1].

8.5 Sistemas de EDOs

Qualquer um dos métodos numéricos que vimos acima funciona igualmente bem para problemas envolvendo sistemas de EDOs na seguinte forma:

$$y_1'(t) = f_1(t, y_1(t), y_2(t), \ldots, y_m(t)) \tag{8.5.1a}$$
$$y_2'(t) = f_2(t, y_1(t), y_2(t), \ldots, y_m(t)) \tag{8.5.1b}$$
$$\vdots$$
$$y_m'(t) = f_m(t, y_1(t), y_2(t), \ldots, y_m(t)) \tag{8.5.1c}$$

onde as funções $f_j(t, y_1, y_2, \ldots, y_m)$ ($j = 1, 2, \ldots, m$) são conhecidas e buscamos as m funções $y_j(t)$ que satisfazem as condições iniciais $y_j(0) = y_{j0}$. Trata-se de uma extensão natural do problema original dado pela Eq. (8.1.5), situação em que $m = 1$, ou seja, apenas uma EDO relacionando uma função desconhecida (y) à variável independente (t).

E por que os métodos numéricos conseguem dar conta desse problema? O segredo está em usar matrizes! Veja que não é nenhum trabalho hercúleo colocar as Eqs. (8.5.1) na seguinte forma matricial:

$$Y'(t) = F(t, Y(t)) \tag{8.5.2}$$

onde $Y(t)$ e $Y'(t)$ são vetores coluna, respectivamente, com as funções $y_j(t)$ e suas derivadas $y_j'(t)$:

$$Y(t) = \begin{pmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_m(t) \end{pmatrix} \quad \Rightarrow \quad Y'(t) = \begin{pmatrix} y_1'(t) \\ y_2'(t) \\ \vdots \\ y_m'(t) \end{pmatrix} \tag{8.5.3}$$

e $F(t, Y(t))$ é outro vetor coluna com as funções $f_j(t, y_j(t))$:

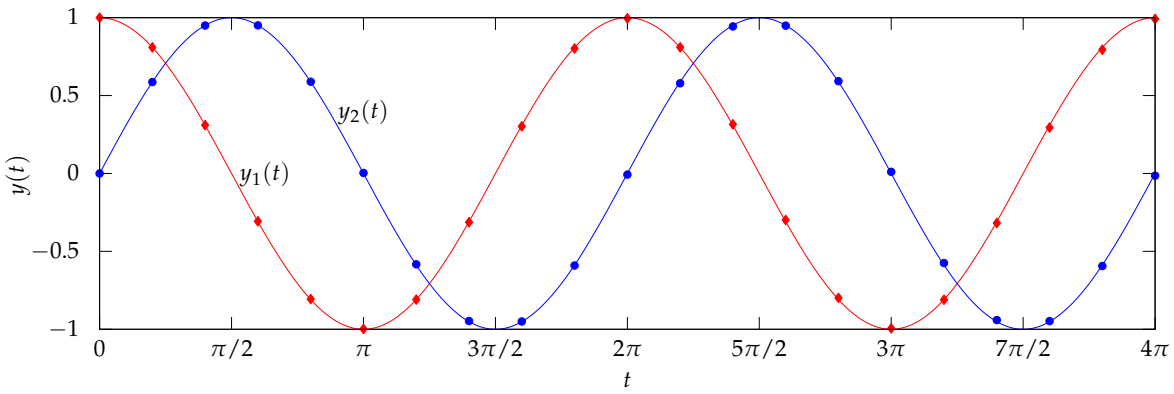
$$F(t, Y(t)) = \begin{pmatrix} f_1(t, y_1(t), y_2(t), \ldots, y_m(t)) \\ f_2(t, y_1(t), y_2(t), \ldots, y_m(t)) \\ \vdots \\ f_m(t, y_1(t), y_2(t), \ldots, y_m(t)) \end{pmatrix} \tag{8.5.4}$$

Repare que o formato da Eq. (8.5.2) não é em nada diferente daquela da Eq. (8.1.5)! Obviamente, precisamos dizer ao nosso método numérico que temos matrizes envolvidas na conversa mas, fora esse detalhe, todo o resto funciona normalmente.

Vejamos um exemplo. Considere o seguinte sistema de EDOs envolvendo duas variáveis:

$$y_1'(t) = -y_2(t) \tag{8.5.5a}$$
$$y_2'(t) = y_1(t) \tag{8.5.5b}$$

 © Luiz T. F. Eleno, v. 0.3, 2021



		RK-4				RK-4		exato	
n	t_n	y_1	y_2	n	t_n	y_1	y_2	$y_1(t_n)$	$y_2(t_n)$
0	0	1.	0.					1.	0.
1	$\pi/5$	0.8091	0.587	11	$11\pi/5$	0.8099	0.5789	0.809	0.5878
2	$2\pi/5$	0.3101	0.9498	12	$12\pi/5$	0.3155	0.9438	0.309	0.9511
3	$3\pi/5$	−0.3066	0.9505	13	$13\pi/5$	−0.2987	0.9488	−0.309	0.9511
4	$4\pi/5$	−0.806	0.5891	14	$14\pi/5$	−0.7986	0.5923	−0.809	0.5878
5	π	−0.998	0.0035	15	3π	−0.9939	0.1050	−1.	0.
6	$6\pi/5$	−0.8095	−0.5829	16	$16\pi/5$	−0.8103	−0.5749	−0.809	−0.5878
7	$7\pi/5$	−0.3128	−0.9468	17	$17\pi/5$	−0.3182	−0.9408	−0.309	−0.9511
8	$8\pi/5$	0.3027	−0.9497	18	$18\pi/5$	0.2948	−0.9479	0.309	−0.9511
9	$9\pi/5$	0.8023	−0.5907	19	$19\pi/5$	0.7949	−0.5939	0.809	−0.5878
10	2π	0.9959	−0.007	20	4π	0.9918	−0.0140	1.	0.

FIGURA 8.5: Solução do problema dado pelas Eqs. (8.5.5) e (8.5.6). A solução exata corresponde às linhas cheias, ao passo que a solução numérica pelo método de Runge-Kutta de ordem 4 (RK-4), com um passo $h = \pi/5$, é indicada pelos pontos.

com as condições iniciais

$$y_1(0) = 1, \quad y_2(0) = 0 \tag{8.5.6}$$

A solução exata do problema é praticamente imediata: $y_1(t) = \cos t$ e $y_2(t) = \operatorname{sen} t$. Para ilustrar a resolução numérica, vamos aplicar o método de Runge-Kutta de ordem 4 à solução do problema no intervalo $0 \leqslant t \leqslant 2\pi$ com um passo $h = \pi/5$. Vamos calcular apenas o primeiro passo como ilustração, pois os demais passos funcionam exatamente da mesma maneira.

Já que pretendemos usar o método de Runge-Kutta de ordem 4, vamos começar investigando o que temos que fazer com a Eq. (8.3.13). Como precisamos substituir as funções $y_j(t)$ pelo vetor coluna $Y(t)$ dado pela Eq. (8.5.3), basta então fazer

$$Y_{n+1} = Y(t_{n+1}) = Y_n + \frac{h}{6}\left(K_1 + 2K_2 + 2K_3 + K_4\right) \tag{8.5.7}$$

onde, assim como as outras grandezas, os Ks são praticamente idênticos à definição anterior — no caso, dada pelas Eqs. (8.3.14) substituindo as funções $f(t, y)$ pelo vetor coluna $F(t, Y)$ da Eq. (8.5.4).

Usando o método de Runge-Kutta de ordem 4 modificado para trabalhar com matrizes e com o passo $h = \pi/5$, conseguimos o resultado mostrado na Figura 8.5. Veja que, apesar de, apenas pelo gráfico mostrado na Figura 8.5, não conseguirmos enxergar muito bem o erro cometido pela aproximação numérica, da tabela que a acompanha percebemos que, como esperado, vamos perdendo precisão à medida em que nos afastamos da condição inicial. Assim, o erro após encerrado o primeiro período (ou seja, em $t = 2\pi$) é de apenas 0.04% em y_1 e de 0.7% e, y_2 (em módulo). Já em $t = 4\pi$, esses mesmos erros passam para 0.8% e 1.4%, respectivamente. Mesmo assim, o erro é tão pequeno que os pontos na Figura 8.5 parecem cair exatamente sobre as curvas da solução exata.

Em resumo: além do fato óbvio de ser necessário alterar a implementação do código para lidar com vetores (matrizes coluna, na verdade), nada precisa ser modificado no algoritmo. Isso facilita enormemente o trabalho, principalmente para EDOs de ordens superiores, que são mais importantes em Física. A 2ª lei de Newton da Mecânica Clássica e a equação de Schrödinger independente do tempo da Mecânica Quântica, por exemplo, são equações de grau 2, ainda que esta última seja, na verdade, um problema de condições de contorno, e não de condições iniciais.

8.5.1 O atrator de Lorenz

Você já ouviu falar do *efeito borboleta*? O termo surgiu após o trabalho divisor de águas do meteorologista estadunidense Edward N. Lorenz [2]. Há um filme de 2004 chamado, precisamente, *O Efeito Borboleta* (*The Butterfly Effect*) que começa com a seguinte citação:

> *It has been said that something as small as the flutter of a butterfly's wing can ultimately cause a typhoon halfway around the world – Chaos Theory.*[5]

Ed Lorenz começou usando *gaivota* no lugar de *borboleta*, ainda em 1963 [3]. No entanto, seguindo algumas sugestões de colegas, logo fez a troca para aumentar o impacto dramático e poético da sentença. Há um outro filme, ainda mais antigo e certamente mais famoso, *Jurassic Park* (1993), em que o galanteador matemático interpretado por Jeff Goldblum usa uma explicação do efeito borboleta para tentar xavecar, sem sucesso, a bela e precavida arqueóloga (Laura Dern).

O que é ainda mais interessante é que o chamado *Atrator de Lorenz*, visto na Figura 8.6 e a ser explicado em breve, realmente parece uma borboleta! Usando o *script* em python fornecido no material de apoio, você pode observá-lo de diversos ângulos e magnificações, ou alterar as condições iniciais e outros parâmetros de cálculo.

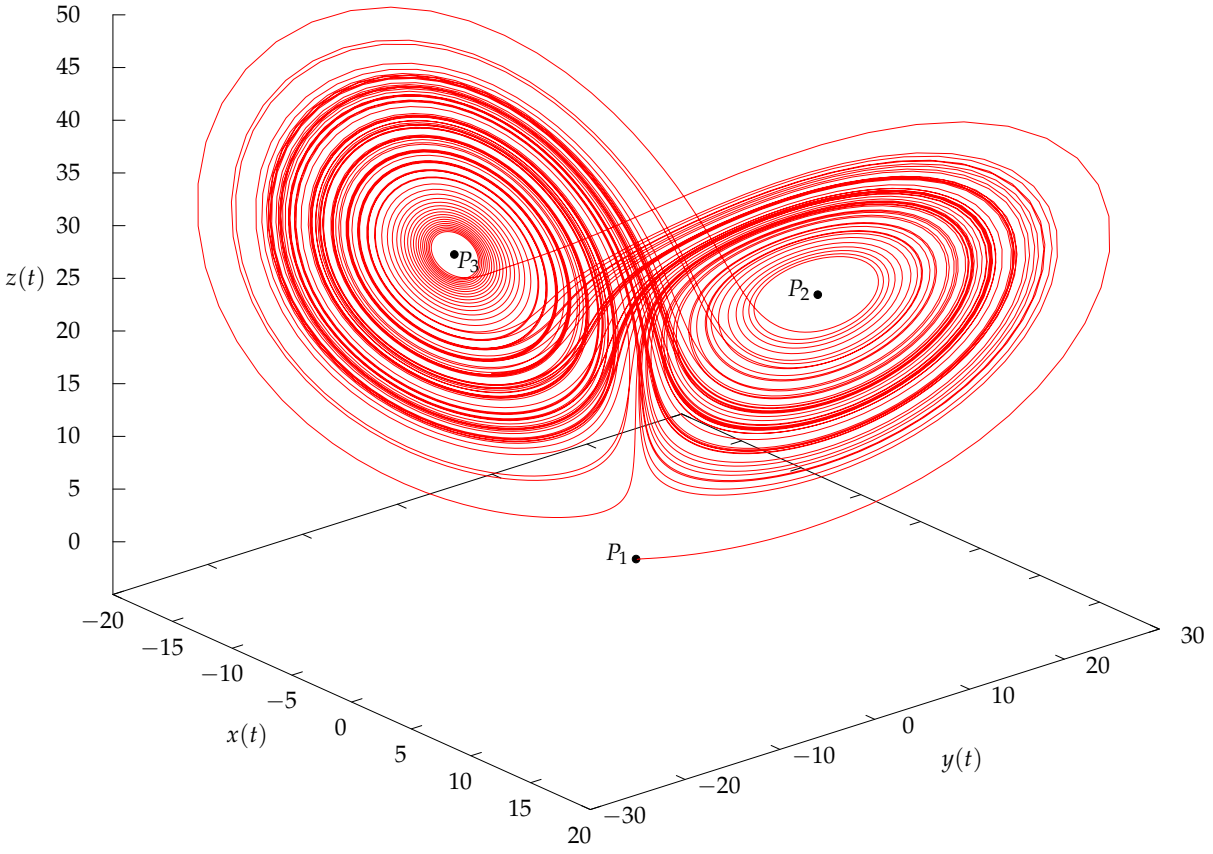


FIGURA 8.6: O atrator de Lorenz com $s = 10$, $r = 28$, $b = 8/3$, para $0 \leqslant t \leqslant 50$ e com as condições de contorno $x(0) = 0.1$, $y(0) = 0$ e $z(0) = 0$. Os pontos P_1, P_2 e P_3 são os pontos estacionários das Eqs. (8.5.8).

O *efeito borboleta* é mais um exemplo de um conceito científico que entrou para a cultura pop do final do século XX para explicar o efeito gigante e imprevisível de qualquer minúscula alteração a que um sistema caótico, como o clima, está sujeito. Cientistas do século XIX e mesmo antes conseguiram descrever com precisão de relógio suiço o movimento de planetas, satélites e cometas, usando condições iniciais e calculando corretamente a dinâmica dos astros por períodos bastante longos. Mesmo fenômenos como eclipses têm a sua periodicidade, chamada de *saros*, com aproximadamente 18 anos, 11 dias e 8 horas, conhecido já pelos antigos babilônios alguns séculos a.C. Foi tão grande o sucesso da Mecânica Clássica em interpretar tais fenômenos usando EDOs que, até a metade do século XX, acreditava-se que a dificuldade em se entender problemas complexos, como o clima, vinha unicamente do enorme número de variáveis. A proliferação dos computadores era uma promessa de trazer tais problemas, até então intratáveis, à alçada de modelos mais complexos que dessem conta de tanta complexidade.

[5] Diz-se que algo tão pequeno quanto o agitar das asas de uma borboleta pode eventualmente causar um tufão do outro lado do mundo – Teoria do Caos.

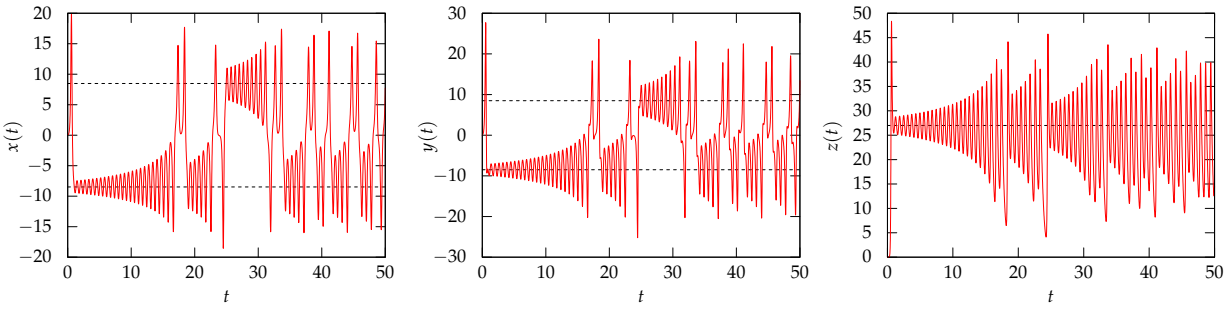


FIGURA 8.7: As três funções $x(t)$, $y(t)$ e $z(t)$ do atrator de Lorenz com $s = 10$, $r = 28$, $b = 8/3$, para $0 \leqslant t \leqslant 50$ e com as condições de contorno $x(0) = 0.1$, $y(0) = 0$ e $z(0) = 0$. As linhas tracejadas horizontais indicam as coordenadas dos pontos estacionários.

Ed Lorenz jogou a sombra de uma nuvem pesada sobre tais esperanças, ainda que tenha criado um novo céu aberto para a Matemática, Física e Engenharia: a *Teoria do Caos*. Num sistema (de EDOs, por exemplo) em que o *comportamento caótico* emerge, qualquer pequena alteração nas condições iniciais tem um enorme efeito sobre a sua evolução temporal, de forma a torná-lo imprevisível — ou, para usar um termo mais adequado, *impredizível*. É o que Ed Lorenz percebeu estudando fenômenos climáticos a partir de modelos de hidrodinâmica. No seu artigo original [2], Lorenz reduziu um sistema de doze EDOs não lineares para um com apenas três:

$$\dot{x} = s(y - x) \tag{8.5.8a}$$
$$\dot{y} = rx - y - xz \tag{8.5.8b}$$
$$\dot{z} = xy - bz \tag{8.5.8c}$$

onde $x = x(t), y = y(t), z = z(t)$ são três funções do tempo t e os parâmetros s, r e b são geralmente escolhidos como $s = 10$, $r = 28$, $b = 8/3$. Apesar de lecionar às vezes a disciplina de Fenômenos de Transporte, minha área de atuação em pesquisa não é esse assunto, e muito menos clima e sua previsão, mas isso pouco importa aqui. Basta saber que $x(t)$ está relacionado a alguma medida da inversão da convecção, ao passo que $y(t)$ e $z(t)$ medem os gradientes de temperatura na horizontal e na vertical, respectivamente. Já os parâmetros s, r e b vêm de propriedades do fluido, da geometria do recipiente que o contém — no caso da atmosfera, o próprio planeta — e do fluxo de calor. Aqui, no entanto, o interesse está em estudar a dinâmica do sistema de Lorenz das Eqs. (8.5.8), sem a preocupação de interpretá-lo fisicamente — até porque ele foi tão simplificado por Lorenz que já não consegue representar fidedignamente a atmosfera, como você pode imaginar.

Pois bem. A Figura 8.6 mostra a solução numérica usando o método padrão em python, a função `odeint` da biblioteca `numpy.integrate`, que, na verdade, é baseada numa antiga implementação em Fortran chamada ODEPACK (https://www.netlib.org/odepack).[6] Mas você pode usar o método de Runge-Kutta de ordem 4, ou o método de Runge-Kutta-Fehlberg, com o mesmo resultado. Além disso, as condições iniciais usadas foram $x(0) = 0.1$, $y(0) =$ e $z(0) = 0$ e acompanhamos a evolução para tempos no intervalo $0 \leqslant t \leqslant 100$ (unidades arbitrárias).

Na Figura 8.7 estão representadas separadamente as três funções $x(t)$, $y(t)$ e $z(t)$ (as três coordenadas da Figura 8.6). Talvez fique mais claro olhando essa figura que o sistema passa de um lóbulo a outro (ou de uma asa da borboleta à outra) do atrator de forma aparentemente caótica — ainda que o sistema seja determinístico, uma vez que veio de um sistema de EDOs muito bem determinado, dado pelas Eqs. (8.5.8). No entanto, há uma sensibilidade extrema às condições iniciais: pequenas alterações na configuração do sistema no instante $t = 0$ levam a uma imprevisibilidade na posição (x, y, z) num instante t posterior. É como a farinha na massa do pão: dois grãos de farinha, inicialmente vizinhos antes da mistura, podem acabar em posições opostas da fornada.

Você tem alguma ideia do significado dos pontos P_1, P_2 e P_3 na Figura 8.6? Eles são os chamados *pontos estacionários* do sistema de EDOS da Eqs. (8.5.8). Um ponto estacionário é aquele no qual as derivadas das funções envolvidas se anulam identicamente. No caso do atrator de Lorenz, isso significa que, nos pontos P_j ($j = 1, 2, 3$), teremos $x' = y' = z' = 0$ (você consegue determinar as coordenadas desses pontos para o atrator de Lorenz?). Assim, em primeiro lugar, se o sistema começa num desses pontos já na condição inicial, ele nunca mais sairá de lá, pois, por exemplo, $x' = 0$ significa que x não varia com t (e analogamente para y e z). Em segundo lugar, há situações em que o sistema tende a algum ponto estacionário para $t \to \infty$. Nesses casos, dizemos que o sistema começa num *regime transiente* até atingir uma condição de *estado estacionário*. Isso é muito comum no tratamento por EDOs da cinética de reações químicas ou de formação de misturas. O atrator de Lorenz da Figura 8.6, por outro lado, não chega ao estado estacionário, mas fica oscilando de forma

[6] mais recentemente, um novo método foi acrescentado à biblioteca `scipy`: a função `solve_ivp`.

não-periódica — *caoticamente* — ao redor de dois deles (P_2 e P_3), apesar de ter começado numa condição de contorno coladíssima ao ponto estacionário P_1. É o que eu vinha dizendo sobre sensibilidade às condições iniciais: pequenas alterações no problema podem levar a efeitos completamente diferentes.

8.6 EDOs de ordem superior

Já fizemos bastante avanço na solução numérica de EDOs: conseguimos resolver sistemas de EDOs com condições inicias, no formato da Eq. (8.5.1), o que nos permite modelar problemas importantes não só em Física e Engenharia, mas em muitas outras áreas, como Biologia e Economia. Mas, falando especificamente do nosso *métier*, não sabemos ainda como resolver numericamente, por exemplo, problemas em Mecânica Classica... Veja que a 2ª Lei de Newton é uma EDO de grau 2, pois envolve a aceleração, que é a taxa de variação da velocidade (ou seja, a derivada segunda da posição). Portanto, se quisermos usar os algoritmos numéricos que vimos anteriormente nesse capítulo, precisamos adaptar EDOs envolvendo derivadas de ordem maior ou igual a 2 para que se encaixem no padrão geral do problema, ou seja, as Eqs. (8.5.1).

Mas já sabemos o que fazer, pois fizemos a mesma coisa logo de cara na seção 2.1. Basta apelidar as derivadas de ordem maior que 1 usando outros nomes, acrescentando assim novas variáveis ao problema. Obviamente, se o problema era unidimensional, passará a ser tratado como um problema envolvendo matrizes e vetores — mas, a essa altura, você já não deve se assustar com isso.

Para ficar mais claro, vamos analisar o seguinte problema: o de uma partícula presa a uma mola de constante k e amortecida por uma força de atrito de um meio viscoso com coeficiente ϵ, como na seção 2.2, mas sentindo também uma força externa senoidal da forma $F_{\text{ext}} = a_0 \operatorname{sen} \beta t$, em que a_0 e β são constantes conhecidas. A 2ª Lei de Newton nos diz que a dinâmica da partícula é dada por

$$m\ddot{x} = -kx - \epsilon\dot{x} + a_0 \operatorname{sen} \beta t \,. \tag{8.6.1}$$

Para colocar o sistema na forma da Eqs. (8.5.2), vamos usar a velocidade $v = \dot{x}$ como se fosse uma outra variável, além da posição $x(t)$, escrevendo assim o vetor $Y(t)$ como

$$Y(t) = \begin{pmatrix} x \\ v \end{pmatrix} \qquad \Rightarrow \qquad \dot{Y}(t) = \begin{pmatrix} \dot{x} \\ \dot{v} \end{pmatrix} \tag{8.6.2}$$

Já sabemos que $\dot{x} = v$; basta então usar $\dot{v} = \ddot{x}$ e a Eq. (8.6.1) para escrever

$$\dot{Y}(t) = \begin{pmatrix} v \\ -(k/m)x - (\epsilon/m)v + (a_0/m) \operatorname{sen} \beta t \end{pmatrix} \tag{8.6.3}$$

ou, de forma um pouco mais elegante,

$$\dot{Y}(t) = \begin{pmatrix} 0 & 1 \\ -k/m & -\epsilon/m \end{pmatrix} Y(t) + \begin{pmatrix} 0 \\ (a_0/m) \operatorname{sen} \beta t \end{pmatrix} \tag{8.6.4}$$

Poderíamos seguir a metodologia do capítulo 2 e resolver o problema algebricamente — coisa que você já viu nos cursos de Física Básica, ainda que provavelmente sem usar matrizes. Você está convidado a brincar com o problema dessa maneira, o que é bastante instrutivo para entender a origem da técnica de separar a solução do problema em uma solução da equação homogênea, ignorando a força externa (já a encontramos na seção 2.2) e em uma solução particular da EDO completa. No entanto, a Eq. (8.6.4) está no formato apropriado para a solução numérica, usando qualquer uma das técnicas que vimos anteriormente — o método de Runge-Kutta de ordem 4, por exemplo.

Para resolver o problema, naturalmente precisamos das condições iniciais e dos valores das constantes do problema. Vamos então adotar $m = 1$, $k = 0.5$, $\epsilon = 0.25$, $a_0 = 1$ e $\beta = 2$ e usar a condição inicial $x(0) = 0$, $v(0) = 0$, ou seja, a partícula está inicialmente em repouso e a mola sem qualquer deformação.

Usando o método de Runge-Kutta de ordem 4, chegamos ao resultado mostrado na Figura 8.8. Como esperado, o oscilador começa a se mover pela ação da força externa, ao mesmo tempo em que briga com a lei de Hooke, que ora o ajuda, ora dificulta seu movimento. Com o tempo, esse regime transiente dá lugar ao regime permanente, em que a frequência de oscilação é completamente ditada pela frequência β da força externa aplicada, combinada ao amortecimento do meio viscoso. Você está convidado a (re)ver a solução algébrica exata em seu livro favorito de Física Básica e conferir que a solução numérica da Figura 8.8 está correta com excelente aproximação. Para efeitos de comparação, a solução exata do problema é

$$x(t) = e^{-\eta t}(a_0 \operatorname{sen} \delta t + b_0 \cos \delta t) + a \operatorname{sen} \beta t + b \cos \beta t \tag{8.6.5}$$

 © Luiz T. F. Eleno, v. 0.3, 2021

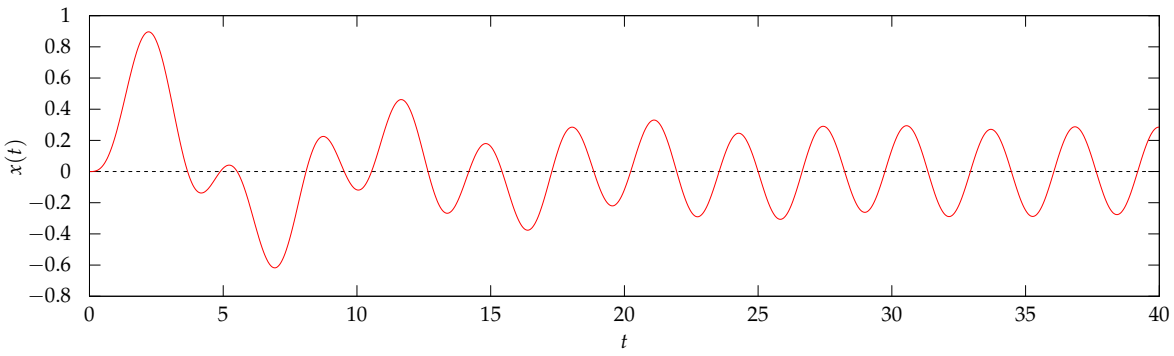


FIGURA 8.8: Solução numérica do problema do oscilador harmônico amortecido e forçado da Eq. (8.6.1) com $m = 1$, $k = 0.5$, $\epsilon = 0.25$, $a_0 = 1$ e $\beta = 2$, para a condição inicial $x(0) = 0$, $v(0) = 0$. A solução foi obtida com uma implementação do método de Runge-Kutta de ordem 4.

sendo

$$\eta = \frac{1}{8}\,, \quad \delta = \frac{\sqrt{31}}{8}\,, \quad a_0 = \frac{113}{25\sqrt{31}}, \quad b_0 = \frac{1}{25}\,, \quad a = -\frac{7}{25}\,, \quad b = -\frac{1}{25}\,. \tag{8.6.6}$$

8.6.1 Problemas celestiais

Vamos investigar agora um exemplo bastante clássico, retirado da chamada *Mecânica Celeste*,[7] que estuda o movimento de planetas, satélites, cometas e outros astros que podemos observar no céu. O problema que queremos usar como exemplo é o *problema de dois corpos*: dois astros (uma estrela e um planeta, por exemplo) atraídos entre si pela Lei da Gravitação de Newton, segundo a qual cada um dos astros, descritos como massas pontuais m_1 e m_2, sentem forças $\vec{F}_1$ e F_2 dadas, respectivamente, por

$$\vec{F}_1 = +G\frac{m_1 m_2}{|\vec{r}_{12}|^2}\,\hat{r}_{12}\,, \qquad \vec{F}_2 = -G\frac{m_1 m_2}{|\vec{r}_{12}|^2}\,\hat{r}_{12} \tag{8.6.7}$$

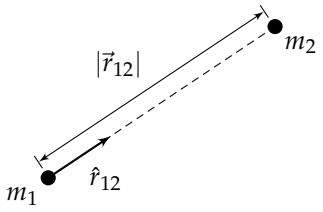


FIGURA 8.9: Duas massas m_1 e m_2 separadas por uma distância $|\vec{r}_{12}|$ e mutuamente atraídas segundo a Lei da Gravitação de Newton.

em que, como na Figura 8.9, $|\vec{r}_{12}|$ é a distância entre os objetos e $\hat{r}_{12}$ é o vetor unitário (também chamado de versor) na mesma direção e sentido que $\vec{r}_{12}$.

Combinando a Lei da Gravitação com a 2ª Lei de Newton, as equações do movimento para as massas m_1 e m_2 são dadas pelas equação abaixo:

$$m_1\ddot{\vec{r}}_1 = +G\frac{m_1 m_2}{|\vec{r}_{12}|^2}\,\hat{r}_{12} \tag{8.6.8}$$

$$m_2\ddot{\vec{r}}_2 = -G\frac{m_1 m_2}{|\vec{r}_{12}|^2}\,\hat{r}_{12} \tag{8.6.9}$$

onde $\vec{r}_1$ e $\vec{r}_2$ são as posições dos astros 1 e 2, respectivamente, no instante t. Do princípio da conservação do momento angular, e percebendo que temos um problema de forças centrais, sabemos que o movimento se dará no plano contendo as duas massas e suas velocidades no instante inicial. Podemos então simplificar o problema e nos restringir ao movimento nesse plano que, por conveniência, chamaremos de xy. As posições das partículas serão então indicadas por $\vec{r}_1 = (x_1, y_1)$ e $\vec{r}_2 = (x_2, y_2)$. Com isso teremos, na verdade, quatro EDOs, uma para cada uma das duas componentes das posições das duas partículas:

$$m_1\ddot{x}_1 = +G\frac{m_1 m_2}{|\vec{r}_{12}|^2}\,\frac{x_2 - x_1}{|\vec{r}_{12}|}\,, \qquad m_1\ddot{y}_1 = +G\frac{m_1 m_2}{|\vec{r}_{12}|^2}\,\frac{y_2 - y_1}{|\vec{r}_{12}|}\,, \tag{8.6.10a}$$

$$m_2\ddot{x}_2 = -G\frac{m_1 m_2}{|\vec{r}_{12}|^2}\,\frac{x_2 - x_1}{|\vec{r}_{12}|}\,, \qquad m_2\ddot{y}_2 = -G\frac{m_1 m_2}{|\vec{r}_{12}|^2}\,\frac{y_2 - y_1}{|\vec{r}_{12}|} \tag{8.6.10b}$$

Cancelando as massas que aparecem nos dois lados do sinal de igualdade e juntando os termos no denominador, simplificamos as Eqs. (8.6.10) para

$$\ddot{x}_1 = +G\frac{m_2}{|\vec{r}_{12}|^3}\,(x_2 - x_1)\,, \qquad \ddot{y}_1 = +G\frac{m_2}{|\vec{r}_{12}|^3}\,(y_2 - y_1)\,, \tag{8.6.11a}$$

[7] o nome vem do clássico em cinco tomos *Mécanique Céleste*, de Pierre-Simon Laplace (1749-1827).

e

$$\ddot{x}_2 = -G\frac{m_1}{|\vec{r}_{12}|^3}(x_2 - x_1)\,, \qquad \ddot{y}_2 = -G\frac{m_1}{|\vec{r}_{12}|^3}(y_2 - y_1) \tag{8.6.11b}$$

onde podemos ainda escrever

$$|\vec{r}_{12}| = |\vec{r}_2 - \vec{r}_2| = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}\,. \tag{8.6.12}$$

Para resolver esse sistema de forma numérica, precisamos montar um sistema de oito EDOs de primeira ordem, acrescentando tambem as velocidades v_{1x}, v_{1y}, v_{2x} e v_{2y} às quatro variáveis originais, x_1, y_1, x_2 e y_2. O sistema que precisamos resolver é então

$$\begin{pmatrix} \dot{x}_1 \\ \dot{v}_{1x} \\ \dot{y}_1 \\ \dot{v}_{1y} \\ \dot{x}_2 \\ \dot{v}_{2x} \\ \dot{y}_2 \\ \dot{v}_{2y} \end{pmatrix} = \begin{pmatrix} v_{1x} \\ +G\frac{m_2}{|\vec{r}_{12}|^3}(x_2 - x_1) \\ v_{1y} \\ +G\frac{m_2}{|\vec{r}_{12}|^3}(y_2 - y_1) \\ v_{2x} \\ -G\frac{m_1}{|\vec{r}_{12}|^3}(x_2 - x_1) \\ v_{2y} \\ -G\frac{m_1}{|\vec{r}_{12}|^3}(y_2 - y_1) \end{pmatrix} \tag{8.6.13}$$

Você pode testar a solução para diferentes massas m_1 e m_2 e algumas condições iniciais. Sugiro trabalhar em *unidades astronômicas* do sistema solar: tome a massa da Terra ($M_\oplus = 5.972 \times 10^{24}$ kg) como unidade fundamental de massa (ou seja, nessas unidades, a massa da Terra é igual a 1), o ano terrestre (365.2425 dias) como unidade de tempo e a distância média Sol-Terra ($d_\oplus = 149.6$ milhões de quilômetros) como unidade de comprimento. Nesse contexto, a constante G é dada por

$$G = 1.19 \times 10^{-4}\, d_\oplus^3 M_\oplus^{-1} \text{ano}^{-2} \tag{8.6.14}$$

Como ilustração, a solução do problema para $m_1 = 6$, $m_2 = 1$ e com as condições iniciais

$$\vec{r}_1(0) = (0,0)\,, \vec{v}_1(0) = (0,0) \tag{8.6.15a}$$
$$\vec{r}_2(0) = (1,0)\,, \vec{v}_1(0) = (0,0.015) \tag{8.6.15b}$$

está representado na Figura 8.10. Repare que nesse gráfico as posições dos dois corpos estão referenciadas ao centro de massa: nesse sistema de coordenadas, a posição $\vec{R}_i$ do corpo i ($i = 1, 2$) é dada por

$$\vec{R}_i = \vec{r}_i - \vec{r}_{cm} \tag{8.6.16}$$

onde $\vec{r}_{cm}$ é a posição do centro de massa:

$$\vec{r}_{cm} = \frac{m_1\vec{r}_1 + m_2\vec{r}_2}{m_1 + m_2}\,. \tag{8.6.17}$$

Da Figura 8.10, vemos que os dois corpos executam um movimento elíptico ao redor do centro de massa, que ocupa um dos focos das elipses, como esperado pelo que se sabe da solução algébrica das equações do movimento. Além disso, o centro de massa se desloca em movimento retilíneo uniforme (não mostrado na Figura 8.10).

É interessante e instrutivo agora observar o que acontece com a energia total do sistema à medida em que o tempo flui. Como sabemos, o sistema de dois corpos gravitacionais é conservativo e, portanto, a energia total é uma constante. Mas podemos fazer a prova calculando a energia a cada etapa da solução numérica. A energia total é a soma das energias cinética e potencial:

$$E = \frac{1}{2}m_1 v_1^2 + \frac{1}{2}m_2 v_2^2 - G\frac{m_1 m_2}{|\vec{r}_{12}|} \tag{8.6.18}$$

O erro relativo (em porcentagem) no problema que estamos acompanhando está mostrado na Figura 8.11. Podemos perceber que a energia não se conserva, mas vai lentamente caindo! Nenhum espanto aqui: o método numérico não tem a obrigação de saber que precisa achar uma maneira de manter a energia constante! No entanto, vemos que o erro, como indica a Figura 8.11, limita-se a um máximo de aproximadamente 0.01% ao final de 800 anos! Isso talvez seja catastrófico se você quer usar o algoritmo para projetar uma missão real ao espaço, mas, para os nossos propósitos aqui, é plenamente aceitável.

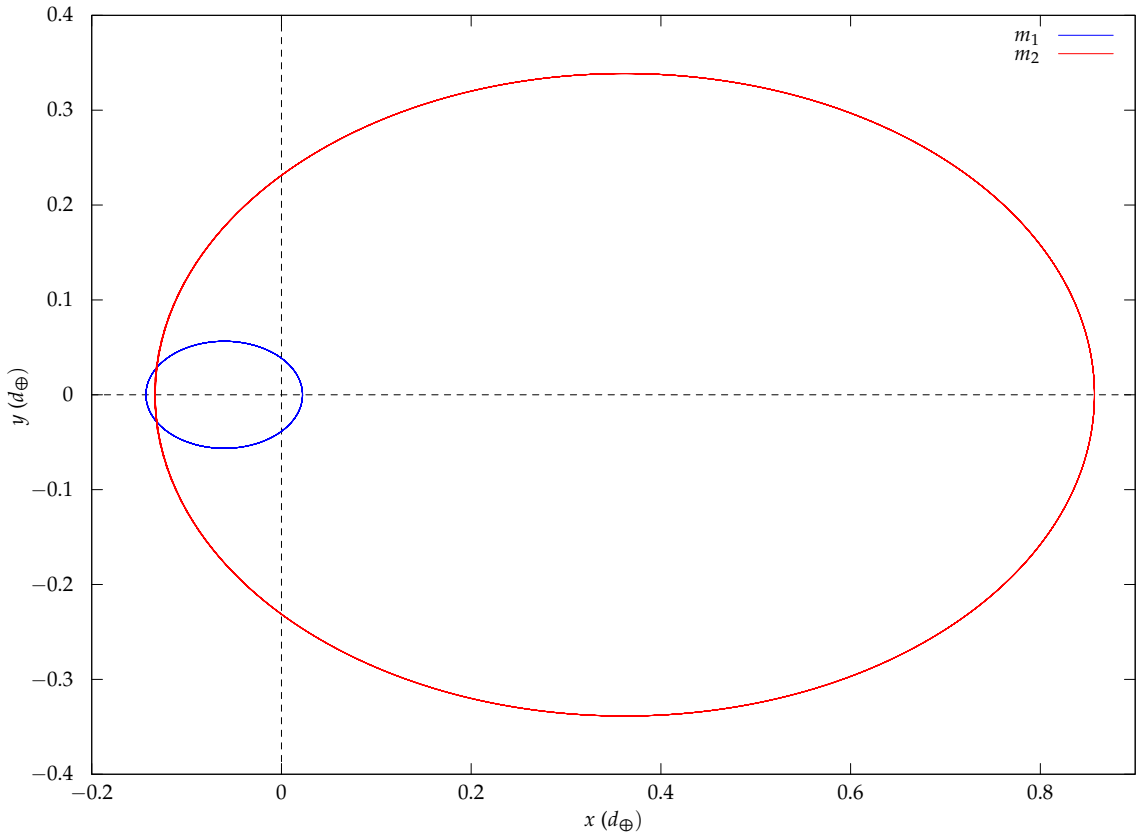


FIGURA 8.10: Movimento de dois corpos de massas $m_1 = 6$ e $m_2 = 1$ (vezes a massa da Terra) sob mútua atração gravitacional, com as condições iniciais dadas pela Eq. (8.6.15). O cruzamento das linhas tracejadas é o centro de massa do sistema, tomado como origem para as posições dos corpos. O problema foi resolvido numericamente usando a função `scipy.integrate.odeint`.

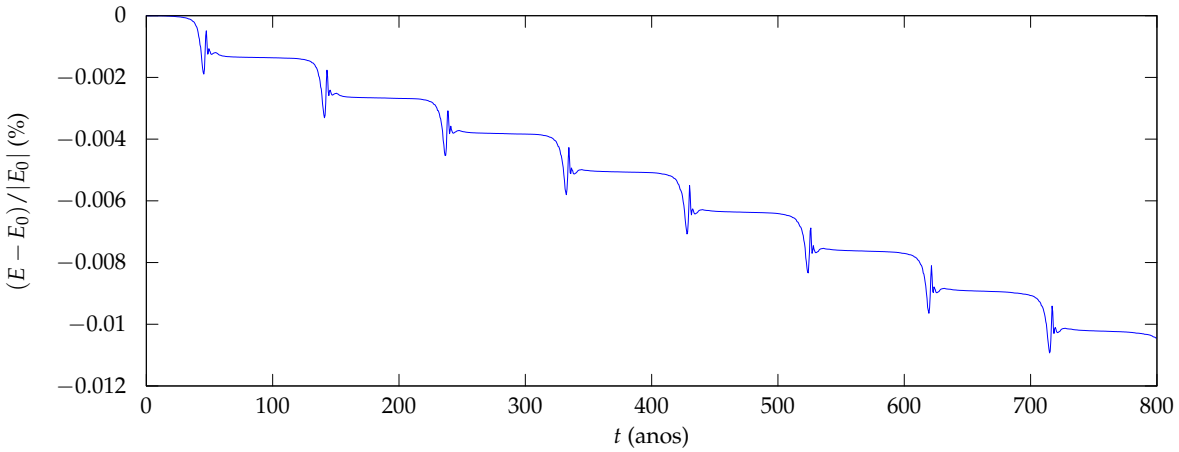


FIGURA 8.11: Erro relativo cometido pela aproximação numérica na conservação da energia no problema de dois corpos gravitacionais.

Além disso, se você acompanhar a animação do problema, encontrada no material de apoio, você vai perceber que as oscilações na energia entre cada patamar, observadas na Figura 8.11, acontecem precisamente no periélio das órbitas, onde a velocidade dos corpos é máxima. Os corpos passam um tempo menor nessas regiões, o que faz com que a distância entre dois instantes consecutivos seja relativamente alto. O algoritmo usado (`scipy.integrate.odeint`) usa passo adaptativo, mas, para aumentar ainda mais a precisão, poderíamos alterar alguns parâmetros de convergência. Deixo isso como exercício aos interessados.

Referências

[1] W. H. Press, S. A. Teukolsky, W. T. Vetterling e B. P. Flannery. *Numerical Recipes 3rd Edition: The Art of Scientific Computing*. New York, NY, USA: Cambridge University Press, 2007.

[2] E. N. Lorenz. "Deterministic Nonperiodic Flow". *Journal of Atmospheric Sciences* 20 (1963), 130–141.

[3] E. N. Lorenz. "The Predictability of Hydrodynamic Flow". *Transactions of the New York Academy of Sciences* 25 (1963), 409–432.

Exercícios

8. 1 — Um modelo para o número de não-conformistas na sociedade é o seguinte:[8] Suponha que uma sociedade tem uma população de $x(t)$ indivíduos no instante t (am anos), e que todo não-conformista que acasala com outro não-conformista tem descendentes também não-conformistas, ao passo que acasalamentos entre conformistas geram uma proporção r de descendentes não-conformistas. Se as taxas de natalidade e mortalidade de todos os indivíduos são, respectivamente, as constantes n e m, e se conformistas e não-conformistas se acasalam aleatoriamente, o problema pode ser espresso pelo seguinte sistema de EDOs:

$$\begin{cases} \dot{x}(t) = (n-m)x(t) \\ \dot{y}(t) = (n-m)y(t) + rn\,[x(t) - y(t)] \end{cases}$$

onde $y(t)$ indica o número de não-conformistas no instante t.

(a) Se $p(t) = y(t)/x(t)$ é a proporção de indivíduos não-conformistas na sociedade no instante t, mostre que podemos combinar as duas equações acima e escrever $\dot{p}(t) = rn\,[1 - p(t)]$.

(b) Encontre $p(t)$ de forma exata em função de n, m, r e da condição inicial $p(0) = p_0$.

(c) Para $p(0) = 0.01$, $n = 0.02$, $m = 0.015$ e $r = 0.1$, encontre $p(t)$ numericamente para $0 \leqslant t \leqslant 50$ anos, comparando com a solução exata.

8. 2 — Imagine uma ilha habitada por coelhoes e raposas. As raposas se alimentam de coelhos e os coelhos comem mato. Considere que existe tanto mato que os coelhos têm abundância de comida. Quando os coelhos proliferam, as raposas florescem e sua população cresce. Mas, quando as raposas se tornam muito numerosas e comem muitos coelhos, elas entram num período de fome e sua população começa a declinar. Como as raposas diminuem, a população de coelhos começa a crescer mais uma vez, propiciando um novo crescimento das raposas. À medida em que o tempo passa, o processo de aumento e diminuição das populações de raposas e coelhos se repete. Problemas desse tipo são estudados por matemáticos, físicos e biólogos e é bastante interessante ver como as conclusões dos dois primeiros, baseadas em modelos matemáticos, confirmam as observações dos últimos. Para descrever esse problema matematicamente, vamos chamar a população de coelhos no instante t de $x(t)$ e a de raposas de $y(t)$. O matemático italiano Vito Volterra (1860-1940) propôs o seguinte sistema de equações diferenciais para as taxas de crescimento do número de coelhos e raposas em função do tempo:

$$\begin{cases} \dot{x}(t) = ax - bxy \\ \dot{y}(t) = -cy + dxy \end{cases}$$

onde a, b, c e d são constantes positivas. O termo ax nos diz que a taxa de crescimento dos coelhos seria proporcional à população de coelhos no instante t, já que a quantidade de mato é ilimitada, o que resultaria num crescimento exponencial do número de coelhos. De modo parecido, o termo $-cy$ nos diz que a taxa de diminuição do número de raposas seria proporcional ao número de raposas, o que resultaria num decaimento exponencial do número de raposas com o tempo. Mas as duas proporcionalidades são quebradas pelos termos $-bxy$ e dxy, que vêm do número de encontros entre coelhos e raposas, que podemos imaginar

[8] adaptado de problema encontrado em *Looking at History Through Mathematics*, N. Rashevsky, 1968.

proporcional ao tamanho das duas populações, isto é, a x e a y, e que, nesses encontros, um coelho é comido, diminuindo a população de coelhos ($-bxy$) e favorecendo o aumento da populaçao de raposas ($+dxy$).

(a) Qual é o *estado estacionário* do sistema, ou seja, quais são os valores de x e y para os quais $\dot{x} = \dot{y} = 0$? Veja que, nessa situação, as populações não se alteram ao longo do tempo.

(b) Faça o gráfico de $y(t)$ em função de $x(t)$ para o caso em que $a = b = c = d = 1$ e na condição inicial $x(0) = 3$ e $y(0) = 1$. Indique o ponto do estado estacionário no mesmo gráfico. É possível atingi-lo com essa condição inicial?

8. 3 — Uma objeto de massa $m = 300\,\mathrm{g}$ é preso a uma mola de constante $k = 25\,\mathrm{N/m}$ e sofre uma resistência viscosa proporcional à velocidade com um coeficiente $\epsilon = 0.6\,\mathrm{N\,s/m}$. Além disso, é impulsionado por uma força externa dada (em N) por $F_e(t) = 50\cos(0.5t)$. As leis de Newton nos dizem que o movimento é descrito pela equação diferencial

$$m\ddot{x}(t) + \epsilon\dot{x}(t) + kx(t) = F_e(t)$$

Se $v(t)$, $a(t)$ e $E(t)$ são, respectivamente, a velocidade, a aceleração e a energia mecânica do objeto no instante t, faça os seguintes gráficos: (a) $x(t)$ vs. t; (b) $v(t)$ vs. t; (c) $a(t)$ vs. t; (d) $E(t)$ vs. t; (e) $v(t)$ vs. $x(t)$ para $0 \leqslant t \leqslant 20\,\mathrm{s}$ e a condição inicial $x(0) = 40\,\mathrm{cm}$, $\dot{x}(0) = 0$.

8. 4 — Nos anos de 1920, Balthasar van der Pol realizou experimentos com novos circuitos elétricos incluindo triodos (tubos de vácuo). Um desses circuitos, conhecido por *oscilador de van der Pol*, é descrito pela equação diferencial

$$\ddot{x} = -x - \varepsilon(x^2 - 1)\dot{x}\,.$$

É interessante que essa equação aparece frequentemente, podendo ser encontrada, por exemplo, na teoria de lasers. Se $\varepsilon = 0$, a equação é simplesmente o oscilador harmônico simples. Mas, quando $\varepsilon \neq 0$, o oscilador de van der Pol se afasta da harmonicidade de forma não-linear. Considerando $\varepsilon = 1$ (que é bem diferente de zero!) e a condição inicial $x(0) = 0.5$, $\dot{x}(0) = 0$, faça um gráfico de (a) $x(t)$ vs. t; (b) $\dot{x}(t)$ vs. t; (c) $\dot{x}(t)$ vs. $x(t)$ para $0 \leqslant t \leqslant 8\pi$.

8. 5 — A termoelasticidade é um fenômeno físico de acoplamento entre propriedades térmicas (capacidade térmica, condutividade térmica, etc.) e elásticas (densidade, constantes elásticas, etc.) de uma material [9]. É o fonômeno responsável, por exemplo, pela expansão térmica. Existem materias termoelásticos bastante complicados! De forma muito simplificada (muito mesmo!), o sistema de equações que descrevem tais materiais é da seguinte forma:

$$\begin{aligned} \rho\ddot{v} &= \mu v - \beta(\dot{\theta} + \tau\ddot{\theta}) \\ c\ddot{\theta} + c\tau\dddot{\theta} &= -\beta\dot{v} + k\dot{\theta} + k^*\theta \end{aligned}$$

onde $\theta(t)$ e $v(t)$ são a temperatura e a velocidade de deformação em função do tempo t, respectivamente, e os demais parâmetros são propriedades físicas como densidade, condutividade térmica, etc. Faça um gráfico de $\theta(t)$ *vs.* t para $0 \leqslant t \leqslant 5$ usando o seguinte conjunto de constantes (em unidades arbitrárias):

$$\rho = 1, \quad \mu = 1, \quad \beta = 1, \quad \tau = 1, \quad c = 1, \quad k = 1, \quad k^* = 1$$

com a condição inicial

$$\theta(0) = 1, \quad \dot{\theta}(0) = 0, \quad \ddot{\theta}(0) = 0, \quad v(0) = 0, \quad \dot{v}(0) = 0\,.$$

[9] Quintanilla, R., "Moore–Gibson–Thompson Thermoelasticity." *Mathematics and Mechanics of Solids* **24** (2019), 4020–31 (doi: 10.1177/1081286519862007).

Parte III

Métodos determinísticos

Capítulo 9

Transformadas de Fourier

> As integrais que obtivemos não são apenas expressões gerais que satisfazem a equação diferencial, elas representam da maneira mais distinta o efeito natural do fenômeno...
>
> Jean-Baptiste-Joseph Fourier

9.1 Revisão: séries de Fourier

Jean-Baptiste Joseph Fourier (1768-1830), físico e matemático francês e ardente adepto da Revolução Francesa, foi o primeiro a perceber contundentemente a utilidade das funções trigonométricas para a solução de problemas envolvendo equações diferenciais, como o fenômeno de condução de calor em meios materiais. Seu *Théorie analytique de la chaleur* (Teoria analítica do calor), de 1822, é um dos grandes clássicos da ciência. O ensaio original, no entanto, data de 1812, quando ganhou o prêmio da Academia Francesa de Ciências.

Resumindo bastante a história:[1] Fourier percebeu que qualquer função periódica pode ser descrita como uma soma de senos e cossenos. Por exemplo, se a função $f(t)$ é periódica com período 2π, tal que $f(t+2\pi) = f(t)$, Fourier propôs a seguinte expansão:

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos nt + \sum_{n=1}^{\infty} b_n \operatorname{sen} nt \tag{9.1.1}$$

em que a_0, a_n e b_n são constantes a serem determinadas usando algumas propriedades das funções trigonométricas, que veremos daqui a pouco. À época, Fourier não se preocupou em demonstrar a validade, nem em garantir a convergência da Eq. (9.1.1) para uma função $f(t)$ qualquer — e foi por isso criticado por uma junta de peso da Academia Francesa de Ciências, composta por Legendre, Lagrange e Laplace. Posteriormente, essa estrada foi pavimentada por outros matemáticos, como Dirichlet e Riemann. Essa área de estudos recebeu o nome de *Análise Harmônica*.

Para determinar os coeficientes a_0, a_n e b_n da Eq. (9.1.1), fazemos uso das chamadas *relações de ortogonalidade* das funções trigonométricas, aliadas ao fato de que a integral das funções seno e cosseno se anulam se o intervalo de integração é um múltiplo inteiro do período dessas funções. Usando as relações de prostaferese encontradas no Apêndice B, você pode verificar que

$$\int_{-\pi}^{\pi} \operatorname{sen} mt \, \operatorname{sen} nt \, dt = \pi \delta_{mn} \tag{9.1.2}$$

$$\int_{-\pi}^{\pi} \cos mt \, \cos nt \, dt = \pi \delta_{mn} \tag{9.1.3}$$

$$\int_{-\pi}^{\pi} \operatorname{sen} mt \, \cos nt \, dt = 0 \tag{9.1.4}$$

para qualquer n e m inteiros, sendo δ_{mn} a função delta de Kronecker, que já encontramos antes na página 32. Apesar de termos utilizado $[-\pi, +\pi]$ como intervalo de integração nas equações acima, qualquer intervalo

[1] Alguns outros detalhes sobre a biografia de Fourier, assim como um desenvolvimento bastante didático sobre o assunto do presente capítulo, são encontrados em DeVries e Hasbun [1], que nos serviu de inspiração em vários aspectos.

de comprimento igual ao período 2π (por exemplo, de 0 a 2π, ou de $-\pi/3$ a $5\pi/3$) daria os mesmos resultados

Com as relações de ortogonalidade, o "truque" de Fourier para achar os coeficientes a_n, por exemplo, é multiplicar ambos os lados da Eq. (9.1.1) por $\cos nt$ (para $n \neq 0$) e integrar num período de $f(t)$, por exemplo, de $-\pi$ a $+\pi$ (antes disso, troque n por m na Eq. (9.1.1):

$$\int_{-\pi}^{\pi} f(t)\cos nt\,dt = \frac{a_0}{2}\int_{-\pi}^{\pi}\cos nt\,dt + \sum_{m=1}^{\infty} a_m \int_{-\pi}^{\pi}\cos mt\cos nt\,dt + \sum_{m=1}^{\infty} b_m \int_{-\pi}^{\pi}\operatorname{sen} mt\cos nt\,dt \tag{9.1.5}$$

As relações de ortogonalidade eliminam quase todas as integrais e a Eq. (9.1.5) se reduz para:

$$\int_{-\pi}^{\pi} f(t)\cos nt\,dt = a_n\pi \tag{9.1.6}$$

que fornece então os coeficientes a_n:

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\cos nt\,dt \tag{9.1.7}$$

Para achar a_0, basta integrar a Eq. (9.1.1):

$$\int_{-\pi}^{\pi} f(t)\,dt = \frac{a_0}{2}\int_{-\pi}^{\pi} dt \tag{9.1.8}$$

que fornece

$$a_0 = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\,dt \tag{9.1.9}$$

Veja que na Eq. (9.1.1) aparece o termo $a_0/2$, e não simplesmente a_0, apenas para deixar as Eqs. (9.1.7) e (9.1.9) parecidas entre si. Por fim, multiplicando a Eq. (9.1.1) por $\operatorname{sen} nt$ e integrando, descobrimos os coeficientes b_n:

$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\operatorname{sen} nt\,dt \tag{9.1.10}$$

Vejamos um exemplo: considere a seguinte função $f(t)$, chamada às vezes de onda quadrada:

$$f(t) = \begin{cases} -1 & ,-\pi \leqslant x \leqslant 0 \\ +1 & ,0 \leqslant x \leqslant \pi \end{cases}, \quad f(t+2\pi) = f(t) \tag{9.1.11}$$

cujo gráfico está esquematizado na Figura 9.1. Veja que $f(t)$ tem descontinuidades em $t = \pm n\pi$ (com n inteiro). Como os coeficientes de Fourier são obtidos por meio de integração, e um número finito de descontinuidades não afeta o valor da integral, podemos continuar usando as relações vistas acima para o cálculo de a_0, a_n e b_n. Usando então as Eqs. (9.1.9), (9.1.7) e (9.1.10), descobrimos que $a_0 = a_n = 0$ e

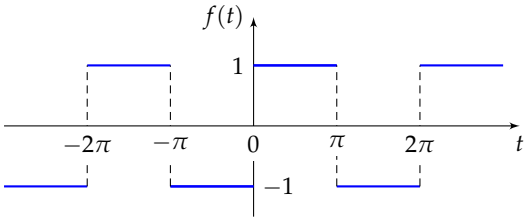


FIGURA 9.1: Função $f(t)$ na forma de uma onda quadrada de período 2π.

$$b_n = \frac{2}{n\pi}(1-\cos n\pi) = \begin{cases} 0 & ,n \text{ par} \\ \frac{4}{n\pi} & ,n \text{ ímpar} \end{cases} \tag{9.1.12}$$

De forma que podemos escrever $f(t)$ como a seguinte série trigonométrica:

$$f(t) = \frac{4}{\pi}\sum_{n=1}^{\infty}\frac{\operatorname{sen} nt}{n} \qquad (n \text{ ímpar}) \tag{9.1.13}$$

O gráfico das somas parciais da Série de Fourier para a função $f(t)$ dada na Eq. (9.1.11) é mostrado na Figura 9.2. A soma parcial $S_n(t)$ corresponde a somar apenas os n primeiros termos da soma da Eq. (9.1.13).

Veja que, nos pontos de descontinuidade da função, por exemplo, em $t = 0$, a Série de Fourier converge para valor médio dos valores da função de cada lado da descontinuidade. Essa é uma propriedade geral da Série de Fourier. Uma outra propriedade importante das séries de Fourier é a seguinte: se a função $f(t)$ é ímpar, ou seja, se $f(-t) = -f(t)$, sempre teremos $a_n = 0$ e a série terá apenas os termos em $\operatorname{sen} nt$. Caso, por outro lado, a função seja par, com $f(-t) = f(t)$, teremos sempre $b_n = 0$ e a Série de Fourier terá apenas os termos em $\cos nt$.

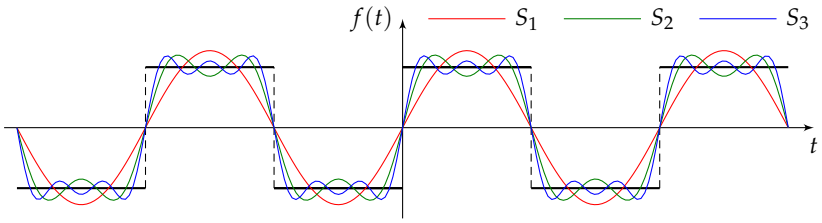


FIGURA 9.2: Somas parciais dos primeiros termos da Série de Fourier para a função onda quadrada.

9.2 Representação complexa

A Eq. (9.1.1) pode ser colocada num formato mais compacto e, para muitas aplicações, também mais interessante, transformando os senos e cossenos em exponenciais complexas usando as relações de Euler do Apêndice B. Vejamos o que acontece: em primeiro lugar, façamos a substituições na Eq. (9.1.1):

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \frac{e^{int} + e^{-int}}{2} + \sum_{n=1}^{\infty} b_n \frac{e^{int} - e^{-int}}{2i} \tag{9.2.1}$$

Agora, vamos juntar tudo numa única soma e agrupar os termos com mesmo expoente:

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \left[\frac{a_n - ib_n}{2} e^{int} + \frac{a_n + ib_n}{2} e^{-int} \right] \tag{9.2.2}$$

e limpar um pouco a notação, redefinindo as constantes:

$$f(t) = A_0 + \sum_{n=1}^{\infty} \left(A_n e^{int} + B_n e^{-int} \right) \tag{9.2.3}$$

Vamos novamente separar a soma em duas parciais, fazendo uma pequena alteração de n para $-n$ no índice da segunda soma:

$$f(t) = A_0 + \sum_{n=1}^{\infty} A_n e^{int} + \sum_{n=-\infty}^{-1} B_{-n} e^{int} \tag{9.2.4}$$

E para que fazer uma barbaridade dessas? A ideia é agora juntar novamente as somas usando um único índice:

$$f(t) = \frac{1}{\sqrt{2\pi}} \sum_{n=-\infty}^{\infty} C_n e^{int} \tag{9.2.5}$$

onde juntamos todos as constantes A_n e B_{-n} numa mesma notação $C_n/\sqrt{2\pi}$. O fator $\sqrt{2\pi}$ é apenas uma escolha para deixar as equações mais bonitas, não se preocupe com ele. Alguns autores (Boas [2], por exemplo) adotam uma notação diferente, tenha cuidado! Repare que juntamos também a constante A_0 (que rebatizamos de $C_0/\sqrt{2\pi}$) à soma na Eq. (9.2.5), que é a expressão procurada: a Série de Fourier em notação complexa.

A aplicação do truque de Fourier fica ainda mais simples em notação complexa: basta multiplicar ambos os lados da Eq. (9.2.5) pelo *conjugado complexo* do termo e^{int}, que é e^{-int}, e integrar num período (mas antes, troque n por m na Eq. 9.2.5):

$$\int_{\pi}^{\pi} f(t) e^{-int}\, dt = \frac{1}{\sqrt{2\pi}} \sum_{m=-\infty}^{\infty} C_m \int_{\pi}^{\pi} e^{imt} e^{-int}\, dt \tag{9.2.6}$$

Vamos resolver a integral do lado direito da equação acima:

$$\int_{-\pi}^{\pi} e^{imt} e^{-int}\, dt = \left. \frac{e^{i(m-n)t}}{i(m-n)} \right|_{-\pi}^{\pi} = \frac{e^{i(m-n)\pi} - e^{-i(m-n)\pi}}{i(m-n)} = \frac{2i \operatorname{sen}(m-n)\pi}{i(m-n)} = 2\pi \frac{\operatorname{sen}(m-n)\pi}{(m-n)\pi} \tag{9.2.7}$$

Mas veja que, como m, n e, consequentemente, $m - n$ são números inteiros, segue que $\operatorname{sen}(m-n)\pi = 0$. Se $m \neq n$, a integral, portanto, se anula. Mas, no caso em que $n = m$, é preciso ter cuidado, pois o denominador também se anula. Mas sabemos que a razão $\operatorname{sen} x/x$ tende a 1 para $x \to 0$. Juntando então os dois casos,

concluímos que

$$\int_{-\pi}^{\pi} e^{imt} e^{-int}\, dt = 2\pi\delta_{mn} \tag{9.2.8}$$

Voltando finalmente à Eq. (9.2.6):

$$\int_{-\pi}^{\pi} f(t) e^{-int}\, dt = \frac{1}{\sqrt{2\pi}} \sum_{m=-\infty}^{\infty} C_m 2\pi\delta_{mn} = C_n\sqrt{2\pi} \tag{9.2.9}$$

e assim conseguimos calcular as constantes C_n:

$$C_n = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} f(t) e^{-int}\, dt \tag{9.2.10}$$

Vamos aplicar essa relação a um exemplo: vamos expandir a função $f(t) = t$, que pode ser chamada de onda triangular, usando a Série de Fourier em notação complexa. Calculando então os coeficientes C_n da expansão usando a Eq. (9.2.10):

$$C_n = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} te^{-int}\, dt \tag{9.2.11}$$

Essa integral pode ser resolvida por partes:

$$\begin{aligned}\int_{-\pi}^{\pi} te^{-int}\, dt &= \left[e^{-int}\left(i\frac{t}{n} + \frac{1}{n^2}\right)\right]\Bigg|_{-\pi}^{\pi} = e^{-in\pi}\left(i\frac{\pi}{n} + \frac{1}{n^2}\right) - e^{in\pi}\left(-i\frac{\pi}{n} + \frac{1}{n^2}\right) \\ &= i\frac{\pi}{n}\left(e^{i\pi n} + e^{-i\pi n}\right) - \frac{1}{n^2}\left(e^{i\pi n} - e^{-i\pi n}\right) = i\frac{\pi}{n} 2\cos n\pi - \frac{1}{n^2} 2i \operatorname{sen} n\pi = i\frac{2\pi}{n}(-1)^n\end{aligned}$$

na qual o último passo vem do fato de que n é inteiro. Com isso, as constantes C_n são dadas por

$$C_n = i\frac{\sqrt{2\pi}}{n}(-1)^n \tag{9.2.12}$$

e a expansão em Série de Fourier, em notação complexa, será dada por

$$f(t) = \frac{1}{\sqrt{2\pi}} \sum_{n=-\infty}^{\infty} i\frac{\sqrt{2\pi}}{n}(-1)^n e^{int} = i \sum_{n=-\infty}^{\infty} \frac{(-1)^n}{n} e^{int} \tag{9.2.13}$$

Para ser coerente, precisaríamos calcular separadamente o caso em que $n = 0$, calculando a integral novamente. Se você fizer isso, verá que $C_0 = 0$. Além disso, para voltar à notação real, podemos usar a relação de Euler na última equação:

$$f(t) = i \sum_{n=-\infty}^{\infty} \frac{(-1)^n}{n}(\cos nt + i \operatorname{sen} nt) \tag{9.2.14}$$

Mas veja que $\cos nt/n$ troca de sinal quando trocamos n por $-n$. Na mesma situação, $\operatorname{sen} nt/n$ mantém o sinal. Por esse motivo, os termos em cosseno se cancelam dois a dois, ao passo que os termos em seno se somam, o que resulta em

$$f(t) = 2i \sum_{n=1}^{\infty} \frac{(-1)^n}{n} i \operatorname{sen} nt = -2 \sum_{n=1}^{\infty} \frac{(-1)^n}{n} \operatorname{sen} nt \tag{9.2.15}$$

ou ainda

$$f(t) = 2\left(\operatorname{sen} t - \frac{\operatorname{sen} 2t}{2} + \frac{\operatorname{sen} 3t}{3} - \frac{\operatorname{sen} 4t}{4} + \ldots\right) \tag{9.2.16}$$

O resultado é coerente com a Série de Fourier real, já que a função $f(t) = t$ é ímpar e, portanto, devem existir apenas os termos em seno.

As primeiras somas parciais estão mostradas na Figura 9.3. Novamente, a expansão converge para o valor médio da função em cada um dos lados das descontinuidades. Você pode perceber que a convergência é bem lenta: a soma dos três primeiros termos descrevem bastante grosseiramente a função.

Talvez você esteja se perguntando há algum tempo: valeu a pena ter usado a Série de Fourier em notação complexa? A álgebra ficou até mais complicada... Teria sido mais fácil calcular apenas os coeficientes b_n da série real, pois sabemos que todos os a_n se anulariam de qualquer jeito, já que a função é impar. A resposta

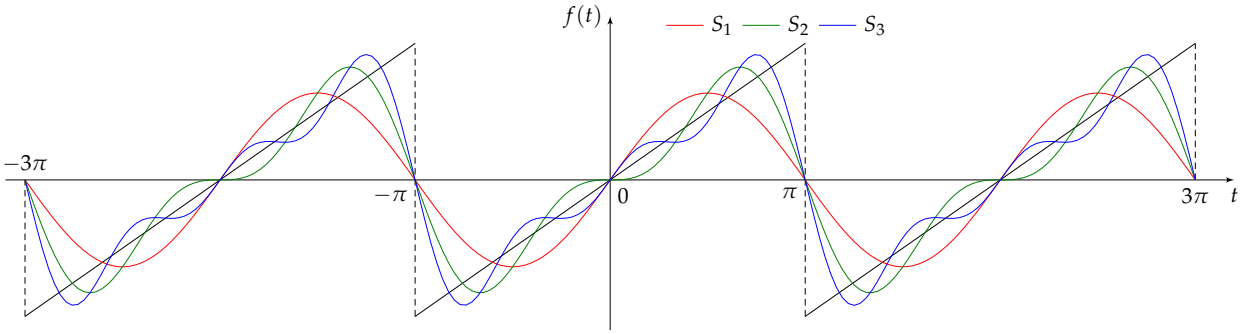


FIGURA 9.3: Três primeiras somas parciais da expansão em Série de Fourier da função onda triangular.

é: realmente, teria sido mais fácil calcular tudo do jeito tradicional, sem usar números complexos. Mas não é essa a utilidade da notação complexa. Ela serve, na verdade, como uma notação mais simplificada e poderosa, que nos permitirá fazer as generalizações que veremos a seguir para chegar ao nosso objetivo final: a Transformada de Fourier.

9.3 Mudando o período

Até aqui, vimos apenas situações em que a função $f(t)$ tem um período 2π. Naturalmente, não vamos escapar agora da tarefa de generalizar para o caso em que $f(t)$ tem um período qualquer. Para facilitar a vida, digamos que esse período seja igual a $2T$, em que T é um escalar qualquer. Isso significa que a função $f(t)$ é tal que

$$f(t+2T) = f(t)\,. \tag{9.3.1}$$

Para facilitar ainda mais, vamos considerar um domínio simétrico em relação a $t = 0$, ou seja, na região $-T \leqslant t \leqslant T$. Fora dessa região, a função se repete periodicamente, como nos dois exemplos já encontrados até aqui, podendo ter alguns pontos de descontinuidade. E, por fim, vamos nos limitar à Série de Fourier em notação complexa.

Será que precisamos deduzir todas as fórmulas desde o início? Não! Se você pensar um pouco, verá que basta alterar o fator dentro dos senos e cossenos (ou nos expoentes das exponenciais complexas) de forma a que tenham o mesmo período que $f(t)$. Veja que isso nos leva a reescrever a Eq. (9.2.5) de uma maneira um pouco diferente:

$$f(t) = \frac{1}{\sqrt{2\pi}} \sum_{n=-\infty}^{\infty} C_n e^{i\pi nt/T} \tag{9.3.2}$$

na qual as exponenciais complexas tem períodos $2T$. As constantes C_n são calculadas pelo truque de Fourier e resultam, em analogia à Eq. (9.2.10), na seguinte equação:

$$C_n = \frac{\pi}{T}\frac{1}{\sqrt{2\pi}} \int_{-T}^{T} f(t) e^{-i\pi nt/T} dt \tag{9.3.3}$$

Cuidado! Alguns autores adotam uma convenção diferente para as constantes pré soma/integral. Não há consenso sobre o assunto — até mesmo por que não faz muita diferença, desde que você seja coerente: adote uma convenção e siga-a até o fim!

Resumindo: podemos adaptar os dois exemplos já vistos, em que o período era igual a 2π, para qualquer outro período desejado, simplesmente adaptando os termos já calculados. Você está convidado a ver o que acontece com as ondas quadradas e triangulares que já encontramos, mas adotando um período correspondente a, por exemplo, $T = 1$. Não precisa calcular todas as integrais novamente!

9.4 A Transformada de Fourier

Como vimos, a Série de Fourier é uma maneira de expandir uma função $f(t)$ dentro de um intervalo. Fora desse intervalo, o que fazemos, na verdade, é repetir periodicamente $f(t)$, de forma que o período da nova função a ser expandida equivale ao intervalo que consideramos inicialmente. Por exemplo, ao expandir em Série de Fourier a função $f(t) = t$, como fizemos na seção 9.2, só consideramos o intervalo $-\pi \leqslant t \leqslant \pi$. Fora desse intervalo, a função não era exatamente a mesma $f(t) = t$, mas cópias periódicas, resultando na chamada onda triangular que vemos na Figura 9.3. Já na seção 9.3, vimos uma maneira de alterar o período — e com isso também a *frequência* — da onda. Assim, podemos expandir (ou contrair) o período da função

considerada o quanto quisermos. Isso sem dúvida é bastante útil, mas não é o fim da história, pois a função $f(t)$ continua periódica. E se quisermos expandir uma função realmente não periódica? Não precisamos ir muito longe: $f(t) = t$ é um exemplo dos mais simples!

A maneira de expandir funções não-periódicas é usar um truque: é só considerar que a sua função não-periódica tem um período infinito. Isso significa que, nas Eqs. (9.3.2) e (9.3.3), temos que fazer $T \to \infty$. Então mãos à obra. Vamos definir inicialmente a frequência angular (ou simplesmente frequência) ω como

$$\omega = \frac{\pi n}{T} \tag{9.4.1}$$

Como n é um número inteiro, a diferença entre dois valores consecutivos de ω é o que chamaremos de $\Delta\omega$, dado por

$$\Delta\omega = \frac{\pi(n+1)}{T} - \frac{\pi n}{T} = \frac{\pi}{T} \tag{9.4.2}$$

Repare que, no limite $T \to \infty$, acontece que $\Delta\omega \to 0$.

Agora, faça as substituições na Eq. (9.3.3) para chegar a

$$C_n = \frac{\Delta\omega}{\sqrt{2\pi}} \int_{-T}^{T} f(t) e^{-i\omega t} dt \tag{9.4.3}$$

Se definirmos agora a seguinte função auxiliar:

$$F_T(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-T}^{T} f(t) e^{-i\omega t} dt \tag{9.4.4}$$

podemos reecrever C_n como

$$C_n = F_T(\omega)\Delta\omega \tag{9.4.5}$$

Muito bem. Veja que, no limite $T \to \infty$, a Eq. (9.4.4) se torna a função $F(\omega)$ dada por

$$\boxed{F(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt} \tag{9.4.6}$$

Essa é a famosa *Transformada de Fourier* da função $f(t)$. Repare que ela é uma função da frequência ω, e não mais da variável t. É comum também encontrar a Transformada de Fourier representada por $\mathcal{F}[f(t)]$:

$$\mathcal{F}[f(t)] = F(\omega) \tag{9.4.7}$$

Uma observação: se t é o tempo, ω é realmente uma frequência (mais especificamente, frequência angular), com unidades coerentes. Nesse caso, dizemos que $f(t)$ é uma função no domínio temporal (*time domain*), enquanto $F(\omega)$ vive no domínio das frequências (*frequency domain*). Mas, nada impede de identificar t com alguma outra variável, como a posição. Nesse caso em particular, ω terá unidades de comprimento de onda recíproco (ou número de onda), sendo alguma forma de medida do *espectro* da função em relação à variável original. Em cristalografia, por exemplo, os padrões de difração são uma medida da intensidade do espectro de espalhamento dos raios X pela nuvem eletrônica dos planos atômicos do material. A variável ω, naquele contexto, que geralmente é substituída por k, define o chamado *espaço recíproco* da rede cristalina, que, por sua vez, mora no espaço direto, ou real.

Mas basta de divagações. Faça agora o seguinte: substitua a Eq. (9.4.5) na Eq. (9.3.2). O resultado será

$$f(t) = \frac{1}{\sqrt{2\pi}} \sum_{n=-\infty}^{\infty} F_T(\omega) e^{i\omega t} \Delta\omega \tag{9.4.8}$$

O que acontece no limite $T \to \infty$, ou seja, se a função é possivelmente não-periódica? Vejamos: já havíamos dito que $\Delta\omega \to 0$ e que $F_T(\omega) \to F(\omega)$. Portanto, a soma da Eq. (9.4.8) acaba virando uma integral:

$$\boxed{f(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F(\omega) e^{i\omega t} d\omega} \tag{9.4.9}$$

A Eq. (9.4.9) é chamada de *Transformada Inversa de Fourier*. É comum também indicar a Transformada Inversa por $\mathcal{F}^{-1}[F(\omega)]$:

$$\mathcal{F}^{-1}[F(\omega)] = f(t) \tag{9.4.10}$$

Repare que, enquanto lidávamos com funções periódicas, os coeficientes C_n eram nosso objetivo. Precisávamos descobrir o valor desses coeficientes para todos os (infinitos) valores de n. Como a Série de Fourier depende dessas constantes para descrever a função $f(t)$, podemos dizer que os C_n são uma representação da função original. Agora, para funções não-periódicas, o objetivo é encontrar a Transformada de Fourier usando a Eq. (9.4.9). Não queremos mais uma infinidade de valores C_n, mas uma única função $F(\omega)$ que nos conta a contribuição de cada frequência ω no intervalo $[\omega, \omega + d\omega]$ à expansão de $f(t)$. Nesse sentido, $F(\omega)$ é uma *distribuição contínua*, ao passo que os C_n eram, de certo modo, uma *distribuição discreta* de frequências. Assim como no caso dos C_n, a Transformada de Fourier é uma representação, no espaço das frequências, da função original no espaço temporal.

Vejamos um exemplo de aplicação da Transformada de Fourier num caso simples. Casos mais complicados fogem do escopo destas notas de aula, ainda mais porque você provavelmente os encontrou (ou encontrará) na disciplina de Física Matemática, ou alguma outra disciplina correlata. O exemplo que veremos aqui é o de um pulso retangular, definido como

$$f(t) = \begin{cases} 1/a\,, & |t| \leqslant a \\ 0\,, & |t| > a \end{cases} \tag{9.4.11}$$

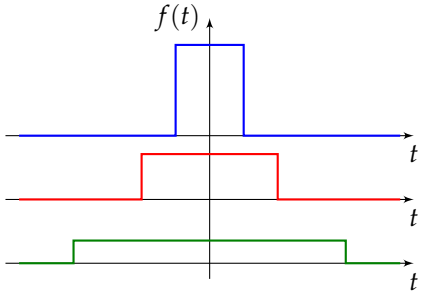


FIGURA 9.4: Pulsos retangulares com diferentes durações e intensidades, mas de mesma área.

onde a é uma constante real positiva, relacionada à duração do pulso, ao passo que $1/a$ é a sua intensidade. Gráficos de $f(t)$, que é evidentemente uma função não-periódica, são mostrados na Figura 9.4, para diferentes valores do parâmetro a. Repare que todos os pulsos têm a mesma área, o que faz com que, se a é grande, o pulso é baixo (baixa intensidade) e largo (longa duração), ao passo que, se a é pequeno, o pulso é quase um pico, de alta intensidade mas curta duração.

Para calcular a Transformada de Fourier da função da Eq. (9.4.11), vamos usar a definição, dada pela Eq. (9.4.6). Repare que, como $f(t)$ é nula fora do intervalo $-a \leqslant t \leqslant a$, o intervalo de integração pode deixar de ser $(-\infty, \infty)$ e passar a $[-a, a]$:

$$F(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-a}^{a} \frac{1}{a} e^{-i\omega t} dt \tag{9.4.12}$$

Qualquer primeiro-anista é capaz de calcular essa integral:

$$F(\omega) = \frac{1}{\sqrt{2\pi}} \frac{e^{i\omega a} - e^{-i\omega a}}{i\omega a} \tag{9.4.13}$$

Você, com o traquejo que já ganhou ao ler o Apêndice B, sabe identificar a relação de exponenciais complexas com funções trigonométricas:

$$F(\omega) = \frac{1}{\sqrt{2\pi}} \frac{2i \operatorname{sen} \omega a}{i\omega a} \quad \Rightarrow \quad F(\omega) = \sqrt{\frac{2}{\pi}} \frac{\operatorname{sen} \omega a}{\omega a} \tag{9.4.14}$$

que é a Transformada de Fourier procurada. Observe que ela é real nesse caso, mas, em geral, é preciso reconhecer que a Transformada de Fourier é uma função complexa. Existem algumas condições para que $F(\omega)$ seja real, ou imaginária, mas não entraremos em tais detalhes no presente texto.

A Eq. (9.4.14) pode ser reescrita como

$$F(\omega) = \sqrt{\frac{2}{\pi}} \operatorname{sinc} \omega a \tag{9.4.15}$$

onde a função $\operatorname{sinc} x$, definida como

$$\operatorname{sinc} x = \frac{\operatorname{sen} x}{x} \tag{9.4.16}$$

é conhecida como *seno cardinal* (apesar de pouca gente usar esse nome) ou simplesmente "sinc". Repare que $\operatorname{sinc} x \to 1$ para $x \to 0$, um dos limites fundamentais do Cálculo.

Gráficos da função $F(\omega)$ dada pela Eq. (9.4.14) são mostrados na Figura 9.5. Podemos usar os gráficos da Figura 9.5 para entender o que acontece quando o pulso é um sinal de áudio e o ouvimos, tentando identificar a frequência, ou, na linguagem da Música, a nota musical correspondente. Quem toca algum instrumento já tentou alguma vez na vida "tirar" uma música de ouvido. Peças musicais mais lentas, com notas de longa duração, são muito mais fáceis de tirar do que um solo irado de guitarra ou um movimento *Allegro Prestissimo* de um concerto para piano de Rachmaninoff. Nosso ouvido (que é um Transformador de Fourier natural) precisa de tempo para identificar a frequência do sinal captado — ou seja, se o pulso é muito breve, há uma

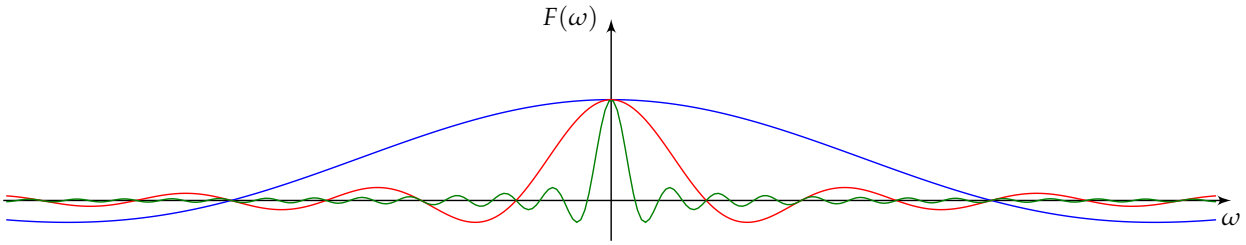


FIGURA 9.5: Transformada de Fourier da função pulso retangular para diferentes durações e intensidades. As cores usadas correspondem aos pulsos mostrados na Figura 9.4.

grande dificuldade em identificar a nota tocada. Isso se reflete na curva azul da Figura 9.5: o pulso, no domínio temporal, equivale a toda uma série de frequências espalhadas pelo espectro da sua Transformada de Fourier. Por outro lado, se o pulso é de longa duração (ainda que de menor intensidade), como no caso da curva verde, o espectro de frequências é mais localizado em $\omega = 0$, com alguns *harmônicos* (picos de menor intensidade) em outras frequências. Na prática, um pulso de áudio tem um caráter ondulatório, com uma frequência diferente de zero, mas a ideia é a mesma: na transformada de Fourier, identificamos mais facilmente a frequência tocada quanto maior a duração do sinal. Identificar as notas pela sua frequência e duração, ao invés de todo o seu espectro ondulatório, é o princípio usado na notação musical tradicional (ver a Figura 9.6).

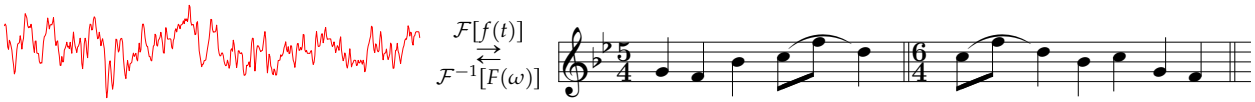


FIGURA 9.6: A notação usada na Música Ocidental é uma codificação via Transformada de Fourier. O instrumentista executa a transformada inversa. À esquerda, uma das notas é mostrada apenas esquematicamente. Ela foi obtida assobiando a primeira nota (um sol) ao microfone e capturando o sinal com a biblioteca `pyaudio`.

Num contexto completamente diferente, o da Mecânica Quântica, as Figuras 9.4 e 9.5 servem ainda para ilustrar outro fenômeno: o Princípio da Incerteza de Heisenberg. Nesse caso, trasformamos uma função no espaço real (das coordenadas espaciais de uma partícula, por exemplo) para o espaço recíproco, ou dos momentos (basicamente, das velocidades). Assim, se conseguimos medir com precisão a posição da partícula (um pulso muito bem localizado em um certo local), não conseguimos medir seu momento com precisão equivalente, e vice-versa.

Apenas mais um exemplo antes de seguir adiante. Vimos que a principal motivação para desenvolver o conceito de Transformada de Fourier foi encontrar uma maneira de determinar o espectro de frequências, $F(\omega)$, de funções não-periódicas. Para isso, usamos o truque de considerar um período infinito para tais funções. Nada nos impede, no entanto, de tentar encontrar a Transformada de Fourier de funções periódicas. Vejamos aonde isso nos leva usando como estudo de caso a função $f(t) = \operatorname{sen} \beta t$, onde β é uma constante que podemos identificar com a frequência (angular) de $f(t)$. A Transformada de Fourier dessa função, de acordo com a Eq. (9.4.6), é dada por

$$F(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \operatorname{sen} \beta t \, e^{-i\omega t} dt \tag{9.4.17}$$

Comece transformando o seno em exponenciais complexas:

$$F(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{e^{i\beta t} - e^{-i\beta t}}{2i} e^{-i\omega t} dt \tag{9.4.18}$$

separando-as então em duas integrais:

$$F(\omega) = \frac{1}{2i\sqrt{2\pi}} \left[\int_{-\infty}^{\infty} e^{-i(\omega-\beta)t} dt - \int_{-\infty}^{\infty} e^{-i(\omega+\beta)t} dt \right] \tag{9.4.19}$$

Se você se lembra das aulas de Física Matemática, talvez se lembre também de uma das representações integrais da *função delta de Dirac*:

$$\delta(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega t} dt \tag{9.4.20}$$

Com isso, a Transformada de Fourier da função seno fica

$$F(\omega) = i\sqrt{\frac{\pi}{2}}\,[\delta(\omega+\beta) - \delta(\omega-\beta)] \tag{9.4.21}$$

O gráfico da Eq. (9.4.21) está mostrado na Figura 9.7. Repare que $F(\omega)$ é complexa nesse caso, e por isso o gráfico, na verdade, mostra $\mathrm{Im}[F(\omega)]$, considerando que os deltas de Dirac são reais. Mas isso não é tão importante quanto o fato de que a Transformada de Fourier de uma função de frequência $\pm\beta$ tem picos exatamente nessa frequência! Porém, não há nada com o que se espantar: Já sabemos que $\operatorname{sen}\beta t$ é precisamente uma combinação linear de duas exponencias complexas, $e^{i\beta t}$ e $e^{-i\beta t}$ — a relação de Euler nos diz isso, e acabamos de usar essa fato na dedução da Eq. (9.4.21). A lição a ser aprendida aqui é: aplicando a Transformada de Fourier a uma função periódica, obteremos um espectro em que a frequência (ou as frequências) de oscilação são predominantes. Quem afina o violão ou a guitarra usando um afinador eletrônico, que tem embutidos circuitos capazes de realizar a Transformada de Fourier de sinais de áudio, faz uso dessa propriedade. Para quem não conhece, o afinador é um aparelho eletrônico capaz de identificar a nota tocada pelo instrumento musical. Ele serve como substituto para um ouvido bem treinado.

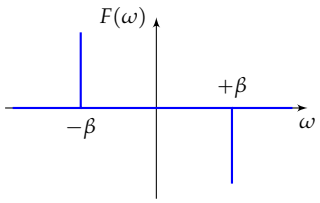


FIGURA 9.7: Transformada de Fourier da função $f(t) = \operatorname{sen}\beta t$. Apenas a parte imaginária está representada.

9.5 Métodos numéricos

A Transformada de Fourier é uma ferramenta poderosa em muitas áreas da Ciência, da Tecnologia e da Arte (Música, como já vimos, mas também no tratamento de imagens, por exemplo). Só conseguiremos aqui ver alguns poucos exemplos da sua vasta gama de aplicações. No entanto, a Transformada de Fourier seria apenas uma curiosidade ou, de modo mais condescendente, teria aplicabilidade limitada, não fosse pela possibilidade de discretizá-la e, especialmente, pela criação de um algoritmo extremamente rápido para calculá-la numericamente, chamado de Transformada Rápida de Fourier. Esse método numérico, geralmente abreviado para FFT (de *Fast Fourier Transform*), não é exagero dizer, revolucionou muitas áreas do conhecimento humano e é um dos responsáveis por você conseguir, por exemplo, assistir a filmes em alta definição no seu celular.

Antes de delinear as ideias por trás da FFT, no entanto, precisamos primeiro entender como calcular numericamente a Transformada de Fourier quando não temos a função $f(t)$ propriamente dita, mas apenas alguns de seus valores — por exemplo, a partir de medidas experimentais de certa grandeza ou quando procedemos a uma discretização dos dados. Nesses casos, precisamos lançar mão de um método que nos fornece, igualmente, a Transformada de Fourier não como uma função $F(\omega)$, mas apenas seus valores em algumas frequências, que chamamos de Transformada Discreta de Fourier.

9.5.1 A Transformada Discreta de Fourier

Vamos recapitular: a Transformada de Fourier, calculada usando a Eq. (9.4.6), repetida aqui para maior conveniência,

$$F(\omega) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} f(t)e^{-i\omega t}dt\,, \tag{9.4.6′}$$

processa um sinal (uma função) $f(t)$ no domínio temporal e obtém sua representação no domínio das frequências, $F(\omega)$. A integral pode ser calculada analiticamente (em alguns casos!), o que fizemos nos dois exemplos vistos anteriormente. Por outro lado, existem diversas técnicas para realizar a integração numérica de funções, coisa que você aprendeu em alguma disciplina de Calculo Numérico: o método dos trapézios, de Simpson, de Romberg, etc. Em todos esses métodos, é preciso discretizar a variável de integração, usando algum passo h (que pode ser fixo ou adaptativo, não importa) e calculando áreas debaixo da curva. Poderíamos usar um método desses para resolver numericamente a Eq. (9.4.6)? Sim, mas existem complicadores. Em primeiro lugar, considere que você precisa começar a discretização em algum valor inicial e parar num valor final. Mas a integral em questão é imprópria: os limites de integração são $-\infty$ e $+\infty$. E, em segundo lugar, você talvez não tenha o sinal $f(t)$ na forma de uma expressão matemática, mas apenas medidas experimentais numa sucessão de valores de t, começando, por exemplo, em $t = 0$. O problema aqui aparece de novo: o que fazer com os valores na região $-\infty < t < 0$, não medidos mas necessários à integração?

Felizmente há maneira de ultrapassar essa cordilheira de dificuldades. A solução para o primeiro problema mencionado no parágrafo anterior (intervalo de integração infinito) é observar o sinal $f(t)$ por um

tempo T longo o suficiente para captar suas principais características, como as principais amplitudes, frequências e períodos embutidos no sinal. Para resolver o segundo problema (o dos valores faltando para integrar em $t < 0$), precisamos recuar alguns passos, voltando ao processo intuitivo que usamos para chegar da Série à Transformada de Fourier. Lembre então que a série de Fourier, em notação complexa, é dada pela Eq. (9.3.2):

$$f(t) = \frac{1}{\sqrt{2\pi}} \sum_{n=-\infty}^{\infty} C_n e^{i\omega t} \tag{9.3.2'}$$

em que os coeficientes C_n são dados pela Eq. (9.3.3), em que a integração é realizada em um período $2T$. No caso da Eq. (9.3.3), fizemos a integral em $-T \leqslant t \leqslant T$, mas essa escolha foi arbitrária. Poderíamos calcular a integral em $0 \leqslant t \leqslant T$, por exemplo, que o resultado seria o mesmo. Na nossa presente situação, vamos supor que $f(t)$ tem um período T (e não $2T$), e adotar $[0, T]$ como intervalo de integração. Com isso, a Eq. (9.3.3) fica reescrita como

$$C_n = \frac{2\pi}{T} \frac{1}{\sqrt{2\pi}} \int_0^T f(t) e^{-i2\pi nt/T} dt \tag{9.5.1}$$

Exatamente como fizemos na seção 9.4, vamos usar a frequência angular ω_n, dada agora por

$$\omega_n = \frac{2\pi n}{T} \tag{9.5.2}$$

com uma diferença entre dois valores consecutivos de ω dada por

$$\Delta\omega = \frac{2\pi}{T} \tag{9.5.3}$$

Com essas substituições, reescrevemos a Eq. (9.5.1) como

$$C_n = \frac{\Delta\omega}{\sqrt{2\pi}} \int_0^T f(t) e^{-i\omega_n t} dt \tag{9.5.4}$$

Até aqui, tudo é muito parecido com o que fizemos na seção 9.4 e serve para ver se você está acordado. Mas agora veja que Eq. (9.5.4) tem tudo para ser integrada numericamente! Vamos então discretizar o domínio $0 \leqslant t \leqslant \mathrm{T}$ em $N+1$ valores de t, igualmente espaçados por um passo $h = T/N$, chamando os valores obtidos de t_j, dados por

$$t_j = jh \qquad (j = 0, 1, 2, \ldots, N) \tag{9.5.5}$$

que fornece $y_j = f(t_j) e^{-i\omega_n t_j}$ para os valores discretizados dentro da integral. Façamos então a aproximação mais grosseira possível: vamos somar a área dos retângulos de largura $h = T/N$ e altura y_j. Com isso, obtemos uma aproximação para a Eq. (9.5.4) dada por

$$C_n \approx \frac{\Delta\omega}{\sqrt{2\pi}} \sum_{j=0}^{N-1} h f(t_j) e^{-i\omega_n t_j} = \frac{\sqrt{2\pi}}{N} \sum_{j=0}^{N-1} f(jT/N) e^{-i2\pi nj/N} \tag{9.5.6}$$

De modo parecido com o que fizemos na seção 9.4, a Transformada Discreta de Fourier é uma sequência de $N+1$ valores F_n, calculados nas frequências $\omega_n = n\Delta\omega$ (cf. a Eq. 9.5.2) simplesmente pela soma que aparece na Eq. (9.5.6):

$$F(n\Delta\omega) = \boxed{F_n = \sum_{j=0}^{N-1} f(jT/N) e^{-i2\pi nj/N}} \qquad (n = 0, 1, 2, \ldots, N) \tag{9.5.7}$$

Repare que $F_{n+N} = F_n$, razão pela qual, dividindo o período T em N intervalos, ou $N+1$ valores de t, como fizemos, obteremos apenas $N+1$ valores F_n que, por sua vez, correspondem a $N+1$ frequências $\omega_n = 2\pi n/T$. Além disso, como $\Delta\omega$ é fixo, a máxima frequência que podemos considerar é dada por $\Delta\omega N$ (na verdade, é a metade desse valor; falaremos nisso mais adiante). Com essa limitação, a Transformada Discreta de Fourier pode perder frequências presentes no sinal original, maiores que esse valor limite, se não adotarmos um valor de N grande o suficiente (ou seja, se não adotarmos um valor pequeno o suficiente para $\Delta\omega$).

Observe também que a Transformada Discreta de Fourier, assim como a Série de Fourier, acaba considerando que a função $f(t)$ é periódica. Fora do intervalo $0 \leqslant t \leqslant T$, a função original é substituída pela continuação periódica tal que transformamos na verdade uma função $f(t)$ tal que $f(t + kT) = f(t)$ para k inteiro e $t \in [0, T]$.

Em resumo, os dois fatores mais importantes para conseguir uma Transformada Discreta de Fourier bem

feira são: o período de observação T, que deve ser o maior possível; e a amostragem de N pontos dentro do domínio temporal considerado, que também deve ser a maior possível.

E quanto à transformada inversa? Ela é um pouco mais complicada de se obter, apesar de você talvez, por analogia com o caso contínuo, desconfiar de como será a sua cara. Para obtê-la, vamos explorar as propriedades de ortogonalidade das somas envolvendo exponenciais complexas, assim como, no caso contínuo, usávamos o truque de Fourier explorando a ortogonalidade dass integrais. Vamos então começar multiplicando a Eq. (9.5.7) por $e^{i2\pi kn/N}$, sendo k um inteiro qualquer (*qualquer* por enquanto) e somar por todos os valores de n:

$$\sum_{n=0}^{N-1} F_n e^{i2\pi kn/N} = \sum_{n=0}^{N-1}\sum_{j=0}^{N-1} f(jT/N) e^{-i2\pi nj/N} e^{i2\pi kn/N} \tag{9.5.8}$$

e rearranjar um pouco o termo do lado direito:

$$\sum_{n=0}^{N-1} F_n e^{i2\pi kn/N} = \sum_{j=0}^{N-1} f(jT/N) \sum_{n=0}^{N-1} e^{i2\pi n(k-j)/N} \tag{9.5.9}$$

A última soma do lado direito é uma progressão geométrica. Veja no apêndice B como simplificá-la para

$$\sum_{n=0}^{N-1} e^{i2\pi n(k-j)/N} = \frac{e^{i2\pi(k-j)} - 1}{e^{i2\pi(k-j)/N} - 1} \tag{9.5.10}$$

Veja que, se $k \neq j$ na expressão acima, o numerador e, portanto, toda a soma, se anula. Por outro lado, se $k = j$, todas as parcelas da soma são iguais a 1 e, consequentemente, a soma resulta em N. Juntando as duas possibilidades no delta de Kronecker, concluímos que

$$\sum_{n=0}^{N-1} e^{i2\pi n(k-j)/N} = N\delta_{jk} \tag{9.5.11}$$

Jogando esse valor na Eq. (9.5.9), ficamos com

$$\sum_{n=0}^{N-1} F_n e^{i2\pi kn/N} = \sum_{j=0}^{N-1} f(jT/N) N\delta_{jk} = Nf(kT/N) \tag{9.5.12}$$

Com isso, obtemos um conjunto de valores f_k do sinal original, que chamamos de Transformada Discreta Inversa de Fourier, dado por

$$f(kT/N) = \boxed{f_k = \frac{1}{N}\sum_{n=0}^{N-1} F_n e^{i2\pi kn/N}} \qquad (k = 0, 1, 2, \ldots, N) \tag{9.5.13}$$

Repare que $f_{k+N} = f_k$, de forma que recuperamos apenas $N+1$ valores da função original, nos pontos $t_k = kT/N$, a partir dos $N+1$ valores da Transformada Discreta de Fourier.

A Transformada Discreta e sua inversa, dadas, respectivamente, pelas Eqs. (9.5.7) e (9.5.13) foram definidas de forma um pouco diferente em relação às suas versões contínuas da seção 9.4 — o fator $\sqrt{2\pi}$ está ausente das versões discretas. Se quisermos comparar as duas versões, precisamos adaptá-las multiplicando-as ou dividindo pelo fator de normalização. Isso, no entanto, não é tão importante, pois as transformadas serão iguais a menos de um fator multiplicativo.

9.5.2 A Transformada Rápida de Fourier (FFT)

Armados com uma implementação da Transformada Discreta de Fourier, poderíamos começar a aplicá-la a alguns exemplos. No entanto, implementar um algoritmo que calcula as transformadas direta e inversa usando as (9.5.7) e (9.5.13) diretamente é enormemente ineficiente. Vamos analisar a Eq. (9.5.7), por exemplo, pois a análise da Eq. (9.5.13) não difere em nada. Precisamos calcular $N+1$ termos F_n, cada um dado por uma soma de N termos. Assim, o número de operações necessário ao cálculo da Transformada Discreta de Fourier deve ser proporcional a $N(N+1)$, que é aproximadamente igual a N^2 para grandes valores de N. Aumentar N em dez vezes, portanto, significaria multiplicar por cem o número de operações e, portanto, também o tempo de cálculo — um pesadelo numérico. Assistir a vídeos em HD pela internet, por exemplo, estaria fora de questão, se isso dependesse de um algoritmo ingênuo como esse.

© Luiz T. F. Eleno, v. 0.3, 2021

A Transformada Rápida de Fourier (FFT, de *Fast Fourier Transform*) é um algoritmo incrivelmente esperto que reduz drasticamente a complexidade numérica do cálculo da Transformada Discreta de Fourier. Na FFT, o número de operações é da ordem de $N \log_2 N$. Isso significa que aumentar em dez vezes a amostragem de pontos implica, se N é grande, num aumento da mesma ordem de grandeza. Existem hoje diversas implementações diferentes para se realizar a FFT e sua inversa. Mas a ideia original veio de Cooley e Tukey [3], num brilhante trabalho de 1965 que abriu a estrada para quase inacreditáveis avanços na informática e nas telecomunicações.

Aqui, estamos mais interessados nas aplicações da FFT do que em sua dedução e implementação. Por isso, vamos descrevê-la apenas em linhas gerais. Vamos então recapitular o problema da Transformada Discreta de Fourier: é preciso fazer muitas contas de acordo com a definição dada por somas de N termos. No entanto, se você observar a Eq. (9.5.7) mais de perto,

$$F_n = \sum_{j=0}^{N-1} f(jT/N) e^{-i2\pi nj/N} \qquad (n = 0, 1, 2, \ldots, N) \tag{9.5.7′}$$

verá que existem muitas exponenciais complexas repetidas. Isso porque existem dois índices nas exponenciais: $e^{-i2\pi nj/N}$ depende de n e j, ambos variando de 0 a $N-1$. Recalculá-los para todos os termos é perda de tempo e deve ser evitado. Além disso, Cooley e Tukey [3] perceberam que, escolhendo um N adequado, poderiam dividir a soma em várias somas parciais, por exemplo, para j par e j ímpar:

$$F_n = \sum_{j\ \text{par}} f(jT/N) e^{-i2\pi nj/N} + \sum_{j\ \text{ímpar}} f(2jT/N) e^{-i2\pi n2j/N} \tag{9.5.14}$$

e continuar fazendo isso recursivamente, sempre separando cada soma em duas parciais, até que cada soma se torne apenas um termo! Para isso, é preciso que N seja uma potência de 2, ou seja, $N = 2^m$ (se não for o caso, basta completar a soma original com zeros até atingir a próxima potência de 2). Assim, na verdade, fazemos apenas uma soma de N termos, aproveitando o cálculo de um dos F_n para os demais. O processo de separar recursivamente as somas em somas parciais adiciona apenas mais algumas operações ao cálculo da FFT. Usando essa estratégia, e com uma contabilidade cuidadosa dos expoentes das exponenciais complexas, Cooley e Tukey [3] mostraram que o método realmente é extremamente rápido, como dissemos no início da seção.

Alguma versão da FFT está implementada em praticamente todas as linguagens de programação. Apenas em python existem pelo menos duas versões: uma na biblioteca `numpy` e outra versão na `scipy.fftpack`.[2] Gostaria de dizer que poderíamos começar a usá-las sem maiores delongas, mas antes precisamos entender alguns efeitos que afetam a acurácia da FFT.

9.6 Propriedades da FFT

Na próxima seção apresentaremos alguns exemplos de aplicação da FFT a alguns problemas interessantes da Física Clássica. Antes da diversão, no entanto, precisamos investigar alguns detalhes numéricos do método, para melhor entender o seu uso e evitar alguns percalços em que os neófitos no assunto (e até mesmo alguns usuários supostamente experientes...) acabam tropeçando. Se quiser acompanhar as próximas seções fazendo também, você mesmo, os cálculos da FFT no computador, consulte o material de apoio para saber como mandar o python fazer isso por você.

9.6.1 Simetria e a frequência de Nyquist

Volte um instante à definição da Transformada Discreta de Fourier dada pela Eq. (9.5.7):

$$F(\omega_n) = F_n = \sum_{j=0}^{N-1} f(jT/N) e^{-i2\pi nj/N} \qquad (n = 0, 1, 2, \ldots, N) \tag{9.5.7′}$$

Como vimos, a FFT é basicamente uma maneira rápida de calcular essa soma quando N é uma potência inteira de dois. Lembre que a Eq. (9.5.7) faz uma amostragem de N pontos de $f(t)$ no intervalo $0 \leqslant t \leqslant T$. Fora desse intervalo, a função é repetida periodicamente e, com isso, estamos sempre transformando uma função periódica — o que não pode ser evitado, apenas mitigado se usamos um período T grande o suficiente. Por esse motivo, a Transformada Discreta de Fourier será também uma função periódica. Além

[2] Documentação encontrada em https://docs.scipy.org/doc/numpy-1.15.0/reference/routines.fft.html e https://docs.scipy.org/doc/scipy/reference/fftpack.html, respectivamente.

disso, ela apresenta uma simetria interessante que não é imediatamente evidente apenas olhando para a Eq. (9.5.7). Essa propriedade nos diz que

$$F_{N-n} = F_{-n} = F(-\omega_n) \tag{9.6.1}$$

Você consegue demonstrar a Eq. (9.6.1)? Troque n por $N-n$ na Eq. (9.5.7) e veja o que acontece:

$$F_{N-n} = \sum_{j=0}^{N-1} f(jT/N) e^{-i2\pi(N-n)j/N} = \sum_{j=0}^{N-1} f(jT/N) e^{i2\pi(-n)j/N} \cancelto{1}{e^{-i2\pi j}} = F_{-n} \tag{9.6.2}$$

O que a Eq. (9.6.1) nos diz é que teremos uma sequência de valores $F(\omega_n)$ para para $0 \leqslant n \leqslant N/2$ que podemos chamar de *região significativa* da transformada. Já os valores de F_n para n no intervalo $N/2 < n \leqslant N$ são apenas cópias dos valores da região significativa, correspondendo a $F(-\omega_n)$. Por isso, poderíamos, se quiséssemos, inverter a ordem dos valores F_n de tal forma que escrevêssemos primeiro os valores para as frequências negativas (os valores F_{N-n}), seguidos dos valores para frequências positivas (F_n). Ou podemos, simplesmente, analisar a região significativa, já que a outra região não traz muita informação nova. A metade do espectro de frequências é conhecida como *frequência de Nyquist*, ω_{Nyquist}, correspondendo a $n = N/2$, ou seja,

$$\omega_{\text{Nyquist}} = \frac{2\pi}{T}\frac{N}{2} = \frac{N\pi}{T} \tag{9.6.3}$$

Esse valor é metade do que dissemos na discussão após a Eq. (9.5.7). Com isso, a Transformada Discreta de Fourier (e a FFT) calcula um espectro de frequências no intervalo $-\omega_{\text{Nyquist}} \leqslant \omega_n \leqslant \omega_{\text{Nyquist}}$. Qualquer frequência maior que essa (em módulo) é perdida pela Transformada Discreta de Fourier.

9.6.2 Subamostragem, *aliasing* e dispersão

Digamos que você esteja interessado em calcular a FFT da função $f(t) = \operatorname{sen} 8t$. Já vimos que a transformada exata é dada pela Eq. (9.4.21) e consiste de dois picos na frequência de vibração (positiva e negativa) que, no presente caso, é ± 8. Para calcular a FFT, precisamos determinar um intervalo de tempo $0 \leqslant t \leqslant T$ que determinará a continuação periódica de $f(t)$. Para simplificar, vamos começar adotando $T = 2\pi$, que é um período de $\operatorname{sen} 8t$, pois $\operatorname{sen}(8t + 2\pi) = \operatorname{sen} 8t$. Como a função é periódica, e como a FFT repete a função periodicamente fora do intervalo considerado, isso é exatamente o que queremos. De fato, escolher T como um múltiplo inteiro do período de $f(t)$ (se $f(t)$ é periódica) é sempre uma boa escolha (veremos em breve o efeito de outras escolhas para T).

Uma vez feita a escolha para T, precisamos fazer uma amostragem dividindo o intervalo $[0, t]$ em N partes iguais que definirão os instantes $t_j = jT/N$. Se você, por pressa, escolher, por exemplo, $N = 4$, a função será calculada em $t_0 = 0$, $t_1 = \pi/2$, $t_2 = \pi$, $t_3 = 3\pi/2$ e $t_4 = 2\pi$ — que fornecem sempre o mesmo valor para $\operatorname{sen} 8t_n$. Com isso, sua amostragem não captou o comportamento da função no intervalo $[0, T]$, pois a FFT (que não conhece $f(t)$, mas apenas os valores $f(t_n)$ que *você* fornece!) pensará que a sua função é constante. Você precisa, portanto, estar atento a efeitos espúrios oriundos de uma *subamostragem* da função.

Um problema relacionado à subamostragem leva a um outro efeito chamado *aliasing*, que não tem uma boa tradução em português nesse contexto. Considere como exemplo novamente a função $f(t) = \operatorname{sen} 8t$. Você já percebeu que usar $N = 4$ é uma péssima escolha e decide aumentar a amostragem para $N = 10$ (ignore por um instante que esse valor não é uma potência inteira de dois; a Transformada Discreta funciona da mesma maneira). Mas isso ainda não é o suficiente! Veja na Figura 9.8 que, nos pontos escolhidos com $N = 10$, as funções $g(t) = -\operatorname{sen} 2t$ e $f(t) = \operatorname{sen} 8t$ fornecem os mesmos valores. Mas já vimos que a máxima frequência captada pela Transformada Discreta de Fourier é a frequência de Nyquist, dada pela Eq. (9.6.3). Escolhendo $N = 10$ num período 2π, portanto, acarreta $\omega_{\text{Nyquist}} = 5$, e perdemos a frequência da nossa função $f(t) = \operatorname{sen} 8t$. Nossa transformada de Fourier pensa que a função é na verdade $g(t) = -\operatorname{sen}(2t)$, cuja frequência é 2. Esse é o fenômeno de *aliasing*, em que uma função se passa por outra, como se adotasse um codinome (*alias*, em inglês). No exemplo, $N = 32$ é a mínima amostragem para que não ocorram efeitos de subamostragem e *aliasing* (por que $N = 16$ não funciona?).

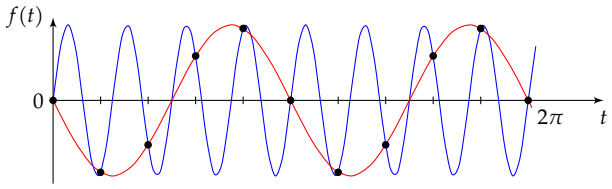


FIGURA 9.8: As duas funções acima, $f(t) = \operatorname{sen} 8t$ (azul) e $g(t) = -\operatorname{sen} 2t$ (vermelho), fornecem os mesmos valores nos pontos mostrados (•), igualmente espaçados de $\pi/5$ ao longo do eixo t no intervalo $[0, 2\pi]$.

Ainda que você tenha usado uma amostragem adequada, podem aparecer problemas se você não escolher bem o período T. Digamos que a sua função seja periódica, por exemplo, $f(t) = \operatorname{sen} 3t$. Até aqui

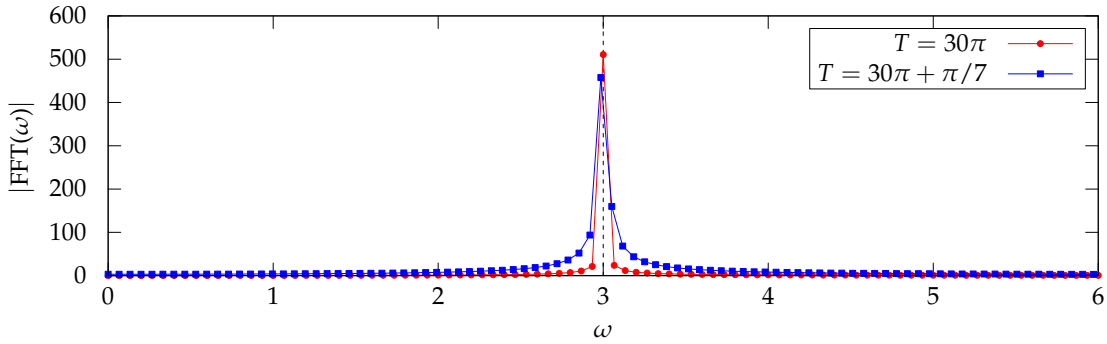


FIGURA 9.9: Transformadas Discretas de Fourier da função $f(t) = \operatorname{sen} 3t$ com uma amostragem de $N = 2^{10}$ mas usando dois períodos: $T = 32\pi$, múltiplo inteiro do período da função, e $T = 32\pi + \pi/7$. A FFT fornece valores complexos, por isso os gráficos mostram a a amplitude espectral $|\text{FFT}(\omega)|$.

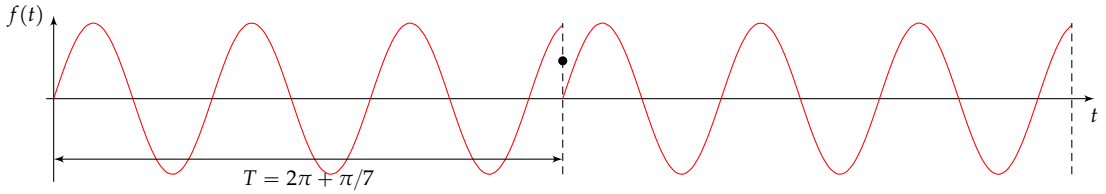


FIGURA 9.10: Para calcular a FFT de $f(t) = \operatorname{sen} 3t$, é preciso fazer a continuação periódica. Se T não é um múltiplo do período da função, aparecerão descontinuidades entre cada período. A Série de Fourier converge para o valor médio (•) nesses lugares. Isso introduz frequências espúrias na FFT.

adotamos sempre $T = 2\pi$, que é um possível período dessa função. Mas e se, mesmo usando um N suficientemente grande, escolhermos um T diferente, que não seja um múltiplo inteiro do período da sua função? Nessas situações, é comum acontecer um fenômeno conhecido como *dispersão* (em inglês, usa-se o termo *leakage*, vazamento). Vamos então fazer dois cálculos da FFT da função $f(t) = \operatorname{sen} 3t$, mostrados na Figura 9.9. Em ambos usamos a mesma amostragem $N = 2^{10} = 1024$, um valor alto o suficiente para evitar os problemas descritos anteriormente. No primeiro cálculo, adotaremos um valor de T bastante alto, digamos, $T = 30\pi$, de forma a observar a função por diversos períodos. Esse valor de T é múltiplo inteiro do período da função. No segundo cálculo, no entanto, vamos usar $T = 30\pi + \pi/7$, que não é um número inteiro de períodos completos, mas um pouco maior. Os resultados de ambos os cálculos estão mostrados na Figura 9.9. Repare que a FFT retorna valores complexos, por isso plotamos a *amplitude espectral* $|\text{FFT}(\omega)|$. Você pode perceber que a curva para $T = 30\pi$ acerta o pico exatamente sobre o valor esperado, $\omega = 3$. Além disso, ela tem uma *dispersão* marcadamente menor que a outra curva, que, além de errar ligeiramente a posição do pico, ainda introduz uma considerável dispersão nos valores de ω. Lembre que a transformada contínua teria um único pico de função delta de Dirac na frequência $\omega = 3$ (e outro em $\omega = -3$, não mostrado), o que significa que a dispersão teórica deveria ser nula. Obviamente, sempre haverá alguma dispersão na FFT, pois a Transformada Discreta de Fourier é, como vimos, uma integração numérica usando a aproximação mais grosseira possível.

Qual o motivo da dispersão quando o período T difere do período da função? A resposta é dada na Figura 9.10, em que vemos um gráfico de $f(t) = \operatorname{sen} 3t$ no período escolhido, $T = 30\pi + \pi/7$, e sua continuação periódica para maiores valores de t. Entre dois períodos, aparecem descontinuidades, como mostra a Figura 9.10. Lembre-se da seção 9.1 que a Série de Fourier converge para o ponto médio entre os valores de $f(t)$ de cada lado da descontinuidade. A Transformada Discreta (e a FFT), que é muito parecida com uma Série de Fourier, tem a mesma tendência. Isso leva à necessidade de mais coeficientes C_n na série de Fourier e, equivalentemente, a mais frequências importantes aparecendo na FFT e dispersando o pico — ou *vazando* para a transformada, usando a analogia feita pelos anglófonos que, como dissemos antes, chamam esse fenômeno de *leakage* (vazamento).

Em conclusão: para se determinar com confiança a Transformada Discreta de Fourier, usando a FFT ou qualquer outro algoritmo, é preciso sempre analisar cuidadosamente os parâmetros N e T, para evitar os problemas que encontramos aqui: subamostragem, *aliasing* e dispersão. Na próxima seção veremos o cálculo da FFT aplicado a alguns fenômenos de interesse, e teremos oportunidades de demonstrar como os fatores que discutimos aqui têm um grande impacto sobre a *qualidade* dos resultados fornecidos pela FFT.

9.7 A FFT em ação

9.7.1 Modos normais

Se você achou que já tinha deixado o capitulo 2 para trás, reveja seus conceitos. Modos normais de oscilação são um prato cheio para ver a FFT trabalhando pois, como vimos, são uma combinação linear de vibrações em várias frequências. A Transformada de Fourier então deveria fornecer picos exatamente nessas frequências, e apenas esses picos.

Vamos fazer o teste com três osciladores acoplados, que vimos ao final da seção 2.6. A Figura 2.11 mostra o intrincado movimento de cada oscilador. É difícil determinar as frequências de vibração apenas olhando para os gráficos da posição e da velocidade. A transformada de Fourier, por outro lado, serve exatamente para isso. Vamos então fazer uma amostragem da posição da primeira partícula, $x_1(t)$, em N pontos num período T, escolhendo-os com cuidado para não perder informação, agora que já sabemos o que pode acontecer e não tivermos cuidado, como vimos na seção 9.6.2. Da Figura 2.11, vemos que os osciladores voltam à posição e velocidades iniciais num instante dado aproximadamente por 57.8 (unidades de tempo arbitrárias). Na FFT, portanto, é melhor usar múltiplos inteiros desse período. Para determinar N, vamos calcular a frequência de Nyquist. A Eq. (9.6.3) fornece $\omega_{\text{Nyquist}} = N\pi/T$, de onde podemos extrair o mínimo valor de N em função do período T adotado. Lembre-se que ω_{Nyquist} é a máxima frequência que a FFT pode detectar. Se adotarmos $N = 2^{10} = 1024$ (um número bastante usual para a FFT) e $T = 20 \times 57.8$ (20 períodos, para ter uma contagem bastante alta), teremos $\omega_{\text{Nyquist}} \approx 2.78$. Se a FFT apresentar três picos, saberemos que usamos um N alto o suficiente (são três modos normais, lembre-se!).

Com os parâmetros estabelecidos, nada nos impede de calcular a FFT. O resultado está mostrado na Figura 9.11. Novamente, como a FFT retorna valores complexos, plotamos o módulo dos valores. Repare também que plotamos a FFT apenas para os valores positivos de ω, já que a curva é simétrica. Realmente vemos apenas três picos na FFT, indicando que o movimento da partícula é composto pela combinação linear de três modos normais com diferentes frequências. Se você voltar às Eqs. (2.6.26) e (2.6.29), verá que essas frequências, em unidades de ω_0 e para $N = 3$ osciladores, são dadas por

$$\begin{aligned} \omega_1 &= \sqrt{2-\sqrt{2}} \approx 0.765 \\ \omega_1 &= \sqrt{2} \approx 1.414 \\ \omega_1 &= \sqrt{2+\sqrt{2}} \approx 1.848 \end{aligned}$$

que são, de fato, as posições dos três picos na Figura 9.11, adotando $\omega_0 = 1$. Obviamente, a transformada exata deveria fornecer três picos de função delta de Dirac centradas nessas frequências. A FFT, por outro lado, sempre apresentará algum grau de dispersão, como podemos observar também na Figura 9.11. Parte da dispersão vem do fato de que as três frequências são incomensuráveis e, portanto, não há um período completo de oscilação, apenas aproximações (o período igual a 57.8, por exemplo, que usamos como base para a FFT). De todo modo, podemos mitigar esse efeito aumentando o período de observação T e a amostragem N, tomando sempre os cuidados necessários em relação a ω_{Nyquist} e mantendo T um número inteiro de períodos do sinal (ou seja, da função $x_1(t)$, a posição do primeiro oscilador).

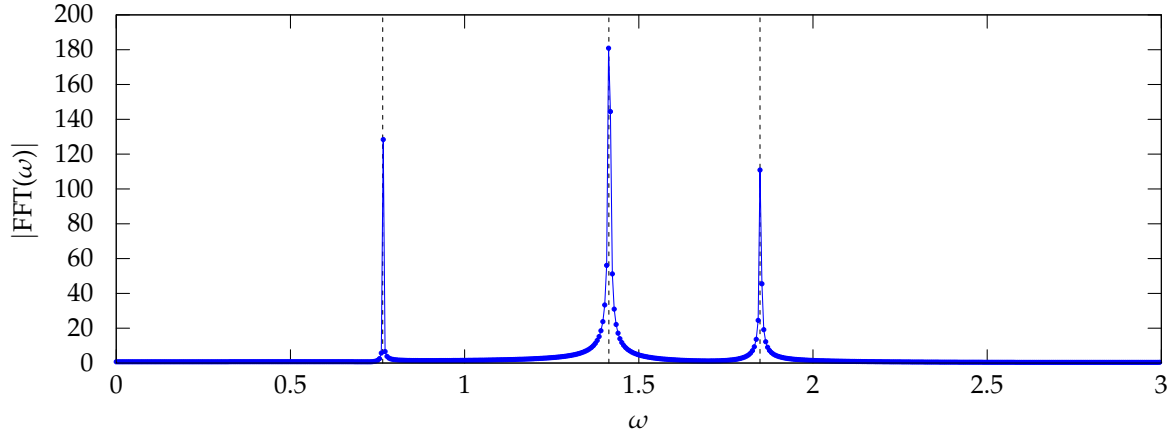


FIGURA 9.11: Transformada Discreta de Fourier da posição $x_1(t)$ do primeiro de um conjunto de três osciladores harmônicos idênticos acoplados, com $\omega_0 = 1$. Usamos como parâmetros $N = 2^{10}$ e $T = 20 \times 57.8$ (unidades arbitrárias). As posições dos picos estão indicadas pelas linhas tracejadas verticais.

9.7.2 O oscilador de van der Pol

Você encontrou o oscilador de van der Pol no exercício 8.4 (caso o tenha resolvido). Trata-se de um oscilador não-linear, o que significa que seu movimento é descrito por uma EDO não-linear. No caso do oscilador de van der Pol, essa EDO é dada por

$$\ddot{x} + \varepsilon(x^2 - 1)\dot{x} + x = 0 \tag{9.7.1}$$

em que ε é um parâmetro numérico variável. Se $\varepsilon = 0$, temos um oscilador harmônico, cuja solução envolve apenas a frequência $\omega = \pm 1$. A Transformada de Fourier nesse caso será, como já sabemos, deltas de Dirac centrados nesses valores. À medida em que ε se distancia de zero, no entanto, mais e mais frequências devem ser ativadas e devem aparecer na FFT.

Vamos usar como exemplo um oscilador de van der Pol para o qual $\varepsilon = 1.5$ e, portanto, com um grau pronunciado de não-linearidade. Para resolver a EDO, vamos adotar como estado inicial a condição $x(1)$, $\dot{x}(0) = 0$ e usar um dos métodos numéricos do capítulo 8 (na verdade, a função `odeint` da biblioteca `scipy.integrate`). Um gráfico da posição do oscilador ao longo do tempo é mostrado na Figura 9.12. Veja que, à exceção dos primeiros instantes após a condição inicial, o gráfico da Figura 9.12, apesar de não ter um aspecto exatamente sinusoidal, tem um claro comportamento periódico, com montes e vales mais parecendo uma fileira de tubarões.

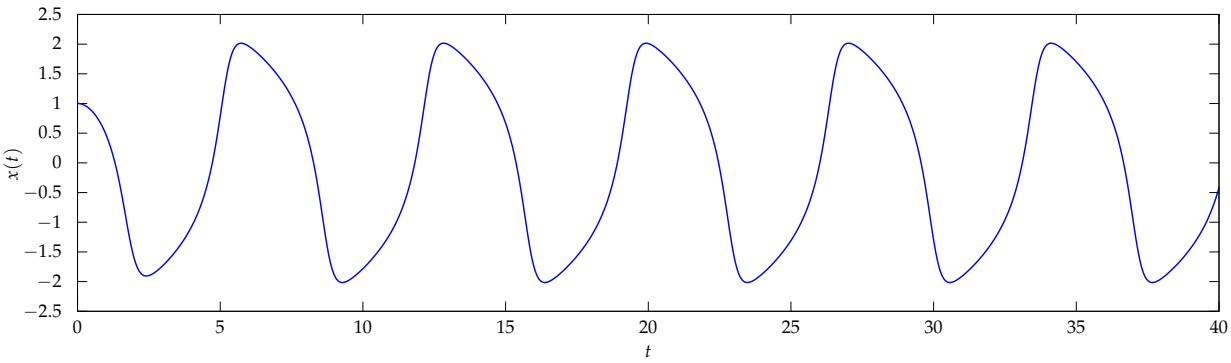


FIGURA 9.12: Posição do oscilador de van der Pol com $\varepsilon = 1.5$ ao longo do tempo, para a condição inicial $x(0) = 1$, $\dot{x}(0) = 0$. A EDO foi resolvida numericamente em python usando a função `odeint` da biblioteca `scipy.integrate`.

O espaço de fases do oscilador, ou seja, um gráfico da velocidade em função da posição numa sucessão de instantes, está representado na Figura 9.13. O conceito de espaço de fases é importante em Mecânica Clássica mas, se você ainda não o encontrou, não faz muita diferença aqui. De qualquer maneira, veja que o oscilador começa seu movimento num regime transiente a partir da condição inicial e chega a um estado estacionário após algum tempo. O estado estacionário forma um laço fechado no espaço de fases, da qual o oscilador não mais escapa. Por uma técnica de tentativa e erro, você consegue sem muito esforço descobrir que o período de oscilação no estado estacionário é de aproximadamente 7.0875 (unidades arbitrárias de tempo). Para isso, resolva a EDO para um período inicial grande de tempo. Com isso, você garante que o oscilador chegou ao estado estacionário. Depois, use a condição atingida ao final do último instante como estado inicial para mais uma sequência de pontos, mas, dessa vez, vá aumentando gradativamente o tempo t até que o oscilador complete um ciclo. O tempo que ele leva para fazer isso é o período do oscilador de van der Pol.

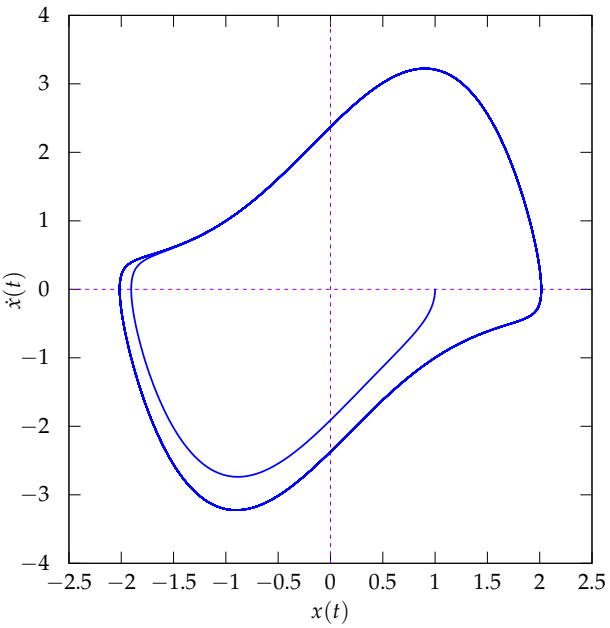


FIGURA 9.13: Espaço de fases do oscilador de van der Pol nas mesmas condições da Figura 9.12.

O estado estacionário é o que importa na maior parte das aplicações. Para estudar o espectro de frequências do oscilador, portanto, precisamos desprezar os primeiros instantes, em que o oscilador ainda está passando pelo regime transiente, e calcular a FFT apenas

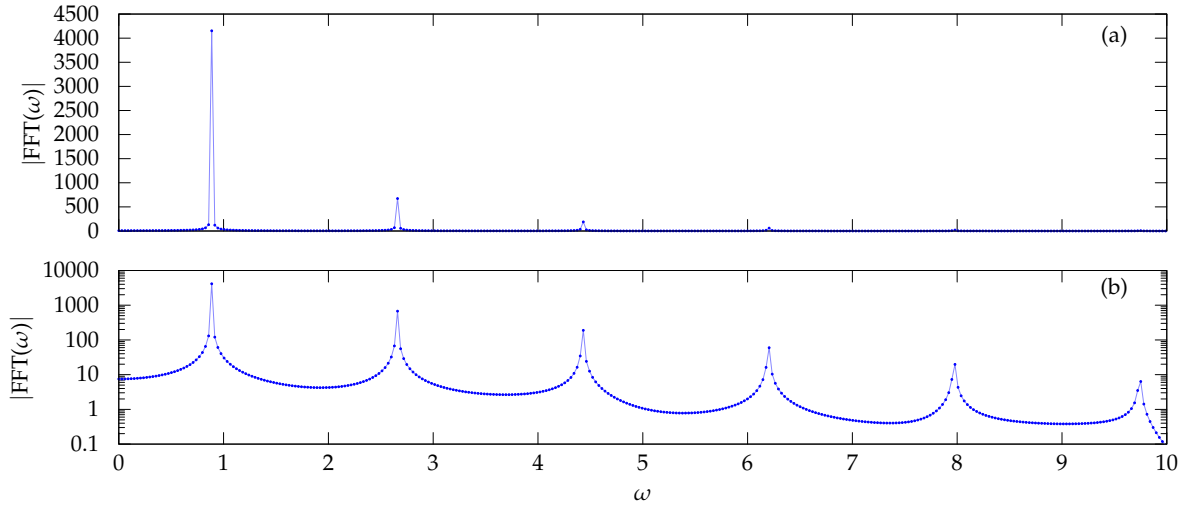


FIGURA 9.14: Transformada Discreta de Fourier, em escalas linear e logaritmica, da posição $x(t)$ do oscilador de van der Pol com $\varepsilon = 1.5$ após atingido o regime permanente. Usamos como parâmetros $N = 2^{12}$ e $T = 30 \times 7.0875$ (unidades arbitrárias).

após esses primeiros passos. Com isso, a FFT mostrará apenas as frequências do regime permanente, que, provavelmente, são as frequências de interesse prático.

A Figura 9.14 mostra o resultado de um cálculo da FFT com um período $T = 30 \times 7.0875$, ou seja, durante trinta ciclos completos do oscilador de van der Pol, com uma amostragem de $N = 2^{12}$ pontos nesse intervalo. Esses parâmetros, como discutimos anteriormente, são adequados para evitar *aliasing* e dispersão espúrias, além de fornecerem um espectro denso de pontos no resultado da FFT. A Figura 9.14a mostra que há uma frequência predominante, ou *frequência principal*, próxima de $\omega = 1$ (na verdade, em $\omega = 2\pi/7.0875 = 0.8865$), com amplitude praticamente uma ordem de grandeza maior que a da próxima frequência, perto de $\omega \approx 2.6$. Há ainda outros picos de intensidade muito menor, quase invisíveis na parte (a) da figura, mas claramente discerníveis quando plotamos o espectro em escala logarítmica, como mostra a Figura 9.14b. A escala logarítmica tem o efeito de diminuir contrastes entre as ordens de grandeza dos dados, o que vem bem a calhar nesse caso, em que as os picos do espectro tem intensidades muito diferentes.

Ao conjunto de frequências de um oscilador, ou de um sinal, de maneira geral, damos o nome de *harmônicos*. O nome é especialmente aplicável a sinais de áudio, como uma nota musical tocada num instrumento qualquer. Os harmônicos de uma nota lá tocada num piano são diferentes dos daqueles da mesma nota tocadas num trompete, por exemplo. Eles ajudam a definir o *timbre* de um instrumento musical.

Vemos da Figura 9.14 que o oscilador de van der Pol, com $\varepsilon = 1.5$, apresenta um conjunto de harmônicos bem-definidos e praticamente igualmente espaçados. São essas as frequências que determinam o comportamento não linear do oscilador. Acima de $\omega = 10$ há outros harmônicos no espectro do oscilador de van der Pol não mostrados na Figura 9.14. Por terem uma intensidade baixíssima, esses harmônicos, na prática, não tem muita importância.

9.7.3 O Atrator de Lorenz e janelamento

Na seção 8.5.1 vimos que o sistema de três EDOs propostas por E. Lorenz leva a um comportamento caótico. Da Figura 8.7 não é possível enxergar um padrão periódico. Dentre as três coordenadas, $z(t)$ é a que mais se aproxima de um comportamento quase-oscilatório. Vamos então investigar seu espectro de frequências usando a FFT.

Antes mesmo de começar o cálculo da FFT, no entanto, chegamos imediatamente a um problema: se $z(t)$ não é periódica, como escolher T para evitar ao máximo a dispersão? Simplesmente usar um T muito grande, com uma amostragem N bastante densa, não é o suficiente, apesar de ajudar. Em casos complicados como esse, uma boa estratégia, a adotar é o *janelamento*. Algumas das janelas mais comumente utilizadas são mostradas na Figura 9.15. O

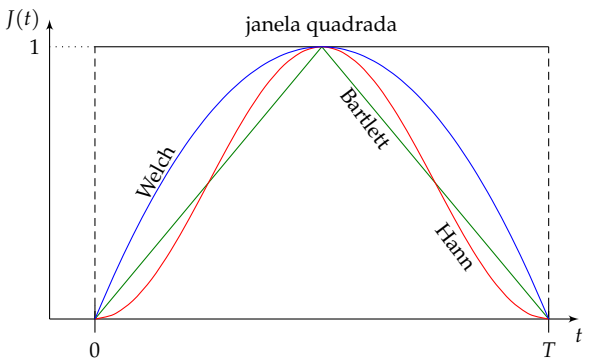


FIGURA 9.15: Algumas janelas comumente usadas para diminuir a dispersão do cálculo da FFT.

janelamento consiste em multiplicar o sinal original por uma função-janela $J(t)$, ou simplesmente *janela*, que tem o efeito de minimizar o efeito das bordas. Repare que todas elas (à exceção da janela quadrada) vão a zero nos limites do intervalo, ou seja, em $t = 0$ e em $t = T$. As expressões matemáticas são dadas a seguir, todas relevantes apenas no intervalo $0 \leqslant t \leqslant T$:

Janela quadrada: Essa janela, na verdade equivale a não fazer nada, pois

$$J(t) = 1 \tag{9.7.2}$$

Janela de Bartlett: é uma janela triangular,

$$J(t) = 1 - \left|1 - \frac{2t}{T}\right| \tag{9.7.3}$$

Janela de Welch: uma janela parabólica,

$$J(t) = 4\frac{t}{T}\left(1 - \frac{t}{T}\right) \tag{9.7.4}$$

Janela de Hann: uma janela um pouco mais complicada,

$$J(t) = \text{sen}^2\,\frac{\pi t}{T} \tag{9.7.5}$$

Aplicar uma janela aos dados significa obter um sinal janelado $s_j(t)$, dado por

$$s_j(t) = s(t)J(t) \tag{9.7.6}$$

sendo $s(t)$ nosso sinal original, possivelmente não-periódico. Repare que que todas as janelas (à exceção da janela quadrada, que não faz nada!), ao serem multiplicadas pelo sinal original, tem o efeito de reduzir as intensidades dos valores do sinal próximos das extremidades. Com isso, o sinal janelado passa a ser uma função contínua e periódica de t. Ao invés de calcular a FFT do sinal original $s(t)$, calculamos a FFT do sinal janelado. Como a descontinuidade nesses pontos é responsável pela dispersão na FFT, o janelamento tem por consequência uma dimimuição da dispersão.

Cuidado! É sempre melhor procurar por periodidicidades no sinal original ao invés de janelar irrefletidamente. Usar um múltiplo inteiro do período do sinal (quando ele existe) é sempre mais eficiente do que usar uma janela sem muita atenção.

Por outro lado, janelar a solução da EDO de Lorenz é adequado, pois ela parece não ter uma periodicidade muito claramente definida. Vamos então calcular a FFT da coordenada $z(t)$, mostrada nas Figuras 8.6 e 8.7, usando uma janela. Existem alguns critérios para se escolher entre as inúmeras janelas propostas na literatura. Todos os critérios são, até certo ponto, subjetivos; uma janela, que fornece bons resultados numa situação particular, pode não ser adequada em outras. Apenas como ilustração, vamos usar a janela de Bartlett, dada pela Eq. (9.7.3), que é uma função triangular.

Assim como fizemos com o oscilador de van der Pol, vamos deixar o oscilador evoluir por alguns instantes, para eliminar o que conseguirmos dos efeitos da condição inicial, e transformar apenas $z(t)$ em instantes posteriores. No entanto, não vamos transformar diretamente $z(t)$, mas um sinal $s(t)$ dado por

$$s(t) = z(t) - \langle z(t)\rangle \tag{9.7.7}$$

onde $\langle z(t)\rangle$ é o valor médio de $z(t)$ no intervalo $[0, T]$ a ser considerado para a FFT. Repare que, definido dessa forma, $\langle s(t)\rangle$, o valor médio de $s(t)$, é nulo. Mas por que fazer essa mudança? Porque, se você olhar para a Figura 8.7, verá que o valor médio de $z(t)$ é bastante diferente de zero. Se você quiser, pode inverter a história e escrever $z(t)$ como

$$z(t) = s(t) + \langle z(t)\rangle \tag{9.7.8}$$

Chamando as transformadas de $z(t)$ e $s(t)$ de $Z(\omega)$ e $S(\omega)$, respectivamente, e usando a propriedade de linearidade da Transformada de Fourier, teremos

$$Z(\omega) = S(\omega) + \langle z(t)\rangle\,\mathcal{F}[1] \tag{9.7.9}$$

sendo $\mathcal{F}[1]$ a Transformada de Fourier da função unitária (repare que $\langle z(t)\rangle$ é uma constante). E qual é a transformada de $f(t) = 1$? Se você fizer a conta, verá que ela é dada por

$$\mathcal{F}[1] = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-i\omega t}dt = \sqrt{2\pi}\,\delta(\omega) \tag{9.7.10}$$

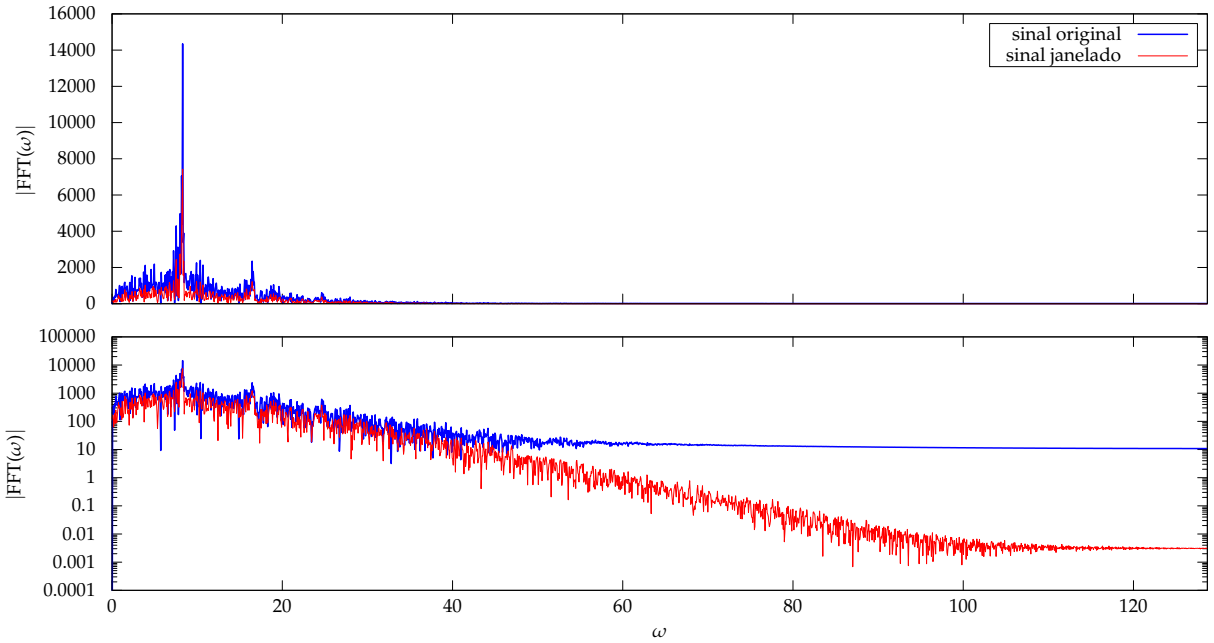


FIGURA 9.16: FFT da coordenada $z(t) - \langle z(t) \rangle$ do sistema de EDOs de Lorenz, cujos detalhes foram mostrados na Figura 8.7. No painel inferior, a escala logarítmica realça frequências de menor amplitude. Estão mostradas as FFTs do sinal original e do sinal com janelamento de Bartlett.

usando a definição da função delta de Dirac encontrada na Eq. (9.4.20). Isso significa que $\mathcal{F}(1)$ consiste num pico bastante pronunciado centrado em $\omega = 0$. Esse pico aparece também na transformada de $z(t)$, pela Eq. (9.7.9). Se não o filtrarmos, removendo-o do sinal, esse pico irá mascarar todos os demais picos da função $z(t)$. A Eq. (9.7.7), por outro lado, garante que o pico em $\omega = 0$ estará ausente da transformada de $s(t)$, cujo valor médio, como dissemos, é zero.

O resultado da FFT do sinal filtrado e janelado está mostrado na Figura 9.16, em escala linear e logarítmica. Adotamos inicialmente valores de N e T de forma a obter uma frequência de Nyquist bastante alta, para conseguir uma visão abrangente do espectro. A FFT do sinal não janelado também está mostrada. Veja que existem vários picos em frequências no intervalo $0 \leqslant \omega \leqslant 40$, o que significa que o sistema de Lorenz tem, de fato, algumas características oscilatórias, com vários picos ao redor de $\omega \approx 8$ e outras regiões com picos de menor intensidade em outras frequências. Para frequências maiores, no entanto, as intensidades vão se reduzindo até que, mesmo em escala logarítmica, as amplitudes atingem uma linha base. Para o sinal não janelado, essa linha base está numa amplitude entre 10 e 100. O sinal janelado, por outro lado, chega a uma linha base bem menor, abaixo de 0.01. Isso indica a eficácia do filtro em eliminar parte da dispersão da FFT. Outros filtros teriam um efeito parecido, com variações que, para o nosso propósito aqui, são irrelevantes.

9.8 Ruído

Como vimos na última seção, a Figura 9.16 mostra uma concentração das maiores amplitudes no intervalo de baixas frequências. No entanto, o sistema de Lorenz exibe frequências num espectro bem amplo antes de atingir a linha base. Sinais cujas transformadas exibem um intervalo grande de frequências importantes são chamados de *ruído* (em inglês, *noise*), à diferença de sinais como uma onda sonora, cuja transformada tende a exibir picos bem centralizados e com pouca dispersão ao redor de frequências bem-definidas. O termo *ruído branco* (*white noise*) é usado para descrever um sinal com frequências de amplitude similar espalhadas por todo o espectro. Já o *ruído vermelho*, em analogia ao espectro da luz visível, é um sinal em cujo espectro predominam as baixas frequências. O espectro do atrator de Lorenz, nesse sentido, contém uma grande quantidade de ruído vermelho, como mostra a Figura 9.16.

Digamos que temos um sinal de ruído branco, $r(t)$. Qual será a sua Transformada de Fourier? Veja que a transformada de um sinal com uma única frequência, como $f(t) = \operatorname{sen} t$, é um pico centrado nessa frequência. O ruído branco, por outro lado, contém, idealmente, todas as frequências em um dado intervalo (na prática, contém muitas e muitas frequências). A transformada nesse caso fornece um grande conjunto de picos, de amplitude parecida e muito próximos um aos outros. No limite, a transformada, portanto, se aproxima de uma linha reta horizontal.

De certa maneira, o ruído está sendo eliminado da tecnologia moderna. Antigamente, um televisor sem antena exibia uma imagem de riscos e pontos, com um chiado característico. Isso é ruído brando, da mesma

maneira que o chiado de uma estação de rádio mal-sintonizada. Ao viajar de avião, você tem problemas para dormir devido ao grande nível de ruído vermelho (de frequências baixas e médias, equivalente a um ruído relativamente grave). Mas o ruído não precisa ser considerado desagradável. Alguns instrumentos de percussão, por exemplo, produzem sons que são praticamente ruído. O som das gotas de uma chuva torrencial caindo sobre o solo também é muito próximo de um sinal de ruído tendendo ao vermelho.

No *site* https://mynoise.net/NoiseMachines/whiteNoiseGenerator.php você pode brincar com um espectro de ruídos e, alterando os diversos canais de frequência (graves, médias e altas), produzir uma grande variedade de sons ruidosos, alguns irritantes, outros bem prazerosos!

9.9 Filtros

Como dissemos no começo da seção 9.5, a Transformada de Fourier, e a FFT em particular, tem uma série de aplicações espalhadas por todos os campos do conhecimento humano. Não temos condição de expandir muito o assunto para além de algumas aplicações bastante elementares. Uma delas é a filtragem de ruídos de sinais de qualquer espécie. Vamos usar a Figura 9.17 como um exemplo. Ela mostra a cotação diária do dólar (USD) em reais (BRL) desde o surgimento da atual moeda brasileira até uma data recente (05/02/2018), com inúmeras valorizações e desvalorizaçõeos causadas pelas flutuações do mercado, planos econômicos, crises e *booms*, ações do Banco Central, etc. Seria pedir muito encontrar padrões periódicos num conjunto tão limitado de dados (pouco mais de vinte anos), e não é essa a nossa intenção aqui. O intuito dessa seção é usar a FFT para *suavizar* sinais contaminados por ruído. Obviamente, a oscilação diária de alta frequência que vemos na Figura 9.17 não seria chamada de ruído por um economista. Do nosso ponto de vista, por outro lado, podemos considerar que a flutuação cambial tem um elemento ruidoso que causa bastante variação de um dia a outro. E se conseguíssemos extrair essa variação da curva, tornando-a mais suave? O objetivo de tal procedimento, para um economista, não teria muita utilidade, mas serve aqui apenas como um exemplo do que podemos fazer usando a FFT com um sinal contaminado por ruído.

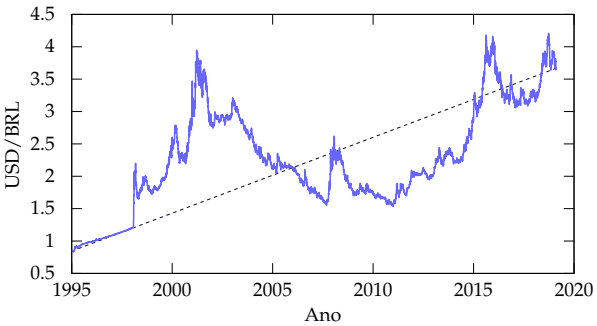


FIGURA 9.17: Cotação do dólar (USD) em reais (BRL) desde o seu surgimento em 1995. A linha tracejada é a linha de tendência dos dados

Muito bem. De posse dos dados sobre a variação cambial do dólar para real, encontrados no material de apoio, você pode calcular a FFT dos dados, diretamente, se quiser, ou aplicar uma janela, com a média subtraída dos dados, etc. O que vamos fazer aqui é diferente. Vamos extrair dos dados uma linha que chamaremos *tendência*, que nada mais é que um segmento de reta ligando o primeiro ao último ponto do conjunto de dados, como mostra a Figura 9.17. Identificando essa linha por $\tau(t)$ e o sinal original (a cotação) por $c(t)$, o sinal que trataremos, $s(t)$, será dado por

$$s(t) = c(t) - \tau(t) \tag{9.9.1}$$

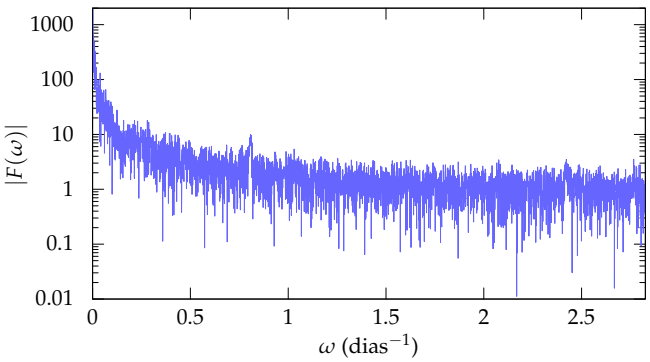


FIGURA 9.18: Transformada Discreta de Fourier dos dados de cotação do dólar mostrados na Figura 9.17

A FFT do sinal $s(t)$ está mostrada na Figura 9.18. Como esperado pela grande flutuação cambial, a FFT é composta de muitas frequências. A frequencia $\omega = 0$ é a de maior intensidade porque, assim como no atrator de Lorenz, a média dos dados não é zero. No entanto, remover a média dos valores não teria aqui o efeito esperado, pois, diferentemente do caso do atrator de Lorenz, aqui, realmente, não há frequências predominantes. Frequências baixas sempre vão dominar o espectro. Por esse motivo, não extraímos a média, mas a linha de tendência.

Agora vamos aplicar um *filtro passa-baixas* (*low-pass filter*) à FFT que acabamos de calcular. Esse filtro deixa passar todas as frequências menores que um *limiar* (*threshold*) pré-estabelecido e atenua ou remove frequências altas. No nosso caso,

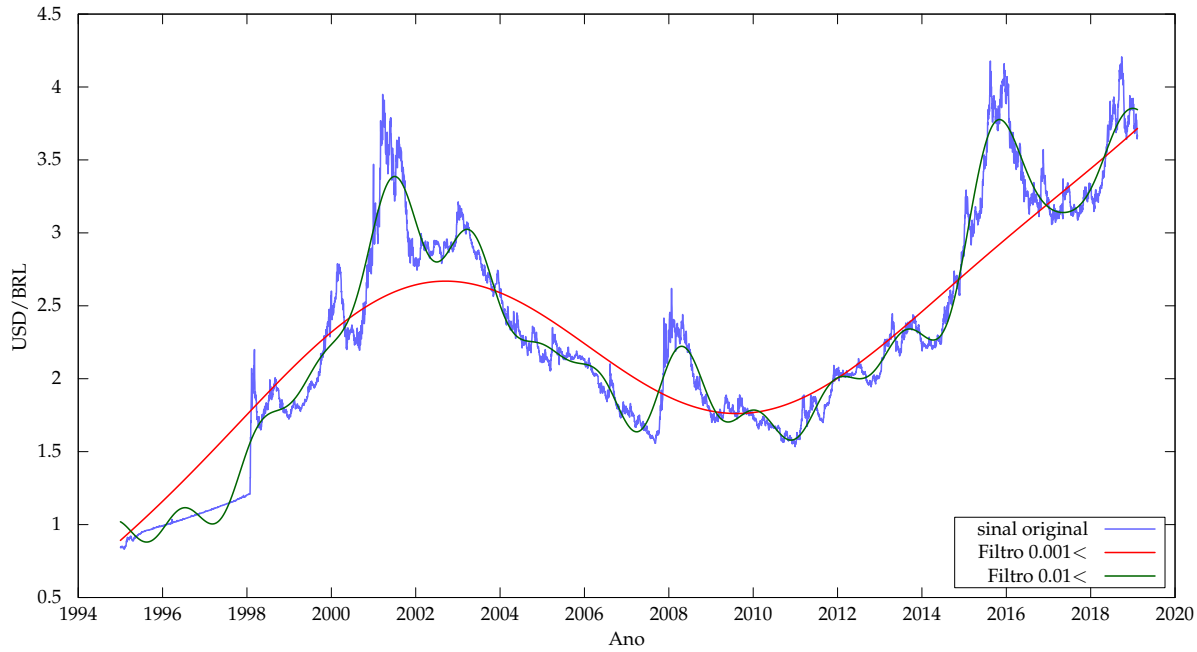


FIGURA 9.19: Sinais suavizados obtidos pela FFT inversa após a passagem da FFT do sinal original por filtros passa-baixas com diferentes limiares.

aplicaremos dois filtros. O filtro passa-baixas 0.01< remove todas as frequências maiores que um limiar de 0.01. O filtro 0.001< faz coisa parecida para frequências maiores que 0.001 unidades. Existem outros filtros que, ao invés de remover de uma vez, apenas atenuam o sinal. Nossos filtros, por sua vez, simplesmente removem o sinal de altas frequências (ou seja, substituem por zero a amplitude calculada pela FFT).

Aproveitemos para dizer que um *filtro passa-altas* (*high-pass filter*) faz o contrário de um filtro passa-baixas: ele remove ou atenua todas as frequências menores do que um certo limiar e deixa passar apenas as frequências altas.

E o que fazer com a FFT filtrada? Podemos usar a Transformada Inversa para obter um sinal filtrado! O filtro passa-baixas remove as frequências altas, responsáveis pelas oscilações de curta período. Isso terá o efeito de suavizar o sinal, como você pode ver na Figura 9.19, onde o resultado dos nossos dois filtros estão mostrados junto ao sinal original $c(t)$. Repare que, como transformamos o sinal $s(t)$, dada pela Eq. (9.9.1), e não o sinal original, $c(t)$, a transformada inversa de $s(t)$ fornece uma versão suavizada de $s(t)$, e não de $c(t)$. Mas basta usar novamente a Eq. (9.9.1) para calcular a versão suavizada de $c(t)$.

Repare na Figura 9.18 que o filtro 0.001< remove quase todas as frequências do espectro. Apenas sobrevivem as frequências baixíssimas, que têm alta intensidade. Correspondentemente, na Figura 9.19 a curva suavizada por tal filtro é bastante suave, e apenas a tendência de longa duração é preservada. Já o filtro 0.01< deixa passar um pouco mais de frequências, e o resultado é uma curva mais sinuosa, que preserva algum comportamento de curta duração observado no sinal original.

Aos interessados no tratamento de dados estatísticos usando a FFT há uma infinidade de referências. A área é bastante vasta e um bom ponto de partida é Press et al. [4]. Uma aplicação totalmente diferente, que também faz amplo uso de filtros, é o tratamento digital de imagens, como veremos na seção 9.11.

9.10 Convolução

Você talvez não se lembre, mas possivelmente já escutou falar em convolução na disciplina de Física Matemática. Pois bem. Quando falamos de filtros, estávamos fazendo uso da convolução, embora seja muito provável que você não tenha percebido. É como quando você era criança e tomava com gosto um guaraná que lhe davam, sem saber que algum remédio estava ali diluído.

Para ver porque e onde usamos a convolução, vamos começar definindo o termo: a convolução de duas funções $f(t)$ e $g(t)$, da variável t no domínio temporal, é indicada por $f \otimes g$ e dada por

$$f \otimes g = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(\tau) g(t-\tau)\, d\tau \tag{9.10.1}$$

onde τ é só a variável de integração (*dummy variable*) e, consequentemente, $f \otimes g$ será uma função de t, assim como as funções originais.

 © Luiz T. F. Eleno, v. 0.3, 2021

Qual é a Transformada de Fourier da convolução de duas funções? Vamos calculá-la usando a definição dada pela Eq. (9.4.6):

$$\mathcal{F}[f \otimes g] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (f \otimes g) e^{-i\omega t}\, dt \tag{9.10.2}$$

Vamos então substituir a Eq. (9.10.1) e ver o que acontece:

$$\mathcal{F}[f \otimes g] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left[\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(\tau) g(t-\tau)\, d\tau \right] e^{-i\omega t}\, dt \tag{9.10.3}$$

Mude agora a ordem das integrais e coloque $f(\tau)$ em evidência:

$$\mathcal{F}[f \otimes g] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(\tau) \left[\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(t-\tau) e^{-i\omega t}\, dt \right] d\tau \tag{9.10.4}$$

Veja agora que podemos simplificar o termo entre colchetes fazendo a substituição $u = t - \tau$, o que fornece

$$\begin{aligned} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(t-\tau) e^{-i\omega t}\, dt &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(u) e^{-i\omega(u+\tau)}\, du = \\ &= e^{-i\omega\tau} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(u) e^{-i\omega u}\, du = e^{-i\omega\tau}\, G(\omega) \end{aligned} \tag{9.10.5}$$

onde $G(\omega)$ é a transformada de Fourier da função $g(u)$ — aqui, u é só uma variável de integração e pode ser trocada de volta por t sem problemas. Agora, substituindo o termo entre colchetes na Eq. (9.10.4) pelo que acabamos de encontrar, teremos

$$\mathcal{F}[f \otimes g] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(\tau) e^{-i\omega\tau} G(\omega)\, d\tau = \left[\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(\tau) e^{-i\omega\tau}\, d\tau \right] G(\omega) \tag{9.10.6}$$

E agora identificamos o termo entre colchetes como $F(\omega)$, a transformada da função $f(t)$, o que nos leva ao *Teorema da convolução de Fourier*:

$$\mathcal{F}[f \otimes g] = \mathcal{F}[f]\mathcal{F}[g] = F(\omega)G(\omega) \tag{9.10.7}$$

Digamos então que temos um sinal $s(t)$ e sua transformada $S(\omega)$. Ao aplicar um filtro $H(\omega)$, estamos multiplicando $S(\omega)$ por uma outra função de ω, $H(\omega)$, e obtemos assim o produto

$$C(\omega) = S(\omega)H(\omega) \tag{9.10.8}$$

Mas sabemos do teorema da convolução que $C(\omega)$ é a transformada da convolução de $s(t)$ com alguma função $h(t)$, ou seja,

$$\mathcal{F}[s \otimes h] = S(\omega)H(\omega) \tag{9.10.9}$$

A transformada inversa então fornece a convolução:

$$\mathcal{F}^{-1}[S(\omega)H(\omega)] = s \otimes h \tag{9.10.10}$$

Assim, o resultado de um filtro é a convolução do sinal original com alguma função $h(t)$ relacionada ao filtro aplicado. Mas qual é a função $h(t)$? Se você precisar dela (e não conhecer $H(\omega)$), basta isolar $H(\omega)$ a partir da Eq. (9.10.10),

$$H(\omega) = \frac{\mathcal{F}[s \otimes h]}{S(\omega)} \tag{9.10.11}$$

e calcular a transformada inversa:

$$h(t) = \mathcal{F}^{-1}[H(\omega)] = \mathcal{F}^{-1}\left[\frac{\mathcal{F}[s \otimes h]}{S(\omega)} \right] \tag{9.10.12}$$

Quando aplicamos um filtro a um sinal, obtemos um novo sinal que é a convolução do original com uma função $h(t)$ dada pela Eq. (9.10.12). A função $H(t)$, que representa um filtro no espaço ω, recebe o nome genérico de *função de transferência no domínio das frequências*. A função $h(t)$ é a função de transferência no domínio temporal e, portanto, representa o filtro no mesmo espaço que o do sinal.

9.11 FFT em duas dimensões: tratamento de imagens

Como último exemplo, vamos investigar a aplicação da FFT ao tratamento de imagens. Muitas técnicas diferentes são usadas para essa finalidade, não necessariamente usando a FFT. Uma visão abrangente do assunto se encontra em Gonzales, Woods e Eddins [5], cujo quarto capítulo serviu de inspiração para a presente seção. Aqui, vamos nos limitar a imagens em branco e preto, também ditas em tons de cinza. Existem tratamentos parecidos para imagens coloridas: os interessados podem consultar o livro apenas citado.

Antes de mais nada, precisamos estender o conceito de Transformada de Fourier para duas ou mais dimensões. Isso é feito facilmente. Por exemplo, a transformada de Fourier de uma função $f(x, y)$, de duas variáveis x e y, é a função $F(u, v)$, de duas outras variáveis u e v, dada por

$$F(u,v) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y) e^{-i(ux+vy)}\, dx\, dy \tag{9.11.1}$$

O fator pré-integral é arbitrário, assim como na transformada unidimensional, ainda que haja um certa lógica que talvez você consiga enxergar. Aqui, x e y são geralmente coordenadas espaciais e u e v, portanto, tem unidades de número de onda.

É igualmente fácil extender a Transformada de Fourier para D dimensões:

$$F(\vec{\omega}) = \left(\frac{1}{\sqrt{2\pi}}\right)^D \int_{\vec{r}} f(\vec{r}) e^{-i\vec{\omega}\cdot\vec{r}}\, d\vec{r} \tag{9.11.2}$$

onde $\vec{r} = (x_1, x_2, \ldots, x_D)$ e $\vec{\omega} = (\omega_1, \omega_2, \ldots, \omega_D)$ são as variáveis no espaço real e no espaço da transformada, respectivamente, e a integral é feita em todo o espaço D-dimensional. A Eq. (9.11.1) é, naturalmente, um caso particular da Eq. (9.11.2).

Usando um procedimento parecido com o que fizemos na seção 9.5.1, é possível implementar uma versão D-dimensional da Transformada Discreta de Fourier. Em duas dimensões, inicie discretizando o domínio da função $f(x, y)$ obtendo uma malha de $M \times N$ pontos (x_m, y_n). A Transformada Discreta de Fourier é então:

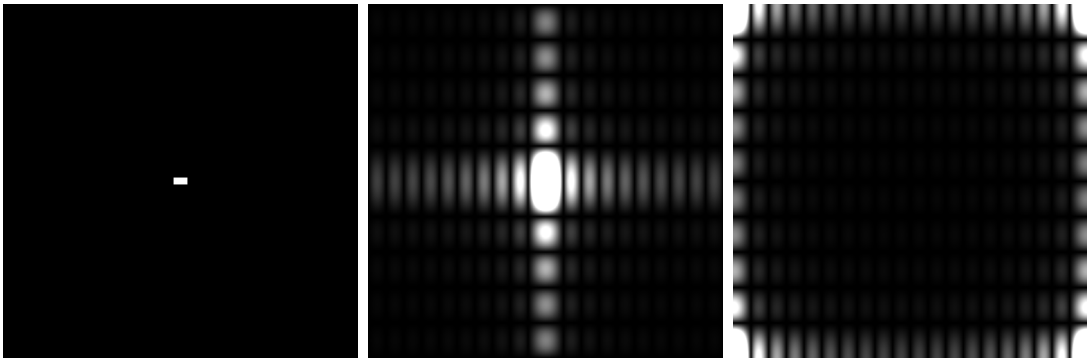
$$F(u_m, v_n) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f(x_m, y_n) e^{-i2\pi(u_m x_m/M + v_n y_n/N)} \tag{9.11.3}$$

que fornece valores numa malha de $M \times N$ pontos (u_m, v_n). Por sua vez, a Transformada Inversa, em sua versão discreta, é calculada usando

$$f(x_m, y_n) = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} F(u_m, v_n) e^{i2\pi(u_m x_m/M + v_n y_n/N)} \tag{9.11.4}$$

O algoritmo FFT para o cálculo da Transformada Discreta de Fourier também é facilmente generalizado para mais dimensões. A biblioteca `scipy.fftpack`, por exemplo, contém funções capazes de realizar a FFT em qualquer dimensão D de maneira bastante eficiente. Vamos usá-las para uma aplicação bastante interessante: o tratamento de imagens.

Vamos começar calculando a FFT de uma imagem bem simples, mostrada na Figura 9.20a. Ela consiste de um pequeno retângulo branco num fundo preto, com o lado maior na horizontal. Matematicamente,



(A) (B) (C)

FIGURA 9.20: A imagem (a) é um pulso retangular em duas dimensões. Sua Transformada de Fourier (b) foi obtida fazendo um deslocamento da FFT original (c).

(A) (B)

FIGURA 9.21: O guaxinim (a), parte da biblioteca `scipy.misc`, é usado como exemplo de imagem. Sua FFT (com deslocamento), em escala logarítmica, resulta no padrão à direita (b).

podemos descrever a imagem como uma função de duas variáveis x (coordenada horizontal da imagem) e y (coordenada vertical) dada por

$$f(x,y) = \begin{cases} 1 & , \quad \text{se } (x,y) \in [-a,a] \times [-b,b] \\ 0 & , \quad \text{caso contrário} \end{cases} \tag{9.11.5}$$

sendo $2a$ e $2b$ a largura e a altura do retângulo branco, respectivamente. A Figura 9.20a é a representação visual dessa função, codificando zero por preto e o valor unitário por branco. Repare que a função da Eq. (9.11.5) é o equivalente bidimensional do pulso retangular dado pela Eq. (9.4.11). A amplitude espectral desse pulso 2D, obtido usando a FFT bidimensional, é mostrada na Figura 9.20b. As coordenadas horizontal e vertical são u e v, respectivamente, e os tons de cinza na imagem codificam os valores mínimo em preto e máximo em branco. Compare a Figura 9.20b com a Figura 9.5: a FFT da função 2D também guarda relação com a função sinc, com um grande pico centrado em $(x,y) = (0,0)$ e um mar de picos e vales, como se fossem ondas se espalhando em todas as direções. Além disso, como o retângulo branco na imagem original é achatado na direção y, a FFT é mais alongada na direção v.

Na verdade, para obter a imagem da FFT mostrada na Figura 9.20b, fizemos um truque. Lembre-se da seção 9.6.1, quando vimos que apenas a primeira metade da FFT contém pontos significantes. Da metade para o final, os valores da FFT são repetições dos valores da primeira metade, tal que $F_{N-n} = F_{-n}$ (no caso unidimensional). O resultado da FFT é, na verdade, a Figura 9.20c. O que fizemos na Figura 9.20b foi um deslocamento (*shift*) da segunda metade da FFT para o começo, em ambas as direções, de forma a obter valores da FFT com a origem no centro da imagem, para que ela coincida com o resultado que seria fornecido pela definição da Transformada de Fourier contínua.

Uma imagem mais interessante é mostrada na Figura 9.21a. Essa foto de um guaxinim faz parte da biblioteca `scipy.misc` e é fornecida como exemplo para os usuários testarem as funções da biblioteca. A imagem, na verdade, é uma matriz de tons de cinza, que fornece um valor entre 0 (preto) e 255 (branco) para cada ponto — ou *pixel* — (x,y) da imagem. A amplitude espectral (com deslocamento, assim como no caso anterior) gerada pela FFT do guaxinim está mostrada na Figura 9.21b. Na verdade, as intensidades foram modificadas por uma escala logarítmica de forma a diminuir as grandes diferenças de intensidade entre diferentes regiões. O pico no centro da imagem é fortíssimo e as intensidades caem muito rapidamente à medida em que você se afasta dele. Sem essa transformação, seria difícil enxergar alguma coisa da FFT além de um pixel branco no centro de um retângulo preto!

9.11.1 Desfoque gaussiano

Usando o teorema da convolução, é possivel criar inúmeros filtros no domínio das frequências que afetarão a imagem original de diversas maneiras. Veremos na presente seção uma maneira de desfocar (*blur*) a imagem usando um filtro gaussiano. A ideia é obter a FFT da imagem, multiplicá-la por uma função gaussiana, e obter a transformada inversa do produto, exatamente como fizemos anteriormente quando descrevemos o uso de filtros.

Uma função gaussiana $G(u,v)$ de duas variáveis u e v é simplesmente o produto de duas gaussianas unidimensionais $G(u)$ e $G(v)$. Lembrando que a função gaussiana de média zero e desvio padrão σ é dada

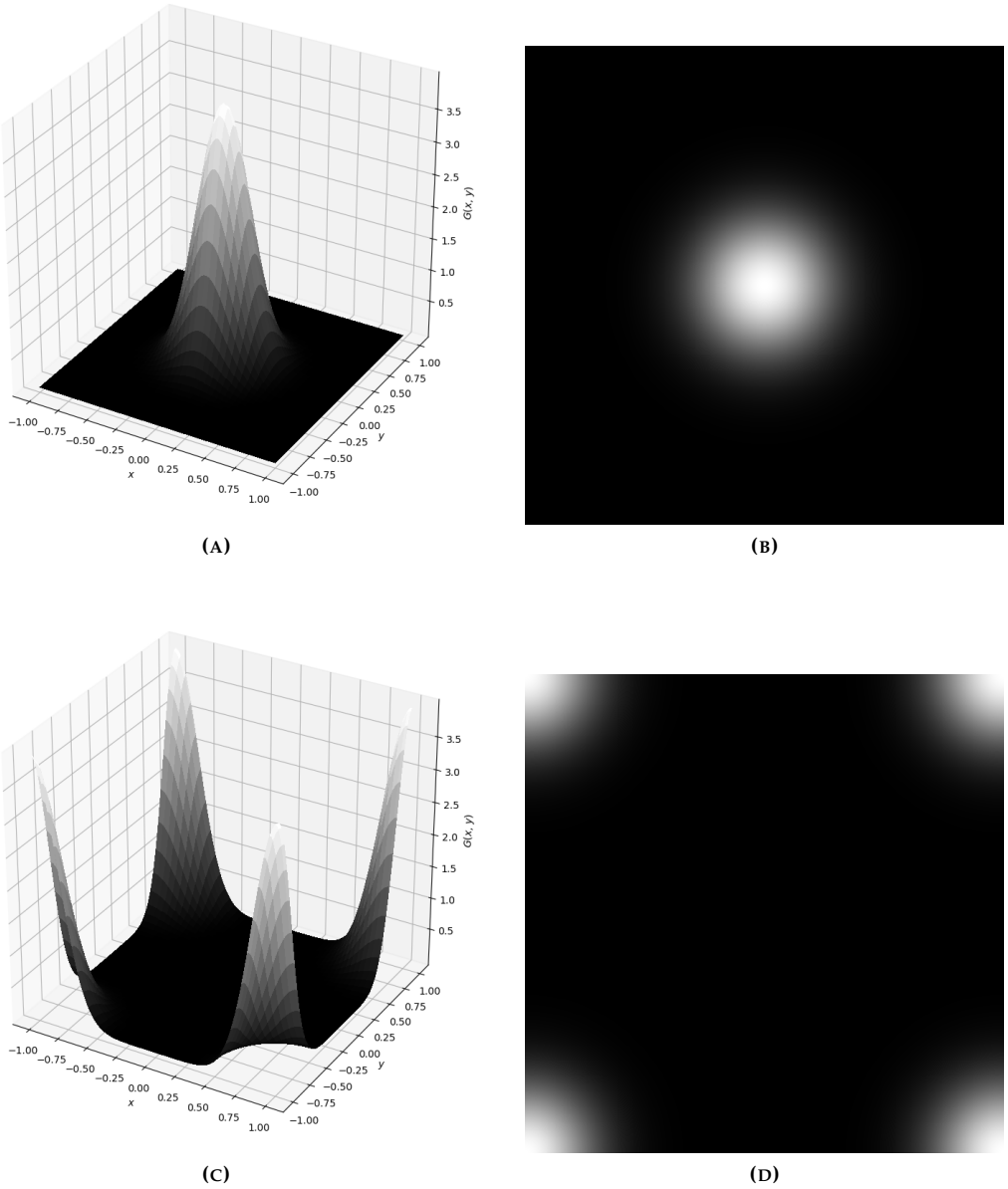


FIGURA 9.22: Gaussiana $G(x, y)$ (a), com desvio padrão $\sigma = 0.2$ e vista superior em (b). Em (c) e (d) estão as versões com deslocamento para uso como filtro de desfoque gaussiano no domínio das frequências.

por

$$G(\omega) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{\omega^2}{2\sigma^2}\right) \tag{9.11.6}$$

Portanto, a gaussiana bidimensional é escrita como

$$G(u, v) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{u^2 + v^2}{2\sigma^2}\right) \tag{9.11.7}$$

Um exemplo, para o caso $\sigma = 0.2$, é mostrado na Figura 9.22a. Por sua vez, a representação da gaussiana como imagem, que equivale à vista superior, é encontrada na Figura 9.22b.

Como discutido anteriormente, por diversos motivos as baixas frequências predominam no espectro da maior parte das imagens. Multiplicar a FFT da imagem original $S(u, v)$ pela gaussiana $G(u, v)$, portanto, terá o efeito de manter a intensidade das baixas frequências sem muita alteração. Por outro lado, frequências mais altas serão atenuadas gradativamente pela gaussiana. Por esse motivo, podemos considerar que esse é um filtro passa-baixas. O efeito de tal filtro será um desfoque da imagem, ou seja, uma perda de nitidez nas transições entre regiões claras e escuras da imagem original. Por quê? Por que transições súbitas precisam de muitas frequências para serem descritas. É a mesma ideia que usamos na seção 9.9 para suavizar uma curva com muito ruído.

(A) (B) (C)

FIGURA 9.23: Aplicação de desfoques gaussianos com diferentes intensidades (b e c), ou seja, diferentes desvios padrão, a uma imagem de alta nitidez (a).

Podemos controlar o grau de desfoque alterando o desvio padrão da gaussiana. Quanto menor o valor de σ, mais concentrado será o pico ilustrado na Figuras 9.22a-b e, consequentemente, maior será o efeito de atenuação da FFT da imagem. Antes de aplicar o filtro, porém, precisamos atentar para o fato de que a FFT gera intensidades deslocadas em relação ao centro. Por esse motivo, o filtro gaussiano precisa ser deslocado antes de ser multiplicado pela FFT da imagem. As Figuras 9.22c-d mostram um exemplo da função $G(x, y)$ deslocada de uma frequência de Nyquist em ambas as direções do domínio das frequências.

Multiplicar a transformada da imagem pela gaussiana (deslocada) gera, pelo teorema da convolução, a transformada da convolução entre a imagem $s(x, y)$ e uma função $g(x, y)$, ou seja:

$$C(u, v) = S(u, v)\, G(u, v) = \mathcal{F}[s \otimes g] \tag{9.11.8}$$

onde $s(x, y)$ é a imagem original e $S(u, v)$ sua a transformada. A imagem convoluida, ou seja, desfocada, será a transformada inversa de $C(u, v)$:

$$s_d(x, y) = s \otimes g = \mathcal{F}^{-1}[C(u, v)] \tag{9.11.9}$$

Uma questão: qual é a função $g(x, y)$? Ora, ela é a transformada inversa de $G(u, v)$. Em Física Matemática, você aprende que a Transformada de Fourier de uma gaussiana unidimensional é outra gaussiana. O mesmo vale para uma gaussiana bidimensional. Você poderia então se perguntar: não teria sido mais fácil multiplicar diretamente a imagem $s(x, y)$ por uma gaussiana $g(x, y)$? Errado! O que você precisa para desfocar é obter a convolução de $s(x, y)$ e $g(x, y)$. Isso é totalmente diferente de multiplicar as duas funções:

$$s(x, y) \cdot g(x, y) \neq s(x, y) \otimes g(x, y) \tag{9.11.10}$$

Calcular uma convolução no espaço (x, y) é extremamente difícil (as integrais envolvidas são extremamente complicadas), e o teorema da convolução é, muitas vezes, nossa única chance de obtê-la. Não o menospreze!

Aplicando um filtro gaussiano à nossa imagem do guaxinim, chegamos às imagens mostradas na Figura 9.23. A imagem original, mostrada na parte (a) da Figura 9.23, é bastante nítida. Aplicando um filtro gaussiano com um alto desvio padrão gera a imagem central: comparando-a com a da esquerda, você perceberá que realmente existe uma perda de nitidez, ou de definição, entre regiões contrastantes da imagem, um resultado que alguns diriam ser mais agradável que o original de alta definição. Por outro lado, uma gaussiana com desvio padrão muito pequeno tem o efeito mostrado na Figura 9.23c, que parece a maneira come eu vejo o guaxinim ao remover (eu, não o guaxinim!) os óculos. Usando desvios ainda menores, você não conseguiria mais identificar o animal na imagem.

Existem muitas outras aplicações da FFT ao tratamento de imagens, que você encontra na literatura especializada [5]. Mas acho que já podemos parar por aqui, já que, como texto introdutório, estamos mais interessados na floresta, e não em cada uma das árvores (ou em toda a fauna, e não só nos guaxinins).

9.12 Comentários finais

Espero que, com os poucos exemplos vistos neste capítulo, você esteja convencido da utilidade da Transformada de Fourier e da FFT para algumas áreas da Física Aplicada e da Computação. Na Medicina moderna, por exemplo, a FFT é fundamental: aliada ao tratamento de imagens, ela é parte essencial da técnica de tomografia computadorizada. Mas os usos e aplicações da FFT não começaram por aqui. Históricamente, a primeira aplicação concreta da FFT foi na solução de equações diferenciais parciais. Esse é o assunto do próximo capítulo, onde a Transformada de Fourier terá novamente um papel relevante.

Referências

[1] P. L. DeVries e J. E. Hasbun. *A first course in Computational Physics*. 2nd. Sudbury (MA): Jones e Bartlett Publishers, 2011.

[2] M. L. Boas. *Mathematical Methods in the Physical Sciences*. 3rd. Hoboken (NJ): Wiley, 2006.

[3] J. W. Cooley e J. W. Tukey. "An algorithm for the machine calculation of complex Fourier series". *Mathematics of Computation* 19 (1965), 297–301.

[4] W. H. Press, S. A. Teukolsky, W. T. Vetterling e B. P. Flannery. *Numerical Recipes 3rd Edition: The Art of Scientific Computing*. 3ª ed. New York, NY, USA: Cambridge University Press, 2007.

[5] R. C. Gonzales, R. E. Woods e S. L. Eddins. *Digital image processing using Matlab*. New Jersey: Prentice Hall, 2004.

 © Luiz T. F. Eleno, v. 0.3, 2021

Capítulo 10

Equações diferenciais parciais

Ciência é uma equação diferencial. Religião é uma condição de contorno.

Alan Turing

10.1 Condução de calor numa barra semi-infinita

10.1.1 Introdução

"Considere uma barra semi-infinita". Tomada ao pé da letra, a frase anterior não faz muito sentido. Uma barra semi-infinita tem só uma extremidade, aquela em que, por comodidade, fixamos a origem de um eixo x. A outra extremidade está tão longe que, para todos os efeitos, nem existe. Qual é a importância prática de um problema aparentemente tão abstrato?

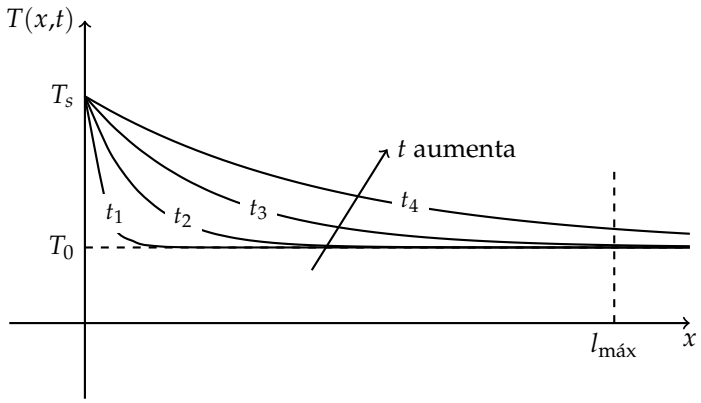


FIGURA 10.1: Perfil de temperaturas ao longo de uma barra semi-infinita.

A importância da aproximação semi-infinita é grande, desde que saibamos como usá-la. Para tentar entender, vamos analisar a Figura 10.1. A nossa barra semi-infinita tem sua extremidade em $x = 0$ e continua indefinidamente para $x > 0$, até $x \to \infty$. Considere que a área da seção transversal é pequena o suficiente para podermos desprezar a transferência de calor em outras direções que não a direção x. Vamos ainda fixar, inicialmente, a temperatura da barra em T_0, completamente homogênea ao longo de todo o seu comprimento (infinito). A partir de um dado instante, que vamos considerar como $t = 0$, a extremidade da barra ($x = 0$) entra em equilíbrio com algum meio externo (localizado em $x < 0$), que está a uma temperatura T_s. Por hipótese, vamos tomar $T_s > T_0$, de forma que na extremidade da barra há um gradiente de temperatura tal que (pela lei de Fourier) o fluxo de calor q'' é positivo, ou seja, calor entra pela extremidade da barra a partir do meio externo. A partir de $t = 0$, a temperatura em $x = 0$ é sempre mantida em T_s. Após um curto intervalo de tempo, o perfil de temperatura ao longo da barra é aquele indicado esquematicamente por t_1 na Figura 10.1. Assim, não houve tempo o suficiente para o fluxo de calor penetrar na barra a não ser numa curta camada próxima à superfície — ou seja, só houve tempo para afetar termicamente uma pequena extensão da barra.

Numa situação prática, digamos que tentamos reproduzir a situação da Figura 10.1 usando uma barra metálica (ou de qualquer outro material) real, de comprimento $l_{\text{máx}} = 1\,\text{m}$. Repare que, agora, a barra tem realmente duas extremidades: uma em $x = 0$, outra em $x = l_{\text{máx}}$. Vamos então submeter a barra às condições descritas no parágrafo anterior, ou seja, vamos aquecer a extremidade da esquerda ($x = 0$) a uma temperatura constante T_s e observar o perfil de temperaturas ao longo da barra em diferentes instantes. Não precisamos fazer nenhuma restrição sobre as condições na outra extremidade, mas, para fixar as ideias,

vamos supor que ela está isolada termicamente. De qualquer modo, após alguns tempo curto, digamos 5 s, apenas uma pequena camada de, digamos, 2 cm foi afetada termicamente. Poderíamos dizer que uma posição foi afetada quando a temperatura atinge naquele ponto um valor 1% maior, por exemplo, que a temperatura inicial T_0. Com isso, podemos considerar a barra de 1 m como semi-infinita no instante $t = 5\,\text{s}$, já que a maior parte do seu comprimento sequer sabe da existência de um gradiente de temperatura. A medida que o tempo passa, uma região cada vez maior da barra é afetada termicamente. Por exemplo, no instante t_2 (ver a Figura 10.1), uma posição mais distante da extremidade chegou a ser afetada termicamente. Em t_3 estamos chegando ao limite de validade da aproximação de barra semi-infinita, se a nossa barra tem comprimento $l_{\text{máx}}$, como indicado na Figura 10.1. Para tempos apenas um pouco maiores, toda a extensão da barra é afetada, e a temperatura da extremidade em $x = l_{\text{máx}}$ começa a aumentar, como no instante t_4 indicado na Figura 10.1. Neste caso, devemos trocar nosso modelo, porque o de barra semi-infinita não serve mais.

10.1.2 Definição do problema em termos de uma equação diferencial

O Quadro 10.1 mostra a tradução do problema em termos da equação diferencial parcial (EDP) da condução de calor em meios materiais usando a condição inicial (CI) e as condições de contorno (CC) descritas na seção 10.1.1. Trata-se de um problema unidimensional, em que a temperatura é descrita em termos das variáveis posição (x) e tempo (t), ou seja, $T = T(x,t)$. Precisamos lembrar que α é a *difusividade térmica* do material da barra. Para os nossos propósitos, vamos considerar α como uma constante, independente da temperatura, do tempo ou da posição.

QUADRO 10.1: Descrição matemática do problema da condução de calor numa barra semi-infinita.

Equação diferencial parcial (EDP):

$$\frac{1}{\alpha}\frac{\partial T}{\partial t} = \frac{\partial^2 T}{\partial t^2} \tag{10.1.1a}$$

Condição inicial (CI):

$$T(x, t = 0) = T_0 \tag{10.1.1b}$$

Condições de contorno (CC):

$$T(x = 0, t) = T_s \tag{10.1.1c}$$

$$T(x \to \infty, t) = T_0 \tag{10.1.1d}$$

O problema é definido apenas para o semi-eixo x positivo.

10.1.3 Transformadas de Laplace

Como não existem condições de contorno bem-definidas em todas as extremidades da barra semi-infinita — ela nem sequer tem uma das extremidades! —, o método de separação das variáveis, válido em outras situações, não funciona aqui. Por isso precisamos lançar mão de um método alternativo que seja mais adequado. O método da *Transformada de Laplace* em relação à variável tempo, cujo domínio de integração é $[0,\infty]$, é exatamente o que precisamos. A definição da transformada de Laplace é a Eq. (10.1.29a), que se encontra no Quadro 10.2 na seção 10.1.6. Usando a definição, uma função $f(t)$ da variável t é Laplace–transformada da seguinte maneira:

$$\mathcal{L}\left[f(t)\right] = \int_0^\infty f(t)\, e^{-st}\, dt \tag{10.1.2}$$

Como t é a variável de integração, eliminamos a variável t quando levamos a cabo a integral. Com isso, a Transformada de Laplace $\mathcal{L}[f(t)]$ da função $f(t)$ será uma nova função, que vamos chamar de $L(s)$, de uma nova variável s:

$$\mathcal{L}\left[f(t)\right] = L(s) \tag{10.1.3}$$

Se temos originalmente uma função de duas ou mais variáveis — como a temperatura $T(x,t)$ no nosso problema —, podemos Laplace-transformar apenas uma das variáveis. No nosso caso, é melhor transformar em relação ao tempo t, de modo que produziremos uma nova função $L(x,s)$ das variáveis x e s, dada por

$$\mathcal{L}_t\left[T(x,t)\right] = L(x,s) = \int_0^\infty T(x,t)\, e^{-st}\, dt \tag{10.1.4}$$

Repare que, na Eq. (10.1.4), indicamos em relação a qual variável Laplace-transformamos a função original usando um sub-índice, $\mathcal{L}_t$.

Para Laplace-transformar derivadas de funções, usamos agora a propriedade dada pela Eq. (10.1.29d). Assim, a transformada da derivada parcial da temperatura em relação ao tempo será escrita em termos da transformada de T e o resultado será:

$$\mathcal{L}_t\left[\frac{\partial T(x,t)}{\partial t}\right] = sL(x,s) - T(x,t=0) \tag{10.1.5}$$

Outra transformada importante para nós é a da derivada segunda da temperatura em relação à posição x. Ela pode ser determinada usando o fato de que x e t são variáveis independentes, e portanto podemos inverter a ordem da derivação parcial (em x) e da integração (em t):

$$\mathcal{L}_t\left[\frac{\partial^2 T(x,t)}{\partial x^2}\right] = \frac{\partial^2}{\partial x^2}\left(\mathcal{L}_t[T(x,t)]\right) \tag{10.1.6}$$

Com isso, a transformada que procuramos é

$$\mathcal{L}_t\left[\frac{\partial^2 T(x,t)}{\partial x^2}\right] = \frac{\partial^2 L(x,s)}{\partial x^2} \tag{10.1.7}$$

e estamos prontos para atacar o problema.

10.1.4 Solução usando Transformadas de Laplace

Encontrando a Transformada

O primeiro passo é achar a transformada da EDP da condução de calor. Usando a Eq. (10.1.1a), portanto, teremos a seguinte igualdade entre transformadas de Laplace:

$$\mathcal{L}_t\left[\frac{1}{\alpha}\frac{\partial T}{\partial t}\right] = \mathcal{L}_t\left[\frac{\partial^2 T}{\partial x^2}\right] \tag{10.1.8}$$

Substituindo agora as Eqs. (10.1.7) e (10.1.5) na Eq. (10.1.8), e usando ainda a propriedade dada pela Eq. (10.1.29c) e a condição inicial, Eq. (10.1.1b), chegamos a

$$\frac{1}{\alpha}\left(sL(x,s) - T_0\right) = \frac{\partial^2 L(x,s)}{\partial x^2} \tag{10.1.9}$$

que pode ser reescrita de forma mais tradicional:

$$\frac{\partial^2 L(x,s)}{\partial x^2} - \frac{s}{\alpha}L(x,s) = -\frac{T_0}{\alpha} \tag{10.1.10}$$

Repare que a derivada que aparece na Eq. (10.1.10) é em relação a x, e não existem derivadas em relação à outra variável, s. Portanto, a Eq. (10.1.10) é uma equação diferencial ordinária (EDO), porque podemos considerar s como constante. Podemos simplificar a notação fazendo

$$\frac{\partial^2 L(x,s)}{\partial x^2} = L'' \tag{10.1.11}$$

e escrevemos a Eq. (10.1.10) na forma

$$L'' - \frac{s}{\alpha}L = -\frac{T_0}{\alpha} \tag{10.1.12}$$

A Eq. (10.1.12) é uma EDO não-homogênea do segundo grau com coeficientes constantes. Além disso, o termo no lado direito é uma constante, de forma que encontramos de imediato uma solução particular:

$$L_{\text{part.}}(x,s) = \frac{T_0}{s} \tag{10.1.13}$$

como pode ser facilmente verificado substituindo-a na Eq. (10.1.12).

Para encontrar a solução geral do problema, precisamos encontrar a solução da EDO homogênea correspondente, ou seja,

$$L'' - \frac{s}{\alpha}L = 0 \tag{10.1.14}$$

 © Luiz T. F. Eleno, v. 0.3, 2021

cuja solução pode ser encontrada em qualquer livro de cálculo, ou usando a estratégia vista no capítulo 2:

$$L_{\text{hom.}}(x,s) = A\exp\left(-\sqrt{\frac{s}{\alpha}}x\right) + B\exp\left(+\sqrt{\frac{s}{\alpha}}x\right) \tag{10.1.15}$$

sendo A e B constantes a serem determinadas e adotamos por hipótese $s > 0$.

A solução geral da Eq. (10.1.10) será simplesmente a soma das equações (10.1.13) e (10.1.15):

$$L(x,t) = L_{\text{hom.}}(x,s) + L_{\text{part.}}(x,s) \tag{10.1.16}$$

ou seja,

$$L(x,s) = A\exp\left(-\sqrt{\frac{s}{\alpha}}x\right) + B\exp\left(+\sqrt{\frac{s}{\alpha}}x\right) + \frac{T_0}{s} \tag{10.1.17}$$

Vamos agora usar uma das condições de contorno, aquela dada pela Eq. (10.1.1d), que nos diz que, para $x \to \infty$, a solução deve ficar limitada. Portanto, precisamos fazer $B = 0$, caso contrário $L(x,t)$ tenderia ao infinito para $x \to \infty$. Isto não pode acontecer, porque a temperatura para $x \to \infty$ tende ao valor T_0, e a transformada não pode divergir. A solução geral então se reduz para

$$L(x,s) = A\exp\left(-\sqrt{\frac{s}{\alpha}}x\right) + \frac{T_0}{s} \tag{10.1.18}$$

Par achar o valor da constante A, precisamos transformar a condição de contorno restante, dada pela Eq. (10.1.1c):

$$\mathcal{L}_t\left[T(x=0,t)\right] = \mathcal{L}_t\left[T_s\right] \tag{10.1.19}$$

Usando as propriedades dadas pelas equações (10.1.29c) e (10.1.29e) na Eq. (10.1.19), descobrimos que

$$L(x=0,s) = \frac{T_s}{s} \tag{10.1.20}$$

Agora fazemos $x = 0$ na Eq. (10.1.18) e, usando ainda a Eq. (10.1.20), determinamos o valor da constante A:

$$A = \frac{T_s - T_0}{s} \tag{10.1.21}$$

Finalmente, juntando todas as informações que conseguimos reunir até aqui, a solução geral da EDP transformada é

$$L(x,s) = \frac{T_s - T_0}{s}\exp\left(-\sqrt{\frac{s}{\alpha}}x\right) + \frac{T_0}{s} \tag{10.1.22}$$

e esta é a transformada de Laplace da temperatura $T(x,t)$ em relação a t.

Encontrando a transformada inversa

Agora precisamos achar a transformada inversa, ou seja, qual temperatura $T(x,t)$ resulta na transformada $L(x,s)$ da Eq. (10.1.22). Antes de começar, vamos escrever a Eq. (10.1.22) de uma forma ligeiramente diferente:

$$L(x,s) = (T_s - T_0)\cdot\frac{e^{-(x/\sqrt{a})\sqrt{s}}}{s} + T_0\cdot\frac{1}{s} \tag{10.1.23}$$

Basta agora lançarmos mão das propriedades da transformada de Laplace do Quadro 10.2 (Seção 10.1.6). Repare que a propriedade dada pela Eq. (10.1.29b) (aditividade da transformada, que é linear) garante que a função $T(x,t)$ será a soma de duas funções, cujas transformadas são as duas funções de s (e de x) no lado direito da Eq. (10.1.23). A propriedade 10.1.29c (homogeneidade da transformação linear) garante ainda que as constantes $(T_s - T_0)$ e T_0, que aparecem multiplicando as parcelas do lado direito da Eq. (10.1.23), serão mantidas sem alteração na expressão para $T(x,t)$. Por fim, com o auxílio das transformadas dadas pelas Eqs. (10.1.29e) e (10.1.29h), a temperatura $T(x,t)$ é dada por

$$T(x,t) = (T_s - T_0)\,\text{erfc}\left(\frac{x}{2\sqrt{\alpha t}}\right) + T_0 \tag{10.1.24}$$

e esta é a solução do problema definido no Quadro 10.1.

Na Eq. (10.1.24), erfc (z) é a *função erro complementar* do argumento z e é definida em termos de erf (z), a

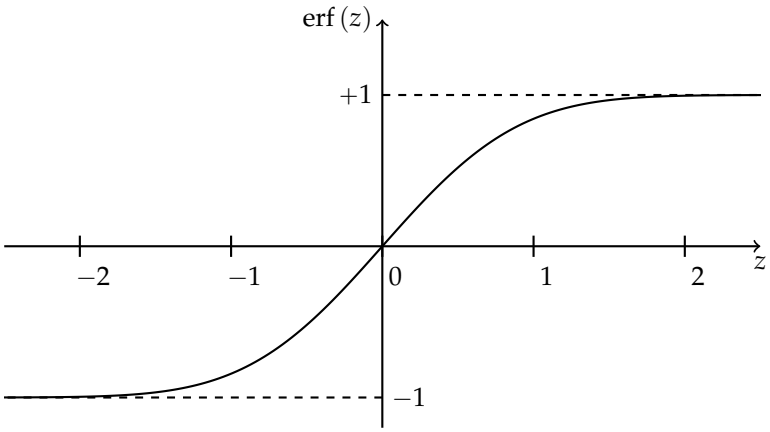


FIGURA 10.2: A função erro, erf (z).

função erro, como

$$\operatorname{erfc}(z) = 1 - \operatorname{erf}(z) \tag{10.1.25}$$

sendo que erf (z) é uma função dada por uma integral da seguinte forma:

$$\operatorname{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-u^2}\, du \tag{10.1.26}$$

Repare que, na Eq. (10.1.26), u é apenas uma variável de integração e não tem significado (em inglês, diz-se que u é uma *dummy variable*). Usando as Eqs. (10.1.25) e 10.1.26) na Eq. (10.1.24) e reagrupando, podemos reescrever $T(x,t)$ em termos da função erro:

$$\frac{T(x,t) - T_s}{T_0 - T_s} = \operatorname{erf}\left(\frac{x}{2\sqrt{\alpha t}}\right) \tag{10.1.27}$$

A Eq. (10.1.27) é a forma mais comum em que a solução do problema é encontrada na literatura.

10.1.5 Limite de validade do modelo de sólido semi-infinito

Para uma discussão da solução, vamos analisar as propriedades da função erro com um pouco mais de detalhes. Um gráfico da função erro se encontra na Figura 10.2. Trata-se de uma função ímpar, ou seja, erf $(-z) = -\operatorname{erf}(z)$, e o comportamento assintótico para $z \to \pm\infty$, quando erf $(z) \to \pm 1$, é claramente observado na Figura 10.2. Quando $z \approx 1{,}83$, o valor da função erro já é erf $(1{,}83) \approx 0{,}99$, ou seja, 99% do valor máximo, que é 1. Se 1,83 é um número difícil de memorizar, um valor mais fácil é $z = 2$. Neste caso, erf $(2) \approx 0{,}995$. Mas, em ambos os casos, podemos aplicar o resultado para estabelecer um limite de validade do modelo de sólido semi-infinito. Vamos supor que nossa barra tenha comprimento $l_{\text{máx}}$. O modelo é válido desde que $z \gtrsim 1{,}83$ (ou $z \gtrsim 2$). Ou seja, a validade do modelo se limita à condição

$$\frac{l_{\text{máx}}}{2\sqrt{\alpha t}} \gtrsim 1{,}83 \tag{10.1.28}$$

Assim, existe um tempo máximo para podermos considerar a barra de comprimento $l_{\text{máx}}$ como semi-infinita. Este tempo corresponde aproximadamente à curva t_3 da Figura 10.1. Para tempos maiores, o problema deve ser tratado por alguma outra aproximação, que leve em conta o comprimento finito do sólido e a existência de uma segunda extremidade em $x = l_{\text{máx}}$.

10.1.6 Apêndice: Propriedades da transformada de Laplace

O Quadro 10.2 traz algumas propriedades úteis da transformada de Laplace. Nas equações a seguir, a é uma constante real, com a restrição de que n é um inteiro positivo na Eq. (10.1.29f) e $a > 0$ apenas na Eq. (10.1.29h).

Referências

[1] J. Nearing. *Mathematical Tools for Physics*. New York: Dover, 2010. URL: http://www.physics.miami.edu/~nearing/mathmethods.

 © Luiz T. F. Eleno, v. 0.3, 2021

QUADRO 10.2: Algumas propriedades da Transformada de Laplace.

Definição:

$$\mathcal{L}\left[f(t)\right] = \int_0^\infty f(t)\, e^{-st}\, dt \tag{10.1.29a}$$

Aditividade:

$$\mathcal{L}\left[f(t) + g(t)\right] = \mathcal{L}\left[f(t)\right] + \mathcal{L}\left[g(t)\right] \tag{10.1.29b}$$

Homogeneidade:

$$\mathcal{L}\left[af(t)\right] = a\,\mathcal{L}\left[f(t)\right] \tag{10.1.29c}$$

Transformada da derivada:

$$\mathcal{L}\left[\frac{df(t)}{dt}\right] = s\,\mathcal{L}\left[f(t)\right] - f(0) \tag{10.1.29d}$$

Transformadas de funções selecionadas:

$$\mathcal{L}\left[1\right] = \frac{1}{s} \qquad (s > 0) \tag{10.1.29e}$$

$$\mathcal{L}\left[t^n\right] = \frac{n!}{s^{n+1}} \qquad (s > 0) \tag{10.1.29f}$$

$$\mathcal{L}\left[e^{at}\right] = \frac{1}{s-a} \qquad (s > a) \tag{10.1.29g}$$

$$\mathcal{L}\left[\operatorname{erfc}\left(\frac{a}{2\sqrt{t}}\right)\right] = \frac{e^{-a\sqrt{s}}}{s} \qquad (s > 0) \tag{10.1.29h}$$

Obs.: Para a definição da função erro complementar, $\operatorname{erfc}(z)$, ver as Eqs. (10.1.25) e (10.1.26).

[2] S. J. Farlow. *Partial Differential Equations for Scientists and Engineers*. New York: Wiley, 1982.

[3] J. H. Lienhard V e J. H. Lienhard IV. *A Heat Transfer Textbook*. New York: Dover, 2011.

Capítulo 11

Cálculo variacional

> Chamei o princípio pelo qual toda pequena variação, desde que útil, é preservada, pelo nome de Seleção Natural.
>
> Charles Darwin

11.1 Introdução

O já antigo mas ainda excelente livro de Courant e Hilbert [1] é a referência básica. Outro texto interessante, mais avançado e voltado especificamente à DFT, é o livro de Engel e Dreizler [2].

11.2 Definição de derivada funcional

$$\Delta x = \frac{x_f - x_i}{n+1} \tag{11.2.1}$$

$$x_k = x_i + k\Delta x\,, \quad y_k = y(x_k), \qquad k = 1, 2, \ldots, n \tag{11.2.2}$$

$$dF = \frac{\partial F}{\partial y_1}dy_1 + \frac{\partial F}{\partial y_2}dy_2 + \frac{\partial F}{\partial y_3}dy_3 + \ldots + \frac{\partial F}{\partial y_n}dy_n \tag{11.2.3}$$

$$dF = \sum_{k=1}^{n} \frac{\partial F}{\partial y_k}dy_k \tag{11.2.4}$$

$$F[y(x)] = \lim_{n\to\infty} F(y_1, y_2, y_3, \ldots, y_n) \tag{11.2.5}$$

$$\delta F = \lim_{n\to\infty} dF = \lim_{n\to\infty} \sum_{k=1}^{n} \frac{\partial F}{\partial y_k}dy_k \tag{11.2.6}$$

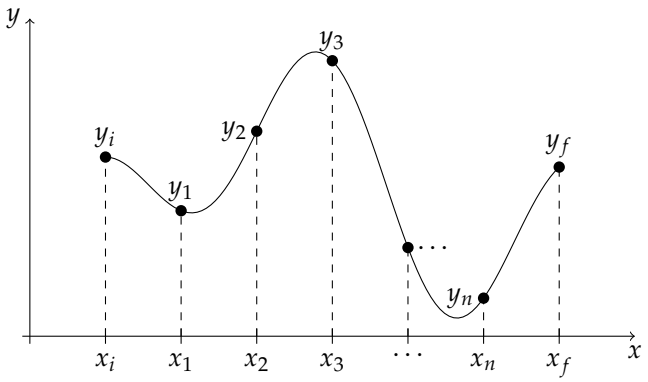


FIGURA 11.1: Discretização da função $y(x)$ em n valores $y_1, y_2, \ldots, y_n$. A função $y(x)$ satisfaz as condições de contorno $y(x_i) = y_i$ e $y(x_f) = y_f$.

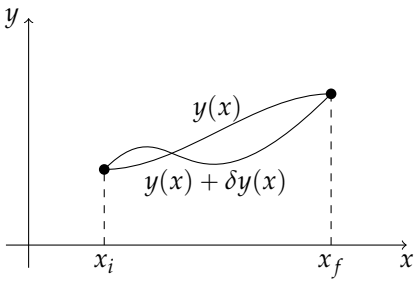


FIGURA 11.2: Ilustração do efeito de uma possível variação $\delta y(x) = \epsilon\eta(x)$ de $y(x)$. Para obedecer às condições de contorno, a função $\eta(x)$ deve se anular nas extremidades x_i e x_f.

$$\delta F = \int_{x_i}^{x_f} \frac{\delta F}{\delta y}\, \delta y\, dx \tag{11.2.7}$$

No caso discreto, um extremo de $F(y_1, y_2, \ldots, y_n)$ é encontrado fazendo $dF = 0$ para qualquer dy_k, o que implica em

$$\frac{\partial F}{\partial y_k} = 0\,, \qquad k = 1, 2, \ldots, n \tag{11.2.8}$$

No limite de $n \to \infty$, o extremo do funcional $F[y]$ é encontrado quando sua derivada funcional se anula:

$$\frac{\delta F}{\delta y} = 0 \tag{11.2.9}$$

Ou seja, qual é a função $y(x)$ que extremiza o funcional $F[y(x)]$? É a função $y(x)$ que anula a derivada funcional na Eq. (11.2.9).

11.3 Epsilons e deltas

$$\delta F[y] = \delta F[y + dy] - \delta F[y] \tag{11.3.1}$$

$$\delta y(x) = \epsilon\, \eta(x) \tag{11.3.2}$$

$$F[y + \delta y] = F[y] + \left.\frac{dF[y + \epsilon\,\eta]}{d\epsilon}\right|_{\epsilon=0} \epsilon + \frac{1}{2!} \left.\frac{d^2F[y + \epsilon\,\eta]}{d\epsilon^2}\right|_{\epsilon=0} \epsilon^2 + \mathcal{O}(\epsilon^3) \tag{11.3.3}$$

$$\delta F[y] = \epsilon \left[\left.\frac{dF[y + \epsilon\,\eta]}{d\epsilon}\right|_{\epsilon=0} + \mathcal{O}(\epsilon) \right] \tag{11.3.4}$$

Por outro lado:

$$\delta F[y] = \int \frac{\delta F[y]}{\delta y(x)}\, \delta y(x)\, dx = \epsilon \int \frac{\delta F[y]}{\delta y(x)}\, \eta(x)\, dx \tag{11.3.5}$$

$$\int \frac{\delta F[y]}{\delta y(x)}\, \eta(x)\, dx = \left.\frac{dF[y + \epsilon\,\eta]}{d\epsilon}\right|_{\epsilon=0} + \mathcal{O}(\epsilon) \tag{11.3.6}$$

$$\boxed{\int \frac{\delta F[y]}{\delta y(x)}\, \eta(x)\, dx = \left.\frac{dF[y + \epsilon\,\eta]}{d\epsilon}\right|_{\epsilon=0}} \tag{11.3.7}$$

Como exemplo, consideraremos o funcional $F[y(x)]$ dado por

$$F[y(x)] = \int [y(x)]^n\, dx \tag{11.3.8}$$

Para calcular a derivada funcional, vamos começar aplicando a Eq. (11.3.7):

$$\int \frac{\delta F[y]}{\delta y(x)}\, \eta(x)\, dx = \frac{d}{d\epsilon} \left[\int [y(x) + \epsilon\eta(x)]^n\, dx \right]_{\epsilon=0} = \int n\, [y(x)]^{n-1}\, \eta(x)\, dx\,, \tag{11.3.9}$$

o que nos leva a concluir que

$$\frac{\delta F[y]}{\delta y(x)} = n\,[y(x)]^{n-1} \tag{11.3.10}$$

11.4 Uma abordagem mais direta para calcular a derivada funcional

$$\delta F = \delta\left(\int y^n\,dx\right) \tag{11.4.1}$$

$$\delta F = \int \delta(y^n)\,dx \tag{11.4.2}$$

$$\delta F = \int n y^{n-1}\delta y\,dx \tag{11.4.3}$$

$$\frac{\delta F}{\delta y} = n y^{n-1} \tag{11.4.4}$$

11.4.1 A menor distância entre dois pontos no plano

Um exemplo interessante é o seguinte funcional:

$$F[y(x)] = \int_a^b \sqrt{1+[y'(x)]^2}\,dx \tag{11.4.5}$$

com a imposição das condições de contorno $y(a) = y_a$ e $y(b) = y_b$. Esse funcional fornece o comprimento da função $y(x)$ entre os pontos (a, y_a) e (b, y_b). Todo mundo sabe que a menor distância entre dois pontos no plano é o segmento de reta entre eles. Além disso, sabemos que pontos de mínimo e máximo são extremos de funções e, portanto, também de funcionais. Logo, se buscarmos pelo extremo do funcional $F[y]$, esperamos encontrar uma linha reta como solução. Vamos comprovar.

$$\delta F = \int_a^b \delta\left(\sqrt{1+y'^2}\right)dx = \int_a^b \frac{y'}{\sqrt{1+y'^2}}\delta y'\,dx \tag{11.4.6}$$

Voltando à notação de epsilons e deltas, fica fácil demonstrar a seguinte igualdade:

$$\delta y' = \delta\left(\frac{dy}{dx}\right) = \frac{d}{dx}(\delta y) \tag{11.4.7}$$

Uma integração por partes da integral mais à direita na Eq. (11.4.6) leva a

$$\delta F = \frac{y'}{\sqrt{1+y'^2}}\delta y\bigg|_a^b - \int_a^b \frac{d}{dx}\left(\frac{y'}{\sqrt{1+y'^2}}\right)\delta y\,dx \tag{11.4.8}$$

Nos limites de integração a e b, $\delta y(a) = \delta y(b) = 0$ pois as extremidades da curva estão fixas. Portanto

$$\delta F = -\int_a^b \frac{d}{dx}\left(\frac{y'}{\sqrt{1+y'^2}}\right)\delta y\,dx \tag{11.4.9}$$

o que nos leva à derivada funcional:

$$\frac{\delta F}{\delta y} = -\frac{d}{dx}\left(\frac{y'}{\sqrt{1+y'^2}}\right) \tag{11.4.10}$$

É fácil verificar que a reta

$$y(x) = c_0 + c_1 x\,, \tag{11.4.11}$$

é a solução geral para o extremo de F, ou seja, a função $y(x)$ que anula a derivada funcional de $F[y]$. As

 © Luiz T. F. Eleno, v. 0.3, 2021

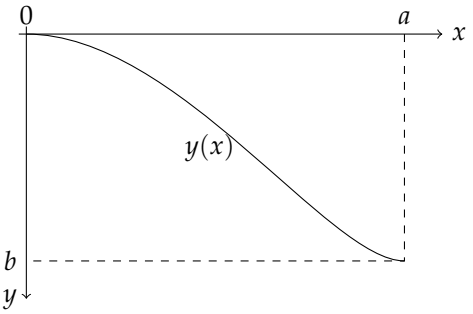


FIGURA 11.3: Sob a ação da gravidade e sem atrito, uma partícula, inicialmente em repouso na origem, desce a rampa descrita pela função $y(x)$ até o ponto (a, b). Qual é a função $y(x)$ para que a partícula chegue ao ponto (a, b) no menor tempo possível?

constantes c_0 e c_1 são encontradas fazendo com que a reta passe pelas pontos (a, y_a) e (b, y_b):

$$y(x) = y_a + \frac{y_b - y_a}{b - a}(x - a) \tag{11.4.12}$$

e o funcional fornece o comprimento do segmento de reta entre as duas extremidades:

$$F[c_0 + c_1 x] = \sqrt{(b-a)^2 + (y_b - y_a)^2} \tag{11.4.13}$$

Da geometria do problema, vemos que o extremo do funcional corresponde a um ponto de mínimo, pois sabemos que a menor distância entre dois pontos no plano é o comprimento do segmento de reta entre eles.

11.4.2 Um clássico problema do Cálculo Variacional

Uma braquistócrona, (do grego "brachistos" mais curto, e "chronos" tempo), ou curva de mais rápida descida, é uma curva entre dois pontos num plano vertical percorrida por uma partícula partindo do repouso do primeiro ponto e chegando no menor intervalo de tempo possível ao segundo, sob a ação da gravidade constante e sem atrito. A Figura 11.3 ilustra a situação.

$$T[y(x)] = \frac{1}{\sqrt{2g}} \int_0^a \sqrt{\frac{1 + y'^2}{y}}\, dx \tag{11.4.14}$$

$$\delta T = \frac{1}{\sqrt{2g}} \int_0^a \delta \left(\sqrt{\frac{1 + y'^2}{y}} \right) dx \tag{11.4.15}$$

$$\delta T = \frac{1}{\sqrt{2g}} \int_0^a \left[\sqrt{1 + y'^2}\, \delta \left(\frac{1}{\sqrt{y}} \right) + \delta \left(\sqrt{1 + y'^2} \right) \frac{1}{\sqrt{y}} \right] dx \tag{11.4.16}$$

$$\delta T = \frac{1}{\sqrt{2g}} \int_0^a \left[\sqrt{1 + y'^2} \left(-\frac{1}{2y\sqrt{y}} \right) \delta y + \frac{y'}{\sqrt{1 + y'^2}} \frac{1}{\sqrt{y}}\, \delta y' \right] dx \tag{11.4.17}$$

$$\int_0^a \frac{y'}{\sqrt{1 + y'^2}} \frac{1}{\sqrt{y}}\, \delta y' = \left[\frac{y'}{\sqrt{1 + y'^2}} \frac{1}{\sqrt{y}} \delta y \right] \Bigg|_0^a - \int_0^a \frac{d}{dx} \left(\frac{y'}{\sqrt{1 + y'^2}} \frac{1}{\sqrt{y}} \right) \delta y\, dx \tag{11.4.18}$$

$$\delta T = -\frac{1}{\sqrt{2g}} \int_0^a \left[\frac{\sqrt{1 + y'^2}}{2y\sqrt{y}} + \frac{d}{dx} \left(\frac{y'}{\sqrt{1 + y'^2}} \frac{1}{\sqrt{y}} \right) \right] \delta y dx \tag{11.4.19}$$

$$\frac{\delta T}{\delta y} = -\frac{1}{\sqrt{2g}} \left[\frac{\sqrt{1 + y'^2}}{2y\sqrt{y}} + \frac{d}{dx} \left(\frac{y'}{\sqrt{1 + y'^2}} \frac{1}{\sqrt{y}} \right) \right] \tag{11.4.20}$$

$$\frac{\sqrt{1+y'^2}}{2y\sqrt{y}} + \frac{d}{dx}\left(\frac{y'}{\sqrt{1+y'^2}}\frac{1}{\sqrt{y}}\right) = 0 \tag{11.4.21}$$

$$y' = p \tag{11.4.22}$$

$$\frac{d}{dx}\left(\,\right) = \frac{d}{dy}\left(\,\right)\frac{dy}{dx} = p\frac{d}{dy}\left(\,\right) \tag{11.4.23}$$

$p = p(y)$

$$\frac{\sqrt{1+p^2}}{2y\sqrt{y}} + p\frac{d}{dy}\left(\frac{p}{\sqrt{1+p^2}}\frac{1}{\sqrt{y}}\right) = 0 \tag{11.4.24}$$

$$\frac{p}{1+p^2}\frac{dp}{dy} + \frac{1}{2y} = 0 \tag{11.4.25}$$

$$1 + p^2 = \frac{c_0}{y} \tag{11.4.26}$$

$$\frac{dy}{dx} = \sqrt{\frac{c_0}{y} - 1} \tag{11.4.27}$$

$$y = 2c_0\,\mathrm{sen}^2\frac{\theta}{2} = c_0(1-\cos\theta) \quad \rightarrow \quad dy = 4c_0\,\mathrm{sen}\,\frac{\theta}{2}\cos\frac{\theta}{2}\,d\theta \tag{11.4.28}$$

$$2c_0\,(1-\cos\theta)\,\frac{d\theta}{dx} = 1 \tag{11.4.29}$$

$$x = 2c_0\,(\theta - \mathrm{sen}\,\theta) + c_1 \tag{11.4.30}$$

Como $y = 0$ para $x = 0$, é preciso que, nesse ponto, $\theta = 0$. Portanto, $c_1 = 0$ e resolvemos o problema em forma parametrizada:

$$\begin{cases} x = r\,(\theta - \mathrm{sen}\,\theta) \\ y = r\,(1-\cos\theta) \end{cases} \tag{11.4.31}$$

sendo $r = 2c_0$. Essa curva é conhecida como cicloide, gerada por um ponto na periferia de um disco de raio r rolando ao longo do eixo x, como ilustrado na Figura 11.4. Para determinar o raio do disco, é preciso impor a condição de contorno do problema original, ou seja, a cicloide precisa passar pelo ponto (a, b). Se $\theta = \theta_f$ quando $(x, y) = (a, b)$, teremos

$$r = \frac{b}{1-\cos\theta_f} \tag{11.4.32}$$

Além disso, para encontrar o valor de θ_f, é preciso que:

$$\begin{cases} a = r(\theta_f - \mathrm{sen}\,\theta_f) \\ b = r(1-\cos\theta_f) \end{cases} \tag{11.4.33}$$

o que gera a equação não-linear

$$\frac{a}{b} = \frac{\theta_f - \mathrm{sen}\,\theta_f}{1-\cos\theta_f} \tag{11.4.34}$$

que só pode ser resolvida numericamente. Apesar disso, podemos calcular o tempo de descida em função de θ_f voltando ao funcional $T[y(x)]$:

$$T = \frac{1}{\sqrt{2g}}\int_0^{\theta_f}\sqrt{\frac{1+(y_\theta/x_\theta)^2}{y}}\,x_\theta\,d\theta \tag{11.4.35}$$

onde fizemos a abreviação

$$x_\theta = \frac{dx}{d\theta}, \qquad y_\theta = \frac{dy}{d\theta}$$

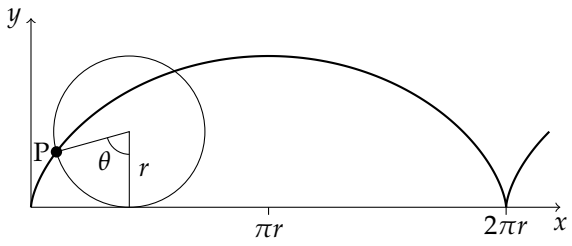


FIGURA 11.4: Uma cicloide gerada pelo ponto P na periferia do disco rolando ao longo do eixo x.

Fazendo o cálculo, descobrimos que

$$T = \theta_f \sqrt{\frac{r}{g}} \tag{11.4.36}$$

Pela situação física do problema, o extremo que encontramos para o funcional, ou seja, a cicloide, deve ser um ponto de mínimo do funcional. Portanto, o tempo dado pela última equação é o menor possível. Qualquer outra função $y(x)$ contínua resultará em um tempo maior que o dado pela Eq. (11.4.36).

Referências

[1] R. Courant e D. Hilbert. *Methods of Mathematical Physics*. Vol. I. Wiley, 1937.

[2] E. Engel e R. M. Dreizler. *Density Functional Theory, an advanced course*. Springer-Verlag, 2011.

Parte IV

Métodos estocásticos

Capítulo 12

Números aleatórios e distribuições

Alea iacta est (o dado está lançado)

Júlio César (100–44 a.c.)

12.1 Variáveis aleatórias

Considere um *espaço amostral* Ω qualquer [1]. O que isso quer dizer? Vejamos um exemplo (não muito criativo): um dado não-viciado com seis faces, numeradas de 1 a 6. O espaço amostral Ω é o conjunto de todos os possíveis resultados do lançamento do dado. Com isso, quero dizer uma das seis possibilidades em que o dado chega ao repouso com uma das faces apoiadas sobre a mesa, mostrando assim, na face oposta (superior) um dos números de 1 a 6 que cada face contém. Cada resultado é o que chamamos de *evento*, e o conjunto de eventos ω forma o espaço amostral Ω. Nesse caso, $\Omega = \{ 1, 2, 3, 4, 5, 6 \}$ e $\omega \in \Omega$ é qualquer um dos valores listados. Nada mais simples. Por *variável aleatória* denotamos uma função $\xi = \xi(\omega)$, que depende de alguma maneira dos eventos do espaço amostral Ω.

Vamos indicar por $P\{ x_1 \leqslant \xi \leqslant x_2 \}$, com $x_1 < x_2$, a probabilidade do evento $\{ x_1 \leqslant \xi \leqslant x_2 \}$, ou seja, a probabilidade de que a variável aleatória ξ apresente um valor no intervalo $[x_1, x_2]$. O conjunto de todos os valores $P\{ x_1 \leqslant \xi \leqslant x_2 \}$ para quaisquer x_1 e x_2 (com $x_1 < x_2$) é a *distribuição de probabilidades* da variável aleatória ξ.

A variável aleatória ξ é dita *discreta* (ou ter uma *distribuição discreta*) se ξ pode assumir apenas valores dentro de um conjunto finito (ou enumerável) de possibilidades x. Nesse caso, podemos indicar as probabilidades de que ξ forneça o valor x por

$$P_\xi(x) = P(\xi = x) \tag{12.1.1}$$

sendo que as probabilidades devem estar normalizadas, ou seja,

$$\sum_x P_\xi(x) = 1 \tag{12.1.2}$$

onde a soma se estende por todos os valores possíveis de x. Para variáveis discretas, vale a relação

$$P\{ x_1 \leqslant \xi \leqslant x_2 \} = \sum_{x=x_1}^{x_2} P_\xi(x) \tag{12.1.3}$$

onde a soma agora é feita para todos os valores de x tais que $x_1 \leqslant x \leqslant x_2$.

Por outro lado, ξ é uma *variável aleatória contínua* (ou tem uma *distribuição contínua*) se ela pode assumir qualquer valor real x, ou seja, $-\infty < x < +\infty$. Nesse caso,

$$P\{ x_1 \leqslant \xi \leqslant x_2 \} = \int_{x_1}^{x_2} p_\xi(x)\, dx \tag{12.1.4}$$

em que $p_\xi(x)$ é a chamada *densidade de probabilidade* da variável aleatória ξ, que também está normalizada no

[1] Uma boa referência introdutória sobre o assunto é Rozanov [1].

intervalo de possíveis valores de x:

$$\int_{-\infty}^{\infty} p_{\xi}(x)\,dx = 1 \tag{12.1.5}$$

Claramente, pela Eq. (12.1.4), temos que, para uma variável contínua,

$$P(\xi = x) = 0 \tag{12.1.6}$$

ao passo que a probabilidade de que ξ esteja no intervalo $[x, x+dx]$ é dada por

$$P\,\{\,\xi \in [x, x+dx]\,\} = p_{\xi}(x) \tag{12.1.7}$$

Ou seja, para variáveis aleatórias contínuas, a probabilidade de que ξ seja exatamente igual a um dado valor x é sempre nula. Podemos falar apenas em uma densidade de probabilidades, $p_{\xi}(x)$, de tal maneira que a integral de $p_{\xi}(x)$ num intervalo $[x_1, x_2]$ nos fornece a probabilidade de que ξ seja encontrada nesse intervalo. Se você já estudou Mecânica Quântica, sabe que $\Psi^*\Psi = |\Psi|^2$, em que Ψ é a função de onda, é uma distribuição contínua de probabilidades. Qual variável aleatória a densidade $|\Psi|^2$ descreve depende do espaço em que ela está escrita: no espaço das posições, a variável aleatória é a posição; no espaço dos momentos — você já entendeu.

12.2 A distribuição uniforme

Vamos supor que um ponto ξ pode ser encontrado com igual probabilidade em qualquer lugar no segmento de reta $[a, b]$ (com $a < b$). Trata-se, portanto, de uma distribuição contínua. A probabilidade de encontrar ξ no intervalo $[x_1, x_2] \in [a, b]$, sendo $x_1 < x_2$, não é difícil de entender, é dada por

$$P\,\{\,x_1 \leqslant \xi \leqslant x_2\,\} = \frac{x_2 - x_1}{b - a} \tag{12.2.1}$$

Veja que podemos escrever

$$P\,\{\,x_1 \leqslant \xi \leqslant x_2\,\} = \int_{x_1}^{x_2} \frac{1}{b-a}\,dx \tag{12.2.2}$$

e que, no caso que $(x_1, x_2) = (a, b)$, a probabilidade está normalizada:

$$P\,\{\,a \leqslant \xi \leqslant b\,\} = \int_{a}^{b} \frac{1}{b-a}\,dx = 1 \tag{12.2.3}$$

Comparando as Eqs. (12.1.4) e (12.2.2), veja que a densidade de probabilidade da distribuição uniforme é escrita na forma

$$p_{\xi}(x) = \begin{cases} 1/(b-a) & \text{, se } a \leqslant x \leqslant b \\ 0 & \text{, se } x < a \text{ ou } x > b \end{cases} \tag{12.2.4}$$

Mais adiante, veremos como fazer o computador gerar números pseudo-aleatórios que, aparentemente, seguem uma distribuição uniforme no intervalo $[0, 1)$. Antes disso, porém, vamos ver como combinar variáveis aleatórias, obtendo outras distribuições de interesse.

12.3 Distribuição de probabilidade conjunta

Considere agora duas variáveis aleatória ξ_1 e ξ_2, que podem ser combinadas como uma nova variável aleatória bidimensional, $\xi = (\xi_1, \xi_2)$. No caso em que as variáveis são discretas, podemos pensar na *distribuição de probabilidade conjunta* de ξ como

$$P_{\xi_1,\xi_2}(x_1, x_2) = P(\xi_1 = x_1, \xi_2 = x_2) \tag{12.3.1}$$

em que x_1 e x_2 podem ser quaisquer eventos do espaço amostral das variáveis ξ_1 e ξ_2, respectivamente. A distribuição de probabilidades para todo um conjunto de valores dentro de um conjunto X é

$$P\{(\xi_1, \xi_2) \in X\} = \sum_{(x_1,x_2)\in X} p_{\xi_1,\xi_2}(x_1, x_2) \tag{12.3.2}$$

Ou seja, $P\{(\xi_1, \xi_2) \in X\}$ fornece a probabilidade de que a variável aleatória $\xi = (\xi_1, \xi_2)$ esteja dentro do conjunto X, e a soma se estende por todos os pares de valores (x_1, x_2) que pertencem a X.

No caso contínuo, podemos falar apenas na distribuição de probabilidades conjuntas dentro de algum conjunto X que, diferentemente do caso discreto, é uma região contínua. A distribuição conjunta será então dada por

$$P\{(\xi_1, \xi_2) \in X\} = \iint_X p_{\xi_1,\xi_2}(x_1, x_2)\, dx_1\, dx_2 \tag{12.3.3}$$

onde a integral dupla deve ser feita no conjunto X em que estamos interessados. A função $p_{\xi_1,\xi_2}(x_1, x_2)$ é a *densidade de probabilidade conjunta*, muito parecida com o seu equivalente unidimensional da seção anterior.

Considere agora que nossas duas variáveis aleatórias ξ_1 e ξ_2 são independentes entre si. Isso significa que a probabilidade de ξ_1 estar no intervalo $x_1' \leqslant x_1 \leqslant x_1''$ não depende da probabilidade de encontrar x_2 no intervalo $x_2' \leqslant x_2 \leqslant x_2''$, e vice-versa. Nesse caso, é esperado que a distribuição conjunta possa ser escrita, no caso discreto, na forma

$$P_{\xi_1,\xi_2}(x_1, x_2) = P_{\xi_1}(x_1) P_{\xi_2}(x_2) \tag{12.3.4}$$

e, no caso, contínuo, como

$$p_{\xi_1,\xi_2}(x_1, x_2) = p_{\xi_1}(x_1) p_{\xi_2}(x_2) \tag{12.3.5}$$

Vejamos dois exemplos clássicos, brevemente explorados por Rozanov [1], para entender como trabalhar com probabilidades conjuntas. O primeiro exemplo é o seguinte: Imagine que dois pontos ξ_1 e ξ_2 podem ser encontrados uniformemente num segmento de reta de comprimento L. A posição do ponto ξ_1 é independente da localização do ponto ξ_2 e vice-versa. Qual é a probabilidade da distância entre os dois pontos ser menor que l?

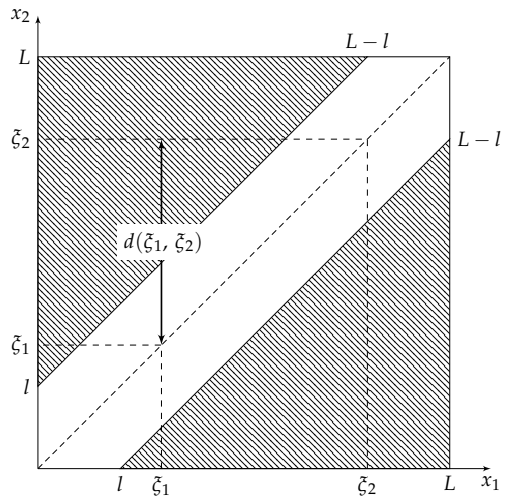


FIGURA 12.1: Os pontos ξ_1 e ξ_2 estão no segmento $[0, L]$. A distância entre eles é $d(\xi_1, \xi_2)$, menor que l apenas na região clara.

A solução do problema fica mais clara usando uma análise gráfica como a mostrada na Figura 12.1. Os pontos ξ_1 e ξ_2, como você pode imaginar, são variáveis aleatórias, que podem assumir quaisquer valores x_1 e x_2 no intervalo $[0, L]$, respectivamente, e estão representados nos eixos x_1 e x_2 da Figura 12.1. A reta $x_1 = x_2$ (a diagonal do quadrado de lado L) nos permite rebater ξ_1 para o eixo x_2 e ξ_2 para o eixo x_1. Com isso, encontramos facilmente a distância entre os pontos, identificada na Figura como $d(\xi_1, \xi_2)$. As regiões hachuradas representam as regiões em que $d(\xi_1, \xi_2) > l$ e são simplesmente a diagonal deslocada de uma distância l para cima ou para baixo. Com isso, a probabilidade de que $d(\xi_1, \xi_2) < l$ é simplesmente a razão entre a área da região central clara e a área de todo o quadrado de lado L. A área da região clara, A_c, é encontrada mais facilmente calculando a área dos triângulos hachurados e subtraindo-a da área do quadrado:

$$A_c = L^2 - 2 \times \frac{1}{2}(L-l)(L-l) = 2Ll - l^2 \tag{12.3.6}$$

Portanto, podemos calcular a probabilidade desejada calculando a relação entre as áreas:

$$P\{\, d(\xi_1, \xi_2) < l \,\} = \frac{A_c}{L^2} = \frac{2Ll - l^2}{L^2} \tag{12.3.7}$$

Veja que podemos ser mais formais e calcular a densidade de probabilidade $p_{\xi_1,\xi_2}(x_1, x_2)$, dada pela Eq. (12.3.5). Como ξ_1 e ξ_2 são distribuídas uniformemente no intervalo $[0, L]$ e são independentes entre si, podemos escrever

$$p_{\xi_1,\xi_2}(x_1, x_2) = p_{\xi_1}(x_1) p_{\xi_2}(x_2) = \frac{1}{L} \times \frac{1}{L} = \frac{1}{L^2} \tag{12.3.8}$$

A probabilidade pedida, então, usando a Eq. (12.3.3), é

$$P\{\, d(\xi_1, \xi_2) < l \,\} = \iint_{A_c} \frac{1}{L^2}\, dx_1\, dx_2 = \frac{1}{L^2} \iint_{A_c} dx_1\, dx_2 = \frac{1}{L^2} A_c \tag{12.3.9}$$

o que fornece o mesmo resultado, portanto, que a Eq. (12.3.7), como esperado.

© Luiz T. F. Eleno, v. 0.3, 2021

O segundo exemplo que veremoso aqui é ainda mais clássico. Ele é conhecido como o oproblema da agulha de Buffon, proposto pela primeira vez no começo do século XVIII por Georges-Louis Leclerc, Comte de Buffon.[2] Trata-se do seguinte: num plano horizontal está traçado um conjunto de retas paralelas, separadas entre si por uma distância d. De uma certa altura, deixa-se cair uma agulha de comprimento l ($l < d$) sobre o plano. Qual é a probabilidade da agulha interceptar alguma das retas, após ela atingir o repouso sobre o plano?

A Figura 12.2a ilustra uma possível situação final, com a agulha AB em repouso no plano da grelha. A distância entre o ponto A e a linha imediatamente acima é y (o ponto A é a extremidade da agulha mais distante da reta considerada). A agulha faz com a direção das linhas um ângulo θ. Repare que só precisamos considerar o espaço entre duas linhas, e basta considerar $0 \leqslant \theta \leqslant \pi/2$. Temos então uma combinação de duas variáveis aleatórias, $\xi_1 = y$ e $\xi_2 = \theta$, que podemos considerar como independentes e uniformemente distribuídas em $[0, d]$ e $[0, \pi/2]$, respectivamente. A densidade de probabilidade conjunta é então, usando novamente a Eq. (12.3.5):

$$p_{\xi_1,\xi_2}(y,\theta) = \frac{1}{d} \times \frac{1}{\pi/2} = \frac{2}{\pi d} \tag{12.3.10}$$

Da Figura 12.2a, a condição para que a agulha intercepte uma das linhas é

$$y \leqslant l \operatorname{sen}\theta \tag{12.3.11}$$

Essa condição corresponde à região hachurada na Figura 12.2b. Então, a probabilidade de interceptação é igual à relação entre a área sob a curva $y = l \operatorname{sen}\theta$ (região hachurada) e a área do retângulo $\{(\theta, y) \,|\, 0 \leqslant \theta \leqslant \pi/2,\ 0 \leqslant y \leqslant d\}$, o que equivale a usar a Eq. (12.3.3):

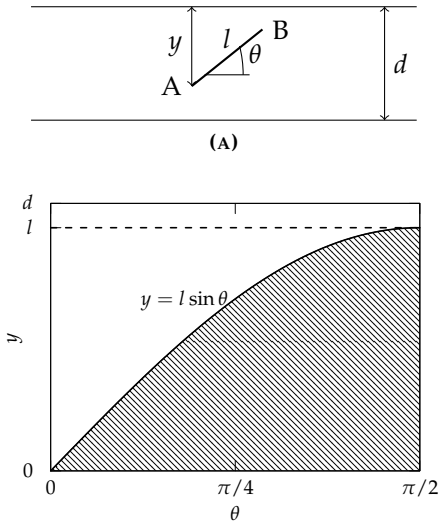


FIGURA 12.2: (a) uma possível posição final da agulha após atingir o repouso. (b) a condição dada pela Eq. (12.3.11) corresponde à região hachurada.

$$P\{\, y \leqslant l \operatorname{sen}\theta \,\} = \int_0^{\pi/2} \int_0^{l \operatorname{sen}\theta} \frac{2}{\pi d}\, dy\, d\theta \tag{12.3.12}$$

que você pode calcular facilmente usando o que aprendeu em alguma disciplina de Cálculo:

$$P\{\, y \leqslant l \operatorname{sen}\theta \,\} = \frac{2}{\pi d} \int_0^{\pi/2} l \operatorname{sen}\theta\, d\theta = \frac{2l}{\pi d} \left[-\cos\theta\right]_0^{\pi/2} = \frac{2l}{\pi d} \tag{12.3.13}$$

É interessante notar que, quando $d = 2l$, a probabilidade será $P\{\, y \leqslant l \operatorname{sen}\theta \,\} = 1/\pi$. Ou seja, se realmente fizéssemos tal experimento, jogando uma agulha inúmeras vezes e calculando a proporção das tentativas em que a agulha cruza alguma linha da grelha, poderíamos determinar estatisticamente o valor de π. Certamente este não é um dos métodos mais eficientes de se fazer isso, e você precisará realmente de muita paciência para isso, mas não deixa de ser um experimento instrutivo... Obviamente, alguém já fez isso e colocou no Youtube! você pode encontrar a discussão neste vídeo: https://www.youtube.com/watch?v=sJVivjuMfWA, no canal Numberphile, que é apoiado pela Univerdade de Nottingham.

12.4 Convolução de distribuições independentes

Vamos considerar agora um problema parecido com os anteriores, mas um pouco mais genérico. Sejam então ξ_1 e ξ_2 duas variáveis aleatórias unidimensionais contínuas e independentes entre si, cujas densidades de probabilidade são $p_{\xi_1}(x_1)$ e $p_{\xi_2}(x_2)$, respectivamente. Considere que tanto ξ_1 quanto ξ_2 sejam variáveis aleatórias de um mesmo espaço, ou seja, os argumentos x_1 e x_2 são eventos ω de um mesmo espaço amostral Ω. uma terceira variável η, definida como

$$\eta = \xi_1 + \xi_2 \tag{12.4.1}$$

(ou qualquer outra função de ξ_1 e ξ_2) é também uma variável aleatória unidimensional, definida no mesmo espaço amostral que ξ_1 e ξ_2. Existe, portanto, uma densidade de probabilidade $p_\eta(t)$, em que t é um evento qualquer do espaço amostral. Mas qual é a relação entre p_η, p_{ξ_1} e p_{ξ_2}?

Para encontrar a expressão procurada, vamos começar tentando calcular a probabilidade de encontrar η no intervalo $[t_1, t_2]$, ou seja, $P\{\, t_1 \leqslant \eta \leqslant t_2 \,\}$. Como ξ_1 e ξ_2 são independentes, podemos escrever imediata-

[2] Georges-Louis Leclerc, Comte de Buffon (17071788), naturalista, matemático, cosmologista e enciclopedista francês.

mente, usando as Eqs. (12.3.3) e (12.3.5),

$$P\{\, t_1 \leqslant \eta \leqslant t_2 \,\} = \iint\limits_{t_1 \leqslant x_1 + x_2 \leqslant t_2} p_{\xi_1}(x_1) p_{\xi_2}(x_2)\, dx_1 dx_2 \tag{12.4.2}$$

em que a integral é calculada para todos os valores de x_1 e x_2 tais que $t_1 \leqslant x_1 + x_2 \leqslant t_2$, para quaisquer valores dados de t_1 e t_2, com $t_1 < t_2$.

A Eq. (12.4.2) pode ser trabalhada em outro formato com a ajuda da Figura 12.3. As variáveis ξ_1 e ξ_2 assumem valores que marcamos nos eixos x_1 e x_2, respectivamente. As retas $x_1 + x_2 = t_1$ e $x_1 + x_2 = t_2$ estão indicadas em linhas cheias e a reta $x_1 + x_2 = t$, com t qualquer valor no intervalo $t_1 \leqslant t \leqslant t_2$, pela linha tracejada. Oo intervalo de integração é exatamente a região entre as retas em linha cheia da Figura 12.3. Veja que nessa região, para qualquer $x_1 = \tau$ ($-\infty \leqslant \tau \leqslant \infty$), teremos $t_1 - \tau \leqslant x_2 \leqslant t_2 - \tau$, e também $t_1 - \tau \leqslant t - \tau \leqslant t_2 - \tau$. Isso nos leva naturalmente à substituição de variáveis

$$x_1 = \tau\,, \qquad x_2 = t - \tau \tag{12.4.3}$$

o que nos leva a um elemento de área

$$dx_1\, dx_2 \quad \to \quad d\tau\, dt \tag{12.4.4}$$

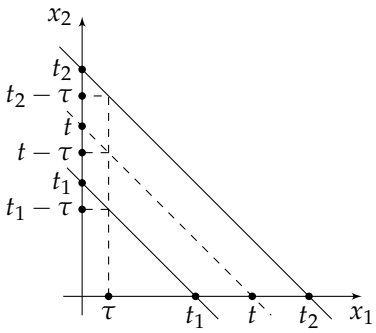


FIGURA 12.3: Para calcular a integral da Eq. (12.4.2), introduzimos as variáveis auxiliares t e τ, cuja interpretação geométrica é mostrada na figura.

já que o jacobiano da transformação é igual a 1. A integral na Eq. (12.4.2), portanto, é reescrita como

$$P\{\, t_1 \leqslant \eta \leqslant t_2 \,\} = \int_{t_1}^{t_2} \int_{-\infty}^{\infty} p_{\xi_1}(\tau) p_{\xi_2}(t - \tau)\, d\tau\, dt \tag{12.4.5}$$

Mas veja que a integral em τ na Eq. (12.4.5), que é dada por

$$p_\eta(t) = \int_{-\infty}^{\infty} p_{\xi_1}(\tau) p_{\xi_2}(t - \tau)\, d\tau \tag{12.4.6}$$

é exatamente a densidade de probabilidade $p_\eta(t)$ procurada! Escrita em função dela, a Eq. (12.4.5) é reescrita como

$$P\{\, t_1 \leqslant \eta \leqslant t_2 \,\} = \int_{t_1}^{t_2} p_\eta(\tau)\, dt \tag{12.4.7}$$

que é estritamente da forma da Eq. (12.1.4).

Mas repare agora na Eq. (12.4.6). Já encontramos uma equação muito parecida com essa na seção 9.10: ela (a menos de um fator $\sqrt{2\pi}$) é a convolução entre as funções p_{ξ_1} e p_{ξ_2}, ou seja,

$$p_\eta(t) = \sqrt{2\pi}\, p_{\xi_1} \otimes p_{\xi_2} \tag{12.4.8}$$

Assim, inesperadamente, a Transformada de Fourier voltam a aparecer (não contavam com a sua astúcia!). Isso porque o teorema da convolução, que vimos na seção 9.10, nos diz que a transformada da convolução é o produto das transformadas, ou seja,

$$\mathcal{F}[p_\eta] = \sqrt{2\pi}\, \mathcal{F}[p_{\xi_1} \otimes p_{\xi_2}] = \sqrt{2\pi}\, \mathcal{F}[p_{\xi_1}] \mathcal{F}[p_{\xi_2}] \tag{12.4.9}$$

e com isso, a transformada inversa fornece a convolução:

$$p_\eta(t) = \sqrt{2\pi} \mathcal{F}^{-1}[\, \mathcal{F}[p_{\xi_1} \otimes p_{\xi_2}]\,] = \int_{-\infty}^{\infty} \mathcal{F}[p_{\xi_1} \otimes p_{\xi_2}]\, e^{i\omega t}\, d\omega \tag{12.4.10}$$

O resultado dado pela Eq. (12.4.6) é completamente geral e vale para a soma de quaisquer duas variáveis aleatórias unidimensionais independentes. Geralmente, calcular diretamente a integral é muito complicado. Da mesma forma, calcular as transformadas e depois a transformada inversa pode ser também um trabalho ingrato. Mas já vimos uma solução no Capítulo 9: a Transformada Discreta. Discretizando os domínios das variáveis ξ_1 e ξ_2, é quase uma tarefa de rotina calcular a convolução entre elas, pelo menos de forma aproximada, usando o algoritmo FFT.

© Luiz T. F. Eleno, v. 0.3, 2021

12.4.1 Um exemplo

Vejamos um exemplo em que calcularemos a densidade de probabilidade dada pela Eq. (12.4.6) de três maneiras diferentes: (1) diretamente, resolvendo a integral; (2) usando o teorema da convolução e obtendo a transformada inversa, de maneira analítica (ou seja, de forma exata); e (3) numericamente, usando a FFT. O exemplo é o seguinte: Considere que ξ_1 e ξ_2 são duas distribuições uniformes independentes, tal que suas densidades de probabilidade são dadas por

$$p_{\xi_1}(t) = \begin{cases} 1/2 & , \text{ se } |t| \leqslant 1 \\ 0 & , \text{ se } |t| > 1 \end{cases} \tag{12.4.11}$$

com uma expressão idêntica para $\xi_2(t)$. Queremos então calcular a densidade de probabilidade da variável aleatória $\eta = \xi_1 + \xi_2$.

Integração direta

Para resolver a integral da Eq. (12.4.6), precisamos determinar os limites de integração em função da variável de integração τ, pois, para o produto se anula em quase todo o eixo τ. Sabemos que $p(\xi_1)(\tau)$ só é diferente de zero no intervalo $-1 \leqslant \tau \leqslant 1$. Já $p(\xi_2)(t-\tau)$ é diferente de zero para

$$-1 \leqslant t - \tau \leqslant 1 \tag{12.4.12}$$

ou seja, para

$$t - 1 \leqslant \tau \leqslant t + 1 \tag{12.4.13}$$

Combinando as duas regiões, teremos que os valores de τ para os quais $p_{\xi_1}(\tau)p_{\xi_1}(t-\tau) \neq 0$ serão encontrados no intervalo

$$\max\{-1, t-1\} \leqslant \tau \leqslant \min\{1, t+1\} \tag{12.4.14}$$

que são os limites de integração. Agora, em relaçao à variável t, veja que a Eq. (12.4.14) só faz sentido para $-2 \leqslant t \leqslant 2$ pois valores de t fora desse intervalo geram limites incoerentes. Além disso, para $-2 \leqslant t \leqslant 0$, os limites de integração são $t - 1 \leqslant \tau \leqslant 1$. Já para $0 \leqslant t \leqslant 2$, os limites são $-1 \leqslant \tau \leqslant t + 1$. Com isso, escrevemos $p_\eta(t)$ como

$$p_\eta(\tau) = \begin{cases} 0 & , \text{ se } t < -2 \\ \displaystyle\int_{t-1}^{1} \left(\frac{1}{2}\right)^2 d\tau = \frac{2+t}{4} & , \text{ se } -2 \leqslant t \leqslant 0 \\ \displaystyle\int_{-1}^{t+1} \left(\frac{1}{2}\right)^2 d\tau = \frac{2-t}{4} & , \text{ se } 0 \leqslant t \leqslant 2 \\ 0 & , \text{ se } t > 2 \end{cases} \tag{12.4.15}$$

O resultado dado pela Eq. (12.4.15) está representado graficamente na Figura 12.4. Concluímos que uma variável aleatória obtida pela soma de duas variáveis aleatórias distribuídas uniformente no mesmo intervalo tem uma densidade de probabilidade triangular. Veja que é preciso ter um enorme cuidado com os limites de integração para calcular a integral de maneira correta. Por outro lado, uma vez determinados tais limites, a integração é quase instantânea, pelo menos para distribuições tão simples quanto as do exemplo. Obviamente, outras distribuições terão suas especificidades que podem simplificar ou complicar o cálculo.

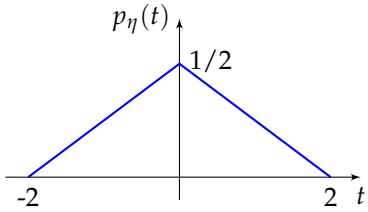


FIGURA 12.4: Densidade de probabilidade da variável aleatória $\eta = \xi_1 + \xi_2$, em que ξ_1 e ξ_2 são distribuídas uniforme e independentemente no intervalo $-1 \leqslant t \leqslant 1$.

Transformada inversa

Agora vamos recalcular tudo de uma outra maneira: determinar, de maneira exata, as Transformadas de Fourier das funções p_{ξ_1} e p_{ξ_2}, multiplicá-las entre si e obter a transformada da convolução. Em seguida, usamos as propriedades da transformada inversa para mostrar que, realmente, a distribuição triangular é a convolução de duas distribuições uniformes idênticas. Pois bem. A Transformada de Fourier da função p_{ξ_1} da Eq. (12.4.11) é calculada usando

$$\mathcal{F}[p_{\xi_1}(t)] = F_{\xi_1}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} p_{\xi_1}(t) e^{-i\omega t}\, dt = \frac{1}{\sqrt{2\pi}} \frac{1}{2} \int_{-1}^{1} e^{-i\omega t}\, dt \tag{12.4.16}$$

Já havíamos calculado, na seção 9.4, a transformada de uma função muito parecida com essa. Ela é facilmente simplificada para

$$F_{\xi_1}(\omega) = \frac{1}{\sqrt{2\pi}} \frac{\operatorname{sen}\omega}{\omega} \tag{12.4.17}$$

Como ξ_1 e ξ_2 têm a mesma densidade de probabilidade, resulta que $F_{\xi_2}(\omega) = F_{\xi_1}(\omega)$. Assim, a transformada da convolução de p_{ξ_1} e p_{ξ_2} será

$$\mathcal{F}[p_{\xi_1} \otimes p_{\xi_2}] = F_{\xi_1}(\omega) F_{\xi_2}(\omega) = \frac{1}{2\pi} \frac{\operatorname{sen}^2\omega}{\omega^2} \tag{12.4.18}$$

Agora, para encontrar a densidade de probabilidade $p_\eta(t)$, usamos a Eq. (12.4.10):

$$p_\eta(t) = \int_{-\infty}^{\infty} \frac{1}{2\pi} \frac{\operatorname{sen}^2\omega}{\omega^2} e^{i\omega t}\, d\omega \tag{12.4.19}$$

A integral na Eq. (12.4.19) é realmente cabeluda (o que, nos meus tempos de criança, significava uma coisa extremamente difícil!), a não ser que você conheça alguns macetes usados na Análise Complexa que você encontra em livros de Física Matemática, como [2]. Vamos deixá-los de lado aqui e adotar outra estratégia: primeiramente, veja que a Eq. (12.4.9) pode ser invertida, ou seja,

$$\mathcal{F}[p_{\xi_1} \otimes p_{\xi_2}] = \frac{1}{\sqrt{2\pi}} \mathcal{F}[p_\eta] \tag{12.4.20}$$

Calcular a transformada da distribuição triangular é fácil: como a Eq. (12.4.15) é diferente de zero apenas no intervalo $-2 \leqslant t \leqslant 2$, a integração se limita a esse intervalo. Então, usando a Eq. (12.4.20) e a definição da Transformada de Fourier, a transformada da distribuição triangular é dada por

$$\mathcal{F}[p_{\xi_1} \otimes p_{\xi_2}] = \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} p_\eta(t) e^{-i\omega t}\, dt = \frac{1}{2\pi} \int_{-2}^{0} \frac{2+t}{4} e^{-i\omega t}\, dt + \frac{1}{2\pi} \int_{0}^{2} \frac{2-t}{4} e^{-i\omega t}\, dt \tag{12.4.21}$$

Com a substituição $t \to -t$ na primeira integral do lado direito, reescrevemos a última expressão como

$$\mathcal{F}[p_{\xi_1} \otimes p_{\xi_2}] = -\frac{1}{2\pi} \int_{2}^{0} \frac{2-t}{4} e^{+i\omega t}\, dt + \frac{1}{2\pi} \int_{0}^{2} \frac{2-t}{4} e^{-i\omega t}\, dt = \frac{1}{2\pi} \int_{0}^{2} \frac{2-t}{4} 2\cos\omega t\, dt \tag{12.4.22}$$

onde usamos algumas propriedades da integração e as relações de Euler do Apêndice B. Essa integral é facilmente calculada por partes:

$$\mathcal{F}[p_{\xi_1} \otimes p_{\xi_2}] = \frac{1}{2\pi} \left.\frac{2-t}{2} \frac{\operatorname{sen}\omega t}{\omega}\right|_0^2 + \frac{1}{2\pi} \frac{1}{2\omega} \int_0^2 \operatorname{sen}\omega t\, dt = \frac{1}{2\pi} \frac{1}{2\omega^2} \left[-\cos\omega t\right]_0^2 = \frac{1}{2\pi} \frac{1-\cos 2\omega}{2\omega^2} \tag{12.4.23}$$

Mas, novamente do Apêndice B, sabemos que $1 - \cos 2\omega = 2\operatorname{sen}^2\omega$. Com isso,

$$\mathcal{F}[p_{\xi_1} \otimes p_{\xi_2}] = \frac{1}{2\pi} \frac{\operatorname{sen}^2\omega}{\omega^2} \tag{12.4.24}$$

Esse resultado é idêntico à Eq. (12.4.18), mostrando que, realmente, a Transformada de Fourier da distribuição triangular é o produto das transformadas de duas distribuições uniformes idênticas. Por outro lado, um(a) leitor(a) mais atento(a) (e mais pedante!) poderia dizer: mas você não mostrou que a Eq. (12.4.19) realmente fornece a distribuição triangular! Realmente, não monstrei, mas digo: espere um momento! Não precisamos calcular a integral cabeluda da Eq. (12.4.19), pois existe um teorema que nos permite evitar osse cálculo! Trata-se do *teorema da integral de Fourier*, que nos diz que a transformada inversa da transformada de uma função $f(t)$ é a própria função original, ou seja:

$$\mathcal{F}^{-1}\left[\,\mathcal{F}[f(t)]\,\right] = f(t) \tag{12.4.25}$$

Esse é um teorema bastante fácil de demonstrar, por isso será deixado como exercício — e você provavelmente já o encontrou em Física Matemática e até mesmo já o usou diversas vezes no presente texto! Por exemplo, para chegar da Eq. (12.4.9) na Eq. (12.4.10), apenas aplicamos, intuitivamente, o teorema dado pela Eq. (12.4.25). Fizemos a mesma coisa várias vezes no Capítulo 9, ao falar, por exemplo, de filtros para a eliminação de ruídos, ao usar um filtro gaussiano para desfocar imagens, etc. Mais uma vez, como eu disse no começo da seção 9.10, você tomou a pílula sem saber!

Transformada Discreta via FFT

Vamos, finalmente, ao terceiro método: calcular tudo numericamente usando a FFT e a FFT inversa! Em comparação aos anteriores, esse método parece brincadeira de criança. Se você lembrar do Capítulo 9, a estratégia para usar a FFT é conseguir bons valores para o período T e o número de pontos N no e com o qual fazer a discretização do sinal original. No presente caso, queremos fazer coisa muito parecida com o que fizemos no caso da transformada contínua, que acabamos de ver: calcular a FFT das densidades p_{ξ_1} e p_{ξ_2}, multiplicá-las entre si e obter a FFT inversa, que será, pelas nossas equações, a densidade triangular p_η.

Para um cálculo rápido, vamos usar $N = 64$ e $T = 6$, com t no intervalo $[-T/2, T/2]$. Com isso conseguimos descrever com folga a parte representativa das distribuições uniformes $p_{\xi_1}(t)$ e $p_{\xi_2}(t)$, dadas pela Eq. (12.4.11), e também da densidade triangular p_η, dada pela Eq. (12.4.15). O resultado é mostrado na Figura 12.5, mas após ter sido feita a normalização dos valores da FFT. Em primeiro lugar, como você se lembra (ou não), a FFT fornece primeiramente valores para frequências no intervalo $0 \leqslant \omega \leqslant \omega_{\text{Nyquist}}$, com $\omega_{\text{Nyquist}} = N\pi/T$, seguidos pelos valores para frequências negativas, ou seja, $-\omega_{\text{Nyquist}} \leqslant \omega \leqslant 0$. Por isso, após calcular a FFT inversa, precisamos fazer um deslocamento (*shift*) de $N/2$ nos valores da transformada inversa. Em segundo lugar, a FFT e a sua inversa não fornecem valores normalizados. Sendo uma densidade de probabilidade, $p_\eta(t)$ precisa estar normalizada (a integral em todo o domínio de t deve ser igual a um). Precisamos então multiplicar todos os valores da FFT por alguma constante para normalizá-la. Tente se convencer, usando as Eqs. (9.4.6) e (9.5.7), que basta multiplicar os valores da FFT inversa por T/N para normalizá-la, sabendo também que basta multiplicar a FFT inversa exata por $\sqrt{2\pi}$, segundo a Eq. (12.4.10), para atingir o mesmo objetivo.

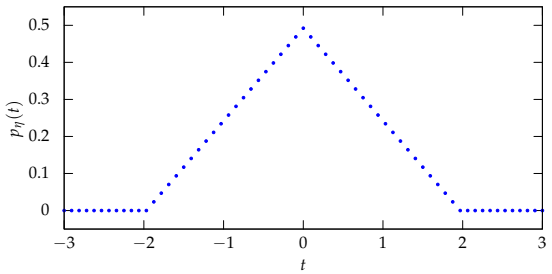


FIGURA 12.5: A convolução de duas variáveis aleatórias uniformes idênticas, via FFT e sua inversa, gera uma densidade de probabilidade triangular.

12.5 Distribuições de probabilidade famosas

Referências

[1] Y. A. Rozanov. *Probability Theory: a concise course*. New York (EUA): Dover, 1969.

[2] M. L. Boas. *Mathematical Methods in the Physical Sciences*. 3rd. Hoboken (NJ): Wiley, 2006.

Capítulo 13

Métodos de Monte Carlo

Apêndices

Apêndice A

Alfabeto grego

É sempre útil ter à mão uma tabela com os nomes das letras do alfabeto grego, onipresentes no linguajar da Física e Matemática — a menos que você já os tenha decorado, e por isso, peço desculpas pela redundância do presente Apêndice.

Àqueles adeptos do LaTeX, é proveitoso também decorar os comandos para gerar os símbolos do alfabeto grego. Eles sempre requerem o uso do modo matemático (*math mode*). Se você não sabe o que é LaTeX, e gostaria de aproveitar para aprender a usar essa poderosa ferramenta e aposentar seu processador de texto convencional, uma boa introdução é uma apresentação, que criei há algum tempo sobre o assunto, encontrada aqui: https://github.com/luizeleno/LaTeX-tutorial-for-newbies.

Maiúscula	Minúscula	Nome	Comandos LaTeX[a]
A	α	alfa	`A`, `\alpha`
B	β	beta	`B`, `\beta`
Γ	γ	gama	`\Gamma`, `\gamma`
Δ	δ	delta	`\Delta`, `\delta`
E	ϵ, ε	épsilon	`E`, `\epsilon`, `\varepsilon`
Z	ζ	zeta	`Z`, `\zeta`
H	η	éta	`H`, `\eta`
Θ	θ, ϑ	teta	`\Theta`, `\theta`, `\vartheta`
I	ι	iota	`I`, `\iota`
K	κ	kapa	`K`, `\kappa`
Λ	λ	lambda	`\Lambda`, `\lambda`
M	μ	mi	`M`, `\mu`
N	ν	ni	`N`, `\nu`
Ξ	ξ	csi	`X`, `\xi`
O	o	ômicron	`O`, `o`
Π	π	pi	`\Pi`, `\pi`
P	ρ	rô	`R`, `\rho`
Σ	σ	sigma	`\Sigma`, `\sigma`
T	τ	tau	`T`, `\tau`
Υ	υ	úpsilon	`Y`, `\upsilon`
Φ	ϕ, φ	fi	`\Phi`, `\phi`
X	χ	chi	`X`, `\chi`
Ψ	ψ	psi	`\Psi`, `\psi`
Ω	ω	ômega	`\Omega`, `\omega`

[a] em *math mode*

Apêndice B

Algumas relações úteis

B.1 Prostaferese

As relações de prostaferese, do grego *prosthesis* (soma) e *aphaeresis* (subtração), são as seguintes relações envolvendo senos e cossenos de dois ângulos a e b:

$$\boxed{\begin{aligned} \operatorname{sen}(a \pm b) &= \operatorname{sen} a \cos b \pm \cos a \operatorname{sen} b \\ \cos(a \pm b) &= \cos a \cos b \mp \operatorname{sen} a \operatorname{sen} b \end{aligned}} \tag{B.1.1}$$
$$\tag{B.1.2}$$

A partir delas, podemos facilmente chegar a inúmeras relações úteis, como os exemplos abaixo:

$$\boxed{\begin{aligned} \operatorname{sen} a \operatorname{sen} b &= \frac{\cos(a-b)}{2} - \frac{\cos(a+b)}{2} \\ \cos a \cos b &= \frac{\cos(a-b)}{2} + \frac{\cos(a+b)}{2} \\ \operatorname{sen} a \cos b &= \frac{\operatorname{sen}(a+b)}{2} + \frac{\operatorname{sen}(a-b)}{2} \end{aligned}} \tag{B.1.3}$$
$$\tag{B.1.4}$$
$$\tag{B.1.5}$$

e

$$\boxed{\begin{aligned} \operatorname{sen} 2a &= 2 \operatorname{sen} a \cos a \\ \cos 2a = \cos^2 a - \operatorname{sen}^2 a &= 2\cos^2 a - 1 = 1 - 2\operatorname{sen}^2 a \end{aligned}} \tag{B.1.6}$$
$$\tag{B.1.7}$$

Dessa última equação, podemos escrever também

$$\boxed{\begin{aligned} \cos^2 a &= \frac{1+\cos 2a}{2} \\ \operatorname{sen}^2 a &= \frac{1-\cos 2a}{2} \end{aligned}} \tag{B.1.8}$$
$$\tag{B.1.9}$$

B.2 Relações de Euler

Podemos usar as relações de Euler entre exponenciais complexas e funções trigonométricas,

$$\boxed{\begin{aligned} e^{ix} &= \cos x + i \operatorname{sen} x \\ e^{-ix} &= \cos x - i \operatorname{sen} x \end{aligned}} \tag{B.2.1}$$
$$\tag{B.2.2}$$

para escrever a relação inversa, ou seja, expressar senos e cossenos em termos de exponenciais complexas, muito úteis na simplificação de algumas expressões:

$$\boxed{\begin{aligned} \cos x &= \frac{e^{ix} + e^{-ix}}{2} \\ \operatorname{sen} x &= \frac{e^{ix} - e^{-ix}}{2i} \end{aligned}} \tag{B.2.3}$$
$$\tag{B.2.4}$$

B.3 Somas

Uma sequência em que cada termo é igual ao anterior multiplicado por uma constante (a *razão* da progressão) é chamada de *série geométrica*. A soma dos n primeiros termos de uma tal progressão é dada por

$$\boxed{\sum_{j=1}^{n} ax^{j-1} = a + ax + ax^2 + \ldots + ax^{n-1} = \frac{a(x^n - 1)}{x - 1}} \tag{B.3.1}$$

Uma variante importante dessa soma, também com n parcelas, é a seguinte

$$\boxed{\sum_{j=0}^{n-1} x^j = 1 + x + x^2 + x^3 + \ldots + x^{n-1} = \frac{(x^n - 1)}{x - 1}} \tag{B.3.2}$$

Nessa última expressão, a soma diverge para $n \to \infty$, a não ser quando $|x| < 1$, quando converge para

$$\boxed{\sum_{j=0}^{\infty} x^j = 1 + x + x^2 + x^3 + \ldots = \frac{1}{1 - x} \qquad (|x| \leqslant 1)} \tag{B.3.3}$$

Vamos usar a Eq. (B.3.1) daqui a pouco. Antes disso, vamos usar as relações de Euler para demonstrar a seguinte identidade:

$$\boxed{\sum_{j=1}^{n} \operatorname{sen} \pi j = 0} \tag{B.3.4}$$

Para isso, primeiramente transforme sen πj em exponenciais complexas:

$$\sum_{j=1}^{n} \operatorname{sen} \pi j = \sum_{j=1}^{n} \frac{e^{i\pi j} - e^{-i\pi j}}{2i} \tag{B.3.5}$$

e separe a soma em duas parciais:

$$\sum_{j=1}^{n} \operatorname{sen} \pi j = \frac{1}{2i} \sum_{j=1}^{n} e^{i\pi j} - \frac{1}{2i} \sum_{j=1}^{n} e^{-i\pi j} \tag{B.3.6}$$

Veja agora que, como j é sempre inteiro, vale a relação $e^{-\pi j} = e^{\pi j} = (-1)^j$. Portanto, as duas somas parciais na Eq. (B.3.6) são idênticas, o que nos leva diretamente à relação procurada, a Eq. (B.3.4). Usando o mesmo procedimento, você pode verificar que

$$\boxed{\sum_{j=1}^{n} \cos \pi j = \begin{cases} 0 & , n \text{ par} \\ -1 & , n \text{ ímpar} \end{cases}} \tag{B.3.7}$$

Uma outra relação útil, e que você pode obter facilmente seguindo a mesma receita usada acima, mais a Eq. (B.3.1), é a seguinte:

$$\boxed{\sum_{j=1}^{n} \cos \frac{2\pi j}{N+1} = -1} \tag{B.3.8}$$

Para demonstrar essa relação, novamente transforme as expressões trigonométricas em exponenciais complexas usando as relações de Euler:

$$\sum_{j=1}^{n} \cos 2\pi k = \sum_{j=1}^{n} \frac{e^{i2\pi k} + e^{-i2\pi k}}{2} \tag{B.3.9}$$

onde fizemos a abreviação $k = j/(N+1)$ para despoluir um pouco as equações. Vamos agora separar a soma

em duas parciais:

$$\sum_{j=1}^{n} \cos 2\pi k = \frac{1}{2}\sum_{j=1}^{n} e^{i2\pi k} + \frac{1}{2}\sum_{j=1}^{n} e^{-i2\pi k} \tag{B.3.10}$$

Veja que cada uma dessas equações é uma progressão geométrica! Podemos então usar a Eq. (B.3.1) para simplificá-las:

$$\sum_{j=1}^{n} \cos 2\pi k = \frac{1}{2}\frac{e^{i2\pi k}(e^{i2\pi kN} - 1)}{e^{i2\pi k} - 1} + \frac{1}{2}\frac{e^{-i2\pi k}(e^{-i2\pi kN} - 1)}{e^{-i2\pi k} - 1} \tag{B.3.11}$$

Vamos rearranjar levemente essa última equação:

$$\sum_{j=1}^{n} \cos 2\pi k = -\frac{1}{2}\left[\frac{e^{i2\pi k(N+1)} - e^{i2\pi k}}{1 - e^{i2\pi k}} + \frac{e^{-i2\pi k(N+1)} - e^{-i2\pi k}}{1 - e^{-i2\pi k}}\right] \tag{B.3.12}$$

Mas veja que $k(N+1) = j$, que é um número inteiro. Por isso, $e^{\pm i2\pi k(N+1)} = e^{\pm i2\pi j} = 1$. Ficamos então com

$$\sum_{j=1}^{n} \cos 2\pi k = -\frac{1}{2}\left[\frac{1 - e^{i2\pi k}}{1 - e^{i2\pi k}} + \frac{1 - e^{-i2\pi k}}{1 - e^{-i2\pi k}}\right] = -1 \tag{B.3.13}$$

como queríamos demonstrar.

Usando a relação acima, mais a Eq. (B.1.9), não é difícil mostrar uma última soma útil e interessante:

$$\boxed{\sum_{j=1}^{n} \operatorname{sen}^2 \frac{\pi j}{N+1} = \frac{N+1}{2}} \tag{B.3.14}$$

© Luiz T. F. Eleno, v. 0.3, 2021

Apêndice C

Algumas linhas (e colunas) sobre matrizes

Estou tentando libertar sua mente, Neo. Mas só posso lhe mostrar a porta. É você quem deve passar por ela.

Morpheus (Matrix)

Temos aqui apenas um esboço dos principais conceitos relacionados a matrizes, utilizados ao longo de vários capítulos. Não se trata de um tratamento abrangente, muito menos rigoroso. O leitor interessado é convidado a consultar a bibliografia sugerida [1–4] ou o livro de Álgebra Linear de sua preferência.

C.1 O que são matrizes?

Uma matriz é, simplesmente uma tabela de números, como você já deve saber. Por exemplo,

$$A = \begin{pmatrix} 1 & 2 & 0 & 8 \\ 3 & 4 & -1 & 1 \\ 5 & -3 & 0 & 1 \end{pmatrix} \tag{C.1.1}$$

é uma matriz com três linhas e quatro colunas. O número de linhas e de colunas de uma matriz é o que chamamos de seu formato, ou de suas dimensões. Por exemplo, a matriz A acima é uma matriz de formato ou dimensões 3×4. É comum omitir as palavras formato ou dimensões, e dizemos simplesmente que A é uma matriz 3×4.

Indicamos qualquer elemento de uma matriz A por A_{ij}, sendo que o primeiro índice (i) representa a linha em que o elemento se encontra e o segundo (j), a coluna. Por exemplo, na matriz A mostrada acima, $A_{23} = -1$ é o elemento na segunda linha e terceira coluna.

C.2 Algumas tipos de matrizes importantes

Se todos os elementos de uma matriz são iguais a zero, ela é chamada de matriz nula. Se uma matriz tem apenas uma linha, ela é chamada de uma matriz (ou um vetor) linha. Se possuir apenas uma coluna, ela é dita uma matriz (ou um vetor) coluna. Já se possuir só uma linha e só uma coluna, a matriz, na verdade, se confunde com um escalar (isto é, um número).

Uma matriz quadrada tem o número de linhas é igual ao de colunas. É comum chamar tal número de ordem da matriz quadrada. Por exemplo,

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 3 & 4 & -1 \\ 5 & -3 & 0 \end{pmatrix} \tag{C.2.1}$$

é uma matriz quadrada de ordem 3. A diagonal principal de uma matriz quadrada corresponde aos termos A_{ii}. No exemplo acima, a diagonal principal é composta pelos elementos $A_{11} = 1$, $A_{22} = 4$ e $A_{33} = 0$. Uma matriz quadrada é simétrica quando $A_{ij} = A_{ji}$ e antissimétrica se $A_{ij} = -A_{ji}$.

Uma matriz diagonal é uma matriz quadrada em que os termos fora da diagonal principal são todos nulos. Em outras palavras, uma matriz A é diagonal quando $A_{ij} = 0$ para $i \neq j$. Já uma matriz identidade é uma matriz diagonal em que os termos da diagonal principal são todos iguais à unidade, ou seja $A_{ii} = 1$. Neste texto, indicaremos matrizes identidade sempre por I.

C.3 Matriz transposta

Se A é uma matriz $m \times n$, a transposta de A, indicada por A^T, é uma matriz $n \times m$ tal que

$$(A^T)_{ij} = A_{ji} \tag{C.3.1}$$

Por exemplo,

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \qquad \Rightarrow \qquad A^T = \begin{pmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{pmatrix} \tag{C.3.2}$$

Repare que a transposição de matrizes é reflexiva, ou seja, a transposta da transposta de A é a própria matriz original A:

$$(A^T)^T = A \tag{C.3.3}$$

Perceba também que, se A é uma matriz quadrada simétrica, vale a relação $A^T = A$. Já se A é antissimétrica, vale $A^T = -A$. Nesse caso, veja também que a diagonal principal de A é nula, pois $A_{ii} = -A_{ii}$ e a única solução possível é $A_{ii} = 0$.

C.4 Soma e produto de matrizes

Para somar duas matrizes A e B e obter a matriz $S = A + B$, fazemos a soma termo a termo, ou seja,

$$S_{ij} = A_{ij} + B_{ij} \tag{C.4.1}$$

Para isso, é necessário que tanto A quanto B tenham o mesmo formato, ou seja, o mesmo número de linhas e de colunas. A matriz resultante (S) terá também as mesmas dimensões de A e B. Repare que vale a propriedade comutativa, ou seja, é sempre verdade que

$$A + B = B + A \tag{C.4.2}$$

desde que A e B tenham o mesmo formato.

Antes de falar de produto de matrizes, vejamos primeiro a operação de multiplicação de uma matriz por um escalar (ou seja, por um número). Esta operação também é simples de definir. Multiplicar a matriz A pelo escalar s resulta na matriz $M = sA$, tal que

$$M_{ij} = sA_{ij} \tag{C.4.3}$$

Ou seja, basta multiplicar todos os elementos de A por s para obter a matriz M.

A multiplicação de matrizes é um pouco mais complicada. O produto de uma matriz A $(m \times p)$ por uma matriz B $(p \times n)$, indicado por AB, é uma matriz P $(m \times n)$ cujos elementos P_{ij} são dados por

$$P_{ij} = (AB)_{ij} = \sum_{k=1}^{p} A_{ik} B_{kj} \tag{C.4.4}$$

A Eq. (C.4.4) noz diz que o elemento P_{ij} de P é obtido pela soma dos produtos, termo a termo, dos elementos da i-ésima linha de A pelos elementos da j-ésima coluna de B. Repare que o número de colunas de A precisa ser igual ao número de linhas de B (p, nos dois casos) para que o produto AB possa ser calculado, de acordo com o esquema

$$P\,(m \times n) = A\,(m \times p) \cdot B\,(p \times n) \tag{C.4.5}$$

onde a seta realça a compatibilidade das matrizes A e B para a multiplicação: o número de colunas de A é igual ao número de linhas de B. O produto $P = AB$ resulta então numa matriz $n \times m$, sendo n o número de linhas de A e m o número de colunas de B. Talvez a seguinte equação ilustre um pouco melhor o processo

de obtenção dos elementos do produto entre as matrizes A e B:

$$\begin{pmatrix} P_{11} & P_{12} & \dots & P_{1j} & \dots & P_{1n} \\ P_{21} & P_{22} & \dots & P_{2j} & \dots & P_{2n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ P_{i1} & P_{i2} & \dots & P_{ij} & \dots & P_{in} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ P_{m1} & P_{m2} & \dots & P_{mj} & \dots & P_{mn} \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1p} \\ A_{21} & A_{22} & \dots & A_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ A_{i1} & A_{i2} & \dots & A_{ip} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m1} & A_{m2} & \dots & A_{mp} \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} & \dots & B_{1j} & \dots & B_{1n} \\ B_{21} & B_{22} & \dots & B_{2j} & \dots & B_{2n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ B_{p1} & B_{p2} & \dots & B_{pj} & \dots & B_{pn} \end{pmatrix}$$

A propriedade comutativa não vale para o produto de duas matrizes, ou seja, em geral

$$AB \neq BA \tag{C.4.6}$$

a não ser em casos particulares (veremos um desses casos ao falar de matriz inversa). Repare que, em alguns casos, o produto BA nem faz sentido! Por exemplo, se A é uma matriz 2×3 e B é 3×4, o produto AB será uma matriz 2×4, mas não existe o produto BA. Em outros casos, AB e BA terão formatos diferentes. Por exemplo, se A é 2×3 e B é 3×2, AB será 2×2 e BA, 3×3. A única situação em que AB e BA sempre existem e têm o mesmo formato é no caso em que A e B são matrizes quadradas de mesma ordem. Mesmo nesse caso, geralmente a Eq. (C.4.6) continua válida, como mostra o exemplo abaixo:

$$AB = \begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 3 & 1 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 3 & 5 \\ 3 & 1 \end{pmatrix} \quad \neq \quad BA = \begin{pmatrix} 3 & 1 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 4 & 6 \\ 2 & 0 \end{pmatrix} \tag{C.4.7}$$

C.5 Matriz inversa

A inversa de uma matriz quadrada A, indicada por A^{-1}, tem a seguinte propriedade:

$$AA^{-1} = A^{-1}A = I \tag{C.5.1}$$

Ou seja, se existe a matriz inversa A^{-1}, a Eq. (C.4.6) deixa de valer e estamos numa situação particular. Vejamos um exemplo:

$$A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \quad \Rightarrow \quad A^{-1} = \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix} \tag{C.5.2}$$

Você pode conferir que

$$\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{C.5.3}$$

Nem toda matriz quadrada possui inversa. Uma matriz quadrada para a qual não existe inversa é chamada de singular. Caso a inversa exista, a matriz é dita invertível ou inversível. Por exemplo, a matriz

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \tag{C.5.4}$$

é singular e consequentemente não possui inversa. Para ver o porquê, é preciso definir alguns outros conceitos que, infelizmente, fogem ao escopo deste apêndice (dica: uma matriz quadrada é singular se, e somente se, seu determinante é nulo).

C.6 Algumas propriedades úteis e importantes

Algumas propriedades de somas, produtos, inversas e transpostas de matrizes são tão úteis que vale a pena elencá-las e demonstrá-las aqui. Utilizamos várias delas em vários capítulos do livro. Vou listar aqui também algumas das propriedades que já mencionei anteriormente, apenas para fins de maior clareza. No que segue, considere que A e B são matrizes para as quais as operações de soma, multiplicação e inversão são possíveis.

P.1 A soma de matrizes é comutativa: $A + B = B + A$.

P.2 Propriedade distributiva da multiplicação em relação à adição: $C(A + B) = CA + CB$.

© Luiz T. F. Eleno, v. 0.3, 2021

Demonstração: faça $P = A + B$, $Q = C(A + B)$, $R = CA$ e $S = CB$. Então

$$Q_{ij} = \sum_k C_{ik}P_{kj} = \sum_k C_{ik}(A_{kj} + B_{kj}) = \sum_k (C_{ik}A_{kj} + C_{ik}B_{kj}) = \sum_k C_{ik}A_{kj} + \sum_k C_{ik}B_{kj} = R_{ij} + S_{ij} \quad \text{(C.6.1)}$$

e portanto $Q + R + S$, o que completa a demonstração. As somas em k são por todas as colunas (linhas) das matrizes do lado esquerdo (direito) dos produtos.

P.3 A multiplicação de matrizes é associativa: $A(BC) = AB(C)$.

Demonstração: faça $P = BC$ e $Q = AB$, R=AP e $S = QC$. Então

$$\begin{aligned} R_{ij} &= \sum_k A_{ik}P_{kj} = \sum_k A_{ik}\sum_l B_{kl}C_{lj} = \sum_k\sum_l A_{ik}B_{kl}C_{lj} = \\ &= \sum_l\sum_k A_{ik}B_{kl}C_{lj} = \sum_l \left(\sum_k A_{ik}B_{kl}\right) C_{lj} = \sum_l Q_{il}C_{lj} = S_{ij} \end{aligned} \quad \text{(C.6.2)}$$

o que nos leva a concluir que $R = S$, completando a demonstração. Repare que, para passar do final da primeira linha da Eq. (C.6.2) para o começo da segunda linha, apenas invertemos a ordem das somas em k e l.

P.4 A transposição é reflexiva: $(A^T)^T = A$.

P.5 A transposta de uma soma é igual à soma das transpostas: $(A + B)^T = A^T + B^T$.

Demonstração: faça $P + A + B$. Então $P_{ij} = A_{ij} + B_{ij}$ e, com isso, $(P^T)_{ij} = A_{ji} + B_{ji} = (A^T)_{ij} + (B^T)_{ij}$, completando a demonstração. Veja que, para obter os índices dos elementos da matriz transposta, basta trocar os índices nos elementos da matriz original.

P.6 A transposta de um produto é o produto das transpostas em ordem contrária: $(AB)^T = B^T A^T$.

Demonstração: Faça o seguinte:

$$(AB)_{ij} = \sum_k A_{ik}B_{kj} \quad \Rightarrow \quad [(AB)^T]_{ij} = \sum_k A_{jk}B_{ki} = \sum_k B_{ki}A_{jk} = \sum_k (B^T)_{ik}(A^T)_{kj} = (B^T A^T)_{ij}\,. \quad \text{(C.6.3)}$$

terminando a demonstração.

P.7 Se A é uma matriz simétrica, vale $A^T = A$.

P.8 Se A é uma matriz antissimétrica, vale $A^T = -A$.

P.9 O produto de uma matriz por sua transposta é uma matriz simétrica: $S = A^T A$ é simétrica.

Demonstração: $S^T = (A^T A)^T = A^T(A^T)^T = A^T A = S$.

P.10 A soma de uma matriz quadrada com sua transposta é uma matriz simétrica: $S = A + A^T$ é simétrica.

Demonstração: $S^T = (A + A^T)^T = A^T + (A^T)^T = A^T + A = A + A^T = S$.

P.11 A diferença entre uma matriz e sua transposta é uma matriz antissimétrica: $\tilde{S} = A - A^T$ é antissimétrica.

Demonstração: $\tilde{S}^T = (A - A^T)^T = A^T - (A^T)^T = A^T - A = -(A - A^T) = -\tilde{S}$.

P.12 Multiplicar uma matriz pela identidade não a altera: $AI = IA = A$.

Demonstração: identificando os elementos da matriz identidade com o delta de Kronecker ($I_{ij} = \delta_{ij}$), teremos

$$(AI)_{ij} = \sum_k A_{ik}\delta_{kj} = A_{ij} \quad e \quad (IA)_{ij} = \sum_k \delta_{ik}A_{kj} = A_{ij}\,. \quad \text{(C.6.4)}$$

O delta de Kronecker, que já encontrarmos em outros capítulos, é definido como

$$\delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} \quad \text{(C.6.5)}$$

P.13 Se A é invertível, existe a inversa A^{-1} tal que $AA^{-1} = A^{-1}A = I$. Caso contrário, A é singular.

P.14 A transposta da inversa é igual à inversa da transposta: $(A^{-1})^T = (A^T)^{-1}$.

Demonstração: uma possibilidade é multiplicar pela identidade "disfarçada":

$$(A^{-1})^T = (A^{-1})^T I = (A^{-1})^T A^T (A^T)^{-1} = (AA^{-1})^T (A^T)^{-1} = I^T (A^T)^{-1} = I(A^T)^{-1} = (A^T)^{-1}. \tag{C.6.6}$$

Obviamente, precisamos supor que A (e A^T) tem inversa.

C.7 Produtos de três matrizes

Estamos agora interessado em encontrar uma expressão para os elementos da matriz $P = ABC$, ou seja, P é o produto de três matrizes, A, B e C, nesta ordem. As matrizes são compatíveis para multiplicação, ou seja, o número de colunas de A é igual ao número de linhas de B, e o o número de colunas de B é igual ao número de linhas de C. Já sabemos, pela propriedade P.3 que podemos escrever

$$P = A(BC) \tag{C.7.1}$$

e, usando a definição de produto entre matrizes dada pela Eq. (C.4.4), os elementos P_{ij} serão dados por

$$P_{ij} = \sum_k A_{ik}(BC)_{kj} \tag{C.7.2}$$

Por sua vez, os elementos $(BC)_{kj}$ podem são calculados por

$$(BC)_{kj} = \sum_l B_{kl} C_{lj} \tag{C.7.3}$$

Juntando agora as duas últimas equações, teremos

$$P_{ij} = \sum_k A_{ik} \sum_l B_{kl} C_{lj} \tag{C.7.4}$$

que também pode ser escrito como

$$P_{ij} = \sum_k \sum_l A_{ik} B_{kl} C_{lj} \tag{C.7.5}$$

que é a equação que buscávamos.

Um outro caso importante em que estamos interessados é quando a matriz C tem apenas uma coluna, ou seja, C é um vetor-coluna. Neste caso $P = ABC$ também será um vetor coluna e, na Eq. (C.7.5), é sempre $j = 1$ e podemos deixá-lo subentendido, escrevendo apenas P_i e C_l. Com isso, podemos escrever

$$P_i = \sum_k \sum_l A_{ik} B_{kl} C_l \tag{C.7.6}$$

para o caso em que P e C são vetores-coluna.

Referências

[1] J. L. Boldrini, S. I. R. Costa, V. L. Figueiredo e H. G. Wetzler. *Álgebra Linear*. 3ª ed. São Paulo: Harbra, 1980.

[2] J. N. Franklin. *Matrix Theory*. Englewood Cliffs (NJ): Prentice-Hall, 1968.

[3] J. M. Anderson. *Mathematics for Quantum Chemistry*. New York: W. A. Benjamin, 1966.

[4] A. J. Pettofrezzo. *Matrices and transformations*. New York: Dover, 1966.

Exercícios

C. 1 — Mostre que, se A é uma matriz simétrica invertível, A^{-1} também é simétrica.

C. 2 — xx

(a) yy

Apêndice D

Diagonalização de matrizes

D.1 Definição

O que entendemos aqui por *diagonalizar* uma matriz quadrada M de ordem n é encontrar as matrizes V e Λ, também quadradas e de mesma ordem que M, que satisfazem a seguinte propriedade:

$$M = V\Lambda V^{-1} \tag{D.1.1}$$

ou, de maneira equivalente.

$$\Lambda = V^{-1}MV \tag{D.1.2}$$

Além disso, a matriz Λ deve ser uma matriz diagonal, ou seja, com todos os elementos fora da diagonal principal nulos.

D.2 Solução algébrica

Da Álgebra linear, sabemos que a os elementos da diagonal principal da matriz Λ são todos os autovalores de M, obtidos resolvendo-se a equação matricial

$$Mv = \lambda v \tag{D.2.1}$$

em que v é um autovetor correspondente ao autovalor λ. O caso mais simples de se lidar é quando temos n autovalores λ distintos. Nessa situação, a matriz V é formada por autovetores v de M, um por coluna na mesma ordem em que a matriz Λ é populada pelos autovalores λ correspondentes a v.

Vejamos um exemplo: a matriz M abaixo

$$M = \begin{pmatrix} 0 & -2 & -2 \\ -2 & -3 & -2 \\ 3 & 6 & 5 \end{pmatrix} \tag{D.2.2}$$

tem três autovalores distintos: $\lambda_\alpha = -1$, $\lambda_\beta = 1$ e $\lambda_\gamma = 2$. Portanto, já encontramos uma matriz Λ possível:

$$\Lambda = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \tag{D.2.3}$$

Basta agora encontrar um autovetor para cada um dos autovalores e formar a matriz V:

$$V = \begin{pmatrix} 0 & -2 & 1 \\ 1 & 1 & 0 \\ -1 & 0 & -1 \end{pmatrix} \tag{D.2.4}$$

cuja inversa é

$$V^{-1} = \begin{pmatrix} 1 & 2 & 1 \\ -1 & -1 & -1 \\ -1 & -2 & -2 \end{pmatrix} \tag{D.2.5}$$

Salientamos que as matrizes Λ e V não são únicas: na primeira, podemos permutar os autovalores nas posições da diagonal principal; na segunda, podemos escolher autovetores diferentes dos que usamos (por exemplo, multiplicar qualquer um deles por uma constante). Não há necessidade de normalizar os autovetores. As propriedades da seção D.1 sempre são satisfeitas.

A primeira dificuldade surge quando M tem menos do que n autovalores distintos. Por exemplo, a matriz abaixo,

$$M = \begin{pmatrix} 0 & -2 & -2 \\ 0 & 1 & 0 \\ 1 & 2 & 3 \end{pmatrix} \tag{D.2.6}$$

tem dois autovalores distintos, $\lambda_\alpha = \lambda_\beta = 1$ (com multiplicidade 2) e $\lambda_\gamma = 2$ (com multiplicidade 1). Quando isso acontece, nem sempre é possível diagonalizar, pois nem sempre é possível encontrar autovetores linearmente independentes correspondendo aos autovalores de multiplicidade maior que 1. Para o exemplo, é possível encontrá-los, o que nos permite escrever as matrizes V e Λ como, por exemplo:

$$V = \begin{pmatrix} -1 & -2 & -2 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}, \qquad \Lambda = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{D.2.7}$$

Por fim, algumas matrizes não são diagonalizáveis. Por exemplo, a matriz

$$M = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \tag{D.2.8}$$

tem um único autovalor $\lambda_\alpha = \lambda_\beta = 1$, com autovetores na forma

$$v = c \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{D.2.9}$$

com c uma constante qualquer. Fica claro que não conseguimos encontrar dois autovetores linearmente independentes e, portanto, é impossível montar uma matriz V que seja inversível e satisfazer as condições da seção D.1.

Para arrematar: matrizes com n autovalores distintos são sempre diagonalizáveis. Por outro lado, matrizes com menos do que n autovalores distintos podem ser diagonalizáveis, mas nem sempre o são.

D.3 Solução numérica em python

A biblioteca `scipy.linalg` tem uma função, `eig`, que retorna os autovalores e autovetores de uma matriz quadrada. Uma outra função, `eigh`, resolve o mesmo problema, mas apenas para matrizes hermitianas, como aquelas que permeiam toda a Mecânica Quântica. Se a sua matriz é hermitiana, use a função `eigh`, pois o cálculo dos autovalores e autovetores será muito mais rápido e eficiente. Encontre mais informações na documentação oficial do `scipy`: https://docs.scipy.org/doc/scipy/reference/generated/scipy.linalg.eig.html e https://docs.scipy.org/doc/numpy/reference/generated/numpy.linalg.eigh.html.

Apêndice E

O que você precisa saber sobre python

Apêndice F

Links para o material de apoio

Nesta seção, você encontra *links* para o material de apoio, citado em vários pontos do texto. A maior parte do material é formado por *scripts* escritos em linguagem python. Todos foram escritos e testados para rodar em python 3.0 ou superior, apesar de, em alguns casos e com algumas pequenas alterações, ser possível rodá-los também em python 2.7. Praticamente todos os *scripts* fazem uso das seguintes bibliotecas:

- numpy;
- scipy;
- matplotlib.

Portanto, é importante instalar uma distribuição python incluindo essas bibliotecas.

Além disso, nas descrições do material encontradas a seguir, suponho que que você tenha uma versão python instalada com as bibliotecas listadas acima, e saiba o que fazer para rodar os *scripts* (mas veja o Apêndice E para algumas dicas). Além disso, alguns códigos fazem uso de outras bibliotecas, descritas caso a caso. Boa diversão!

F.1 Fractais

O arquivo `Samamba.py` cria todos os fractais (e mais alguns) oriundos de transformações afins vistos no capítulo 1. O *docstring* da classe `Samamba` diz o seguinte:

```
Samamba

uma interface genérica para
belas imagens de fractais
gerados por transformações afins

Prof. Luiz T. F. Eleno
v. 2018.1
```

Rodando o arquivo, você terá acesso a uma interface gráfica capaz de gerar e salvar as imagens. Para essa última função, no entanto, é preciso ter instalada a biblioteca `pillow`, uma versão mais amigável da biblioteca `PIL` (Python Imaging Library). Você está convidado a testar os diferentes fractais incluídos nos exemplos, além de inventar os seus próprios fractais personalizados!

F.2 Modos Normais

A classe `Nosciladores_exato.py` contém a implementação da classes `N_osc`, que resolve de forma exata (a menos da precisão numérica) o problema de N osciladores acoplados e presos a paredes rígidas e imóveis nas extremidades. como visto na seção 2.6. Você pode usar a classe para resolver todos os exercícios do capítulo 2.

F.3 O atrator de Lorenz

Você pode usar o *script* `lorenz.py` para gerar o atrator de Lorenz da seção 8.5.1. É necessário instalar a biblioteca `mpl_toolkits.mplot3d` para gerar os gráficos 3D. O resultado pode ser visto de diversos pontos de vista em várias diferentes ampliações. Além disso, você pode alterar os parâmetros s, r e b das equações de Lorenz, assim como as condições iniciais, para testar o efeito de diferentes configurações do problema.

F.4 Espectro de frequências de um sinal de áudio

tone generator http://www.szynalski.com/tone-generator

Apêndice G

Respostas para exercícios selecionados

 © Luiz T. F. Eleno, v. 0.3, 2021

Índice de autores

Índice de assuntos