Brian D. Wood

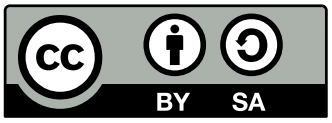
arXiv:2305.05036v1 [math.HO] 8 May 2023

Introduction to Engineering Mathematics and Analysis

用数学语言对物理系统建模

May 10, 2023

ISBN-13 978-1-955101-33-2



本作品采用知识共享署名 4.0 国际许可协议授权。
https://creativecommons.org/licenses/by-sa/4.0/

Text Revision 1.3

引用格式:

Wood, B. D. (2023). Introduction to Engineering Mathematics and Analysis: Modeling Physical Systems Using the Language of Mathematics. First Edition, Rev. 1.3. Oregon State University, Corvallis OR. DOI: https://doi.org/10.5399/osu/1152.

序言

> 指出没有人会以不懂文学为荣，已经几乎成了一句老生常谈；但以不懂科学为荣，甚至自豪地声称自己不擅长数学，却是社会所能接受的.....
>
> ——理查德·道金斯

为什么要学习数学?

科学家和工程师按定义都是问题求解者。但仔细想想，他们是如何成为问题求解者的，这倒是件有趣的事。一个人早期的学术生涯中，或许上过一门课名里带有“问题求解”几个字，但学生通常并不是在这类课程中学会如何真正解决问题的。工程师和科学家如何学到他们最终掌握的知识，是教育研究的一个领域。这个问题过于宽泛，不适合在此讨论；不过，探讨数学在教育过程中的应用，则在本文的讨论范围之内。

“可是，我什么时候才会用到 这个？”这句反复出现的话，对任何学过（或教过）数学的人来说都很熟悉。它凸显了数学教学方式中的一个特殊问题。例如，学习诗歌时，学习者通常不会惊呼：“可是，我什么时候才会用到 这首诗？”思考数学为何被视为与诗歌等学科不同，便能看出问题的端倪。学习者往往期待 数学能在非常实际的意义上派上用场。这种期待可以理解，因为我们早期接受的大多数数学训练（例如学习加减乘法、解决基本的代数问题）都有显而易见的实用价值。通常，直到引入更高层次的数学思维（例如形式化的线性代数、微积分、微分方程）时，人们才开始质疑数学是否有用。

Most of us study mathematics because it is *useful*. Few would argue that knowing how to add, subtract, multiply, and divide are skills that are not useful. Similarly, few would argue, for example, that spreadsheets are not useful; and spreadsheets are primarily a tool based in mathematical operations. Most engineers and scientists have had the opportunity and need to apply the various rules of algebra to solve *actual real-world problems*. So, there are subject areas for which the tools of mathematics are generally agreed to be useful.

One of the problems that occurs in the study of *higher mathematics* is that it is not made clear *in what sense* it is useful. In part, this is because the utility of higher mathematics is not as directly obvious as it is for more basic mathematical reasoning. It is worth making this as concrete as possible, however, to help better understand why the study of higher mathematics can be worth the effort that one has to put in.

学习数学，部分而言，就是学习算法思维。然而除此之外，学习数学还能帮助我们学习并实践算法思维。即使你从未应用过所学的任何数学工

具，算法思维的过程也能广泛应用于问题求解。因此，学习数学可以帮助你成为更出色的问题解决者，即使解决问题时并未用到数学！这为普遍学习高等数学提供了充分的理由。学习数学不仅能教会人们如何真正理解新的数学概念（根据职业道路的不同，这些概念可能会在应用中出现），还能帮助培养新的问题求解技能，其用途远远超出数学方法本身的应用。

因此，即使在修完涵盖这些主题的课程后，你再也不计算导数，也不再求解偏微分方程，学习这些内容本身也会带来益处。各个层次的数学学习都有助于培养并强化你思考问题的能力，以及生成算法方法来解决所遇问题的能力。思考问题的过程未必会涉及你所学的形式数学，但学习数学的经历必定会对这一过程产生深远影响（而且是积极的！）。

目录

Preface v

1 语言、数学与模型 .1

1.1 什么是语言？.2 1.2 什么是数学？.3 1.3 什么是模型？.4 1.3.1 模型的用途是什么？.4 1.3.2 可以构建哪些类型的模型？.5 1.4 建模过程.7 1.5 模型与单位.9 1.6 数学、模型、决定论与复杂性.10 1.6.1 模型中的决定论.11 1.6.2 模型中的复杂性.12 1.6.3 模型中的简洁性.15 1.6.4 模型过拟合.16 1.7 模型构建示例：Attneave 的猫.18 1.8 习题.21

2 数学定义、概念与回顾 .25

2.1 集合与集合构造表示法.25 2.1.1 数.27 2.1.2 欧几里得空间.30 2.1.3 实数线上的区间.32 2.1.4 复数.35 2.2 函数.38 2.2.1 有界性.40 2.2.2 连续性.42 2.2.3 函数等价类.49 2.3 线性空间、线性算子与线性代数.50 2.3.1 线性算子.52 2.3.2 线性代数.58 2.4 微积分.61 2.4.1 导数.61 2.4.2 偏导数.66 2.4.3 积分.66 2.5 数列与级数.72 2.5.1 数列.73

2.5.2 级数.76 2.5.3 级数收敛.77 2.5.4 幂级数.79 2.6 泰勒定理与泰勒级数.81 2.6.1 泰勒级数.81 2.6.2 泰勒级数构造.85 2.6.3 幂级数与泰勒级数的一致收敛和逐点收敛.86 2.6.4 *逼近理论中的泰勒级数.88 2.7 *泛函与积分变换.91

3 第一章第二阶常微分方程.101
3.1 术语.101 3.2 引言.105 3.3 一阶常微分方程.105 3.3.1 检验微分方程的解.106 3.3.2 可直接积分的一阶常微分方程.108 3.3.3 不可分离线性一阶常微分方程.119 3.4 常系数二阶常微分方程.125 3.4.1 可直接积分的二阶常微分方程.126 3.4.2 关于二阶常微分方程解的思考.128 3.4.3 常系数线性二阶常微分方程：齐次情形.131 3.4.4 常系数非齐次二阶常微分方程的解 141

4 傅里叶实数集级数第一部分：入门概念 .155
4.1 术语.155 4.2 复习：幂级数.157 4.3 复习：泰勒级数.158 4.4 三角级数.163 4.5 傅里叶级数.164 4.5.1 再论正交性.165 4.5.2 傅里叶正弦级数.168 4.6 傅里叶余弦级数.174 4.7 基函数比较.179 4.8 傅里叶级数收敛：狄利克雷条件.180 4.9 频谱.181 4.10 区间变换.185 4.11 关于零对称区间上的傅里叶级数.188 4.11.1 奇函数与偶函数.188 4.11.2 区间 $I = [-L,L]$ 上的傅里叶级数.190 4.11.3 周期延拓.197 4.12 傅里叶级数收敛性简介.199 4.13 *适当基的作用不容忽视.200

5 阶跃函数 and Delta Functions . 211
5.1 术语.212 5.2 阶跃函数与狄拉克δ函数：基本思想.214 5.2.1 点源的数学建模.214 5.2.2 前进方向：极限行为.216 5.3 阶跃函数和狄拉克函数的构造.217 5.4 δ序列.222 5.5 狄拉克δ函数的性质.223 5.6 筛选性质的解释.224 5.7 *狄拉克δ序列极限的可计算示例.226 5.8 阶跃函数及其（广义）导数.228 5.9 狄拉克δ函数有傅里叶级数表示吗？.228 5.10 阶跃函数和狄拉克δ函数的一些恒等式.230

6 偏微分方程、守恒律与本构关系导论 e
Eq方程式 . 235
6.1 引言.235 6.2 术语.235 6.3 偏导数与偏微分方程的类型.237 6.3.1 偏导数.237 6.3.2 线性偏微分方程类型的特征描述.238 6.4 偏微分方程在科学与工程学中的起源：守恒公理.240 6.5 微分守恒方程的推导：连续介质力学导论243 6.5.1 守恒公理文字表述的推导.244 6.6 本构方程与通量定律.248 6.6.1 通过对流通量得到的一阶双曲型偏微分方程.250 6.6.2 通过梯度通量定律得到的二阶（抛物型）偏微分方程.252 6.7 辅助条件.255 6.7.1 偏微分方程的阶数与必要辅助条件的确定.255 6.7.2 辅助条件的类型.256 6.7.3 一维空间和一维时间下辅助条件的应用.257 6.8 获得偏微分方程解的变换.261 6.8.1 对称性与变换.261 6.8.2 变换示例.262 6.9 作为抛物型偏微分方程稳态解出现的椭圆型偏微分方程.263*附录：波动方程的推导.265

7 分离变量法（SOV）.277
7.1 引言.277 7.2 术语.278 7.3 热方程/扩散方程.280 7.3.1 热方程/扩散方程的一个性质：方差递减.281 7.3.2 热方程/扩散方程的分离变量法.282 7.4 波动方程.291 7.4.1 波动方程的一个性质：能量守恒.292

7.4.2 波动方程的分离变量法.293 7.5 拉普拉斯方程.301 7.5.1 拉普拉斯方程的一个性质：最大值原理.301 7.5.2 拉普拉斯方程的分离变量法.303 7.6 线性叠加.307

8 傅里叶变换 . 311
8.1 引言.311 8.2 术语.312 8.3 回顾傅里叶级数：复傅里叶级数.313 8.4 作为变换的傅里叶级数.317 8.4.1 线性变换.317 8.4.2 有限傅里叶变换.318 8.5 傅里叶变换.320 8.5.1 傅里叶积分的推导.321 8.5.2 傅里叶变换的记号与限制.323 8.6 傅里叶变换恒等式.324 8.6.1 平移函数的傅里叶变换.324 8.6.2 缩放函数的傅里叶变换.324 8.6.3 导数的傅里叶变换.325 8.7 卷积.329 8.7.1 卷积的定义与实例.330 8.8 傅里叶变换中的卷积.333 8.9 傅里叶变换的扩展.337 8.9.1 狄拉克δ函数的傅里叶变换.337 8.9.2 海维赛德函数的傅里叶变换.338 8.9.3 常函数的傅里叶变换.340 8.9.4 sin(ax) 和 cos(ax) 的傅里叶变换.341 8.10 *双曲型方程再论.346

9 拉普拉斯变换 .361
9.1 引言.361 9.2 术语.361 9.3 拉普拉斯变换.362 9.4 关于拉普拉斯变换函数的一些说明.365 9.5 拉普拉斯逆变换.368 9.5.1 *布罗姆维奇积分.369 9.5.2 Post反演定理.369 9.6 Heaviside 阶跃函数和狄拉克δ函数的拉普拉斯变换.371 9.7 不连续函数的拉普拉斯变换.372 9.8 导数的拉普拉斯变换.375 9.8.1 时间卷积.378 9.9 卷积与拉普拉斯变换.380 9.10 使用拉普拉斯变换求解常微分方程.383 9.11 使用拉普拉斯变换求解偏微分方程.387

10 入门读物 on Feedforward Neural Networks: An Analytical Approach 403
10.1 引言.403 10.2 术语.404 10.3 前馈神经网络图形表示的解读.411 10.3.1 更多符号说明.413 10.4 人工神经网络的历史与用途.414 10.5 前馈神经网络的一些原理.416 第一部分：线性前馈网络.419 10.6 用于线性回归的前馈网络.419 10.6.1 用前馈神经网络对数据进行线性回归的示例.419 第二部分：非线性前馈网络.423 10.7 ReLU 函数.424 10.8 ReLU 的代数性质.425 10.8.1 平移.426 10.8.2 加法.427 10.8.3 标量乘法.428 10.8.4 ReLU 函数的乘积.429 10.8.5 ReLU函数的复合.430 10.9 ReLU 函数微积分.431 10.9.1 ReLU 函数的积分.431 10.9.2 ReLU函数的导数.432 ~~10.10 *ReLU* 函数的多维形式.433~~ 10.11 关于ReLU网络与逼近理论的简要说明.434 10.11.1 脊函数.436 10.12 用于非线性回归的前馈网络.437 10.12.1 使用 ReLU 的单隐层前馈神经网络.437 10.13 非线性前馈神经网络对已知函数的逼近.447

References 457

因此，从这个意义上说，数学似乎既超越了语言，又不及语言：
虽然它的语言表达能力有限，但它似乎也包含着一种与艺术和
音乐有某些共同之处的思维形式。

A. Ford and F.D. Peat
from “The Role of Language in Science”

“All models are wrong, but some are useful.”

George Box, statistician.

“The best material model of a cat is another, or preferably the same, cat.”

Arturo Rosenblueth and
Norbert Wiener,
from “The Role of Models in Science”

1

语言、数学与模型

在开始探讨科学与工程中用于建模的各种数学方法之前，值得花一点时间讨论一些关于数学建模构成要素的、更具哲学意味（或着眼于“宏观图景”）的观点。数学建模这一主题涵盖的内容很多数学建模，其中一些内容最好通过实例和应用来理解。不过，关于一般意义上的建模，也有一些值得思考的高层次问题。接下来的材料提出了其中一些问题，其讨论比本书其余部分更偏向科学哲学。尽管这些材料带有一定的哲学色彩，但这绝不意味着它们在应用层面上不重要。理解数学建模过程中隐含的内容，实际上非常有用。例如，明确指出那些经常未被察觉的假设，是成为优秀问题求解者的关键一环！在问题求解过程中，只要可能，解释自己的假设总是个好主意；或许不必像后文材料那样深入到哲学细节，但应解释到足以帮助他人（或自己）理解问题是如何解决的程度。

由于语言（尤其是数学这种语言）是建模 的一部分，因此花一点工夫尝试准确理解语言究竟是什么，是值得的。下一节将开始探讨这个问题。虽然这个问题广泛涉及一些深奥的哲学领域，但接下来的概要讨论至少有助于揭示这些问题，同时避免过度陷入哲学结构之中。本章整体上汇集了有关建模的概念，尤其是运用数学进行建模的概念。接下来的内容有助于更好地理解建模过程，以及数学如何帮助建模者实现这一目标。

1.1 什么是语言?

本导论章节讨论语言，旨在探讨语言、数学与模型之间的连接。我们无法深入探讨语言 这一主题，但即便只是粗略概览，也有助于实现本导论的目的。

如果你曾经学过语言，可能会熟悉符号、句法和语义这些概念。符号是我们表达语言（书面或口头）的方式。句法说明如何将一组词语（符号）按照规则组合起来，以形成以下任一内容：(i) 陈述/断言，(ii) 问题，(iii) 命令，或 (iv) 感叹句（大多数语言都有这四种形式）。另一方面，语义旨在为句法赋予意义。

语言可分为自然语言和形式语言。自然语言是人类凭直觉和经验发展起来、用于彼此交流的语言。这类语言可以是口语、书面语，或两者兼有。自然语言往往具有很强的表达力；换言之，它们能够表达从主观（“你喜欢这幅画吗？”）到客观（“烤面包机坏了”）的各种概念。自然语言的难点之一在于，很难将其规则编纂成文。因此，尽管自然语言的表达力很强，但要明确掌握其全部规则却很困难。大多数试图熟练掌握一门外语的人最终都会遇到这个问题；虽然达到能够顺畅交流的熟练程度并不难，但要习得日常语言的细微之处（例如，俚语、习语和屈折变化的用法都很有挑战性），从而真正做到流利，却异常困难。

形式语言是专门制定出来的一种语言，具有一套明确规定其使用方式的规则。由于一些显而易见的原因，这类语言往往主要是书面语言。在这些情况下，人们以牺牲复杂规则下的广泛表达能力为代价，换取受限的表达能力和极其明确的规则。形式语言中最熟悉的例子，大概就是 Python 或 C++ 这样的计算机语言了。不过，形式语言的出现远早于计算机。举例来说，一阶（谓词）逻辑这种形式语言由一位名叫 C.S. Pierce 的数学家于19世纪末发现，并由著名数学家戴维·希尔伯特在约1915年将其发展为大致沿用至今的形式（Ewald，2022）。人们普遍认为，一阶逻辑足以对所有普通数学进行公理化。事实上，这表明，由于数学源自一阶逻辑的语言，数学本身具备形式语言所需的全部特性。语言与数学之间的关系本身就是一个研究课题；它十分引人入胜，因为它将如此迥异的学科交织在一起。感兴趣的学生可查阅参考文献中所列的 Rosen（2019，第13章）相关著作，了解更多内容。

虽然我们将在后续内容中讨论建模，但值得指出的是，20世纪最著名的哲学家之一路德维希·维特根斯坦曾提出，语言本身反映了现实。一些学者认为，这意味着语言本身就是一种模型。例如，Mendie 和 Udofia（2019）在阐述维特根斯坦的哲学思想时写道：“……一个命题（语言）是现实的图像，而一个命题（语言）也是我们想象中的现实模型。”我们不会进一步深入探讨语言作为一种用于创建其他模型的模型所涉及的形而上学层面。不过，在下文更广泛地讨论模型这一概念时，这仍是一个值得思考的概念。

1.2 什么是数学?

以（非常粗略的）定义语言 开篇，介绍数学与建模的教科书，似乎有些奇怪。但由于数学本身就是一种语言，这样做有助于我们确立学习数学和进行数学思维的方法。与许多宽泛的概念一样，要定义 数学是什么，实际上颇具挑战性。某种事物我们如此熟悉、经常使用，却同时又难以定义，这种情况颇为特别。不过，经历这种困难的并非只有我们。数学家、哲学家、逻辑学家、语言学家以及不同学科的科学家，长久以来一直在尝试定义什么 才是数学（例如，可参见著名数学家 R. Courant 所著的《什么是数学？》（Courant 和 Robbins，1996），了解一位专家的观点）。

从我们的角度来看，对于“什么是数学？”这个问题，我们会采取务实的态度。我们关注这个问题，主要是希望对它有所了解，从而进一步认识数学的用途和局限。这也有助于我们更好地理解如何运用数学（以及数学思维），成为更擅长解决问题的人。

在上述内容中，我们介绍了将数学视为一种语言的观点。无需过分纠结于细节（例如，数学是一种思维方式，还是其具体表达形式，关于这一讨论可参见 Ford 和 Peat（1988）），从哲学角度来看，数学是一种形式语言，这一观点已得到相当广泛的认同（Gowers 等，2008，第 I.2 节）。

数学作为一种语言，听起来或许主要是学术问题，但实际上，它与数学学习者的现实生活息息相关。一般而言，学习一门新语言被视为一项重大任务，往往需要经过多年的练习才能达到一定的熟练程度。而且，尽管数学是一种形式语言（其规则定义明确），要熟练掌握这种严格遵循规则的语言同样是一项挑战。因此，许多人觉得数学很难，也就不足为奇了——从非常现实的意义上说，学习数学就是学习一门外语。数学恰好是一种结构有限（且要求严谨）的形式语言，但归根结底，它仍然是一门语言。作为一门语言，即使从学术角度对它进行广泛研究，也未必就能特别擅长运用它。事实上，和许多语言一样，掌握数学需要练习。

对许多学生来说，将数学视为一种语言，多少会让人感到安心。首先，这意味着即使已经学习数学多年，也可以不再执着于数学理应很容易这种想法。精通数学很像是学习运用和理解语言的新方式，即便这种语言是你已经掌握了一些的。例如，能流利使用现代书面英语和口语的人，初次接触莎士比亚的语言时仍可能感到吃力。又或者，人们可能会觉得，要学习具有特定学科含义的语言颇具挑战性，例如哲学、历史或心理学中使用的术语。

这里的重点在于，将数学视为一种语言，能够（也应该）改变人们学习新的数学思想的方式。作为一种语言，要理解数学，就需要练习（例如，明确地解题）和学习。不过，一般来说，并不存在“懂数学的人”和“不懂数学的人”之分，正如我们不会把人分成“懂西班牙语的人”和“不懂西班牙语的人”。将数学视为一种语言，可以让人灵活地重新训练自己的数学思维方式，也能让人认识到，练习（以及随之而来的错误！）是学习这门语言不可或缺的一部分。

当人们接受学习数学就像学习一门外语一样时，这种认识可以为一个通常令许多人缺乏热情的学科带来全新的视角。有时，学习的过程可能涉及*un*学习那些并无助益的思维模式。在学习过程中，这种情况十分常见。

数学，并且希望“数学作为一种语言”这一视角能够成为一种组织工具，帮助人们以全新的方式理解和探索这一主题。

1.3 什么是模型?

要定义模型由什么构成，是一项艰巨的任务。若查阅科学哲学相关论述，就会发现人们对此完全没有共识。幸运的是，我们定义模型的目的更为实际，而非哲学层面。因此，就本文而言，我们可以如下定义模型。

定义 1.1（模型）。模型是一种概念或物理抽象，用于预测和解释通常更复杂的特征（某种事物）、现象（存在的某种自然实体，例如引力）或过程（例如扩散传质），这些特征、现象或过程被认为会在物理世界中发生。通常，人们认为模型将一组输入（自变量）与一组输出（因变量）联系起来（或对这种联系加以解释），但对此应作非常宽泛的理解。重要的是，模型可以描述可能仅是假设性的过程。

着重强调“物理世界中的特征或过程”是有意为之。当然，人们可以设想用于描述形而上学现象的模型。这里，我们采用形而上学一词最基本的含义，即超出目前我们所理解的物理定律可能描述的范围。例如，人们可能会建立一个模型来解释鬼魂的存在；然而，这种模型几乎肯定不在我们目前所理解的物理学范畴之内。根据语境，这仍会留下许多模糊地带：例如，社会文化现象属于科学研究范畴，那么针对这类现象建立的模型是否值得我们考虑？在这里，我们可以说，只要存在一个不违背物理定律、能够解释所关注现象的模型，那么这个模型就是值得讨论的。这一点或许在阅读下文提供的经验模型讨论后会更容易理解。

为使表述更加紧凑，今后我们可以将特征、现象或过程统称为物理系统。系统的定义如下。

定义 1.2（系统）。物理系统，或简称系统，是由相互作用或相互关联的元素（物理特征、现象或过程）组成的整体，这些元素按照一组规则（这些规则可能完全已知，也可能并不完全已知）运行，从而构成一个统一的整体。系统可分为离散系统或连续系统。对于离散系统，每个元素都是彼此独立的，元素总数原则上是一个确定的整数。对于连续系统，每个元素都被定义为整体的一部分（但定义方式并不唯一）；不过，这种划分仅是概念上的，不能将系统视为各个独立部分之和。

1.3.1 模型的用途是什么？

模型的构建源于特定的需求或目的。本文的大多数读者应该已经接触过模型，也很可能能说出模型有用的一些充分理由。模型构建之所以开展，原因如下，但此列表未必详尽无遗。

1. **问题概述与简化**。生成模型的一个常见原因是简化原本复杂得令人难以应付的系统。理想气体的例子（下文也将用来说明复杂性）就很有代表性。在接近标准温度

和压力时，对一定体积气体中的所有分子进行动量平衡，可以非常准确地求出压力。当然，这种计算相当复杂，因为一摩尔气体含有6.02×10^{23}个分子。另一种方法是使用理想气体定律$P = 1/V(nRT)$来表示压力，前提是已知V、n、R和T。显然，使用理想气体定律计算压力，比进行动量平衡计算更高效。

2. **为了理解一个系统**。有时，人们构建模型是为了更好地理解一个系统。我们仍以理想气体为例。理想气体定律最初是通过经验方法建立的（我们将在下文讨论经验模型），随着科学的发展，人们希望进一步了解气体的行为方式。统计力学是科学的一个分支，它计算大量遵循力学定律（牛顿定律或量子力学定律）的物体的统计行为。事实证明，如果对固定体积、已知温度下的大量分子应用牛顿定律，就可以推导出理想气体定律。这是通过对定义明确的微观模型求平均值，证明一条经验宏观定律的例子。这里使用“宏观”和“微观”这两个词，只是为了相对地体现两个模型所研究的长度尺度差异。由此，理想气体定律既与最初促成这条经验定律的实验一致，也与基于牛顿定律的概念模型一致。理想气体定律能够通过这两种独立的方法得出，这一事实既增强了人们对它的信心，也为我们的模型提供了额外的解释能力。

3. **用于预测系统行为**。模型最常见的用途之一是预测系统行为。这样的实例有很多，例如，可以考虑一个简单的问题：确定将用作人行桥的桁架所受的力。虽然实际设计一座桥是一个多步骤的过程，但至少可以从对拟建桥梁结构应用牛顿定律开始，以确定典型负荷下桁架各构件中的力的分布。这是本科物理学或静力学课程中经常进行的分析。掌握力的分布信息后，就可以开始确定需要哪些材料（材料类型、材料形状和尺寸），才能使结构正常发挥作用。这种方法的优点在于，可以对桁架的各种方案进行预测。不同的结构设计、不同的材料以及不同的负荷条件，都可以作为各种假设情景进行分析。这既实用又高效。利用模型预测系统行为，意味着无需实际建造许多不同的桁架，再逐一进行测试。这种建模方法的优点应该显而易见。

1.3.2 可以构建哪些类型的模型？

Restricting ourselves now to models that have some relevance to a physical system, we can identify several kinds of useful models and put them in categories. Almost every attempt to categorize broad concepts leads to some lack of distinction for particular cases, but the general organization of ideas is still a useful one. Hence, we categorize models as follows.

1. **物理模型**。本章开篇引用的这句话“猫最好的模型就是另一只猫，最好还是同一只猫”，既带有几分幽默，也旨在说明物理模型的一个重要特征。顾名思义，物理模型是由物质（在某些意义上也包括能量）构成，用来表现某种特征、现象或过程的模型。物理模型可以是模拟物；

也就是说，它可以是一种模拟在物理性质上截然不同的现象的现象。例如，稳定地下水流和电流可以用同一个微分方程来描述。因此，过去研究人员曾使用电学类比模型来模拟地下水流。物理模型也可以基于相似性的概念构建。在这种情况下，思路是建立一个代表实际物理系统、但尺度不同的物理模型。通过假设模型与现实之间一个或多个无量纲数相同，来确保物理模型与现实具有相似性。相似性这一概念似乎是由流体力学家Osborne Reynolds在研究渠道中的湍流时提出的（Reynolds，1883年）。在极端情况下，物理模型可以就是研究对象本身。在俄勒冈州立大学，我们拥有H.G. Andrews实验林，其中包含若干实验流域（流域是一个地理界定区域，降雨和降雪会汇集并流向出流口）。在这种情况下，“流域最好的模型就是另一个流域”。

物理模型的一个关键方面是，模型的输入 就是物理模型本身，以及一组特定的初始条件、边界条件和参数条件。物理模型的输出 几乎总是随时间、空间或两者变化的测量值 序列。原则上，物理模型对于其假设等内容是透明的，因此有时被称为白箱 模型（这与下文所述的黑箱 模型相对）。

2. **概念模型**。概念模型主要通过使用上文定义的某种语言 来表征。我们最感兴趣的概念模型，是那些可以使用数学语言定义的模型。不过，这并不是唯一有用的概念模型类型！使用常规英国英语（或任何其他自然语言）对系统进行解释，使用流程图，甚至使用绘图，都可以被归类为概念模型。概念模型的主要区别特征是：（i）它不是 所建模系统的直接物理模拟物；（ii）它使用某种符号语言来表达（在这里，我们将“符号语言”一词的使用范围扩展到包括流程图或描述性插图等图形表示）。使用数学（或形式逻辑）作为语言来表达的概念模型有一个优势，即模型及其模型输入、结构和输出（或结果）的交流，对于概念模型而言尽可能清晰且无可争辩。一些概念模型会明确处理不确定或模糊的数据；这类模型有时被称为灰箱模型。

3. **经验模型**。经验模型是主要基于将模型输入与模型输出联系起来建立的模型，而不一定需要理解将输入转换为输出的中间机制。因此，这类模型有时也称为黑箱模型（这一术语的起源尚不确定，但大约在 1960 年，R. Ashby 和 N. Wiener 在建模领域使其广为人知）。经验模型可能是科学中最古老的建模方式，并且已在自然科学的许多不同领域取得成功。人们可能倾向于认为，经验模型在某种程度上不如非经验模型。这通常并不正确！例如，可以考虑最初的“摩擦定律”：其最简单的形式指出，两个曲面之间的摩擦力与它们之间的法向力成正比。当然，这一直是一个实用的模型，至今仍被广泛应用。它也主要基于经验主义。虽然我们现在确实知道，观察到的摩擦力源自曲面上的微观变化（并且取决于研究尺度，还源自其他微观力），但我们很少尝试明确地对此类现象进行建模。相反，长期积累的经验性证据有力地表明，这一经验模型既有效又准确。

经验主义甚至是科学哲学的一个分支。众所周知，恩斯特·马赫（马赫数以他命名）是一位坚定的经验主义者，他认为科学应尽可能建立在严格可观测的事物之上，不应通过不可观测的量来进行解释。虽然经验

模型在自然科学中仍发挥着作用，但严格遵循经验主义哲学已不再流行；部分原因是这种方法曾遭遇一些挫折。例如，众所周知，马赫反对路德维希·玻尔兹曼等人提出的物理学原子理论。当时他的反对理由是，原子无法被直接观测，因此假定原子存在并不符合科学精神。当然，如今我们能够常规地直接测量原子（例如通过原子力显微镜），这番观点现在看来似乎有些倒退。因此，我们必须在对周遭宇宙进行纯概念建模与纯经验建模之间取得某种平衡。无法曾经 被测量的纯概念模型（例如弦理论）多少有些空洞。而纯经验模型若不敢解释导致观测值出现的现象，就无法像原本可能做到的那样推动科学进步。

4. **数字模型**。数字模型之所以最后介绍，是因为它们作为模型具有独特的地位。数字模型就本文所讨论的范畴而言，是指依靠数字计算机处理输入数据并生成输出数据的模型。颇为有趣的是，这类模型在某些方面是上述三种模型的混合体。首先，计算机本身就是一种物理装置；与用于地下水流的模拟电气模型颇为相似，数字计算机依靠物理硬件来构建模型。其次，数字模型通常是概念模型的近似。例如，当程序在计算机上使用求根方法时，其底层理论完全是概念性的；然而，实际执行会受到计算机有限算术精度的限制，因此只是一种近似。最后，请注意，计算机是由处理器、内存、存储设备等组成的复杂机器。很少有计算机用户真正详细了解计算机的工作原理。即使有人确实详细了解计算机的工作原理，通常也无法观察计算算法将输入数据变换为输出数据时发生的所有过程。从这个意义上说，我们必须在某种程度上将数字模型视为经验模型，原因至少在于它在功能上是一个“黑箱”。尽管存在这些局限，数字模型确实彻底改变了自然科学领域的建模。最近，机器学习的出现，在某些方面让科学回归经验根基：它使用户能够借助计算算法拟合或分类数据，从大量数据中获得经验性认识。尽管人们正努力让人类能够理解和解释这类模型，但目前大多数模型仍应主要被视为经验模型。这并不意味着它们没有用处；它只意味着我们有时能给出解释，却缺乏更深层的理解。

1.4 The Modeling Process

One of the beneficial uses of mathematics for scientists and engineers is that it is a natural language for problem solving; although it is hardly true that all problem solving requires mathematics. The process of problem solving involves the use of a language – which may or may not be formal language for expressing mathematics or logic– so that the problem solving process can be represented. Assuming that a problem has been identified, solving the problem using a language (including the language of mathematics involves) at least the following steps.

1. **抽象**。 确定问题后，需要用一些比实际系统更简单的术语来描述它。如果要用数学方法求解，就需要将简化后的系统“转译”为适当的数学表达式。抽象通常以普遍化的方式进行（例如，用变量而不是数来表示常数），这样抽象表示就能推广到所关注的具体情形之外。这是数学的一项强大特征：一个数学陈述可以表示问题的抽象，并实际上涵盖一整类相似问题。

表示问题抽象的数学陈述，实际上可以表示一整类相似问题。举例来说，考虑一根一端铰接、另一端自由且自由端承受负荷的梁发生小挠度（d）的问题。在适当的假设下，该问题的解取决于负荷的模（例如W）、梁的长度（L）、杨氏模量（E）以及梁横截面的惯性矩（I）。在这种情况下，可以用数学方法求解一类问题。具体而言，解的形式为

$$d = \frac{WL^3}{3EI} \tag{1.1}$$

对于W、L、E和I的所有可能（且合理）取值均成立。这是一条非常有力的陈述，涵盖了范围极广的各种物理情形。这也是抽象成为一个重要概念的原因之一；对于经过适当抽象的问题，其结果可以具有比问题的具体实例（例如W、L、E和I均为固定值时）更广泛的普适性。

在可能的情况下，抽象这一步应明确说明所采用的假设。这一步十分重要，因为它能让我们识别模型中哪些变量被视为重要，哪些变量则被假定为不重要。如果需要修订模型（见下文第4步），识别出未纳入模型的变量，就能更容易地建立一个预测能力可能更强的修订模型。

用一种语言陈述问题是解决问题不可或缺的一环。用数学陈述问题尤其有力。例如，这能让人们将问题陈述（及其附带的简化）以相当精确的方式传达给他人。由于数学语言经过形式化和简化，因此具有很高的保真度（一旦形成某个数学陈述，其含义便相当明确），但其表达力不及大多数自然语言（也就是说，人们通常不会用数学语言写诗）。

2. **Model building via algorithmic processing.** There is also a deductive process that must occur so that the problem stated can eventually be used to determine a *solution*, assuming that one exists. The solution process is example of an algorithm, and the mental process of actually working through the steps to the solution is, as mentioned in the introduction, *algorithmic thinking*. In reality, there is no need use mathematics specifically for problem solving. One could, for example, use modern English as a language and accomplish the same thing. However, the compactness of the symbols used in mathematics make it a much more convenient tool for problem solving. Another important feature is that, because mathematics is a highly constrained language used for a specific purpose, it is much less ambiguous than modern English would be.

3. **计算与具体实现**。问题一经提出，并确定了解题过程，就可以计算出一个特解。以上述梁挠度为例，具体的计算步骤是推导挠度公式，然后将特定情况下的 W、L、E 和 I 代入公式。结果将是一个具体的数值，以指定的单位表示预测的挠度。尽管模型计算的结果 是具体的（也就是说得到了一个答案），但这些结果可能无法很好地反映所求解物理问题更为重要的行为（即模型在“现实世界”中的应用）。因此，具体实现包含几个步骤：首先，必须实际使用模型计算（或测量）模型得出的结果。其次，假设模型对应于一个真实的物理系统，那么模型结果应与现实足够吻合，才能认为模型可接受。这里的足够吻合 表示特定应用所要求的保真度水平，不能脱离对该应用保真度需求的了解而单独确定。例如，人们可能希望

确保梁的设计不会失效，因为失效通常会造成严重后果；因此，建模可能需要较高的保真度。而另一个模型，比如预测你下次遛狗时狗会吃多少零食的模型，即使预测有误，也可能不会造成什么负面后果（不过，你的狗或许不这么认为）；因此，较低的保真度也是可以接受的。

4. 修订。 如果模型满足针对预期应用预先确定的某种保真度测度（如上文所述），则该模型是完备的。通常，这意味着与数据相匹配，不过也可能存在其他判定标准。我们可能会发现，最初的建模工作未能达到预期的保真度目标。如果出现这种情况，那么在抽象步骤中确定的假设和被忽略的变量，就能为构建修订后的模型提供很大帮助，使其在采用特定的保真度测度进行评估时具有更强的预测能力。在获得一个成功的模型之前，可能需要多次重复执行上述步骤 1 至 4。

修订步骤还可能涉及一个称为模型验证的过程。模型验证是确定特定模型是否适用于预期目的的过程；因此，它与修订 过程显然有关。通常，模型验证要求模型的某项性能度量指标达到规定的准确度或容差水平。例如，性能度量指标可以是：概念模型对某些观测数据的描述，其与观测数据之间的平方偏差和小于某个值 V；该值由外部约束条件决定（例如设计的成本效益、安全需求等）。对于许多应用而言，模型验证是建模过程中的重要步骤，也是建模领域中的一个完整子学科。

上述步骤描述了建模的过程。科学家和工程师时常使用模型。每当我们构建一个旨在体现更复杂系统基本特征的简单系统时，我们就是在建立一个模型。具体采用哪种模型，取决于可用的工具以及所要表示的系统类型。

1.5 模型与单位

最后，如果不讨论单位问题，本节讨论就仍不完整。纯数学无需引入单位也能自成体系，但在对物理系统进行建模时，使用单位至关重要。尽管我们大多数人都知道单位一词的含义，给它下个定义仍然很有用。使用一套单位，可以识别并传达所建模的物理量的类别；也可以用某种标准化的方式确定这些量的模。形式上，定义如下。

定义 1.3（基本单位）。一组基本单位为各种物理量提供标准化的名称和模。它们构成测量系统的基础，其他物理量则可以表示为这组基本单位的倍数。

An important concept here is that there is a set of *base* units which all other *units* are constructed from. For example, in the International System of Units (SI) , the fundamental units are:

1. The meter (symbol: m), used to measure length.
2. The kilogram (symbol: kg), used to measure mass.
3. The second (symbol: s), used to measure time.
4. The ampere (symbol: A), used to measure electric current.

5. 开尔文（符号：K），用于测量温度。6. 摩尔（符号：mol），用于测量物质中的物质量或粒子数量。7. 坎德拉（符号：cd），用于测量光强度。

单位的写法有一些值得注意之处。许多单位以人名命名，但这类单位的全称不大写。不过，相关单位的符号通常要大写。因此，我们写作 1 安培 $= 1$ A。这样合理吗？其实，这个问题并不重要：既然已经确立了一套规范，那么在这里该怎么做就很明确了（即使这种写法为什么如此并不清楚！）。有兴趣的学生可以查阅美国国家标准与技术研究院（NIST）度量衡办公室（https://www.nist.gov/pml/owm/writing-si-metric-system-units），进一步了解单位名称和惯例中那些复杂之处。

这里要说明的要点是，这个基本单位集合可以描述所有可能存在的其他物理量。举一个具体的例子，我们可以考虑电压的概念（在许多方面，它与流体系统中的压力类似）。我们习惯用伏特表示电压（例如，在北美洲，单相家庭电路的电压约为 120 伏特，频率为 60 秒$^{-1}$）。不过，伏特并非基本单位。请注意以下关系：一伏特等于一焦耳每库仑。一焦耳等于一牛顿米。一库仑等于一安培秒。因此

$$1 \text{ volt} = 1 \frac{\text{J}}{\text{C}} = 1 \frac{\text{N}\cdot\text{m}}{\text{A}\cdot\text{s}} = 1 \frac{(\text{kg}\cdot\text{m}/\text{s}^2)\cdot\text{m}}{\text{A}\cdot\text{s}} = 1 \frac{\text{kg}\cdot\text{m}^2}{\text{A}\cdot\text{s}^3} \tag{1.2}$$

On the farthest right expression in Eq. (1.2), the volt is expressed in its base SI units. The list of seven units provided above are sufficient to express *all possible physical quantities that are known today*. While one might be tempted to believe that "units" would not continue to be researched today, it turns out that it is still a subject of much discussion! There are efforts currently to establish all units in terms of "universal constants" (e.g., the gravitational constant, G, or Plank's constant $\hbar$), and thus the topic is still one of evolving research (e.g., see Bordé (2005)).

1.6 数学、模型、决定论与复杂性

When developing models, one usually makes certain kinds of very basic assumptions about the ultimate utility or purpose of the model. For example, quite frequently we *assume* that the models that we generate will allow us to make some kind of *prediction*; this is indeed often the purpose that we bother to generate models in the first place.

尽管人类数千年来一直成功地运用模型，但在（相对而言）较近的时期，我们发现，有时我们建立的模型所提供的信息，在定性（以及定量）上都与我们通常期待从模型中获得的信息不同。更具体地说，我们了解到，并非所有模型都可以被认为是确定性的 （这个词将在下文作更详细的定义）。相反，我们发现，有些模型的结果对其模型参数（例如密度等物理属性）、初始状态下的条件或边界条件极为敏感，以至于在实际中我们无法作出确定性的预测。这并不意味着这类模型毫无用处。但这确实意味着，人们期望从这类模型中获得的信息，与期望从表现出确定性行为的模型中获得的信息有所不同。关于决定论及其相关概念复杂性的讨论，将在下文展开。

1.6.1 模型中的决定论

决定论是希腊哲学家在公元前7世纪和6世纪发展出的一种思想框架。简言之，这种哲学认为因果关系紧密相连，原则上可以得到理解。它假设，如果观察者掌握了关于某个对象的足够的信息，就可能预测该对象随后的一举一动。到了17世纪中后期，著名的哲学家和科学家（如勒内·笛卡尔和艾萨克·牛顿爵士）将这一观点确立为物理世界的自然组成部分。然而，将决定论的思想推进到现代科学话语中的，是数学家兼科学家皮埃尔-西蒙·拉普拉斯。拉普拉斯明确指出（拉普拉斯，马尔基斯·德）

> 因此，我们应当把宇宙的当前状态视为其先前状态的结果，同时也是将要出现的状态的原因。假设有这样一种智能：它能够理解自然界中一切起作用的力，以及构成自然界的各种存在物的相应状态——这种智能足够强大，能够对这些数据进行分析——那么，它就能用同一个公式概括宇宙中最大天体的运动和最轻原子的运动；对它而言，没有任何事情是不确定的，未来就像过去一样，尽在其眼前。

这种观点几乎被视为一条物理定律。然而，在19世纪末，亨利·庞加莱向世人指出，系统或许并不总是如人们所理解的那样具有确定性。庞加莱发现，有些微分方程系统（这些方程反过来又是物理系统的模型）中，最终的解对初始条件非常敏感。庞加莱的观察促使他提出了如下表述

> 一个我们无法察觉的极小原因，决定了一个不容忽视的显著结果，于是我们说，这一结果是由偶然造成的。假如我们准确地知道自然定律以及宇宙在初始时刻的状态，就能准确预测同一宇宙在随后某一时刻的状态……然而，事情并非总是如此，初始条件的微小差异可能会导致最终现象出现极大差异。前者中的一个微小误差，会使后者中的误差变得巨大。如此一来，预测便不再可能，我们面对的就是随机现象。

多年来，人们不断研究并完善了这一观测；如今，我们将对这类敏感系统的研究（以及其他一些名称）称为混沌理论。简而言之，这意味着数学系统和物理系统都可能存在这样的情况：尽管我们竭尽全力，采用最先进的技术，仍然无法充分了解系统，以至于无法确定地预知其未来行为。如今，由于我们对量子力学已有一定了解（哪怕只是浅显的了解），我们对这一想法多少有些熟悉；然而，许多人仍会惊讶地发现，即使是非量子力学系统（或宏观系统），其行为也可能无法预测。

The non-deterministic behavior of physical and mathematical systems led to somewhat of a revolution in the way that physicists, scientists, and mathematicians thought about and represented models of processes that occur in our world. What we once thought of as *knowable* if given the proper kinds and amount of information was now understood to be knowable only within certain bounds. Even without the strangeness embodied in quantum mechanical systems, we found that some large-scale physical systems (such as the weather) could be so sensitive to the initial conditions or other physical properties, that their long-term behavior could not be modeled with certainty. Primarily, this kind of behavior was observed for *nonlinear* models. The concept of nonlinearity is explained in more detail in the next chapter. In short, however, one can think of nonlinearity in this application as meaning that small changes to the system can yield arbitrarily large responses. This does not happen in so-called *linear* systems, where small changes to the system always yield a response that is small in some sense.

The gradual unraveling of the idea that the physical world and associated models of it should be *deterministic* represented a significant change in the way that scientists and mathematicians thought about the models

they used. Because part of the purpose of a model is to describe the complicated universe by collecting and utilizing understandable ideas, it seemed that the lack of determinism created a problem. This problem was specifically that *some* reasonable models of physical processes yielded behavior that was inherently non-predictive in some ways. The consideration of this problem led to new research that eventually became known as *complexity science*(Phelan, 2001), or simply *complexity*.

1.6.2 模型中的复杂性

在英语中，当我们说某事复杂时，通常是指它“不容易理解”。这可能是技术性的评论，也可能是对现实世界中其他因素的描述（例如，我们可以毫无技术含义地说，中东的地缘政治局势很复杂，而无意涉及复杂性的任何技术概念）。同样，在数学中，complex 也可以表示涉及具有虚部的数。在建模中，我们通常用“复杂”一词表达的既不是上述含义（尽管确实有些情况下，这些定义中的多个可能同时适用！）。虽然本文只会探讨少数可视为复杂的模型，但理解复杂性对于建模的意义至关重要。

模型的复杂性在不同学科中有不同含义。然而，几乎所有与建模相关的定义，都涉及回答以下问题：要解释我们的模型，并说明它如何用于预测我们试图描述的行为，有多困难？我们也可以把“有多困难”换成“需要多少信息”，使问题更具体一些（不过，这还需要进一步定义信息，而我们在此不打算详细讨论！）。回答这个问题的一个关键概念，是尝试界定模型及其解中的规律性（Phelan，2001）。复杂性科学引入了识别和研究规律性的新方法，甚至适用于非确定性系统；重要的是，这些方法代表了看待问题的新视角（Phelan，2001）。人们通常会想到在非确定性系统中寻找规律性，但这个思路对确定性系统同样有用。举例来说，假设我们想测量10,000个滚珠轴承的重量。一种确定性的做法是测量每个轴承，但也可以根据实际所需的信息采用其他策略。如果需要每个轴承的实际重量，我们确实必须逐个称量（并用标签做好记录）。然而，假设我们只需以特定置信度了解重量（比如95%的置信度）？那么，我们可以测量随机选取的轴承子集的统计数据，并对平均值和标准差进行统计推断（通过足够数量的测量值确保准确度达到95%）。当然，这样做时，我们对系统的规律性作出了一个强假设。举例来说，就轴承而言，我们假设这10,000个轴承的总体具有足够的规律性，因此可以用标准统计学方法对其进行概括。这是一个关于系统规律性的建模假设，它使我们能够提出一个用于测量轴承重量的确定性模型（尽管所得模型是统计模型，但按照我们对模型的定义，它在确定性意义上仍是确定性的），而不必逐个测量轴承。因此，对模型中的规律性作出假设并加以探寻，可以成为一种非常有力的工具！

那么，我们该如何定义“复杂性”这个词呢？同样，目前还没有一个得到普遍认可的统一定义。不过，在本文中，我们将对建模中的复杂性作如下定义。当应用于建模时，“复杂的”（或“复杂性”）一词意味着以下至少一种情况成立。

1. **大量变量**。有些模型包含的自由度（即变量）太多，以至于实际中无法或不适合求解该系统。一个例子是这样一种模型：将一摩尔理想

气体视为遵循经典力学（牛顿定律）的粒子集合，并处于固定温度下。虽然这种模型确实是确定性的，而且对于某些气体和状态也相当准确，但很少有人会认为，追踪每个分子的 3 个动量变量和 3 个位置变量（总计 $(3+3)\times 6.02\times 10^{23}\approx 3.6\times 10^{24}$ 个变量！）是切实可行的计算。不过，借助统计工具，并假设气体分子的速度表现出某种规则性，我们实际上可以通过考虑大约 10^{23} 个粒子所遵循的牛顿定律，推导出理想气体定律。这里规则性假设的作用显而易见：我们不必使用数量级为 10^{24} 的变量来描述理想气体，而可以使用经典理想气体定律$pV = nRT$。对于本例中处于固定温度下的一摩尔气体，规则性假设使我们能够用变量（p，V）和一个常数（R）来描述该系统。这大大降低了解释该系统模型（或该模型所蕴含的“信息”）的难度！

2. **Non-deterministic behavior.** A second meaning for complexity in a model is the condition where the model behavior is so sensitive to initial, boundary, or other conditions (e.g., the exact size of a particular parameter) that it is nearly impossible to predict the behavior of the model in a classical sense. Such models are called *chaotic*, as discussed above. Colloquially, we can think of such models as ones that will predict a very different outcome with even a very tiny change in the conditions (initial condition, parameters, etc) the describe the model. Because we seldom know initial, boundary, or parameters for real systems with high accuracy, the resulting model has very little predictive value because small errors in the conditions can yield wildly different results. In short, many systems in science and engineering subscribed to the rule that “small perturbations create small effects” (and this is always true in a sense for linear problems). For chaotic problems, this rule is no longer true.

3. **涌现行为**。这一概念并不容易进行具体描述，但如今，在描述某些模型（通常是非线性模型）的行为时，涌现 一词已经十分常见，因此有必要理解涌现行为 的本质特征。用通俗的英国英语来说，行为描述的是这样一个过程：模型从一个并不十分“有趣”的状态出发，转变为具有高度结构性、信息含量或呈现出未必能够预期的行为的状态。换言之，模型以一种出人意料的方式从一种原型过渡到另一种原型。虽然这很难算是一个具体定义，但似乎也不存在具体的定义。最早对这类现象开展详细研究的人之一，是著名的英国数学家艾伦·图灵（他主要因20世纪40年代和20世纪50年代在计算机领域开创性的工作而闻名）。在图灵的应用中，涌现行为由一个反应-扩散方程给出，并被提出作为解释生物学中某些图案如何形成的一种可能机制（例如猎豹身上的斑点，或某些贝壳上的图案）。为了纪念这一发现，这些图案有时被称为图灵斑图。对于感兴趣的读者，Krause等人（2018）对一般意义上的涌现行为进行了综述。

In each of these kinds of complexity, the search for regularity in the models and their solutions can yield understanding in areas where it would be otherwise lacking.

The fairly well known example of complexity in mathematics can be found in the example of certain fractals. As a specific example, one can point to the *Julia set*, which is a chaotic function of the initial point selected for iteration. Most of us have seen images of the Julia set, and these images have almost reached the level of pop culture; a cursory examination of Fig. 1.1 might suggest that it is a familiar looking plot. The Julia set is usually computed via iteration. Even a small change in the initial condition for these iterations will yield a dramatically different set of results for the “shape” of the resulting Julia set. An example of high

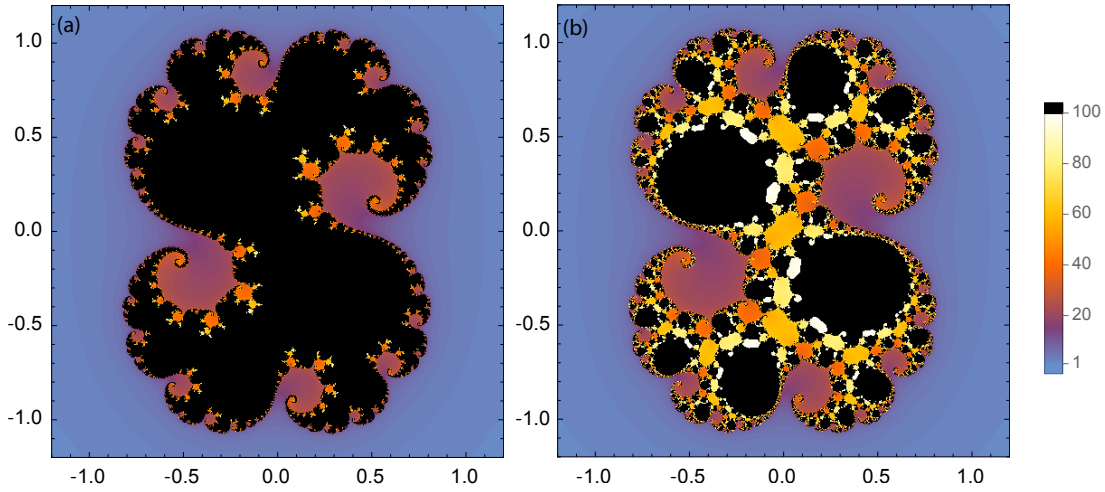


图 1.1：朱利亚集的两种表征。通过选取一组初始值网格（x,y），并计算$z_{j+1} = z_j^2 + c$直至收敛，可以迭代确定该集合，其中$z_j = x_j + iy_j$。c的值是一个形式为$c = c_x + ic_y$的常数。选择不同的c值，就会得到不同的朱利亚集。上方图像是通过计算每个收敛值$z_\infty = x_\infty + iy_\infty$在网格上的模，并计算其模$m$得到的，其中$m = (x_\infty^2 + y_\infty^2)^{\frac{1}{2}}$。(a) $c = 0.274 - i0.008$。(b)$c = 0.275 - i0.008$。对于这些图像，c的实部变化 0.4%，就会导致所得函数的模发生显著变化。

朱利亚集的敏感性如图 1.1 所示。在这里，模型中某个参数仅发生 0.4% 的变化，就导致集合的模（以颜色表示）发生了剧烈变化。

作为复杂性的第二个例子，我们可以考虑关于涌现行为这一概念所给出的第三个定义。非线性反应扩散系统是最古老的一类能够展现涌现行为的系统实例。方程 (1.3)-(1.4) 所描述的扩散反应系统称为具有菲茨休-南戈莫反应的图灵模型；这些方程表示两种化学物种的扩散混合与反应（其物种浓度由变量u_1和u_2表示），它们能够在稳态下于空间中形成自组织非均匀图案。

$$\frac{\partial u_1}{\partial t} = D_1 \left(\frac{\partial^2 u_1}{\partial x^2} + \frac{\partial^2 u_1}{\partial y^2} \right) + (u_1 - u_2) - u_1^3 + \kappa \tag{1.3}$$

$$\frac{\partial u_2}{\partial t} = D_2 \left(\frac{\partial^2 u_2}{\partial x^2} + \frac{\partial^2 u_2}{\partial y^2} \right) + (u_1 - u_2) \tag{1.4}$$

图 1.2 绘出了u_1在三个时刻的空间分布。在零时刻（左图）u_1的浓度（以颜色表示）在空间中是随机的。到了中间时刻（中图），扩散与非线性反应共同作用，空间结构开始显现。最后，在近稳态条件下（右图），空间图案已清晰形成，呈现出u_1的高浓度岛状区域。虽然u_2的条件未绘出，但其初始条件和演变过程相似。

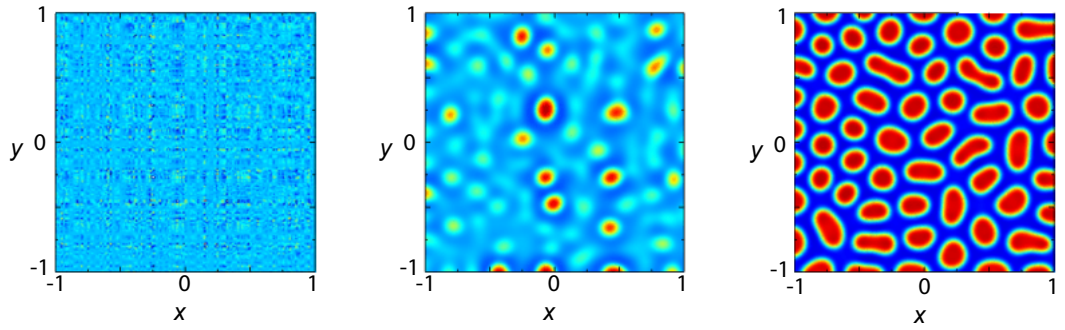


图 1.2：方程 (1.3)–(1.4) 所给出的反应扩散系统中物种u_1 的三种空间分布。这些图像表示物种u_1 的浓度，分别对应$t = 0$、100 和 1000 秒，展示了涌现图案形成。（左）在$t = 0$时，两种物种的初始配置是空间中的一个随机函数。（中）$t = 100$ 秒时的图案。（右）$t = 1000$ 秒时的图案，代表接近稳态条件。模型采用以下参数运行：$D_1 = 0.00028$ 厘米^2/秒，$D_2 = 0.005$ 厘米^2/秒，$\tau = 0.1$，以及$\kappa = -0.5$ 毫摩尔/(厘米^3·秒)。色标中，红色表示浓度高，深蓝色表示浓度低。Hubodeker 创作的“图灵分岔”图像采用 CC-BY-SA-3.0 许可。

1.6.3 Simplicity in Models

将简洁性应用于建模时，它并非“复杂性”的反义词。模型简洁性所指的是，最有效的模型只施加捕捉所建模现象所必需的约束，不多不少。事实证明，这一思想在科学哲学中由来已久。模型应尽可能简单这一观点，至少自 13 世纪初以来便已存在（Franklin，2001）。这一原则通常被称为奥卡姆剃刀（这里的“剃刀”是一个形象的说法，表示这一思想会剔除不必要的部分），其名称来自一位名叫奥卡姆的威廉的英国方济各会修士。虽然奥卡姆并非这一思想的创始人，但他确实使其在科学哲学中广为流传。他的一种表述版本（实际上见于其著作中）如下：

> Plurality must never be posited without necessity

In short, the idea is this: Given two competing models for a system, each with similar predictive power, the simpler model should be preferred. Sometimes this idea is also called the *principle of parsimony*, for obvious reasons.

The statement of Occam's razor has been repeated many times over the decades. Some famous examples are as follows.

- Isaac Newton stated in his famous text the *Principia* "We are to admit no more causes of natural things than such as are both true and sufficient to explain their appearances."
- Einstein can be quoted as stating (in a lecture at the University of Oxford, 1933) "It can scarcely be denied that the supreme goal of all theory is to make the irreducible basic elements as simple and as few as possible without having to surrender the adequate representation of a single datum of experience."

- 一位名叫罗杰·塞欣斯的作曲家曾错误引用爱因斯坦的话：“我还记得爱因斯坦的一句话，这句话当然也适用于音乐。 [爱因斯坦] 大意是说，一切都应尽可能简单，但不能过于简单！”。
- 诺贝尔奖得主、医学研究人员西奥多·伍德沃德曾给出一个令人难忘的说法。他用一句格言阐释了这一概念：“听到马蹄声时，想想是马，而不是斑马。”尽管这一原则原本旨在用于疾病诊断，但它形象地概括了奥卡姆剃刀的核心思想。

虽然这一切听起来多少有些定性，但事实证明，它可以调整后应用于定量环境。20世纪50年代中期，一位名叫埃德温·杰恩斯（Edwin Jaynes）的统计学家向物理学界展示了一个令人惊讶的结论：奥卡姆剃刀以最大熵原理的形式表述时，可以使人们准确地推导出统计力学结果。用传统语言来说，杰恩斯证明，在某些统计力学计算中应用最大熵这一统计学概念，等同于引入最少量的假设。杰恩斯还证明，在经典统计力学中所需的传统假设，可以通过在已知约束条件下最大化熵来取代。这样做将产生一个最大限度地避免不合理假设的解。这在我们思考系统建模的方式上带来了某种程度的革命，最终也成为奥卡姆剃刀在实践中的一种切实体现。因此，最初作为一种哲学原理（奥卡姆剃刀）提出的思想，最终被证明能够以可量化的方式发挥实际作用。

1.6.4 过拟合模型

建模中一种与简约性概念相关的现象是模型过拟合问题。术语 过拟合 源自统计学和数据科学等相关学科。从概念上讲，这个想法很简单。过拟合是指对一组特定的现象数据进行非常精确的建模，但这种建模方式无法很好地推广到同一现象的其他数据集。通常，这是由于没有遵循奥卡姆剃刀原则所导致的。通过一个例子来解释这个问题最为简单。

流体中运动物体所受的阻力已被研究了数百年，人们常常发现阻力与速度的平方之间存在关系。假设我们开展两组相互独立的实验，在风洞中测量物体在不同速度下所受的阻力；我们将这两组实验称为实验 A 和实验 B。我们使用实验 A 来校准一个模型，以便预测某一速度下的阻力。然后，我们使用根据实验 A 校准的模型，来预测实验 B 中测得的数据所反映的行为。

在图 1.3 中，我们绘出了原始数据（包括 95% 置信区间），并用两种不同的多项式拟合数据。首先，根据我们对阻力定律的了解，二次式应能较好地拟合这些数据。拟合得到的二次式（约束为处处为正）似乎与数据吻合得不错，而且对每个数据点而言，其结果都落在 95% 置信区间内，因此可以接受。在图 1.3(c) 中，我们展示了用五阶多项式（约束为处处为正）拟合数据的结果。包括原点在内，总共有六个数据点。总是可以找到一个 n 阶多项式，使其经过 n 个 $+1$ 点。因此，这种情况下的拟合是完美的。但这是最优拟合吗？这在某种程度上取决于我们的目标。首先，我们知道阻力定律通常是速度的二次函数，因此五阶多项式与我们的“先验数据”相悖；这些数据表明，我们可能会预期得到二次形式。虽然五阶多项式确实完美地拟合了数据，但我们不禁要问，曲率上的额外变化是否代表了任何物理现象，还是仅仅非常精确地表征了我们知道实验中必然存在的实验误差。

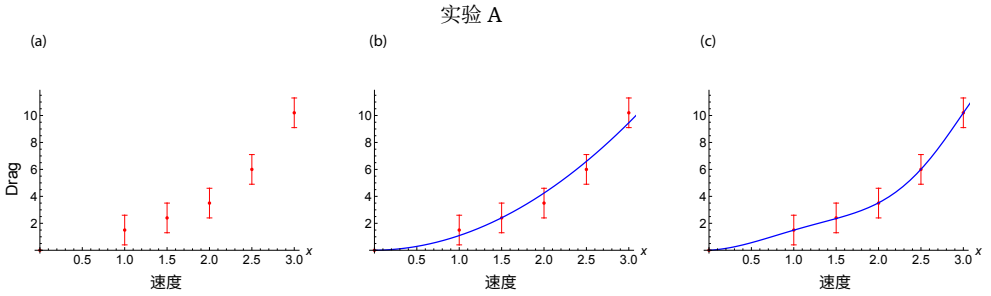


图 1.3：阻力实验的结果，实验 A。(a) 实验 A 的阻力数据。(b) 形式为 $F_D = 1.05U^2$ 的二次式最小二乘拟合（限制为正值）。(c) 完美拟合 6 个数据点的五阶多项式；$F_D = -0.137553x^5 + 1.17801x^4 - 2.89406x^3 + 2.66579x^2 + 0.785579x$

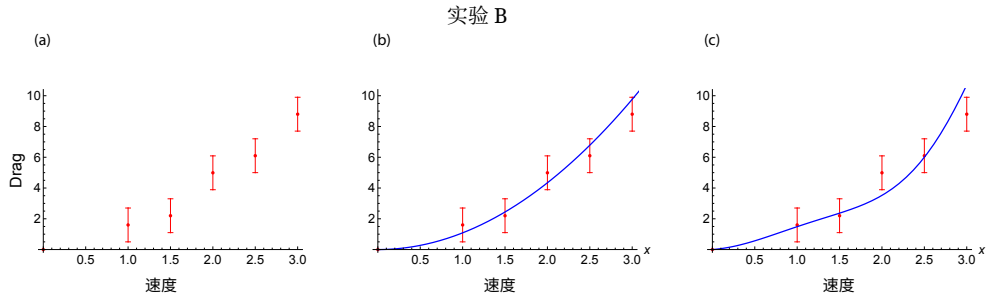


图 1.4：阻力实验的结果，实验 B。(a) 实验 A 的阻力数据。(b) 拟合模型 A 并在模型 B 上测试的二次式最小二乘拟合（限制为正值）。(c) 模型 A 的五阶多项式拟合，并在模型 B 上测试。可以看出，对于实验 B 的数据，二次式模型实际上比五阶多项式模型拟合得更好。这是过拟合造成的。

开始回答这些问题的一种方法，是看看这个模型的泛化能力如何。换句话说，我们有两个由实验 A 得出的模型。如果这些模型的泛化能力良好，就意味着即使我们再采集一组实验数据（其中也包含测量误差），模型仍能较好地拟合数据（也就是说，它能够“解释”这些数据）。图 1.4 展示了使用实验 A 得到的模型拟合实验 B 数据的结果。图 1.4(a) 展示实验 B 的数据，(b) 和 (c) 部分则展示二次式和五阶多项式模型对数据的拟合结果。此时，五阶多项式拟合得更好已不再显而易见。事实上，由于多项式中增加了额外项，它对数据的解释反而不如二次模型准确（也就是说，有两个数据点不在五阶模型的 95% 置信区间内）。

这是一个过拟合问题的好例子。这里发生的情况是，我们增加了不必要的自由度（即额外的常数及其伴随的多项式函数 x^3, x^4,x^5）。尽管这些额外的自由度使我们能够针对一组特定的阻力数据生成更准确的模型，但该模型未能很好地泛化。这是因为该模型过于具体。

对我们已有的数据来说。我们不必要地给模型增加了额外的自由度，这与奥卡姆剃刀相悖。这是因为我们忽略了先验数据（也就是说，我们知道阻力模型通常是二次式的）。然而，借助最大熵的形式体系，我们也可以考察许多这类实验，并发现五阶模型产生的误差并不服从正态分布（这是二次拟合下的预期结果），而是服从某种偏态分布（因为它会在速度值域的中间部分系统性地低估阻力）。可以证明，偏态分布的熵更低，而正态分布的熵更高。简而言之，分布的熵越高，描述该分布所需的参数就越少。因此，奥卡姆剃刀会建议我们选择熵最大的分布（假设最少），在本例中也就是二次模型。有趣的是，奥卡姆剃刀在模型拟合问题上也能提供定量见解。事实上，这类熵技术已成为蓬勃发展的机器学习领域中的重要工具。

1.7 模型构建示例：Attneave 的猫

The topic of how geometric shapes are processed by the human visual cortex has been a topic of both qualitative and quantitative research for many decades (Köhler and Held, 1949), and it continues to be an active area of research for a variety of reasons ranging from the psychological (e.g., understanding the structure-function of the brain) to the practical (predicting how best to present information visually to an active user, such as in a "heads up" display).

In 1954, a psychological researcher named Fred Attneave proposed a model where he postulated that in 2-dimensional contour figures, the regions of high curvature would contain the greatest amount of information. While this theory is still being discussed in the literature (Torfs et al., 2010; De Winter and Wagemans, 2008), it appears to be one component of a possibly more complex understanding of human perception and vision.

考察第 1.3 节中列出的问题求解步骤如何应用于 Attneave 的猫，是很有意思的。下面的列表将讨论对每个步骤的思考。

1. **抽象化。** Attneave 的研究中所涉及的抽象化步骤具有重要意义。Attneave 并不是试图建立猫本身的模型，而是试图建立猫的图像模型。抽象化步骤是 Attneave 提出（并尝试回答）"究竟是什么使一幅图像能够被识别"这一问题的过程。具体而言，Attneave 希望剔除所有不必要的内容，从而使注意力能够集中在一两个元素上。最终，他选择将图像表示为轮廓，其中直线段通过高曲率线段连接起来。在本例中，抽象化步骤的主要输出是形成了一个相对简单的假设：高曲率区域是轮廓图像中信息量最大的部分。在抽象化步骤中，明确自身的假设，以及在模型修订时可能需要考虑的其他变量或替代变量（如果确有必要），非常重要。在 Attneave 的论文中，这一点体现在讨论和结论部分，他指出其他信息（例如纹理、颜色、对比度和相似性）也可能具有重要意义，但他的研究结果主要聚焦于曲率。明确所采用的简化方式是抽象化过程中必不可少的一步。模型很少能够通过一次努力就完成；更常见的情况是，需要经过多次迭代。为了便于进行这些迭代，清楚了解抽象化步骤中模型包含了哪些内容、舍弃了哪些内容是很有帮助的。随后，可以在修订过程中重新审视这些变量，并决定是否需要引入其中一个或多个变量，以达到所期望的模型保真度。

2. **Model building via an algorithmic process.** The algorithm used to build the models is slightly more complex than that used simply to draw Attneave's cat. First, Attneave wanted to test the hypothesis that

curvature was indeed important. This was done by experiments with a large number of subjects who looked at simple closed curves (think potato-shapes here), where subjects were asked to place points on the curve such that they could re-draw the curve using only the points selected. From this, he was able to determine that his abstraction (segments with high curvature contain the most information) appeared to have validity. This provided good evidence from which he could generate a "model building" algorithm. Without belaboring it too much, the algorithm would look something like the following:

a. 对于给定的轮廓图像，确定曲率最大的区域。Attneave 并未详细研究轮廓图像本身的生成过程。不过，他大致勾勒了如何利用边缘检测消除冗余信息，以确定图形轮廓可能绘制的位置。

b. 找出轮廓图形中曲率最大的区域。你可能在微积分课程中学过，曲率是沿曲线移动时切向量的变化率。对于平面曲线，这可以表示为

$$\kappa = \frac{|x'y'' - y'x''|}{\left(x'^2 + y'^2\right)^{\frac{3}{2}}} \tag{1.5}$$

虽然 Attneave 没有通过这种方法计算局部曲率，但计算起来并不困难。图 1.5(a) 展示了绘制出的 Attneave 的猫的一个版本。图 1.5(b) 则用色标沿着这个名为 Attneave 的猫的图形轮廓标出了曲率的绝对值。虽然这幅图像是对轮廓图像进行后期处理后的版本，但仍然可以清楚地看出图形中曲率较大的区域。

c. 保留曲率最高的区域，将轮廓的其余部分替换为连接每个高曲率线段端点的直线段。Attneave 并未详细说明具体方法，但可以设想使用局部曲率阈值处理，来决定保留哪些部分、舍弃哪些部分。同样，参照图 1.5(b)，可以看出紫色区域代表轮廓中曲率几乎为零的部分。

尽管这确实是一个非常简单的算法，但它仍然可以应用于几乎任何轮廓图像，而且基本不会产生不确定性。有些图像（例如曲率恒定的圆）几乎不会发生变化，而另一些图像（例如猫的图像）则会发生显著变化。

3. **具体实现**。实现过程，顾名思义，就是对构建好的模型进行"实现"。也就是说，在确定合适的抽象和算法之后，使用该算法处理输入数据（在本例中为轮廓图像），生成输出（在本例中为简化后的轮廓线）。

4. **Revision.** While no specific revision step is conducted in this particular modeling exercise, note that a number of important assumptions were identified by Attneave in his original (1954) paper. Attneave's paper has been revisited many times by other researchers, and some of these have found that the neglected variables listed by Attneave were found to increase the descriptive ability of the models that they form. So, although model revision was not a component of Attneave's model (at least, no discussion of alternative models was discussed in the paper; whether or not revisions were made in his initial work is not known), because he did describe a number of assumptions and neglected variables, other researchers have been able to extend his model to improve (to some extent) Attneave's original model.

Attneave's cat provides a useful, and novel example of the problem solving process, with a focus on model development. It is good to think about mathematics as being only one component of problems solving



Fig. 1.5: Attneave's cat. (a) A smooth contour outline of a cat. (b) The post-processed contour outline of Attneave's cat, expressed as straight lines connected by segments of high curvature. This figure is generated by applying the algorithm described in the text. (c) For reference, the rainbow color scale indicates the (normalized, absolute value) curvature associated with the outline of the figure in (b). Areas of high curvature are yellow to red; areas of low curvature are blue to purple.

and modeling. For many practical scientific and engineering problems, however, mathematics is the natural language for expressing models. A solid background in mathematics will both provide useful modeling tools, as well as experience with the problem solving process in general. Even if you never again use mathematics in problem solving (although this is hard to imagine for most scientists and engineers), the process of learning mathematics is, itself, a compelling method to practice the problem of problem solving. Problem solving skills, once learned, have considerable transferability. Learning and practicing mathematics will inevitably make you a better problem solver, regardless of the kind of problems that you are presented with.

1.8 习题

1. 本章开篇引用了这样一句话：“猫的最佳模型是另一只猫，最好就是同一只猫。”这句话虽然有些幽默（而且提到了猫），却也蕴含着重要意义。请结合对系统及其抽象的思考，写几句话解释这句话想要说明什么。

2. 本章开篇还引用了另一句话：“所有模型都是错误的；有些模型是有用的。”请结合对系统及其抽象的思考，写几句话解释这句话想要说明什么。具体来说，请解释“所有模型都是错误的”何以成立？有没有任何一个“模型”是正确的？（提示：参见习题1中重复引用的那句话。）

3. 几千年来，圆一直令世人着迷。巴比伦人和埃及人都曾使用圆周率的近似值；圆周率是圆的周长与直径之比。不过，真正开启这一研究（可以这么说）的是叙拉古的阿基米德（约公元前287—前212年）。他的想法是在圆内作正多边形。他推测，边数越多，这些多边形的面积就越接近外接圆的面积。利用现代三角学不难证明

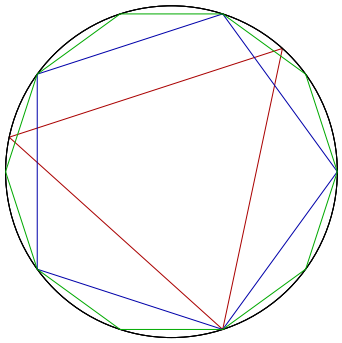


图 1.6：内接于圆的多边形。

$$A(r,n) = nr^2 \cos\left(\frac{\pi}{n}\right) \sin\left(\frac{\pi}{n}\right) \qquad (1.6)$$

其中 $n \geq 3$ 是内接多边形的边数。我们可以将$A(r,n)$视为圆面积的一个模型，其中圆的半径为r。计算半径为 $r = 1$厘米的圆在n 分别等于 3、4、5、10、30 和 50 时的面积，并计算这些取值下的相对误差（$A(1,n) - \pi$）$/\pi$。然后回答以下问题。

a. 这个模型看起来是一个可靠且有用的模型吗？ b. 你能证明当n 任意增大时，这个模型会精确收敛到圆的面积吗？（提示：当 θ 很小时，$\sin(\theta) = \theta$）。c. 假设你有一个特定的应用场景，打算用$A(r,n)$来估算圆的面积。为了选定要使用的n 值，你需要考虑哪些因素？

4. **混沌行为**。 对于某些类型的现象，从昆虫种群随时间的变化到钻井钻头的磨损速率（参见 Li 和 Yorke，1975），以下模型可以描述离散时间区间内的状态关系。

$$x(n+1) = rx(n)\,[1 - x(n)] \qquad (1.7)$$

其中 r 是增长率参数，n 表示时间周期的数量（每个周期具有相同的区间，例如 Δt），而 $x \in [0,1]$。形如 $x(n+1) = F(n)$ 的方程（如上面的方程）将系统的当前状态 与前一状态联系起来。它们称为 差分 方程；具体而言，方程 (1.7) 有时称为逻辑函数 方程。

从 20 世纪 60 年代开始，一些研究人员逐渐意识到，这类简单的方程有时会产生出乎意料的复杂（以及本章前文所述意义上的复杂）行为。特别是，对于某些r取值，方程（1.7）会对其初始值变得非常敏感。这个表达式非常简单，可以轻松地用计算机语言编写代码，甚至可以通过电子表格程序进行计算。使用你选择的方法，计算并绘制给定$x(1)$时，直至$(n+1) = 50$的解。换句话说，给定第一个值后，你需要计算方程（1.7）的一组包含 49 个值的序列。应使用的参数$x(1)$和r请参见下表。

Case	x(1)	r
1	0.5	1
2	0.5	2
3	0.5	3
4	0.5	3.99
5	0.5001	3.99

表1.1：使用逻辑函数差分方程的模型值。

确定解后，比较情形1和情形2。它们之间有什么不同？情形3又表现出不同的行为——你会如何描述情形3随时间（即步数）变化的行为？最后请注意，情形4和情形5相似，只是初值相差0.001（约为初值的0.2%）。这两种情形随时间表现出相似的行为，还是不同的行为？

5. **概率模型**。用于理解概率性结果的概念（例如掷硬币或投掷六面骰子的结果）是最早应用模型进行深入研究的物理—概念系统之一。该主题的历史表明，提高赌博胜算是认真研究这一主题的动机之一。统计学家A. Hald（2003）曾指出

> 直到 16 世纪初，意大利数学家才开始讨论机会游戏中各种结果出现的可能性。他们依据的基本思想是，单场游戏的各种可能结果出现的概率相同。

概率论中的问题可能颇具挑战性，但在这里，我们提出一个简单的系统，用于思考其中涉及的建模概率。具体来说，假设有一个瓮（不知为何，这类问题中经典的容器就是瓮！），里面有三只球：一只黄色、一只红色和一只蓝色。问题是：如果从瓮中一次取一只球，共取三只，那么第一只取出的球是黄色球的概率是多少？

这个问题实际上并非要计算其中涉及的概率（大多数人凭经验无需多想，就能猜出概率是 $1/3$），而是要建立一个解释该概率的模型。关于概率有许多不同的“哲学观点”，但我们在此不作深入讨论。这里提出如下方法：瓮中有三只球，每次取一只。我们假设球出现的顺序很重要。可能的排列只有有限种（例如，蓝、黄、绿是一种；绿、黄、蓝则是另一种，与前一种不同）。如果我们假设三只球的所有结果出现的可能性都相同，那么列举所有可能结果，就能帮助我们回答关于各个结果概率的问题。对于这个问题，请完成或回答以下内容。

a. Enumerate each possible outcome (hint: there are 6 total possible outcomes).
b. Of the total number of outcomes, how many of them represent the cases where the yellow ball is drawn first?
c. What is the ratio of the number of cases where the yellow ball is drawn first to the total number of cases?
d. In this model, what is the primary modeling *assumption* that allows us to actually compute the probabilities?
e. Are there any additional assumptions regarding the process described that should be true so that the analysis is valid?

6. Consider the situation described in problem 5. Now suppose that you want to know the probability of the selecting the balls in the following order: red, blue, yellow. Can you describe the modeling process (i.e., write a short narrative) that explains how you arrived at your result?

7. 考虑问题 5 中描述的情形。现在假设你想知道出现以下结果的概率：红球或黄球先被选中。该结果的概率是多少？你能否描述建模过程（即写一段简短的叙述），说明你是如何得出结果的?

8. 考虑问题5中描述的情形。现在假设你想知道以下结果发生的概率：选出的前两个球分别是红球和黄球，但顺序不限（即先选出黄球再选出红球，或先选出红球再选出黄球）。这一结果对应的概率是多少？你能否描述建模过程（即写一段简短的叙述），解释你是如何得出这一结果的?

9. 考虑问题 5 中描述的情形。现在假设你实际上总共有四个球：黄色、红色、蓝色和绿色。你能否使用问题 5 中描述的方法来确定黄色球最先被选中的概率？虽然这种方法很方便，因为它可以让我们轻松计算任何可能结果的概率，但你能否看出这种方法（即明确列出每一种可能事件）有哪些缺点?

10. 请尽可能简洁、清楚地用文字描述对一个数据集过拟合 的问题。

11. 假设你进行了两次实验，用红外测温法在十个等间隔的时刻测量温度升高值（$T - T_0$）。数据集见表 9。

时间	$(T - T_0)$ 实验 1	$(T - T_0)$ 实验 2
0	-0.110	-0.110
1	1.100	0.832
2	1.835	5.035
3	2.347	6.547
4	2.009	6.809
5	5.771	10.771
6	6.208	11.008
7	11.198	13.765
8	12.010	15.211
9	16.986	17.215
10	20.420	19.200

表 1.2：两次测量温度的实验。

a. 从数据集 1 开始，分别拟合一条直线（$f_1(x) = ax$）和一个二次式（形式为 $f_2(x) = ax^2 + bx$），并计算这两条拟合曲线对应的决定系数（也称为相关系数）r^2。你可以使用大多数电子表格程序或计算机语言来完成。哪条曲线拟合得更好?

b. 现在使用拟合前一个数据集所得的函数（f_1 和 f_2）来拟合实验 2 的数据。同样，计算两条拟合曲线的 r^2 值。哪条曲线拟合得更好？如果你打算根据实验 1 建立一个模型，并将其用于拟合后续实验的数据，根据你目前掌握的信息，你会选择哪一个——f_1 还是 f_2?

12. **信息**。文中提出了信息这一概念。理解模型的信息量的一种方式，是考虑描述它有多困难。为了具体说明这一点，请考虑一个由数字组成的二进制系统。考察以二进制格式表示的 5 位整数。每个这样的数都具有“XXXXX”的形式，其中每个“X”都是“1”或“0”。如果某位是“1”，则将一个单位的 $2^{(n-1)}$ 加起来，其中 n 表示从最右侧开始该数字所在位置的数值。因此，二进制数字“10010”表示十进制数 $2^4 + 2^1 = 17$。这样一组数字所包含的信息，就是数字的个数（这接近于该系统可表示的最大数的对数——例如 $\log_2 11111 = 4.95$，约为 5 个二进制数字或位）。请据此回答以下问题。

a. 二进制数10010包含多少信息？11111呢?

b. 哪个包含的信息更多——二进制数11111还是十进制数31？ c. 假设我发送一条消息，其中包含一个二进制数字，可以是1或0。收到消息的人获得了多少比特的信息？ d. 考虑下面两个由5000个数字组成的二进制数实例。情况(i)：每个数字都通过抛硬币随机生成，并根据结果写成1（正面）或0（反面），直到生成5000个数字。情况(ii)：一个由数字“1”重复5000次组成的数。（提示：能否用某种方式压缩第二个数，用少于5000个数字进行传递，且不丢失任何数字？第一个数呢？） e. 在你上一题的答案中，这两个数的规律性这一概念是否会影响你对它们的处理方式?

13. As briefly mentioned in this chapter, some philosophers view language itself as a model of reality. In a few sentences, describe what this might mean– that language is itself a model. Reference the discussion on the properties of models and modeling, and attempt to describe the process of *abstraction*, *model building*, *realization*, and *revision* as it might apply to a natural language by providing examples.

14. Read the paper titled *What is complexity science, really?* by S.E. Phelan (Phelan, S.E.: What is complexity science, really? Emergence 3(1), 120–136 (2001)). Phelan identifies three philosophical frameworks in which *science* might be defined and interpreted. What are these? Does the paper suggest that any particular framework is best?

15. Read the paper titled *This is not a universe: Metaphor, language, and representation* by Liliane Papin (Papin, L.: This is not a universe: Metaphor, language, and representation. PMLA 107(5), 1253–1265 (1992)). Papin discusses some of the challenges that language poses in science. In your own words, what do you think is the main point of the paper? In other words, what is the message that Papin is attempting to communicate regarding language and science?

“一条糟糕的评论，就像用所有最好的食材烤好一个蛋糕，却有人一屁股坐在上面。”丹妮尔·费尔南德斯·多米尼克·舒莱因-斯蒂尔，有史以来第四畅销的小说作家

2

数学定义、概念与回顾

Almost everyone who uses this text will have some background in mathematics; it is assumed that this background includes advanced algebra, introductory calculus, linear algebra, and an introductory course in ordinary differential equations. The purpose of this chapter is to review a host of definitions, concepts about mathematics, and some of the basic results learned in previous coursework. The material presented here is not meant to be exhaustive, but is focused primarily on elements that will be useful in the remainder of the text.

由于本章的内容针对的是本书后续部分中出现的概念，因此本章更像是对一些重要主题的汇集，而非围绕一两个主要观点展开的阐述。极少数情况下，本章会引入本科数学课程中可能未涉及的内容；不过，这些内容所需的背景知识仅限于本书中所介绍的内容。对于不需要复习数学定义、概念，或线性代数与微积分基础知识的读者，本章中的许多主题都可以放心略过。

2.1 集合与集合构造表示法

虽然我们不会广泛使用集合这一概念，但会使用集合构造表示法，作为讨论实数轴上的区间（或更高维的欧几里得 [空间，有时称为笛卡尔] 2维空间和三维空间）的一种方便、紧支撑的方式。本文采用的版本比更一般情形下可能采用的版本简单。例如，集合构造表示法中有表示“且”和“或”的符号；为了尽量减少新符号，我们选择直接使用这些词。以下是我们将使用的集合构造表示法元素。下文可能会进一步定义其中一些内容（例如“集合”一词）；此处的目的仅仅是介绍这种符号表示。

1. **Set**. A set is denoted using curly braces, “{ }”. Thus $\{a, b, f\}$ is a set containing three items, a, b, and f.

2. **"Such that" symbol**. In set builder notation, the colon ":" is used to indicate the concept of "such that" or "with the property that" or sometimes "as follows". When one sees the colon after a variable (say, the variable x), it might be useful to think of this as indicating "create the set of x values such that..."
3. **"属于"符号**。我们用"$\in$"表示"属于"。稍后我们会更完整地定义元素。
4. **"不属于"符号**。我们用"$\notin$"表示"不属于"。
5. **谓词**。谓词在英国语法中，是句子中用来陈述主语某些情况的部分。在集合论中，"谓词"一词有相同的含义。因此，谓词可以理解为必须应用的一条规则或公式。假设谓词由方程（或不等式）Φ给出，并且要求我们构造满足 Φ 的所有x 值组成的集合。具体来说，假设对应关系为 $\Phi \Leftrightarrow x > 1$。我们可以将这样的集合记为$A$，并写成如下形式

$$A = \{x : x > 0\} \tag{2.1}$$

这读作："构造集合A ，其中A 等于所有满足$x > 0$的数x 组成的集合。"这里隐含着一个前提：我们知道x 应该是哪种数。假设x 指的是整数。如果我们想在集合表示中明确写出这一点，可以将其改写如下。

$$A = \{x : x > 0 \text{ and } x \in \mathbb{Z}\} \tag{2.2}$$

这里， $\mathbb{Z}$ 是用于表示整数集的特殊保留符号。

尽管我们不会经常使用，但可以用 $\forall$ 表示"对于所有"，用 $\exists$ 表示"存在"。例如，大于零的偶整数可以用集合构造表示法描述为

$$\{n : (\exists k) k \in \mathbb{Z} \text{ and } k > 0 \text{ and } n = 2k\} \tag{2.3}$$

这可以读作："按如下方式创建所有n 值组成的集合：存在k 个值，它们是大于零的整数。n 值的集合通过将这些k 值组成的集合乘以二来得到。"至此，我们结束了通过集合构造表示法所采用的概念和符号的介绍。

Now that we have the notation established, we can proceed to define some of the basic features of sets (we will be ultimately be primarily interested in the applications of set notation to intervals on the real line).

Definition 2.1 (Element). An *element* of a set (also called a *member* of a set) is the generic name associated any one of the objects contained by the set. Often the symbol x is used to denote the element of a set (although any symbol could be used).

Definition 2.2 (Set). A *set* is a collection of objects; more specifically for our purposes, a collection of mathematical objects. Sets can contain a finite number of objects, or an infinite number of objects.

Definition 2.3 (subset). A *set* where the elements of of the set are also elements of some given set. To be a *proper subset*, the subset must not contain all of the elements of the given set.

Definition 2.4 (universal set). Many problems are created when attempting to define a completely general *universal set*; in the broadest sense, they can lead to paradoxes in formal set theories. For our purposes (which is primarily to discuss intervals on the real line), we adopt a universal set that is a well-defined entity that avoids such problems. Specifically, we take U to be the real line in one dimension (or the 2-D plane or 3-D space in higher dimensions). One may also *define* U to be some subset I of the real line, as long as all discussions relate to intervals that are proper subsets of I.

Even though our use of set theory will be minimal, it is still useful to define the basic operations of the *union* and *intersection* of sets. These are as follows.

定义 2.5（集合的并集）。两个集合 A 和 B 的并集，是由属于其中任一集合（不重复）的元素构成的集合 C。通常记作 $=A\cup B$。若集合中的元素用 x表示，则可写为 $A\cup B=\{x:\ x\in A$ 或 $x\in B\}$。对于三个或更多集合，可通过每次对两个集合进行运算并迭代完成。

定义 2.6（集合的交集）。两个集合 A 和 B 的交集，是由同时属于这两个集合的元素C 构成的集合。通常记作 $=A\cap B$。也可写为 $A\cap B=\{x:\ x\in A$ 且 $x\in B\}$

请注意，集合的并集和集合的交集分别对应于且和或这两个英语连词。具体来说，语句$D=A\cup B\cup C$ 且 $x\in D$ 意味着x 属于集合A、集合B 或集合C （其中可包含这些集合之间的任意交集）。相比之下，语句$D=A\cap B\cap C$ 且$x\in D$ 意味着x 分别属于A、B和C。

定义 2.7（集合的补集）。假设 $C=A\cup B$。全集 U中C的补集由U 中所有不属于集合C的元素组成。这通常写作 $C'=U\backslash C$ （其中反斜杠读作减法或移除）。也可以写作 $C'=\{x:\ x\in U$ 且$x\notin A$ 且$x\notin B\}$。如果U定义为实数轴（或高维空间）的某个子集，那么有时补集称为相对于 U 的补集。

These concepts are presented graphically in Fig. 2.1. One additional definition is made here, in part because it helps to clarify a notation that is often observed in the mathematical literature.

Definition 2.8 (Cartesian product of sets). The Cartesian product of two sets A and B, is the set of all ordered pairs (a,b) where $a\in A$ and $b\in B$.

$$A\times B=\{(a,b):a\in A\ \text{ and }\ b:\in B\}. \tag{2.4}$$

This latter notation will arise when we discuss the Euclidian plane and Euclidian space.

2.1.1 Numbers

数的概念直观而自然，我们从小就开始用手指数物体，并学习计数。然而，数本身也是一个有趣的话题；甚至有一个数学分支称为数论，专门研究整数。研究整数听起来可能并不特别令人兴奋，但实际上这是一个研究非常深入的领域！例如，数论已被用于寻找非常大的素数（即不能由两个更小的数相乘得到的数）。这类数在公钥加密方案和其他类型的计算机安全应用中具有很大价值。事实上，目前有几项悬赏（约15万美元，截至2021年），用于寻找拥有1000万或1亿位数字的素数。目前已知的最大素数（截至2022年1月）是$282{,}589{,}^{933}-1$，它在十进制下有24,862,048位数字。

It turns out that the discussion of the various of *sets of numbers* one might encounter in applied mathematics is a good way to discuss sets and intervals in a way that has familiarity.

The most frequently used sets of numbers are as follows.

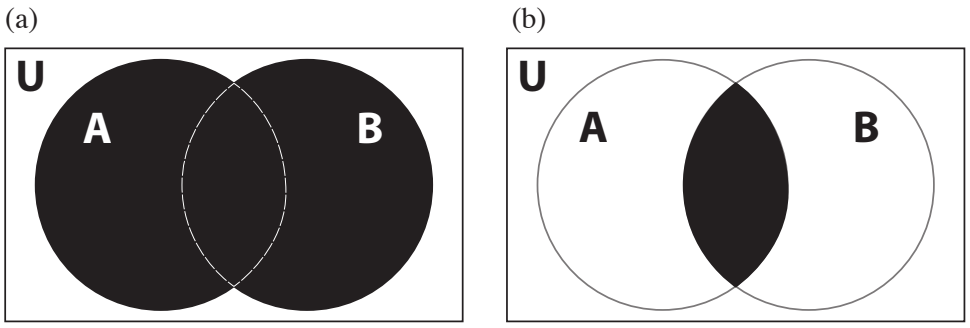


图 2.1：（a）两个集合的并集。（b）两个集合的交集。在这些图中，A 和 B 是集合，并且这两个集合包含于全集U中。因此，A 和 B 是U 的子集。

- $\mathbb{N}$– **“自然”数**。它们是正整数的集合 $\mathbb{N} = \{1,2,3,4,\ldots\}$。某些定义包含零；为表示这种情况，通常使用符号 $\mathbb{N}_0 = \{0,1,2,3,\ldots\}$。

- $\mathbb{Z}$ – **The integers**. These are the set of all *positive or negative* integers, including 0 $\mathbb{Z} = \{\ldots -2,-1,0,1,2,\ldots\}$.

- $\mathbb{Q}$ – **The rational numbers**. The “$\mathbb{Q}$” here stands for *quotient*, or, in other words, the ratio of two numbers. The *rational* numbers are formed by the set of all possible ratios of all integers p and q in $\mathbb{Z}$ so that all pairs for a ratio p/q which is a member of $\mathbb{Q}$. Note that there is a restriction; $q \neq 0$. Also, because we can have $q = 1$, the integers are a subset of the rational numbers.

- $\mathbb{P}$ – **The irrational numbers**. In short, the set of all numbers that cannot be expressed as a rational number, i.e., they are not elements of $\mathbb{Q}$.

- $\mathbb{R}$ – **The real numbers**. The real numbers are the set of all things that we might think of as conventional numbers; in other words, it contains as subsets all of the sets of numbers defined previously. For our purposes, the real numbers can be though of most simply as the *union* (or the combining of) the rational and the irrational numbers (in mathematics this might be written $\mathbb{R} = \mathbb{Q} \cup \mathbb{P}$). The real numbers are what we use *in principle* when doing computations relating to physical systems. The real line can also be though of as the 1-dimensional Euclidian space, where the Euclidian spaces are defined in the next section.

- $\mathbb{C}$ – **复数**。复数是数这一概念的推广，包含实数集 部分和虚数 部分。我们大多数人在数学学习生涯的早期就接触过复数这一概念；然而，它们常常引发强烈的不安与困惑。不过，与许多数学构造一样，复数的“现实性”并不十分重要。重要的是，它们满足某些代数上的必要条件，从而保证数学上的一致性（它们构成一个交换 代数）；而且从实用角度来看，它们是有用的。

图 2.2 展示了这些数系（作为集合）之间的关系。请注意，对于这些数集中的每一个，值 $\pm\infty$ 都不属于该数集！有些数集（称为扩展实数集）包含值 ∞，但本文不会用到这些集合。这些集合的一种伸展是更高维欧几里得空间的集合；下文将对此进行讨论。

这些数集本身具有有趣的性质和历史。例如，无理数最早由古希腊人发现。特别是，人们认为，一位名叫梅塔蓬图姆的希帕索斯（Hippasus）的毕达哥拉斯学派成员（即追随毕达哥拉斯学说的希腊数学团体成员）最先证明了这一点。奇怪的是，虽然正是毕达哥拉斯定理使希帕索斯得以证明无理数的存在，但这一发现却与毕达哥拉斯学派认为所有数学都可以通过整数之比来表示的信念完全相矛盾。据说，希帕索斯是在海上发现这一惊人事实的，而他的毕达哥拉斯学派同伴对这一发现感到极其愤怒，竟然将他扔进了海里！尽管这个故事的真实性确实值得怀疑，但可以肯定的是，人们对数的认识所发生的变革，往往都会遭到相当大的抵制。

虽然许多人认为复数（将在下文复习）对具体思维方式提出了挑战，但即使是实数，也比人们预想的更加奇特，在哲学层面上也更具挑战性。现代数学中一些最具争议的问题，源于试图理解实数集合，而其中许多挑战又通过无理数子集产生。特别是集合的极限和完备这两个关键概念——我们稍后将（简要地）涉及这些主题——直到19世纪和20世纪才得到充分理解（详见 Snow（2003））。在转向其他主题之前，下面的例子说明了实数概念比表面上看起来更加复杂。

Example 2.1 ($0.999\ldots = 1$*).* We have all be exposed the idea of repeating decimal numbers, but the characteristics of such numbers can be elusive. For example, consider the sequence of numbers (we will define sequences more formally in §2.5) that approach the value 1 as follows $s = \{0.9, 0.99, 0.999, 0.9999, \ldots\}$, or more generally $s = \{0.(9)_n, n \in \mathbb{N}\}$. Clearly this sequence gets closer and closer to 1 as we add more repeats of the numeral 9 (i.e., increase n). Also, if we fix any small number, ε, then, no matter how small epsilon is, we can always take n large enough so that we get closer than within ε to the value 1

$$1 - 0.(9)_n < \varepsilon \tag{2.5}$$

In some sense, then, is $0.(9)_n$ equal to 1? This seems like it would be a curious thing, but we can show that it is true. The following is an informal illustration. There are proper, concrete proofs for the somewhat *ad hoc* illustration below. Nonetheless, the point is made, and the result is correct, as odd as the result may seem!

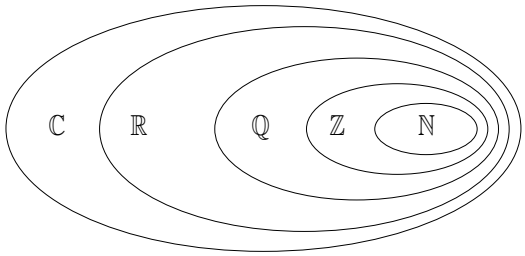


图 2.2 从集合的角度看数系。在这幅图中，每个椭圆代表一个集合；包含在另一个较大的椭圆中的椭圆是包含该集合的子集。

$$\begin{aligned} x &= 0.9999\ldots \\ 10x &= 9.999\ldots \\ 10x &= 9 + 0.9999\ldots \\ 10x &= 9 + x \\ 9x &= 9 \\ x &= 1 \end{aligned} \tag{2.6}$$

The explanation for this involves the idea of convergence of a sequence and completeness of a set to fully understand. For now, we will put this notion aside and simply recognize that real numbers (particular those that are irrational or are repeating decimals) are trickier to understand than rational numbers are! There is even some more philosophical discussion about in what sense real numbers are (objectively) real. Because of the various concepts of infinity associated with the real numbers, they sometimes illustrate behavior that is counter intuitive (the fact that $0 \times \infty$ is not defined is one such example; it is easy to construct examples where $0 \times \infty$ is equal to any number that you choose).

2.1.2 欧几里得空间

欧几里得空间或许可以用较为通俗的方式描述为我们都熟悉的三维空间。在继续之前，为了理解欧几里得的定义，我们必须先了解什么是公理。

定义 2.9（公理）。公理是关于物理系统或数学系统的一种陈述，人们假定它为真，但无法在该系统内部证明它为真。有时，公理也称为定律、原则或公设；这些名称的不同用法源于历史原因或学科术语差异（例如物理学与数学之间的差异）。

现在我们可以定义欧几里得空间了。欧几里得空间是我们在普通几何中习惯思考的那种空间。欧几里得空间遵循欧几里得《几何原本》中的五条公理，具体如下。

1. 任意两点之间都可以画一条直线。

2. Any terminated straight line may be extended indefinitely.
3. A circle may be drawn with any given point as center and any given radius.
4. All right angles are equal.
5. For any given point not on a given line, there is exactly one line through the point that does not meet the given line.

where the language here is not identical to that of Euclid, but the concepts are (in particular, this version of axiom 5 is known as Playfair's axiom (Playfair, 1795, p. 7)). So, why invoke Euclid at this juncture? It turns out that most of the mathematics you have done have been done in Euclidan space. In one dimension, Euclidian space is the real line; in two dimensions, the plane; and, in three dimensions what we often use as our model of space. Using set notation, in 1-, 2-, and 3-dimensions, we have

$$\text{1-dimension} \qquad \mathbb{R} = \mathbb{R}^1 = \{x : x \in \mathbb{R}\} \tag{2.7}$$

$$\text{2-dimensions} \qquad \mathbb{R} \times \mathbb{R} = \mathbb{R}^2 = \{(x,y) : x \in \mathbb{R} \text{ and } y \in \mathbb{R}\} \tag{2.8}$$

$$\text{3-dimensions} \qquad \mathbb{R} \times \mathbb{R} \times \mathbb{R} = \mathbb{R}^3 = \{(x,y,z) : x \in \mathbb{R} \text{ and } y \in \mathbb{R} \text{ and } x \in \mathbb{R}\} \tag{2.9}$$

Here, the Cartesian product defined above has been used in defining $\mathbb{R}^2$ and $\mathbb{R}^3$.

有时，欧几里得几何也称为平坦空间，也就是说，它没有曲率。物理学家（尤其是阿尔伯特·爱因斯坦）已经证明，空间实际上并非平坦，而是会因引力作用而弯曲。尽管如此，对于大多数地面上的系统（即规模足够小的系统），只要不需要极其精确的测量值，我们就可以将空间视为欧几里得空间。这些空间的共同之处在于，我们对几何的传统认识（平行线不相交；所有直角的角度度量都相等）在这一模型中依然成立。

需要注意的是，坐标系独立于其所描述的底层空间。我们用来描述欧几里得空间的传统坐标系是笛卡尔坐标系，它以最早对其进行描述的法国哲学家兼数学家勒内·笛卡尔命名。事实上，欧几里得空间有时也称为笛卡尔空间。不过，需要谨慎区分。笛卡尔坐标系就是我们熟悉的直角坐标系（图 2.3(a)）。坐标系可能十分方便，但不应将它与欧几里得空间本身混为一谈。坐标系是一种便利工具，可为每个点赋予唯一的标签；但坐标系的几何结构不一定反映底层空间的几何结构。例如，我们也可以采用熟悉的柱坐标系（图 2.3(b)）来描述欧几里得空间。比如，在对欧几里得空间中一个本身呈圆柱形的物体进行建模时，这样做会很方便。尽管这两种坐标系不同，底层空间（以及假定在其中成立的几何原则）仍保持不变。

虽然这并不是本文材料中的重要概念，但底层空间本身也可能不是平坦的（即它不是欧几里得空间）。例如，球形空间是你可能通过地球仪类比而听说过的一种空间；在这样的空间中，必须修改欧几里得几何的五条公理，才能建立一个一致的系统。非欧几里得空间（例如刚才提到的球形空间，或双曲型空间）的发现，是19世纪初由欧洲和俄罗斯的数学家完成的。这些独特几何空间的发现最终传播到了大众文化中，并真正改变了数学家、科学家和广大公众看待科学与科学发现的方式。作为其对大众文化影响的一个例子，非欧几里得几何的思想广为人知，以至于作家 H. P. Lovecraft 将这些独特几何用于描述异世界的

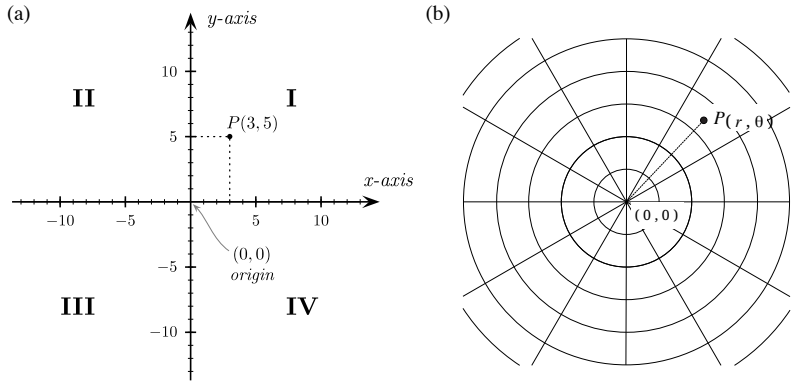


图 2.3：二维（2-D）欧几里得空间的两种坐标系。(a) 笛卡尔坐标系（CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=603366。(b) 柱坐标系。无论使用坐标系 对欧几里得空间中的点进行标记，空间本身都保持不变。因此，欧几里得几何五公理适用于这两种坐标系。

他虚构作品中的设定。与集合论的发展（将在下一节末尾简要讨论）一样，非欧几何的发现也助推了数学革命。这场革命始于19世纪初，一直持续到20世纪中期。

2.1.3 实数线上的区间

区间 的概念可以用集合的语言来描述，而区间这一概念也是我们会频繁使用的。下面将讨论开区间、闭区间、有限区间和无界区间的概念。类似的概念也适用于高维情形，但此处暂不讨论这些推广。同样，我们会讨论集合记号和集合论（除了介绍集合存在且具有元素这一基本概念，以及并集和交集这两种基本运算外，不作进一步探讨）。毫不夸张地说，集合论本身构成了现代数学的基础。它也是数学中最难、争议最大的领域之一，其中的争论（或方法与观点上的分歧）至今仍在不断引发讨论和研究。我们不会深入探讨集合论本身，也不会详述它为何会成为数学中如此棘手的组成部分，但会在本节末简要介绍造成部分困难的缘由。

定义 2.10（区间）。一个区间是一个集合，包含两个指定实数a 和b之间的所有实数；假设$a < b$。区间可以是开区间、闭区间，或开闭混合区间。闭区间包括端点，例如，

$$I = \{x : a \leq x \leq b\} \text{ usually denoted } I = [a, b]$$

而开区间则不包含端点

$$I = \{x : a < x < b\} \text{ usually denoted } I = (a,b)$$

关于像 (*a*,*b*) 这样的开区间，有一点颇为有趣：它们有最大或最小实数。根据定义，开区间不包含端点 *a* 和 *b*。例如，开区间 (1,2) 包含所有大于 1 且小于 2 的数，但不包含 1 或 2。因此，开区间还可以这样描述。暂且只考虑由*a*指定的区间最小值部分。现在，假设我们在区间内选取一个尽可能靠近*a*的数；称这个数为*c*。根据定义，$c - a > 0$。接着，我们可以选取一个更小的数，它位于*a*和*c*之间的中点，并称其为$d = (c + a)/2$。这个数也在区间内，因为$d - a = \frac{1}{2}(c - a) > 0$。这个论证可以无限次重复，因此开区间中不存在最小的数。对于开区间中的最大数，也可以作类似论证。

区间也可以采用显而易见的半开（或半闭，两者含义相同）形式，例如，$I = \{x : a < x \leq b\}$ 是半开区间，因为不包含下界。.

Finally, note that intervals can be *unbounded*. An unbounded interval is defined as follows.

Definition 2.11 (Unbounded interval). An *unbounded interval* is a set containing all of the real numbers greater than or less than some specified real number, *a*. Such intervals can be open or closed; thus $I = [a\ \infty)$ is considered unbounded and closed, whereas $I = (a\ \infty)$ is unbounded and open.

考虑到上述为区间定义的各种可能情况（并定义空区间和退化区间），实数轴上的区间可以分为以下十一种不同类型（参见 Craig（1969，第3章））。

空：

$$[b,a] = (b,a) = [b,a) = (b,a] = (a,a) = [a,a) = (a,a] = \{\} = \varnothing$$

退化的：

$$[a,a] = \{a\}$$

Proper and bounded:

$$\begin{aligned}
\text{Open:} &\quad (a,b) = \{x : a < x < b\} \\
\text{Closed:} &\quad [a,b] = \{x : a \leq x \leq b\} \\
\text{Left-closed, right-open:} &\quad [a,b) = \{x : a \leq x < b\} \\
\text{Left-open, right-closed:} &\quad (a,b] = \{x : a < x \leq b\}
\end{aligned}$$

Left-bounded and right-unbounded:

$$\begin{aligned}
\text{Left-open:} &\quad (a,+\infty) = \{x : x > a\} \\
\text{Left-closed:} &\quad [a,+\infty) = \{x : x \geq a\}
\end{aligned}$$

Left-unbounded and right-bounded:

$$\begin{aligned} \text{Right-open:} \quad & (-\infty, b) = \{x : x < b\} \\ \text{Right-closed:} \quad & (-\infty, b] = \{x : x \leq b\} \end{aligned}$$

两端无界（开区间和闭区间）：

$$(-\infty, +\infty) = \mathbb{R}$$

最后这个区间（实数轴）既被列为开集，也被列为闭集。这是一个技术细节，却在网上的数学论坛上引发了无休止的讨论。若不引入额外的数学结构（例如拓扑学），很难解释清楚。不过，简而言之，可以这样描述：实数轴包含所有可能的实数，因此相对于实数而言，它是闭的。然而，实数轴上不存在最大实数或最小实数，因此（如同对开区间的讨论）实数轴也是开的。虽然这听起来像个悖论，但它也说明，数学中并非总需要二元选项（“真”或“假”）；有时还存在一个第三选项，即“既非真，也非假”。数学中有句俗语：“门必定是开着或关着的，不可能既开又关；而一个集合可以是开的、闭的、既开又闭，或两者皆非！”（Munkres，2014年，第91页）。这听起来或许有些模棱两可，但事实上，这些概念都有明确的定义，也很有用。

关于集合论的简短说明。我们在上文使用了集合记号，但并未真正涉及集合论本身。你可能早在小学时就学过一些集合论知识——那个阶段通常会讨论如何将对象归为一组（形成一个集合），然后比较不同集合的性质。表面看来，这似乎简单得令人误以为没有难度。然而，建立一套健全的集合论的目标，却引发了数学史上最有趣的数学研究时期之一。此处的健全有特定含义：这意味着该理论必须提出足够的公理，使其既完备又一致。在此语境中，完备一词基本上是指“系统中的每个真命题都能由公理证明”。一致一词基本上是指“系统中的每个恰当命题（或者说问题）都能被证明要么为真要么为假”。人们希望，由于集合论如此基础，那么至少原则上，一个足够强大的集合论版本就能成为推导全部数学的基础（至少原则上如此）。然而，出乎意料的是，这样的集合论最终被证明是不可能存在的。

集合论的早期研究由许多数学家完成；但尤其是两位德国数学家——Richard Dedekind 和 George Cantor——在约 1870 年至 1900 年间的工作，为后来有时被称为朴素集合论的理论铺平了道路。关于这段历史，一个简明扼要的介绍可见于（Ferreir´os，2022）。这里的修饰词“朴素”并非贬义；它仅仅意味着该理论并未以形式逻辑的语言表述。尽管这种集合论极其有用，Cantor 还是惊讶地发现它并不一致——换言之，他发现可以在该理论中构造出悖论。在此不再赘述集合论的全部历史；由于无法建立无悖论的集合论，20 世纪初的数学陷入了一场危机。这场危机虽然酝酿已久，但 Bertrand Russel 在 1901 年对该悖论的描述，使其表现得最为明显。罗素悖论是一种数学陈述，与所谓的说谎者悖论非常相似，后者提出了“这句话不是真的”这一相互矛盾且自指的陈述。

这种悖论会对数学产生深远而持久的影响。20世纪20年代，数学家恩斯特·策梅洛和阿伯拉罕·弗兰克尔最终在一定程度上修正了集合论（在数学中称为ZFC集合论）。不过，这里必须对“一定程度上”这一说法加以限定。ZFC避免了早期版本的集合论所遭受的那类悖论。然而，在20世纪30年代，一位名叫库尔特·戈¨德尔的数学家证明了一个相当出人意料的结果。戈¨德尔大致证明了，.....

后续：任何复杂到足以描述普通整数算术的公理系统，要么不完备，要么不一致。换句话说，如果我们的公理是一致的（即不会导致悖论或不可判定陈述），那么在这些公理的每个模型中，都存在一个为真但不可证明的陈述。虽然这一结果在某种程度上打击了建立一套涵盖全部数学的总体理论这一想法，但其实际影响仍存在争议。首先，G¨odel所使用的证明属于一种称为一阶逻辑的特定类型。因此，更强大的逻辑语言是否也会遭遇同样的命运，尚不清楚。其次，尽管首要的集合论ZFC根据G¨odel的定理必然是不一致的，或者包含为真但不可证明的陈述，但尚不清楚这是否具有任何实际 意义。例如，如果我陈述“所有非翼龙类都不是非恐龙”，我当然说出了某个真命题。无论如何，如果我对恐龙（无论会飞与否）的任何事情都不感兴趣，那么这句话即使为真，对我也没有价值。因此，尽管在ZFC中证明所有为真 的陈述或许是不可能的，但到目前为止，这些无法触及的真理尚未造成任何实际问题（可能是因为，正如这个例子一样，它们传达的真理与我们无关）。而且，迄今为止，没有人发现ZFC框架内存在不一致的陈述。因此，ZFC一直是数学中的有用工具。对更强大集合论的探索仍是数学研究中一个受到关注的课题。

2.1.4 Complex Numbers

Unlike, say, the natural numbers (which can be illustrated by collecting actual objects), the complex numbers are a purely mathematical construct. That does not mean that they are not useful or interesting, however. There are many examples of concepts that exist only mathematically, but are nonetheless useful for many practical applications.

The complex numbers are an extension of the real numbers. Although almost everyone reading this text has probably encountered them previously, it is useful to recap their basic properties.

Definition 2.12. A *complex number* assumes that there exists a mathematical object, called the unit imaginary number, i, such that $i = \sqrt{-1}$ so that $i^2 = -1$. Every complex number consists of two parts, a *real* part, and an *imaginary* part that is proportional to i. The conventional form for a complex number is $a + bi$, where a and b are real numbers. The set of all complex numbers is usually denoted by $\mathbb{C}$.

The rules for addition, subtraction, and multiplication are slightly modified from those of real numbers as follows. First, addition and subtraction are done by adding and/or subtracting the real and imaginary parts of a complex number independently. Therefore

$$(a+ib)+(c+id) = (a+c)+i(b+d)$$
$$(a+ib)-(c+id) = (a-c)+i(b-d)$$

Multiplication of two complex numbers is defined as follows

$$(a+ib)\times(c+id) = ac+ibc+iadi-bd$$

or, equivalently, grouping terms

$$(a+ib)\times(c+id) = (ac-bd)+i(bc+ad)$$

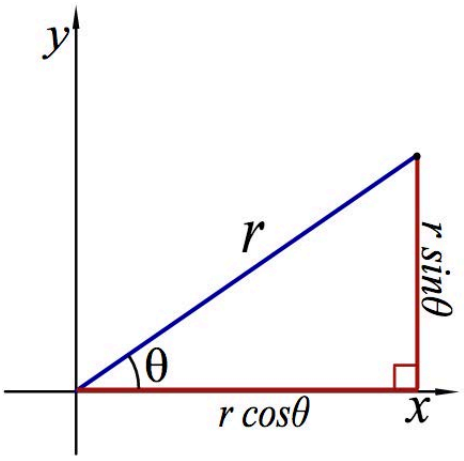


Fig. 2.4 The complex plane. Complex numbers can be given a meaningful interpretation as vectors on a plane.

The complex numbers are often represented as vectors on a plane, where the real part is plotted on the horizontal axis, and the complex part on the vertical axis. In this representation, a complex number would be represented by a pair of points, i.e., $x+iy \Leftrightarrow (x,iy)$. There are some advantages and disadvantages of this, but it generally improves interpretation for problems with physical significance. When this formalism is adopted, a complex number is frequently represented typographically as a vector, as in $\mathbf{z} = (x,iy)$. The representation of the complex numbers in this manner requires mildly re-defining the vector dot product. We define this in the following.

定义 2.13（复共轭）。对于复数 $\mathbf{z} = a+ib = (a,ib)$，其 复共轭定义为 $\bar{\mathbf{z}} = a-ib = (a\ -ib)$。

定义 2.14（复数点积）。对于两个向量 $\mathbf{z}_1 = (a_1,ib_1)$ 和 $\mathbf{z}_2 = (a_2,ib_2)$，$\mathbf{z}_1$ 与 $\mathbf{z}_2$ 之间的复数点积定义为

$$\mathbf{z}_1 \cdot \bar{\mathbf{z}}_2 = (a_1a_2 + b_1b_2)$$

which has the advantage of being a single real number. In particular, this means that a vector dotted with itself is defined by

$$\mathbf{z} \cdot \bar{\mathbf{z}} = (a,ib) \cdot (a,-ib) = a^2 + b^2$$

and the magnitude of a complex vector is defined by a rule that looks much like the rule for computing the length of a vector on a plane

$$\|\mathbf{z}\| = \sqrt{\mathbf{z} \cdot \bar{\mathbf{z}}} = \sqrt{a^2+b^2}$$

The recognition that complex numbers can be treated as vectors on a plane implies that they can be conveniently represented in polar coordinates. In fact, there are some significant reasons for doing so. To start, define the length of the vector by the number $r = \|\mathbf{z}\| = \sqrt{a^2+b^2}$. Then, we have the following relationships between the (a,b) and $(r\ \theta)$ coordinate systems (Fig. 2.4)

$$r = \sqrt{a^2 + b^2}$$

$$a = r\cos(\theta) \qquad ib = ir\sin(\theta) \tag{2.10}$$

$$\theta = \tan^{-1}\left(\frac{b}{a}\right) \tag{2.11}$$

or,

$$\mathbf{z} = r(\cos(\theta), i\sin(\theta)) = r(\cos(\theta) + i\sin(\theta)) \tag{2.12}$$

In this representation, r is sometimes called the *magnitude* or *modulus* or *absolute value* of the complex number, and θ is called the *argument* of the complex number.

你可能以前见过欧拉公式；它导出了著名的关系$e^{i\pi} = -1$（回想一下，e 是自然对数的底）。不过，更一般地说，欧拉公式的形式为

$$e^{i\theta} = \cos\theta + i\sin\theta \tag{2.13}$$

尽管这个公式看起来确实非同寻常，但一旦我们接受虚数是实数的一种合理伸展，这个结果的证明就会变得相当简单（我们在研究无穷级数时会进一步探讨这一点）。目前，我们先接受这个公式，不作证明。不过请注意，它可以将复数$\mathbf{z}$ 以一种特别简洁的形式表示，如式（2.12）所示：

$$\mathbf{z} = re^{i\theta} \tag{2.14}$$

有趣的是，最后这个公式使复数的对数有了意义。对式 2.14 两边取自然对数，得

$$ln(\mathbf{z}) = ln(re^{i\theta}) \tag{2.15}$$

$$= ln(r) + i\theta \tag{2.16}$$

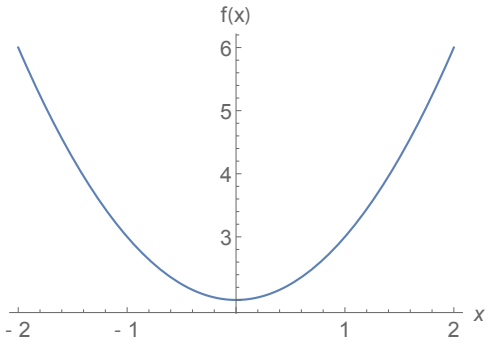
注意到$e^{i\theta} = e^{i(\theta+2k\pi)}$，$k = 0$，1，2，.....，显然，复数的对数通常有无限多个表示

$$ln(\mathbf{z}) = ln(r) + i(\theta + 2k\pi),\ k = 0, 1, 2, \ldots \tag{2.17}$$

Although complex numbers have a helpful representation as vector quantities on the complex plan, they are technically just an extension of the real number system. Therefore, it is not common to adopt a bold-face type to represent them (as we have done above). In general, complex numbers are set in regular, italicized script, (e.g., z). Generally, the context prevents there from being any confusion. In future uses of complex numbers, we will not use bold-faced script to represent them. Thus, the equation above would be more properly written

$$z = re^{i\theta} \tag{2.18}$$

Fig. 2.5 A function.



2.2 Functions

We all have been introduced to the concept of functions. When most people in college mathematics think of a function, the first thing that comes to mind is a relationship that looks something like

$$f(x) = x^2 + 2, \qquad x \in [-2, 2] \tag{2.19}$$

In this description, we are given a *domain* (the numbers such that $-2 \leq x \leq 2$- recall, square brackets indicate *inclusive* interval notation), and a *range* for the function (the values $f(x)$ for the defined domain). A graph of this function is plotted in Fig. 2.5. Although this is how we typically think of a function, we can define it more generally.

Definition 2.15. A *function* (or sometimes *mapping*) is a relationship between two sets, A and B, such that each element of A is *uniquely* associated with an element of B.

The word *uniquely* appearing in our definition here is really important: a function is always single-valued.

This definition requires a bit of additional explanation for our purposes. First of all, for us, the sets involved for a function are almost always intervals of the real number line. We even give these functions additional names and descriptions, as follows. We usually refer to functions as having *independent* and *dependent* variables. The independent variables are the *inputs* to the function (or more properly, the set that is input to the function; this is the set A in the definition). Conversely, the *dependent* variables are the set of values that are produced by the set of independent variables upon application of the function (this is the set B in the definition). Each member of the set of independent variables is mapped to *exactly one* member of the set of independent variables (this is the *uniquely* part of the definition). Fig. 2.6 gives a pictorial representation of the process.

The concept of functions that strictly increase or decrease is one that is used routinely in mathematics; so much so that they have a special name: *monotonic*. These functions are defined as follows.

Definition 2.16. A function $f(x)$ is said to be *monotonically increasing* on and interval, I, if the following are true:

1. $f(x_1) \leq f(x_2)$ for all $x_1 < x_2$ in I.
2. $f(x)$ is not the constant function ($f(x) = C$, C a constant).

类似地，如果满足以下条件，则称函数$f(x)$单调递减：

1. 对于I中的所有$x_1 < x_2$，$f(x_1) \geq f(x_2)$。

2. $f(x)$ is not the constant function.

显然，一般而言，函数既不是单调递增的，也不是单调递减的，而是在其整个定义域的子域上结合了这两个概念。常函数在其定义域的任何子域上都既不是单调递增的，也不是单调递减的。

There is one final class of functions that requires definition because the term arises so frequently in applied mathematics. These are the *algebraically homogeneous*, or simply *homogeneous*, functions. Note that later on, the concept of homogeneous differential equations will be discussed; while the concepts are related, one should not confuse the two. Here, we define algebraically homogeneous functions as follows:

Definition 2.17 ((Algebraically) homogeneous function). An algebraically homogeneous function is one with multiplicative scaling behaviour: if all its arguments are multiplied by some factor, α, then the effect on the function is that its value is multiplied by some power of the factor α. In other words, homogeneous functions display the following behavior

$$f(\alpha x) = \alpha^n f(x)$$

Here, n is a real number, and it called the degree of the function.

Example 2.2 (Homogenous functions.). Below are a few examples of homogeneous functions.

1. All polynomials of the form $f(x) = x^n$ are homogeneous. To see this, just note the following: $f(\alpha x) = (\alpha x)^n = \alpha^n x^2$. In this case, the degree of the function is equal to the degree of the polynomial.

2. 含有加性常数的函数都不是齐次函数。这一点也很容易证明。假设$f(\alpha x) = \alpha^n f(x)$ 是一个齐次函数。现在考虑函数 $g(x) = f(x) + c$。函数 $g(x)$ 不可能是齐次函数，因为乘法标量 α 无法缩放常数c。像$g(x)$ 这样的函数（即含有加性常数的函数）称为仿射函数。这仅仅意味着，这两个函数通过对坐标（自变量）进行线性缩放，再进行平移（加法标量）而相互关联。

3. The function $f(x) = 1/x^2$ is homogeneous. To see this, note

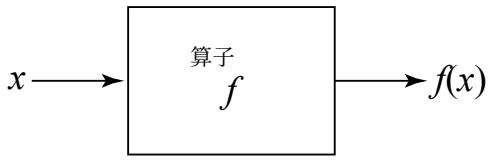


Fig. 2.6 An function, in conceptual form. A function can also be considered a type of *operator*.

$$
\begin{aligned}
f(\alpha x) &= \frac{1}{(\alpha x)^2} \\
&= \frac{1}{\alpha^2}\frac{1}{x^2} \\
&= \alpha^{-2}x^2
\end{aligned}
$$

因此，这个函数是齐次的，次数等于 -2。

2.2.1 有界性

在科学与工程中，我们通常将涉及的集合视为实数轴上的区间。为说明这一点，请看上文公式 (2.19) 给出的示例。在这里，我们可以用集合构造表示法将输入 集A定义为$A = \{x:\ x \in [-2,2]\}$（复习一下，其含义是：“集合 A 是由所有满足x 介于 -2 和 2 之间（包含端点）的数构成的”）。我们可以将集合B视为函数f所定义的所有数。回想一下，我们有$f(x) = x^2 + 2$。输出集B定义为$B = \{f:\ f(x) = x^2 + 2，\text{对所有} x \in A\}$均成立。还要注意，这些区间要么是闭区间，要么是开 区间。闭区间和闭 区间（或一般而言，集合）的概念实际上比人们想象的更为深奥复杂（详见 §2.1.3）。现在，我们将自变量的区间 与这些区间上的函数 结合起来。下面是一个示例。

例 *2.3*（作为区间的函数定义域）。函数

$$f(x) = \frac{\sin(x)}{x},\ x \in (0,\pi] \tag{2.20}$$

定义域为 $D = (0,\ \pi]$。该函数在此定义域上单调递减。虽然这个函- 数看起来并没有什么特别之处，但它在 $x = 0$ 处确实有一些奇怪的行为。下图给出了该函数的图像。

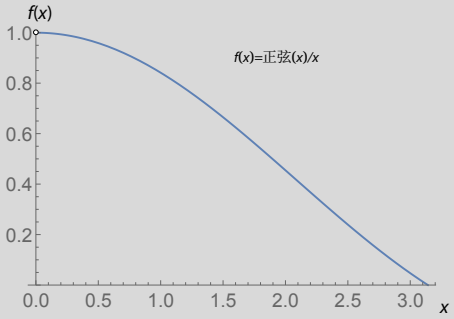


Fig. 2.7: The function $f(x) = (\sin x)/x$.

该函数在 0 与 π 之间的所有点处都有明确的定义（即 $D = (0, \pi]$），但在未定义 于$x = 0$处。这类描述中会出现一个难点概念，在提出以下问题时就会遇到：构成定义域的集合的最小值是多少？严格来说，这个函数的定义域没有最小值！数 0 不在定义域内（因此，定义域描述中使用了“(”）。然而，对于任意小的数 ε，例如 $\varepsilon = 1 \times 10^{-100000000}$，函数值 (正弦$x$)/$x$ 非常接近 1。它只在恰好 $x = 0$时未定义。在这种情况下，由于严格来说不能使用最大值 或最小值，我们会改为说数 0 是定义域的最大下界 或下确界 。这里需要这样理解“最大”：在所有使 $f(x)$有定义的区间的下界中，$x = 0$ 是其中最大的一个。任何更大的值都会落在定义域内，因此不属于下界。

上面的例子表明，通常无法确定一个开集合中的最小数。不过，我们往往可以确定构成该集合下界的最大数（在上面的例子中，这个数是 0）。由此得到以下定理（我们不会证明它！），我们主要将其作为定义使用。

定理 2.1（最小上界，最大下界）。如果一个区间有下界，那么它也有一个最大下界（*g.l.b.*，或下确界）。也就是说，存在唯一的数 z_-，它是该区间的下界，且没有比它更大的数是下界。任何大于 z_- 的数要么 *(i)* 在该区间内，要么 *(ii)* 大于或等于该区间的某个上界。

关于上界也可以作出类似的陈述。如果一个区间有上界，那么它也有一个最小上界（*l.u.b.*，或上确界）。最小上界是唯一的数 z_+，它是该区间的上界，且没有比它更小的数是上界。

为了让这一概念更加具体，可以考虑我们上面讨论过的区间：$(0, 1]$。对于这个区间，最大下界是数 0，尽管数 *0* 本身并不属于该区间。最小上界是数 1。显然，没有任何小于 1 的数，z_0，可以作为该区间的上界，因为$z_0 < 1$ 而根据定义，1 属于该区间。

我们在应用中遇到的许多函数都是映射，它们永远不会趋近于无穷大的值。这在一定程度上是因为物理系统实际上不会出现具有无限模的现象。然而，在许多情况下，将某种现象建模为趋向无穷大值的情况仍然可能很有用。此外，某些纯数学概念要求定义趋向无穷大的函数（例如，正切函数，$\tan(x) = \sin(x)/\cos(x)$，随着 $\cos(x)$ 趋近于零，必然趋向于无穷大）。在数学术语体系中，确立函数有界（以及相应地，无界）的概念是很有用的。

Definition 2.18. A function, f, is *bounded* on some interval $I = [a, b]$ if there is some number, M such that $|f(x)| \leq M$ for all $x \in I$.

任何不有界的函数都是无界函数。有界函数和无界函数的示例见图 2.8。在该图中，(a) 部分包含一个有界函数。图中的灰色虚线表示该函数的绝对值。位于$x = 10$ 处的水平线表示界定该函数的 M 的最小值。该函数绝对值的所有值都小于或等于M，因此该函数是有界的。注意，通常没有必要像这里这样确定M的最小值，以证明函数的有界性。即使选择$M = 12$，仍然可以证明该函数的所有值都小于$M = 12$，因此该函数仍然是有界的。在图 2.8(b) 中，一个

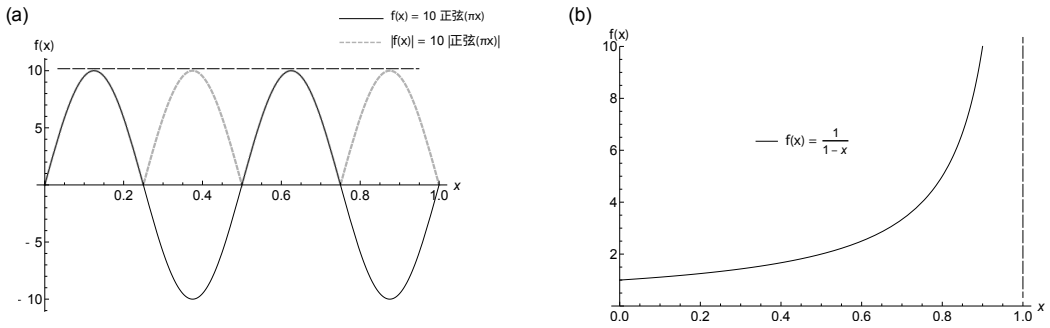


图 2.8：（a）函数 $f(x) = \sin(4\pi x)$ 在实数轴上所有可能的区间上都有界。（b）函数 $f(x) = 1/(1-x)$ 在区间 $x \in [0,1)$ 上无界，当 $x \to 1$。

函数$f(x) = 1/(1-x)$ $0 \le x < 1$ 函数被绘出。当$x \to 1$时，该函数在此区间上是无界的。对于M的任意取值，我们总能找到一个充分接近 1 的x 值，使得$f(x)$ 的值大于M。因此，尽管定义区间不包含 1 这个值（在这种情况下，人们会发现由 $1/0$ 定义的无穷，这有时称为奇点），该函数仍然是无界的！注意，虽然我们通常认为像$f(x) = 1/(1-x)$ 这样的无界函数在奇点处取无穷值，但更恰当地说，该函数在$x = 1$处是未定义的。原因如下：在$x = 1$这一点之前，我们可以设想一个定义良好的极限行为。换句话说，正确的说法是

$$\lim_{x \to 1} \frac{1}{1-x} = \infty \tag{2.21}$$

因为它表示形如 $1/(1-x)$ 的数的极限，其中 x 是小于 1 的任意实数。比值 $1/(1-x)$ 总是有定义的，因为在所定义的区间内，x 永远不会等于 1。这与令$x = 1$形成对比；在这种情况下，函数的值为 $f(x = 1) = 1/0$。形式 $1/0$ 并非有时（错误地）表示的“无穷”。严格来说，表达式 $1/0$ 根本没有定义；它是一种未定义运算。虽然最后这段讨论或许有些过于技术性，但在应用数学中，有不少实例表明，这个看似技术性的要点对于正确理解问题至关重要。

2.2.2 Continuity

Functions are not necessarily always continuous. For example, consider the following function. Assuming $x \in [-2,2]$, define the function

$$f(x) = \begin{cases} 5 & x = 1 \\ x^2 + 2, & \text{otherwise} \end{cases} \tag{2.22}$$

This function is plotted in Fig. 2.9. Although it is true that this function requires some extra handling (for the point $x = 1$), the end result is not too dissimilar to what we are used to seeing. However, it does indicate that it might be useful to further characterize functions on the basis of how continuous and smooth they are.

A few comments about the various labels that are applied to functions to describe how smooth they are are provided in the following.

函数 $f(x)$ 如果随着自变量 x 的变化而逐渐变化，则称其为连续 函数。对于一元函数，可以将其理解为：无需将笔尖离开纸面，就能用笔在纸上画出该函数。更正式地说，使用令人畏惧的 δ-ε 辐角，可以给出更具体的表述。不同于大多数 δ-ε 辐角，我们将分步骤缓慢进行。首先，假设我们关注函数在点 $x = a$处的连续性。现在，我们考察 a周围距离为 $\pm\delta$以内的所有点。我们可以将这些点表示为满足 $a - \delta < x < a + \delta$的所有 x值，或者等价地$|x - a| < \delta$。现在，对于我们在x–轴（定义域）上找到的所有这些点，我们可以计算绝对值 $|f(x) - f(a)|$。假设我们这样做，并发现存在一个数 $\varepsilon > 0$使得

$$|f(x) - f(a)| < \varepsilon \text{ for all } x \text{ such that } |x - a| < \delta \tag{2.23}$$

现在，对于连续函数，我们预期，缩小区间$a - \delta < x < a + \delta$（也就是让 δ变小）时，也能相应地找到一个新的 ε 值，使其满足（2.23）中给出的判据。换句话说，随着 $|x - a|$ 变小， $|f(x) - f(a)|$也会变小。这本质上就是函数连续性的定义方式。正式来说，我们将论证倒过来（这在数学定义和证明的表述中很常见.....），如下所示：

定义2.19. 函数f在点$x = a$ 处连续 ，当且仅当（i）a 属于f定义域中的某个开区间，并且（ii）无论我们选取多小的 ε ，总能找到a周围的某个区间（即$a - \delta < x < a + \delta$），使得 $|f(x) - f(a)| < \varepsilon$，无论我们选取的 ε有多小。

Definition 2.20 (Continuous Functions). A function is *continuous* on and interval, I, if it is continuous for *each open sub-interval* of I.

In essence, the "for each open sub-interval" part of the definition is to avoid any possible problems with defining continuity at the least upper or greatest lower bound of the interval. In short, we don't need to worry about continuity at the end points of our domain. This is designed specifically to avoid potential ambiguities as were discussed for the function $f(x) = 1/(x-1)$ $0 \le x < 1$ in Section 2.2.1.

These definitions can be a bit tricky to apply, and they are not the only way to define continuity. Alternatively, one can insist that the following limits, taken from the left and right sides of a must be valid for each point in a domain for a function to be considered continuous.

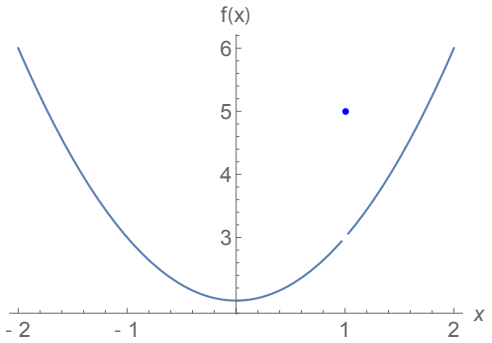


Fig. 2.9 A function. This function has a single point of discontinuity.

$$\lim_{x^+ \to a} |f(x) - f(a)| \implies 0 \tag{2.24}$$

$$\lim_{x^- \to a} |f(x) - f(a)| \implies 0 \tag{2.25}$$

And sometimes this alternative is easier to use in applications. Note. The long double arrow ($\implies$) in mathematics should be interpreted as meaning "this implies". Regardless of which method is used, the results are the same. This is easiest to see with a simple example, as illustrated in Fig. 2.10.

Example 2.4 (Continuity of Functions).

In Fig. 2.10, it is apparent from inspection that the function $g(x) = \frac{1}{2}x^2$ (the orange line on the Fig. 2.10) is continuous. Suppose we consider the point $x = 1$. Clearly, we can pick a small number, say $\varepsilon = 1/1000$, such that there is a small enough interval around $x = 1$ ($1 - \delta < x < 1 + \delta$) where $|f(x) - f(1)| < 1/1000$. In fact, we can compute what this interval is as follows

$$|f(1+\delta) - f(1)| < 1/1000$$

Accounting for the absolute value, there are two options

$$\tfrac{1}{2}(1+\delta)^2 - \tfrac{1}{2} < 1/1000 \quad \text{or} \qquad \tfrac{1}{2} - \tfrac{1}{2}(1-\delta)^2 < 1/1000$$

Solving these inequalities for δ, we find

$$|\delta| < \sqrt{\frac{1002}{1000}} - 1 \qquad |\delta| < 1 - \sqrt{\frac{998}{1000}}$$

The first of these is the smallest, so it is safe to take δ equal to that value.

You can check this directly by noting that

$$\left| f\left(1 + \sqrt{\frac{1002}{1000}} - 1\right) - f(1) \right| = \frac{1}{1000}$$

$$\left| f\left(1 - \sqrt{\frac{1002}{1000}} - 1\right) - f(1) \right| = 0.000999 < \frac{1}{1000}$$

No matter how small we make ε, we can always find a value of δ so that the inequality is valid.

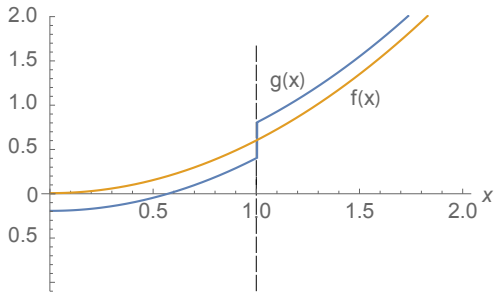


Fig. 2.10 Two functions, f and g. The function g has a discontinuous jump in the value of $\frac{1}{2}$ at $x = 1$.

Now consider the function f (the blue line on Fig. 2.10). Here, our scheme works reasonably well until we set ε to be less than about $1/2$. Because there is a jump of $1/2$ right at $x = 1$, we can easily find an ε that breaks our definition. For example, set $\varepsilon = 1/10$. Now we are looking for the values of x in an interval $a - \delta \leq x \leq a + \delta$, such that $|f(x) - f(a)| < 1/10$. Of course, we can see just by looking at the graph that there is no such interval. In any small interval around $x = a$, the value of $|f(x) - f(a)|$ is *at least* $\frac{1}{2}$, but definitely never smaller than that value.

The definition of continuous, then, aligns with our intuitive notion, even if the definition itself takes a bit of thinking to fully understand.

具有跃变的函数，例如图 2.10 中的函数，被称为不连续函数。当存在有限个不连续点时，这些函数有时被称为分段连续函数。换句话说，分段连续函数是在其中具有有限个跃变且不会在任何地方爆炸到$\pm\infty$的函数。这是一个足够重要的概念，值得给出一个具体定义。

Definition 2.21 (Piecewise Continuous Function). A *piecewise continuous function* is a function that is at continuous everywhere, except at a *finite* number of points.

Within continuous functions, the functions are sometimes by their derivatives. As an example, examine Fig. 2.11. This function is defined by

$$f(x) = \begin{cases} 2x+2 & 0 \leq x < 2 \\ 4\left(x - \frac{1}{2}\right) & 2 \leq x < 3 \\ 10(x-2) & 3 \leq x \leq 4 \end{cases} \tag{2.26}$$

In general, the word *smooth* is used to indicate a function that has a derivative at each point in its domain. Thus, the function given by Eq. (2.19) is smooth; the function given by Eq. (2.22) is technically *non-smooth*, but this arises because of the discontinuity imposed by a single point. Extending the idea of piecewise functions, we can call the function given by Eq. (2.29) *piecewise smooth*. Most of the functions that we study in this text will be of this kind.

定义2.22（分段光滑函数）。分段光滑函数是这样一种函数：该函数的一阶导数处处有明确定义（即导数存在且不是无限值），但有限个点除外。

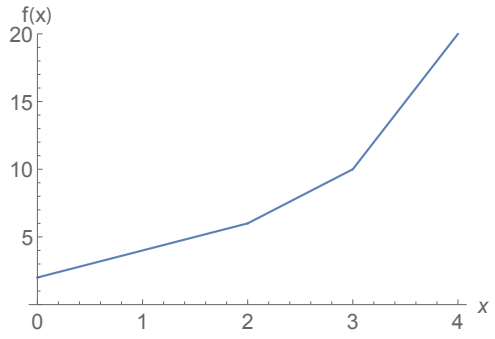


图 2.11 一个在两个点处导数不连续的连续函数。这类函数有时记作C^0，表示它是连续的，但在定义域内并非处处都有导数。

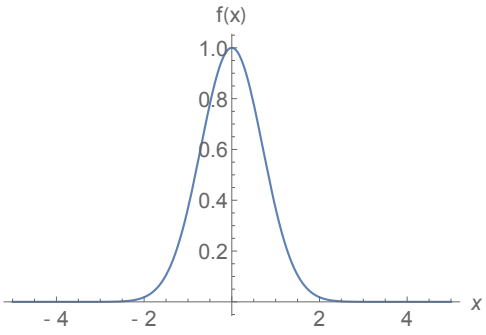


图 2.12 一个高度连续的函数，$f(x) 1/\sqrt{\pi} \quad -^{2}$ exp(x)。该函数具有任意阶导数，因此称为C^∞ 函数。它同时也是解析函数（不过请注意，存在C^∞函数并非解析函数！）。

We can further characterize smoothness by how many derivatives a function has. If a function has an infinite number of derivatives that exist, then it is called C^∞. An example here is the function $f(x) = \exp(-x^2)$. This function is smooth (it is plotted in Fig. 2.12), and is often called a *Gaussian* function. We can define particular C^∞ functions that are called *analytic* functions; however this will have to wait until after the review of derivatives and integration is presented.

有些函数只有有限阶导数存在。举个例子，考虑函数

$$f(x) = \begin{cases} \frac{x^2}{2} & 0 \leq x < 1 \\ -\frac{x^2}{2} + 2x - 1 & 1 \leq x \leq 2 \end{cases} \tag{2.27}$$

这个函数看起来 相当平淡无奇——它在图 2.13 中以橙色曲线绘出。然而，它的导数（以蓝色绘出）在$x = 1$处有一个尖点。因此，虽然它的一阶导数存在且连续，但在$x = 1$处，二阶导数并未定义。为体现这一点，只有一个连续导数的函数称为C^1。虽然我们还没有复习过导数，但这里先给出这一概念的一般化形式。

Definition 2.23 (n^{th}-order continuous function). Suppose a function f is such that the first n derivatives (i.e., f', f'', f''' … $f^{(n)}$) are both *bounded* and exist (i.e., there are no points such that the derivative of order

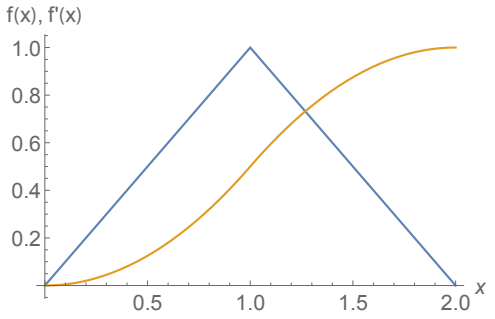


图 2.13 一个连续函数（橙色），其一阶导数（蓝色）也是连续的。然而，二阶导数在 $x = 1$ 处不连续。这个函数称为C^1函数，表示它在定义域内处处具有连续的一阶导数。

$(n-1)$ generates a cusp or other discontinuity). Then the function is continuous in the derivative up to order n, or, more simiply, C^n.

有一类特殊的函数，称为解析函数，它们具有许多有趣的性质。我们熟悉的许多函数都是解析函数，例如：

1. 所有有限次多项式。2. 指数函数和对数函数。3. 三角函数。

在数学发展的早期（例如，直到 20 世纪初），解析函数与函数是同义词。从那时起，函数的概念大为扩展；本文后续内容将介绍一些特殊实例，它们在定义域的某些部分（或处处，例如图 2.15 所示的托梅函数）并非解析函数。

定义 2.24（解析函数——定义 1）。设函数f 定义在定义域D上。如果对每个闭区间K （它是D 的子集，即$K \subset D$），都存在一个常数C，使得对每个点$x \in K$ 以及每个非负整数k，下列界成立，那么函数f 在D 上处处是解析函数：

$$\left|\frac{d^k f}{dx^k}(x)\right| \leq C^{k+1}k! \tag{2.28}$$

While this definition does make a mathematical mouthful, it is relatively easy to understand intuitively. What the definition is trying to tell us is that, for any closed finite interval that is part of the domain of the function, that the function, nor any of its derivatives, go to infinity in that domain. There are some technical issues regarding the domain of the function (and the closed subsets that are selected from it) that will not be covered here. In general, however, we can think of a function as being analytic at a point in its domain if its value and the value of all of its derivatives do not tend toward infinity.

As a final note about the classification of functions, there is one additional classification that is useful to know about. There are, in one sense, two different kinds of functions that we use in common practice. There are polynomial functions, and the roots of polynomial functions to start with. A polynomial function is a polynomial (of any finite order) whose coefficients are also polynomials. For example, a polynomial function of order 4 is defined by

$$a_4(x)x^4 + a_3(x)x^3 + a_2(x)x^2 + a_1(x)x + a_0(x) = 0$$

对于此类函数，其根会生成新的函数 $f(x)$，其中涉及（有理）分数次幂（包括负幂）。这类函数称为<text inline>代数</text>代数 函数，因为它们仅使用代数运算法则，通过一个有限 算法即可定义。以下均为代数函数的实例。

$$f(x) = x$$
$$f(x) = x^{-1} = \tfrac{1}{x}$$
$$f(x) = \sqrt{x}$$
$$f(x) = \frac{\sqrt{1+x^3}}{x^{3/7} - \sqrt{7}x^{1/3}}$$

这些函数的优点在于，只需描述生成它们的多项式，就能以数学方式完整地刻画它们。代数函数是仅涉及有限个项，并且只使用加法、减法、乘法、除法以及有理分数次幂运算的表达式。从某种意义上说，只要我们理解了定义这些函数的运算，就能充分“理解”这些函数。这实际上是构造性的——对于许多代数函数实例，我们可以给出一个算法（包含有限个项），说明该函数是什么以及如何计算其函数值。重要的是，所有代数函数都由某个多项式的根给出；然而，对于次数为 5 或更高的有理系数多项式，并非所有这类多项式的根都能表示为代数函数（对此的证明称为阿贝尔–鲁菲尼定理）。

然而，事实证明，还有许多有趣且实用的函数并不是任何多项式方程的根。常见的实例包括正弦x、余弦x、自然对数x以及e^x函数。这类函数被称为（没错，这确实就是它们的名称）超越函数。这一名称背后的含义是，这些函数超越了我们通常称为代数的学科所能描述的范围。超越函数将在我们后续学习微分方程时出现，其中一些将被证明是描述此类方程的解所必不可少的。所有超越函数都是解析的 （我们稍后将通过泰勒级数详细定义这一术语），因此它们都有收敛的泰勒级数表示。

There are functions that are even more . . . *interesting* to attempt to defined because their structure begins to challenge the concept of function altogether. Consider the following function (sometimes called the modified Dirichlet function or the Thomae function)

$$f(x) = \begin{cases} 0 & x = 0 \\ \frac{1}{q} & x \in \mathbb{Q},\ x = p/q \\ 0 & x \in \mathbb{P} \end{cases} \tag{2.29}$$

It is understood here that $x = p/q$ is expressed in fully reduced fraction form. This is a really unusual function, and clearly it has no simple, closed algebraic formula. A plot of this function appears as Fig. 2.10. In addition to its strange definition and look, it also has some other unusual characteristics. It turns out that the real numbers are “denser” than the rational numbers. In a sense, there are more (many, many more) irrational numbers than rational ones. Thus, this function is continuous (and has a derivative) at each point where x is an irrational number. It is discontinuous at every point where x is rational. It is not terribly important to understand from this example all of the details of why this kind of behavior exists (although

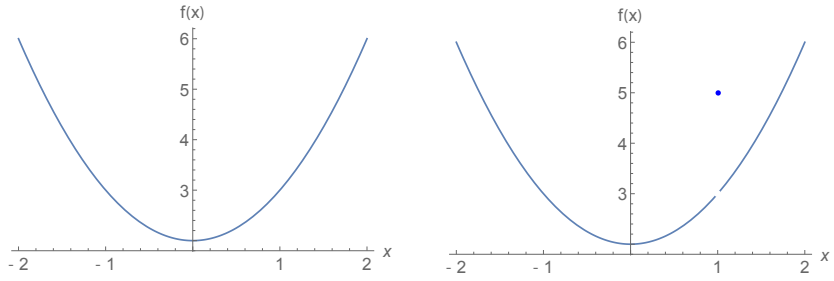


图 2.14：属于同一等价类的两个函数。第二个函数仅在一个点上与第一个函数不同。

our discussion of intervals above is suggestive), but the essential idea is to build an intuition that the real (irrational) numbers constitute a much larger (denser) set than the rational numbers do.

2.2.3 函数等价类

稍后讨论傅里叶级数的概念时，理解函数的等价类这一概念会很有帮助。简而言之，如果两个函数f 和g 至多在有限个或可数个点上彼此不同，则认为它们属于同一个等价类。图 2.14 给出了两个属于同一等价类的函数示例。这两个函数的数学表征如下：

$$f(x) = x^2$$
$$g(x) = \begin{cases} x^2 \ for\, x \neq 1 \\ 5 \ for\, x = 1 \end{cases}$$

注意，对于g （右端的函数），在$x = 1$附近有以下极限：

$$lim_{x \to 1^+} \ g(x) = 1$$
$$lim_{x \to 1^-} \ g(x) = 1$$

其中，1^+ 表示从右端趋近的极限，而 1^- 表示从左边趋近的极限。请务必记住，这里的极限是在趋近值$x = 1$的意义下存在的，但并不真正达到该值。如果需要复习单侧极限，本科微积分教材一定会涵盖这部分内容。

It turns out that, from the perspective of integration (including Riemann integration– see §2.4.3), the presence of a single discontinuous point does not affect the result of the integration. Speaking colloquially, a single point has no measure, so it does not add or subtract from the integral. The following definitions makes this more formal.

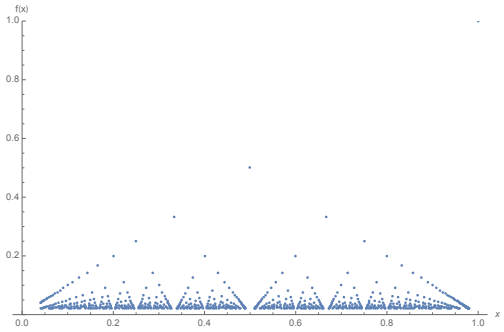


图 2.15 展示托梅函数一组取值的图，其中 $2 \leq p \leq 50$, $2 \leq q \leq 50$。这看起来可能非常奇怪，但它符合函数的判定标准。

Definition 2.25 (Removable discontinuity). A function is said to have a set of removable discontinuities if (a) the number of discontinuous points is finite, and (b) the left and right-hand limits at each point of discontinuity are equal. The value of the function at the discontinuous points can be taken to be the value of the limit at that point.

定义 2.26（函数的等价类）。如果满足以下条件，则称两个函数 f 和 g 属于同一个等价类：(a) 函数 f 和 g 都只有可去不连续点；(b) 对于不位于 f 或 g不连续点处的所有点，$f - g = 0$。或者，条件 (b) 也可以表述为："移除所有不连续点后，$f - g = 0$"。

2.3 线性空间、线性算子与线性代数

在本材料中，我们只有少数几次会用到线性代数，但简要复习一下还是很有必要的，而且这也让我们能够讨论线性算子这一概念。接下来材料中遇到的许多问题都涉及线性算子，因此花些时间理解它们是什么是值得的。

Operators are just what they sounds like: They are mathematical constructs that operate on something (say, a function) to generate something else (for example, a different function) (Fig. 2.6). This is a very general concept, and it is difficult to give it a precise meaning. As an example, all functions are operators. To make this concrete, take a look at the following example.

Example 2.5 (Functions as operators.).

We are all familiar with the idea of a *function*. In the notation that we have are probably most familiar with, a function is defined as a one-to-one (i.e., unique) mapping between two sets. In a more familiar setting, we think of a function as having a *range* (which is usually an interval on the real line, plotted on the horizontal axis by convention) and a domain (which is usually some portion of the real line, plotted on the vertical axis by convention). For example, a properly defined function would be any of the following:

$$f(x) = 5x, \qquad 0 < x < 10 \tag{2.30}$$

$$g(x) = \text{正弦}[2\text{正弦}(2\ \text{正弦}\{2\text{正弦}(x)\})] \qquad \infty < x < \infty \tag{2.31}$$

$$h(x) = \sin[1/x] \qquad -0.1 < x < 0.1 \tag{2.32}$$

These functions are plotted in Fig. 2.16.

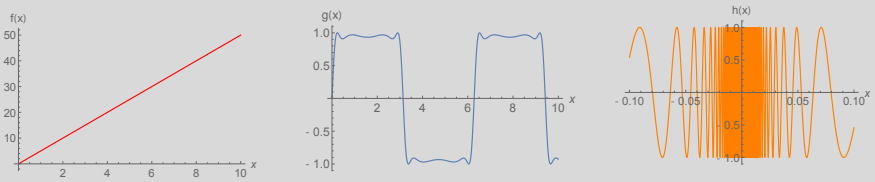


图 2.16：例 1.2 的图。三个函数。

需要认识到，关于这些函数，一个重要的特点是：它们是作用于其自变量的运算。从某种意义上说，函数独立于其自变量而存在。因此，一旦我定义了 独立地 于其自变量的$f(x)$，例如式（2.30）中的定义，那么我实际上所做的就是定义一种运算，即作用于一个集合（自变量），以生成另一个集合（函数值）。通常，当我们将函数视为算子时，会考虑以下两点：（1）函数用于定义运算；（2）该运算作用于输入到函数中的任意对象。在这种意义下，输入到函数中的对象通常称为辐角。辐角可以是在某个区间上定义的简单自变量（这是我们对函数的常规用法），也可以是其他集合。

For example, suppose we define the following set: $X = \{x : x = 1,3,4,5,42\}$. We would read this in words as "The set X is the set that contains values of x such that x is equal any of the values 1, 3, 4, 5, and 42". The set X has a finite number of values in it, but we can still use any of the functions above as operations on this set. Let's interpret $f(X)$ the following way:

$$f(X) = \{f(1), f(3), f(4), f(5), f(42)\}$$

或者，使用表达式$f(x) = 5x$计算其值

$$f(X) = \{5,\ 15,\ 20,\ 25,\ 210\}$$

We can even think about putting functions in other functions (creating a as arguments when we use the operator idea. As an example, consider the function $p(z) = z^2 + 1$. Now, can we make sense of the operations $f(p)$ and $h(p)$? Sure, we need only apply the operations to the argument of the function, regardless of what the argument is.

$$\begin{aligned} f(p) &= 5p(z) \\ &= 5(z^2+1) \\ &= 5z^2+5 \end{aligned} \qquad\qquad \begin{aligned} h(p) &= \sin(1/p(z)) \\ &= \sin\left(\tfrac{1}{z^2+1}\right) \end{aligned}$$

When we interpret a function $f(x)$ this way, we sometimes write the function without the independent variable (with the idea that the independent variable can be anything), as in simply f. This is just a notational device used to be more compact, there is usually no deeper meaning associated with it.

2.3.1 Linear Operators

Now that we understand the basic idea of an operator, we can consider what a *linear operator* is. This is actually pretty simple at this point.

Definition 2.27 (Linear Operator). A *linear operator*, $\mathscr{L}$, is an operator that subscribes to the properties of *additivity* and *homogeneity*. Specifically, this means

$$additivity: \mathscr{L}(f+g) = \mathscr{L}(f) + \mathscr{L}(g) \tag{2.33}$$

$$homogeneity: \mathscr{L}(\alpha f) = \alpha \mathscr{L}(f) \tag{2.34}$$

Linear operators are a subset of all possible operators. In fact, most interesting phenomena in engineering and physics are nonlinear in general. Linearity exists only as an approximation to the more general nonlinear behavior. In this text, we will be generally concerned with (although not exclusively!) linear operations and linear operators.

As discussed above, functions can themselves be though of as operators. Thus, a linear function would subscribe to the properties of additivity and homogeneity above. As and example, the following function is linear

$$f(x) = 5x \tag{2.35}$$

To show that it is linear, note that we can prove both additivity and homogeneity in one step. If we make the substitution $x \to \alpha y + \beta z$, then we find

$$\begin{aligned} f(\alpha y + \beta z) &= 5(\alpha y + \beta z) \\ &= \alpha 5y + \beta \\ &= \alpha f(y) + \beta f(z) \quad (\because \text{ linear}) \end{aligned} \tag{2.36}$$

illustrating that the function is indeed linear.

线性的某个特性可能会造成一些混淆。我们习惯于将诸如$y = mx + b$线性这样的函数称为线性函数。然而，不难证明，这个函数并不满足可加性和齐次性的性质（试试看！）。这种失败有一个微妙的原因。从非常实际的意义上说，$y = mx + b$实际上是更基本函数$y = mx$的平移。请看图 2.17。图中所示的每个函数都是函数$y = \frac{1}{2}x$的平移。这类平移称为仿射变换。事实证明，当我们称一个函数线性时，实际上是指基本函数本身（其中$b = 0$）是线性的。常数项仅表示对这个更基本函数的一种仿射变换。事实上，对于图 2.17 中定义的每条直线，我们都可以通过对坐标系进行变换，简单地消除常数项b。例如，如果我们将坐标系

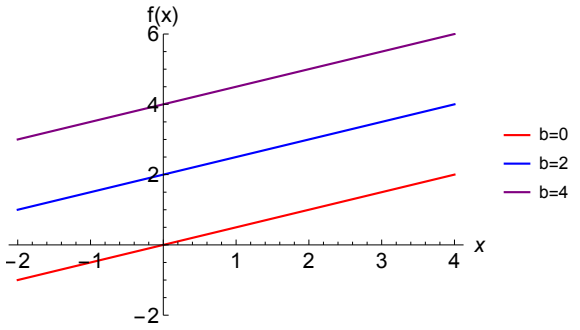


Fig. 2.17 Translations of the function $y = \frac{1}{2}x$.

沿竖直方向移动 2 的距离，那么图中的蓝色直线就会具有$b = 0$。因此，就线性而言，从某种意义上说，一个函数的所有仿射变换都是相同的。为了判断线性，第一步应当是先对坐标系进行仿射变换，使$b = 0$。更实际地说，在检查算子（包括函数）的线性时，我们基本上可以忽略常数。

Example 2.6 (A familiar example: The derivative as an operation.). The process of differentiation is one example of a familiar operation. Consider, for example, the following.

$$\mathscr{L}(f(x)) = \frac{d}{dx}(f(x))$$

在本例中，算子通过导数记号 $\frac{d}{dx}$ 来定义。当我们说 $\mathscr{L} \equiv \frac{d}{dx}$ 时，只是在定义这样一个事实：此处的“抽象”算子实际上就是求导运算，记作 $\frac{d}{dx}$。 —

Example 2.7 (The derivative as a linear *operator.).* Is the derivative operator a *linear* one? We can check. Recall the definition.

$$\mathscr{L}(\alpha f(x) + \beta g(x)) = \alpha\mathscr{L}(f(x)) + \beta\mathscr{L}(g(x)) \tag{2.37}$$

Since our operator in this case is just $\frac{d}{dx}$, we need only to check that the linearity identity is met. In other words, we need to evaluate

$$\begin{aligned}\mathscr{L}(\alpha f(x) + \beta g(x)) &= \tfrac{d}{dx}(\alpha f(x) + \beta g(x)) \\ &= \frac{d}{dx}(\alpha f(x)) + \frac{d}{dx}(\beta g(x)) \\ &= \alpha\frac{d}{dx}(f(x)) + \beta\frac{d}{dx}(g(x))\end{aligned}$$

This is exactly the form that a linear operation must take as defined by Eq. (2.37), so the derivative is a linear operator.

A compound linear operator is just the sum or one or more linear operators. The easiest way to understand these is through some simple examples. Although we will not use operator notation extensively in this text, it is a useful notation to understand.

Example 2.8 (Compound linear operators). Compound linear operators act as follows.

$$(\mathscr{L}_1+\mathscr{L}_2)\,(f(x)) \equiv \mathscr{L}_1(f(x))+\mathscr{L}_2(f(x)) \tag{2.38}$$

It is important to recognize that, although the notation looks as though it indicates multiplication, it does not! Here is a more complex looking example that helps to understand the idea better.

$$\begin{aligned}\left(\frac{d}{dx}+\frac{d}{dy}\right)(x^2y) &= \frac{d}{dx}(x^2y)+\frac{d}{dy}(x^2y)\\ &= 2xy+x^2\end{aligned}$$

Note that the operators *do not* act like multiplication!

$$\begin{aligned}\left(\frac{d}{dx}+\frac{d}{dy}\right)(x^2y+y^3) &= \frac{d}{dx}(x^2y+y^3)+\frac{d}{dy}(x^2y+y^3)\\ &= 2xy+x^2+3y^2\end{aligned}$$

函数也可以被视为算子；但只有齐次函数才是线性算子。 s.

Example 2.9 (Functions as operators.). Consider the functions defined in the previous example.

$$f(x) = 5x, \qquad 0<x<10 \tag{2.39}$$
$$g(x) = \sin[2\sin(2\sin\{2\sin(x)\})] \qquad \infty<x<\infty \tag{2.40}$$
$$h(x) = \sin[1/x] \qquad -0.1<x<0.1 \tag{2.41}$$

Are they linear operators? To check this, we need only try the operations on the quantity $\alpha y+\beta z$ to see if the conditions given by Eqs. (2.33)-(2.34) are met. For the generic operator L, we will substitute our particular functional operators. Thus, for the example of f, we have the result

$$\begin{aligned} f(\alpha y+\beta z) &= 5(\alpha y+\beta z) \\ &= \alpha 5y+\beta 5z \\ &= \alpha f(y)+\beta f(z) \end{aligned}$$

这表明f是线性的。对于函数g，我们得到

$$\begin{aligned} g(\alpha y+\beta z) &= \sin[2\sin(2\sin\{2\sin(\alpha y+\beta z)\})] \\ &\neq \alpha g(y)+\beta g(z) \end{aligned}$$

so g is not a linear operator (and also not a linear function!). The final function h is left to the reader to check.

There are two more concepts that are useful to introduce when discussing the (somewhat abstract) concept of linear operators. These are the *identity operator* and the *inverse operator*.

定义 2.28（恒等算子）。恒等算子 $\mathscr{I}$ 是指满足以下条件的任意算子：对于 $\mathscr{I}$ 定义域中的所有可接受对象（例如数、函数、向量等）f，均有 $\mathscr{I}(f) = f$。

Example 2.10 (Identity operators.). Here are three common examples of identity operators that you have seen before.

1. The number "1" is the identity operator in arithmetic. For any number a, we have $1 \cdot a = a$.
2. We can define an identity function $g(f)$ in a domain $D = \{0 < x < 1\}$ as follows: $g(f(x)) = f(x)$ for all x in D. It is not too difficult to recognize that this operator is equivalent to multiplying a function by the number 1.
3. consider the following matrix multiplication (Note: linear algebra is covered in the next section).

$$\mathbf{I} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

现在，对于每一个可能的二维向量 $\mathbf{a}$，我们都有

$$\mathbf{I} \cdot \mathbf{a} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$$

Therefore, $\mathbf{I}$ is an identity operator for the domain of vectors of dimension two.

We won't be using the *theorem-proof* format very often in this material, but occasionally it is helpful, especially when the proofs are short and clever. The definition of the inverse operator can be done this way.

Theorem 2.2 (Inverse operator). *A linear operator can have an inverse, $\mathscr{L}^{-1}$, only if $\mathscr{L}(x) = 0$ implies that $x = 0$.*

Proof. If $\mathscr{L}(x) = y$ then the inverse of $\mathscr{L}$ is the mapping which takes y back to x. (AN ASIDE: Here, it might be helpful to think of a conventional function, such that $\mathscr{L}(x) = y(x)$, and x is the domain of the horizontal axis. For example, $\mathscr{L}(x) = x^2$ *for* $0 < x < 1$ associates each independent value in the domain x with a unique value for the result (which we call $y(x)$) in the range.)

现在假设 $\mathscr{L}(\mathrm{x_1}) = \mathrm{y_0}$ 且 $\mathscr{L}(\mathrm{x_2}) = \mathrm{y_0}$。那么，根据线性性，$\mathscr{L}(\mathrm{x_1} - \mathrm{x_2}) = 0$。以下结论必有一个成立：(a) $\mathrm{x_1} - \mathrm{x_2} = 0$（即，$\mathrm{x_1}$ 和 $\mathrm{x_2}$ 是同一个值）；或者 (b) $\mathrm{x_1} - \mathrm{x_2} \neq 0$，因此定义域中的两个 x 值经过映射后，在值域中对应同一个 $\mathrm{y_0}$。对于反函数，定义域和值域的角色互换。因此，对于反函数，y 构成定义域，而 $\mathscr{L}^{-1}(\mathrm{y}) = \mathrm{x}$ 是值域。然而，根据定义，一个函数的定义域中的一个值只能对应值域中的一个值。因此，$\mathrm{x_1}$ 和 $\mathrm{x_2}$ 不相等的情况不可能成立（否则反函数就不是函数）。所以必有 $\mathrm{x_1} = \mathrm{x_2},$，这意味着 $\mathscr{L}(\mathrm{x}) = 0$ 仅当 $\mathrm{x} = 0$时才成立。■

不必过于深入技术细节，就我们的目的而言，如果能够证明以下条件，逆线性算子 $\mathscr{L}^{-1}$ 就存在：

$$\mathscr{L}^{-1}(\mathscr{L}) = \mathscr{L}(\mathscr{L}^{-1}) = \mathscr{I} \tag{2.42}$$

并且 $\mathscr{L}^{-1}$ 绝不会是多值的。

通常，实例确实有助于让这些概念更加清晰。

例*2.11*（逆算子。）

1. 假设我们有 $f(x) = x^2$，其中 $\{x:\ 0 < x < 1\}$。这个函数有逆吗？试试这些函数

$$y = f(x) = x^2 \tag{2.43}$$
$$x = g(y) = \sqrt{y} \tag{2.44}$$

现在注意，$\mathrm{f(g(y))} = \mathrm{f}(\sqrt{\mathrm{y}}) = \mathrm{y}$。但是，根据公式（2.43），我们有 $\mathrm{y} = \mathrm{x}^2$；因此，$\mathrm{f(g(y))} = \mathrm{x}^2$。现在注意，恒等算子满足 $\mathscr{I}(\mathrm{f(x)}) = \mathscr{I}(\mathrm{x}^2) \equiv \mathrm{x}^2$。所以，根据定义，我们必须有 $\mathrm{f(g(x))} = \mathrm{x}^2 \equiv \mathscr{I}(\mathrm{f(x)})$。因此，f (g(x)) 是恒等算子，并且根据定义，g 是 f 的逆。

Functions that are inverses of one another have an interesting graphical feature. Consider the two functions above; if we plot them both as functions of x (i.e., we plot the functions $y = x^2$ and $y = \sqrt{x}$, we obtain the plot below. Functions that are inverses of one another have reflective symmetry about the line $y = x$.

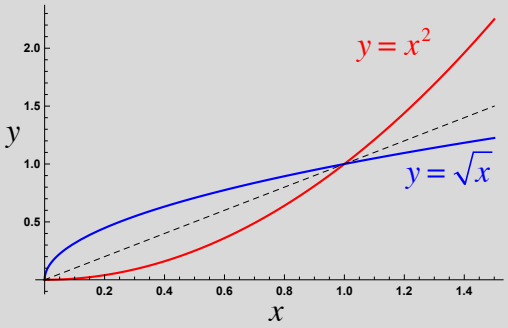


Fig. 2.18: Inverse functions.

There is one caveat that needs to be attended to when discussing the idea of linear operators and linear equations. The idea of *linearity* is usually associated with only the *operator* part of the equation. In particular, what this means is that constants involved in a function are excluded. This is by convention, not necessity. Consider the following linear equation.

$$f(x) = 3x + 12 \tag{2.45}$$

Obviously the operator f must be linear, because $f(x)$ is a line! However, if we attempt our normal process for checking linearity, we encounter a problem

$$f(\alpha y + \beta z) = 3(\alpha y + \beta z) + 12 \qquad = \alpha 3y + \beta 3z + 12 \tag{2.46}$$

$$\neq \alpha f(y) + \beta f(z) \tag{2.47}$$

那么，哪里出了问题？这个问题的答案有点棘手。从严格意义上讲，我们不应将常数视为算子的一部分。换句话说，假设我们定义 $L(x) = 3x$。那么，上面的方程可以写成

$$f(x) = \mathscr{L}(x) + 12 \tag{2.48}$$

And the actual question can be posed as "is the operator $\mathscr{L}$ linear"? And the answer to that question is obviously yes. The important thing to remember here is that the *linearity of the equation* has only to do with the operators involved, not with any particular constants that may be tacked on to the operators. Thus, a linear equation is any equation whose operator (which, by definition does not include additive constants) is linear.

2.3.2 Linear Algebra

Now that we know what a linear operator is, we can discuss linear algebra. Most people's experience with linear algebra leaves them without an intuitive notion of what is really going on. Because linear algebra is such a mathematically rich topic– it is the first place where more formal mathematical analysis can sensibly be done in terms of vector spaces – it tends to favor mathematical detail over process. At any rate, for the material covered here, a simple review of linear algebra is sufficient. To start, we will consider a set of linear functions; for the purposes of generating a concrete example, consider the following

$$f_1(x,y,z) = 3x+2y+z-5=0 \tag{2.49}$$
$$f_2(x,y,z) = x+y-z-1=0 \tag{2.50}$$
$$f_3(x,y,z) = 2x+2y+3z-10=0 \tag{2.51}$$

显然，这是一个方程组，而且显然它们是线性的。这些方程有解吗？嗯，我们知道，如果这三个方程是线性无关的，那么它们一定有解。也就是说，没有任何一个方程是另外两个方程的线性组合；或者说，不存在系数a，$b \neq 0$使得$f_1 = af_2 + bf_3$。请注意，这一个条件就足够了（为什么？）。假设这些方程是线性无关的，并且方程的数量与变量的数量相同，那么这个方程组就有解（这是由线性代数基本定理得出的结论——我们将不对此进行证明！）

矩阵代数之所以诞生，原因之一就是把这些方程全部写下来有些重复。因此，我们可以通过定义矩阵乘法来简化符号表示。首先请注意，上面的方程组可以写得更加简洁（省略不必要的函数标签）

$$3x+2y+z=5 \tag{2.52}$$
$$x+y-z=1 \tag{2.53}$$
$$2x+2y+3z=10 \tag{2.54}$$

Now, suppose we define four vectors, as follows

$$\vec{r}_1 = (3,2,1)$$
$$\vec{r}_2 = (1,1,-1)$$
$$\vec{r}_3 = (2,2,3)$$
$$\vec{x} = (x,y,z)$$

回顾一下，两个向量之间的点积定义为 $\mathbf{a}\cdot\mathbf{x} = (a_1,a_2,a_3)\cdot(x,\ y,\ z) = a_1x + a_2y + a_3z$。注意到这一点后，我们可以将方程更紧凑地写成

$$\vec{r}_1\cdot\vec{x} = 5$$
$$\vec{r}_2\cdot\vec{x} = 1$$
$$\vec{r}_1\cdot\vec{x} = 10$$

Or, noting $\vec{x} = (x,y,z)$, using conventional notation for vectors

$$\begin{bmatrix}\mathbf{r}_1\\ \mathbf{r}_2\\ \mathbf{r}_3\end{bmatrix}\cdot\begin{bmatrix}x\\ y\\ z\end{bmatrix}=\begin{bmatrix}5\\ 1\\ 10\end{bmatrix} \tag{2.55}$$

This essentially *defines* matrix multiplication by the notation

$$\begin{bmatrix}3 & 2 & 1\\ 1 & 1 & -1\\ 2 & 2 & 3\end{bmatrix}\begin{bmatrix}x\\ y\\ z\end{bmatrix}=\begin{bmatrix}5\\ 1\\ 10\end{bmatrix} \tag{2.56}$$

其中，为完成乘法，需要将每一行（视为向量）与向量 $\vec{x}=(x,y,z)$ 作点积。这是记住如何进行矩阵乘法的最简单方法！

为了解决这个问题，我们可以采用多种方法。最简单的方法就是按照如果仍然将所有内容写成包含三个未知数的三个方程时所采用的方法来处理：同时从方程中消去变量。简而言之，我们可以进行以下任意一种运算：

1. 将任意一行乘以一个常数。 2. 将任意两行相加，并用结果
替换这两行中的任意一行。 3. 交换任意两行。

这里有一点需要注意——无论我们对矩阵的行做什么，都需要对右端的向量做同样的操作（毕竟它们是方程！）最简单的理解方式就是做一个例子。我们可以用一种极其有条理的算法化方式来完成，但对于小矩阵，通常最好是寻找简化的机会。首先，将第 2 行乘以 -2，然后加到第 3 行，再将第 3 行除以 5。得到。

$$\begin{bmatrix}3 & 2 & 1\\ 1 & 1 & -1\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}x\\ y\\ z\end{bmatrix}=\begin{bmatrix}5\\ 1\\ 8/5\end{bmatrix} \tag{2.57}$$

Now, multiply row three by -1 and add to row 1; then add row 3 to row 2. Now we have

$$\begin{bmatrix}3 & 2 & 0\\ 1 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}x\\ y\\ z\end{bmatrix}=\begin{bmatrix}17/5\\ 13/5\\ 8/5\end{bmatrix} \tag{2.58}$$

Clearly, we are close. Multipy row 2 by -2 and add to row 1.

$$\begin{bmatrix}1 & 0 & 0\\ 1 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}x\\ y\\ z\end{bmatrix}=\begin{bmatrix}-9/5\\ 13/5\\ 8/5\end{bmatrix} \tag{2.59}$$

As the last step, multiply row 1 by -1, and add it to row 2

$$\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}x\\ y\\ z\end{bmatrix}=\begin{bmatrix}-9/5\\ 22/5\\ 8/5\end{bmatrix} \tag{2.60}$$

The matrix on the left-hand side of this expression is called the *identity* matrix, denoted $\mathbf{I}$. It is the matrix version of "1". Carrying out the matrix multiplication on the left-hand side leads to the solution

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -9/5 \\ 22/5 \\ 8/5 \end{bmatrix} \tag{2.61}$$

As a final note regarding linear algebra, there are a few more words to say about determining whether or not a set of equations is solvable or not. Above, we mentioned that the equations needed to be linearly independent. That is actually not an easy thing to check (even though it is easy to define). It turns out that there is a characteristic number for a set of equations that indicates whether or not there is a solution to them. This number is called the determinant, and it turns out to be somewhat difficult to define. The following serves as a decent definition for the general case of an $n \times n$ matrix.

Definition 2.29. The determinant of a $n \times n$ matrix can be found by

1. Putting the matrix in upper or lower triangular form (i.e., conducting row reduction).
2. Once in triangular form, multiplying the values on the diagonal gives the determinant.

Here, *triangular form* just means eliminating all of the entries in the matrix either above or below the diagonal. There are a few caveats when using this approach to compute the determinant. While you can always add a multiple of one row to another row, for other row-reduction operations more care is needed. The following list provides guidance for the reduction process.

Table 2.1: A list of operations that can be done for computing the determinant, and the associated effect of the operation.

	Type of operation	**效果**
1	Add a multiple of one row to another row	无影响
2	Multiply a row by a constant, c	Determinant is multiplied by c
3	Interchange two rows	Determinant changes sign

Example 2.12 (Determinants.). Consider the matrix we just examined in the material above

$$\mathbf{A} = \begin{bmatrix} 3 & 2 & 1 \\ 1 & 1 & -1 \\ 2 & 2 & 3 \end{bmatrix} \tag{2.62}$$

To find its determinant, we first perform row operations (but not interchanging any two rows, and not multiplying any row by a constant) to eliminate the entries above or below the diagonal to make a triangular matrix. A quick glance at the matrix $\mathbf{A}$ indicates that either way will be reasonably easy. One set of steps is as follows: (a) Multiply row 2 by -2 and adding it to row 3 (replacing row 3 with the sum), (b) multiply row 1 by $-1/3$ and adding it to row 2 (replacing row 2 with the sum) should do it. The result is

$$\mathbf{A} = \begin{bmatrix} 3 & 2 & 1 \\ 0 & 1/3 & -4/3 \\ 0 & 0 & 5 \end{bmatrix} \tag{2.63}$$

That wasn't too bad. According to our definition, the determinant is just the product of the entries on the diagonal.

$$\det(\mathbf{A}) = |\mathbf{A}| = 3 \times 1/3 \times 5 = 5 \tag{2.64}$$

这里主要出于参考目的，展示了行列式的两种符号表示形式（“det”和竖线）。

行列式有几个实用之处。对我们而言，它很有用，因为它可以告诉你一组方程是否存在冗余（也就是说，它们是否线性无关）。

定义 2.30。 方阵 **A** 线性无关当且仅当它具有非零行列式，$\det(A) \neq 0$。

顺便说一下，当且仅当 （有时写作*iff*）表示结论在两个方向上都成立。以上面的例子来说，它的意思是：“如果一个方阵的行列式非零，那么它就是线性无关的”并且 “如果一个方阵线性无关，那么它的行列式非零。”

有一种求解线性方程组的方法称为克莱姆法则，它只涉及计算行列式。对于较小的（2×2 或 3×3）矩阵，如果行列式不难计算，这种方法就很有用。问题中会进一步讨论这种方法。

2.4 Calculus

这部分微积分复习与前面的章节一样，篇幅较短，重点在于复习对后续内容有用的知识点，而不是对整个学科进行全面总结。复习的合理起点是导数的定义。

2.4.1 The Derivative

Definition 2.31 (The Derivative of a Function). A derivative of a continuous (at least C^1, as defined above) function, f, at a point x in the domain of the function is defined by the limit

$$f'(x) = \lim_{t \to x} \frac{f(t) - f(x)}{t - x} \tag{2.65}$$

or, equivalently, we can define

$$f'(x) = \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x} \tag{2.66}$$

There are at least two common notations for the derivative: f' and $\frac{df}{dx}$. For functions of a single variable, there is little chance for confusion. Sometimes, for purposes of clarity or presentation, one of the two is preferable to the other, especially when multiple functions of different dependent variables are considered.

Note that the definition of the derivative automatically provides the definition of higher-order derivatives. For example, consider the function $g(x) = f'(x)$. Then we have

$$g'(x) = (f'(x))' = \lim_{\Delta x \to 0} \frac{g(x+\Delta x) - g(x)}{\Delta x}$$

Or, establishing the notation for the second derivative

$$f''(x) = \frac{d^2 f}{dx^2} = \lim_{\Delta x \to 0} \frac{f'(x+\Delta x) - f'(x)}{\Delta x}$$

Or, applying the original definition of the derivative given by Eq. 2.66, we find

$$f''(x) = \frac{d^2 f}{dx^2} = \lim_{\Delta x \to 0} \frac{f(x+2\Delta x) - 2f(x+\Delta x) + f(x)}{(\Delta x)^2}$$

This provides some explanation for why the second derivative is denoted by $\frac{d^2 f}{dx^2}$.

Generally, we do not derive the derivatives of functions from first principles, except perhaps in our introductory course on calculus. However, it is useful to recall how this is done.

Example 2.13 (Derivatives.). Computing derivatives directly from the definition of the derivative is not tremendously difficult, but sometimes it does require a little creativeness in determining the limit. As an example, let's look at how to find the derivative of the function $f(x) = x^2$.

$$\begin{aligned} f'(x) &= \lim_{t \to x} \frac{f(t) - f(x)}{t - x} \\ &= \lim_{t \to x} \frac{t^2 - x^2}{t - x} \\ &= \lim_{t \to x} \frac{(t-x)(t+x)}{t - x} \\ &= \lim_{t \to x} (t + x) \\ &= 2x \end{aligned}$$

Using the definition of the derivative, it is also possible to derive what the derivative of a *product of two functions* is. This is called the *product rule for differentiation* or the *Leibniz rule for differentiation*.

定理 2.3（微分的乘积法则）。

$$\frac{d}{dx}[f(x)g(x)] = f'(x)g(x) + f(x)g'(x) \tag{2.67}$$

Proof. The proof for this is just a straightforward application of the definition of the derivative

$$\begin{aligned}\frac{d}{dx}[f(x)g(x)] &= \lim_{t\to x}\frac{f(t)g(t)-f(x)g(x)}{t-x}\\ &= \lim_{t\to x}\frac{f(t)g(t)+[f(x)g(t)-f(x)g(t)]-f(x)g(x)}{t-x}\\ &= \lim_{t\to x}\frac{[f(t)-f(x)]g(t)}{t-x}+\lim_{t\to x}\frac{[g(t)-g(x)]f(x)}{t-x}\\ &= \frac{df(x)}{dx}g(x)+f(x)\frac{dg(x)}{dx}\end{aligned}$$

■

还有一条至关重要的求导规则，即导数复合规则，更常被称为求导链式法则。当一个函数嵌套在另一个函数中时，这条规则十分有用；换句话说，这种函数就是复合函数。

定义 2.32（复合函数）。复合函数是指自变量本身也是一个函数的函数。

例如，考虑位置与速度之间的关系。假设有一个位置函数（在一维空间中），满足 $x = x(t)$。那么，速度取决于位置和时间，$v(x(t), t)$，这是一个复合函数。

链式法则提供了一种计算复合函数导数的方法。链式法则的证明相当复杂，因此此处不予介绍，不过几乎所有微积分入门教材中都有相关内容。然而，这一结果十分重要，在后续内容中会频繁用到。

定理 2.4（求导链式法则）。假设有一个函数 f，其辐角是另一个函数 g。假设 $g(t)$ 在点集 $T = \{t : a < x < b\}$ 上可导，并且 $f(y)$ 在点集 $Y = \{g(t)\}$ 上可导。复合函数 $f(g(t))$ 可导，其导数为

$$\frac{d}{dt}f(g(t)) = \frac{df(g)}{dg}\frac{dg(t)}{dt}$$

Keeping the functions g and f straight sometimes causes confusion. Some examples are helpful.

例 *2.14*（乘积法则，速度—位置示例。）

Above, we mentioned the relationship between position and velocity as an example that could be considered a composite function problem. This is one example where one can compute derivatives both with and without the use of the chain rule. In essence, it provides a way to validate that the chain rule leads to correct results.

Suppose you go out to run along the street, but in a very strange way (why you decide to do this is open to discussion). As you run along, you run in a way such that your position is the cube of the time you have been running (obviously, you can't keep this up forever, but for a short time it is possible.) In particular, suppose your position, $x(t)$, is given by $x(t) = \frac{1}{3}t^3$. Now, we know that velocity is defined as the time rate of change of position, and acceleration is the time rate of change of velocity, so

$$v(t) = \frac{dx(t)}{dt} \qquad a(t) = \frac{dv(t)}{dt} \qquad (A)$$
$$= t^2 \qquad = 2t$$

从这里开始，我们将进行一些看起来有些不寻常的操作，但这是验证链式法则所必需的。首先，请注意，我们可以将 x 表示为关于 v 的函数，如下所示（注意，$v = t^2$ 隐含着 $t = v^{\frac{1}{2}}$）。

$$x(v(t)) = \frac{1}{3}t^3 = \frac{1}{3}v^{\frac{3}{2}} \qquad (B)$$

Now, we have expressed x as a composite function, $x(v(t))$. We can compute the time derivative of this function by

$$\frac{dx(v(t))}{dt} = \frac{dx}{dv}\frac{dv}{dt} \qquad (C)$$

We can compute dx/dv by

$$\frac{dx}{dv} = \frac{d}{dv}\left(\frac{1}{3}v^{\frac{3}{2}}\right) = \frac{1}{2}v^{\frac{1}{2}}$$

and dv/dt has already been computed above in Eq. (A). Combining these, we find

$$\frac{dx(v(t))}{dt} = \frac{dx}{dv}\frac{dv}{dt}$$
$$= \left(\frac{1}{2}v^{\frac{1}{2}}\right)(2t)$$

and substituting $v = t^2$ gives the result

$$\frac{dx(v(t))}{dt} = t^2$$

这与上文公式 (A) 中给出的结果完全相同。

Example 2.15 (Product Rule.). Some kinds of catalyst can be deactivated by chemicals produced during the catalytic reaction, or by external factors such as UV light. This process is sometimes called *catalytic poisoning*. Similar kinds of deactivation can happen to the enzymes in cell systems in biological reactors.

Suppose that an experiment is run, and it is determined that the amount of product begin produced, under conditions of deactivation (whether cells or catalyst), is given by

$$c(t) = c_0 \exp\left[-(k_0 - k_1 t)t\right]$$

换言之，类一级速率过程具有一个随时间变化的有效速率常数（即 $k_{e\,有效} = k_0 = k_1 t$）。求反应速率 $c'(t)$。

解。

Using the product rule requires that we first identify the composite functions. Often, in practice, this is not done explicitly; rather, people just keep mental note of which function is which. However, it is instructive to explicitly identify the functions when there is any potential for confusion. For this problem, take

$$\begin{aligned} g(t) &= -(k_0 - k_1 t)t \\ f(g) &= c_0 \exp(g) \end{aligned}$$

We need to recall the definition for the derivative of the exponential

$$\frac{d}{dg} exp(g) = exp(g)$$

The exponential is the only function (except the function 0) whose derivative is the same as the starting function! With this, we have all we need.

$$\begin{aligned} \frac{dg(t)}{dt} &= -(k_0 - \frac{k_1 t}{2}) \\ \frac{df(g)}{dg} &= \exp(g) \end{aligned}$$

So, the result is ...

$$\begin{aligned} \frac{d}{dt} f(g(t)) &= \frac{df(g)}{dg}\frac{dg(t)}{dt} \\ &= c_0 \exp(g)\left[-(k_0 - \frac{k_1 t}{2})\right] \end{aligned}$$

Substituting the function $g(x)$ (from above) into this result and rearranging gives the final result

$$\frac{d}{dt} f(g(t)) = -c_0(k_0 - \frac{k_1 t}{2}) \exp\left[-(k_0 - k_1 t)t\right]$$

2.4.2 偏导数

There is not much more to say regarding derivatives, except to explain the notion of a derivative when a function has multiple independent variables. So, with out delay, we will define the partial derivative of a function with two independent variables. The case of additional independent variables is identical, so no more than two is required for the definition.

定义 2.33（偏导数）。对于含有两个自变量的函数，偏导数是指对其中一个变量求导，同时将另一个变量视为常数。假设$f(x,t)$ 是定义在定义域 Ω 上的<style id='x'>x 和t 的连续（至少为C^1）函数（Ω 可能是形状不规则的定义域，因此我们不再尝试对其作更详细的集合描述）。f 关于各变量的偏导数为

$$\frac{\partial f(x,t)}{\partial x} = \lim_{\Delta x \to 0} \frac{f(x+\Delta x,t) - f(x,t)}{\Delta x}$$
$$\frac{\partial f(x,t)}{\partial t} = \lim_{\Delta t \to 0} \frac{f(x,t+\Delta t) - f(x,t)}{\Delta t}$$

对于三个或更多变量的函数，也有类似的结果；根据上述实例，其推广应当相当直观。

定义 2.34（两个或三个变量函数的链式法则）。回顾一下，复合函数是指自变量本身也是一个函数。当一个函数依赖于两个或更多个本身也是函数的变量时，链式法则可用于按以下形式求导

$$\frac{d}{dt} f(x(t),y(t)) = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}$$

请注意，在此定义中，$x(t)$ 和 $y(t)$ 的导数并非偏导数，而是常规导数。这是因为 $x(t)$ 和 $y(t)$ 都是单变量函数，即关于变量 t 的函数。因此，对于这些量，常规导数才是正确的导数形式。

2.4.3 Integration

Integration is a simple concept, but it turns out to be quite deep in actual applications. In introductory calculus, the concept of the Riemann integral is introduced. This integral applies to most functions that are encountered in science and engineering; in particular, it is useful for evaluating functions that are *piecewise smooth* as defined previously. This is not the only kind of integral that can be defined. For example, consider the following function (the Dirac function)

$$f(x) = \begin{cases} 1 \text{ if x is irrational on} x \in [0,1] \\ 0 \text{ if x is rational on } x \in [0,1] \end{cases}$$

This function is not piecewise smooth because it has an infinite number of holes in it (i.e., it has a hole at every possible fraction between 0 and 1!) The conventional Riemann integral cannot be used for such a function. However, there are more general forms of the integral (such as the Lebesgue integral) that can be used to measure such functions. We will not explore the Lebesgue integral in this text, but we will introduce (broadly) the ideas behind it at the end of this section on integration.

就我们的目的而言，这里不会详细回顾积分的发展过程，而是介绍积分的一些重要性质。

2.4.3.1 黎曼和、积分与微分

虽然这里不需要全面讲解积分理论，但有必要回顾一下连续函数常规定积分的定义。这里的主要目的是复习积分定义的基本思想，而不是尽可能地构造最一般的概念。因此，目前只需考虑连续函数即可（不过，该定义也适用于只有有限个间断点的函数）。假设我们要计算一个单自变量函数 f 在某个区间 $x \in [a,b]$上的积分。回想一下，这对应于 a 到 b 之间的曲线下面积。如果给我们这样一条曲线，并要求用图形方式计算面积，我们可能会尝试构造一系列矩形来近似这条曲线，从而估算面积（见图）。假设我们这样做，并使每个矩形在 x–轴上具有相同的宽度。令 $X = \{x_1, x_2, x_3 \ldots x_N\}$，使得对所有 $i = 1$ 至 $N - 1$，$x_{i+1} - x_i = \Delta x$（一个常数）；因此，$\Delta x = (b-a)/(N-1)$。下面的和是函数 f 的积分的一个近似

$$f_{\Delta x} = \sum_{i=1}^{i=(N-1)} f(x_i^*)\Delta x \tag{2.68}$$

其中，x_i^*是x 的任意取值，且满足$x_i \le x_i^* \le x_{i+1}$。因此，在此计算中，$f(x_i^*)$是$x_i$ 与x_{i+1}之间矩形高度的估计值。因此，定义$f_{\Delta x}$ 的和是曲线下面积的估计值。随着 Δx 逐渐变小，这个和自然会越来越准确，因为在每个矩形中，恒定的高度越来越能表征该函数。于是，黎曼和定义为

$$\int_a^b f(x)\,dx = \lim_{N\to\infty}\left(\sum_{i=1}^{i=(N-1)} f(x_i^*)\Delta x\right) \tag{2.69}$$

图 2.19 展示了逐步细化黎曼和的过程。

这里有几点值得指出。首先，注意积分符号实际上源自一个拉长的“S”，以提醒我们积分是由求和产生的。其次，积分中的dx 有特定且可解释的含义。它通常称为微分，但微分的概念并不像表面看起来那样直观。事实上，即使在微积分诞生多年之后，微分的概念仍不甚明晰。虽然有许多方法可以使微分的概念在形式上明确（例如，可将其定义为从实数集到实数集的线性映射，参见 Protter 等人（2012，第 7 章）），但就我们的目的而言，只需按如下方式理解即可。

微分是对函数变化量的非零近似，其中因变量的变化量可以小到所需的程度，从而使近似达到期望的准确度。更正式地说，对于函数$g(x)$，其微分定义为

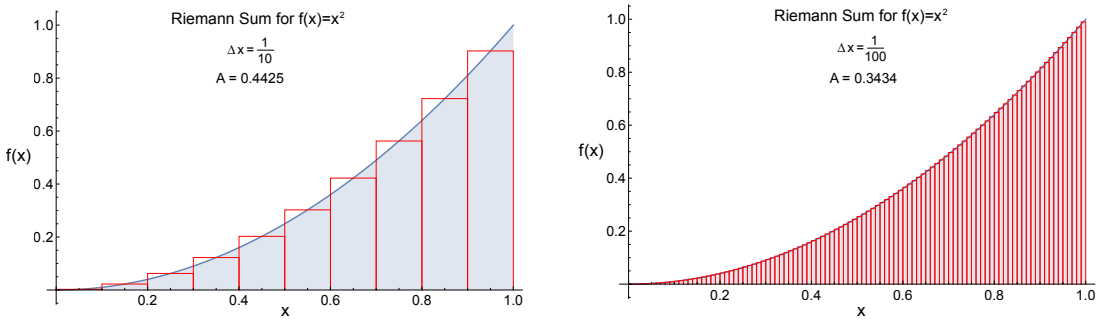


Fig. 2.19: Two examples of Riemann sums. For these example, x_i^* was taken at the midpoint of each rectangle. On the left, a divisions of the domain with $N = 10$, gives a total of 0.4424; this represents a 32% over-estimate of the actual value. On the right divisions of the domain with $N = 100$, gives a total of 0.4424; this represents only a 3% over-estimate of the actual value.

$$dg = g'(x)dx \tag{2.70}$$

where dg stands for a "small" change in the function g, and dx stands for a small change in the independent variable. In this case, the meaning of "small" is a qualified one; it means small enough such that the error involved is less than some specified error. Importantly, the concept is that the error can be driven to zero as dx approaches zero. While we have not yet reviewed the concept of the Taylor series (this appears in the material following), one can use a Taylor series to define the concept. Recall

$$g(x+\Delta x) = g(x) + \Delta x g'(x) + O[(\Delta x)^2] + \ldots \tag{2.71}$$

假设 $\Delta x \ll 1$，则 $(\Delta x)^2 \ll (Deltax)$，相对于含有 Δx 的项，$(\Delta x)^2$ 及更高次项可以忽略。重新整理可得

$$g(x+\Delta x) - g(x) + = \Delta x g'(x) \tag{2.72}$$

现在，我们注意到，符号 dx用于表示 Δx，其条件是 Δx可以任意小（这实际上就是它的定义）；类似地，$dg = g(x+\Delta x) - g(x)$。这就给出了我们所需的结果。

$$dg = g'(x)\Delta x \tag{2.73}$$

注意，如果我们取函数 $g(x) = x$，最终会得到关系 $dx = Deltax$。这解释了为什么在积分的表征中使用 dx。按照惯例，等式 $dx = Deltax$ 意味着微分的定义为

$$dg = g'(x)dx \tag{2.74}$$

This is more than simply a formal manipulation of symbols. Once we have defined differentials to be quantities whose error can be made as small as we like, then the resulting structure is essentially a linear

one. Some powerful methods in the analysis of, for example, non-Euclidian geometry. They also arise in the study of differential equations.

2.4.3.2 微积分基本定理

The fundamental theorem of calculus says some really important things. Primarily, it tells us the following.

Theorem 2.5. *Suppose a smooth function, F, is defined on the interval $[a,b]$. Because the function is smooth, it has a derivative, $F' = f$. Then*

$$\int_{t=a}^{t=b} f(t)dt = F(b) - F(a)$$

or, equivalently,

$$\int_{t=a}^{t=b} \frac{dF(t)}{dt}dt = F(b) - F(a)$$

将这一定理推广到分段连续函数上很简单。只需分别计算每个（有限个）连续区间上的积分即可。

这是一个非常强大的定理，它实质上将求积分的问题转化为求导数的问题。也就是说，如果给定一个函数f，并且我们恰好知道一个导数等于f的函数F，那么就可以利用这一知识计算f的积分。因此，函数F有时被称为f的原函数。尽管这听起来有些循环论证，但事实并非如此。许多已知积分之所以为人所知，是因为我们已经找到了相应函数的原函数。

对许多人来说，第一门微积分课程会学习许多求函数原函数的“技巧”。我们会把大部分内容留在过去，认为大家都会记住（或者能够查到）遇到的大多数基本积分和导数。不过，有一种“技巧”在许多应用中都非常有用，而且我们有机会多次使用它。这就是分部积分法。

Theorem 2.6. *Let f and g be smooth functions on an interval $x \in [a,b]$. Then*

$$\int_{x=a}^{x=b} f(x)\frac{dg(x)}{dx}dx = f(x)g(x)\Big|_{x=a}^{x=b} - \int_{x=a}^{x=b} \frac{df(x)}{dx}g(x)dx$$

通常，这条规则写成一种便于记忆的形式。

$$\int_{x=a}^{x=b} udv = uv\Big|_{x=a}^{x=b} - \int_{x=a}^{x=b} vdu$$

Integration by parts is particularly useful under the following circumstances: (i) there is an integrand of the form of a nth-order polynomial times some function that we know how to ingegrate n times, or (ii) there is an integrand that contains an th-order derivative times some function that we know how to differentiate n times, and we would like to eliminate the derivative. This will be made clearer in the following example.

Example 2.16 (Integration by parts.). Integrate the following functions over the interval $x \in [0,1]$

$$(a) \qquad f(x) = x\sin(\pi x)$$

$$(b) \qquad g(x) = \frac{df}{dx}\sin(\pi x) \quad \text{given } f(0) = 0$$

解。对于 (a)，令 $u = x$，$dv =$ 正弦$(\pi x)dx$。则，$du = dx$ 且 $v =$ $-\pi$余弦(πx)（注意$d/dx(-$余弦 $x) =$正弦 x）。然后

$$\begin{aligned}\int_0^1 x\sin(\pi x)\,dx &= \frac{x}{\pi}\cos(\pi x)\Big|_0^1 - \int_0^1 -\cos(\pi x)dx \\ &= \frac{1}{\pi} + \sin(\pi x)\Big|_0^1 \\ &= \frac{1}{\pi}\end{aligned}$$

For (b), let $u = \sin(\pi x)$, $dv = \frac{df}{dx}dx$. Then, $du = 1/\pi cos(\pi x)$ and $v = f(x)$ (fundamental theorem of calculus). Then

$$\begin{aligned}\int_0^1 \frac{df}{dx}\sin(\pi x)\,dx &= \frac{x}{\pi}\cos(\pi x)f(x)\Big|_0^1 - \int_0^1 f(x)\frac{1}{\pi}\cos(\pi x)dx \\ &= -\int_0^1 f(x)\frac{1}{\pi}\cos(\pi x)dx\end{aligned}$$

In this latter example, we can go no further than this without knowing more about f; however, we have *eliminated* the derivative, which is frequently a useful operation.

2.4.3.3 Lebesque Integration

In the material above, we presented an example of the Dirac function that was not integrable using the conventional (Riemann) integral. Here, consider the following variation of that function

$$f(x) = \begin{cases} x \text{ if x is irrational on} x \in [0,1] \\ 0 \text{ if x is rational on } x \in [0,1] \end{cases}$$

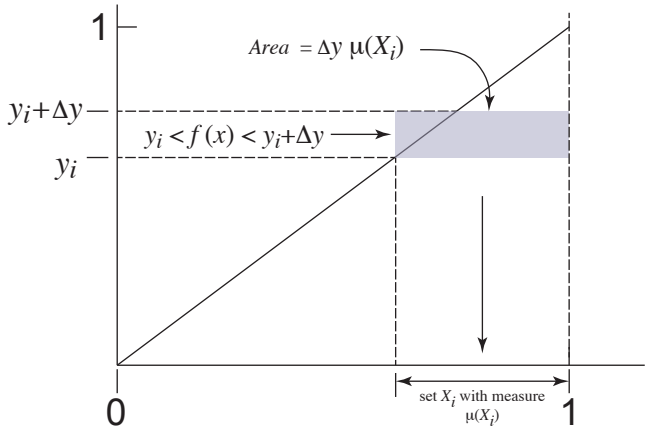


图 2.20 勒贝格积分中值域的划分。

This is a function that maps all possible values of the independent variable x on the interval $X = \{x : x \in [0,1]\}$ to some real number (in this case, this real number is also on the interval $f(x) \in [0,1]$). In the mathematical literature, this function (mapping) is sometimes written as "$f : X \to \mathbb{R}$", which just means (in written English) "each element of X is assigned exactly one value in the real numbers".

20 世纪初，一位名叫 Henri Lebesgue 的法国数学家以稍微抽象一些的方式思考积分，从而推广了积分的概念。他没有考虑先将定义域划分，再通过各种极限运算为积分赋值的"良好"函数（对于严格为正的函数，这个值对应于曲线下方的"面积"），而是转而考虑划分值域。图 2.20 展示了这一过程的一个例子，其中函数 $f(x)$ 的值域和定义域都在区间 $[0,1]$ 上。此时，不再是对某个区间上的竖直小块求和，而是对与 $x \in [0,1]$ 对应的值域适当部分上的"水平"小块求和。

如果我们能为区间 测度 x_i 赋予一个合理的测度，就能计算每个这样的小块的面积；经过适当的极限过程后，将这些面积相加便得到积分。现在的问题是，如何理解为区间 x_i 赋予测度。对于黎曼可积函数，这个测度对应于我们在黎曼和中通常考虑的测度。

Measure Theory is the discipline within mathematics that deals with the question of "how do we assign metrics to mathematical quantities?" We are all familiar with the conventional Euclidian metric in $\mathbb{R}^3$– it is just the *length* of what we think of as vectors. So, for a point (x,y,z) with origin $(0,0,0)$, the Euclidian metric is

$$d(x,y,z) = \sqrt{x^2+y^2+z^2} \tag{2.75}$$

对于更奇特的集合（例如实数轴上的有理数），人们逐渐发展出许多巧妙的方法，为这类集合赋予合理的"测度"概念。特别是，勒贝格积分所使用的测度会将零测度赋予孤立点（也就是说，数轴上的一个点被赋予零距离，这在一定程度上符合我们的直觉）。事实证明，有理数的稠密程度远不如无理数——也就是说，从某种意义上说，无理数的数量远多于有理数。因此，根据勒贝格测度，有理数.....

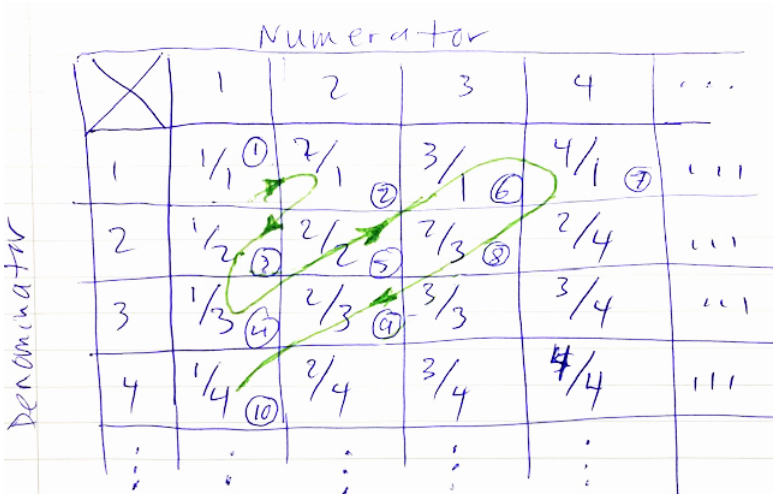


图 2.21 利用整数枚举有理数。

数看起来像是具有零测度的孤立点。虽然我已经简要说明了这样一种概念确实存在，但要实际演示如何通过构造来实现，远远超出了我们在此所能完成的范围。不过，这类考量并非毫无直觉。可以建立有理数与实数之间的一一对应关系；如果你从未见过这种构造，图 2.21 展示了辐角的核心。因此，虽然有理数可以与整数一一对应，但无理数不可以。而且，只有这两种可能——一个数要么是有理数，要么是无理数。因此，无理数比有理数稠密得多。由此，我们完全可以将有理数视为以孤立的空洞（点）渗透在所有可能实数构成的集合中。由于这些点没有测度，因此任意实区间的测度，例如 $[a,b]$，就是 $\mu = (a-b)$。同样，这看起来或许像是我们一无所获，但事实并非如此！对于像 $[0,1]$ 上的无理数这样的集合，传统的黎曼和积分并不存在——在这种情况下，确实无法考虑划分区间（如果区间端点之一落在有理数上，会怎样？当你让 Δx 趋近于零时会发生什么——它必然会经过有理数！这类问题还有很多.....）。然而，采用勒贝格积分法后，这就不再是问题了。

For the purposes of this text, this is all we need to discuss about the Lebesgue integral (although there is much more that can be said about it!) It is worth noting, however, that you may encounter this kind of integration in the mathematical literature. If you do, you have most of the idea behind what such integrals actually represent, and you may still be able to read through the material. As a final note, in the mathematics literature, one often encounters the concept of the integral of the square of a function. This is usually called the L_2 metric, and it is understood that it is measured relative to the Lebesgue measure (hence the "L" in the symbol for the metric).

2.5 Sequences and Series

Infinite sequences and series are incredibly useful, and they are the true workhorses of applied mathematics. Many important functions (especially many of the transcendental ones that we use routinely, such as $\sin(x)$ and $\exp(x)$) are frequently *defined* and/or *computed* by use of an infinite series. The coverage of sequences

and series here is necessarily limited. However, in the chapters following, we will have ample opportunity to revisit these concepts.

2.5.1 数列

数列与级数相关。理解数列的一种方式是将其视为一个数或函数表达式的有序列表；因此，数列将整数（定义域）映射到另一个集合（该集合中的元素可以是常数、整数的函数，或整数加上其他变量的函数）。因此，将无穷数列中的每一项相加，是构成无穷级数的一种方式。虽然在微积分入门课程中，无穷数列通常被视为仅以整数为自变量的函数，但它们也可以是其他自变量自身的函数。我们将在研究幂级数和泰勒级数时，以及在后续章节学习傅里叶级数时，使用此类数列的和。

为了具体说明数列的概念，我们给出如下定义。

定义 2.35（有限数列）。有限数列是一个对象（元素）列表，其索引取自自然数的一个连续子集（$\mathbb{N} = \{1,2,3,4..N\}$或 $\mathbb{N}_0 = \{0,1,2,3,4..N\}$）。在本文中，含有$N$项的数列（如果第一个索引为 0，则为$N+1$项）记作$A = (a_1, a_2, a3, ..., a_N)$（如果第一个索引为 0，则记作$A = (a_0, a_1, a2, ..., a_N)$）。位置$i$上的元素（其中$i \in \mathbb{N}$或 $\mathbb{N}_0$）称为该数列的第i^{th}项。无论元素如何定义，数列都被视为定义在自然数上的函数。

Note– the primary feature of sequences is that they are a *list* indexed by the real numbers. One can think of this as a column of values on a spreadsheet program, where each entry has a unique number. Also like a spreadsheet program, the objects in the list may be functions of the index (its numbered location), or of both the index and additional independent variables (representing, for example, spatial coordinate or time).

Definition 2.36 (Infinite Sequence). A *sequence* is a list of objects (elements) indexed by the natural numbers, ($\mathbb{N} = \{1,2,3,4\ldots\}$ or $\mathbb{N}_0 = \{0,1,2,3,4\ldots\}$). This list may be specified by a functional rule, or by listing the elements explicitly. A common notation for an infinite sequence is $A = (a_n)_{n=1}^{\infty}$ (or $A = (a_n)_{n=0}^{\infty}$ if the first index is zero).

关于有限数列的所有说明同样适用于无穷数列。特别是，以特定方式收敛于某个指定函数的函数无穷数列，在应用分析中是相当常用的概念。我们将在后续内容中进一步讨论这一概念。

Example 2.17 (Examples of sequences). Sequences are just lists of objects. Thus, they represent a fairly general mathematical concept. Here are a few examples.

1. The following is a somewhat boring finite sequence: $A = (1,1,1,1,3,1)$. It has a finite number of terms (six), and each element is the same, except for the fifth term.
2. The following is an infinite sequence containing all of the even numbers greater than zero: $A = (2,4,6,8,\ldots)$.

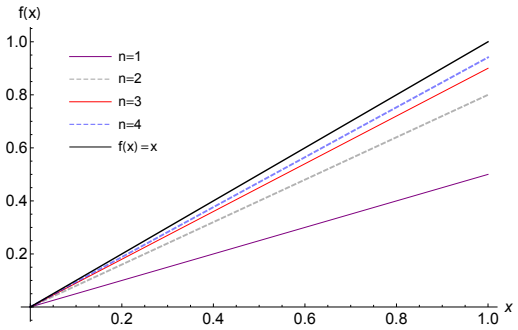


图2.22 由函数 组成的x的序列。在这种情况下，序列本身仍以实数n为索引。该序列由$a_n(x)$ =给出

3. Here are two other ways to denote the very same sequence of the previous example: $A = (2n)_{n=1}^{\infty}$ and $S = (2n)_{n\in\mathbb{N}}$.

4. 这是一个表示熟悉的无理数的无穷数列：A =(3，1，4，1，5，9，2，6，5，3，5，...)

5. 这是一个项在n增大时收敛于

$1 \quad \to \infty \quad = 1 + 1/ \quad {}_{n=1}^{\infty}$ =的数列：$A(n^2)$，，，，

$(2\ {}^{5}_{4}\ {}^{10}_{9}\ {}^{17}_{16}\ {}^{26}_{25}\ \cdots)$

，。我们可以计算该数列的以下极限：$\lim_{n\to\infty}(a_n) = 1 + \lim_{n\to\infty}({}^{1}_{n^2}) = 1$。

6. 以下例子是一个同时依赖于整数集合（它们构成该序列）和另一个自变量的函数（为具体起见，我们可以将其视为空间）。

$$A = (1/(1+1/n^2)x)_{n=1}^{\infty}$$

因此，逐项计算时，该序列可表示为 $A = (x/2, x/(1+1/4)x, x/(1+1/9), x/(1+1/16)\ldots)$。函数的前四个序列（$n = 1, 2, 3$ 和 4）的图像见图2.22。稍加思考便可发现，该序列最终收敛于函数$f(x) = x$。

在上面的实例中，我们引入了无穷数列收敛的概念。具体来说，我们需要明确数列收敛意味着什么。我们需要考虑的问题是，如何证明一个无穷数列确实趋近于我们定义的某个极限。这有些棘手。假设有一个记作 $A = (a_n)_{n\in\mathbb{N}}$ 的数列。仔细想想，并不存在一个实际的数a_∞。对于无穷数列，我们真正能问的只有：“当n 变成任意大的整数时，会发生什么？”这正是极限概念的核心。虽然在本节之前的内容中，我们讨论了趋近于有限值的极限，但尚未讨论无穷极限。

A *mathematical* discussion of limits inevitably an ε-δ-like argument. In this case, in stead of having a δ interval in which some error measure must be small, we instead have an integer for which all values of the sequence indexed by this integer or higher are sufficiently close to a limit. To make this clear, here is the definition of the limit of a sequence. Here, for clarity, we focus on sequences that are a function of only n (unlike the final case described in the previous Example); however, extension of these concepts can be done on a pointwise basis for sequences of functions.

Definition 2.37 (Limits of a sequence: Sequence convergence). Suppose $(a_n)_{n=1}^{\infty}$ is a sequence. Then, the limit, L, exists if the sequence gets arbitrarily close to L as n increases. In other words, for every $\varepsilon > 0$ that can be chosen, no matter how small the value of ε, then it is also true that there is an integer, N, such that

$$if\ n \geq N,\ \ then\ \ |a_n - L| < \varepsilon \tag{2.76}$$

在这种情况下，我们称 $(a_n)_{n=1}^{\infty}$ 有一个极限，L，并记作

$$\lim_{n\to\infty} a_n = L \tag{2.77}$$

如果一个序列不收敛，是否意味着它必然发散？答案是否定的，反例 $(a_n)_{n=1}^{\infty} = (-1)^n$，$n = 1, 2, 3 \ldots$便说明了这一点。这个序列既不收敛也不发散；它只是在 -1 和 1 之间周期性地振荡。我们可以给出发散序列的定义，具体如下。

定义 2.38（发散序列）。设 $(a_n)_{n=1}^{\infty}$ 是一个序列。那么，如果a_n 随着n任意增大而趋于无穷，就称该序列发散。换句话说，对于任意数 $M < \infty$，无论M多大，总是存在一个整数N 使得

$$if\ n \geq N,\ \ then\ \ |a_n| > M \tag{2.78}$$

在这种情况下，我们称 $(a_n)_{n=1}^{\infty}$ 发散（即，极限趋于 $\pm\infty$）。

$$\lim_{n\to\infty} a_n = \infty\ \ or\ \ \lim_{n\to\infty} a_n = -\infty \tag{2.79}$$

用这个定义判断一个序列是否收敛时，一个问题是必须先知道极限L，才能检验该定义。法国数学家奥古斯丁-路易·柯西在19世纪初的一部著作中描述了一种解决这一问题的方法（尽管从严格意义上说，最先描述这一方法的是数学家伯纳德·波尔查诺（来自如今属于捷克共和国的一片地区）！）。柯西（和波尔查诺）的一个有趣见解是：不从绝对意义上定义收敛，而是根据各项彼此之间的关系来定义。在给出定义之前，请注意其中使用了逻辑术语“当且仅当”（iff）。这仅表示该定义正反两个方向都成立。定义陈述之后还会对此作进一步解释。

定义 2.39（柯西序列）。若无穷数列 $(a_n)_{m=1}^{\infty}$ 对任意可能的 $\varepsilon > 0$，都存在一个正整数N 使得，则称其为柯西序列。

$$\|a_n - a_m\| < \varepsilon \text{ for all } m > N \text{ for all } n > N$$

Just a note about the terminology “iff”. In the defintion above, the “iff” indicates that if you specify a value for ε, then you can always find values for $n, m > N$ such that the proof is true. Also, if you specify a value of N, then for all $n, m > N$ you can always find a value of ε such that the proof is true. So, the proof works both directions: an N is implied to exist if the sequence is Cauchy. In the other direction, if an N is specified, then a value of ε is supposed to exist if the sequence is Cauchy.

The usefulness of Cauchy sequences can be found in the following theorem, versions of which were proven by both Cauchy and Bolzano.

定理 2.7（序列收敛的柯西收敛准则）。序列（a_n）$_{n=1}^{\infty}$ 收敛的充要条件是它是柯西序列。

这里还需要对“充要”一词作一些补充说明。从某种意义上说，这一术语与“当且仅当”有关。“充要”这一术语用于说明定理的结论是否总能推出初始陈述（即证明是否可以反向成立）。简而言之，“充要”意味着证明沿正向总是成立的（所有柯西序列都收敛），但沿反向只在某些情况下成立（并非所有收敛序列都是柯西序列）。在本例中，“充分”表示，观察到收敛这一事实可以作为序列可能是柯西序列的证据（因此，一个不收敛的序列不可能是柯西序列），但这一证据不足以证明该结论成立。还必须补充其他条件（例如上面给出的定义），才能确定一个收敛序列是否为柯西序列。

In summary, the concept of a Cauchy sequence is a powerful one because it allows one to investigate convergence properties of a sequence without first knowing the limit. This closes an important logical gap (i.e., if one knows the limit of a sequence, then it is already obvious if the sequence converges or not!), and was an important landmark in abstract mathematical analysis.

2.5.2 Series

With the concept of sequences defined, it is relatively straightforward to define an *infinite series*. In short, an infinite series is the sum of some infinite sequence. However, it is frequently useful to think about them in the following sense.

Definition 2.40 (Series). Suppose $(a_n)_{n=0}^{\infty}$ is a sequence. Now, define the *partial sums* of the sequence by

$$S_N = \sum_{n=0}^{n=N} a_n$$

i.e.,

$$\begin{aligned} S_0 &= a_0 \\ S_1 &= a_0 + a_1 \\ S_2 &= a_0 + a_1 + a_2 \\ S_3 &= a_0 + a_1 + a_2 + a_3 \\ &\dots \end{aligned}$$

Then, the S_n are known as the partial sums of the series. As we allow $N \to \infty$, the partial sums define the following infinite series

$$S = \sum_{n=0}^{\infty} a_n \tag{2.80}$$

There are two important things to note here.

1. As noted above, it is not necessary that the lower bound of a series start at 0. However, most series used routinely in applied mathematics start at either 0 or 1. If starting at some integer other than 0, the definitions above would be modified in the obvious way.

2. Not all series necessarily *converge*. In fact, it turns out that even some *non-convergent* series are useful in applied mathematics. In fact, if you have ever applied Stirling's approximation for the factorial

$$n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n \left(1+\frac{1}{12n}+\frac{1}{288n^2}-\frac{139}{51840n^3}-\frac{571}{2488320n^4}+\cdots\right) \tag{2.81}$$

then you have used a non-convergent series. The concepts of convergence are one of the most important in applied mathematics, so some review of important results will be presented in the material following.

3. 无穷级数通常用于表示完整的函数。正如在序列相关内容中所述，我们可以将级数中的各项视为一个函数序列的和，该函数序列同时依赖于 n 和其他一些自变量。

2.5.3 Series Convergence

To start, it is important to define what it means for a series to converge. As we have seen above, the list of partial sums form a sequence, so the study of convergence of sequences and series are substantially intertwined.

定义 2.41（无穷级数的收敛）。假设我们按照上述方式定义无穷级数 S_n 的部分和。那么，如果部分和在 $N \to \infty$ 趋于无穷时趋向于一个固定的极限L，则称该级数收敛 。

$$S = \lim_{N\to\infty} s_N = \sum_{n=0}^{N} a_n = L \tag{2.82}$$

定义2.42（无穷级数的绝对收敛）。假设我们按照上文所述定义无穷级数 S_n 的部分和。那么，如果部分和在 $N \to \infty$ 趋于无穷时趋向某个固定极限 L，就称该级数绝对收敛。

$$S = \lim_{N\to\infty} s_N = \sum_{n=0}^{N} |a_n| = L \tag{2.83}$$

有许多收敛判别法可用于判断一个序列（或级数）是否收敛。以下判别法适用于各种级数（特别注明的除外），包括：(1) 积分判别法，(2) 比较判别法，(3) 极限比较判别法，(4) 比值判别法，(5) 根值判别法，以及 (6) 莱布尼茨收敛判别法。下面列出这些判别法，不予证明。

以下列出相关定义和收敛定理，不予证明。这些定理都是众所周知的，你可能在学习微积分时已经接触过。任何一本涵盖无穷级数的微积分入门教材中都能找到这些定理的证明。

定理 2.8（积分判别法）。设$(a_n)_{n=1}^{\infty}$是一个非负（即每一项均为正数或零）序列，并设f是定义在 $[0,\infty)$ 上的连续单调递减函数，使得

$$f(n) = a_n \ \ for\, n \geq 1 \tag{2.84}$$

然后这个级数

$$\sum_{n=1}^{n=\infty} a_n \tag{2.85}$$

converges if and only if the integral

$$\int_{x=1}^{\infty} f(x)dx \tag{2.86}$$

is finite.

Theorem 2.9 (Comparison Test). *Suppose it is known that the series* $\sum_{n=0}^{\infty}|b_n|$ *converges. If, for a second series* $\sum_{n=0}^{\infty}|a_n|$ *the condition* $|a_n| \leq |b_n|$*, then the second series in* a_n *also converges.*

定理 2.10（极限比较判别法）。假设级数 $\sum_{n=0}^{\infty}|b_n|$ 收敛。对于另一个级数，$\sum_{n=0}^{\infty}|a_n|$，我们希望确定其收敛性质。如果条件

$$\lim_{N\to\infty}\left|\frac{a_n}{b_n}\right| = L \tag{2.87}$$

where L is some finite number, then the series associated with a_n *also converges.*

Theorem 2.11 (Ratio Test). *Suppose* a_n *is not equal to zero for all values of n. We can say that a series converges (absolutely) if*

$$\lim_{n\to\infty}\frac{|a_{n+1}|}{|a_n|} < 1 \tag{2.88}$$

This latter expression is called the (general) ratio test (which you may have learned in a course on calculus). Note that if $r = 1$*, it is not obvious whether the series converges or diverges.*
The converse is also true; that is, a series is said to diverge *if*

$$\lim_{n\to\infty}\frac{|a_{n+1}|}{|a_n|} > 1 \tag{2.89}$$

If the limit of this quantity is exactly equal to 1, then nothing can be said about convergence of the series.

Theorem 2.12 (Root Test). *Suppose that one computes the limit*

$$\lim_{n\to\infty}(|a_n|)^{\frac{1}{n}} = L \tag{2.90}$$

且 $|(L|)$ 是一个满足 $|(L < 1|)$ 的有限数。因此，该级数绝对收敛。如果 $|(L > 1|)$，则该级数发散。如果 $|(L = 1|)$，则根值判别法无法判断该级数是否收敛。

定理 2.13（莱布尼茨收敛判别法）。

Suppose the sequence $(|a_n|)_{n=1}^{n=\infty}$ *is*

1. *A monotonically decreasing sequence (*$|a_{n+1}| < |a_n|$ *for all values of n), and*
2. $\lim_{n\to\infty} a_n = 0$

Then the alternating series

$$\sum_{n=0}^{\infty}(-1)^n a_n \tag{2.91}$$

$$\sum_{n=0}^{\infty}(-1)^{n+1} a_n \tag{2.92}$$

converge.

2.5.4 Power Series

A *power series* is just a label applied to a particular kind of series. A power series is defined by any series of the form

$$f(x) = \sum_{n=0}^{\infty} c_n x^n \tag{2.93}$$

One very significant distinction that separates power series from series more generally is that power series are defined by the sum of a *sequence of functions*. Note that comparing a power series with the general series described above, we have the important correspondence

$$a_n = c_n x^n \tag{2.94}$$

Note that for a power series (as with series in general) the lower index can start at any number, although this number will generally be 0 or 1. Power series are one of the most frequently used devices in applied mathematics. Note that the power series can also be shifted to be defined around any point, a, in the domain as follows

$$f(x) = \sum_{n=0}^{\infty} d_n (x-a)^n \tag{2.95}$$

where now we have the correspondence

$$a_n = d_n (x-a)^n \tag{2.96}$$

这种移位形式的幂级数是一种方便的符号表示方式。然而，请注意，原则上可以展开该级数并重新组合各项，从而恢复方程 (2.93)。因此，方程 (2.93) 和 (2.95) 表示同一个级数的两种不同的表达方式（即，它们可以是等价的）。

与上述讨论类似，可以将幂级数的部分和定义为截断到某个编号为N的有限项处的和。具体而言，注意到上述定义的幂级数的前几个部分和为

$$
\begin{aligned}
S_0 &= c_0 \\
S_1 &= c_0 + c_1 x \\
S_2 &= c_0 + c_1 x + c_2 x^2 \\
S_3 &= c_0 + c_1 x + c_2 x^2 + c_3 x^3 \\
&\dots
\end{aligned}
$$

It is useful to compare these partial sums with those given in Definition 2.40.

In the next example, we provide the power series representation of a few familiar functions.

Example 2.18 (Power Series.). Trigonometric functions are examples of transcendental functions (as introduced above). By definition, this means that there is no *finite* set of algebraic steps that one can use to generate such functions (e.g., there is no polynomial that will reproduce them exactly). If you think about trigonometric functions, you will realize that you have probably never been asked to compute the value of, say, the sine function yourself. This is because the sine function must be defined by an algorithm that has, in principle, an infinite number of steps!

虽然像正弦(x)和余弦(x)这样的函数无法用有限个代数步骤表示，但它们可以用无限多个代数步骤表示。以下是这些函数的幂级数。

$$\cos(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \dots \tag{2.97}$$

$$\sin(x) = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots \tag{2.98}$$

诚然，原则上，这些函数的值由无限多个项定义，但级数中的项通常会随着n越来越大而逐渐减小其模。从实用角度来说，这意味着在许多情况下，只需使用有限个项，就能在关注的点处获得对级数的良好表征。例如，下图绘出了使用$N = 2$，5，10，和 20 项对正弦函数进行近似的结果。显然，即使对于$N = 2$，幂级数在约为 $0 < x < 1.5$ 的区间内，仍能很好地近似该函数。

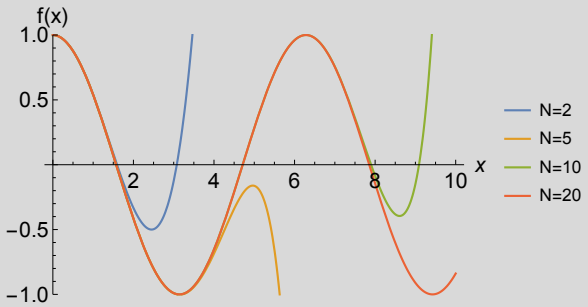


Fig. 2.23: The sine function approximated by its power series, using an increasing number of terms.

2.6 Taylor's Theorem and Taylor Series

泰勒级数是一种特殊的幂级数，以使其闻名的英国数学家布鲁克·泰勒命名。泰勒的过人之处在于，他推导出了如何为任何已知其导数的函数计算幂级数各项的方法。有趣的是，牛顿本人也提出了我们如今称为泰勒级数的级数（Ferraro 等人，2008），不过他关于这一主题的原始手稿在其有生之年一直未曾发表。更令人称奇的是，一位名叫 Colin Maclaurin 的苏格兰数学家在一本数学教科书中推广了牛顿未发表的无穷级数方法（并得到牛顿认可）；在这种形式下，级数围绕零展开，称为麦克劳林 级数。牛顿、Maclaurin 和泰勒在同一时期都是伦敦皇家学会的成员。Maclaurin 和泰勒都认识牛顿（O'Connor 和 Robertson，2017 年 5 月），而且 Maclaurin 和泰勒很可能也彼此相识。尽管这三位科学家兼数学家彼此相识，他们在级数方面的许多工作都各自包含独到的贡献，而且在许多情况下，同样的想法甚至被不同的人独立"重新发现"。这或许凸显了这样一个观点：科学（尤其是在通信不像今天这样即时的时代）往往就是一个不断发现与重新发现的过程。有句格言大意是："所谓发现者，通常不过是那个能够将其描述得足够清楚、从而无需后人再次发现的人。"事实证明，泰勒级数是解析数学中最重要的发展成果之一；因此，这些思想经过一段时间逐渐形成，并非完全出自泰勒一人之手，也就不足为奇了。

2.6.1 Taylor Series

泰勒级数是幂级数的一种特殊形式。它与一般幂级数的**区别**在于，其系数 由待表示为级数形式的函数的各阶导数组合确定。首先，我们通过构造性定义来定义泰勒级数，如下所示。

Definition 2.43 (Taylor Series). For a function that has an infinite number of derivatives at a point, $x = a$, that function has a Taylor series expansion that represents the value of the function at a location $x = (a + \Delta x)$ as follows (where $\Delta x = [x - a]$ or, equivalently, $x = a + \Delta x$)

$$\underbrace{f(a+\Delta x)}_{f(x)} = f(a) + \Delta x \left.\frac{df}{dx}\right|_{x=a} + \frac{1}{2!}(\Delta x)^2 \left.\frac{d^2f}{dx^2}\right|_{x=a} + \frac{1}{3!}(\Delta x)^3 \left.\frac{d^3f}{dx^3}\right|_{x=a} + \ldots$$

or, equivalently, the following notation is frequently used

$$f(x) = f(a) + (x-a)f'(a) + \frac{1}{2!}(x-a)^2 f''(a) + \frac{1}{3!}(x-a)^3 f'''(a) + \ldots$$

Finally, note this last form can be written in the familiar summation notation for the Taylor series

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n$$

where $f^{(n)}$ is the n$^t h$ derivative of the function f, and the factorial of an integer, n, is defined by $n! = n \cdot (n-1) \cdot (n-2) \ldots 2 \cdot 1$, and $0! = 1$. Note that with this definition, we have that a Taylor series is a (shifted) power series of the form

$$d_n = \frac{f^{(n)}(a)}{n!}$$

其中，回顾一下，系数 d_n 是为式 (2.95) 给出的移位幂级数定义的，这些系数与非移位幂级数的系数不同。

此定义并未说明该级数是否收敛；收敛问题将在后续内容中讨论。上文关于级数的导论材料将序列部分和的概念定义为级数的一种类型。现在，这一概念扩展到泰勒级数的部分和；这些部分和各自表示一个完整的函数，而不只是一个数。

定义 2.44（部分和）。泰勒级数的部分和是截断到某个正整数N处的和，使得

$$f_N(x) = \sum_{n=0}^{N} \frac{f^{(n)}(a)}{n!}(x-a)^n$$

注意，每个无穷级数也可以用其部分和序列来描述。这一点很重要，因为它展示了如何用一列函数f_n来描述一个函数，并使其收敛到所需的函数f。下一个例子将对此进行说明。不过，首先我们需要讨论哪些类型 的函数具有收敛的泰勒级数。这些函数就是前文所述的解析函数 。这里将给出“解析”一词的第二个（也更根本的）定义。

The Taylor series provides us with an intuitive tool to define *analytic* functions (which were defined previously in §2.2.2). This second definition is as follows. Note that $x = a + \Delta x$ implies that $\Delta x = (x-a)$. Thus for $a = 0$, we have $\Delta x = x$ (a Maclaurin series).

Definition 2.45 (Analytic Functions– Definition 2). An *analytic function* is a function which (i) is defined (i.e., it has a computable value) on an open interval (I) of the real line (I is the set of values, x, such that $x \in (a,b)$; or, in set notation $I = \{x : a < x < b\}$), (ii) its Taylor series converges to a definite value on the interval I, and (iii) the definite value it converges to is equal to $f(x)$ for all values of x in the interval I.

Admittedly, that is quite a mouthful of requirements; however, many functions that of are of interest in science and engineering are analytic. For example, all polynomials on a finite domain are analytic. The functions $\sin x$, $\cos x$, and $\ln x$ are analytic on appropriate domains. So, although the notion of being analytic seems stringent on the surface, most of the functions we commonly encounter are analytic on some portion of their domains.

One of the nice properties of all analytic functions is that *all* of the derivatives of an analytic function are bounded. In other words, no derivative of an analytic function can grow to infinity anywhere in the domain where the function is analytic. This is strongly suggested by the fact that all analytic functions have convergent Taylor series (i.e., any derivative that approached infinity would prevent the Taylor series from converging to a finite value). Nonetheless, the following theorem is stated without proof.

Theorem 2.14 (Derivatives of analytic functions). *Let x_0 be a point, and r a positive value. Let a function, f be analytic in an interval around the point x_0; that is, for some positive r, the function f is analytic in $\{I : x_0 - r < x < x_0 + r\}$ Then there exists a positive real number M, such that for every $n \in \mathbb{N}$*

$$|f^{(n)}(x)| \leq \frac{Mrn!}{(r - |x - x_0|)^{(n+1)}}$$

In other words, the derivatives of f, $f^{(n)}$, are bounded in the interval, and the derivative cannot grow arbitrarily large within the interval.

下一个例子将说明所有多项式都是解析函数。

Example 2.19 (Polynomials are Analytic.). Suppose we have a polynomial on an open interval other 大于 $I = (-\infty, \infty)$. The polynomial is analytic everywhere in its domain.

This is not a proof, but it could easily be turned into on. Instead, consider a concrete polynomial, for example

$$f(x) = x^3 + 2x^2 + 2x + 1$$

defined on the *open* interval $0 < x < 1$. Suppose we want to determine the Taylor series for this polynomial around the point $a = 0$. We can expand this polynomial in a Taylor series around the point $x = a = 0$ to determine the value of the function at the point $x = (a + \Delta x)$. To do so, we first need to compute all of the derivatives of f. This seems as though it might be nearly impossible, until we note

$$\begin{aligned}
f(x) &= x^3 + 2x^2 + 2x + 1 \\
f'(x) &= 3x^2 + 4x + 2 \\
f''(x) &= 6x + 4 \\
f'''(x) &= 6 \\
f^{(4)}(x) &= 0 \\
f^{(5)}(x) &= 0 \ldots
\end{aligned}$$

现在，根据该公式，泰勒级数可表示为

$$f(a + \Delta x) = f(a) + \Delta x \frac{df}{dx} + \frac{1}{2!}(\Delta x)^2 \frac{d^2 f}{dx^2} + \frac{1}{3!}(\Delta x)^3 \frac{d^3 f}{dx^3} + \ldots$$

Noting that, for this example, $\Delta x = x - a = x - 0 = x$, then we have

$$\begin{aligned} f(0) &= 1 \\ f'(0) &= 2 \\ f''(0) &= 4 \\ f'''(0) &= 6 \\ f^{(4)}(0) &= 0 \\ f^{(5)}(0) &= 0 \ldots \end{aligned}$$

$$f(x) = f(0) + x \left.\frac{df}{dx}\right|_{x=0} + \frac{1}{2!}x^2 \left.\frac{d^2f}{dx^2}\right|_{x=0} + \frac{1}{3!}x^3 \left.\frac{d^3f}{dx^3}\right|_{x=0} + \ldots$$

Substituting the derivatives above yields

$$f(x) = 1 + x(2) + \frac{1}{2!}x^2(4) + \frac{1}{3!}x^3 6 + 0 + 0 + 0 + \ldots$$

Simplifying terms, we find

$$f(x) = x^3 + 2x^2 + 2x + 1$$

which was the original polynomial. Thus, this polynomial is its own Taylor series! This is actually true for all polynomials, and it is not hard to prove using the principle of induction. The ancillary conclusion that can be reached is that all polynomials defined on a (open, non-infinite) domain are analytic functions.

Example 2.20 (Partial Sums Example.).

In the previous example, the polynomial $f(x) = x^3 + 2x^2 + 145x - 1$, when expanded around $x = 0$, had the following partial sums

$$\begin{aligned} f_0(x) &= 1 \\ f_1(x) &= 2x + 1 \\ f_2(x) &= 2x^2 + 2x + 1 \\ f_3(x) &= x^3 + 2x^2 + 2x + 1 \\ f_4(x) &= x^3 + 2x^2 + 2x + 1 \\ &\ldots \end{aligned}$$

It is interesting to see how the sequence $(f_1, f_2, f_3, f_4, f_5, \ldots)$ converges to the function $f(x)$ exactly in this instance. To see this, we plot each of the partial sums on $x \in [0, 1]$ giving

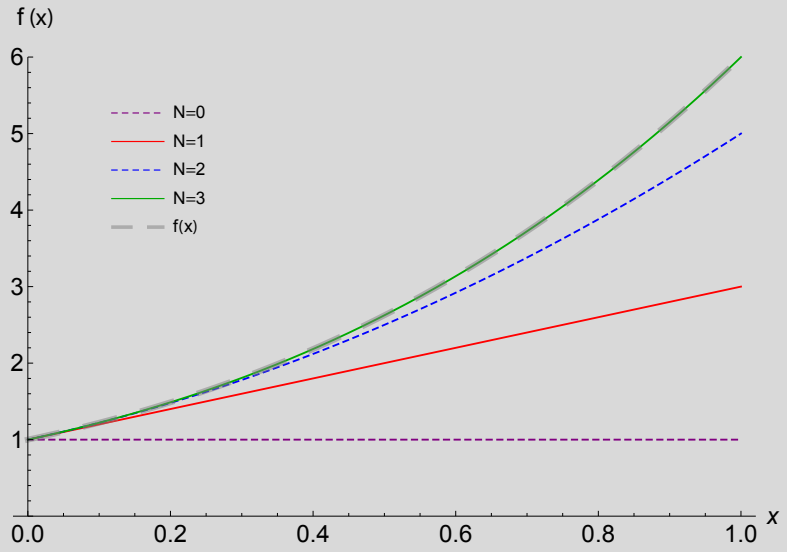


Fig. 2.24: The partial sums (Taylor polynomials) of the Taylor series of $f(x) = x^3 + 2x^2 + 2x + 1$.

2.6.2 Taylor Series Construction

So, we have established that the Taylor series can represent all analytic functions, and we examined a particular example of a polynomial that illustrated that the series actually did what we hoped it would. One might, at this juncture, wonder: *why does the Taylor series work?* It turns out, that this is reasonably easy to prove. Again, we will not do this as a formal proof, but rather as a construction that illustrates the method. A formal proof can be constructed from this outline.

To start, recall the fundamental theorem of calculus for the function f in the form

$$\int_{x=a}^{x=a+\Delta x} \frac{df}{d\xi} d\xi = f(a+\Delta x) - f(a) \tag{2.99}$$

where here we have used ξ as the variable of integration. This is already reasonably suggestive. Rearranging, note that we have

$$f(a+\Delta x) = f(a) + \int_{x=a}^{x=a+\Delta x} \frac{df}{d\xi} d\xi \tag{2.100}$$

Now, we use integration by parts once (setting $u = df/d\xi$, $dv = d\xi$) to give

$$f(a+\Delta x) = f(a) + (a+\Delta x)\frac{df}{dx}\bigg|_{a+\Delta x} - a\,\frac{df}{dx}\bigg|_{a} - \int_{x=a}^{x=a+\Delta x} \xi \frac{d^2 f}{d\xi^2} d\xi \tag{2.101}$$

Now note, also by the fundamental theorem of calculus

$$\int_{x=a}^{x=a+\Delta x} (a+\Delta x)\frac{d^2 f}{d\xi^2} d\xi = (a+\Delta x)\,\frac{df}{dx}\bigg|_{a+\Delta x} - (a+\Delta x)\,\frac{df}{dx}\bigg|_{a} \tag{2.102}$$

Rewriting this as

$$(a+\Delta x)\,\frac{df}{dx}\bigg|_{a+\Delta x} = \int_{x=a}^{x=a+\Delta x} a\frac{d^2 f}{d\xi^2} d\xi + (a+\Delta x)\,\frac{df}{dx}\bigg|_{a} \tag{2.103}$$

最后，将其代入式（2.101），得到

$$f(a+\Delta x) = f(a) + \Delta x \frac{df}{dx}\bigg|_{a} - \int_{x=a}^{x=a+\Delta x} (\xi - a)\frac{d^2 f}{d\xi^2} d\xi \tag{2.104}$$

重复这一过程n 次，即可得到直到n项的泰勒公式。

2.6.3 Uniform Versus Pointwise Convergence for Power and Taylor Series

Both power and Taylor series are unique to the study of series because they involve an infinite sum of functions indexed by the integers ≥ 0. Assuming, for example, that we have a single independent variable in space, these series can then be functions of both x in some interval I and $n \in \mathbb{N}_0$. More explicitly, we have a sequence of functions whose partial sums are

$$f_n(x) = \sum_{n=0}^{N} d_n (x-a)^n \tag{2.105}$$

or, assuming a Taylor series where $d_n = \frac{f^{(n)}(a)}{n!}$

$$f_n(x) = \sum_{n=0}^{N} \frac{f^{(n)}(a)}{n!} (x-a)^n \tag{2.106}$$

Thus, as far as convergence is concerned, we must consider the convergence of these series for each point x as $n \to \infty$. There are many senses (or modes) of convergence that have been adopted in mathematics. Here, the discussion will be focused on two that are particularly relvant to series representations of functions; they apply to the Taylor series described in this section as well as the Fourier series that will be introduced in future chapters.

Before stating the definitions for these two kinds of convergence, it is possible to generate some intuition about them. The convergence known as *uniform* is defined by the idea that one can make the series approximation to the function exhibit a maximum error, ε, for *every* point in the domain. In contrast, pointwise

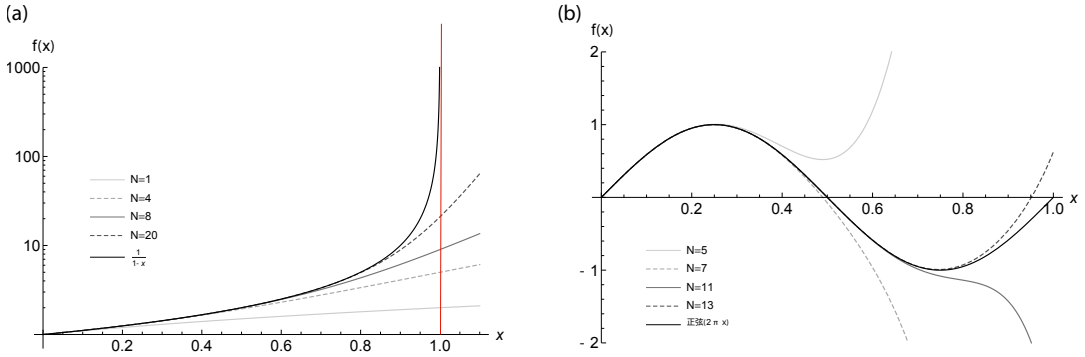


图 2.25：级数收敛类型。(a) 一致收敛的级数。(b) 逐点收敛的级数。

convergence is not as strict. It suggests that, while every point in the domain must converge, one cannot necessarily ascribe a singe error, ε, which is the maximum error. Instead, the error associated with each point must be determined independently. Before continuing, recall that a sequence of functions $f_n(x)$ (e.g., the partial sums of a Taylor series) is said to converge if

$$\lim_{n\to\infty} f_n(x) = f(x) \tag{2.107}$$

现在，我们可以给出逐点收敛和一致收敛的定义了。

定义 2.46（函数序列的逐点收敛）。一个函数序列，$f_n(x)$，如果对于某个区间 I 中的所有 $x \in I$，以及对于每一个给定的误差 $\varepsilon(x)$，我们总能找到一个N 值，使得它收敛于逐点，则称该函数序列在该区间上逐点收敛。

$$|f_n(x) - f(x)| < \varepsilon(x), \text{ for all } n \geq N \tag{2.108}$$

这里需要注意的关键一点是，所需的 N 值取决于位置 x，以及在该点选取的 ε(x) 的值。对于这种情况，无法保证存在一组适用于所有点的 ε(x) 和 N。思考这一点可能有些令人困惑，不过一个例子会有所帮助。图 2.25(a) 展示了函数 $\mathrm{f(x)} = 1/(\mathrm{x}-1)$ 在区间 $\mathrm{x} \in [0,1)$ 上的泰勒级数近似。从图中可以清楚地看出，增加部分和中的项数会使级数的近似效果更好。不过，请注意以下几点。在点 $\mathrm{x} = 1$ 处，该函数趋于无穷。没有多项式会在某个有限的 x 值处趋于无穷；因此，点 $\mathrm{x} = 1$ 附近的误差情况很复杂。虽然对于任何 $\mathrm{x} < 1$（无论我们离点 $\mathrm{x} = 1$ 多近），总能找到一个足够大的 N 值，使误差小于我们选取的 ε(x) 的任意值，但只有在指定 x 和 ε(x) 后，我们才能确定这个 N 值。不过，不存在适用于所有点的有限 N 值！例如，假设我们找到能为 $\mathrm{x} = 999/1000$ 提供 $\varepsilon < \frac{1}{10^6}$ 的最小 N 值。这个 N 值无疑会很大；然而，如果我们选择 $\mathrm{x} = 9999/10000$，它也不会给出 $\varepsilon < \frac{1}{10^6}$。因此，所需的 N 值总是可以达到的，但它也始终取决于所选的 x 和 ε(x) 的值。

—

与逐点收敛不同，一致收敛可以定义如下。

定义 2.47（函数序列的一致收敛）。一个函数序列 $f_n(x)$ 如果在某个区间 I 上一致收敛，是指对于所有 $x \in I$，以及对于每一个给定的误差 ε，我们总能找到一个 N 值，使得

$$|f_n(x) - f(x)| < \varepsilon, \text{ for all } x \in I \text{ and for all } n \geq N \tag{2.109}$$

请注意，这里收敛定义域的区别。这里可以选定一个 ε 值，并且可以保证存在足够大的N，使得级数近似与函数之间的差值对定义域中的每一个点都同时小于 ε。图 2.25(b) 绘制了一个一致收敛的函数（正弦函数）。虽然函数的泰勒多项式近似总是存在误差，但只要N 足够大，就可以使定义域I 中任意位置的误差都小到我们所希望的程度。换句话说，误差有一个具体的上界，而且这个上界只取决于N的大小，而不取决于所选的位置x。

在研究函数的傅里叶级数表示时，我们还会再次讨论一致收敛与逐点收敛这一主题。虽然这两种收敛方式之间的区别看似细微，但实际上却会产生重大影响！从实际应用的角度来看，一致收敛意味着，通常只需有限项，就能找到一个良好的函数近似，使其在研究域内的所有位置都适用。对于仅逐点收敛的函数而言，从应用角度对其进行近似则可能要复杂得多。

2.6.4 ⋆The Taylor Series in Approximation Theory

Taylor series, and power series in general, do not always converge; even if they converge, they may converge only on limited domains. For the series

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n \tag{2.110}$$

there are three possibilities for convergence, as follows.

1. The series converges only for $x = 0$ (in which case, it converges to zero there).
2. The series converges for all x.
3. 存在一个数 实数集，使得 $\sum_{n=0}^{\infty} f(n)(a)\, n!\,(x-a)^n$ 在 $|x| <$ 实数集 时收敛，而在 $x >$ 实数集时发散。对于x =实数集，该级数可能收敛，也可能发散。

Most of the useful Taylor series, as one might imagine, are convergent. This means that they either (1) converge everywhere on the real line, or (2) converge in some interval $|x| < R$. Convergent Taylor series are interesting in that they look like polynomials of infinite-order. In fact, there is no such thing as an infinite-order polynomial (the proper term there would be a power series or Taylor series), but every finite approximation of a Taylor series is some finite order polynomial. To see that, suppose that we set $c_n = f^{(n)}(a)/n!$, and that we conduct the expansion about $x = 0$ (we can do the case for $a \neq 0$, and all the following results are the same, it is just more complicated to explain). Then, the Taylor series takes the form

$$f(x) = \sum_{n=0}^{\infty} c_n x^n \tag{2.111}$$

We can make finite approximations to every Taylor series, simply by truncating it at some value N. Then, we have the partial sums discussed earlier, which we represent by

$$f_N(x) = \sum_{n=0}^{N} c_n x^n \tag{2.112}$$

These partial sums represent a *sequence of functions*, and for convergent Taylor series this sequence of functions gets to be a better and better approximation for f as N increases.

假设 we consider the first five terms in a Taylor series of the form of Eq. (2.112). A little thought will indicate that this series must take the form

$$f_5(x) = c_0 + c_1 x + c_2 x^2 + c_3 x^3 + c_4 x^4 + c_5 x^5 \tag{2.113}$$

但是，正如上文所述，这只是一个多项式！显然，若泰勒级数在某个区间上收敛，就意味着被近似的函数可以用多项式表示。这还意味着，只要在展开式中取足够多的项，我们就能让这种多项式表征的误差尽可能小（这是由收敛的定义直接得出的）。观察到的收敛类型取决于所研究的函数，不过无论如何，我们总能找到一个（有限）多项式，以任意小的误差近似该函数。

像式（2.113）这样的表达式称为泰勒多项式。这之所以有趣，是因为它实际上表明，任何具有收敛泰勒级数的函数，都可以展开为基函数，而这些基函数都是多项式。具体而言，基函数的集合为

$$1,\ x,\ x^2,\ x^3,\ x^4,\ x^5 \ldots \tag{2.114}$$

暂且假设我们用一个符号名称来称呼这些函数。令$p_0 = 1, p_1(x) = 1, p_2(x) = x^2$, 等等。。那么，函数集合 $\{p_i\}$ 足够"丰富"，这些基函数的线性组合能够近似任何具有收敛泰勒级数的函数。根据上文中展开为五项的例子，我们得到的线性组合为

$$f_5(x) = c_0 p_0(x) + c_1 p_1(x) + c_2 p_2(x) + c_3 p_3(x) + c_4 p_4(x) + c_5 p_5(x) \tag{2.115}$$

This is quite an amazing result. In essence it says that every function with a Taylor series that converges on some interval can be approximated to any accuracy that we like by some polynomial! An example will be helpful here.

例 *2.21*（通过泰勒级数进行多项式近似。）

For this example, we consider the familiar function $f(x) = e^x$. The Taylor series for this function (computed for a=0) is easily calculated, and fairly well known.

$$e^x = 1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3 + \frac{1}{24}x^4 + \frac{1}{120}x^5 + \ldots \tag{2.116}$$

Once nice property of the Taylor series for e^x is that it converges for every x on the real line. It is interesting to see how well different polynomials provide estimates for e^x. Suppose we consider the

interval $0 \leq x \leq 3$. In the plot below, the first 6 Taylor polynomials are plotted for the exponential series. To be clear, the polynomials are explicitly listed below. Note that in each case, the function is a polynomial of the form $\sum_0^N c_i p_i(x)$, where the $p_i(x)$ are the polynomial basis functions, $p_0(x) = 1$, $p_1(x) = x$, $p_2(x) = x^2$, etc.

$$
\begin{aligned}
f_0(x) &= 1 \\
f_1(x) &= 1 + x \\
f_2(x) &= 1 + x + \frac{1}{2}x^2 \\
f_3(x) &= 1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3 \\
f_4(x) &= 1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3 + \frac{1}{24}x^4 \\
f_5(x) &= 1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3 + \frac{1}{24}x^4 + \frac{1}{120}x^5
\end{aligned}
$$

From this plot, it is clear that the sequence of Taylor polynomials are functions, and this sequence of functions appears to converge to the function $f(x) = e^x$ (as we had expected). Additionally, we see that as the order of the Taylor polynomial increases, the approximation improves. For the Taylor polynomial $f_5(x)$, the fit is quite good over the entire range plotted, with the maximum error being only about 8%.

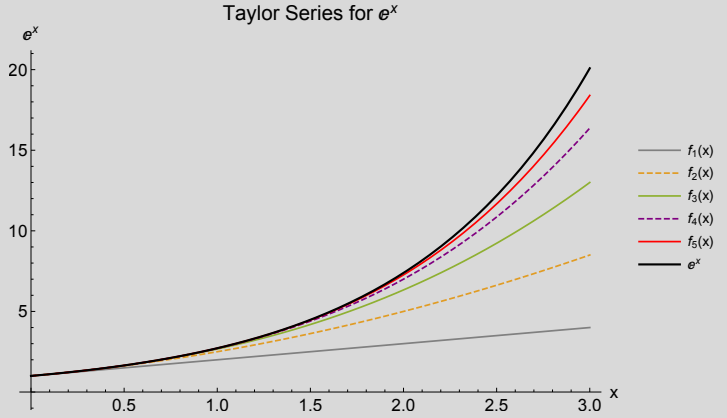


Fig. 2.26: The first six Taylor polynomials for the function $f(x) = e^x$ plotted for the interval $0 \leq x \leq 3$.

It is really quite remarkable to realize that any analytic function can be approximated by a polynomial to any accuracy that we might desire. The example of Taylor polynomials provides two interesting concepts worth reflecting on.

1. Functions can be created by linear combinations of an entire set of other functions called *basis functions*. Here, the basis functions were the polynomials $p_0(x) = 1$, $p_1(x) = x$, $p_2(x) = x^2$. The linear combination was formed by taking the appropriate weighting of these functions, and summing them up. This is done

automatically during the process of computing the Taylor series. However, it is useful at this juncture to realize that we can take any analytic function and, using a Taylor series, *decompose* this function into the weighted sum of polynomial basis functions. We will return to this idea of decomposing functions in terms of basis functions when we discuss Fourier series.

2. 其次，我们还涉及了数学中所谓的逼近理论。逼近理论关注的是如何用基函数的线性组合来表示一个函数，就像我们用泰勒多项式所做的那样。关于用多项式逼近函数，有一个更强的论断，称为魏尔斯特拉斯逼近定理。它同样提出用多项式来逼近连续函数，但对于这些近似如何收敛到函数这一行为，它能够作出比泰勒级数通常所能给出的更强论断。

这段讨论涉及了逼近理论。在某些意义上，更广义的逼近理论构成了现代实分析（以及复分析！）的基础（也就是说，几乎涵盖了应用数学中所有可以想象的主题）。本文后文将介绍傅里叶级数。傅里叶级数是整个应用数学中最重要的工具之一；归根结底，它也是对有界区间上函数的逼近理论的研究。学习数学时，寻找规律十分重要。一种类型中已充分理解的方法（例如，使用泰勒级数的逼近理论），在研究具有相似结构的新问题时（例如，使用傅里叶级数研究有限区间上的逼近理论），可以成为极其有用的模拟物。

2.7 *泛函与积分变换

In this section, a concept that has already been "seen" by most students is re-evaluated for the purposes of later use. In particular, this section is concerned with the notion of *functionals*, which, in short, are just functions of functions. Recall that a *function* uniquely maps every element of a set A (the domain) to a single element in set B (the range). We usually think of the set A as being some portion of the real line (or of $\mathbb{R}^2$, or the complex plane, etc.) defined by a number. Similarly, we usually think of the output of the function as being a number; thus a function maps numbers (or vectors of numbers in multiple dimensions or for complex numbers) to numbers.

For a *functional*, the same definition is true, except both the input set, A, and the output set, B, are not numbers, but entire *functions*. The following is a definition

定义 2.48（泛函）。一个函数，其定义域是一个函数集合，而其值域是另一个函数集合（包括常函数），则称为泛函（参见 Courant 和 Hilbert，1953 年，第 IV 章）。

Some examples will make the notion of a functional much clearer.

Example 2.22 (A simple functional.).

Suppose we consider the set of all possible functions, $f(x;a) = ax^2$, where a is some constant parameter, and $x \in [-1,1]$. In this notation, we have indicated the parameters associated with a function

by listing them (preceded by a semi-colon) with the independent variables. This is a common notation that is used when the parameters of a function might also be considered to vary.

We can define the following functional, $F[f(x;a)]$ that relates each function f to a single real number as follows.

$$F[f(x;a)] = \int_{x=-1}^{x=1} ax^2\,dx$$

In this case, we can even determine the particular real number that each function, f, is mapped to by the functional F.

$$\begin{aligned} F[f(x;a)] &= \int_{x=-1}^{x=1} ax^2\,dx \\ &= a\,\frac{x^3}{3}\Big|_{x=-1}^{x=1} \\ &= \frac{2}{3}a \end{aligned}$$

The action of the functional, F is now clear: it maps every function of the form $f(x) = ax^2$ *uniquely* to the number F[f(x;a)] = 2/3 a. The mapping is unique because each value of a defines a different member of the set of parabolas we have defined, and each such parabola is linked to a real number that depends only on a.

According to our definition, functionals can also map *complete functions* to *new functions*. An example here can be seen by defining the conventional moving-box average.

Example 2.23 (A functional resulting in a new function: Moving average).

Suppose we define the following box function

$$B(y) = \begin{cases} 2 & 0 \le y \le \frac{1}{2} \\ 0 & otherwise \end{cases}$$

Now, for any function f, we can define the following averaging operation given by a functional, F

$$F[f(x),x] = \int_{y=-\frac{1}{4}}^{y=\frac{1}{4}} f(x+y)B(y)\,dy$$

This functional has been constructed to (1) for any point x, take the (uniformly weighted) average of the function $f(x)$ between $x-\frac{1}{4}$ and $x+\frac{1}{4}$, and (2) assigns this new value to the point x. (Note: this means that the function needs to be defined on the interval $x-\frac{1}{4} < x < x+\frac{1}{4}$). Thus, the function assigns a uniformly weighted average of the function near x to to the point x. To make this a bit more clear, the following figures can help. First, note that the averaging function is plotted in Fig. 2.27.

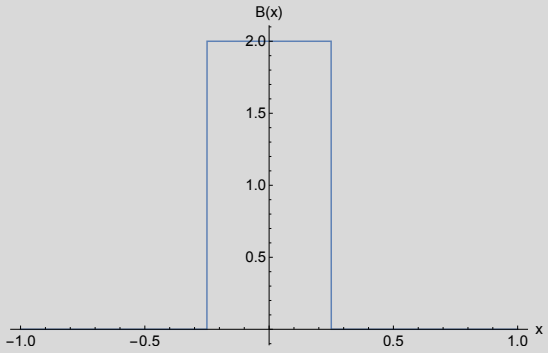


图 2.27：平均函数；一个简单的均匀盒，其面积等于 1。在图 2.28 中，我们绘制了一个受到随机噪声影响的大致线性函数。通常，若要从函数中去除噪声，可以将平均函数作用于含噪函数，以平滑随机波动。在本例中，我们将使用上面定义的盒函数。

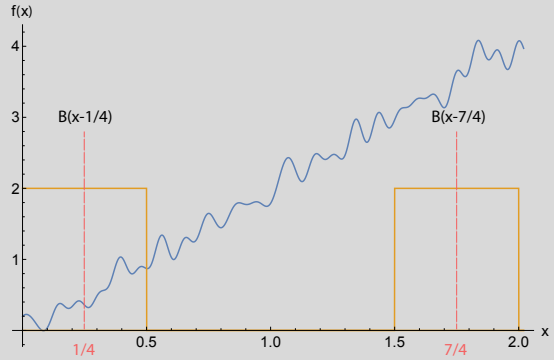


图 2.28：区间 $x \in [0,2]$ 上待平滑处理的含噪函数（蓝色）。图中绘出了盒函数 $B(x-1/4)$ 和 $B(x-7/4)$ 以供参考。请注意，在本例中，泛函 $F(x)$ 仅在 $\frac{1}{4} \le x \le \frac{7}{4}$ 上有定义，因为否则盒函数将超出该函数的定义域。

Upon applying the moving box average via the functional F defined above, we have as a result a smoothed version of our function. This smoothed function can be observed in Fig.

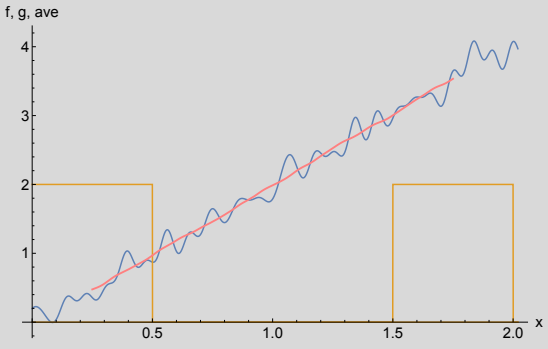


Fig. 2.29: A noisy function (in blue) on the interval $x \in [0,2]$ to be smoothed. The line in red indicates the smoothed function. Note that the smoothing is only defined between $\frac{1}{4} \leq x \leq \frac{7}{4}$, because it is only within these bounds that the averaging function is within the domain of the noisy function $f(x)$.

This second example is one where the functional specifically takes the form

$$g(x) = F[f]|_x = \int_{y=a}^{y=b} K(x,y) f(y)\, dy \tag{2.117}$$

这里使用了另一种符号表示（$F[f]|_x$），以表示<函数>结果泛函的自变量。这不应与函数 f 的自变量混淆（在本例中，该自变量由积分号下的变量 y 表示）。这类积分变换在应用数学中得到了广泛应用。在这种用法中，函数 $K(x,y)$ 通常称为积分的核函数 。由带有核函数运算的积分所定义的泛函，最著名的用途之一就是傅里叶变换和拉普拉斯变换。这两种积分变换方法将在后续章节中详细介绍。

习题

1. 使用集合构造表示法写出以下陈述：“构造由整数中小于数值三的数 n 组成的集合”。2. 使用集合构造表示法写出以下陈述：“构造由整数中小于数值三且大于负二十七的数 n 组成的集合”。3. 完全平方数是自然数的平方；我们假设 0 是完全平方数。使用集合构造表示法，构造所有完全平方数（包括零）组成的集合的表达式。4. 假设我们希望定义区间 I $=$ (a,b]。使用集合构造表示法，以 I $=$ {x : ... 的形式为该区间构造集合表达式。5. 假设 $I_1 = (0,1]$ 且 $I_2 = [1,2]$。现在定义 $I_3 = I_1 \cup I_2$。使用集合构造表示法，以 $I_3 =$ {x : ... 的形式描述 I_3。6. 假设 $I_1 = (0,1]$ 且 $I_2 = [1,2]$。现在定义 $I_3 = I_1 \cap I_2$。使用集合构造表示法，以 $I_3 =$ {x : ... 的形式描述 I_3。7. 假设 $I_1 = (0,1]$ 且 $I_2 = (1,2]$。现在定义 $I_3 = I_1 \cap I_2$。使用集合构造表示法，以 $I_3 =$ {x : ... 的形式描述 I_3。8. 求一个 2×2 矩阵的行列式相对直接。考虑矩阵

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

where a_{ij} indicates (by convention) the value of the ith row and the jth column.
现在，执行以下操作：(i) 使用行变换将矩阵化为上三角形式；(ii) 根据三角矩阵对角线上的产物计算行列式。结果应为 det(**A**) $= a_{11}a_{22} - a_{12}a_{21}$。

9. As long as we have computed the determinant for a 2×2 matrix, we may as well learn about the solution to a set of two equations in two unknowns. There is a method, known as *Cramer's rule* that allows you to find the solution to *any* set of linear equations as long as you can find the appropriate determinants. It turns out that for an $n\times n$ system of equations, you need to find n determinants. In general, this is not so easy to do; for small matrices, however, it can be very convenient.
For two equations in two unknowns, the coefficient matrix, **A** is 2×2. For a 2×2 system of the form

$$\mathbf{A}\cdot\vec{x} = \vec{b}$$

or

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

Cramer's rule says that the solution can be found from the following determinants where the right-hand side has been substituted into the columns of **A**

$$x = \frac{1}{\det(\mathbf{A})}\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix} \qquad y = \frac{1}{\det(\mathbf{A})}\begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}$$

where both $\det(\cdot)$ and $|\cdot|$ indicate the determinant. Suppose we have the set of equations

$$2x + 4y = 2$$
$$2x + 2y = \frac{3}{2}$$

or, in matrix form

$$\begin{bmatrix} 2 & 4 \\ 2 & 2 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 \\ \frac{3}{2} \end{bmatrix}$$

Use Cramer's rule to solve for x and y.

10. Linearity is an important concept to be familiar with. For the following operators, $\mathscr{L}(\cdot)$, determine if they are linear or nonlinear

a. $\mathscr{L}(x) = ax^2 + bx$b. $\mathscr{L}(x) = \sqrt{a^2 x}$c. $\mathscr{L} = d^2 dx_2 \mp a\, ddx$ ，因此 $\mathscr{L}(f) = d^2 f$ $dx_2 + a^d \underline{f}\, dx$d. $\underline{\mathscr{L}(f) = \int x f(x) dx}$

11. 积分有两种方法：使用不定积分，或使用定积分。微积分基本定理将两者联系起来。记住微积分基本定理的一种方法，是用两种方法计算曲线下面积。请试试看。对于函数 $f(x) = x$ 在区间 $x \in [1,2]$ 上，用两种方法计算面积。

a. 计算不定积分。结果会包含一个未知常数；将它记作C。当我们指定要计算哪个区间的面积时，C 的值便确定了。假设我们要计算以下两个区间的面积：(i) $0 \leq x \leq 1$，以及 (ii) $0 \leq x \leq 2$。通过画出函数的草图，然后利用简单几何方法计算相应区域的面积，求出区间 (i) 和 (ii) 上曲线$f(x) = x$ 下面的面积。b. 现在，按以下步骤计算这个问题中不定积分的数值。你知道每个区间所对应的准确面积，也已经在 (a) 部分求出了带未知常数C 的不定积分。对于区间 (i) 和 (ii)，求出未指定常数 C 的值。两个区间算出的值应该相同。c. 注意，通过将上述两个面积相减，我们可以得到区间 $1 \leq x \leq 2 \Leftrightarrow x \in [1,2]$ 的面积。在草图上画出该区域，并根据 (b) 部分的结果计算其数值面积。

d. Finally, compute the definite integral

$$\int_{x=1}^{x=2} x\, dx \tag{2.118}$$

You should (kind of obviously perhaps at this point) get exactly the same value. The fundamental theorem of calculus states that the area for some fixed interval is just the different between two indefinite integrals, as you have shown.

12. 使用分部积分法计算下列积分（即求出一个不再含有积分的函数表达式），其中 $x \in [0,1]$。注意：为了得到最终结果，你可能需要多次使用分部积分法！

    a. $f(x) = x$正弦(πx)
    b. $f(x) = x$余弦(πx)
    c. $f(x) = x^2$ 正弦(πx)
    d. $f(x) = x^2$ 余弦(πx)

13. For the functions f and g in Fig. 2.10, consider the domain $\mathbf{x} \in [0, 1.5]$. What are the *ranges* of the two functions (approximate- estimate as best you can)? Are either of the functions piecewise continuous? Are either of the functions piecewise smooth?

14. The exponential function has the unusual property that its Taylor series converges everywhere on the real line. Given that information, is the following function *analytic* on the domain $D = [0,1]$?

$$f(x) = \exp(\pi(x-1)), \tag{2.119}$$

    Why or why not?

15. **Euler's rule**. Euler's rule, given by

$$e^{ix} = \cos x + i \sin x$$

    is actually pretty easy to prove. What you need to know to do that is the following (and that $x^0 = 1$ and $0! = 1$ by definition)

$$\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!} = \frac{x^0}{0!} + \frac{x^1}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \dots$$
$$\cos x = 1 - \frac{1}{2!}x^2 + \frac{1}{4!}x^4 - \dots$$
$$\sin x = x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 - \dots$$

    使用 t hese series, in addition to the properties of the imaginary number $i = \sqrt{-1}$ to prove Euler's rule.

16. Every polynomial is its own Taylor series. To see this, determine the Taylor series for the polynomial

$$y(x) = 3x^3 + 4x^2 + 2x + 5$$

    Around the point $x = 0$. Start by first computing $y(0)$, $y'(0)$, $y''(0)$, $y'''(0)$, and $y^{(\mathrm{IV})}(0)$. Then use the formula for the Taylor series to arrange these terms in the proper form, and compute the coefficients for each power of x.

17. Consider the following function, defined piecewise over the interval $0 \le x \le 2$.

$$f(x) = \begin{cases} 0 \text{ for } 0 \leq x \leq 1 \\ x \text{ for } 1 < x \leq 2 \end{cases}$$

a. Is this function continuous on the interval $x \in [0,2]$?

b. Is this function C^1 continuous on the interval $x \in [0,2]$?

c. Can this function be described as piecewise C^1 continuous? Why or why not?

d. Does this function have a conventional derivative defined for all $x \in (0,2)$?

18. 考虑以下在区间 $0 \leq x \leq 1$ 上分段定义的函数。

$$f(x) = \begin{cases} 2 \text{ for } 0 \leq x \leq \frac{1}{2} \\ -2 \text{ for } \frac{1}{2} < x \leq 1 \end{cases}$$

a. 该函数在区间 $x \in [0,1]$ 上连续吗？ b. 该函数在区间 C^1 上连续吗？ c. 该函数能否描述为分段C^1 连续？为什么能或不能？ d. 该函数可积。要对它积分，只需将积分定义域分成两部分，一部分覆盖 $x \in [0, \frac{1}{2}]$，另一部分覆盖 $x \in (\frac{1}{2}, 1]$。绘制该函数的积分关于 x 的函数图，其中 x 表示积分的上限。你需要将问题分成两种情况：一种是 $x \leq \frac{1}{2}$，另一种是 $x > \frac{1}{2}$。

19. 请用自己的话（也可以配图）描述你对微分 dy的理解。它与导数$y'(x)$有什么关系？其中x 是自变量。

20. 以下哪些函数是齐次的？请展示你的解题过程。

a. $f(x) = x$ b. $f(x) = x^2$ 正弦(πx) c. $f(x) = x^{\pi}$ d. $f(x) = \sqrt{5}x$ e. $f(x) = x^2\sqrt{x}$

21. Does the following series converge? Which test for convergence can be used to establish the convergence or divergence of this series?

$$P(n) = \sum_{n=1}^{\infty} \frac{1}{n^2} \tag{2.120}$$

22. Is the following function analytic at $x = 0$? Is it continuous at $x = 0$? Assume that the domain is $-\infty < x < \infty$.

$$f(x) = \exp\left(-\frac{1}{x^2}\right) \tag{2.121}$$

23. 泰勒级数是所谓渐近展开的核心。对于这类展开，当展开针对的x 值远小于收敛半径R时，仅需少数几项，这些级数通常就能在有限定义域内对函数给出非常精确的表征。

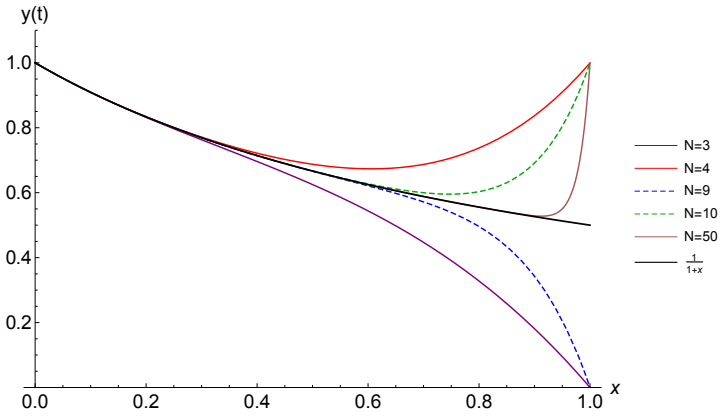


图 2.30 在区间 $x \in [0,1]$ 上绘制的 $f(x) = 1/(1+x)$ 的若干泰勒多项式。

一个常用的渐近级数示例是用于该函数的级数。

$$f(x) = \frac{1}{1+x},\ 0 \le x < 1 \tag{2.122}$$

此问题的收敛半径为 $|x| < 1$。
对于此问题，请完成以下各项。

a. 求该函数的一般泰勒级数。注意：此函数的收敛半径为 $R < 1$，且在边界 $R = 1$ 处不收敛。换言之，此级数在 $|x| < 1$ 时收敛。b. 明确计算该级数的前五项（即求出它的泰勒多项式）。使用 Mathematica（或任何方便的软件）绘制级数近似 $f_5(x)$ 和原函数 $f(x) = 1/(1-x)$ 在区间 $x \in [0,1]$ 上的图像。c. 此级数在 $|x| < 1$ 时收敛。下图展示了 $N = 3, 4, 9, 10$ 和 50 对应的泰勒多项式。无论 N 多大，每个多项式的值都是 1 或 0，取决于最后一项是偶数项还是奇数项。因此，无论计算多少项，该级数在 $x = 1$ 处都不收敛。可以证明，该级数在所有小于 1 的 x 值处确实收敛。该级数在定义域 $x \in [0,1$ 中的收敛是一致收敛 还是逐点收敛？请用一句话说明你的选择。

24. **The Divide and Conquer method for finding the square root.** The ancient Babylonians had a method to compute the square root of any positive real number, a. It was an iterative algorithm that started with an initial *guess* of the root, and then provided the method to find the exact root. Mathematically, what they did was to develop the following *implicit* relationship for the square root.

$$
\begin{aligned}
x^2 &= a && (by\ definition)\\
2x^2 &= x^2 + a && (add\ x^2\ both\ sides)\\
2x &= x + \frac{a}{x} && (divide\ by\ x)\\
x &= \frac{x + a/x}{2} && (solve\ for\ x,\ implicit\ equation)
\end{aligned}
$$

Here, a is the starting number, and x is the square root being sought.
Given an initial guess, x_1, then a better estimate could be found by

$$x_2 = \frac{x_1 + a/x_1}{2} \tag{2.123}$$

or, in general

$$x_{i+1} = \frac{x_i + a/x_i}{2}, \qquad i = 1,2,3,\ldots (or\ i \in \mathbb{N}) \tag{2.124}$$

使用此方法为 $\sqrt{2}$ （即一个 $= 2$）构造一系列估计值，从初始猜测 $x_1 = 1$ 开始。你可以使用计算器（或电子表格）完成计算，但请展示每一步的结果。还要计算每一步估计值 i 的相对误差，$\varepsilon_i =$ (估计值$_i$ −实际值)/实际值，其中估计值$_i$ 是上述算法第 i^{th} 步计算出的值，而实际值是使用计算器或电子表格计算出的值。当相对误差小于 $1/100$ 时，停止算法。

“将数学问题划分为线性问题和非线性问题，就好比将整个宇宙划分为香蕉和非香蕉。”

—Author Unknown.

3

一阶和二阶常微分方程

3.1 Terminology

在学习常微分方程（ODEs）时，有一些定义和术语值得提前明确。其中一些内容在复习章节中已经出现过。学习过程中重复讲解某些概念实际上并非多余！我们大多数人并不能仅仅通过看一遍就掌握某个知识点。一个概念或定义越是基础、越是重要，就越值得反复讲解，最好从略有不同的角度进行阐述，以适应人们对该主题的不同思维方式。

• **定义域**。 某个定义域 是可为特定应用选取的可能值的集合。定义域可以表示为离散值的集合，也可以表示为连续区间。不属于定义域的值通常对相关应用无效。例如，假设进行了一项实验，在不同的时间$t1$,$t2$,$t3$,$t4$测量了四个温度值（例如，分别用$T1$,$T2$,$T3$,$T4$表示）。那么，值$t1$,$t2$,$t3$,$t4$就表示所收集实验结果的定义域。

• **值域**。 值域 是与某个应用中的特定定义域相关联的可能值的集合。例如，上文所述温度实验的值域就是测得的温度：$T1$，$T2$，$T3$，$T4$。这些温度中的每一个都与定义域中的一个点唯一对应。

• **自变量**。 自变量是指在对某个过程进行数学描述时，代表原因或输入的变量。原则上，这些变量不依赖于任何其他量——换言之，我们可以从任何有效的定义域中自由选取它们。在上文的温度实验示例中，自变量是用于测量温度的时刻：

t_1, t_2, t_3, t_4。所选的测量时刻仅受过程本身之外因素的限制。这些因素可能是便利性，也可能是物理条件（例如，如果系统在 5 分钟后就达到稳定状态，人们可能不愿意在 4 天内进行 10,000 次测量）。有时会将自变量与因变量列表“关联”起来，以便明确它们之间的关系。因此，在上面的例子中，可以写作$T(t)$，表示温度是时间的函数。请注意自变量与函数定义域之间的关系。

• **因变量**。指定自变量后所计算出的变量称为因变量。在实际应用中，因变量是我们所关注的变量，代表希望预测的物理量。请注意因变量与函数值域之间的关系。

• **导数**。导数是因变量随其某个自变量变化而变化的微分变化率。如果只有一个因变量和一个自变量，那么导数就对应于我们通常所说的在两个坐标轴上绘制的简单函数的斜率（纵轴表示自变量，横轴表示因变量）。在微积分中，我们都学过导数表示为

$$\frac{dy}{dt} \equiv \lim_{\Delta t \to 0} \frac{\Delta y}{\Delta t} \tag{3.1}$$

$$\equiv \lim_{\Delta t \to 0} \frac{y(t+\Delta t) - y(t)}{\Delta t} \tag{3.2}$$

请注意，尽管 $\frac{dy}{dt}$ 这种符号表示被广泛使用，但也经常使用撇号符号表示导数（见下文 3.1 节），此时导数表示为$y'(t)$。

导数的阶数仅仅表示对一个函数求导的次数。因此，当一个函数被求导两次时，记作$y''(t)$ 或 $\frac{d^2y}{dt^2}$，这也称为二阶导数。

• **Ordinary differential equation.** An *ordinary* differential equation (ODE) is one in which there is only one independent variable. All derivatives appearing in the equation are taken with respect to the single independent variable. The term ordinary is used in contrast with the term partial differential equation which may be with respect to more than one independent variable. Note that one can have more than one independent variable in the case of coupled systems of equations. For example, consider the conventional multiplicative reaction rate for fully mixed chemical reactors. Suppose there are two chemical species, y_1 and y_2 who must come in contact for the reaction $y_1 \xrightarrow{k_1} y_2$ to happen. The reactions can be represented by

$$\frac{dy_1}{dt} = -k_1 y_1 y_2 \tag{3.3}$$

$$\frac{dy_2}{dt} = k_1 y_1 y_2 \tag{3.4}$$

Although there are *two coupled* differential equations here, there is still only *one* independent variable. Thus, this set of equations are still ODEs.

- **Order of an ODE.** The *order* of a differential equation is the order of the highest derivative in the equation. Examples.

$$y'' + 2y' = 0 \qquad (second - orderODE) \tag{3.5}$$
$$y''' + 3y = 42 \qquad (third - order\ ODE) \tag{3.6}$$
$$y^{(iv)} = f(t) \qquad (fourth - order\ ODE) \tag{3.7}$$
$$\frac{dy^2}{dt^2} + \left(\frac{dy}{dt}\right)^3 = 0 \qquad (second - order\ ODE) \tag{3.8}$$

Note that in many texts, the prime notation is altered using Roman or Arabic numerals in parentheses to signify derivatives of order four or more. Thus $y''''(t) = y^{(iv)}(t) = y^{(4)}(t)$. You can see why this might be the case- writing a 7th-order derivative by $y''''''(t)$ starts to get to be a bit ridiculous.

- **Homogeneous ODE.** A *homogeneous* ODE is one in which there are no terms that do not involve the independent variable. For example, examine the following

$$y'' + 3y' + 7y = 0 \qquad homogeneous \tag{3.9}$$
$$y'' + 3y' + 7y = 5t \qquad nonhomogeneous \tag{3.10}$$
$$y''(t) + 3ty'(t) + 7t^2y(t) = 0 \qquad 齐次 \tag{3.11}$$
$$y'(t) + 7y(t) = \pi \qquad nonhomogeneous \tag{3.12}$$
$$\tag{3.13}$$

更一般地，阶为n 的线性齐次方程形式为

$$a_n(t)y^{(n)}(t) + a_{n-1}(t)y^{(n-1)}(t) + \ldots + a_1y(t) = 0 \tag{3.14}$$

其中$a_n(t)$是该常微分方程的系数（这些系数本身可能是自变量t的函数），而y 是因变量。

- **线性** **常微分方程**。如果一个线性常微分方程 的齐次 部分满足以下两个性质，则称其为线性常微分方程 ：

1. 如果函数$y(t)$是该常微分方程的一个解，那么将该解乘以任意常数 α后，会得到一个同样是解的函数y_2 ；换句话说，$y_2(t) = \alpha y(t)$是一个解。有些令人困惑的是，这一性质本身有时被称为代数齐次性；该概念已在第2.2节中定义并复习过。

2. 如果有两个线性无关解，分别记为y_1 和y_2, ，那么二者的线性组合为$y(t) = y_1(t) + y_2(t)$；这个线性组合也是该常微分方程的一个解。这一性质有时称为可加性。

- **非线性常微分方程**。任何不是线性的常微分方程都是非线性常微分方程。

- **解析函数**。 解析函数是指在局部范围内可以表示为收敛幂级数的函数。换句话说，对于由自变量 x 给定的任意点，函数 $f(x)$ 都可以表示为一个在某个收敛半径内收敛的幂级数。

- **Transcendental function.** A transcendental function is an analytic function that can not be specified by a (finite) polynomial. In other words, it cannot be expressed exactly by a finite sequence of algebraic operations. Such functions are often approximated by the appropriate truncation of an infinite series representation.

• **附加条件**。附加条件是指在自变量的指定值处给定因变量（或其某个导数）的条件 。

• **初始条件**。初始条件是在特定时间给定的附加条件；该时间被假定为过程的初始时间或起始时间。例如，要确定反应过程中的浓度，需要知道某一时刻的初始浓度。随后可以利用常微分方程，确定自初始条件给定以来，浓度随时间变化的情况。

• **边界条件** 边界条件是初始条件的空间对应项。例如，绝热杆中的稳态热流将产生温度分布，该温度分布是两端所规定的辅助条件的函数；这些就是边界条件。

注 *3.1*（导数的符号表示）。
导数有多种常用符号表示，这源于微积分在不同人手中几乎同时发展而形成的漫长历史。在微积分的发展过程中，最著名的两位人物（这也理所当然）是牛顿和莱布尼茨（也拼作莱布尼茨）；两人都在 1660 年代发展了微积分的基本思想。牛顿喜欢在变量上方加点来表示导数；因此，函数$y(t)$对t 的导数会表示为 $\dot{y}(t)$；或者，当表达式中自变量显而易见时，只写作 $\dot{y}$。莱布尼茨采用了一种更直观、但稍显繁琐的符号表示；他使用$\frac{dy}{dt}$来表示导数。这种符号表示的一个优点是，它始终 能清楚地表明自变量是什么！尽管这只是一种符号表示，但英国的数学家与欧洲大陆的数学家之间实际上围绕这一问题存在着显著的竞争。让情况更复杂的是，还出现了其他符号表示，包括法国数学家 Lagrange 的符号表示（他引入撇号来表示导数，如$y'(T)$），以及瑞士数学家 Euler 的符号表示（他引入$Dy(t)$来表示导数）。这一问题在 1830 年代初期得到了一定程度的（小小的）解决：当时，剑桥大学（英格兰）成立了一个自称为分析学会 的团体，专门致力于推广莱布尼茨的导数符号表示。虽然脱离上下文来看，这似乎有些荒谬，但这一转变确实有其意义。由于莱布尼茨的符号表示在欧洲各地广泛使用，英国学生难以理解欧洲数学家最新开展的工作。这被视为让英国落后于人，因此分析学会应运而生。改用莱布尼茨的符号表示有助于促进英国与欧洲之间更好的交流。

In the present day, one finds both the Lagrange prime notation, and the Leibniz notation used frequently. In addition, in some disciplines such as continuum mechanics and physics one still finds the dot notation used as a reference to the influence that Newton had on the development of these areas.

3.2 引言

As soon as one defines the derivative, it turns out that one also poses a *differential equation*. A differential equation is exactly what it sounds like: it is a mathematical expression that involves the derivative of some unknown function rather than, say, a purely algebraic expression of some unknown function. The word *ordinary* is used as a prefix to describe differential equations that contain exactly one independent variable. Differential equations that involve two or more independent variables are called *partial differential equations*; those will be the subject of later chapters.

微积分领域的德国数学家戈特弗里德·威廉·莱布尼茨因于 1675 年正式写下第一个有解的微分方程而受到赞誉（Sasser，1992）；然而，艾萨克·牛顿早在 1671 年就有了求解实例，尽管这些实例直到 1736 年才发表（Krishnachandran，2020）。莱布尼茨的领悟本质上就是本引言第一句话所阐述的内容。假设我们有一个函数 $y(x) = 1/2x^2$。该函数的导数由下列方程给出：

$$\frac{dy}{dx} = x \tag{3.15}$$

从“正向”角度思考这一点，就是求函数 导数 $y(x) = 1/2x^2$, ；此外，还有第二种理解方式。这种理解方式是将方程 (3.15) 视为一个含有未知函数 y 的方程，并对其求解。换句话说，问题是：“在其定义域内的每一点，是否都存在某个函数 $y(x)$，使得它的导数等于给定的量 $dy/dx = x$？”当然，我们知道这个问题的答案，因为我们最初就是根据对 $y(x)$ 的了解推导出它的。

虽然从反向角度看待导数问题似乎微不足道，但这种视角的转变实际上具有深远的意义。一旦接受这样一种观点，即可以通过求解以导数形式表达的方程来确定完整的函数，那么一个充满新型方程（及问题）的可能性世界便会豁然开启。我们不再只是简单地了解函数及其导数之间的对应关系，而是可以完全根据某种物理量的变化率来构建问题，然后在不一定事先知道解的情况下，着手“寻找”解。在科学和工程学问题中运用这一技能具有极强的实用价值，尤其适用于各种物理量的变化率通过某种守恒定律给出的那些问题。

我们对常微分方程的复习仅涵盖以下内容：（1）线性一阶常微分方程和（部分）非线性一阶常微分方程；（2）常系数线性二阶常微分方程。系数非常数的线性二阶常微分方程这一更一般的情形是一个有趣的课题（在学习 Sturm-Liouville 问题时，我们也会遇到这类方程），但该课题涉及级数解；这些解的求解过程可能会有些繁琐，而且并不能实质性地增加有助于深入理解常微分方程的新概念，因此我们决定在本次复习中省略这一部分。对于常系数线性二阶问题的解法有了基本了解后，有兴趣的学生只需稍微自学一下这一课题，就能相当容易地掌握如何处理系数非常数的问题。

3.3 First-Order ODEs

First order ODEs are those which contain derivatives of order one as the highest order of derivative in the problem. The most general form for a first-order ODE is (in the independent variable x)

$$\frac{dy}{dx} = f(x,y) \tag{3.16}$$

or, equivalently

$$y'(x) = f(x,y) \tag{3.17}$$

这里，未知函数 $y(x)$ 是一个函数。因此，如果用通俗英语来表述方程 (3.17)，它表达的是："y 是这样一个函数：在x的定义域内每一点，自变量x 对应的一阶导数都是函数 f(x,y)。"由于$f(x,y)$ 可以是y的任意（合理）函数，因此整个方程可以是线性的，也可以是非线性的。例如，若选择$f(x,y) = \alpha y(x) + \beta y^2(x) + x$，则可根据线性的常规判据确定，对于这一选择的$f(x,y)$，该问题是非线性的。

3.3.1 检验微分方程的解

一阶微分方程出现在许多科学家和工程师熟悉的应用中。例如，全混间歇反应器中的经典一级反应问题就是一个例子。若c表示浓度，那么我们认为反应器的时间历程$c(t)$由初始配置（初始条件）c_0, 和速率表达式共同确定。两者共同构成对系统的描述，其形式为

$$\frac{dc(t)}{dt} = -k_1 c(t) \tag{3.18}$$

$$c(0) = c_0 \tag{3.19}$$

在该表达式中，k_1 是一级反应速率常数。由方程 (3.18)–(3.19) 给出的方程组表示了化学物种 c 的质量衡算随时间t变化的一级微分方程。当问题以时间为自变量时，方程 (3.19) 通常称为初始条件；当空间为自变量时，则称为边界条件。出于我们的研究目的，有时我们也将它们统称为辅助条件，因为这些表达式提供了有关问题的附加数据，而这些数据必须通过数学之外的相关考量来确定。稍后我们将对此进行更详细的讨论。

请注意，按照惯例，变量c 称为因变量，因为其值被视为时间的函数。同样，t 称为自变量，因为其值可以自由选取。有时，为了明确这种依赖关系，会用自变量来标记因变量；因此，可以写成$c(t)$，以确保c 与t 之间的关系清晰明确。

虽然我们尚未讨论如何求解，但我们可以回想一下，这个问题的解是时间的指数函数

$$c(t) = c_0 exp(-k_1 t) \tag{3.20}$$

通过注意以下几点，我们可以检验这是否为一个解。

1. Taking the derivative of both sides of Eq. (3.20) and then substituting the result for both the derivative and the solution back into Eq. (3.18), we can determine that the equation is *valid*.

2. We also note that at $t = 0$, we have that $exp(0) = 1$, so that $c(0) = c_0$ and thus meets the initial condition. This verifies that the solution given by Eq. (3.20) is both a solution, and meets the ancillary conditions required for the solution.

下面将详细说明这一过程中的各个步骤。

例 *3.1*（判断给定的解是否有效）。你通常会想要检查一个提出的解是否确实有效——换句话说，它是否同时满足辅助条件（初始条件或边界条件）和常微分方程本身。这个解可能是给定的，也可能是你自己推导出来的。解题时，最好按照以下步骤检查解，以确认解是正确的！无论常微分方程和辅助条件是什么，检查解的步骤都相同。对于具有一个指定初始条件的一级反应问题，步骤如下。

•**确定解所需的导数**。在我们的例子中，方程是一阶的，因此只需要一阶导数。

$$c(t) = c_0 exp(-k_1 t) \tag{3.21}$$
$$c'(t) = -k_1 c_0 exp(-k_1 t) \tag{3.22}$$

• **将解及其导数代入常微分方程**。将解及其一阶导数（如上所求）代入常微分方程，我们得到

$$-k_1 c_0 exp(-k_1 t) = -k_1 c_0 exp(-k_1 t) \tag{3.23}$$

Clearly, the two sides of this equation are equal, and this verifies that the proposed solution is indeed correct.

• **确定是否满足辅助条件**。在这种情况下，我们只有一个辅助条件（初始条件）。将初始条件中指定的时间（$t = 0$）对应的自变量值代入，我们得到

$$c(t = 0) = c_0 exp(-k_1(0)) \tag{3.24}$$
$$c(t = 0) = c_0 \tag{3.25}$$

Clearly, the ancillary condition is met for the time specified. Since ODE is validated and the specified condition is met, then Eq. (3.20) is validated as being the particular solution we are looking for.

3.3.2 可直接积分的一阶常微分方程

Integrable linear first-order ODEs represent an interesting case because of their direct link to calculus. You may recall from your first course in calculus the much lauded (but rarely remembered) *fundamental theorem of calculus*. There are a number of presentations for the theorem; two of them are given below.

Theorem 3.1 (Fundamental theorem of calculus). *Suppose f is a continuous function on any open interval I of the real line. Then, there exists a function F defined by*

$$F(t) = \int_x f(x)\,dt + C \tag{3.26}$$

called the antiderivative *or* indefinite integral *of f on the interval I. Here, C is usually called a* constant of integration*. The function F has the property that*

$$F'(t) = f(t) \tag{3.27}$$

for all values of t in the interval I.

The interesting thing to note here is that Eq. (3.27) is actually a first-order ODE! Therefore, one interpretation of the fundamental theorem of calculus is that it defines the relationship between the integral and the derivative by solving the first-order ODE given by Eq. (3.27).

The primary feature of the fundamental theorem of calculus is that it establishes the idea that differentiation is the inverse operation to integration (hence, the term *antiderivative* is often used for indefinite integrals). In a more applied context, there is a corollary that is perhaps more familiar; this corollary relates the indefinite integral to the definite integral (i.e., the area under a curve in a closed interval).

Corollary 3.1 (The definite integral). *If f is a continuous function on the closed interval $[a,b]$, and F is its antiderivative, then the area under the function f in the interval is given by a quantity called the* definite integral*, and it is defined by*

$$\int_{t=a}^{t=b} f(t)\,dt = F(b) - F(a) \tag{3.28}$$

With this tool in hand, we are now ready to consider the solution to the first case of ODEs. The simplest first-order ODEs are linear, directly integrable problems. For these problems, the function f depends on *only* the independent variable. We use the terminology $f(y,t) = Q(t)$

$$\frac{dy(t)}{dt} = Q(t) \tag{3.29}$$

Note that, although we are using the independent variable t here, this variable could stand for anything (not necessarily time!). In applications, we will often use the variable t when dealing with problems where the independent variable is meant to represent time, and the variables x in problems where it represents space. However, in the study of ODEs it is understood that there is only *one* independent variable, and the particular symbol used to represent that variable has no inherent significance on its own.

求解方程 (3.29) 是微积分基本定理的一个直接应用。将方程两边同时除以 $P(t)$，并对两边积分，可得

$$\int \frac{dy(t)}{dt}\,dt = \int Q(t)\,dt \tag{3.30}$$

由微积分基本定理可立即得到结果

$$y(t) = \int Q(t)\,dt + C_1 \tag{3.31}$$

Where here, we have chosen to put the constant of integration, C_1, on the right-hand side of the result (N.B., the sign of the arbitrary constant at this juncture is somewhat irrelevant; when it is evaluated on the basis of an ancillary condition, the correct sign will be imposed).

Example 3.2 (First-order separable equations). The conventional problems of computing the position, velocity, and acceleration of an object are usually expressed by systems of first-order, separable ODEs. Recall from your physics classes that, if position is expressed as a function of time, $x(t)$, we have:

$$v(t) = \frac{dx(t)}{dt} \qquad a(t) = \frac{dv(t)}{dt} = \frac{d^2x(t)}{dt^2}$$

We can use these relationships to solve interesting dynamical problems. Consider the following. Suppose we want to determine the velocity as a function of time for a skier starting from a standstill, and accelerating down slope as shown in the figure. We also want to know how long in seconds it takes to get to the bottom of the slope, and the velocity of the skier when they get there.

Assumptions. It is always a good idea to *list your assumptions* when solving problems. That way, it is clear under what circumstances the result you obtain is valid. It also provides you with something to look back on for problem improvement/refinement should your result come out less reasonable than you would have hoped. For this problem, we will make some *very rough* approximations that may not be altogether reasonable. These are (1) the resistance between the skis and the snow will be neglected, and (2) air resistance will be neglected. Note: if we wanted to include such effects, we might hope to find an "overall" drag term that takes the form $drag\ force = \alpha v^2(t)$.

To start the problem, we do a simple force balance on the skier, noting that $v(t=0) = x(t=0) = 0$. The acceleration in the down slope direction is found to be

$$F = ma(t) \Rightarrow a(t) = F/m$$

But, we can compute the force to be

$$F = mg\sin\theta$$

Combining these, we have an expression for the acceleration

$$a(t) = F/m = g\sin\theta$$

请注意，在这种情况下，加速度实际上是一个常数。为了求出速度，我们只需进行积分。该积分可以通过定积分（将初始条件纳入积分上下限）来计算，也可以通过不定积分（将积分常数通过初始条件确定，并将其作为约束条件）来计算。

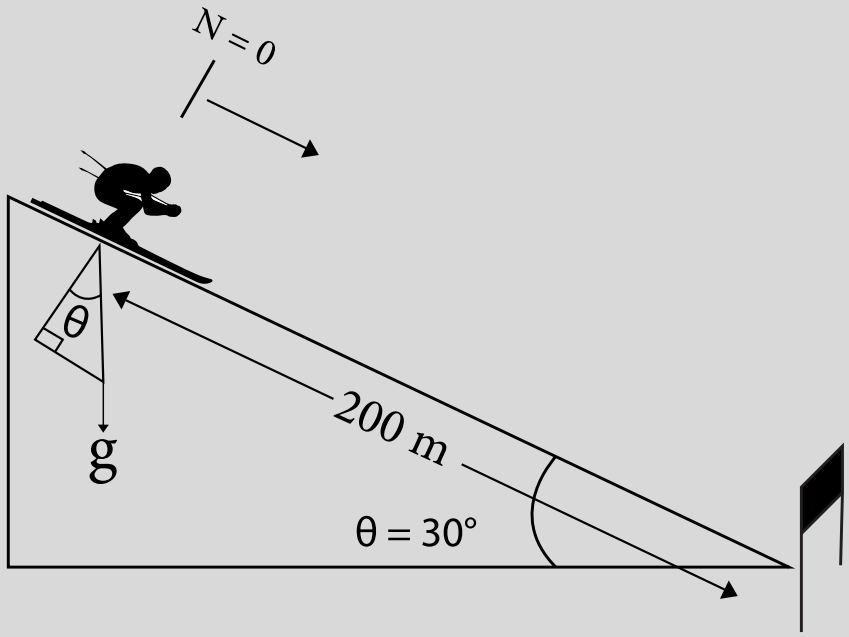


Fig. 3.1: A skier.

$$\text{Recalling} \quad a = \frac{dv(t)}{dt}$$

Upon substituting for a

$$\begin{aligned} \frac{dv(t)}{dt} &= g\sin\theta \\ \int_t \frac{dv(t)}{dt}dt &= \int_t g\sin\theta dt \\ v(t) &= gt\sin\theta + C_1 \end{aligned}$$

Using the fact that $v(0) = 0$, then $C_1 = 0$.

$$v(t) = gt\sin\theta$$

位置作为时间函数可通过类似的积分求得。

$$\text{Recalling} \quad v(t) = \frac{dx(t)}{dt}$$

Upon substititing for $v(T)$

$$\begin{aligned} \frac{dx(t)}{dt} &= gt\sin\theta \\ \int_t \frac{dx(t)}{dt}dt &= \int_t gt\sin\theta dt \\ x(t) &= \tfrac{1}{2}gt^2\sin\theta + C_2 \end{aligned}$$

Using the fact that $x(0) = 0$, then $C_2 = 0$.

$$x(t) = \tfrac{1}{2}gt^2\sin\theta$$

or, solving for t

$$t = \left(\frac{x(t)}{\frac{1}{2}g\sin\theta}\right)^{\frac{1}{2}}$$

With this last expression, we can solve for the time given the final position $x(t_f) = 200$m. Using $g = 9.81$ m/s, the result is $t = 9.0$ s. From the expression for $v(t)$, we can find the velocity at the end of the run to be 44 m/s (equivalent to 99 miles per hour!).

当然，这将是在坡度为 30° 的 200 米长滑雪坡道上所能达到的绝对最大值。需要注意的是， 30° 的坡度相当于 67% 的坡度；这实际上相当陡峭！

3.3.2.1 一阶可分离方程

提醒一下，一阶常微分方程的最一般形式为

$$\frac{dy}{dt} = f(y(t),t) \tag{3.32}$$

where $f(y(t),t)$ can be *any* linear or nonlinear function of $y(t)$. In general, there is no solution to all problems of this form. However, for a reasonably common subset of this equation, solutions do exist; for example, if f is a linear function of y then solutions are certain to exist. For example

$$\frac{dy}{dt} = y(t) \tag{3.33}$$

is a linear problem. It represents the statement that we are looking function whose derivative is the same as the function itself. A little thought will indicate that the exponential is the only such function (i.e., $y(t) = e^t$).

这种形式称为可分离一阶常微分方程，当函数f可以表示为两个函数的产物时，其中一个函数仅显式地涉及t，另一个函数仅显式地涉及y。换句话说，该问题具有以下形式

$$\frac{dy}{dt} = r(y)Q(t) \tag{3.34}$$

In general separable equations like this can be linear or nonlinear, and solutions for both cases are possible assuming that we can complete the necessary integration. First-order ODEs are one of the few areas of mathematical modeling where one can obtain solutions for relatively general conditions. Most physical systems actually contain some nonlinearity. It is only our models that approximate such systems as being linear. Nonlinear problems are the rule rather than the exception.

Nearly all discussions of problems of this form write it in the following way, where $R(y) = 1/r(y)$.

$$R(y)\frac{dy}{dt} = Q(t) \tag{3.35}$$

There are few nonlinear problems where a general solution exists. Problems of the form of Eq. (3.35) do have solutions in general, and therefore represent a really unusual situation in mathematics! A few notes are in order here. First, equations of the form of Eq. (3.35) can, in principle, be linear (take $R(y) = 1$) or nonlinear; regardless, the solution method discussed in the following is valid. Second, in the material that follows, it is important to be aware of potential confusion between independent and dependent variables. One can compute an integral that is written in terms of dependent variables. This can happen when a transformation of variables allows an integration of the dependent variable to be expressed entirely in terms of the dependent variable.

这个圆盘 ussion can be put in context with an example solution for a particular nonlinear first-order ODE. 一般形式 (abstracted) case will be developed following the example.

Example 3.3 (First-order nonlinear separable equations). Consider the following nonlinear ODE (offered without a physical motivation).

$$y^2\frac{dy}{dt} = t^2 \tag{3.36}$$

This expression is clearly nonlinear in y. Just as a reminder about testing for nonlinearity: the two required properties are (1) additivity and (2) homogeneity as discussed in §2.3.1; failing either one of these tests indicates nonlinearity. If we substitute $y = \alpha y_1(t)$ into our problem we find

Checking for nonlinearity

$$(\alpha y_1)^2\frac{d(\alpha y_1)}{dt} = t^2 \tag{3.37}$$

$$\alpha^2 y_1^2 \alpha\frac{dy_1}{dt} = t^2 \tag{3.38}$$

$$\alpha^3 y_1^2\frac{dy_1}{dt} = t^2 \tag{3.39}$$

因此，该问题不满足齐次性，属于非线性问题。尽管如此，由于它具有式 (3.34) 的形式，因此该问题是可分离的。为了解此方程，我们首先对两边关于t进行积分。由此得到

$$\int y^2 \frac{dy}{dt} dt = \int t^2 \, dt \tag{3.40}$$

In Section 2.4.3.1, the differential was defined. There, we showed that it was appropriate to make the following correspondence

$$dy = y'(\tau) d\tau \tag{3.41}$$

or, equivalently

$$dy = \left(\frac{dy}{d\tau} \right) d\tau \tag{3.42}$$

while it is tempting to think of this as if the differential quantities dt can be "cancelled out", this is not the most useful way of thinking about this expression. In reality, we *must* mean something different from that interpretation because the derivative dy/dt represents the *limit of a ratio as dt tends to zero*. However, we cannot think of dy or dt individually as being defined by such a limit; otherwise, these two quantities would be identically zero, and Eq. 3.42 would not be particularly useful. Instead, we think of this relationship between differentials as being described as follows.

当自变量 dt 的增量足够小时，函数 $y(t)$ 将发生 dy 量的变化，并伴有误差 ε。通过减小dt，可以使这种近似中的误差 ε 任意小。

With this recognition, we are in a position to solve Eq. (3.40). First we make the substitution $dy = (dy/dt)dt$ to give

$$\int y^2 dy = \int t^2 \, dt \tag{3.43}$$

Integrating this, we find

$$\frac{1}{3} y^3 = \frac{1}{3} t^3 + C_1 \tag{3.44}$$

Noting that $3C_1$ is just a new constant, we can write this in the form

$$y^3 = t^3 + C_2 \tag{3.45}$$

There are no ancillary conditions given for this problem, so this solution is complete. Usually, if it is possible, we try to express a solution in its *explicit* form. Doing so here yields

$$y(t) = (t^3 + C_2)^{\frac{1}{3}} \tag{3.46}$$

Now that we have seen a concrete example, it is pretty easy to see how these problems can be solved generally. Here is the solution method, but done with arbitrary functions. The problem is

$$\frac{dy}{dt} = r(y)Q(t) \tag{3.47}$$

$$y(t_0) = y_0 \tag{3.48}$$

or, recalling $R(y) = 1/r(y)$

$$R(y)\frac{dy}{dt} = Q(t) \tag{3.49}$$

$$y(t_0) = y_0 \tag{3.50}$$

现在，我们只需对等式两边关于t积分，即得

$$\int R(y)\frac{dy}{dt}\,dt = \int Q(t)\,dt \tag{3.51}$$

如前一个例题所述，在第 1 章（第 2.4.3.1 节）的积分理论复习中，我们回顾了微分的概念。在那里，我们指出，导数是通过两个趋于零的量之比的极限过程来定义的，而微分则被定义为一个非零量。本质上，要理解微分，我们假设可以不通过极限过程来处理 “dy” 和 “dt”，而是将它们视为两个量的比值，用以近似导数，并带有误差 ε；该误差可以小到我们所希望的程度。基于这一模型，我们可以恰当地将一阶导数几乎视为代数量。因此，通常可以写成

$$dy = \frac{dy}{dt}dt \tag{3.52}$$

得到

$$\int R(y)dy = \int Q(t)\,dt + C \tag{3.53}$$

其中，最后的常数表示由所给辅助条件（若有）确定的积分常数。

Example 3.4 (A first-order, linear equation). Suppose we have the problem

$$y\frac{dy}{dt} = t^2 \tag{3.54}$$

附加条件为

$$y(0) = 1 \tag{3.55}$$

What is the explicit solution for $y(t)$?

Solution. We can use the expression given by Eq. ((3.53)) to solve this problem. Clearly, the

solution is given by

$$\int y\,dy = \int t^2\,dt + C \tag{3.56}$$

or, upon computing the integrals

$$y^2(t) - y^2(0) = \tfrac{2}{3}t^3 + C \tag{3.57}$$

Noting the initial condition, this can be expressed

$$y^2(t) = (1 + \tfrac{2}{3}t^3) \tag{3.58}$$

where we must have $C = 0$ (it is easy to verify that this expression matches the ancillary condition at $t = 0$). This is an implicit equation for y, since the expression does not provide us with the value of y, but, rather, the value of y^2. In this case we are lucky– the expression is a quadratic, and we can use the quadratic formula (or, in this case, simply taking the square root of both sides of the expression) to get an explicit expression for y.

$$y(t) = \pm\sqrt{1 + \tfrac{2}{3}t^3} \tag{3.59}$$

The derivative of the solution is

$$\frac{dy}{dt} = \pm\,\tfrac{1}{2}(1 + \tfrac{2}{3}t^3)^{-\frac{1}{2}}\,2t^2 \tag{3.60}$$

And with this information, it is easy to verify that the original nonlinear ODE is also met by the 解答 on.

Note that we can also solve this problem using *definite* integration. When we do so, we no longer have a constant of integration because this constant is automatically accounted for by the information in the bounds. To solve the problem using definite integrals, we integrate in time from where information is known (at t_0) to any unknown time (symbolized by simply t) as follows

$$\int_{y(0)}^{y(t)} y\,dy = \int_{\tau=0}^{\tau=t} \tau^2\,d\tau \tag{3.61}$$

请注意，这里我们使用积分变量 τ 来表示积分中的时间变量。请记住，这是积分中常用的符号表示，有助于避免积分变量与积分上下限中的变量混淆。对其进行积分后，我们得到解。

$$y(t)\big|_{y(0)}^{y(t)} = \tfrac{2}{3}\tau^3\big|_{\tau=0}^{\tau=t} \tag{3.62}$$

Noting the initial condition $y(0) = 1$, we have the solution

$$y(t) = \pm\sqrt{1 + \tfrac{2}{3}t^3} \tag{3.63}$$

which is identical to the solution via indefinite integration.

The previous problem was one that happened to be linear in the function y (Question: can you show this?). For nonlinear separable equations, the solution to nonlinear problems proceeds essentially the same way. The only real issue one faces with separable problems is whether or not the two sides, once separated, can be integrated.

在下一个示例中，我们将使用分离变量法求解一个以因变量为未知量的非线性问题。对于这个问题，无法将结果表示为显式形式（即 $y(t)$ = 函数(t) 的形式）。因此，我们将结果保留为隐式形式。

示例 *3.5*（米氏动力学）。米氏动力学描述的是反应速率受导致反应发生的化合物总数（或总量）限制的动力学过程。例如，由载体颗粒表面催化剂引起的反应（形式为 $A \xrightarrow{\text{催化剂}} B$）有时会受到反应速率限制，因为一旦所有催化剂位点都被反应物（A）占据，反应便无法继续。反应速率为

$$\frac{dc_A}{dt} = -\frac{kc_A}{c_A + K_A}$$

其中，k 是反应速率常数（mol/(L·s)），而K_A 是半饱和常数（mol/L）。米氏动力学有时也称为饱和动力学，因为随着反应物物种浓度增加，动力学反应速率会达到最大值。图中对此进行了说明

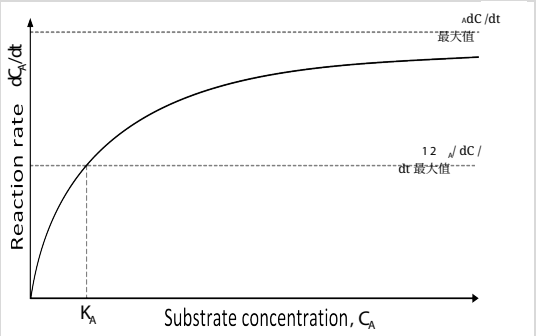


图 3.2：米氏动力学。通过反应速率与浓度曲线的几何特征，可以看出这种动力学的两个有趣特点。当浓度远大于半饱和常数K_m时，反应速率近似等于动力学速率参数k。半饱和常数K_A 可通过先求出最大反应速率的一半，再找到浓度轴上对应的值来确定。该值等于K_A。

在假设该速率定律成立的情况下，确定 $c_A(t)$ 的函数关系，以及初始条件$c_A(0) = C_0$。c_A 达到其初始值的一半需要多长时间？解。首先进行分离变量法，得到如下形式

$$-\frac{c_A + K_A}{c_A} dc_A = kdt$$

$$-\int_{C=c_A(0)}^{C=c_A(t)} \frac{C + K_A}{C} dC = \int_{\tau=0}^{\tau=t} kd\tau$$

(Note that here, we have been careful to use variables of integration when writing out the *definite integral*. Sometimes this is helpful to do; it reinforces what the correct bounds for the integration are.

然而，为了简洁和方便，一旦熟悉了不定积分，就没有必要在积分时明确切换到一个单独的哑变量。这样得到的表达式虽然有些“不严谨”，但只要清楚自己在做什么，通常不会造成任何问题。)

Performing the integrations, we find

$$-C|_{c_A(0)}^{c_A(t)} - K_A \ln C|_{c_A(0)}^{c_A(t)} = kt$$

Evaluating this result at the bounds, and using the properties of logarithms, the final result is

Performing the integrations, we find

$$t = \frac{1}{k}\left[C_0 - c_A(t) + K_A \ln \frac{c_A}{C_0}\right]$$

As a check on this result, note two properties that we would expect from this solution (1) The solution meets the initial condition at $t = 0$ (i.e., for $t = 0$, $c_A(0) = C_0$), and (b) the result is positive for all values of c_A; we expect this latter property because time can not be a negative quantity! Also note that argument of the logarithm is *dimensionless*. It is a good practice to assure that in your *final results* all logarithmic arguments should be dimensionless! They arguments may not be dimensionless during intermediate steps, but if they cannot be made to be free of units at the end of the problem, this may indicate that there has been an error. Generally speaking, it is improper to take the logarithm of a quantity with units.

To solve the second part of the problem, we substitute $c_A(t) = C_0/2$, and solve for time; the result is

$$t = \frac{1}{k}\left[\tfrac{1}{2}C_0 + K_A \ln 2\right]$$

Example 3.6 (Solution to a linear first-order ODE by separation of variables). Suppose we have the classical first-order reaction in a completely-mixed reactor problem discussed above, and specified by the set of equations

$$\frac{dy}{dt} = -k_1 y \tag{3.64}$$

$$y(t=0) = c_0 \tag{3.65}$$

What is the solution?

Answer. Separating variables we have

$$\frac{1}{y}\,dy = -k_1\,dt \tag{3.66}$$

Integrating both sides of this immediately yields

$$\int \frac{1}{y}\,dy = -k_1 \int dt \tag{3.67}$$

or

$$\ln(y) = -k_1 t + C \tag{3.68}$$
$$y(t) = C\exp[-k_1 t] \tag{3.69}$$

Substituting the ancillary equation into this result, we find

$$y(0) = C \tag{3.70}$$
$$or \tag{3.71}$$
$$C = c_0 \tag{3.72}$$

从而使最终结果成为熟悉的随时间呈指数衰减的函数。

$$y(t) = c_0\exp[-k_1 t] \tag{3.73}$$

Example 3.7 (Solution to a nonlinear first order ODE by separation of variables). Suppose instead of the classical first-order reaction in a completely-mixed reactor problem, we have a second-order problem specified by

$$\frac{dy}{dt} = -k_1 y^2 \tag{3.74}$$
$$y(t=0) = c_0 \tag{3.75}$$

What is the solution?

Answer. Separating variables as above, we have

$$\frac{1}{y^2}\,dy = -k_1\,dt \tag{3.76}$$
$$\tag{3.77}$$

Integrating both sides of this immediately yields

$$\int \frac{1}{y^2}\,dy = -k_1 \int dt \tag{3.78}$$

or

$$-y^{-1} = -k_1 t + C \tag{3.79}$$

Upon solving explicitly for $y(t)$ we have

$$y(t) = -\frac{1}{-k_1 t + C} \tag{3.80}$$

最后，将辅助方程代入这一结果，我们得到

$$y(0) = -\frac{1}{C} \tag{3.81}$$

$$or \tag{3.82}$$

$$C = -\frac{1}{c_0} \tag{3.83}$$

so that the final result is the following function of time

$$y(t) = \frac{c_0}{k_1 c_0 t + 1} \tag{3.84}$$

Noting that at $t = 0$, we have $y(0) = c_0$, we find that the ancillary condition is met. Also noting that the derivative is given by

$$\frac{dy}{dt} = -\frac{k_1 c_0}{(k_1 c_0 t + 1)^2} \tag{3.85}$$

it is easy to validate that the solution not only meet the initial condition, but it meets the original ODE also.

3.3.3 不可分离线性一阶常微分方程

在前面的章节中，一阶常微分方程通过直接应用微积分基本定理求解。这种方法要求方程是可分离的。然而，存在大量无法直接分离的方程。这类方程的标准形式为

$$\frac{dy}{dt} + P(t)\, y = Q(t) \tag{3.86}$$

There is no obvious way to integrate both sides of this equation to get a solution. Doing so would result in an integral with integrand $P(t)y(t)$; because we do not *know* $y(t)$, we are stuck with an implicit integral solution for $y(t)$. While technically not a disaster, it is not a solution that is easy to compute the values of the dependent variable.

To handle problems of this general form for linear first-order ODEs requires first transforming this problem. This idea is a key element in analytical modeling and in mathematical analysis generally; it is important enough to highlight the idea specifically.

Key Idea 1.1. When encountering a problem that one does not know how to solve, it is worthwhile considering if it is possible to transform the problem into a form that does have a known solution.

This idea behind this statement is simultaneously extremely useful, and potentially frustrating! It suggests what must be attempted, but it provides no information about how this is to be accomplished. This brings up a second key idea that arises in mathematical analysis.

Key Idea 1.2. Mathematical analysis often involves the creative use of intuition or experimentation to find solutions.

An example of these two ideas is exactly what is embodied by the following solution to the general first-order linear ODE. It is often most useful to see such developments as an example before stating the more general and abstract result.

Example 3.8. Suppose we have a continuously-stirred tank reactor (CSTR) undergoing a first-order reaction process $A \to B$ with kinetic rate constant k_1. If the flow rate in and out are both equal to Q, the concentration of species A is c_0 at the inlet, and the initial concentration of species A is $y_A(0) = 0$ then the net balance for the concentration of species A reactor is given by the ODE

$$\frac{dy_A}{dt} = \frac{Q}{V}c_0 - \frac{Q}{V}y_A - k_1 y_A \tag{3.87}$$

$$y_A(0) = 0 \tag{3.88}$$

Here, we note that the quantity $\frac{Q}{V}$ is related to the *residence time* for the reactor, τ

$$\frac{Q}{V} = \frac{1}{\tau} \tag{3.89}$$

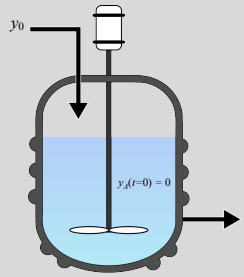


图 3.3：连续搅拌釜式反应器。槽内初始时物种的浓度 A 和 B均为零，但含有悬浮的催化剂（以不会离开系统的颗粒形式存在）。在时间 $t = 0$，物种 A 开始进入反应器，催化反应随之进行。

为了解决这个问题，我们首先将方程改写为以下形式

$$\frac{dy_A}{dt} + (\tau^{-1} + k_1)y_A = \tau^{-1}y_0 \tag{3.90}$$

现在，很明显，这个表达式并不属于上述材料中已经讨论过的形式。然而，我们注意到，如果令 $\beta = (\tau^{-1} + k_1)$，然后将该表达式两边同时乘以 $e^{\beta t}$，我们就有

$$\frac{dy_A}{dt}e^{\beta t} + y_A\beta e^{\beta t} = e^{\beta t}\tau^{-1}y_0 \tag{3.91}$$

Although it may not be obvious that this has improved things, it has once we make an observation: The left-hand side of the expression is now a total derivative. In other words, we note

$$\frac{d}{dt}(y_A e^{\beta t}) = \frac{dy_A}{dt}e^{\beta t} + y_A\beta e^{\beta t} \tag{3.92}$$

But this is exactly the left-hand side of the equation above! Thus, we can substitute as follows

$$\frac{d}{dt}(y_A e^{\beta t}) = e^{\beta t}\tau^{-1}y_0 \tag{3.93}$$

Now, we have transformed the original equation into one we know how to solve- this expression is separable in the conventional way that was discussed previously. If this is unclear, we can make the following temporary transformation of variables

$$\psi(t) = y_A e^{\beta t} \tag{3.94}$$

Then, we have

$$\frac{d\psi}{dt} = e^{\beta t}\tau^{-1}y_0 \tag{3.95}$$

Separating variables and integrating, we find

$$\int d\psi = \int e^{\beta t}\tau^{-1}y_0\,dt + C \tag{3.96}$$

这就得到了

$$\psi(t) = \frac{1}{\beta}e^{\beta t}\tau^{-1}y_0\,dt + C \tag{3.97}$$

Returning to the original variables, we have

$$y_A \exp\left((\tau^{-1} + k_1)t\right) = \frac{1}{(\tau^{-1} + k_1)}\exp\left((\tau^{-1} + k_1)t\right)\tau^{-1}y_0 + C \tag{3.98}$$

Finally, solving for y_A gives the expression

$$y_A(t) = \frac{1}{(1 + \tau k_1)}y_0 + \exp\left(-(1 + \tau k_1)\frac{t}{\tau}\right)C \tag{3.99}$$

令$t = 0$，这样我们就可以利用初始条件求解C

$$C = -\frac{1}{(1+\tau k_1)} y_0 \tag{3.100}$$

Thus, the final solution is

$$y_A(t) = \frac{y_0}{(1+\tau k_1)} \left[1 - \exp\left(-(1+\tau k_1)\frac{t}{\tau}\right)\right] \tag{3.101}$$

In Example 3.8, the function $P(t)$ that was incorporated into the exponential was a simple constant, $\beta = \tau^{-1} + k_1$. Recall, however, the general form is

$$\frac{dy}{dt} + P(t)\, y = Q(t) \tag{3.102}$$

解 here is actually no more difficult than when $P(t)$ is a constant. For this more general case, take

$$s(t) = \int P(t)\, dt \tag{3.103}$$

这种变换可以通过不定积分来完成。不定积分的常数可以取为零（当我们使用积分因子时，**始终**如此），因为任何这样的常数最终都可以从所得的方程中消去。注意，

$$\frac{ds}{dt} = P(t) \tag{3.104}$$

By the fundamental theorem of calculus.

Now, multiplying Eq. (3.102) through by $\exp[\beta(t)]$ we find

$$\frac{dy}{dt}\exp[s(t)] + P(t)\exp[s(t)]\, y = \exp[s(t)]Q(t) \tag{3.105}$$

This expression can be rewritten as

$$\frac{d}{dt}\left(\exp[s(t)]\, y\right) = \exp[s(t)]Q(t) \tag{3.106}$$

Now, set $\psi(t) = \exp[s(t)]\, y$, separate variables, and integrate both sides as *definite* integrals. This results in

$$\int_{\psi(t'=t_0)}^{\psi(t'=t)} d\psi = \int_{t'=t_0}^{t'=t} \exp(s(t'))Q(t')dt' \tag{3.107}$$

Integrating this yields

$$\psi(t) - \psi(t_0) = \int_{t'=t_0}^{t'=t} \exp(s(t'))Q(t')dt' \tag{3.108}$$

Upon returning to the original variables, we find

$$\exp[s(t)]\, y(t) - \exp[s(t_0)]\, y(t_0) = \int_{t'=t_0}^{t'=t} \exp[s(t')]Q(t')dt' \tag{3.109}$$

并求解 this for $y(t)$ gives us the general formula for the solution to the separable first-order linear ODE

$$\boxed{y(t) = \exp[-s(t)]\left[\int_{t'=t_0}^{t'=t} \exp[s(t')]Q(t')dt' + y(t_0)\exp[s(t_0)]\right]} \tag{3.110}$$

其中，回忆一下

$$s(t) = \exp\left(\int P(t)dt\right) \tag{3.111}$$

This expression gives the solution, then, to *any* linear non-separable first-order ODE.

The entire analysis above can also be done using indefinite integration. The entire process is exactly the same up to Eq. 3.106. From there, we can integrate both sides as indefinite integrals, but remembering to add the appropriate constant of integration. Thus, we would have the result (starting right after Eq. 3.106)

$$\int d\psi = \int \exp(s(t))Q(t)dt \tag{3.112}$$

$$\psi(t) = \int \exp(s(t))Q(t)dt + C_1 \tag{3.113}$$

Where C_1 is a constant of integration from the integration of the left-hand side. Note, that the integral on the right-hand will also generate a constant of integration. However, the two constants can be combined, so we can ignore the constant coming from the remaining integral. Recalling that $\psi(t) = \exp[s(t)]y$, then this can be simplified to give $y(t)$ as follows

$$\exp[s(t)]y(t) = \int \exp(s(t'))Q(t')dt' + C_1 \tag{3.114}$$

$$\boxed{y(t) = \exp[-s(t)]\left[\int \exp(s(t'))Q(t')dt' + C_1\right]} \tag{3.115}$$

Where, again, remember that *only one constant*, C_1, is generated for this problem in total, which is consistent for the fact that it is a first-order problem. By comparing the two boxed results above, it is clear that $C_1 = y(t_0)\exp[s(t_0)]$. However, some people prefer to use the simpler looking version given by Eq. (3.115), and then simply evaluate the unknown constant using the ancillary data. The results will be the same regardless of which approach is taken.

Example 3.9 (CSTR with catalyst deactivation.). Solid catalysts are used in a huge variety of reactors, from high-tech reactors to make pharmaceuticals to reactors for ion-exchange in the treatment of drinking water. One potential problem with solid catalysts is that they can become deactivated over time. There are any number of deactivation processes (e.g., particle sintering, catalyst poisoning, and catalyst coking are three types of deactivation that are common); all of them reduce catalyst effectiveness, and have significant economic impact.

One model for catalyst inactivation involves an inverse-time function that decreases the net rate of reaction with increasing time. Suppose we have a reaction that proceeds in the presence of a catalyst with the simple form $A \xrightarrow{C} B$, where A is the reactant, B is the product, and C is the catalyst. Returning to the case of a CSTR from the previous example, the molar mass balances for a catalytic reaction with deactivation for chemical species A can be specified by

$$\frac{dy_A}{dt} = \tau^{-1} y_0 - \tau^{-1} y_A - \left(\frac{1}{1+k_d t}\right) k_1 y_A \tag{3.116}$$

$$y_A(0) = 0 \tag{3.117}$$

$$y_B(0) = 0 \tag{3.118}$$

$$\tag{3.119}$$

We have assumed that the catalyst concentration is constant in the CSTR (catalyst neither enters nor leaves the reactor). Note that because initially there is no species B in solution, the concentration of species B at any time can be computed from how much of species A has been converted. While it is possible to also get the concentration of species B at any time, for now we will focus only on the balance for species A.

Here, we can see that Eq. (3.116) is identical to Eq. (3.87), except the reaction term includes an additional function of time that multiplies the first-order reaction. This function decreases the net rate of reaction as time increases. This problem would be somewhat intimidating if we had not previously developed the integration factor formula given by Eq. (3.110) (and even with it, it may still be challenging!) Rewriting the mass balance equation, we find

$$\frac{dy_A}{dt} + \left[\tau^{-1} + \left(\frac{k_1}{1+k_d t}\right)\right] y_A = \tau^{-1} y_0 \tag{3.120}$$

In this form, it is clear that we have

$$s(t) = \int \left[\tau^{-1} + \left(\frac{k_1}{1+k_d t}\right)\right] dt \tag{3.121}$$

or, after completing the integral

$$s(t) = \left[\frac{t}{\tau} + \frac{k_1}{k_d}\ln(1+k_d t)\right] \tag{3.122}$$

This makes rather short work of writing down the solution; recalling Eq. (3.110)

$$y(t) = y(t_0)\exp[s(t_0)]\exp[-s(t)] + \exp[-s(t)]\int_{t'=t_0}^{t'=t} \exp[s(t')]Q(t')dt' \tag{3.123}$$

Making the appropriate substitutions (and noting that $y(t_0) = 0$), we have the solution

$$y_A(t) = \exp\left[-\left(\frac{t}{\tau} + \frac{k_1}{k_d}\ln(1+k_d t)\right)\right]\int_{t'=0}^{t'=t} \exp\left[\frac{t}{\tau} + \frac{k_1}{k_d}\ln(1+k_d t)\right]\frac{y_0}{\tau}dt' \tag{3.124}$$

Computing this last integral is somewhat of a challenge; however, the integral is tabulated in tables, and can be computed by symbolic mathematics programs like Mathematica. The result is

$$y_A(t) = \tau e^{-\frac{1}{k_d\tau}}\left(\frac{1}{k_d\tau}\right)^{-\frac{k_1}{k_d}}\left[\Gamma\left(\frac{k_1+k_d}{k_d}, -\frac{k_d t+1}{k_d\tau}\right) - \Gamma\left(\frac{k_1+k_d}{k_d}, -\frac{1}{k_d\tau}\right)\right] \tag{3.125}$$

Plots of the solution for species A and B are provided in Fig. 3.4.

伽马函数，Γ 是一种超越函数，它将阶乘运算推广到实数（或复数）范围。该函数应用广泛，通常作为内置函数存在于 Mathematica 等符号数学程序以及 MATLAB 等解释型编程语言中。

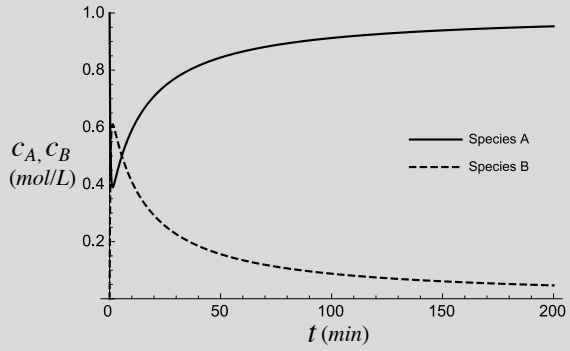


Fig. 3.4: Solutions for species A and species B as a function of time for a catalytic reaction with deactivation. For this solution, $k_1 = 2\ min^{-1}$, $k_d = 0.1\ min^{-1}$, $\tau = 1\ min^{-1}$, $y_0 = 1\ mol \cdot m^{-3}$.

Just for completeness, note that the two Gamma functions are defined by the following integrals

伽马
$$\Gamma_1(x) = \int_{t=0}^{t\to\infty} t^{x-1} e^{-t} dt \tag{3.126}$$

不完全伽马函数
$$\Gamma_2(a,x) = \int_{t=x}^{t\to\infty} t^{a-1} e^{-t} dt \tag{3.127}$$

More information on the Gamma function can be found in many texts. The Gamma function can be evaluated using common computational platforms such as MATLAB or Mathematica, both of which have built-in routines for computing these functions.

3.4 常系数二阶常微分方程

We will cover exactly two kinds of second-order ODEs: (1) directly integrable second-order ODEs, and (2) 一般的秒[内容不完整]nd-order ODEs with constant coefficients.

Linear second-order ODEs are functions of a single independent variable that take the form

$$\frac{d^2y}{dx^2} + b(x)\frac{dy}{dx} + c(x)y = g(x) \tag{3.128}$$

In general, solutions of these equations (where the coefficients are functions rather than constants) require series solution methods which we will not be reviewing. However, there are a few other interesting problems that are solvable. These are the (1) directly integrable problems, and (2) the case of linear second-order ODEs with constant coefficients.

3.4.1 可直接积分的二阶常微分方程

If a second-order can be put in the form

$$\frac{d}{dx}\left(p(x)\frac{dy}{dx}\right) = h(x) \tag{3.129}$$

then, in principle, it can be solved directly by two integrations through the application of the fundamental theorem of calculus (FTC).

Suppose the ODE above is defined on $x \in [a,b]$, it has two appropriate ancillary conditions imposed, and $p(x) \neq 0$ anywhere in the domain. Then, integrating both sides of Eq. (3.129) gives (using the fundamental theorem of calculus), and then dividing by $p(x)$

$$\frac{dy}{dx} = \frac{1}{p(x)}\int_x h(\xi)d\xi \tag{3.130}$$

where, here we have used a dummy variable of integration to avoid confusion in later developments. A second integration gives a solution for $y(x)$

$$y(x) = \int_x \frac{1}{p(x)}\int_x h(\xi)d\xi dx \tag{3.131}$$

Note that these two integrations will generate two constants of integration.

Example 3.10 (Directly Integrable Second-Order ODE.). Suppose we have an integrable of the form

$$\frac{d}{dx}\left((x^2+1)\frac{dy}{dx}\right) = 3x^2, \;\; x \in [0,1]$$

with

$$y(0) = 0, y(1) = 0$$

What is the solution?

Solution. To start, integrate both sides as an indefinite integral and divide through by x^2+1

$$\frac{dy}{dx} = \int_x \frac{x^3+c_1}{x^2+1}\,dx$$

现在，这个积分并不容易求解（至少表面上看是如此）。可以尝试使用部分分式法（通常在微积分入门课程中介绍）对其进行积分。另一种方法是查阅积分表，也可以使用各种符号数学语言来求出其值。

this integral. We are going to take the latter option here, noting that the integral can be computed as

$$\frac{dy}{dx} = \frac{1}{2}[x^2 + 2c_1 \arctan(x) - \ln(1+x^2)]$$

A second integration yields

$$y(x) = c_1\left(x\arctan(x) - \frac{1}{2}\log\left(x^2+1\right)\right) + \frac{x^3}{6}$$
$$-\frac{1}{2}x\log\left(x^2+1\right) + x - \arctan(x) + c_2$$

Applying the first ancillary condition, $y(0) = 0$, we find $c_2 = 0$. Applying the second ancillary condition $y(1) = 0$ gives

$$0 = c_1\left(\frac{\pi}{4} - \frac{1}{2}\log(2)\right) - \frac{\pi}{4} + \frac{7}{6} - \frac{\log(2)}{2}$$

or

$$c_1 = \frac{\frac{\pi}{4} - \frac{7}{6} + \frac{1}{2}\log(2)}{\pi - 2\log(2)}$$

So, the final solution is, somewhat unbelievably, the function

$$y(x) = \left(\frac{\frac{\pi}{4} - \frac{7}{6} + \frac{1}{2}\log(2)}{\pi - 2\log(2)}\right)\left(x\arctan(x) - \frac{1}{2}\log\left(x^2+1\right)\right) + \frac{x^3}{6}$$
$$-\frac{1}{2}x\log\left(x^2+1\right) + x - \arctan(x)$$

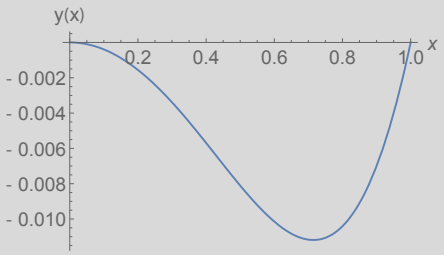


图 3.5：解 \(y(x)\) 的图像

It is possible to verify that $y(0) = y(1) = 0$ for this result, suggesting strongly that it is a correct one.

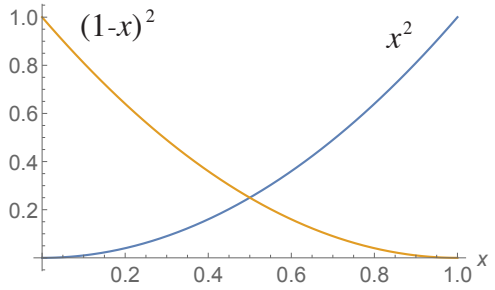


Fig. 3.6 The functions $y(x) = x^2$ and $y(x) = (1-x)^2$

3.4.2 思考二阶常微分方程的解

Before doing any additional analysis, it is worth doing a little thinking to build our intuitive understanding of problems of this sort. So, let's think about this problem entirely backward for a few moments. To start, consider the function

$$y(x) = \alpha y_1(x) + \beta y_2(x) \quad x \in [0,1] \tag{3.132}$$

其中

$$y_1(x) = x^2$$
$$y_2(x) = (1-x^2)$$

This function is clearly the sum of two independent functions; these independent functions are plotted in Fig. 3.6

These functions have some interesting features. First, note that y_1 is zero at one end, and 1 at the other; y_2 has the same feature, but at opposite ends of the domain. Also, the derivative y_1 is zero at $x = 0$, and the derivative of y_2 is zero at $x = 1$. These features give lots of flexibility if one wants to develop new functions by creating linear combinations of these two. For example, suppose I want a function on $[0,1]$ that is (1) a quadradic, (2) has the value 3 on the left-hand side, and has the value of 2 on the right-hand side. A little though will indicate that we can *create* this function by the following linear combination

$$\begin{aligned} y(x) &= 2y_1(x) + 3y_2(x) \\ &= 2x^2 + 3(1-x)^2 \\ &= 5x^2 - 6x + 3 \end{aligned}$$

If we want to be really clever, we can even indicate the functional *value* at one end, and the *slope* at the other end. To do this, first note

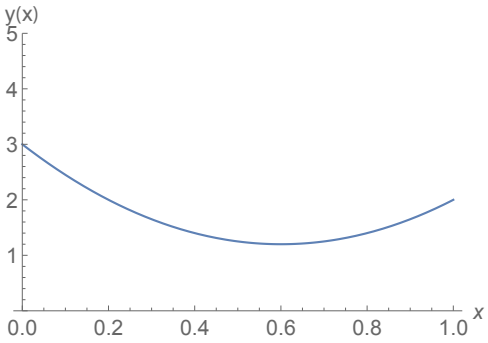


Fig. 3.7 The functions $y(x)$ composed specifically so that $y(0) = 3$ and $y(1) = 2$.

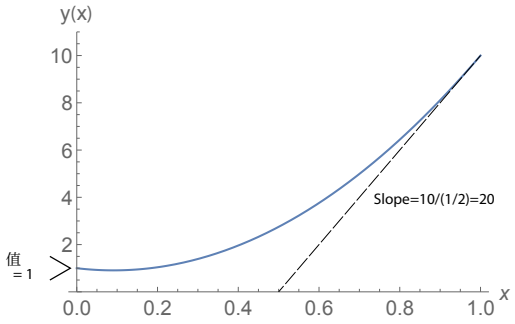


Fig. 3.8 The functions $y(x)$ composed specifically so that $y(0) = 1$ and $y'(1) = 20$.

$$y'(x) = 2\alpha x - 2\beta(1-x) \quad x \in [0,1]$$

So, now to meet two conditions (one for the value at one end, one for the slope at the other end) we can solve these two equations simultaneously. For example, suppose we wanted $y(0) = 1$, $y'(1) = 20$. This implies

$$\begin{aligned} 1 &= 0 + \beta(1-0) \rightarrow \beta = 1 \text{（来自}y(x)\text{的方程）} \\ 20 &= 2\alpha + 0 \qquad\qquad \rightarrow \alpha = 10 \text{（来自}y'(x)\text{的方程）} \end{aligned}$$

到目前为止，我们还没有讨论过常微分方程的解。然而，我们刚刚考察过的函数与这里的内容密切相关。为了说明这一点，我们只需考虑以下非齐次问题的解。正如上一节所述，这是一个可以直接积分求解的问题，而且其形式特别简单。

$$\frac{d^2y}{dx^2} = k_0$$

where k_0 is some specified constant.

Without any ancillary information, we can still solve this problem. To do so, we integrate both sides to give

$$\int_x \frac{d^2y}{dx^2}dx = \int_x k_0 dx \Rightarrow$$
$$\frac{dy}{dx} = k_0 x + C_1$$

and, integrating a second time, we find

$$\int_x \frac{dy}{dx}dx = \int_x (k_0 x + C_1)dx \Rightarrow \tag{3.133}$$
$$y(x) = \frac{k_0}{2}x^2 + C_1 x + C_2 \tag{3.134}$$

事实证明，我们之前考察过的函数 $y(x) = \alpha x^2 + \beta(1-x)^2$, 与这个解等价。要看出这一点，请注意，如果我们展开并重新整理方程（3.132），就得到。

$$y(x) = (\alpha + \beta)x^2 - 2\beta x + \beta$$

比较两个解在 x^2 项处的值，我们发现，要使两个解相匹配，必须有$k_0 = 2(\alpha + \beta)$。假设我们重新考察$y(0) = 3, y(1) = 2$的情形。用这些约束求解方程（3.134）得到

$$y(0) = 3 \Rightarrow C_2 = 3$$
$$y(1) = 2 \Rightarrow C_1 = -(1 + k_0/2)$$

我们对两个方程之间相容性的约束不允许 α 和 β 取任何 值；它们的值与k_0的关系为 $2(\alpha+\beta) = k_0$。这意味着，我们的边值问题$y(0) = 3$, $y(1) = 2$，其中 $\alpha = 3, \beta = 2$，对应于$k_0 = 10$。因此，我们得到$C_1 = -6$。容易证明，这个解与图 3.7 中绘出的解相同。或者，我们可以说，我们绘出的第一个解（当时我们强制右端的值为 2，左端的值为 3）

$$y(x) = 2x^2 + 3(1-x)^2$$

Corresponds to the second-order ODE

$$\frac{d^2y}{dy^2} = 10$$
$$y(0) = 3$$
$$y(1) = 2$$

So, why did we bother with this long example? Well, for one thing, it showed us, empirically at least, that we do indeed generate two constants of integration for second-order ODEs. It also showed us that we can think of the solution to such problems as the linear combination of *two linearly independent* solutions. In the example above, we thought of these two solutions as $y_1(x) = x^2$ and $y_2(x) = (1-x)^2$.

3.4.3 Linear Second-Order ODEs with constant Coefficients: Homogeneous Case

Now, let's examine the general case of linear second-order equations with constant coefficients. Recall the general form for a *linear* second-order ODE given by Eq. (3.128). When the term $c(x) \equiv 0$, then the equation is called *homogeneous*. Sometimes, functions like $g(x)$ are called *source* terms. Because functions like this are not dependent on y, and because the entire term is independent from y, the action of these functions can be though of as "sources" or "sinks" that drive the problem.

We will study the particular *homogeneous* case where the coefficients of the differential equation are constants rather than functions. This means that the equations of interest to us take the form

$$\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = 0 \tag{3.135}$$

首先，我们注意到每一个 具有常系数的二阶齐次常微分方程都可以写成这种形式。其次，我们注意到，像这样的方程的解必然包含两个未知常数。我们可以认为，每进行一次为了逆转导数而需要执行的"积分"，都会产生一个未知常数。这里，我们对"积分"一词的使用稍微宽泛了一些，但这是一种正确的理解。由于导数的最高阶数为二阶，这意味着需要进行两次积分。最后，根据我们之前的经验，我们可能会预期，该方程的解是两个线性无关解的线性组合。我们尚未证明这一点（稍后我们会提供某种形式的证明），但这一想法与我们迄今为止的观测结果是一致的。

We expect that there are two linearly independent solutions to the general homogeneous problem with constant coefficients, and we know that we can make new solutions by making linear combinations of them. However- as to what the solutions actually might be, we have not much to go on. We also know that for the directly integrable homogeneous problem

$$\frac{d^2y}{dx^2} = 0 \tag{3.136}$$

解必须是形如 $y(x) = C_1x + C_2$ 的线性解。因此，无论我们为方程 (3.135) 找到何种更一般的解，它都必须将线性情形作为可能的情况之一（最终我们将看到，确实如此）。

因此，为了继续推导，我们将采用一种在数学中不常讨论的方法：运用直觉。若观察方程（3.135），我们马上会注意到，我们要寻找某个函数$y(x)$，使得至少可以对它求导，将导数乘以常数，再把它们相加得到零。事实上，对于$b = 0$的情况，我们甚至知道，必须有一个函数，其二阶导数是原函数的某个常数倍。显然，所有多项式都不可能符合要求。我们立刻能想到的、可能具有这种性质的函数只有少数几种。首先，我们知道，像正弦x和余弦x这样的函数具有这种性质，因此它们似乎是合理的候选函数。另一个可能具有这种性质的函数是带有复指数的指数函数。事实上，由于欧拉公式（$e^{ix} = $余弦$x + i$正弦$x$），指数函数甚至涵盖了正弦$x$和余弦$x$函数。因此，直接猜测（尽管这是一个有根据的猜测）尝试令$y(x) = e^{sx}$（其中s可能是实数，也可能是复数），似乎是一个不错的起点。尝试这种方法后，我们得到以下结果

$$\frac{d^2}{dx^2}e^{sx} + b\frac{d}{dx}e^{sx} + ce^{sx} = 0$$
$$s^2e^{sx} + bse^{sx} + ce^{sx} = 0$$

到目前为止，开局不错。现在，我们只需确定在什么条件下这个表达式才可能成立。由于s^2e^{sx}在任何点都不可能为零（即使s是复数），因此我们可以放心地将等式两边同时除以这个值。这样就大大简化了问题

$$s^2+bs+c=0$$

我们要确定：给定b和c的常数（实数）值后，上述方程是否存在关于s的解。但现在这个表达式只是关于s的一个二次方程，因此我们可以用求根公式来求解。回想一下，解为（注意，对于ax^2+bx+c的形式，我们有$a=1$）

$$s_1=\frac{-b+\sqrt{b^2-4c}}{2} \tag{3.137}$$

$$s_2=\frac{-b-\sqrt{b^2-4c}}{2} \tag{3.138}$$

这个结果告诉我们，除重根的情况（稍后会讨论）外，$y(x)$确实有两个线性无关解。它们是$y_1(x)=e^{s_1x}$和$y_2(x)=e^{s_2x}$。我们可以将这两个解线性组合，得到可能的最一般解。为此，我们需要下面证明的叠加原理。

定理 3.2（二阶常微分方程的叠加原理）。设 $y_1(x)$ 和 $y_2(x)$ 是该常微分方程的解

$$\frac{d^2y}{dx^2}+b\frac{dy}{dx}+cy=0$$

那么，线性组合 $y(x)=\alpha y_1(x)+\beta y_2(x)$ 也是一个解。

证明。令 $L=\frac{d^2}{dx^2}+b\frac{d}{dx}+c$。那么 $L[\alpha y_1+\beta y_2]=\alpha L[y_1]+\beta L[y_2]$，根据线性性可得。由于 $L[y_1]$ 和 $L[y_2]$ 都是解，它们都等于零。这就证明了 $L[\alpha y_1+\beta y_2]=0$。■

根据以上所有信息，我们现在可以将通解写成如下形式

$$y(x)=\alpha e^{s_1x}+\beta e^{s_2x} \tag{3.139}$$

稍加思考这个解，就会发现我们实际上还能进一步说明。二次方程的根可能有以下几种情况。

1. 情形1。根 s_1 和s_2 是互不相同的实数。2. 情形2。根 s_1 和 s_2 均为实数，但并不互不相同（$s_1=s_2$）。3. 情形3。根是如下形式的共轭复数：$s_1=u+i\lambda$，$s_2=u-i\lambda$。

这就是仅有的三种选择。由于方程 秒^2+赫维赛德阶跃函数秒 $+c=0$ 的根基本上决定了解的类型，因此这个方程通常被称为 特征方程。接下来最简单的方法是依次考虑每一种选择。

3.4.3.1 情形 1：根为实数且互异

For this case, the solution is relatively uncomplicated; it is just a linear combination of two exponential functions with real exponents. Thus, the solution is

$$y(x) = \alpha e^{s_1 x} + \beta e^{s_2 x} \tag{3.140}$$

其中 α 和 β 是由两个辅助条件确定的常数。

例 *3.11*（河流中的污染物降解）。当有机碳排入溪流时，随着碳被溪流水中的微生物降解，会导致溪流中的氧气含量降低。假设一家蔬菜加工厂获准将其废水排入一条溪流，有机碳浓度为 20 毫克/升（见图）。假设有机碳持续排放（每天 24 小时），立即混合，并且这种易降解有机碳在溪流中的有效浓度为 $u(x = 0) = u_0 = 2$ 毫克/升，位于排放点处。同时假设对流-弥散-反应方程适用，且所有过程均已达到稳态。

碳将在下游多远处影响氧气水平？我们将“影响”氧气水平解释为有机碳浓度等于或高于 $u = 0.05$ 毫克/升。假设速度为 $v_0 = 1.2$ 千米/小时，弥散系数为 $D_0 = 0.7$ 千米^2/小时，且一阶降解速率等于 $k_1 = 0.3$ 小时$^{-1}$。根据问题的物理组织情况，确定一组合理的边界条件，以便求得解。

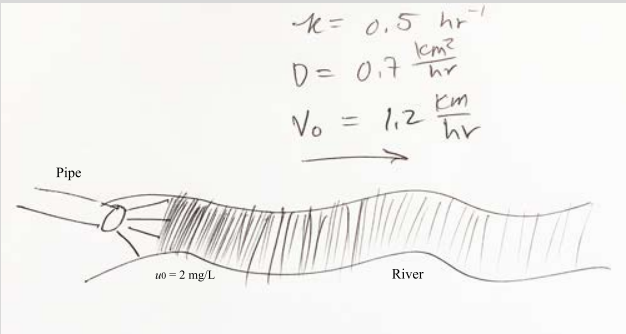


图 3.9：一条溪流。由人工精心绘制。**S**

m

<escape>**解答**。问题陈述表明，适当的平衡方程应采用以下形式。</escape>

$$D_0 \frac{d^2u}{dx^2} - v_0 \frac{du}{dx} - k_1 u(x) = 0$$

Or, upon rearranging a bit,

$$\frac{d^2u}{dx^2} - \frac{v_0}{D_0}\frac{du}{dx} - \frac{k_1}{D_0}u(x) = 0$$

Taking $a = 1$, $b = -v_0/D_0$, and $c = -k_1/D_0$. Using the quadratic formula, we have the roots

$$s_1 = \frac{1}{2}\left(\frac{v_0}{D_0} + \sqrt{\left(\frac{v_0}{D_0}\right)^2 + 4\frac{k_1}{D_0}}\right)$$
$$s_2 = \frac{1}{2}\left(\frac{v_0}{D_0} - \sqrt{\left(\frac{v_0}{D_0}\right)^2 + 4\frac{k_1}{D_0}}\right)$$

通解为:

$$u(x) = \alpha \exp\left(\frac{v_0}{2D_0}x\right)\exp\left(\frac{1}{2}x\sqrt{\left(\frac{v_0}{D_0}\right)^2 + 4\frac{k_1}{D_0}}\right)$$
$$+ \beta \exp\left(\frac{v_0}{2D_0}x\right)\exp\left(-\frac{1}{2}x\sqrt{\left(\frac{v_0}{D_0}\right)^2 + 4\frac{k_1}{D_0}}\right)$$

为了确定解，我们需要施加一些辅助条件，以便更充分地定义问题（在本例中，由于这些条件适用于空间，我们可以称它们为边界 条件）。$x = 0$ 处的条件并不难确定。显然，我们希望此处的浓度为u_0, ，因此可以设$u(0) = u_0$。至于第二个辅助条件，情况就不一定那么明了。不过，稍加思考便可得出以下结论：源下游的浓度只能降低 ，不能增加。这是因为该问题本质上是恒定速度下的守恒传输（伴有弥散展宽），再加上衰减反应。或者，更简单地说，我们可以施加这样的条件：源下游的浓度不能变得任意高，尽管它可以趋于零。这两个条件都与这样一个观点一致：第一个指数函数项的常数必须为零。为什么？因为如果它不为零，那么随着沿流向远离源，浓度将呈指数增长并变得极大——根据我们对该问题的直觉，这在物理 上说不通。更“数学化”、但含义相同的说法是：我们预期当x 变得任意大时，碳的浓度会变得任意小。无论采用哪种说法，我们现在都得到了这个解。

$$u(x) = \beta \exp\left(\frac{v_0}{2D_0}x\right)\exp\left(\frac{1}{2}x\sqrt{-\left(\frac{v_0}{D_0}\right)^2 + 4\frac{k_1}{D_0}}\right)$$
$$u(0) = u_0$$

Evaluating this solution at $x = 0$ and using the boundary condition at $x = 0$, we rapidly find the final solution

$$u(x) = u_0 \exp\left(\frac{v_0}{2D_0}x\right) \exp\left(-\frac{1}{2}x\sqrt{\left(\frac{v_0}{D_0}\right)^2 + 4\frac{k_1}{D_0}}\right)$$

A plot of this function appears below.

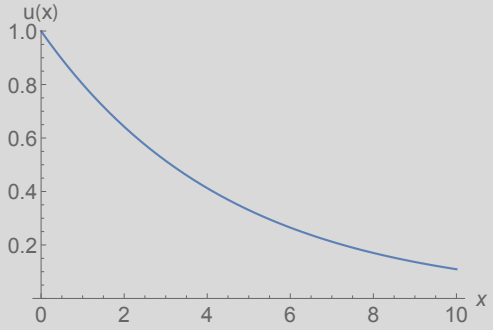


图3.10：解$u(x)$的图像

The last part of the question asks for the distance at which the carbon content becomes acceptable- i.e., less than 0.05 mg/L. We can determine this by setting

$$\frac{u(x)}{u_0} = \exp\left(\frac{v_0}{2D_0}x - \frac{1}{2}x\sqrt{\left(\frac{v_0}{D_0}\right)^2 + 4\frac{k_1}{D_0}}\right)$$

Or, taking the natural log of both sides

$$\ln\left(\frac{u(x)}{u_0}\right) = \left(\frac{v_0}{2D_0}x - \frac{1}{2}x\sqrt{\left(\frac{v_0}{D_0}\right)^2 + 4\frac{k_1}{D_0}}\right)$$

Solving for x gives

$$x = \ln\left(\frac{u(x)}{u_0}\right)\left(\frac{v_0}{2D_0} - \frac{1}{2}\sqrt{\left(\frac{v_0}{D_0}\right)^2 + 4\frac{k_1}{D_0}}\right)^{-1}$$

并且，将上述参数值代入后，我们得到该解。

$$x = \ln\left(\frac{0.05}{2}\right)\left(\frac{1.2\ km/h}{2\times 0.7\ km^2/h} - \frac{1}{2}\sqrt{\left(\frac{1.2\ km/h}{0.7\ km^2/h}\right)^2 + 4\frac{0.5\ h^{-1}}{0.7\ km^2/h}}\right)^{-1}$$

$$x = 16.62\ km$$

3.4.3.2 Case 2: The Roots are Real, but Not Distinct

This is the case where the terms in the radical of the characteristic equation $s^2+bs+c=0$, is zero. Or, equivalently, we have that $b^2=4c$, so that the only root is $s=-b/2$. This presents a little bit of a problem, because we know that we need *two independent* equations to form the general solution to a second-order ODE.

解决这个问题的方法是求出第二个解。回想一下，我们最初寻找指数解的方式在某种程度上是经验性的；我们考虑了哪些类型的函数可能以某种方式组合，从而使二阶齐次常微分方程的原型可能成立。指数函数就是一个候选函数，最终证明它确实可行。我们知道，第一个解会生成如下形式的指数函数：$y_1(x)=\alpha e^{-b/2x}$。由于缺乏明确的方向，我们可以尝试寻找一个与第一个解成比例的第二个解。例如，我们可以提出这样的形式：

$$y_2(x)=f(x)e^{-b/2x}$$

And hope that this works. Well, to be quite honest, this is roughly how the process unfolded. If we are guessing a functions that might work for $f(x)$, we might try the simplest thing we can think of.

$$y_2(x)=xe^{-b/2x}$$

Of course, this suggestion wouldn't have been made if it wasn't going to work. To see that it does, we need only take some derivatives

$$\begin{aligned}
y_2(x) &= xe^{-xb/2}\\
y_2'(x) &= e^{-xb/2}-x\frac{b}{2}e^{-xb/2}\\
y_2''(x) &= -\frac{b}{2}e^{-xb/2}-\frac{b}{2}e^{-xb/2}+x\left(\frac{b}{2}\right)^2e^{-xb/2}\\
&= -be^{-xb/2}+x\frac{b^2}{4}e^{-xb/2}
\end{aligned}$$

Now, substitute this into

$$\frac{d^2y}{dx^2}+b\frac{dy}{dx}+cy=0 \tag{3.141}$$

给出

$$-be^{-b/2x}+x\frac{b^2}{4}e^{-b/2x}+be^{-b/2x}-x\frac{b^2}{2}e^{-b/2x}+cxe^{-b/2x}=0$$

And, after simplifying, this is

$$\left[-\frac{b^2}{4}+c\right]xe^{-b/2x}=0$$

Finally, recalling that for this case we have $b^2=4c\Rightarrow -b^2/4+c=0$, we have proved that y_2 is a solution.

Example 3.12 (Case 2 Example). Consider the following steady diffusion problem

$$\begin{aligned} D\frac{d^2u}{dx^2} &= 0 \\ u(0) &= u_0 \quad \text{(This is a specified concentration boundary condition)} \\ -D\left.\frac{du}{dx}\right|_{x=L} &= 0 \quad \text{(This is a } \textit{no-flux} \text{ boundary condition.)} \end{aligned}$$

Show that this solution is a special case of the general solution for Case 2.

解答。对于这个例子，我们有 $a = 1$ 和 $b = c = 0$。因此，特征方程的唯一根为 秒$_1$ =秒$_2 = 0$。这是一个重根（尽管它为零），这意味着通解为

$$\begin{aligned} u(x) &= \alpha e^0 + \beta x e^0 \\ &= \alpha + \beta x \end{aligned}$$

Note that this solution is actually a line, which is indeed a special case of the general exponential solution. Using the two boundary conditions, we have

$$\begin{aligned} u_0 &= \alpha \\ 0 &= \beta \end{aligned}$$

Thus, our final solution is

$$u(x) = u_0 \text{ (a constant)}$$

3.4.3.3 Case 3: The Roots Are Complex Conjugates

For this case, the characteristic equation $(s^2 + bx + c)$ is such that $b^2 - 4c < 0$; therefore, one is faced with complex roots coming from the quadratic equation. i.e.,

$$\begin{aligned} s_1 &= -\frac{b}{2} + \frac{1}{2}\sqrt{b^2 - 4c} \\ s_2 &= -\frac{b}{2} - \frac{1}{2}\sqrt{b^2 - 4c} \end{aligned}$$

因为$b^2 - 4c < 0$，根式会产生一个虚数。为方便起见，令$u = -b/2$ 且$i\lambda = 1/2\sqrt{b^2 - 4c}$。于是，这两个根可以表示为以下形式

$$\begin{aligned} s_1 &= \mu + i\lambda \\ s_2 &= \mu - i\lambda \end{aligned}$$

which is a complex conjugate pair ($s_1 = \overline{s_2}$). Our general solution is, therefore

$$\begin{aligned} y(x) &= \alpha e^{\mu x + i\lambda x} + \beta e^{\mu x - i\lambda x} \\ &= e^{\mu x}\left(\alpha e^{i\lambda x} + \beta e^{-i\lambda x}\right) \end{aligned} \tag{3.142}$$

回顾欧拉公式，我们可以将两个线性无关解写为

$$\begin{aligned} y_1(x) &= e^{\mu x}e^{ix} = e^{\mu x}\left[\cos(\lambda x) + i\sin(\lambda x)\right] \\ y_2(x) &= e^{\mu x}e^{-ix} = e^{\mu x}\left[\cos(\lambda x) - i\sin(\lambda x)\right] \end{aligned}$$

Now, if we were satisfied with solutions that, in general, contained complex numbers (and, in some cases, such as quantum mechanics, we might very well be!), then we would be done. However, for the vast majority of problems in science and engineering, the problems are such that we expect a real number as an answer. This is actually no problem. We can, if we like, construct two new solutions from the existing one by creating linear combinations of them. Recall, that when we make linear combinations, we can multiply either of the two equations by any constant we like- including complex numbers. Suppose, then, that we construct two new independent solutions as follows

$$\begin{aligned} y_A(x) &= e^{\mu x}\frac{1}{2}\left[y_1(x) + y_2(x)\right] \\ &= e^{\mu x}\frac{1}{2}\left[(\cos(\lambda x) + i\sin(\lambda x)) + (\cos(\lambda x) - i\sin(\lambda x))\right] \\ &= e^{\mu x}\frac{1}{2}(2\cos(\lambda x)) \\ &= e^{\mu x}\cos(\lambda x) \end{aligned}$$

and

$$\begin{aligned} y_B(x) &= e^{\mu x}\frac{i}{2}\left[y_1(x) - y_2(x)\right] \\ &= -e^{\mu x}\frac{i}{2}\left[(\cos(\lambda x) + i\sin(\lambda x)) - (\cos(\lambda x) - i\sin(\lambda x))\right] \\ &= -e^{\mu x}\frac{i}{2}(2i\sin(\lambda x)) \\ &= e^{\mu x}\sin(\lambda x) \end{aligned}$$

With a little thought, we might have realized this from the start. Our two independent solutions can be considered to be either $e^{\mu x}\sin x$ and $e^{\mu x}\cos x$ or $e^{\mu x}e^{ix}$ and $e^{\mu x}e^{-ix}$. It actually makes no difference mathematically- one set of solutions can always be constructed from the other. However, for applications where we desire solutions that are real functions (rather than complex functions), we are better off using the trigonometric solutions. Thus, we can take our general solution to be in this case

$$\begin{aligned} y(x) &= \alpha y_A(x) + \beta y_B(x) \\ &= \alpha e^{\mu x}\cos(\lambda x) + \beta e^{\mu x}\sin(\lambda x) \end{aligned}$$

Or, factoring out the exponential

$$y(x) = e^{\mu x}(\alpha\cos(\lambda x) + \beta\sin(\lambda x)) \tag{3.143}$$

示例 *3.13*（谐振子。）谐振子 只是描述某种周期性弹跳的事物时用到的一个非常复杂的术语。这可以是时钟的摆、弹簧上的一个重物，或者是首席执行官或辉瑞（口号：“辉瑞品质！”）在八月下旬一个没有月亮的星期三夜晚，在莫哈韦沙漠里莫名其妙地在蹦床上弹跳。假设我们讨论的是弹簧上的重物（抱歉）。起初，弹簧被拉长了 10 厘米。假设重物的质量为 1千克，弹簧的弹簧常数为 $K = 3$ 牛顿/厘米。假设该系统是理想的，因此永远不会损失能量（诚然，这不可能；但在较短的时间内，往往非常接近现实！）。描述重物振荡的函数是什么？假设运动发生在无摩擦水平面上（例如，在空气曲棍球桌上），这样就不必担心引力。

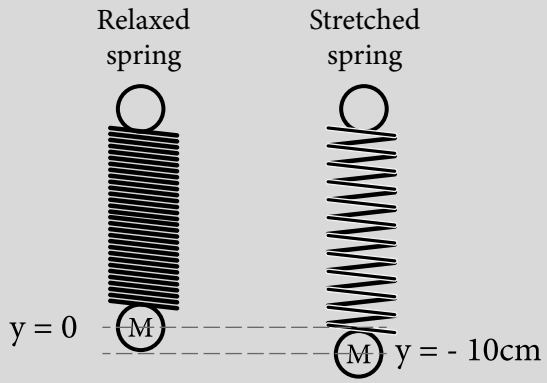


Fig. 3.11: A weight on a spring- with motion in the horizontal plane. On an air hockey table.

Solution. A simple force balance on the mass just as it is let go is as follows

$$m\frac{dv}{dt} = K\Delta y$$

where dv/dt is the acceleration (time rate of change of velocity), g is the gravitational constant (9.81 m/s^2), and Δy is the initial displacement $\Delta y = 0-(-10)\ cm = 10\ cm$. Because the equilibrium position is $y(0) = 0$, we have that $\Delta y = -y(t)$. Recall that $v(t) = dy/dt$. With this in mind, our force balance can be written (with two ancillary conditions based on the physical system)

$$\frac{d^2y}{dt^2}+\frac{K}{m}y(t)=0$$

condition 1	$y(0)=-10$	initial position at y=-10 cm
condition 2	$y'(0)=0$	zero initial velocity

The characteristic equation for this problem is $s^2+K/m=0$ ($a=1$, $b=0$, $c=K/m$). Thus the two roots are

$$s_1=0+\frac{1}{2}\sqrt{-4\frac{K}{m}}=i\sqrt{\frac{K}{m}}$$
$$s_2=0-\frac{1}{2}\sqrt{-4\frac{K}{m}}=-i\sqrt{\frac{K}{m}}$$

Using the notation established above, we have $\mu=0$ and $\lambda=\sqrt{K/m}$. Therefore, the solution that we want is

$$y(t)=e^0\left(\alpha\cos(t\sqrt{K/m})+\beta\sin(t\sqrt{K/m})\right)$$

Taking the derivative (for use with the second ancillary condition)

$$y'(t)=\left(-\sqrt{K/m}\alpha\sin(t\sqrt{K/m})+\sqrt{K/m}\beta\cos(t\sqrt{K/m})\right)$$

Using the two ancillary conditions, we have

$$-10=\alpha\cos(0)+\beta\sin(0)\Rightarrow\alpha=-10$$
$$0=-\sqrt{K/m}\alpha\sin(0)+\sqrt{K/m}\beta\cos(0)\Rightarrow\beta=0$$

So, our solution is

$$y(t)=-10\cos(t\sqrt{K/m}) \tag{3.144}$$

Does this solution make sense? We can check. First of all, we should check to see if the initial conditions (the two ancillary conditions) are met. Note: $y'(x)=-10\sqrt{K/m}\sin(t\sqrt{K/m})$.

$$y(0)=-10\cos(0)$$
$$y'(0)=-10\sqrt{K/m}\sin(0)=0$$

因此，是的，我们的解确实满足初始条件。很容易验证，它也满足常微分方程（求导数，并将结果代回原常微分方程即可）。下面给出了 $0<t<60$ 秒时的振荡图像。

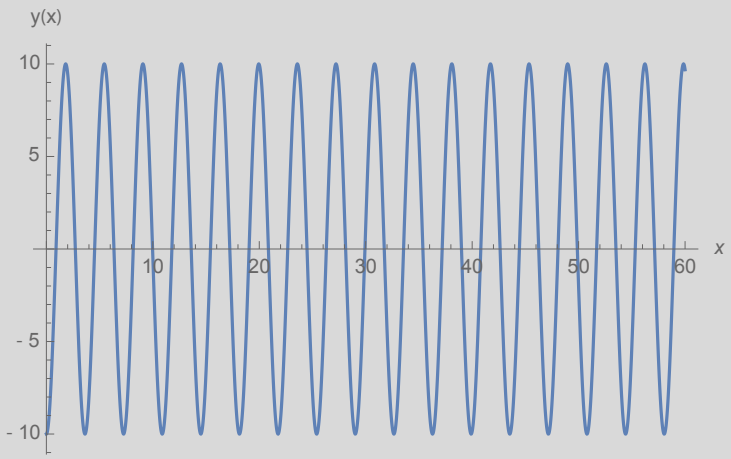


Fig. 3.12: Plot of solution for the weight on a spring.

3.4.4 Solutions for Nonhomogeneous Second-Order ODEs with Constant Coefficients

We will study only one approach for nonhomogeneous second-order ODEs. In particular, we are getting around to solving the problem given by

$$\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy(x) = g(x) \tag{3.145}$$

As a matter of notation, we call any solution to this problem *a particular* solution, $y_p(x)$.

The method that we are going to use is called *variation of parameters*, and it was developed by two somewhat famous mathematicians: Leonard Euler (of Euler identity fame), and Joseph-Louis Lagrange. The basic idea behind the method is as follows. Suppose we have the two solutions to the homogeneous second-order equation

$$\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy(x) = 0 \tag{3.146}$$

We call the general solution to this problem the *homogeneous* solution . Because we will make use of $y(x)$ to indicate the solution to Eq. (3.145), we will use $y_h(x)$ to indicate the homogeneous solution

$$y_h(x) = \alpha y_1(x) + \beta y_2(x)$$

Recall that we previously defined the operator notation for general second-order ODEs to be

$$L = \frac{d^2}{dx^2} + b\frac{d}{dx} + c = 0 \tag{3.147}$$

so that $L[y_h] = 0$ gives us Eq. (3.146). This is really nothing but a shorthand notation. So, now we consider the problem $L[y_p] = g(x)$– the nonhomogeneous problem. There are an infinite number of particular solutions, y_p. To see this, suppose we have *any* particular solution to Eq. (3.145). We need only examine the sum

$$\begin{aligned} y(x) &= y_p(x) + \gamma y_h(x) \\ &= y_p(x) + \alpha y_1(x) + \beta y_2(x) \end{aligned}$$

其中 γ 是一个常数。然后注意

$$L[y(x)] = L[y_p(x)] + \gamma L[y_h(x)]$$

但是，根据定义 $L[y_h] = L[\alpha y_1(x) + \beta y_2(x)] = 0$，并且 $L[y_p] = g(x)$，所以

$$L[y(x)] = g(x)$$

因此，如果我们有任何特解，就总是可以加上齐次解的任意倍数，从而得到一个新的解。之所以可能需要加入齐次解，具体原因在于，我们找到的特解可能不满足必要的辅助条件。然而，我们知道可以在齐次解中找到 α 和 β 的特定值，以满足可能的边界条件。因此，我们可以利用所添加的齐次解来调整特解，使其满足必要的边界条件。

参数变易法可以用格林函数理论来解释，但我们这里不会采用这种方法（首先，这需要对分布理论有一定了解）。另外，我们来思考一下所求解 $y(x)$ 的性质。显然，$y(x)$ 不可能是 $y_1(x)$ 或 $y_2(x)$ 的常数倍，否则它就会是齐次方程的一个解。换句话说，我们意识到

$$\frac{y(x)}{y_1(x)} \neq C_1 \qquad \frac{y(x)}{y_2(x)} \neq C_2$$

其中C_1 和C_2 为常数。此时的论证似乎简单得近乎荒谬。如果上述比值不可能等于常数，那么它们就必定等于函数。具体而言，是x的函数，因为这是我们所考虑的唯一自变量。因此，我们可以这样论证：

$$\frac{y(x)}{y_1(x)} = u_1(x) \qquad \frac{y(x)}{y_2(x)} = u_2(x)$$

For some (currently unknown) functions u_1 and u_2. This is equivalent to stating that a solution for $y_p(x)$ must be of the form

$$y_p(x) = u_1(x)y_1(x) + u_2(x)y_2(x)$$

Now, if you really think about it, at this juncture we haven't really done much. In fact, if we have any three bounded and continuous functions, f, g, and h, it is not difficult to show that we can always find two functions u_1 and u_2 such that $h = u_1 f + u_2 g$. The real brilliance of Eq. (3.4.4) is that we already know that $L[y_1] = L[y_2] = 0$. So, working with Eq. (3.4.4) to find the functions u_1 and u_2 is going to be made much simpler by this fact. This becomes clearer when we actually try to do this. We are looking for the solution

$$L[y_p] = g(x)$$

or, substituting

$$L[u_1(x)y_1(x) + u_2(x)y_2(x)] = g(x)$$

Or, writing this all out more painfully

$$\frac{d^2}{dy^2}[u_1y_1 + u_2y_2] + b\frac{d}{dy}[u_1y_1 + u_2y_2] + c[u_1y_1 + u_2y_2] = g(x) \tag{3.148}$$

Let's start by noting

$$\frac{d}{dy}[u_1(x)y_1(x)] = u_1'y_1 + u_1y_1'$$
$$\frac{d}{dy}[u_2(x)y_2(x)] = u_2'y_2 + u_2y_2'$$

因此,
使得

$$\begin{aligned}\frac{d}{dy}[u_1(x)y_1(x) + u_2(x)y_2(x)] &= u_1'y_1 + u_1y_1' + u_2'y_2 + u_2y_2' \\ &= (u_1'y_1 + u_2'y_2) + (u_1y_1' + u_2y_2')\end{aligned}$$

and

$$\begin{aligned}\frac{d^2}{dy^2}[u_1y_1 + u_2y_2] &= \frac{d}{dy}[(u_1'y_1 + u_2'y_2) + (u_1y_1' + u_2y_2')] \\ &= (u_1'y_1 + u_2'y_2)' + u_1'y_1' + u_1y_1'' + u_2'y_2' + u_2y_2''\end{aligned}$$

Substituting these results back into Eq. (3.148), we have

$$\begin{aligned}&(u_1'y_1 + u_2'y_2)' + u_1'y_1' + u_1y_1'' + u_2'y_2' + u_2y_2'' \\ &+b[(u_1'y_1 + u_2'y_2) + (u_1y_1' + u_2y_2')] \\ &+c[u_1y_1 + u_2y_2] = g(x)\end{aligned}$$

From here, we collect all of the terms that are multiplied by u_1 or u_2. This result is

$$\begin{aligned}&u_1(y_1'' + by_1' + cy_1) + u_2(y_2'' + by_2' + cy_2) \\ &+u_1'y_1' + +u_2'y_2' \\ &+b(u_1'y_1 + u_2'y_2) \\ &+(u_1'y_1 + u_2'y_2)' = g(x)\end{aligned}$$

This simplified immediately because y_1 and y_2 are solutions to the homogeneous problem $y'' + by' + cy = 0$, leaving us with

$$u_1'y_1' + +u_2'y_2' + b(u_1'y_1 + u_2'y_2) + (u_1'y_1 + u_2'y_2)' = g(x)$$

At this juncture, we have one equation, but two unknown functions u_1 and u_2. This kind of a problem has an infinite number of solutions; we only need one solution, and any particular solution will do. So, we are free to add any additional equation that we like (as long as it is independent) to make our system solvable. A particularly convenient requirement is to set $u_1'y_1 + u_2'y_2 = 0$; you can see that doing so is reasonably clever because it eliminates *two* groups of terms in the equation above! In summary, then, we have the two equations

$$\begin{aligned} u_1'y_1 + u_2'y_2 &= 0 \\ u_1'y_1' + +u_2'y_2' &= g(x) \end{aligned}$$

which is two equations in the unknowns u_1' and u_2'. In fact, we can arrange this in a matrix

$$\begin{bmatrix} y_1 & y_2 \\ y'_1 & y'_2 \end{bmatrix} \begin{bmatrix} u'_1 \\ u'_2 \end{bmatrix} = \begin{bmatrix} 0 \\ g(x) \end{bmatrix}$$

and then use Cramer's rule to solve for u_1' and u_2'. First note that

$$\begin{vmatrix} y_1 & y_2 \\ y'_1 & y'_2 \end{vmatrix} = y_1y_2' - y_2y_1'$$

So, Cramer's rule gives us

$$\frac{du_1}{dx} = \frac{1}{y_1y_2' - y_2y_1'} \begin{vmatrix} 0 & y_2 \\ g(x) & y'_2 \end{vmatrix} \qquad \frac{du_2}{dx} = \frac{1}{y_1y_2' - y_2y_1'} \begin{vmatrix} y_1 & 0 \\ y'_1 & g(x) \end{vmatrix}$$

or

$$\frac{du_1}{dx} = \frac{-y_2g(x)}{y_1y_2' - y_2y_1'} \qquad \frac{du_2}{dx} = \frac{y_1g(x)}{y_1y_2' - y_2y_1'}$$

虽然这并非必要，但我们可以通过对该表达式两边关于 x 进行积分，将其写成积分形式。请注意，由于我们要寻找的是<escape>*any* </escape>特解，因此可以舍弃此时积分过程中产生的任何常数。

$$u_1(x) = \int_x \frac{-y_2g(x)}{y_1y_2' - y_2y_1'}\,dx \qquad u_2(x) = \int_x \frac{y_1g(x)}{y_1y_2' - y_2y_1'}\,dx$$

This completes our solution. However, to summarize, recall that the two solutions above allow us to generate the following particular solution

$$y_p(x) = u_1(x)y_1(x) + u_2(x)y_2(x)$$

To this particular solution, we must add the general form of the homogeneous solution to get our final solution for the nonhomogeneous equation

$$y(x) = y_p(x) + y_h(x)$$

or

$$y(x) = y_p(x) + \alpha y_1(x) + \beta y_2(x)$$

最后这一步至关重要！如果不加上齐次解，我们就没有两个自由常数可通过两个辅助条件来确定。提醒一下，由于 $L[y_h] = 0$，加上齐次解并不会影响我们满足非齐次二阶常微分方程

$$\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy(x) = g(x)$$

因为加　　上齐次解只会给这个常微分方程加上零！这里举个例子会很有帮助。

例 *3.14*（非齐次二阶常微分方程的解）。前一个例子中的谐振子很适合用于本例，部分原因是我们已经有了齐次解。受迫谐振子是指被迫以不同于其固有频率振荡的谐振子。考虑前一个例子中描述的空气曲棍球桌上的系统。假设我们的重物由磁性材料制成，并且我们在桌子下方放置了一个巨大的电磁阵列，它能产生一种随时间变化的磁力，其变化规律为

$$g(x) = -10\cos(10t)$$

这与我们上面求得的固有频率不同，后者等于 -10余弦$(\sqrt{K/m}t) = -10$余弦$(\sqrt{3}t)$。这样做会发生什么？现在我们要求解的方程组为

$$\frac{d^2y}{dt^2} + \frac{K}{m}y(t) = -10\cos(10t)$$

condition 1	$y(0) = -10$	initial position at y=-10 cm
condition 2	$y'(0) = 0$	zero initial velocity

Recall, our homogeneous solution for this case was

$$y(t) = \alpha\cos(t\sqrt{K/m}) + \beta\sin(t\sqrt{K/m})$$

注意，此处我们有 $y_1(t) =$ 余弦$(t\sqrt{K/m})$、$y_2(t) =$ 正弦$(t\sqrt{K/m})$。按照上面的例子，我们现在要寻求一个具有以下形式的非齐次问题的解。

$$y_p(t) = u_1(t)\cos(t\sqrt{K/m}) + u_2(t)\sin(t\sqrt{K/m})$$

为了求得这个解，我们需要函数 y_1, y_2, y_1', y_2' 和$g(t)$。此外，组合式$y_1y_2' - y_2y_1'$也很有用。

$$\begin{aligned}
y_1 &= \cos(t\sqrt{K/m}) \\
y_2 &= \sin(t\sqrt{K/m}) \\
y_1' &= -\sqrt{K/m}\sin(t\sqrt{K/m}) \\
y_2' &= \sqrt{K/m}\cos(t\sqrt{K/m}) \\
y_1y_2' - y_2y_1' &= \sqrt{K/m}\cos^2(t\sqrt{K/m}) + \sqrt{K/m}\sin^2(t\sqrt{K/m}) = \sqrt{K/m} \\
g(t) &= -10\cos(10t)
\end{aligned}$$

此时，求解过程只剩下一堆繁琐的计算工作。

$$\begin{aligned}
\frac{du_1}{dt} &= \frac{10\sin(t\sqrt{K/m})\cos(10t)}{\sqrt{K/m}} \\
\frac{du_2}{dt} &= \frac{-10\cos(t\sqrt{K/m})\cos(10t)}{\sqrt{K/m}}
\end{aligned}$$

These can be integrated using standard trigonometric techniques, or by looking the integrals up on a table. The results are

$$\begin{aligned}
u_1(t) &= 10\sqrt{\frac{K}{m}}\left(-\frac{\cos\left(t\left(\sqrt{\frac{K}{m}}-10\right)\right)}{2\left(\sqrt{\frac{K}{m}}-10\right)} - \frac{\cos\left(t\left(\sqrt{\frac{K}{m}}+10\right)\right)}{2\left(\sqrt{\frac{K}{m}}+10\right)}\right) \\
u_2(t) &= -10\sqrt{\frac{K}{m}}\left(\frac{\sin\left(t\left(\sqrt{\frac{K}{m}}-10\right)\right)}{2\left(\sqrt{\frac{K}{m}}-10\right)} + \frac{\sin\left(t\left(\sqrt{\frac{K}{m}}+10\right)\right)}{2\left(\sqrt{\frac{K}{m}}+10\right)}\right)
\end{aligned}$$

Well, worse things have happened... Recalling that $K/m = 3$ and simplifying, our final solution is

$$y(t) = \alpha\cos\left(\sqrt{3}t\right) + \beta\sin\left(\sqrt{3}t\right) + \frac{30}{97}\cos(10t)$$

and its derivative is

$$y'(t) = -\sqrt{3}\alpha\sin\left(\sqrt{3}t\right) + \sqrt{3}\beta\cos\left(\sqrt{3}t\right) - \frac{300}{97}\sin(10t)$$

From the first ancillary condition we have

$$-10 = \alpha\cos(0) + \frac{30}{97}\cos(0) \Rightarrow \alpha = -\frac{1000}{97}$$

From the second ancillary condition we have

$$0 = \sqrt{3}\beta\cos(0) \Rightarrow \beta = 0$$

Apparently, the solution we want is

$$y(t) = -\frac{1000}{97}\cos\left(\sqrt{3}t\right) + \frac{30}{97}\cos(10t)$$

习题

练习题

对于下面的问题，请记住，$y = y(x)$ 是因变量，而 x 是自变量。另请注意，将哪个变量称为自变量并不重要。因此，有些问题可能会用其他自变量和因变量来表述（例如，$u(t)$ 和 t）。

对于以下各题，使用分离变量法求出<escape></*escape*>通解<*escape*></escape>。同时列出每种情况下的因变量。只有在能够明确求出因变量，且不需要付出特别大的计算工作量时，才要求明确求出因变量（例如，不要求求出超过二次的多项式解的根）。

1. $\frac{dy}{dx} = \frac{x}{y}$
2. $\frac{dy}{dx} = y^2 + 1$
3. $\frac{dy}{dx} = x^2$
4. $\frac{dy}{dx} = \sqrt{x}$
5. $\frac{dy}{dx} = \sin x - \cos x$
6. $\frac{dy}{dx} = y^2 + y$
7. $\frac{1+e^t}{1-e^{-y}}\frac{dy}{dt} + e^{t+y} = 0$
8. $\exp(t+y)\frac{dy}{dt} = \frac{t}{y}$, for $t > 0$, $y(t) > 0$
9. $z^3 \frac{du}{dz} = \sqrt{z^2 - u^2 z^2}$
10. $u'(t) = 4tu^2(t)$
11. $y'(t) = mc^2 t^{\pi} y(t)$ where m and c are constants

For the following problems, determine the *general* solution by whatever method is required. If ancillary conditions are specified, then also find the *particular* solution for those conditions.

12. $\frac{dy}{dx} + \frac{1}{x^2}y = 0, x \in (0,1]$
13. $\frac{dy}{dt} + 2y = 4,\ y(0) = 0,\ x \in [0,2]$
14. $\frac{dy}{dx} - ay = f(x),\ x \in [a,b]$其中，$f(x)$ 是某个可积函数。你将无法显式求解它！你必须保留其积分形式。
15. $\frac{dy}{dz} = 2y + z,\ z \in [0,\infty]$
    $y'(0) = 1$
16. $y' + x^5 y = x^5$
17. $y'(t) = y(t) + \frac{1}{1-e^{-t}}$
18. $\frac{dy}{dt} + y\cos t = \cos t,\ t \in [0,\infty], y(0) = 5$
19. $\frac{dy(t)}{dt} - y(t) - t\sin(t) = 0$

For the following problems, determine the *general* solution for the second order ODE.

20. $y''(t) - 5y'(t) - 14y(t) = 0$
21. $y''(x) - 25y(x) = 0$
22. $\frac{d^2y}{dx^2} + 2\frac{dy}{dx} - 24y = 0$
23. $\frac{d^2y}{dx^2} - 6\frac{dy}{dx} + 9y = 0$
24. $\frac{d^2y}{dx^2} + 10\frac{dy}{dx} + 25y = 0$
25. $\frac{d^2y}{dx^2} + 2\sqrt{2}\frac{dy}{dx} + 2y = 0$

26. $\frac{d^2y}{dx^2}+3\frac{dy}{dx}+3y=0$

27. $\frac{d^2y}{dx^2}+2\frac{dy}{dx}+5y=0$

28. $6\frac{d^2y}{dt^2}-4\frac{dy}{dt}=0$

29. $6\frac{d^2y(t)}{dt^2}-4y(t)=0$

For the following problems, determine the *particular* solution for the second order ODE and the two given ancillary conditions.

30. $\frac{d^2y}{dx^2}-2\frac{dy}{dx}-15y=0,\ y(0)=1,\ y'(0)=-1,$ $x\in[0,\infty]$

31. $\frac{d^2y}{dx^2}+16y=0,\quad y(0)=3,\quad y'(0)=12,$ $x\in[0,\infty]$

32. $\frac{d^2y}{dx^2}-10\frac{dy}{dx}+25y=0,\ y(0)=0,\ y(1)=0,$ $x\in[0,1]$

33. $4y''(t)-4y'(t)+y(t)=0,\quad y(0)=1, y'(0)=0, t\in[0,\infty]$

34. Show that if $y_1(x)$ and $y_2(x)$ are each solutions to the ODE

$$\frac{d^2y}{dx^2}+b\frac{dy}{dx}+cy=0$$

那么，对于任意常数C_1和C_2,，函数$y(x)=C_1y_1(x)+C_2y_2(x)$也是一个解。首先，将这个表达式直接代入常微分方程，替换其中的$y(x)$。将含y_1 的项与含y_2的项分别合并。

对于以下各题，求非齐次二阶常微分方程的解。如果给出了辅助条件，则求出未知系数。

35. $y''(t)-y(t)'-2y(t)=2e^{-t},\ y(0)=1,$ $y'(0)=0$

36. $\frac{d^2y}{dx^2}+2\frac{d^2y}{dx^2}+y=3e^{-t}, y(0)=0, y'(0)=1$

37. $\frac{d^2y}{dx^2}-\frac{d^2y}{dx^2}+\frac{1}{4}y=4e^{t/2}, y(0)=1, y'(0)=1$

应用与更具挑战性的问题

1. 重做滑雪者问题（例题 3.2），但作如下修改：将雪面摩擦和空气阻力造成的合阻力都用一个与速度成正比的项表示：$F_{drag}=-\alpha mv(t)$。将此力加入力平衡中，然后用 $\alpha=1\ \mathrm{s}^{-1}$重新求解该问题。

2. 对于这个涉及连续搅拌釜式反应器（体积为 $\forall$）、流量为Q且发生一级反应的问题，确定进水浓度y_{in} 突然变化时，连续搅拌釜式反应器中的浓度将如何变化，已知其初始浓度为y_0。

a. The mass balance and initial condition are given by

$$\frac{dy}{dt}+(\frac{1}{\tau}+k)y=\frac{1}{\tau}y_{in} \tag{3.149}$$

$$y(t=0)=y_0 \tag{3.150}$$

Recall here that $\tau=\forall/Q$ (the hydraulic residence time), $\forall$ is the volume of the reactor, Q is the flow rate at both the inlet and the outlet, y_{in} is the influent concentration, y_0 is the initial concentration in the tank, and k is the first-order reaction rate constant.

求解该问题，求出 $y(t)$。你可以考虑进行代换 $\beta = (\frac{1}{\tau} + k)$，以简化分析过程（并避免处理复杂代数运算时出现错误），但如果采用这种方法，切记最后要换回原始变量。

b. 对于条件 $\forall = 10\ \mathrm{m}^3$, $Q = 1\ \mathrm{m}^3$/小时、$k = 0.4$ 小时$^{-1}$、$y_0 = 2$ 摩尔/m^3 和 $y_{in} = 1$ 摩尔/m^3, , 绘制该问题在 $0 \leq t \leq 100$ 小时内的解。在同一张图上，绘制参数相同但 $k = 0$ 小时$^{-1}$ 的解。评论这两个解的主要区别。每个问题的稳态解是什么（$t \longrightarrow \infty$）？

3. 催化剂有时会发生降解，原因可能有很多。例如热降解、被其他反应性化学物质中毒（这些物质会改变催化位点的氧化态或不可逆地结合到位点上），以及处理物流中的组分造成的污垢（对催化剂产生物理改变）。

假设我们有催化反应 $A \xrightarrow{C} B$，其中物种 A 在活性催化位点（以物种C表示）存在时转化为物种 B。假设催化剂由于低pH下发生的反应过程而降解。

We can model a catalytic reaction by creating balances for both the catalyst (y_C) and the chemical being catalysed (chemical species A). A possible set of balances for catalytic degradation during the production of the three chemical species (reactant A, product B, and catalyst C) is as follows.

$$\frac{dy_C(t)}{dt} = -k_C y_C(t) \tag{3.151}$$

$$y_C(0) = C_0 \tag{3.152}$$

$$\frac{dy_A(t)}{dt} = -k_A y_A(t) y_C(t) \tag{3.153}$$

$$y_A(0) = A_0 \tag{3.154}$$

$$\frac{dy_B(t)}{dt} = k_B y_A(t) y_C(t) \tag{3.155}$$

$$y_B(0) = B_0 \tag{3.156}$$

We can solve this problem in a step-wise fashion. First, note that the balance for the catalyst, C, does not depend at all on the concentrations of the other two species. Thus, an explicit expression for the concentration $y_C(t)$ is straightforward to evaluate. For the problem, do the following.

a. 求解催化剂的平衡方程，以确定$y_C(t)$ 的表达式。 b. 将$y_C(t)$ 的表达式代入物种A的平衡方程右端。求解该表达式以得到$y_A(t)$。 c. 确定$y_A(t)$ 和$y_C(t)$ 的解后，最后一步是将它们代入物种B的平衡方程右端。求解所得表达式以得到$y_B(t)$。你可能需要使用 Mathematica 等符号积分软件来计算相关积分。 d. 在区间 $0 \leq t \leq 20$ 分钟内绘制所得函数，使用以下常数列表。为使图更易于查看，请绘制归一化值y_A/A_0, y_B/A_0, 和y_C/C_0。

$$k_A = 0.2 \text{ min}^{-1} \qquad A_0 = 10 \text{ mmol/L} \tag{3.157}$$

$$k_B = 0.3 \text{ min}^{-1} \qquad B_0 = 0 \text{ mmol/L} \tag{3.158}$$

$$k_C = 0.25 \text{ min}^{-1} \qquad C_0 = 1 \text{ mmol/L} \tag{3.159}$$

4. 有时，问题是可分离的，但所得的解在因变量中是隐式的 。例如，考虑以下问题：在间歇反应器中，通过酶动力学使某种化学物种发生降解。如果适用米氏动力学，则充分搅拌的间歇系统的一个初值问题可表述为

$$\frac{dy}{dt} = -\mu \frac{y}{y+K} \tag{3.160}$$

$$y(0) = y_0 \tag{3.161}$$

其中y 是化学物种的浓度（质量/体积），k 是反应速率参数（质量/时间），而K 是半饱和常数（质量/体积）。对于这个问题，请完成以下内容。

a. 用简单的分离变量法求解此题。通过使用定积分求解，或在使用不定积分时根据初始条件求出积分常数，将初始条件纳入计算。 b. 结果以变量 t 表示时是显式的，但以变量 y 表示时是隐式的。假设我们有 $y_0 = 20$毫摩尔/升， $\mu = 0.1$ 毫摩尔·秒$^{-1}$，以及 $K = 8$ 毫摩尔/升。假设你想求出浓度介于 $0 \leq y \leq 20$ 之间时，约在 $0 \leq t \leq 50$ 秒的时间范围内的方程解。你该如何做到这一点？提示：试着调换因变量和自变量的角色，把时间视为因变量。。c. 使用你方便使用的绘图软件绘制结果。如果你使用 Mathematica，可以使用绘图命令 ParametricPlot。你可以在 Mathematica 的帮助页面中查看如何使用该命令。

5. Suppose we have a completely stirred reactor that, because of a problem with the pump, the flow rate decreases in time. Specifically, assume

$$Q(t) = \frac{Q_0}{\alpha + \gamma t} \tag{3.162}$$

Assume that the initial condition is $y(0) = 0$. Where α and γ are parameters that control the decrease in flow rate over time. The revised mass balance equation can be taken for this case as

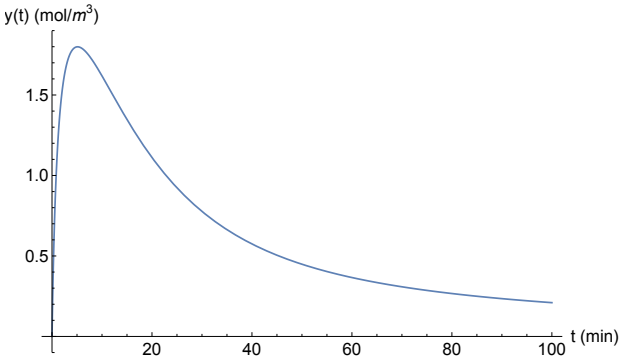
$$\frac{dy}{dt} = \frac{1}{\tau}\left(\frac{1}{\alpha+\beta t}\right) y_{in} - \frac{1}{\tau}\left(\frac{1}{\alpha+\beta t}\right) y(t) - k y(t) \tag{3.163}$$

a. Begin by rewriting this expression so that you have a single term on the left-hand side that can be associated with $P(t)$, and a single term on the right hand side associated with $Q(t)$. Separate from the ODE, write down the explicit functions representing $P(t)$ and Q(t).

b. Determine the integrating factor by first computing $s(t)$ from

$$s(t) = \int P(t) dt \tag{3.164}$$

图3.13 应用题5的解。



Then, compute explicit expressions for $e^{s(t)}$ and $e^{-s(t)}$. You should need to use the logarithmic identity $a\ln(x) = \ln(x^a)$.

c. Finally, using Mathematica (or similar symbolic integration software), determine the result

$$y(t) = e^{-s(t)}\left[\int Q(t)e^{s(t)}dt + C_1\right] \tag{3.165}$$

计算这个积分颇具挑战性，并将得到一个用前文介绍过的伽马函数表示的解。请注意，在 Mathematica 中，你需要通过对常数和变量施加条件来计算这个积分，如下所示。

```
Integrate[Q(t), {t, 0, t}, 假设 -> {{t, k, t} ∈ 实数集 && {t, k, tau} ≥ 0}]
```

（此处将用适当的表达式代入$Q(t)$）。

d. 注意到初始条件$y(0) = 0$，求C_1的适当取值。

e. 最后，使用数值 $k = 0.1$ 分钟$^{-1}$ 和 $\tau = 0.5$ 分钟、$a = 1$ 分钟以及$b = 1$，使用任意合适的软件绘制解在值域 $0 \leq t \leq 100$ 分钟内的图像。所得图像应与图 3.13 中的曲线相似。所得函数是否符合你对废水浓度随时间变化情况的直觉？为什么符合或不符合？

6. 弥散是化学物种输运过程中发生的展宽，其成因包括速度场的变化以及分子扩散的影响。有时，对于平推流反应器，人们会对反应器进行无弥散模型化；其理念是，减少展宽是平推流反应器的目标之一。

对于这个问题，假设我们有一个长度为 1 m 的平推流反应器，该反应器能够催化反应物 A 转化为产物B，并分别在考虑和不考虑水动力弥散过程的情况下，对系统的稳态进行建模。暂时假设我们只关注物种 A （反应物）的浓度演化。该问题的轮廓如下所示。

Without hydrodynamic dispersion, the steady state differential mass balance equation is given by the first order expression

$$-u_0\frac{dy}{dx} = ky(x) \tag{3.166}$$

$$y(0) = y_0 \tag{3.167}$$

Including hydrodynamic dispersion, the steady state differential mass balance equation is given by the first order expression

$$D\frac{d^2y}{dx^2} - u_0\frac{dy}{dx} = ky(x)$$
$$y(0) = y_0$$
$$y'(L) = 0$$

（注意，由于现在的问题是二阶的，我们还需要一个额外的辅助（边界）条件来确定常数。）首先，对于每个问题，先将方程两边除以使最高阶导数 的系数等于 1 的常数。这一步通常能让问题更容易理解。请记住，在解完之前不要代入任何数值。

a. 求解表示无弥散对流的第一个问题。注意，要解这个问题，考虑系统的停留时间会很有帮助。系统长度为L 单位，流体速度为u_0。一份流体从$t = 0$ 时进入系统，到离开系统所需的时间为

$$L = u_0\tau \tag{3.168}$$

其中， τ 定义为平推流反应器的停留时间。显然，停留时间就是长度除以速度，即 $\tau = L/u_0$。还要注意，在这个问题中，变量x 和t 之间存在关系。任何位置x 都可以通过关系$t = x/u_0$转换为与之对应的停留时间（截至该位置）。因此，对于纯对流系统，可以在时间和空间位置之间进行转换。

b. Using $D = 1\times10^{-9}$ m^2 s^{-1}, $u_0 = 1\times10^{-8}$ m s^{-1}, $k = 1\times10^{-8}$ s^{-1}, and $y_0 = 100$ mmol/L, solve this problem over the interval $0 < x < L$, with $L = 1$ m. Plot the two solutions using whatever software is convenient.

c. How do the two solutions compare? Why do you think they are different from one another?

7. Steady state heat transport in a radiator fin, assuming that the fin is thin compared to its length, can be described by the following second-order ODE. Here, we are imposing boundary conditions at the left of a specified temperature ($y(0) = y_0 = 400$ K), and the right boundary as a zero gradient condition.

$$K\frac{d^2y}{dx^2} - \frac{h}{\ell}y = -\frac{h}{\ell}y_\infty$$
$$y(0) = y_0$$
$$\left.\frac{dy}{dx}\right|_{x=L} = 0$$

Physically, this equation represents the temperature, y, in the fin assuming that there is both heat conduction, and convective cooling by an external air source blowing over the radiator fin.

Assume the fin is $L = 0.1$ m long, with thermal conductivity $K = 385$ W m^{-1}K^{-1}, and with heat transfer coefficient $h = 150$W m^{-2}K^{-1}. The thickness of the fin is $\ell = 0.002$ m. Assume the far-field temperature, y_∞ is 293 K. Solve this problem. Plot the solution over the domain and range $0 < x < L$ and $293 < T < 400$. Please remember not to substitute any numeric values into the solution until the solution is complete.

"Fourier's Theorem . . . is not only one of the most beautiful results of modern analysis, but may be said to furnish an indispensable instrument in the treatment of nearly every recondite question in modern physics."

Lord Kelvin, Treatise on Natural Philosophy (1912)

4

Fourier Series Part I: Introductory Concepts

Fourier series are one of the most widely used, and powerful tools in all of applied mathematics, with applications that range from signal analysis, to compression algorithms for making smaller computer files. Before investigating Fourier series in detail, it will be useful to revisit some basic concepts about infinite series in general. Again, it is helpful to start with some terminology.

4.1 Terminology

In the study of Fourier series, a number of new concepts arise. For convenience, some of the more important definitions and vocabulary are summarized here.

- **Fourier series.** A series expressed in terms of a weighted sum of sine or cosine functions (or both). The sine and cosine functions are of increasing frequency, and the weight functions represent the corresponding amplitude for that function. As an example, a general Fourier sine series on $x \in [0,1]$ is given in terms of the amplitudes B_n and frequencies (modulated by the integer n) as follows

$$f(x) = \sum_{n=1}^{\infty} B_n \sin(nPix)$$

- **振幅。** 周期函数的振幅是其垂直方向模的最大值。对于形式为B正弦x的正弦函数，系数B表示其振幅。例如，如果$B = 2$，那么随着x变化，该函数在 ± 2 之间振荡。

- **Period.** A periodic function, $f(x)$, is said to have a period P_λ if $f(x) = f(x+nP)$ for all positive integers n. In other words, the period is the distance (space or time) along the $x-$axis that defines the

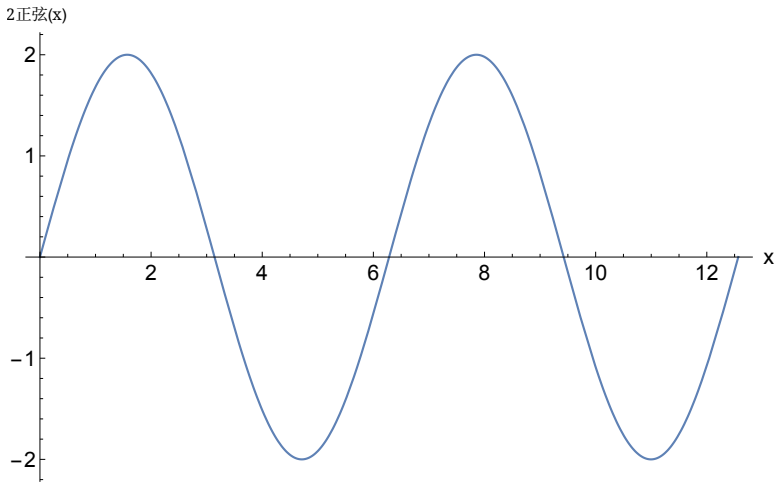


图 4.1 周期函数$f(x) = $正弦$(x)$在区间 $0 \le x \le 4\pi$ 上的图像。该函数的振幅为$B = 2$。该函数的周期为$P = 2\pi$；函数每隔 2π 的整数倍就会重复，因此对于所有正整数 n，$f(x) = f(n2\pi x)$。

repeating unit of the periodic function. When the independent variable is space, sometimes P_λ is called the *wavelength* and given the symbol λ. The subscript λ is added to the the symbol for the period (P_λ) and the frequency (F_λ, defined below) so that these symbols are defined uniquely (and as a reminder of their connection with the concept of wavelength). For sine and cosine functions in the form $\sin(n\pi x)$ or $\cos(n\pi x)$, the period is given by

$$P_\lambda = \frac{2\pi}{n}$$

• **频率**。如果自变量是时间，周期有时也用频率表示。由于频率一词经常使用，即使自变量不是时间，人们也习惯谈论频率。频率最适合理解为周期函数在单位长度或单位时间内完成的完整周期数。周期P_λ与频率F之间的关系为

$$F = \frac{1}{P_\lambda}$$

因此，如果周期为$P_\lambda = 2\pi$，频率为$F_\lambda = 1/(2\pi)$。这表示自变量每经过 2π 个单位，就会出现一个完整周期。对于形式为正弦$(n\pi x)$或余弦$(n\pi x)$的正弦函数和余弦函数，其周期为

$$F = \frac{1}{P_\lambda} = \frac{n}{2\pi}$$

The SI units for frequency are the Hertz (Hz) which measures the number of periods completed per second. However- take caution! The terminology *period* is sometimes used even when the independent variable is something other than time.

• **收敛**。级数的收敛是一个复杂的主题。我们曾简要讨论过泰勒级数的收敛概念。对于傅里叶级数，我们所研究的函数都将在某种有用的意义下收敛。对于连续函数，级数将逐点收敛。这意味着对于任意 x

chosen in the domain, the partial sums evaluated at x_0, $(S_n(x_0))$, get as close to the value of the function at x $(f(x_0))$ as we like by increasing n sufficiently. For discontinuous functions, the Fourier series converge in a sense that will be discussed later in the chapter.

- **谱**。我们将发现，函数f的傅里叶级数通常是无穷级数，由正弦函数、余弦函数或二者的加权和构成。对于每个这样的函数，级数都以整数n为索引，如上所述，n 同时表示频率和振幅。于是，函数f的谱由振幅相对于n值的离散（但无限.....）图像给出。
- **收敛**。级数（或序列）随着项数趋于无穷而达到有限极限的性质。在第2章中，讨论了级数收敛的两种方式——逐点收敛和一致收敛，其中一致收敛是更严格的一种方式。

4.2 复习：幂级数

我们大多数人都熟悉无穷级数，因为微积分入门课程会介绍它们。理解泰勒级数有一种非常直观的方式。首先，为了铺垫相关内容，考虑将多项式拟合到一组数据的问题。例如，我们都知道，两点可以唯一确定一条直线；类似地，三点可以确定一个二次式，四点可以确定一个三次式，依此类推。以下定理从更一般的角度支持了这一结论：可以用多项式精确拟合$n+1$个点（我们陈述该定理，但不予证明）

定理 4.1。 对于任意一组 k 个 $+1$ 点 (x_i, y_i)，其中

- i 是一个指标，满足 $i = 1, \ldots k+1$
- 对于所有 $i \neq j$，都有 $xi \neq x$ j

那么存在一个唯一的 次数至多为 k 的多项式，使得对于所有 $i = 1, \ldots, k+1$，$f(x_i) = y_i$。

这个定理引出了一个有趣的问题。假设我们有一个定义在某个区间$I = (a,b)$上的连续函数$f(x)$（假设各阶导数也都存在且有界，因此$f(x)$是一个C^∞ 函数）。现在，假设我们在区间I内取样 n 个$n+1$点。显然，通过拟合一个 n 阶多项式来匹配这$n+1$个采样点，我们可以得到函数$f(x)$的某种近似。实际上，可以使用 Lagrange 多项式插值等方法来实现这一点（这里不作讨论）。

Now, by definition, all polynomials are of finite order. The order of the highest power can be as *large* as you like, but it cannot be infinity; thus there are a finite number of terms. The interesting question that might occur after thinking about this problem is whether or not one can consider a kind of "infinite" polynomial that could fit continuous curves exactly. Being somewhat loose, we might wonder of something like an "infinite" polynomial would correspond to fitting an "infinite" number of points on the interval I. We would expect such a polynomial to provide an exact (in some sense) representation of the continuous function $f(x)$.

While this is not necessarily a rigorous mathematical way of posing the question, it does offer a kind of analogy that provides some basis for working with what are called power series. Recall, a polynomial of order n is given by

$$f(x) = \sum_{n=0}^{n=N} a_n (x-b)^n$$

where, here a *shift* of the amount b has been included in the polynomial. A power series, then, is the generalization of this idea to the limit $n \to \infty$.

$$f(x) = \sum_{n=0}^{n=\infty} a_n (x-b)^n$$

请注意，该级数中的任意有限项都是一个多项式；然而，幂级数将求和扩展到无穷。尽管幂级数与多项式具有相似性，但在适当的极限下，幂级数和多项式确实会表现出一些截然不同的行为。具体而言，我们注意到以下几点。

1. 首先，所有多项式在每个有限区间内都收敛到一个有限值，$I = (a,b)$，其中a和b是实数。2. 反过来，并非每个幂级数都收敛。有限多项式在任意有限区间上都有有限值，而“无限”多项式可能在任何区间上都不收敛。3. 不难证明，当x趋于正无穷或负无穷时，任意多项式都会趋于正无穷或负无穷。换句话说，对于一个$f(x)$为n^{th}阶的多项式，我们有

$$\left| \lim_{x \to \infty} f(x) \right| = \infty$$

4. 与n阶多项式相比，幂级数在x趋于无穷时，可以收敛到任意值，包括有限值。从某些方面来看，这实际上是一个相当奇怪的结果。例如，假设表示$f(x)$的幂级数具有以下行为

$$\lim_{x \to \infty} f(x) = \lim_{x \to \infty} \sum_{n=0}^{n=\infty} a_n (x-b)^n = 0$$

当然，确实存在具有这种行为的幂级数（例如，函数$f(x)$ =指数函数$(-x^2)$的幂级数就具有这种行为）。然而，对于仅使用N项的幂级数的任意有限近似，无论N有多大，我们都有

$$\lim_{x \to \infty} \left| \sum_{n=0}^{n=N} a_n (x-b)^n \right| = \infty$$

这是非常奇怪的现象，因为当$x \to \infty$时，幂级数趋于零；然而，该级数的每一个有限近似都会发散至$\pm\infty$。不过，概念从有限情形伸展到无限情形时，往往会出现违反直觉的现象！

4.3 Review: Taylor Series

The Taylor series should be a familiar concept from introductory calculus. It was named after the English mathematician Brook Taylor who worked on problems of calculus in the early 1700's. Although primarily known for his work on calculus, his interest ranged widely (as was typical of the times), even authoring a treatise under the unusual title *On the Lawfulness of Eating Blood* which was discovered, upon his demise, to be among his unpublished papers.

The Taylor series is actually derivable directly from the definition of the power series. Note, if we take the nth derivative of the power series we find the following

$$\begin{aligned}
&\textit{first derivative} && f'(x) = a_1 + 2a_2(x-b) + 3a_3(x-b)^2 + \dots \\
&\textit{second derivative} && f''(x) = 2a_2 + 6a_3(x-b) + \dots \\
&n^{th}\textit{derivative} && f^{(n)}(x) = n!a_n + (n+1)!a_{n+1}(x-b) + \dots
\end{aligned}$$

Starting from this result, if we now set $x = b$, we find that all of the terms except the first are identically zero. The coefficients for the power series above are then defined by

$$\begin{aligned}
f'(b) &= a_1 \\
\frac{1}{2!}f''(b) &= a_2 \\
\frac{1}{n!}f^{(n)}(b) &= a_n
\end{aligned}$$

代入 tuting this result into the power series above, gives us the *Taylor series* around the point $x = b$

$$f(x) = \sum_{n=0}^{n=\infty} (x-b)^n \frac{f^{(n)}(b)}{n!}$$

Sometimes the series written around $b = 0$ is called the *Maclaurin* series. Setting $b = 0$ gives this series

$$f(x) = \sum_{n=0}^{n=\infty} x^n \frac{f^{(n)}(0)}{n!}$$

Any function that has a convergent Taylor series a set of points, $x \in X$ is called *analytic* on X. Many series (such as the exponential function) converge for all possible values of x, and are thus analytic everywhere. Any finite truncation of a Taylor series is called a *Taylor polynomial*. Note that every finite truncation of a Taylor series is actually a polynomial! Also true is that analytic functions converge nicely: as one increase the number of terms in the sum approximating the function f, the result gets uniformly closer to the actual value of f. An immediate consequence of these properties is that, for every function that is *analytic* in some domain, X, there is a polynomial that can represent that function as closely as we like. Essentially, this follows from the definition of an analytic function, and the Taylor polynomial.

实例 ***4.1***（泰勒级数/麦克劳林级数实例）。函数 $f(x) = \exp(x)$ 是一个值得探讨泰勒级数展开的有趣例子，部分原因在于它在定义域 $x \in (-\infty, \infty)$ 内处处恰好收敛。开始分析时，请注意，我们首先需要求出 $f(x)$ 的无限多个导数。这看起来或许是一项艰巨的任务，但函数的导数往往呈现出可以加以利用的规律（使用归纳原理）。考虑以下内容

$$
\begin{aligned}
f(x) &= \exp(x) \\
f'(x) &= \exp(x) \\
f''(x) &= \exp(x) \\
&\dots
\end{aligned}
$$

这里，我们不难看出其中的规律。对于任意 n 阶导数，其导数由方程 $f^{(n)}(x) = \exp(x)$ 给出。现在，假设我们在点 $x = 0$ 处展开级数（即选择平移参数 b 为零，这也正是麦克劳林级数的定义）。那么，每个导数的值都为 1：$f^{(n)}(0) = 1$。由此得到的泰勒级数为

$$
\begin{aligned}
\exp(x) &= \sum_{n=0}^{n=\infty} x^n \frac{1}{n!} \\
&= 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} + \frac{x^6}{6!} + \dots
\end{aligned}
$$

As a second case, suppose we would like to find the Taylor series for $\sin(x)$.

$$
\begin{aligned}
f(x) &= \sin(x) \\
f'(x) &= (-1)^2 \cos(x) \\
f''(x) &= (-1)^3 \sin(x) \\
f'''(x) &= (-1)^4 \cos(x) \\
&\dots
\end{aligned}
$$

In principle, we can define the Taylor series to be expanded around any real number. And there are good reasons to do that in some cases; for example, if you know you need a particular expansion to be accurate in a specific interval, $I = (a_1, a_2)$, it would be good to choose $a_1 < b < a_2$ for improved accuracy in the expansion. In this example, no particular point was stated, so we adopt $x = 0$ for convenience. Under those conditions, we find

$$
\begin{aligned}
f(0) &= 0 \\
f'(0) &= 1 \\
f''(0) &= 0 \\
f'''(0) &= -1 \\
f^{(4)}(0) &= 0 \\
f^{(5)}(0) &= 1 \\
&\dots
\end{aligned}
$$

In other words, we want only the *odd* terms, and those terms alternate sign. A little thought, will indicate that the function $\text{odd}(n) = 2n+1$, $n = 0, 1, 2, \dots$ counts by odd numbers only. To switch signs we multiply by powers of -1; we can define a function that alternates sign by $\text{alt}(n) = (-1)^n$. Putting this together, we find the following result valid for any value of n

$$f^{(2n+1)}(0) = \text{alt}(n) = (-1)^n$$

Note the following examples

$$\begin{array}{lll} n=0 & 2n+1=1 & f^{(1)}(0) = (-1)^0 = +1 \\ n=1 & 2n+1=3 & f^{(3)}(0) = (-1)^1 = -1 \\ n=2 & 2n+1=5 & f^{(5)}(0) = (-1)^2 = +1 \end{array}$$

Recalling the definition of the Maclaurin series

$$f(x) = \sum_{n=0}^{n=\infty} x^n \frac{f^{(n)}(0)}{n!}$$

and noting again that only the odd terms are non-zero, the series representation for $\sin(x)$ is

$$\begin{aligned} f(x) &= \sum_{n=0}^{n=\infty} x^{2n+1} \frac{f^{(2n+1)}(0)}{(2n+1)!} \\ &= \sum_{n=0}^{n=\infty} x^{2n+1} \frac{(-1)^n}{(2n+1)!} \\ &= x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \ldots \end{aligned}$$

Plots of finite Taylor polynomial (Taylor series truncated at order n) are given in the figure below.

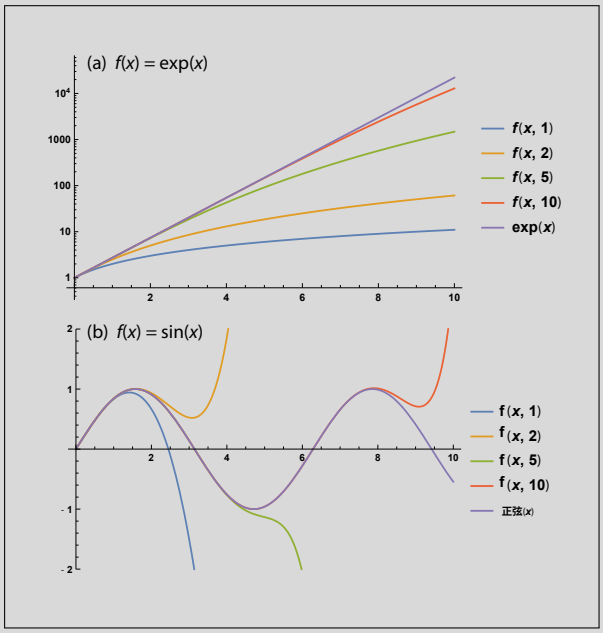


Fig. 4.2: (a) Exponential, and (b) Sine functions plotted for 1, 2, 5, and 10 terms in the approximating Taylor polynomial.

特定函数的泰勒级数通常可以直接获得，而且根据定义直接计算它们也并非难事。不过请注意，所有泰勒级数都有一个收敛半径。换句话说，泰勒级数会在某个定义域内收敛（即随着求和中项数的增加，趋向于一个确定的数值，而这个确定的数值等于f在该点的值）。通常很容易判断一个级数是否收敛，但要确定其收敛半径则困难得多。本章不讨论级数的收敛性质，但如果已知，我们会谨慎列出其收敛半径。以下给出一些众所周知的泰勒级数：

$$
\begin{aligned}
e^x &= \sum_{n=0}^{\infty} \frac{x^n}{n!} && = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots && (converges\ for\ all\ x)\\
\log(1+x) &= \sum_{n=1}^{\infty} (-1)^{n+1} \frac{x^n}{n} && = x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots && (converges\ for\ |x| < 1)\\
\frac{1}{1-x} &= \sum_{n=0}^{\infty} x^n && = x + x^2 + x^3 + \ldots && (converges\ for\ |x| < 1)\\
\frac{1}{(1-x)^2} &= \sum_{n=1}^{\infty} n x^{n-1} && = 1 + 2x + 3x^2 + 4x^3 + \ldots && (converges\ for\ |x| < 1)\\
\sin x &= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1} && = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots && (converges\ for\ all\ x)\\
\cos x &= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n} && = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots && (converges\ for\ all\ x)
\end{aligned}
$$

像这样的级数最有用的性质之一，是可以用来计算超越函数的值（回顾第1章的内容，超越函数简而言之是一种无法表示为有限多项式的解析函数）。因此，泰勒级数是计算超越函数近似值的少数可用方法之一。想一想，除此之外你还能如何计算例如 $\sin \pi/12$ 的值？你可以像古希腊人那样通过图形估算（比如画一个巨大的单位圆），但要使用实数进行计算，就需要级数解（或其他算法）。我们已经习惯于按下计算器上的“sin”按钮，但实际上，计算 sin 函数的真实值非常困难！在遥远的过去（计算器出现之前——天哪！），人们使用级数解来计算有用的超越函数（如 sin、cos、自然对数等）的值，并将结果刊载在非常厚重的书中。工程师和科学家每当想知道 sin 或 cos 的某个特定值时，都必须在巨大的表格中查找这些数值。让我们为计算器欢呼吧！

总之，关于级数的这番讨论旨在强调几个关键概念。

1. Power series can be interpreted as an extension of the concept of polynomials to infinite degree.
2. Taylor and Maclauren series are a special cases of power series.

3. Power series may converge only on some finite interval; it is also possible that they do not converge at all.
4. Because the sum is an infinite one, the limiting behavior of a series (e.g., as $|x| \to \infty$) might not be well approximated by a finite truncation of the series.
5. Infinite sums frequently thwart intuition. Infinite sums can create behavior that would otherwise not be expected!

4.4 Trigonometric Series

Power series are not the only kind of series that provides useful results. In fact, expansions in *trigonometric series* are probably more widely used than any other kind of series expansion. A trigonometric series is one that contains a trigonometric function that varies with the index of the sum. A few examples illustrate some examples of trigonometric series on the interval (0,1). In each of these examples, n is an integer greater than or equal to zero.

$$(a) \qquad B_n = -50\cos\left(\frac{n\pi}{r} + \frac{\pi}{4}\right) \qquad f(x) = \sum_{n=1}^{\infty} B_n \sin(n\pi x)$$

$$(b) \qquad A_n = -50\cos\left(\frac{n\pi}{r} + \frac{\pi}{4}\right)) \qquad f(x) = \sum_{n=1}^{\infty} A_n \cos(n\pi x)$$

$$(c) \qquad B_n = \frac{\sin\left(\frac{\pi n}{2}\right)}{n} \qquad f(x) = \sum_{n=1}^{\infty} B_n \sin(n\pi x)$$

$$(c) \qquad A_n = \frac{\text{正弦}(\frac{n}{2}\pi)}{n} \qquad f(x) = \sum_{n=1}^{\infty} A_n \cos(n\pi x)$$

The plots of these series appear in Fig. 4.3 where here we have used A_n to indicate the coefficients of cosine series, and B_n to indicate the coefficient for sine series.

As you can see in these results, trigonometric series can allow the representation of some very peculiar functions. Some of the interesting features about the particular functions plotted above are as follows.

- Although the two functions used to create the new functions are C^∞ (i.e., infinitely smooth analytic functions), the series seem to give functions appear to be, in some cases, non-differentiable at every point.
- Although the two functions used to create the new functions are C^∞ (i.e., infinitely smooth analytic functions), the series seem to give functions appear to be, in some cases, discontinuous.
- The sine and cosine series give very different series solutions even when $A_n = B_n$.

Although all of these examples are interesting, it is not clear that any of them correspond to any classical function that we recognize- either polynomial or transcendental. So, although we certainly can generate many interesting trigonometric series, the questions arises: “if I am given a specific, known (polynomial

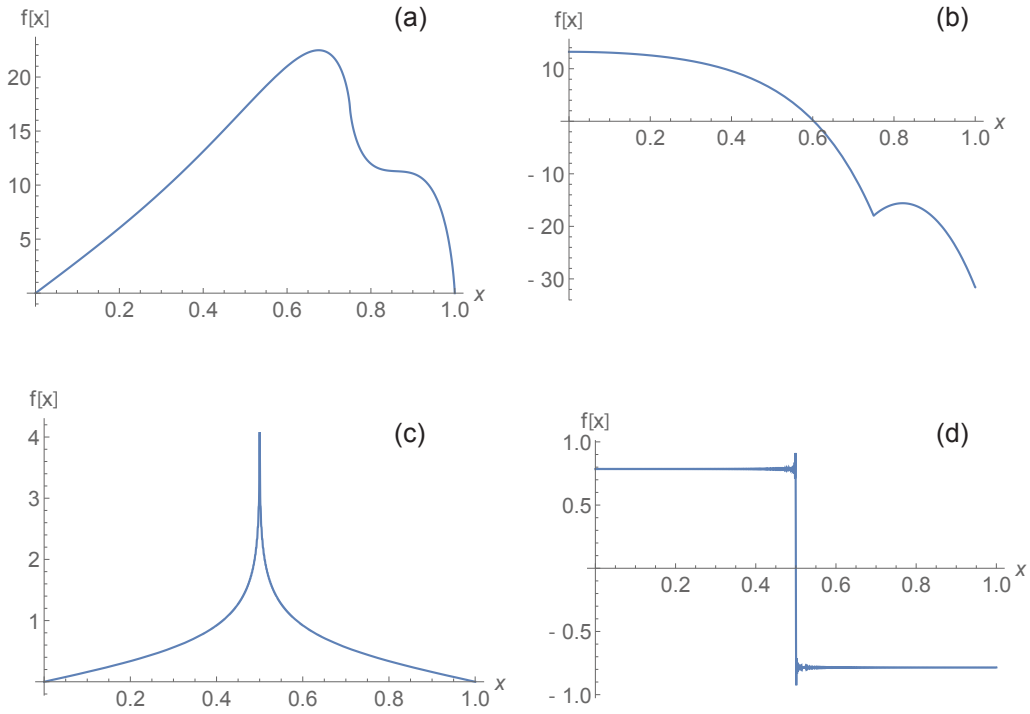


Fig. 4.3: The functions given by (a)-(d) above.

or transcendental function) on a finite interval, can I determine a trigonometric series for that?" This is a question that occurred to Joseph Fourier, a French engineer, in the early 1800's. The answer to this question, remarkably, is a resounding yes for almost any function that one can imagine. In fact, modern mathematics was dramatically shaped by the quest for the answer to this question. Not only did it cause mathematicians to re-think the notion of what a function is, but it caused them to refine the mathematical methods that have led to modern mathematical analysis.

4.5 Fourier Series

对于幂级数，可以将其理解为在一个无限多项式集合中进行展开（尽管严格来说，并不存在无限阶多项式）。对于三角级数，其级数展开自然是三角函数。

为了推导泰勒级数，我们能够建立一种方案，其中我们（1）提出一个包含无限多个（未知）系数的无穷级数展开，a_n，以及（2）确定这些系数的】【。

infinite number of coefficients by finding repeating patterns in the derivatives needed, so that any derivative of order n could be explicitly computed if n were specified.

对于傅里叶级数，我们将采用类似的方法。不过，未知参数不再是无限多个导数的函数，而是可以表示为无限多个积分。乍看之下，这听起来似乎很糟糕，就像泰勒级数的情形一样；但我们会发现，积分中存在重复模式，因此可以推导出关于级数指标n的积分闭式表达式。

傅里叶级数发展的一个关键环节，是将正交性的概念推广到连续函数。乍看之下，这或许不太容易凭直觉理解，但通过类比我们熟悉的有限向量正交概念，就能让它变得直观。

4.5.1 重新审视正交性

在数学学习过程中，我们大多数人都曾接触过两个向量正交这一概念。图 4.4 展示了二维和三维空间中的垂直向量。这里有两个相关概念，有助于我们定义并描述二维和三维空间中的垂直向量。

1. 首先，如果两个向量 **a** 和 **b** 的点积为零，则称它们互相垂直。也就是说，当 $\mathbf{a}\cdot\mathbf{b}=0$时，两个向量互相垂直。回想一下，两个向量的点积等于它们的模之积乘以两者之间夹角的余弦，因此我们有

$$\mathbf{a}\cdot\mathbf{b}=\|\mathbf{a}\|\|\mathbf{b}\|\cos\theta$$

其中，θ是两个向量所在平面内它们之间的夹角。注意，当且仅当 $\theta=\pi/2=90°$时，该值为零。

2. 其次，存在一组相互垂直的基向量，可以将任意向量表示为这些基向量的线性加权和。这些向量通常用符号**i**、**i**和**i** 表示，分别对应x–、y–和z– 方向上的单位向量。因此，在笛卡尔坐标系中，我们有

$$\mathbf{i}=(1,0,0)\qquad \mathbf{j}=(0,1,0)\qquad \mathbf{k}=(0,0,1)$$

利用这些基向量，任意向量$\mathbf{a}=(a_1,a_2,a_3)$都可以表示为以下基向量的加权和：$\mathbf{a}=a_1\mathbf{i}+a_2\mathbf{j}+a_3\mathbf{k}$。

In summary then, every vector can be decomposed into a linear weighted sum of its basis vectors, and each of these basis vectors is, by definition, perpendicular to the others. These ideas are relatively familiar and easy to grasp in 2- and 3-dimensions. Of course, these concepts can be extended to any finite number of dimensions. We cannot visualize such extensions, but mathematically the concepts remain valid. We will discuss in what sense two *functions2* can be perpendicular by making an analogy of functions being approximated by finitely-long vectors.

现在，考虑函数 有限近似 $f(x)=\sin(\pi x)$ 和 $g(x)=\sin(2Pix)$ 在区间 $I=[0,1]$ 上的两个有限近似。在这种情况下，我们所说的有限近似是指

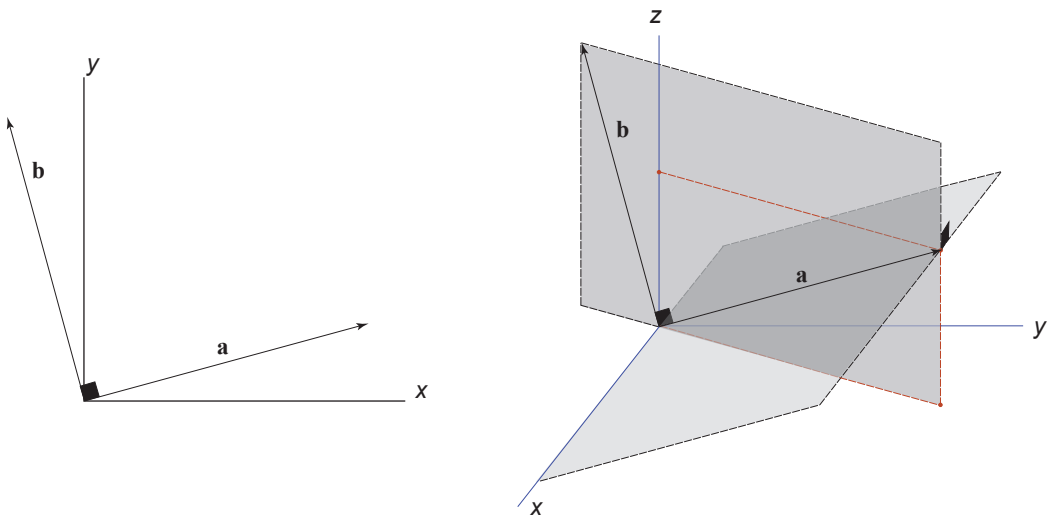


图 4.4：二维和三维中的垂直向量**a**和**b** 。

1. Segmenting the interval into N pieces represented by the vector $\mathbf{x}$, where the spacing is given by $\Delta x_i = 1/N$, and the components of the vector are given by the recursive relationship $x_{i+1} = x_i + \Delta x_i,\ i = 0,1,2,\ldots,N$.

2. 确定 $\sin(\pi x_i)$ 和 $\sin(2\pi x_i)$ 的值，其中$i = 0$，1，2，.....，N。

作为一个简单（且粗略）的例子，考虑 N = 4 的情况。在此情况下，x 的取值为 x =(0,1/4,1/2, 3/4, 1)。为便于符号表示，将 $\sin(\pi x_i)$ 函数的取值向量定义为 ${}_{1\pi}$S =$\sin(\pi x_i)$，i = 0，1，2，3，.....，N。因此，在此例中，结果为 ${}_{1\cdot\pi}$S =(0, $\sin(\pi/4)$, 1, $\sin(\pi/4)$, 0)。我们对 $\sin(2\pi x_i)$ 也采用类似的定义，使得 ${}_{2\pi}$S =(0, 1, 0, −1, 0)。这两个正弦函数及其离散近似的点如图 4.5 所示。

Although this is a rather crude approximation to the two sine functions, note that for our carefully chosen sampling points, we have the following result

$$\begin{aligned} {}_{1\pi}\mathbf{S} \cdot {}_{2\pi}\mathbf{S} &= (0, \sin(\pi/4), 1, \sin(\pi/4), 0) \cdot (0, 1, 0, -1, 0) \\ &= 0 \cdot 0 + \sin(\pi/4) \cdot 1 + \sin(\pi/4) \cdot (-1) + 0 \cdot 0 \\ &= 0 \end{aligned}$$

So, even though this is an approximate representation of the two sine functions, we find that the finite vectors representing the functions themselves are orthogonal. This is a bit more compelling when we use a larger number of points, say $N = 100$. For that case, the approximate discrete functions (each with 101 points) are illustrated in Fig. 4.6.

对于这个离散表示，我们可以看出，所用点的数量足够多，确实捕捉到了两个正弦函数的特征。我们有一长串数，分别对应 **x**、 ${}_{1\pi}\mathbf{S}$ 和 ${}_{2\pi}\mathbf{S}$，如下所示。

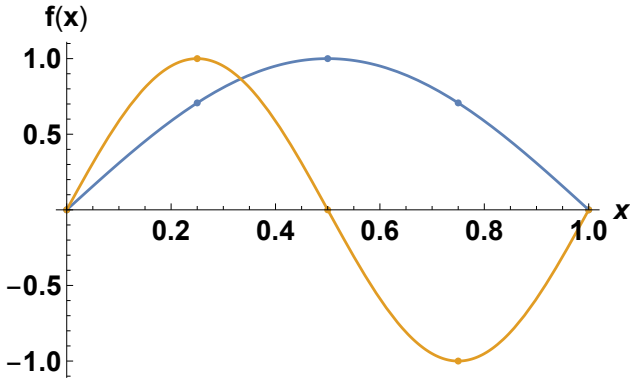


Fig. 4.5: The functions $f(x) = \sin(\pi x)$ and $g(x) = \sin(2\pi x)$, approximated by four discreet sampling intervals (defined by 5 points, $\mathbf{x} = (0, 1/4, 1/2, 3/4, 1)$). This representation results in 5-dimensional vectors representing the values of the sine functions. The continuous representation of these two functions is shown for comparison.

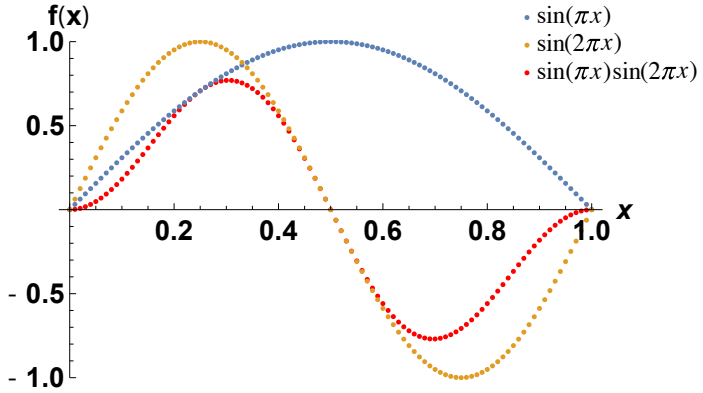


图 4.6：函数 $f(x) =$ 正弦(πx) 和 $g(x) =$ 正弦$(2\pi x)$，通过四个离散采样区间（由 101 个点定义）进行近似。这种表征得到 101 维向量，用于表示正弦函数的取值。

$$\begin{aligned}
\mathbf{x} &= (0, 1/100, 2/100, 3/100, \ldots, 1) \\
{}_{1\pi}\mathbf{S} &= (0, \sin(\pi/100), \sin(2\pi/100), \ldots, 0) \\
{}_{2\pi}\mathbf{S} &= (0, \sin(2\pi/100), \sin(4\pi/100), \ldots, 0)
\end{aligned}$$

While we can verify that ${}_{1\pi}\mathbf{S} \cdot {}_{2\pi}\mathbf{S}$ by direct computation, a look at the graph in Fig. 4.6 makes this somewhat unnecessary. The red points plotted in this graph represent the product of the two sine functions. A little though will indicate that the symmetry of the problem guarantees that the sum of the red points (which represents the dot product ${}_{1\pi}\mathbf{S} \cdot {}_{2\pi}\mathbf{S}$) will sum to zero.

While this is an approximate method to understand how two continuous functions can be though of as being orthogonal, it also is to some extent rigorous. For example, consider the quantity

$$\sum_{i=0}^{i=N+1} {}_{1\pi}\mathbf{S}\Delta x_i = \sum_{i=0}^{i=N+1} \sin(\pi x_i)\Delta x_i$$

In the appropriate limit, this gives

$$\lim_{N\to\infty}\left(\sum_{i=0}^{i=N+1} \sin(\pi x_i)\Delta x_i\right) = \int_0^1 \sin(\pi x)\,dx$$

And, in a similar fashion, we have

$$\sum_{i=0}^{i=N+1} {}_{1\pi}\mathbf{S}\cdot{}_{2\pi}\mathbf{S}\Delta x_i = \sum_{i=0}^{i=N+1} \sin(\pi x_i)\sin(2\pi x)\Delta x_i$$

In the appropriate limit, this gives

$$\lim_{N\to\infty}\left(\sum_{i=0}^{i=N+1} \sin(\pi x_i)\sin(2\pi x)\Delta x_i\right) = \int_0^1 \sin(\pi x_i)\sin(2\pi x)\,dx = 0$$

特别地，很容易验证这个最后的积分恒等于零。因此，两个连续函数彼此正交这一概念，确实与有限维向量正交的情形存在直接且可证明的联系。

4.5.2 Fourier Sine Series

Fourier series are a special kind of trigonometric series that have somewhat astounding properties. Fourier series use only the sine and cosine functions to expand a function $f(x)$ as a series. In starting the investigation of Fourier series, we will consider first only the interval $I = \{x : x \in [0,1]\}$; this interval is sometimes called the *unit interval*. We can expand the definition of the Fourier series to other intervals once we understand how they work on the unit interval.

To start, consider the most basic question that we can ask. If we have a simple, analytic function on an interval $x \in [0,1]$, can we determine the Fourier series for it? To be concrete, let's suppose we have a specific function, say $f(x) = e^{-x}$, and we decide we would like to find a sine series for that function. The question is, can we find a series of the form

$$f(x) = \sum_{n=0}^{\infty} B_n \sin(n\pi x) \tag{4.1}$$

Or, being specific to the example given, $f(x) = e^{-x}$

$$e^{-x} = \sum_{n=0}^{\infty} B_n \sin(n\pi x) \tag{4.2}$$

其中，对于这两个表达式，n 是一个整数，而B_n 是一个无限常数序列（即序列$B_n = (B_1, B_2, B_3,$ $B4,\ldots)$.关于这样一个提出的级数，我们需要探讨两个主要问题，分别是：

1. 是否存在一种方法（算法或构造性证明），可以让我们确定无限常数序列 B_n?

2. 如果我们能够求出常数 B_n，能否证明所得级数收敛到函数 $f(x)$?

继续之前，我们先说明几点。首先，由于 $\sin(0\pi x) = 0$，严格来说，我们不必从 $x = 0$ 开始级数；也可以从 $x = 1$ 开始。按照惯例，正弦级数通常写作

$$e^{-x} = \sum_{n=1}^{\infty} B_n \sin(n\pi x)$$

不过，从 $n = 0$ 开始级数也没有错！函数 $\sin(0\pi x)$ 是正弦级数的一个基函数；只是保留它并不会增加任何内容。其次，请注意下面这个恒等式（我们在上一节讨论正交性时提到过）：

$$\int_{x=0}^{x=1} \sin(n\pi x)\sin(m\pi x)dx = \begin{cases} 0 \; if \; m \neq n \\ \frac{1}{2} \; if \; m = n \end{cases} \tag{4.3}$$

在这些表达式中，n 和 m 都是整数，但不一定相等。由于它与有限长度向量情形下的点积有直接的类比，我们把满足类似于式（4.3）条件的两个函数称为正交 。有了这些信息，我们现在来看一个相当惊人的结果。如果将式（4.2）两边乘以 $\sin(m\pi x)$ 并积分，可得

$$\int_{x=0}^{x=1} e^{-x}\sin(m\pi x)dx = \sum_{n=1}^{\infty} B_n \int_{x=0}^{x=1} \sin(n\pi x)\sin(m\pi x)dx$$

where we have taken the integral inside the sum (and this is always allowable for our purposes; the same is not necessarily true for differentiation!) The reason that this is interesting, is because of what it does to the right-hand side of the equation. When we started, we had an infinite number of values of B_n to contend with. But, because the integral on the right-hand side is only non-zero when $m = n$, then we must have the following

$$\int_{x=0}^{x=1} e^{-x}\sin(n\pi x)dx = B_n \int_{x=0}^{x=1} \sin^2(n\pi x)dx$$

In other words, all of the terms in the series, after integration, are zero, except for the *one term* where n and m are equal. For that one term, the integral becomes the integral of $\sin^2(n\pi x)$ over the interval $x \in [0,1]$. By the identities above, this is exactly $\frac{1}{2}$. Thus, we end up with the result (rearranging a little) that allows us to compute B_n:

$$B_n = 2\int_{x=0}^{x=1} e^{-x}\sin(n\pi x)dx$$

这个积分可以通过分部积分来求解（然后利用它生成一个重复函数这一事实，并将其移到等式一边），也可以查积分表，或者使用 Mathematica 等软件。结果为

$$B_n = \frac{\pi n - e\pi n\cos(\pi n)}{\pi^2 n^2 + 1}$$

尽管这一结果看起来有些奇怪，但我们似乎已经计算出了B_n 在n所有可能取值下的值。换句话说，我们已经得到了计算级数结果所需的全部系数。我们的级数显然具有如下形式

$$e^{-x} = \sum_{n=1}^{\infty} \frac{\pi n - e\pi n\cos(\pi n)}{\pi^2 n^2 + 1}\sin(n\pi x)$$

这有可能是真的吗？有没有办法把一组正弦曲线赋予适当的权重后相加，从而形成指数函数？好吧，让我们来找出答案。下面，我计算了B_n 的前 10 个值。

$$\begin{aligned} B_n = &\{0.790704, 0.196239, 0.287042, 0.099972, 0.173461, 0.0668818, \\ &0.124146, 0.050223, 0.0966368, 0.0402013\} \end{aligned}$$

这些数值表示了每个正弦函数的振幅 或高度，直到$n = 6$。要观察函数本身，我们只需计算$B_n \sin(n\pi x), n = 1,2,3,4,5,6$ 的值。这些函数绘制在图 4.8 中。作为术语说明，请注意，函数 $\sin(n\pi x)$ 称为基 函数。对于每个这样的基函数，整数n 控制基函数的频率 ，而相应的系数B_n 控制基函数的振幅 。

It is clear from the plot in Fig. 4.7 that each of the terms in the series is a sine function, but each with a different amplitude, and each with a frequency that increases with n. The most remarkable thing, however, is what happens when we add these components together. Adding the first 6 terms (as we have above) gives us an approximation to the function $f(x) = e^{-x}$ that is not necessarily good (although, it is not necessarily bad either!). If we compute the first 50 terms, things begin to look much nicer. Finally, with the first 100 terms, the function and the series are almost indistinguishable at most points.

The set of functions $E = \{0,\ \sin(\pi x),\ \sin(2\pi x),\ \sin(3\pi x)\ \ldots\}$ can be thought of as a set of *orthogonal basis functions* from which new functions can be built. There is an analogy with *orthogonal basis vectors* here that was introduced in §4.5.1.

For infinite-dimensional vectors (functions), the equivalent of the dot-product is an integral that is called the *inner product* or sometimes the L^2 (pronounced "ell-two") inner product. Although we will not make extensive use of it, there is even a specific symbol used in mathematics to indicate the L^2 inner product. On the interval $x \in [0,1]$ we would have

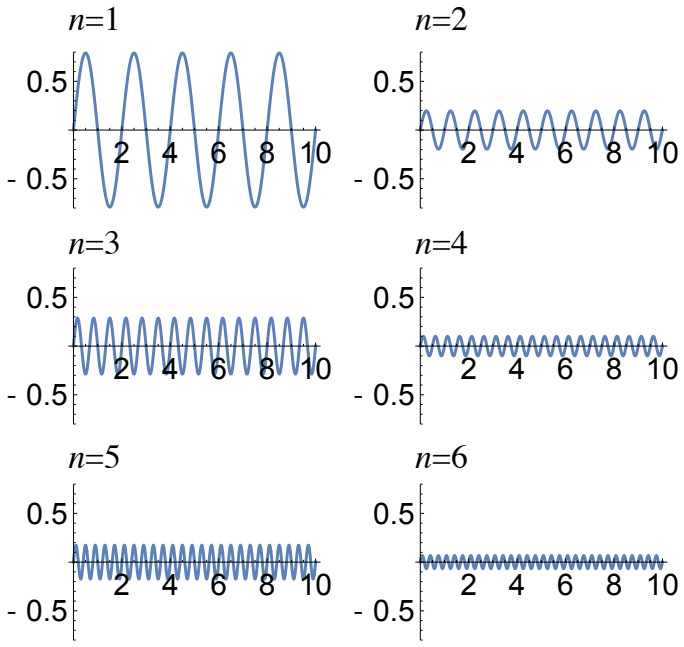


Fig. 4.7: The functions $B_n \sin(n\pi x)$ for $n = 1,2,3,4,5,6$.

$$\langle \sin(n\pi x), \sin(m\pi x) \rangle = \int_{x=0}^{x=1} \sin(n\pi x) \sin(m\pi x)\, dx$$

To summarize the main results for the sine series, we have the following. For any smooth function, $f(x)$, on $x \in [0,1]$, the Fourier sine series is given by

$$f(x) = \sum_{n=0}^{\infty} B_n \sin(n\pi x) \tag{4.4}$$

$$B_n = 2 \int_{x=0}^{x=1} f(x) \sin(n\pi x) dx, \quad n = 0,1,2,3,\ldots \tag{4.5}$$

or, equivalently, because $\sin(0\pi x) = 0 \Rightarrow B_0 = 0$

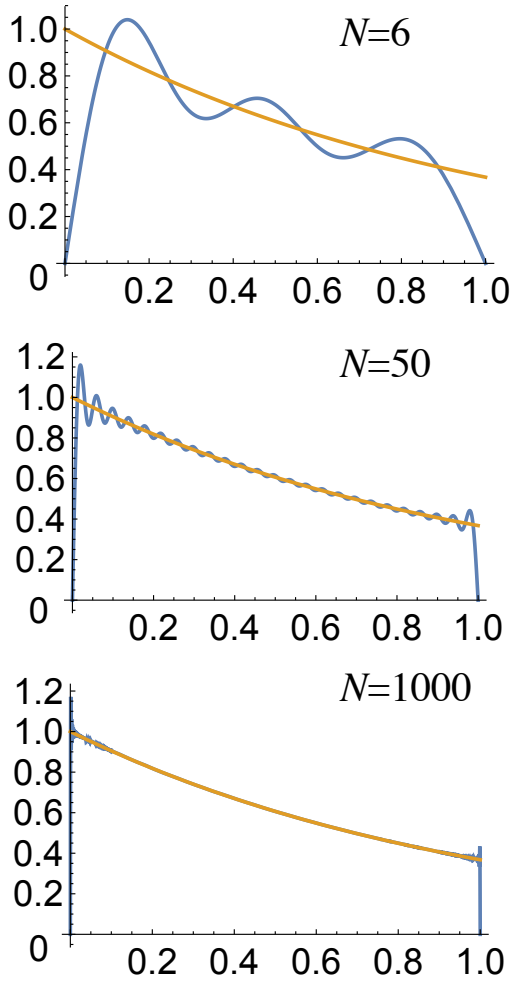


Fig. 4.8: The partial sum of the series, for the total number of terms, N, equal to $N = 6, 50,$ and 1000.

$$f(x) = \sum_{n=1}^{\infty} B_n \sin(n\pi x) \tag{4.6}$$

$$B_n = 2 \int_{x=0}^{x=1} f(x) \sin(n\pi x) dx, \quad n = 1, 2, 3, \ldots \tag{4.7}$$

We can think of the set of functions $E_{sin} = \{0, \sin(\pi x), \sin(2\pi x), \sin(3\pi x) \ \ldots\}$ as being the basis functions from which any function on the unit interval $I = [0, 1]$ can be reconstructed. The reconstruction requires an infinite sum to be made, and the sum is a weighted one, where the weights (amplitudes) are given by the values of B_n. Note that there is an almost exact correspondence here to the case of reconstructing an arbitrary

有限向量**a**，通过计算一组加权基函数的和得到：$\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$。

例 *4.2*（傅里叶正弦级数示例：*Heaviside* 函数）。

定义一个在某个点（例如x_0）左侧等于零、在x_0右侧等于 1 的函数，可能很有用。这种函数称为阶跃函数或*Heaviside* 函数。本章“阶跃函数与δ函数”（第 5 章）将进一步讨论此函数。现在，我们按如下方式定义该函数。

$$H(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases} \tag{4.8}$$

注意，阶跃函数可以按标准方式平移；也就是说，在x_0左侧为零、右侧为一的阶跃函数表示为$H(x - x_0)$。

尽管该函数不是连续函数，我们仍然可以求出它的傅里叶级数。这引出了一个关于傅里叶级数收敛的有趣事实，而这一点最初并未被理解：连续且光滑函数的无穷和可能是不连续的！我们暂且不进一步讨论这一点，继续接受这样的观点：即使是不连续函数，也可以有定义良好的傅里叶级数，并且该级数在某种（尚待说明的）意义下收敛到该函数。假设我们将海维赛德函数向右平移二分之一。回顾其定义：

$$f(x) = \sum_{n=0}^{\infty} B_n \sin(n\pi x) \tag{4.9}$$

$$B_n = 2\int_{x=0}^{x=1} H(x - \tfrac{1}{2}) \sin(n\pi x) dx, \;\; n = 0, 1, 2, 3, \ldots \tag{4.10}$$

Thus, noting that this integral is identically zero from zero to $\frac{1}{2}$, and unity for $x \geq \frac{1}{2}$ we have

$$B_n = 2\int_{x=\frac{1}{2}}^{x=1} \sin(n\pi x)\, dx \tag{4.11}$$

This integral is not difficult to do; the result is

$$B_n = \frac{2}{n\pi}\left[\cos\left(\frac{n\pi}{2}\right) - \cos(n\pi)\right] \tag{4.12}$$

该函数及其近似曲线绘制于图 4.9 中。

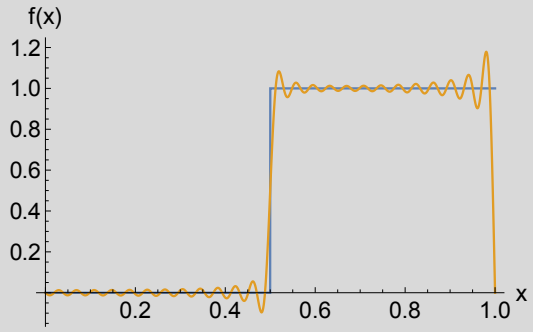


Fig. 4.9: The Heaviside function $H(x-\frac{1}{2})$ (blue) and the approximating cosine series with $n=50$ terms (orange).

You may notice that at the discontinuities, this approximation seems to overshoot and undershoot the actual value by a factor of nearly 10%; this is illustrated by the "spikes" observable at $x=1/2$ and $x=1$. The reason for this phenomenon is somewhat complicated, but it is a consequence of the use of a finite number of terms to approximate the series. While the overshoot and undershoot are not eliminated by the use of more terms, the *width* of the spikes diminishes to zero as $n\to\infty$. This is known as *pointwise* convergence (as discussed in the review given in Chapter 2– see §2.6.3); recall that this is distinctly different from *uniform* convergence. While we will not study convergence extensively in this chapter, it is important to understand the concept of convergence, and how different modes of convergence might look on graphs of the function with increasing n.

4.6 Fourier Cosine Series

在上一节中，我们已经证明，对于"良好"的光滑函数，我们只需运用微积分，就能够推导出傅里叶正弦级数的表征。正如你可能想到的，对于余弦级数也可以采用同样的方法，而且整个推导过程大致相同。

For the sine series, we found that the set of basis functions $E_{\sin}=\{\sin 0\pi x,\ \sin(\pi x),\ \sin(2\pi x),\ \sin(3\pi x)\ \ldots\}$ allowed us to construct a sine series on $x\in[0,1]$ of the form

$$f(x)=\sum_{n=1}^{n\to\infty}B_n\sin(n\pi x)$$

并且，利用函数 $\sin(n\pi x)$ 和 $\sin(m\pi x)$ ($m,n\in\mathbb{N}$) 的正交性，我们能够为所有 振幅系数 B_n 确定一个积分方程。注意，尽管我们没有将其包含在求和中，但从技术上讲，零函数是这组基函数中的函数之一（因为它是与 $\sin(0\pi x)$ 相关联的函数）。由此类比，我们可能会猜想，余弦级数的基函数集合可能是 $E_{\cos}=\{\cos\ 0\pi x,\cos(\pi x),\cos(2\pi x),\cos(3\pi x)\ \ldots\}$。这指出了正弦级数与余弦级数之间一个重要的区别：与$n\ =0$ 相关联的基函数对于余弦级数来说是非零的！具体来说，注意 $\cos(0\pi x)=1$ 这是一种非对称性，会带来一些麻烦，

主要是因为我们总是不得不牢记这一事实。不过，我们仍然可以像之前一样继续进行。首先，我们建议采用以下形式的余弦级数

$$f(x) = \sum_{n=0}^{n\to\infty} A_n \cos(n\pi x) \tag{4.13}$$

此时，级数必须从零开始。同样，按照惯例，这个级数通常会写成允许求和从$n = 1$开始的形式（正如正弦级数的写法一样）。有点荒唐的是，我甚至打算在级数中保留 $\cos(0\pi x)$ 函数的原样。

$$f(x) = A_0' \cos(0\pi x) + \sum_{n=1}^{n\to\infty} A_n \cos(n\pi x) \tag{4.14}$$

这里，我采用了一种不常见的符号表示A_0'来表示傅里叶余弦级数的第一项。这样做的原因是，在本文中，我们将以一种非传统的方式表示这一项，但这种方式很实用，而且能够减少错误。下文将对此作进一步讨论。

与正弦级数完全类似，余弦函数的内积也满足正交条件。请注意。

$$\int_{x=0}^{x=1} \cos(n\pi x)\cos(m\pi x)dx = \begin{cases} 0 & if\ m \neq n \\ \frac{1}{2} & if\ m = n \neq 0 \\ 1 & if\ m = n = 0 \end{cases} \tag{4.15}$$

$$\tag{4.16}$$

注意，$n = m = 0$的情况会得出一个与其他情况（$n,m > 0$）不同的结果。与正弦级数的处理方式相同，现在我们将推导一种确定余弦级数中A_n（以及A_0'）取值的方法。首先，我们可以将方程（4.14）的两边乘以$\cos(m\pi x)$，然后进行积分。由此可得

$$\int_{x=0}^{x=1} f(x)\cos(m\pi x)\,dx = A_0' \int_{x=0}^{x=1} \cos(0\pi x)\cos(m\pi x)\,dx + \sum_{n=1}^{n\to\infty} \int_{x=0}^{x=1} A_n \cos(n\pi x)\cos(m\pi x)\,dx$$

Using the results from Eq. (4.6), we have (and, specifically, recalling that every term in this integrated version of the series is *zero* except for the one term where $n = m$) the following two cases.

- Case 1. $n = m = 0$

$$\int_{x=0}^{x=1} f(x)\cos(0\pi x)\,dx = A_0'$$

因此，
使得

$$A_0' = \int_{x=0}^{x=1} f(x)\,dx$$

- 情形2. $n = m$, 以及 $n,m > 0$

$$\int_{x=0}^{x=1} f(x)\cos(n\pi x)\,dx = \frac{1}{2}A_n$$

以便

$$A_n = 2\int_{x=0}^{x=1} f(x)\cos(n\pi x)\,dx$$

这里，我们使用 $m = n$ 来用 n 表示最终结果（同样，这是惯用的符号表示）。现在，如果我们查看 A_0' 和 A_n 的结果，就会注意到一个有点令人烦恼的问题。A_n 的通式实际上也适用于 A_0'，只是结果会大一倍。因此，有两种同样合理的处理方式。

1. 我们分别记住 A_0' 和 A_n （$n > 0$）的公式，然后继续往下。2. 我们只记住A_n的公式，并认为它也适用于$n = 0$；但这样一来，我们还必须将A_0' 重新定义如下

$$A_0' = \frac{A_0}{2}$$

事实上，两种方法都可以，而且都要求你记住两件不同的事情。按照惯例，大多数教材都采用A_0这种写法。不过在实际应用中，这并不是最佳方法。相反，我更倾向于将傅里叶余弦级数的第一项视为函数在该区间上的平均值 。对于单位区间，区间长度为$L = 1$。因此，我们得到相应的平均值

$$A_0' = \frac{1}{1}\int_{x=0}^{x=1} f(x)\,dx$$

Later, when we allow the intervals to be of arbitrary length, $I = [0, L]$, the first term in the Fourier cosine series will *still be* the average value, computed for this more general case by

$$A_0' = \frac{1}{L}\int_{x=0}^{x=L} f(x)\,dx$$

In summary, we have

$$f(x) = \frac{A_0}{2} + \sum_{n=1}^{n\to\infty} A_n\cos(n\pi x) \tag{4.17}$$

$$A_n = 2\int_{x=0}^{x=1} f(x)\cos(n\pi x)\,dx,\ \ n = 0, 1, 2, \ldots \tag{4.18}$$

或者，更确切地说

$$f(x) = A_0' + \sum_{n=1}^{n\to\infty} A_n \cos(n\pi x) \tag{4.19}$$

$$A_0' = \int_{x=0}^{x=1} f(x)\,dx \tag{4.20}$$

$$A_n = 2\int_{x=0}^{x=1} f(x)\cos(n\pi x)\,dx, \quad n = 1,2,\ldots \tag{4.21}$$

where we *always* compute the first term in any cosine series as the average value over the appropriate interval.

These results are essentially identical to those for the sine series, except that we always will have to deal with the $n = 0$ term of the cosine series (which is identically zero for the sine series) whenever we use the cosine series. This all makes cosine series just a little less fun, but ultimately it will be worth it. We will discuss later reasons that we might prefer one series over another.

Example 4.3 (Fourier Sine and Cosine Series Compared). The function $f(x) = x$ is a frequent example for sine and cosine series, mostly because the expressions for A_n and B_n are integrable. To start, let's compute the sine series. This is given by

$$B_n = 2\int_{x=0}^{x=1} x\sin(n\pi x)\,dx$$

Integrating by parts gives

$$B_n = -2\frac{\cos(n\pi)}{n\pi} = -\frac{2(-1)^n}{n\pi}$$

So, the result is

$$f(x) = \sum_{n=1}^{n\to\infty} B_n \sin(n\pi x)$$
$$x = \sum_{n=1}^{n\to\infty} -\frac{2(-1)^n}{n\pi}\sin(n\pi x)$$

Now, for the cosine series we have

$$A_n = 2\int_{x=0}^{x=1} x\cos(n\pi x)\,dx$$

分部积分得

$$A_n = \frac{2(-1+(-1)^n)}{n^2\pi^2}$$

注意，对于$n = 0$，这会得到 0/0，这是未定义的。为了避免处理这个问题，我们回到级数中第一项的原始定义

$$A_0' = \int_{x=0}^{x=1} x\,dx = \frac{1}{2}$$

So, in conclusion, we have

$$f(x) = A_0' + \sum_{n=1}^{n\to\infty} A_n \cos(n\pi x)$$
$$x = \frac{1}{2} + \sum_{n=1}^{n\to\infty} \frac{2(-1+(-1)^n)}{n^2\pi^2} \cos(n\pi x)$$

Thus, we have two *different* series for the same function. Is there any practical difference between them? We can check by plotting these functions up for $n = 10$ and comparing.

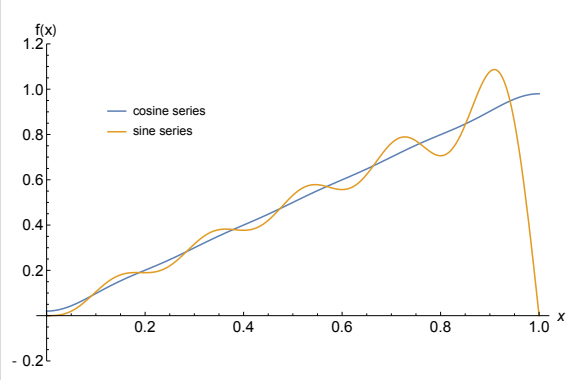


Fig. 4.10: Cosine and sine series for the function $f(x) = x$, each with 10 terms.

我们有两个表示该函数的不同级数，从观察中可以清楚地看出，这两个级数的行为截然不同。首先，使用余弦基函数的级数似乎比使用正弦级数的级数收敛得好得多。其次，正弦级数在$x = 1$处为零，这并不是我们想要的结果。随着项数增加，我们可以看到，这种行为始终没有真正消失。下面是这两个级数的图像，其中每个级数的项数总计为 100

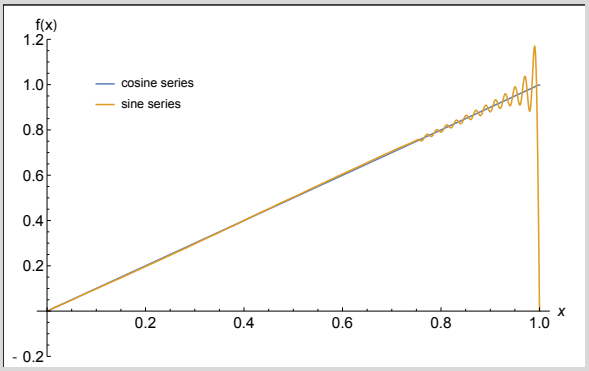


Fig. 4.11: Cosine and sine series for the function $f(x) = x$, each with 100 terms.

如你所见，使用正弦级数时，在$x = 1$附近会出现一些略显奇怪的现象。对此，至少部分原因是：正弦级数被迫在 π 的所有倍数处为零；因此，当$x = 1$时，$\sin(n\pi x)$中的每一项都恒等于零。正弦级数所能做到的最好结果，就是试图构造一个不连续点，使其从左侧接近$x = 1$时，从 1 骤然下降到零。事实上，它确实做到了这一点。随后还会讨论由此产生的其他一些问题（例如，可以观察到使用正弦级数时，在$x = 1$附近出现一些振荡）。

4.7 Comparison of Basis Functions

Both the Fourier sine and cosine series have a similar features. For each, we can think of any function f on a finite interval $D = [a, b]$ as being decomposed into its constituent amplitudes and frequencies, as represented by the weighted sine or cosine functions. These functions form a *basis*, as mentioned above, although it is a basis with an infinite number of basis "vectors" (in this case, the basis vectors are the sine and cosine functions).

了解正弦级数与余弦级数之间的一些相似之处很有帮助。表 4.1 列出了以n为变量的基函数。两类基函数之间有着相当明显的相似之处，$n = 0$的情况除外。在这种情况下，两者并不相似。对于正弦函数，我们有 $\sin(0) = 0$；而对于余弦函数，$\cos(0) = 1$。因此，严格来说，当$n = 0$时，正弦函数确实有一个基函数；然而，由于该函数只是零常函数，它对相应的级数没有任何贡献。对于余弦级数，我们有另一个常函数 $\cos(0) = 1$。这个函数确实会对级数有所贡献，因为它使级数可以平移一个常数。这实际上能让某些问题更容易处理；例如，$f(x) = 1$ 的余弦级数只有$n = 0$ 项（其他项均为零）。而$f(x) = 1$ 的正弦级数展开则复杂得多。方程 (4.22)–(4.23) 明确体现了这一区别。

sine series:
$$1 = \sum_{n=1}^{\infty} B_n \sin(n\pi x) = \sum_{n=1}^{\infty} \frac{1-(-1)^n)}{n\pi} \sin(n\pi x) \tag{4.22}$$

cosine series:
$$1 = A_0' + \sum_{n=1}^{\infty} A_n \cos(n\pi x) = 1,$$
$$(\text{i.e., } A_0' = 1, A_n = 0 \text{ for } n > 0) \tag{4.23}$$

表 4.1：正弦和余弦傅里叶级数的基函数。

n	$\sin(n\pi x)$	$\cos(n\pi x)$
0	0	1
1	$\sin(\pi x)$	$\cos(\pi x)$
2	$\sin(2\pi x)$	$\cos(2\pi x)$
3	$\sin(3\pi x)$	$\cos(3\pi x)$
4	$\sin(4\pi x)$	$\cos(4\pi x)$
⋮	⋮	⋮
n	$\sin(n\pi x)$	$\cos(n\pi x)$

4.8 Fourier Series Convergence: The Dirichlet Conditions

毫不夸张地说，对傅里叶级数收敛所需条件的理解，是推动现代数学分析诸多领域发展的动因之一。要全面理解傅里叶级数的收敛性，需要在数学中引入新的工具和定义，而这些工具和定义对于解析函数并非必需。例如，傅里叶级数的存在使得人们开始质疑函数究竟应如何界定。举例来说，Heaviside 函数不是解析函数（因此没有泰勒级数表达式），但可以通过傅里叶级数构造出来。这促使人们重新审视哪些映射可以合理地称为函数。无穷级数的收敛意味着什么，也引发了类似的概念危机。对于解析函数，泰勒级数收敛到相应函数这一概念，人们已经有了较为充分的理解。然而，傅里叶级数却带来了一些概念上的疑惑。再次以 Heaviside 函数为例，从傅里叶正弦级数展开的例子可以清楚地看到，可以用傅里叶级数构造出对 Heaviside 函数的近似 。然而，傅里叶级数展开的每一项，都是一个带权的C^∞光滑正弦函数。于是便产生了如下概念上的矛盾：光滑函数的和怎么会生成一个存在不连续点（且在该不连续点处没有导数）的函数？这类认识促使人们更深入地理解级数收敛的含义。人们也由此认识到，处理无限 和时，结果可能会出乎直觉。

The conditions for which Fourier series can be understood to converge have been extended well beyond even the cases that originally caused consternation with mathematicians; interested readers can find very

general treatments in the texts by Zygmund (1955), Lighthill (1970), and Hardy and Rogosinski (1999). For our purposes, the most general possible extensions are not necessary. One of the very first researchers on the topic of convergence of Fourier series was a German mathematician named Peter Gustav Lejeune-Dirichlet (for whom the "Dirichlet" boundary condition is named), who in 1829 published a set of criterion for the convergence of Fourier series that covered a wide class of practical functions. Dirichlet not only determined the conditions such that certain Fourier series converge, but he also developed a more modern concept of the function. Notably, the Dirichlet conditions can be described concisely, and have a good physical interpretation. Dirichlet's conditions can be stated as follows.

1. 函数$f(x)$定义在值域$x \in [-L,L]$,上，并且在该区间内有界。注意，在半区间$x \in [0,L]$上定义的函数可以通过周期延拓扩展到零点两侧的对称区间。

2. The function $f(x)$ has, at most, a finite number of points $x_{0,i}, i = 1 \ldots N$ of discontinuity (i.e., a countable number of points where the left and right limits at $x_{0,i}$ approach $f(x_{-,i})$ and $f(x_{+,i})$ respectively). At such points, $f(x_{0,i})$ is given the value
$$f(x_{0,i}) = \frac{f(x_{-,i}) + f(x_{+,i})}{2} \tag{4.24}$$
that is, just the average value at the discontinuity.
3. The function $f(x)$ has a countable number of maxima or minima on the interval $x \in [-L, l]$.

The first condition is sometimes stated by requirement that the function be *absolutely integrable*, i.e.,

$$\int_{-L}^{L} |f(x)|\, dx < \infty$$

and this definition will prove to be useful when the topic of Fourier transforms arises in Chapter 8. Requirement 3 above is sometimes extended to functions that have an infinite number of maxima and/or minima on the interval and have a property known as *bounded variation*. However, this extension was not part of Dirichlet's conditions, and this extension was due to the French researcher Camille Jordan (Jordan, 1881). For the material following, we will assume that we investigate at least meet the Dirichlet conditions (unless otherwise specified).

4.9 The Spectrum

The spectrum of a Fourier series is simply the collection of amplitude coefficients A_n or B_n plotted versus the index, n. Examination of the spectrum of a Fourier series has many practical applications (such as filtering in the Fourier domain), and it also helps us to better understand the physical meaning of the Fourier series.

假设我们进一步考察该函数的正弦级数和余弦级数。

$$f(x) = x, \quad x \in [0, 1] \tag{4.25}$$

我们在上文已经考察过这一点。回顾一下，对于该函数的正弦和余弦级数，我们得到了以下关于n的振幅列表，其中$n \in \mathbb{N}$（即$n = 1, 2, 3, \ldots$）。

余弦

$$\begin{cases} A_0' = \dfrac{1}{2} \\ A_n = \dfrac{2(-1+(-1)^n)}{n^2\pi^2} \quad n = 1,2,3,\ldots \end{cases}$$

sine

$$B_n = \frac{-2(-1)^n}{n\pi} \quad n = 1,2,3,\ldots$$

There is an interesting fact about these two series. Although the functions that they represent are infinite dimensional (in the *uncountable* infinity sense of the word infinite), the series representations are *countably* infinite. In other words, if you know all of the amplitudes, you can reconstruct the function. In a sense, this represents a form of compression of the information embedded in a function. In fact, Fourier series are sometimes used for exactly that purpose (e.g., in the compression of images in the JPEG format!)

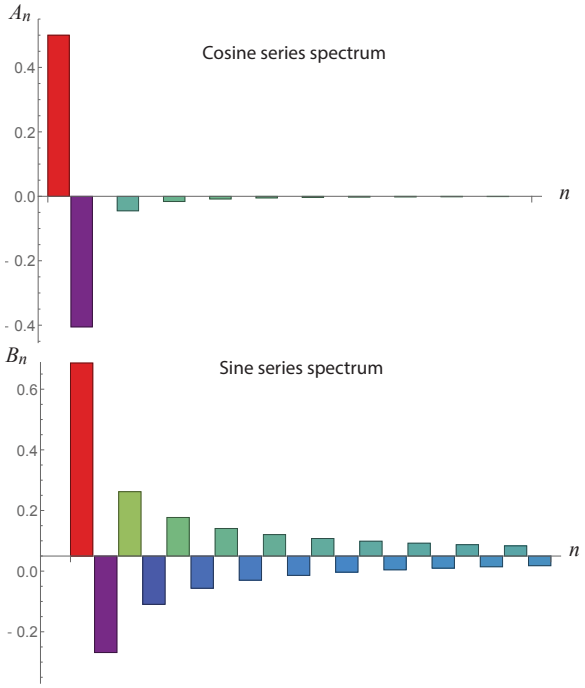


Fig. 4.12 The spectra for the cosine series (top) and sine series (bottom). The first 20 amplitudes are shown.

The amplitudes of the sine and cosine functions are useful for a number of reasons. The main one is that they can give you some sense for how quickly a function converges. The graph of A_n or B_n versus n represents the graph of the amplitudes, and it is called the *spectrum* of a sine or cosine series. As an example, let's look at the two spectrum plots for the sine and cosine series for $f(x) = x$ (Fig. 4.12).

观察这两个频谱，可以发现几个重要的差异。最主要的差异是，这个函数的余弦级数由前两个或三个（非零）振幅主导；相比之下，其余振幅都很小，并且衰减（变小）得非常快。对于正弦级数，虽然前两个振幅（按模计）确实最大，但它们并没有完全主导其他振幅，

而且其衰减似乎也没有那么快。事实上，对于正弦级数而言，前 20 个振幅中是否有任何一个可以忽略不计、相对于前两个来说都不明确。这为我们判断两个级数的收敛情况提供了一些指引。观察这些图，我们或许会认为，即使余弦级数只包含振幅A_0, A_1 和 A_3 (A_2，其余所有偶数序号的振幅均为零)，也能得到良好的结果。而正弦级数则不能如此。为验证这一点，图 4.13 展示了由前三个（非零）振幅构成的余弦级数和正弦级数的图像。观察图 4.13 的结果可知，我们关于振幅相对大小及其衰减速率的判断大体上是正确的。

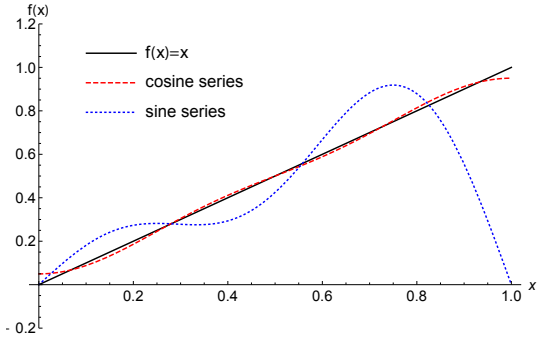


Fig. 4.13 Plots of the cosine and sine series, using only the first three nonzero amplitudes.

频谱也有重要的实际应用。每当处理包含周期分量的时间序列问题时（例如，数周内的潮高、数日内的大气压、一个月内每小时采集的温度），频谱都能告诉你哪些频率（频率 $=n/2$）很重要。同样，也可以利用频谱分析来帮助分析噪声数据，甚至构建滤波器以去除噪声。下一例将对此进行探讨。

例 *4.4*（直流电压的噪声滤波）。假设你在实验室工作，有一台设备通过直流（DC）信号来控制开关。当设备检测到较大的电压尖峰（例如，$20 < V < 50$ 伏特）时，就会启动；检测到较小的电压尖峰（$2 < V < 10$）时，就会关闭。小于约 $2V$ 的电压变化会被忽略。要让这台设备正常工作，你需要一个稳定、干净的直流电压源，能够输出各种幅度的电压尖峰。假设你有一个直流电源，并且大约每 1 秒发出一次低压和高压尖峰。你测量其电压输出，并预期每隔约 1 秒会出现一组由 $25V$ 和 $5V$ 电压尖峰组成的序列。你记录这些信息（示波器是一种可以完成这项工作的设备），却发现直流信号实际上非常不干净，受到各种噪声的干扰。假设它看起来像图 4.14 中的曲线。这相当糟糕，因为信号中噪声太多，以至于有时一些较小的脉冲会产生超过 $10V$的电压（因此，它们不会关闭设备！）

你向当地那位无所不能的能工巧匠展示了你的结果，对方表示，你只需要一个直流滤波器，就能去除来自整流器的高频干扰（整流器是一种将出口的交流电流转换为直流电流的设备）。只要你能提出需要滤除哪些频率，他们就可以为你制作一个。

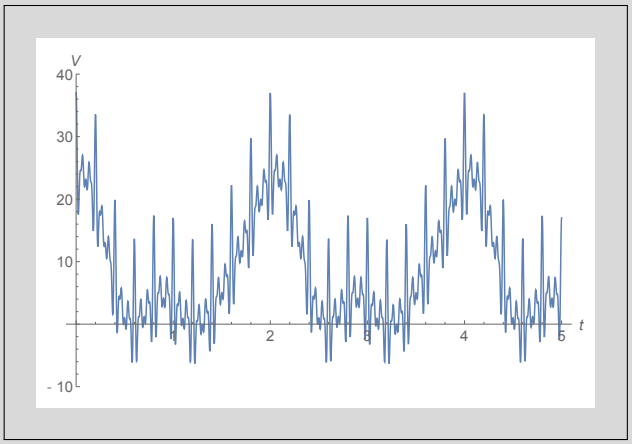
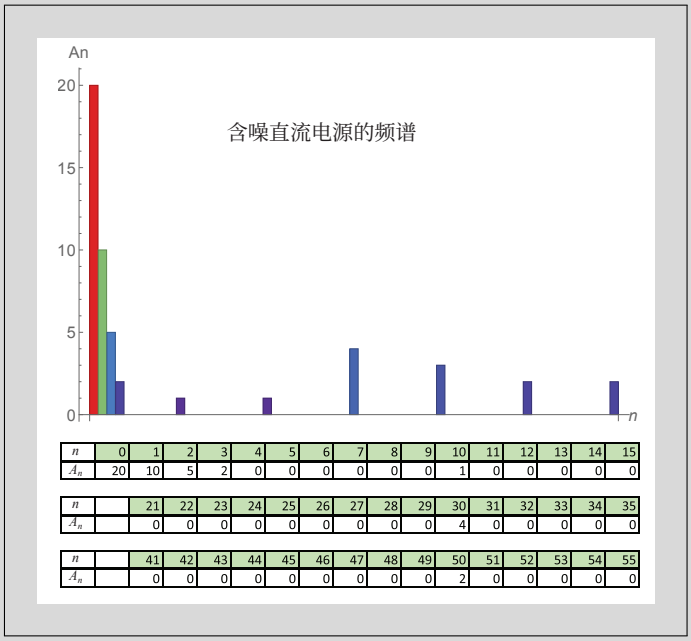


图 4.14：含噪电压-时间信号。



n	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
A_n	20	10	5	2	0	0	0	0	0	0	1	0	0	0	0	0

n	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35
A_n	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0

n	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55
A_n	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0

图 4.15：含噪电流数据的离散余弦变换频谱。

了解了傅里叶级数的相关知识后，你认为真正需要的是直流电流的频谱。直接根据原始数据计算频谱的方法有很多；最流行的一种（这里不会详细介绍！）称为**快速傅里叶变换**，简称 FFT。实际上，你将使用一种快速算法来计算所谓的离散余弦变换（DCT），这样得到的就只有余弦幅频信息。作为 Mathematica 高手，你计算出了频谱，结果如图 4.15 所示。看到这个频谱后，你立刻发现了问题！你预期信号的周期约为 1 秒（即 $n \approx 2$，但你看到有多个分量的n 远大于这个值。解决办法似乎很明确：你只需继续开发

一种可以“滤除”所有高频的滤波器。观察图形后，你认为可以安全地滤除所有满足$n > 10$的值（这意味着你决定保留频谱图中的第一个小波峰）。为了检查这是否足够，你甚至根据这种滤波计算了新的信号。基本上，你可以通过取A_n的前四个非零值来重构正确的余弦信号。你得到的结果是

$$V(t) = 9 + 10\cos(\pi x) + 5\cos(2\pi x) + 2\sin(3\pi x) + \cos(10\pi x)$$

绘制该图可以让你评估滤波器的效果。图形如图 4.16 所示。观察这些结果后，你认为所建议的截止滤波器可行。你会得到幅值正确的“开”和“关”峰值，而且电压波动足够小，可以忽略不计。

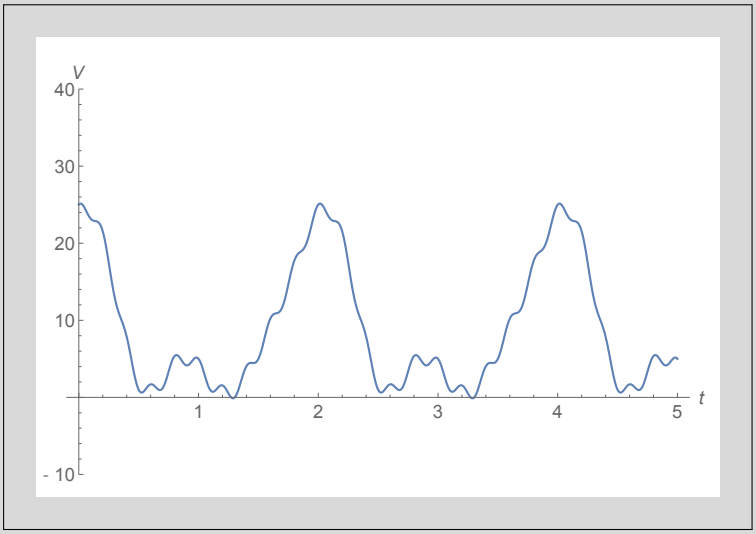


图 4.16：滤波效果良好的电压信号！

4.10 区间变换

在上面的实例中，我们只考察了区间$x \in [0,1]$上的傅里叶级数。当然，我们可能希望在更一般的区间上定义傅里叶级数，例如$x \in [0,L]$。这种变换其实并不难。仔细想想，变换到新区间，基本上就是沿$x \in [0,1]$ 轴拉伸或压缩相应的函数，使其现在出现在$x \in [0,L]$上；函数在竖直方向上的变化保持不变（只是映射到这些拉伸或压缩后的坐标上）。图 4.17 大致展示了这一过程。

Making this change in the domain is not as difficult as it might seem. To start, think about the following variable

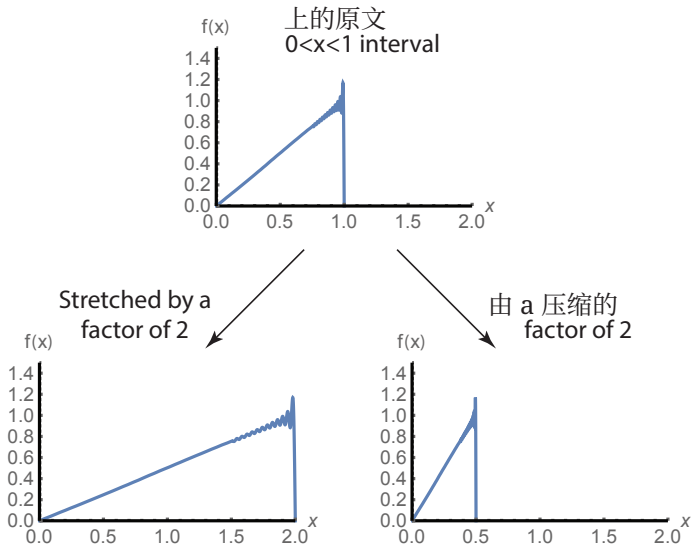


Fig. 4.17 Rescaling the domain of the Fourier series.

$$x = \frac{z}{L}$$

where $z \in [0, L]$. While z on the right-hand side goes from 0 to L, the variable x on the left-hand side only goes from 0 to 1. We have essentially *mapped* the $[0, L]$ onto the $[0, 1]$ interval with this transformation. Why is this useful? Well, let's consider the Fourier sine series for the function $f(x) = x$ that we derived above for $x \in [0, 1]$. The result was

$$f(x) = \sum_{n=1}^{n\to\infty} -\frac{2(-1)^n}{n\pi} \sin(n\pi x), \quad x \in [0, 1]$$

Now, consider the series

$$f\left(\frac{z}{L}\right) = \sum_{n=1}^{n\to\infty} -\frac{2(-1)^n}{n\pi} \sin\left(n\pi\frac{z}{2}\right), \quad x \in [0, 2]$$

This is basically the same series, except the interval is now twice as long. In fact, recognizing that the variable z is just a symbol for the independent variable, we could re-label the variable in this new function back to x if we like. Also, we technically do not need to put parameters like L inside our notation for the function, just the list of independent variables. To make this clear, we can write

$$f\left(\frac{x}{L}\right) = g(x)$$

So, our longer-interval series can be equivalently written (in terms of the independent variable x)

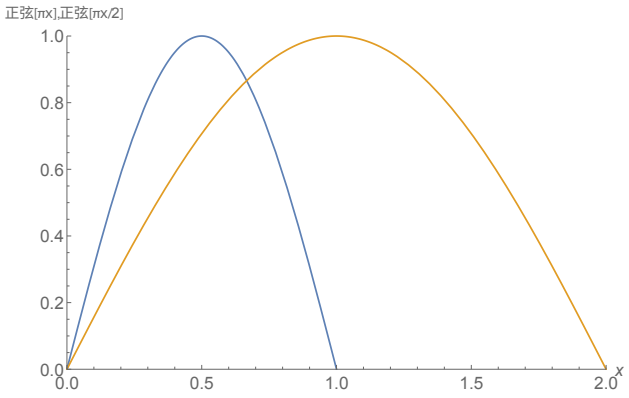


图 4.18 两个正弦函数，其中 $n = 1$。其中一个的距离是另一个的两倍，但二者都对应正弦级数的第一个频率。

$$g(x) = \sum_{n=1}^{n\to\infty} -\frac{2(-1)^n}{n\pi} \sin\left(n\pi\frac{x}{2}\right), \quad x \in [0,2]$$

Note, this is no longer the Fourier series for $f(x) = x$; it is the Fourier series for we have defined the Fourier series for a *new* function

$$g(x) = \frac{1}{2}x$$

如果考虑这里发生的情况，$g(1)$ 的值与$f(1/2)$ 的值相同；类似地，$g(2)$ 的值与$f(1)$ 的值相同。因此，最终结果是，原函数f 现在被拉伸到一个长度是原来两倍的区间上。这实际上就是图 4.17（上图和左下图）所绘制的情况。因此，改变区间的过程其实就是这么回事。

如果考虑这里发生的情况，其实并没有那么复杂。简而言之，我们用一组新的基函数$E_{\sin} = \{0, \sin(1\pi x/L), \sin(2\pi x?) \cdots, x \in [0,L]\}$，替换了这组正交基函数$E_{\sin} = \{0, \sin(1\pi x), \sin(2\pi x) \cdots, x \in [0,1]\}$。

Now, we have on little detail to clean up. Suppose you are given a series on an interval that is not $x \in [0,1]$ to begin with. You could (a) convert the function to an interval that covered $x \in [0,1]$ by an appropriate transformation of variables, and and carry on as we did above, or (b) we could just re-derive the expressions for B_n for the larger interval. Generally, the second of these two options is going to be much more convenient. To do this, we just re-create the process that we did with the unit interval. We end up with the following integrals we need to evaluate

$$\int_{x=0}^{x=L} \sin(n\pi x/L)\sin(m\pi x/L) = \begin{cases} 0 & n \neq m \\ \frac{L}{2} & n = m \end{cases} \tag{4.26}$$

最终，我们发现，对于 $x \in [0,L]$

$$B_n = \frac{2}{L}\int_{x=0}^{x=L} f(x)\sin\left(n\pi\frac{x}{L}\right)dx$$

Note that this reduces to the same expression for the unit interval when $L = 1$!

4.11 Fourier Series on Symmetric Intervals around Zero

尽管我们一直在诸如 $I = [0,L]$,这样的正区间上研究傅里叶正弦级数和余弦级数，但实际上，傅里叶正弦级数和余弦级数的自然定义域是关于零对称的区间，即 $I = [-L,L]$。为了理解这一点，我们主要只需要掌握两个概念。

1. The concept of *odd* versus *even* functions.
2. The concept that the functions $sin(n\pi x/L)$ and $\cos(n\pi x/L)$ each repeat with a period of $2L$ (N.B., for $n = 1$, the functions repeat exactly once in an interval of $2L$; for $n > 1$ the functions repeat multiple times in an interval of $2L$).

We start the discussion with definitions of odd and even functions.

4.11.1 Even and Odd Functions

So far, we have discussed only Fourier series on the interval $[0,L]$, for some positive value for L. In general, there is absolutely nothing preventing a Fourier series to be defined on *any* finite interval $I = \{x : a < x < b\}$ where a and b are any real numbers. However, including the origin in the domain has the potential to create some technical problems for finding Fourier sine and cosine series. Before starting the discussion on Fourier series on more general intervals, we will discuss the concept of *odd* and *even* functions, and their relationship to Fourier series.

Definition 4.1. Suppose a function f is defined on some *symmetric interval* around origin ($x = 0$), i.e., $I = [-L,L]$, where a and is some positive real number. Then the function may be *even*, *odd*, or *neither even nor odd* as follows.

- The function f is said to be *odd* on I if $f(x) = -f(-x)$ for all $x \in I$.
- The function f is said to be *even* on I if $f(x) = f(-x)$ for all $x \in I$.
- If neither of these two conditions is true, then f is neither even nor odd on the interval I.

The concepts of even and odd functions are more than just a obscure mathematical property. It turns out that, for every single function defined on a symmetric interval around zero can be decomposed into an even part and and odd part whose sum gives the original function. The proof of this is intuitive and clever, so it will be explained in a few lines.

定理 4.2。在以零为中心的对称区间 I 上，任何既非偶函数也非奇函数的函数 f，都可以分解为该区间上两个函数之和，其中一个是偶函数，另一个是奇函数。

Proof. Define two new functions as follows

$$f_{even}(x) = (f(x) + f(-x))/2 \qquad f_{odd}(x) = (f(x) - f(-x))/2$$

容易验证，$f_{even}(-x) = f_{even}(x)$，而且$f_{odd}(-x) = -f_{odd}(x)$，因此这两个函数分别是偶函数和奇函数。也容易验证，$f_{even}(x) + f_{odd}(x) = f(x)$。■

事实证明，这种将函数分解为偶函数部分和奇函数部分的方式也是唯一的；也就是说，不存在其他这样的分解，只有我们上面构造的这一种。图 4.19（(b) 和 (d)）给出了正弦函数和余弦函数在区间 $-1 < x < 1$ 上的实例。这种图示呈现方式使我们注意到，上述泛函定义得出了偶函数和奇函数的几何性质，这些性质可以从它们的图中观察到。特别要注意以下内容

- 余弦函数是偶函数（$\cos(-\pi x) = \cos(\pi x)$）。从图形上看，我们可以通过将 $[0,1]$ 上的余弦函数关于纵轴反射，得到 $[-1,0]$ 上的余弦函数。
- The sine function is an *odd* function ($\sin(-\pi x) = -\sin(\pi x)$). Graphically, we can generate the sine function on $[-1,0]$ by reflecting the sine function on $[0,1]$ about the vertical axis, and then a second time about the horizontal axis.

There are a few additional features about even and odd functions that should be mentioned. The first is an assessment of what happens when one multiplies even and odd functions. As an example of even and odd functions to keep in mind, think of the even function $f_{even}(x) = x^2$ and the odd function $f_{odd}(x) = x^3$ on $x = [-1,1]$. Now we state the following theorems (without proof)

定理 4.3（偶函数与奇函数的乘积）。假设我们有一个偶函数 $f_{even}(x)$ 和一个奇函数 $f_{odd}(x)$ 定义在区间 $[-L,L]$ 上。那么，这些函数的乘积具有以下性质

1. $f_{even} \times f_{even}$ *is an even function.*
2. $f_{even} \times f_{odd}$ *is an odd function.*
3. $f_{odd} \times f_{odd}$ *is an even function.*

For easy visualization, these results are summarized graphically in Fig. 4.20.

Theorem 4.4 (The integral of an even and odd functions). *Suppose we have and even function $g_{even}(x)$ and an odd function $g_{odd}(x)$ defined on the interval $[-L,L]$. Then the integral of g_{even} over the interval $[-L,L]$ is non-zero. The integral of g_{odd} over the interval $[-L,L]$ is zero.*

The reason that these two theorems are important, is because they explain what we can and cannot do with Fourier sine and cosine series. In particular, recalling that the sine series is odd, and the cosine series is even, then we have the following results. Suppose $f(x)$ function on the interval $I = [-L,L]$. Then, the following things must be true

1. If $f(x)$ is even, then it can be expanded only as a cosine series. A sine series will not work because each integral would involve an even times an odd function, which is odd; the resulting integrals defining B_n would be zero.
2. If $f(x)$ is odd, then it can be expanded only as a sine series. A cosine series will not work because each integral would involve an odd times an even function, which is odd; the resulting integrals defining A_n would be zero.

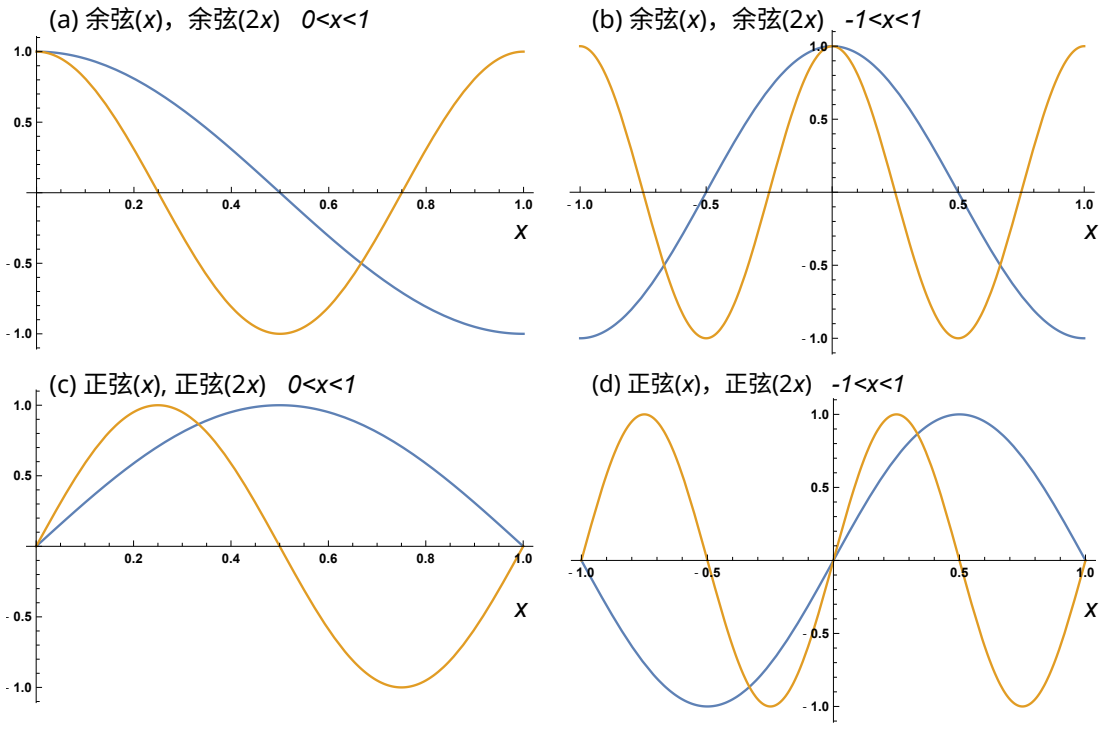


Fig. 4.19: Sine and cosine functions on the intervals $0 < x < 1$ and $-1 < x < 1$. From the definitions in the text (and by examining parts (b) and (d) of this figure), it should be clear that cosine is an even function and sine is and odd function on the interval $-1 < x < 1$.

3. 如果$f(x)$既不是偶函数也不是奇函数，那么可以将其视为两个函数之和：$f(x) = f_{even}(x) + f_{odd}(x)$。傅里叶级数必须同时包含正弦分量和余弦分量，才能表示整个函数。

了解了奇函数与偶函数之后，我们可以考虑求一般函数在关于零对称的区间上的傅里叶变换。

4.11.2 区间$I = [-L,L]$上的傅里叶级数

Now consider the symmetric interval $I = [-1,1]$. It turns out that the relevant basis functions for this are $\sin(n\pi x)$ and $\cos(n\pi x)$, just as they were for the unit interval $I = [0,1]$. To see this more clearly, examine the behavior of both the sine and cosine functions on the interval $I = [-1,1]$ as shown in Fig. 4.19. It is easy to see that the interval $I = [-1,1]$ allows the sine and cosine functions to go through an integer number of periods; in fact, the functions on the interval $[-1,0]$ are just reflections of the functions on the interval $[0,1]$

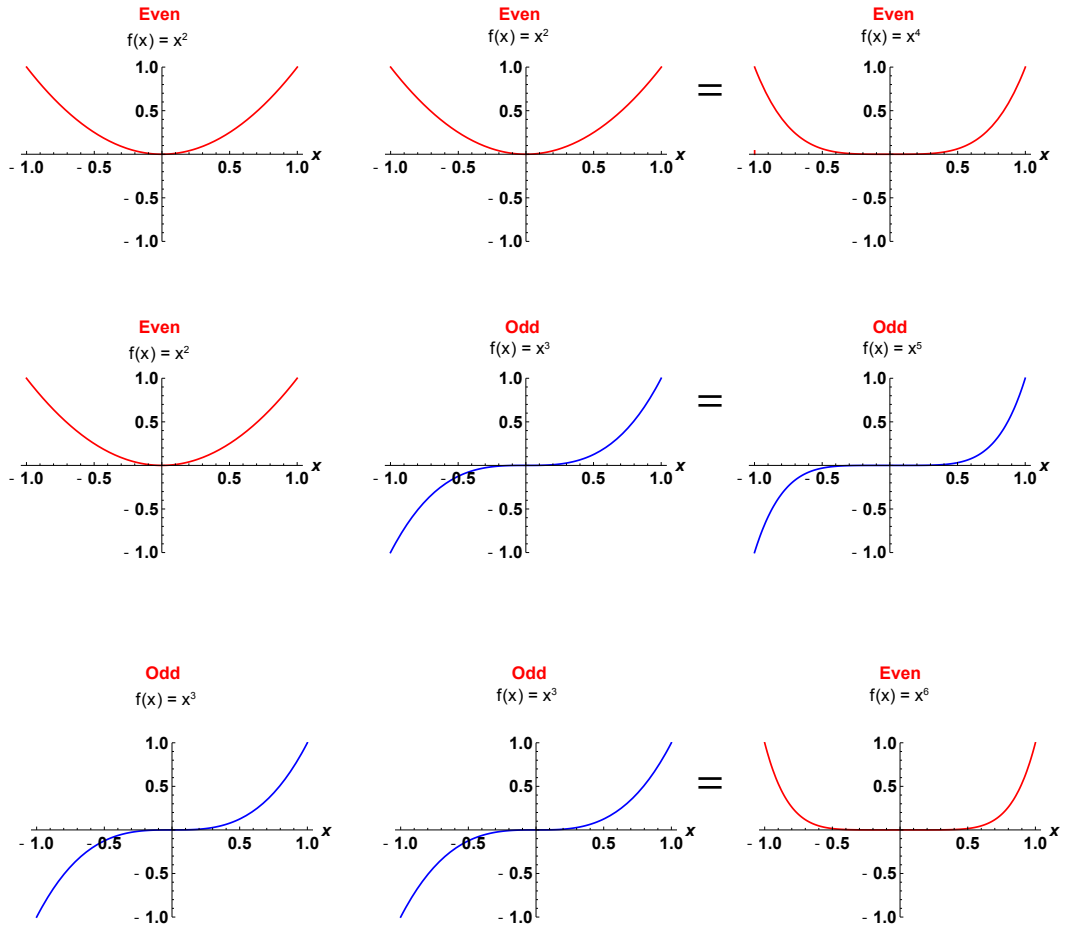


Fig. 4.20: The result of multiplying even and odd functions. In this figure, the multiplication happens from left to right, across rows.

as discussed above. The basis functions on the interval $I = [-1, 1]$ are exactly the same as those on $I = [0, 1]$; specifically, the functions $\sin(n\pi x)$ and $\cos(n\pi x)$.

关于这种区间变换，有一个稍微容易混淆的地方需要注意。对于对称区间 $I = [-L, L]$，我们只需将自变量除以 L，就可以将该区间映射到区间 $I = [-1, 1]$。换句话说，新的基函数变为 $\sin(n\pi x/L)$ 和 $\cos(n\pi x/L)$。这一点值得注意，因为我们仍然是除以 L，而 L 仅对应完整区间宽度的一半（区间宽度为 $2L$）。这与在 $I = [0, L]$ 上进行区间变换的情况不同，在后一种情况下，我们除以的是整个区间的宽度，即 L。

理解 $I=[0,1]$ 上的傅里叶级数与 $I=[-1,1]$ 上的傅里叶级数之间关系的最佳方式，是将它们视为相同的级数，只不过我们忽略了区间的负半部分。事实上，当我们计算函数 $f(x)$ 在区间$I=[0,1]$上的傅里叶级数时，所做的正是这件事。如果将 $f(x)$ 展开为正弦级数，那么我们实际上忽略了该解在区间 $[-1,0]$ 上的奇函数伸展。如果将 $f(x)$ 展开为正弦级数，那么我们实际上忽略了该解在区间 $[-1,0]$ 上的偶函数伸展。通过一个例子，可以使这一点更加清楚。不过，在继续之前，我们先注意一般区间 $I=[-L,L]$ 上的以下积分。

$$\int_{x=-L}^{x=L}\sin(n\pi x/L)\sin(m\pi x/L)=\begin{cases}0 & n\neq m\\ L & n=m\end{cases} \tag{4.27}$$

$$\int_{x=-L}^{x=L}\cos(n\pi x/L)\cos(m\pi x/L)dx=\begin{cases}0 & if\ m\neq n\\ L & if\ m=n\neq 0\\ 2L & if\ m=n=0\end{cases} \tag{4.28}$$

Note that the interval is now $2L$, but the bounds of the integral go from $-L$ to L. This sometimes creates some confusion, so be aware of the details here.

These two integrals are still *orthogonal*, but the case for $n=m$ now evaluates to twice the value we had for the unit interval. This makes sense, since the total domain of integration is twice that of the unit interval. Following this process through to the evaluation of A_n and B_n leads to the (hopefully unsurprising) results

$$\begin{aligned}A_0' &= \frac{1}{2L}\int_{-L}^{L}f(x)dx\\ A_n &= \frac{1}{L}\int_{-L}^{L}f(x)\cos(n\pi x/L)dx,\ \ n=1,2,3,\ldots\\ B_n &= \frac{1}{L}\int_{-L}^{L}f(x)\sin(n\pi x/L)dx\ \ n=1,2,3,\ldots\end{aligned}$$

with

$$f(x)=A_0'+\sum_{n=1}^{\infty}A_n\cos(n\pi x/L)$$

or

$$f(x)=\sum_{n=1}^{\infty}B_n\sin(n\pi x/L)$$

Recall, it is easy to produce the formula for the first term of the cosine series (A_0') if you remember that it is just the *average* of the function.

Example 4.5 (Sine series for $f(x) = x$ on $I = [0,1]$ versus on $I = [-1,1]$). Lets return to the familiar case of the sine series expanded fo the function $f(x) = x$. To begin, lets compute the sine series for this function on the interval $I = [0,1]$. Following the details above, we have

$$B_n = 2\int_0^1 x\cos(n\pi x)dx = \frac{-2(-1)^n}{n\pi}$$

Which gives us the series

$$x = \sum_{n=0}^{n=\infty} \frac{-2(-1)^n}{n\pi}\sin(n\pi x), \quad \text{for } 0 < x < 1$$

这应该很熟悉；我们之前已经计算过这个级数。现在，让我们尝试将之前针对单位区间所做的步骤应用于区间 $I = [-1,1]$。我们从 $I = [-1,1]$ 上的级数展开定义开始。

$$x = \sum_{n=0}^{n=\infty} B_n \sin(n\pi x), \quad \text{for } -1 < x < 1$$

As done previously, we use orthogonality to determine the B_n. Multiplying both sides of the last equation by $sin(m\pi x)$ and integrating gives us

$$\int_{-1}^1 x\sin(m\pi x)\,dx = \sum_{n=0}^{n=\infty} B_n \int_{-1}^1 \sin(n\pi x)\sin(m\pi x)\,dx$$

and, using the orthogonality of the sine functions, only one term in the sum on the right hand side is non-zero; this corresponds to the case where $n = m$

$$B_n = \int_{-1}^1 x\sin(m\pi x)\,dx$$

Finally, computing this integral gives

$$B_n = \frac{-2(-1)^n}{n\pi}$$

which gives us the series

$$x = \sum_{n=0}^{n=\infty} \frac{-2(-1)^n}{n\pi}\sin(n\pi x), \quad \text{for } -1 < x < 1$$

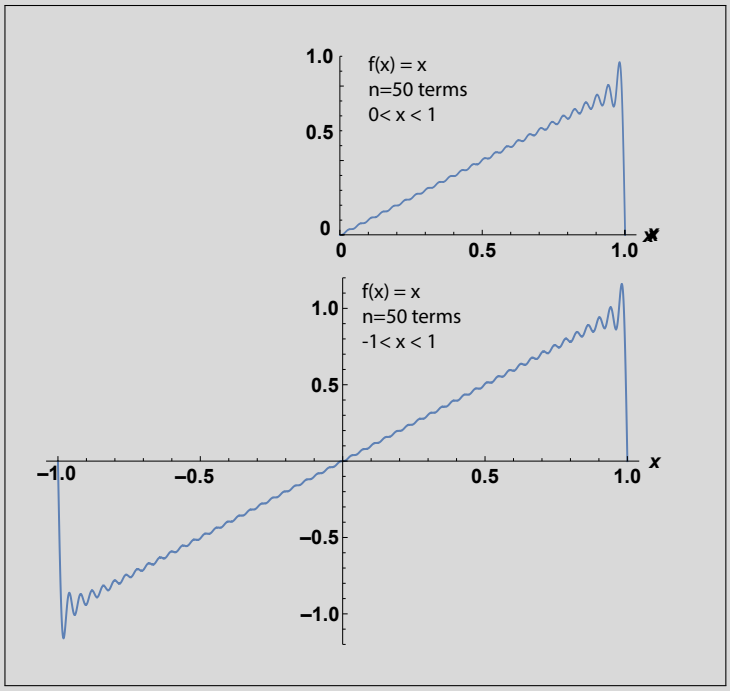


Fig. 4.21: Comparison of the Fourier sine series for the function $f(x) = x$ on the domains $I_1 = [0,1]$ and $I_2 = [-1,1]$. Note that sine is and odd function, thus the function on the symmetric domain $I_2 = [-1,1]$ is also odd.

Comparing this with the result for the unit interval, we find that the two series are identical! Thus, when we expand a function as a sine series on the interval $I = [0,1]$, we really are generating the same result as for the interval $I = [-1,1]$, it is just that we are ignoring the negative portion of the solution (because, by definition of the domain being $I = [0,1]$, it is not in the domain that is of interest to us). A plot can help compare the solution over the two domains; such a plot given in fig. 4.21.

The case where a function is neither even nor odd is an interesting one. To start, we will *attempt* to expand the function using only a sine series and then only a cosine series. We will find that *neither* series is able to expand the function by itself. However, the *sum* of the two series will reproduce the function. In short, the sine series reproduces the *odd* part of the function, and the cosine series reproduces the *even* part of the function. An example is really useful here.

Example 4.6 (Cosine and Sine series for $f(x) = e^{-x}$ on $I = [-1,1]$).

The function $f(x) = e^{-x}$ is neither even nor odd on $I = [-1,1]$; this is easy to see by looking at a plot of the function on this interval.

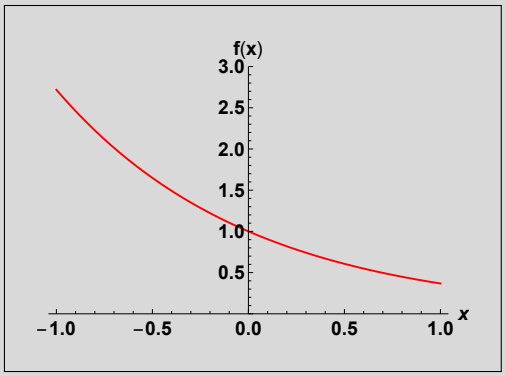


Fig. 4.22: The function e^{-x} is neither even nor odd.

In principle, we could *split* the function $f(x) = e^{-x}$ into its even and odd parts, and then compute the Fourier series for each of these parts. However, the computation of the series coefficients A_n and B_n will actually take care of this for us; the multiplication by cosine or sine filters out the even and odd parts of the function (respectively) during the computation of these integrals. So, let's try computing the two parts as separate operations, starting with the cosine series. For reinforcement of the ideas of orthogonality, this example will be shown with all steps developed (rather than simply relying on the formulas derived earlier).

$$e^{-x} = \sum_{n=0}^{n=\infty} A_n \cos(n\pi x), \quad \text{where for this problem } L = 1.$$

Multiplying both sides by $\cos(m\pi x)$ and integrating gives

$$\int_{-1}^{1} e^{-x}\cos(m\pi x)dx = \sum_{n=0}^{n=\infty} A_n \int_{-1}^{1} \cos(n\pi x)\cos(m\pi x)dx$$

For the integral on the right, there are three cases that can occur: (1) $m \neq n$, in which case the integral is zero, (2) $n = m \neq 0$, in which case the integral is equal to 1, and (3) $n = m = 0$, in which case the integral is equal to 2 (the length of the interval!) Regardless of case, for each value of m chosen, there is only *one* nonzero value of the sum as n goes from zero to infinity. Thus, we can rewrite this equation as (and, remembering that because $n = m$, we are free to set the index to n everywhere)

$$\int_{-1}^{1} e^{-x}\cos(n\pi x)dx = A_m \int_{-1}^{1} \cos(n\pi x)\cos(n\pi x)dx$$

We treat the $n = 0$ case independently. For that case, we find that the integral on the right gives the result of 2, and, solving for A_0', we find

$$A_0' = \frac{1}{2}\int_{-1}^{1} e^{-x}\cos(n\pi x)dx = \sinh(1)$$

For the remaining values of n, we find

$$A_n = \int_{-1}^{1} e^{-x}\cos(n\pi x)dx = \frac{(-1)^n(-1+e^2)}{e(1+n^2\pi^2)}$$

Together, these yield the series

$$f_{even}(x) = \sinh(1) + \sum_{n=1}^{n=\infty} \frac{(-1)^n(-1+e^2)}{e(1+n^2\pi^2)}\cos(n\pi x)$$

Here, the function is labeled "$f_{even}(x)$ because it is not in fact the function $f(x) = e^{-x}$; this can be seen in Fig. 4.23. The associated Fourier cosine series *must be* an even function, and we can verify that it is. Conceptually, this funtion represents the "even" component of the function $f(x) = e^{-x}$.

For the sine series, we will not reproduce the entire sequence of operations leading to the resulting series. Instead, we perform operations analogous to those above, and find the result (presented previously)

$$B_n = \int_{-1}^{1} e^{-x}\sin(n\pi x)dx = \frac{(-1)^n(-1+e^2)n\pi}{e(1+n^2\pi^2)} \tag{4.29}$$

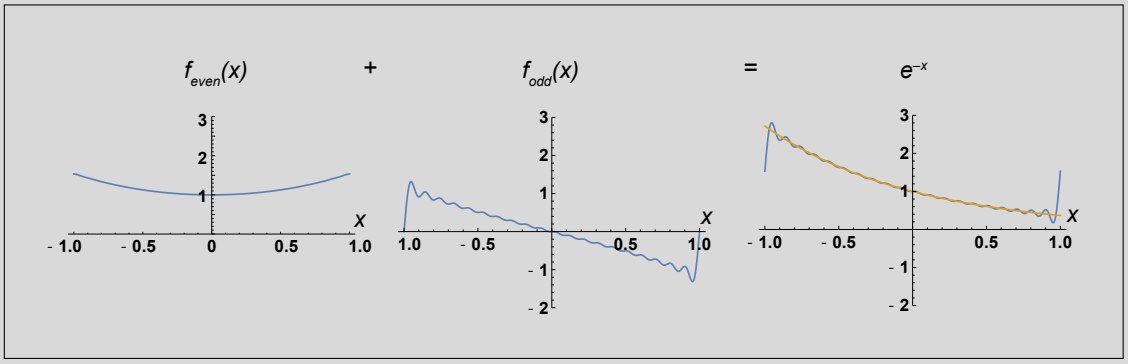


Fig. 4.23: The function e^{-x} is the sum of an even (cosine) Fourier series plus and odd (sine) Fourier series

And, the corresponding series is

$$f_{odd}(x) = \sum_{n=1}^{n=\infty} \frac{(-1)^n(-1+e^2)n\pi}{e(1+n^2\pi^2)}\sin(n\pi x)$$

该函数也绘于图 4.23 中。观察可知，正弦级数实际上是一个奇函数。

最后请注意，这两个级数的和实际上还原了我们要找的函数

$$\begin{aligned} e^{-x} &= f_{odd}(x) + f_{even}(x) \\ &= \sinh(1) + \sum_{n=1}^{n=\infty} \frac{(-1)^n(-1+e^2)}{e(1+n^2\pi^2)} \cos(n\pi x) + \sum_{n=1}^{n=\infty} \frac{(-1)^n(-1+e^2)n\pi}{e(1+n^2\pi^2)} \sin(n\pi x) \end{aligned}$$

Figure 4.23 illustrates the sum of these two series, showing that the sum of the sine and cosine series expansions for the function results in a total series that reproduces the function.

From the example above, we can draw some conclusions.

1. 关于零对称的区间上的函数可以是偶函数、奇函数，或者既非偶函数也非奇函数。对于后一种情况，总是 可以将该函数唯一地分解为偶部分和奇部分，即 $f(x) = f_{even}(x) + f_{odd}(x)$。
2. Even functions are represented exclusively by Fourier cosine series.
3. 奇函数仅由傅里叶正弦级数表示。
4. For functions that are neither even nor odd, we need to sum both the Fourier cosine and sine series so that both the even and odd components of the function are represented. This expression takes the general form

$$\begin{aligned} f(x) &= f(x) + f_{even}(x) + f_{odd}(x) \\ f(x) &= A_0' + \sum_{n=1}^{n\to\infty} A_n \cos(n\pi x) + \sum_{n=1}^{n\to\infty} B_n \sin(n\pi x) \end{aligned}$$

4.11.3 Periodic Continuation

The examples above illustrated that even if we compute a Fourier series for $x \in [0, L]$, the Fourier series is actually defined for $x \in [-L, L]$. The shape of the function in the negative part of the interval will depend on whether a sine or a cosine series is used to expand the function. This is exactly what we saw in Example 4.5.

事实证明，区间 $[-L,L]$ 上的傅里叶级数实际上是一个周期条件函数。也就是说，在区间 $[-L,L]$, 之外，该函数会通过生成在 $[-L,L]$上定义的函数的副本而不断重复。

As an example, consider our cosine and sine series developed for $x \in [0, 1]$ in Example 4.3. Recall, the two functions that we found for the unit interval were

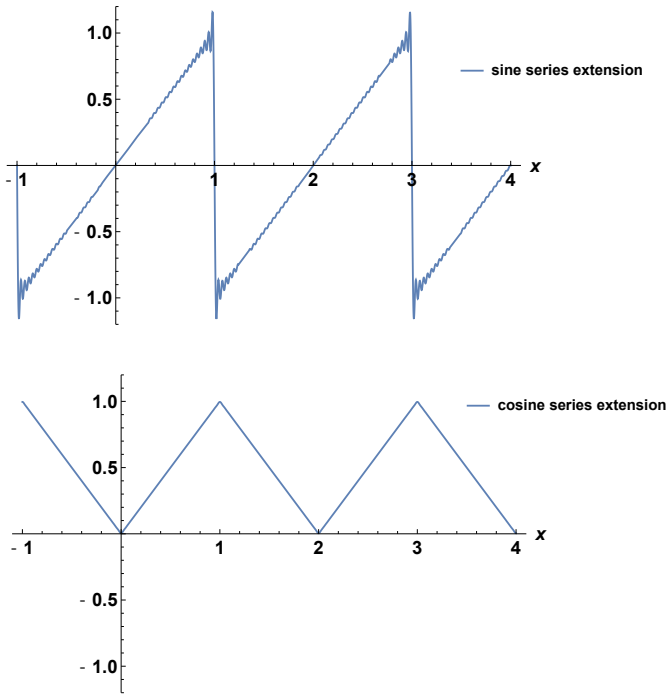


Fig. 4.24 The extensions of the Fourier sine and cosine series for the function $f(x) = x$, x to the interval $x \in [-1, 4]$.

cosine function

$$x = \frac{1}{2} + \sum_{n=1}^{n\to\infty} \frac{2(-1+(-1)^n)}{n^2\pi^2} \cos(n\pi x) \tag{4.30}$$

sine function

$$x = \sum_{n=1}^{n\to\infty} \frac{-2(-1)^n)}{n\pi} \sin(n\pi x) \tag{4.31}$$

现在，当我们在$x \in [-1,4]$ 上绘制它们（图4.24）时，会发现每个函数都会重复$x \in [-1,1]$上所表示的内容。这是合理的——每个基函数（正弦(πx), 正弦$(2\pi x)$...）都对任意x值有定义，而且按照构造，每个三角函数的周期都是 $2/n$。因此，所涉及的最低频率函数是正弦(πx)和余弦(πx)，这两个函数的周期都恰好为2。所有其他函数的重复频率更高，但每个函数在长度为2的区间内重复n 次。因此，由这些函数求和构造出的函数也必然以2为周期重复。比较图4.24中$f(x) = x$ 的正弦级数和余弦级数，可以清楚地看出事实确实如此。这个例子也再次凸显了奇函数（正弦）级数展开与偶函数（余弦）级数展开之间的差异。

4.12 傅里叶级数收敛简介

在微积分中，研究无穷级数时会涉及收敛问题。判断某个级数是否会以某种方式收敛，是应用数学中的核心问题。

毫不夸张地说，对傅里叶级数收敛性质的研究实际上开启了数学分析的新时代。傅里叶级数刚被提出时，傅里叶在没有给出证明的情况下（约从 1807 年开始，当时他首次发表了研究结果）便声称，这种方法适用于任何函数。不过需要指出的是，当时人们甚至还没有明确界定何为恰当的函数，也没有清晰定义积分。了解这一背景后，我们就能更好地理解傅里叶过于热切的宣称。

自傅里叶提出他的理论以来，人们逐渐认识到，该理论比最初设想的要稳健得多。尽管收敛的细节可能涉及非常专业的数学概念，但可以说，自然出现在物理学和工程学中的许多函数，都会以某种有用且直观的意义收敛。

最早对傅里叶级数收敛性进行的严肃研究由 P. Dirichlet 完成。他于 1829年 提出了关于傅里叶级数收敛性的第一个定理。按照现代标准，这个定理的限制远比技术上所需的严格，但它仍是理解傅里叶级数收敛性的一个有用切入点。狄利克雷提出的条件足以保证至少逐点收敛，但其严格程度超过了技术上所必需的程度。换言之，有些函数的级数虽不满足这些条件，却仍能以某种明确的意义收敛。尽管如此，狄利克雷条件涵盖了许多在应用问题中会出现的重要函数。虽然我们稍后才会更详细地讨论傅里叶级数的收敛性质，但这里先列出狄利克雷条件如下。

Theorem 4.5 (Convergence of Fourier Series: Dirichlet conditions.). *The Fourier series for a function f converges p*如果在区间 I 上逐点满足以下条件，则称其为逐点收敛。

1. *The function is* bounded *over the domain I. This is sometimes stated more rigorously by the idea that the function f is* absolutely integrable*, which means*

$$\int |f(x)|\,dx < \infty$$

2. 该函数是分段 C^1 连续的。根据定义，这意味着存在有限个点，在这些点处，右导数与左导数不相等（即导数存在有限个不连续点）。

3. *The function f has only a* finite *number of maxima or minima in the interval. In other words, the function can not oscillate infinitely fast anywhere in the interval.*

As a note, the finite number of maxima or minima prohibits functions such as $f(x) = \frac{1}{\sin(1/x)}$ from meeting the conditions because such functions oscillate infinitely fast as $x \to 0$. If the the three conditions above are met, then the function f can be said to have a Fourier series that converges everywhere in the domain, I. The series converges pointwise to its actual value $f(x)$ except at points of discontinuity. At points of discontinuity, x_d the series converges to the average value on the two sides of the discontinuity, i.e.,

$$f(x_d) = \frac{1}{2}[f(x_d^+) + f(x_d^-)]$$

Here, $f(x_d^+)$ represents the value of the function at x_d approaching from the right, and $f(x_d^-)$ represents the value of the function at x_d approaching from the left.

4.13 *适当基大有不同

本节稍微偏离了主题，但很值得一读。现在我们已经掌握了傅里叶级数，可以讨论这样一个观点：在表示特定函数时，有些基比其他基更合适。本节篇幅虽短，目的并非特别严谨，而是让学生有机会思考 不同类型的级数表示在特定应用中可能各有优劣。

我们已经学习了两种展开。首先，我们讨论了以下函数的幂级数 m

$$f(x) = f(x_0) + (x-x_0)f'(x_0) + \tfrac{1}{2!}(x-x_0)^2 f''(x_0) + \ldots \tag{4.32}$$

现在，尽管我们不总是把幂级数看作多项式展开，但它实际上确实是！假设我们在$x_0 = 0$附近展开。很明显，所得级数看起来像是由单项式 展开而成（变量x 的某个幂就是一个单项式）。

$$f(x) = f(0) + xf'(0) + \tfrac{1}{2!}x^2 f''(0) + \ldots \tag{4.33}$$

在这里，我们可以把函数 1、x、x^2, 等看作展开所用的基函数，而$f(0)$、$f'(0)$、$f''(0)$等则是展开的系数。

Now, lets think about the series expansion of two different functions on $x \in [0,1]$: $f(x) = \sin(\pi x)$ and $g(x) = x^2$. In the material in this chapter, we have learned how to compute the Fourier series for each of these functions. And we already know how to expand functions as power series. Here, we will just list the results of the expansions for comparison. The first seven terms in both expansions for the functions $f(x)$ and $g(x)$ are given below.

$f(x) = \sin(\pi x)$:

$$\text{傅里叶} \quad f(x) = 1\cdot\sin(\pi x) + 0 + 0 + 0 + 0 + 0 + \ldots \tag{4.34}$$

$$\textit{Power} \quad f(x) = 0 + \pi x + 0 + \frac{\pi^3 x^3}{6} + 0\frac{\pi^5 x^5}{120} + 0 + \ldots \tag{4.35}$$

$g(x) = x^2$:

$$\textit{Fourier} \quad g(x) = \frac{\left(2\pi^2-8\right)}{\pi^3}\sin(\pi x) - \frac{1}{\pi}\sin(2\pi x) + \frac{\left(18\pi^2-8\right)}{27\pi^3}\sin(3\pi x) - \frac{1}{2\pi}\sin(4\pi x) + \frac{\left(50\pi^2-8\right)}{125\pi^3}\sin(5\pi x) - \frac{1}{3\pi}\sin(6\pi x) + \ldots \tag{4.36}$$

$$\textit{Power} \quad g(x) = 0 + 0 + x^2 + 0 + 0 + 0 + \ldots \tag{4.37}$$

Plots of these series appear in Fig. 4.25. For the function $f(x)$, it is clear that the Fourier series is very efficient at capturing the function (in this case, it is exact), and this requires only one non-zero term. However, the power series expansion contains significant deviations near $x = 1$ even with six terms. The converse can be

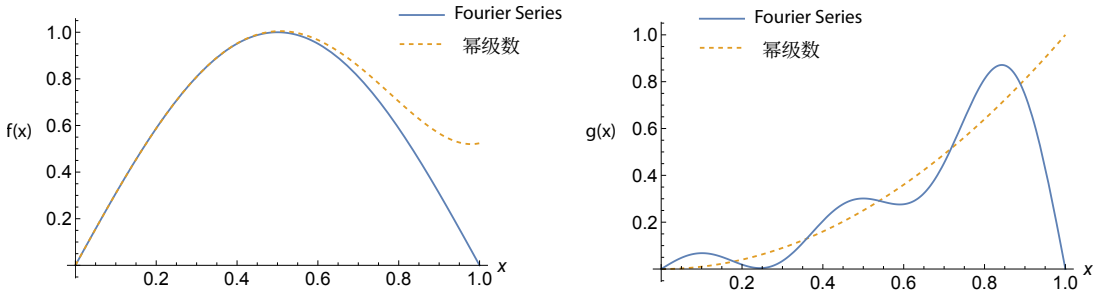


Fig. 4.25: Comparison of the series expansions for $f(x) = \sin(\pi x)$ and $g(x) = x^2$ by the Fourier series and power series expansions.

see for the function $g(x)$. For that function, the power series gives an exact result, whereas the Fourier series gives an approximation that is not particularly good when only the first six terms are used.

So, it becomes clear that different kinds of basis functions can give dramatically different results. When one is attempting to*fit* an expansion to a function, there are choices among expansions. While we have examined only two here, there are many other kinds of basis functions that can be used to approximate a target function. When one would like a good and efficient approximation, exploring the opportunities for choice of basis functions to use can make a difference!

习题

练习题

1. We first encountered the idea of expanding a function in terms of a set of basis functions when we examined the Taylor series for analytic functions (Section 2.6.4). In that material, it was illustrated that a Taylor series truncated at $n = N$ was called a Taylor polynomial of order N. In such expansions, we have a sum of terms proportional to the following basis set of polynomials

$$= \{1, x, x^2, x^3, x^4, \ldots\}$$

where we use the notation $p_0(x) = 1$, $p_1(x) = x$,$p_2(x) = x^2$, etc. Thus, a Taylor series can be represented by

$$f(x) = \sum_{n=0}^{\infty} d_n p_n(x)$$

or, equivalently

$$f(x) = \sum_{n=0}^{\infty} d_n x^n$$

where d_n represents the appropriate coefficient of term n of the Taylor series, and p_n are the basis functions.
The finite Taylor polynomials of order N are given by

$$f_N(x) = d_0 \cdot 1 + d_1 x + d_2 x^2 + d_3 x^4 + \ldots + d_N x^N$$

It is clear from this representation that a Taylor polynomial is just a linear combination of the polynomials in the set $\mathscr{P}$. Although the polynomials themselves are not linear (except, of course, for the cases $N = 0$ and $N = 1$), their sum is in fact a linear combination of the basis functions.
These polynomials can be very useful approximations for functions that otherwise have no easy analytical expression. For example, the function

$$f(x) = \frac{1}{1-x} \tag{4.38}$$

converges on the interval $-1 < x < 1$ (i.e., its radius of convergence, R, is $|x| < 1$ When doing analysis, encountering functions like this can create serious difficulties because they are not linear. However, for x small enough, one can use the first few terms of a Taylor series to very accurately represent the function. This gives us a good way to introduce the idea of generating an approximation to an analytic function using a polynomial (extracted from its Taylor series).
For this problem, do the following to illustrate how a linear combination of basis functions can generate an approximation to a function that is analytic in some domain. Do the following.

a. 求出该函数的一般泰勒级数。注意：该函数的收敛半径为实数集 < 1，且在边界实数集 $= 1$ 处不收敛。换句话说，该级数在 $|x| < 1$ 时收敛。

b. Compute the first five terms of this series explicitly (i.e., work out the Taylor polynomial for it). Using Mathematica (or whatever software is convenient for you) plot the series approximation $f_5(x)$ and the original function $f(x) = 1/(1-x)$ over the interval $x \in [-1,1]$; recall, $f_5(x)$ represents the sum of the series for $n = 0$ to $n = 5$ (i.e., the first six terms of the series).
Compute the fractional error by

$$error(x) = [f_5(x) - f(x)]/f(x)$$

and plot this on the same plot with $f_5(x)$ and $f(x)$.

c. For x small enough, the first two terms in the Taylor series will be accurate, i.e., the Taylor series can be given by

$$f_1(x) = 1 + x$$

Determine the maximum value for x_{max} such that the $f_1(x)$ approximation for $f(x)$ is less than 1% within the domain $0 \le x \le x_{max}$. Note– you do not need to find an exact value of x_{max} such that the error is 1%! Just find a reasonable value for x_{max} where the maximum error is in the range of 0.08 to 1.2%. Hint: Try plotting the difference $[f_1(x) - f(x)]$ to find the needed value of x_{max}.

2. Much like the fact that a polynomial is its own Taylor series, any trigonometric function in terms of the Fourier basis functions is its own Fourier series. To see this, consider the following function

$$f(x) = 6\sin(5\pi x), \;\; x \in [0,1]$$

Using the definitions we have learned for the Fourier series, show that this function is its own Fourier series expansion.

对于以下问题，求所示函数在区间 傅里叶正弦级数 $x \in [0,1]$ 上的傅里叶正弦级数。你无需重新推导系数 B_n 的公式（当然，如果这样更适合你，也欢迎你重新推导）。使用 50 项绘制函数图像。另请绘制前 50 项对应的谱。

3. $f(x) = 1$ 4. $f(x) = x$（你需要使用分部积分） 5. $f(x) = x^2$（这里需要使用两次分部积分） 6. $f(x) = \sqrt{x}$（可以使用符号数学程序计算定义系数B_n的积分） 7. $f(x) =$ 正弦(x)（可以使用符号数学程序计算定义系数B_n的积分）

8. $f(x) = \cos(x)$
(A symbolic mathematics program can be used to compute the integral defining the coefficients B_n)

9. $f(x) = \begin{cases} 1 \text{ for } x \le \frac{1}{2} \\ 0 \text{ for } \frac{1}{2} < x \le 1 \end{cases}$

10. $f(x) = \begin{cases} 0 \text{ for } x \le \frac{1}{2} \\ 1 \text{ for } \frac{1}{2} < x \le 1 \end{cases}$

11. $f(x) = \begin{cases} x \text{ for } x \le \frac{1}{2} \\ 0 \text{ for } \frac{1}{2} < x \le 1 \end{cases}$

12. $f(x) = \begin{cases} 1-2x \text{ for } x \le \frac{1}{2} \\ 0 \text{ for } \frac{1}{2} < x \le 1 \end{cases}$

13. $f(x) = \begin{cases} 0 \text{ for } x < \frac{1}{3} \\ 1 \text{ for } \frac{1}{3} \le x \le \frac{2}{3} \\ 0 \text{ for } \frac{2}{3} < x \le 1 \end{cases}$

For the following problems, find the *Fourier cosine series* on the interval $x \in [0,1]$ for the function indicated. You do not have to re-derive the formula for the coefficients B_n (although you are certainly welcome to if that works better for you). Plot the function using 50 terms. Also plot the associated spectrum for the first 50 terms.

14. $f(x) = 1$

15. $f(x) = x$
(you will have to use integration by parts)

16. $f(x) = x^2$
(you will have to use integration by parts twice here)

17. $f(x) = \sqrt{x}$
(A symbolic mathematics program can be used to compute the integral defining the coefficients A_n)

18. $f(x) = \sin(x)$
(A symbolic mathematics program can be used to compute the integral defining the coefficients A_n)

19. $f(x) = \cos(x)$
(A symbolic mathematics program can be used to compute the integral defining the coefficients A_n)

20. $f(x) = \begin{cases} 1 \text{ for } x \le \frac{1}{2} \\ 0 \text{ for } \frac{1}{2} < x \le 1 \end{cases}$

21. $f(x) = \begin{cases} 0 \text{ for } x \le \frac{1}{2} \\ 1 \text{ for } \frac{1}{2} < x \le 1 \end{cases}$

22. $f(x) = \begin{cases} x \text{ for } x \le \frac{1}{2} \\ 0 \text{ for } \frac{1}{2} < x \le 1 \end{cases}$

23. $f(x) = \begin{cases} 1-2x \text{ for } x \le \frac{1}{2} \\ 0 \text{ for } \frac{1}{2} < x \le 1 \end{cases}$

24. $f(x) = \begin{cases} 0 \text{ for } x < \frac{1}{3} \\ 1 \text{ for } \frac{1}{3} \le x \le \frac{2}{3} \\ 0 \text{ for } \frac{2}{3} < x \le 1 \end{cases}$

For the following problems, find the *Fourier series* for the function on the interval indicated. Use the full series expansion (involving both sine and cosine components) unless otherwise noted.

25. $f(x) = 1, x \in [-1,1]$

26. $f(x) = x, x \in [-1,1]$

27. $f(x) = |x|, x \in [-1,1]$

28. $f(x) = \exp(2x), x \in [-2,2]$.

29. $f(x) = \exp(2x), x \in [0,2]$; do this as a sine series.

30. $f(x) = \exp(2x), x \in [0,2]$; do this as a cosine series.

31. $f(x) = \frac{1}{1+x}$

Applied and More Challenging Problems

32. **偶延拓与奇延拓**。区间 $x \in [0,L]$ 有时称为半区间，因为该区间上的傅里叶正弦级数和余弦级数也都定义在区间 $[-L,L]$上。事实上，两者都.....

Fourier sine and cosine series are defined on $[-L,L]$, and they are periodic with period $2L$. For the sine series, the extension to $x \in [-L,L]$ should result in an odd function since sine is an odd function on this interval. For the cosine series, the extension to $x \in [-L,L]$ should result in an even function since sine is an even function on this interval.

对于以下问题，请在指定区间$x \in [0,L]$上求正弦和余弦傅里叶级数。然后，利用你得到的结果，证明这两个级数在 $[-L,L]$ 上有定义，并将其绘制在$x \in [-L,L]$ 上，取$N = 50$ 项。你会发现，正弦级数由函数在$x \in [0,L]$, 时的奇函数伸展定义，而余弦级数由函数在$x \in [0,L]$时的偶函数伸展定义。

a. $f(x) = x, \ \ x \in [0,1]$

b. $f(x) = x^2 \ \ x \in [0,2]$

c. $f(x) = \begin{cases} 1 \text{ for } x \le \frac{1}{2} \\ 0 \ \text{ for } \frac{1}{2} < x \le 1 \end{cases}$

d. $f(x) = \begin{cases} 0 & 0 < x < \frac{1}{4} \\ 8x-2 & \frac{1}{4} \le x \le \frac{1}{2} \\ 6-8x & \frac{1}{2} < x \le \frac{3}{4} \\ 0 & x > \frac{3}{4} \end{cases}$

(Note: This implicitly indicates that the function is defined on the interval $x \in [0,1]$.)

33. **周期延拓**。在后续周期区间$x \in [-L,L]$ 上对正弦和余弦傅里叶级数进行这种伸展，称为周期延拓。由于正弦和余弦函数都是周期函数，周期为 $2L$，因此，将针对$x \in [-L,L]$ 求得的级数向正方向或负方向延伸时，解会重复出现，这是因为底层三角函数具有周期性。

For the following problems, develop *both the Fourier sine and cosine series* for the interval $x \in [0,L]$ indicated. Then, show that the that the series is actually a periodic one that repeats every $2L$ by plotting the series for $x \in [-5L,5L]$. Use $N = 50$ terms.

a. $f(x) = x, \ \ x \in [0,1]$

b. $f(x) = x^2 \ \ x \in [0,2]$

c. $f(x) = \begin{cases} 1 \text{ for } x \le \frac{1}{2} \\ 0 \ \text{ for } \frac{1}{2} < x \le 1 \end{cases}$

d. $f(x) = \begin{cases} 0 & 0 < x < \frac{1}{4} \\ 8x-2 & \frac{1}{4} \le x \le \frac{1}{2} \\ 6-8x & \frac{1}{2} < x \le \frac{3}{4} \\ 0 & x > \frac{3}{4} \end{cases}$

(Note: This implicitly indicates that the function is defined on the interval $x \in [0,1]$.)

34. Consider the following function, defined piecewise over the interval $0 \le x \le 1$ (Fig. 4.26).

$$f(x) = \begin{cases} 0 & 0 < x < \frac{1}{4} \\ 16x - 4 & \frac{1}{4} \le x \le \frac{1}{2} \\ 12 - 16x & \frac{1}{2} < x \le \frac{3}{4} \\ 0 & x > \frac{3}{4} \end{cases}$$

Find the Fourier sine series for this function. Plot the function using the first 100 terms in the series.

35. Consider the following function, defined piecewise over the interval $0 \le x \le 1$ (Fig. 4.26). Note that this function is the derivative of the piecewise function given in problem 34.

$$f(x) = \begin{cases} 0 & 0 < x < \frac{1}{4} \\ 16 & \frac{1}{4} \le x \le \frac{1}{2} \\ -16 & \frac{1}{2} < x \le \frac{3}{4} \\ 0 & x > \frac{3}{4} \end{cases}$$

Find the Fourier sine series for this function. Plot the function using the first 100 terms in the series.

36. There are many requirements as to when a Fourier series can be differentiated term-by-term. While this question is a complex one in general, what we can say is that on the interval $x \in [0,1]$, if $f(0) = f(1) = f'(0) = f'(1) = 0$ *and* the function f is piecewise C^2, then both the Fourier sine and cosine series for $f(x)$ can be differentiated to obtain the appropriate series for $f'(x)$. Show that this is the case by doing the following.
    a. Compute the termwise derivative of the Fourier series for the function given in problem 34.
    b. Plot the function using the first 100 terms in the resulting series.
    c. Compare this sum to the exact solution given by the piecewise function given in 35.
37. For some functions, the evaluation of the sine or cosine series requires extra care and thought because the function $f(x)$ itself is one of the series terms. For instance, consider the following piecewise function

$$f(x) = \begin{cases} \sin(2\pi x) & 0 < x \le \frac{1}{2} \\ 0 & \frac{1}{2} < x \le 1 \end{cases}$$

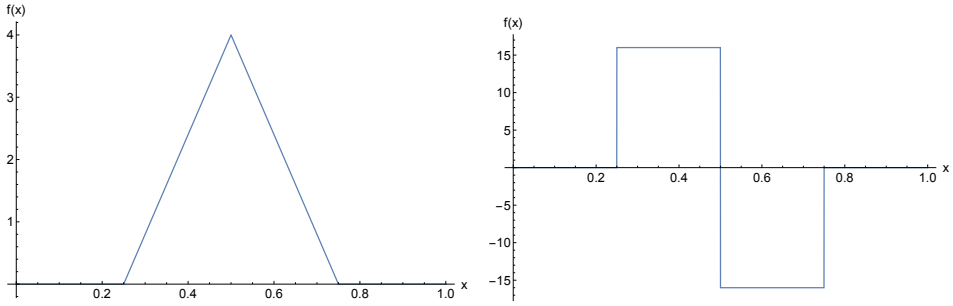


Fig. 4.26: The piecewise function defined for problems 34 (left) and 35 (right).

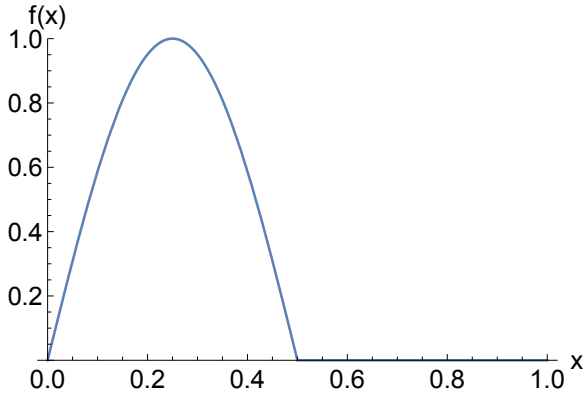


Fig. 4.27 The piecewise function defined for problems 37.

For this function, plotted in Fig. 4.27, the set of coefficients for the Fourier sine series, B_n, contain an apparent singularity (the denominator goes to zero). Thus, the problem requires some additional handling. To see one resolution to this problem, complete the following steps.

a. Compute the Fourier sine series for this problem. Verify that the result for the coefficients is

$$B_n = \frac{4\sin\left(\frac{1}{2}n\pi\right)}{(-4+n^2)\pi}$$

Note that for $n = 2$, this result presents a problem. The numerator is zero, and the denominator is zero, leading to a $0/0$ indefinite form.

b. 处理这个问题有几种方法。虽然B_n 的结果是正确的，但不定式的出现是因为我们计算的是一般情形下的积分。如果直接考虑$n = 2$ 的情形下的积分，就可以避免不定式，因为我们可以进行一些显而易见的化简，而在考虑更一般的情形时，这些化简并不明显。具体来说，注意 $n = 2$ 情形下的积分是

$$\int_0^{\frac{1}{2}} \sin(2\pi x)\sin(2\pi x)\,dx = \int_0^{\frac{1}{2}} \sin^2(2\pi x)\,dx$$

Once one combines the two functions into $\sin^2(2\pi x)$, the indefinite form no longer occurs. Show that the integration above leads to the result $B_2 = \frac{1}{2}$.

c. There is a second, equivalent way to handle this problem. Recall from your study of calculus L'Hôpital's rule, which states that for an $0/0$ indefinite form

$$\lim_{n\to n_0} \frac{f(n)}{g(n)} = \lim_{n\to n_0} \frac{f'(n)}{g'(n)}$$

Even though the function B_n is not a continuous one, we can certainly treat it as though it were. Show that, treating n as a continuous variable, you can use L'Hôpital's rule on B_n in the form

$$B_2 = \lim_{n\to 2} \frac{4\sin\left(\frac{1}{2}n\pi\right)}{(-4+n^2)\pi}$$

to show that $B_2 = \frac{1}{2}$.

d. 使用 $N = 10$ 项绘制傅里叶正弦级数，并绘制分段函数以比较二者 .

38. 函数 $f(x) = \exp(-(ax)^2)$（这是统计学和物理学中使用的经典高斯函数）具有一个奇特的性质。它的谱与函数本身的形状非常接近（假设 a 足够大，使得当 $x \to L$ 时，$f(x) << 1$；当 $x) \in [-L,L]$。对于本题，假设所关注的区间为 $x \in [-1,1]$。

a. 要看出频谱与原函数的形状相似，请计算此函数在该情形下的傅里叶余弦级数。为此，你需要注意以下积分。

$$\int_0^1 \exp(-a^2x^2)\cos(n\pi x)\,dx = \begin{cases} \frac{1}{2a}\sqrt{\pi}\,\mathrm{erf}(a), & n=0 \\ \frac{1}{a}\sqrt{\pi}\exp\left(-\frac{\pi^2n^2}{4a^2}\right)\mathrm{erf}(a), & n>0 \end{cases}$$

b. 假设 $a = 3$，则 $f(x) =$ 指数函数$(-(3x)^2)$。在两个独立的图上绘制：(i) $f(x)$ 在 $x \in [0,1]$ 上的原函数；(ii) $B_n = B(n)$。对于后一个函数，我们可以使用以下绘图技巧，将其扩展为外观更光滑的插值版本。技巧如下：不要绘制 $B(n)$，而是绘制系数 $B(n/100)$。对于此例，在区间 $n \in [0,600]$ 上绘制 $B(n/100)$。这样应能得到一张看起来光滑的插值图（尽管 B_n 的值会在横轴上每隔 100 个单位的除法处出现）。

在 Mathematica 中，绘制后一个函数的一条命令如下：

```
DiscretePlot[B[n/100], {n, 1, 600}, 范围大小 -> 完整, 颜色函数 -> "彩虹", 绘图值域 ->全部, 坐标轴标签 ->{"n", "B(n)"}, 图像大小 -> 大, 绘图值域 ->全部]
```

39. Proving that that the sine and cosine functions are *orthogonal* is something that was not shown in the text, but it is not difficult to illustrate. To do so, one needs to use the following trigonometric identities

$$\cos(A)\cos(B) = \frac{1}{2}\left[\cos(A-B)+\cos(A+B)\right]$$
$$\sin(A)\sin(B) = \frac{1}{2}\left[\sin(A-B)+\sin(A+B)\right]$$

Using these identities, show that the following are true.

$$\int_{x=0}^{x=1} \cos(n\pi x)\cos(m\pi x)\,dx = \begin{cases} \frac{1}{2} \text{ for } n=m \\ 0 \text{ for } n\neq m \end{cases}$$
$$n = 1,2,3,\ldots$$
$$\int_{x=0}^{x=1} \sin(n\pi x)\sin(m\pi x)\,dx = \begin{cases} \frac{1}{2} \text{ for } n=m \\ 0 \text{ for } n\neq m \end{cases}$$
$$n = 1,2,3,\ldots$$

NOTE: Please work this out carefully by hand.

“对我们来说，任何描述自然基本定律的方程都必然具有极高的数学美，这是一种近乎信仰的行为。这是一种非常有益的宗教信念，可以被视为我们取得诸多成功的基础。”

保罗·狄拉克谈量子力学定律，其中狄拉克δ函数作为一个重要特征出现

5

阶跃函数与δ函数

关于什么构成函数的概念，我们在第2章中曾作过简要介绍。虽然这一概念本身看似足够简单，但这在很大程度上是因为我们通常在整个教育过程中一直接触着这一概念。然而，“函数”这个词本身直到17世纪末才出现；当时，数学家戈特弗里德·莱布尼茨（因微积分而闻名）和约翰·伯努利（瑞士数学家，其儿子丹尼尔因伯努利原理、流体力学以及下文将讨论的伽马函数而闻名）开始更正式地发展函数这一概念。甚至到了19世纪初，许多人仍认为函数这一概念只适用于解析函数（见第2章）。

许多数学家都为对函数定义概念的推广作出了贡献。然而，毫不夸张地说，正是 Joseph Fourier 的理论推动了现代对函数进行定义的努力。通过使用傅里叶级数（见第 4 章），Joseph Fourier 能够构造出与数学家过去习惯研究的函数大不相同的函数。例如，有些傅里叶级数定义的函数在函数的定义域内存在值的不连续点（即跃变）。更令人担忧的是，这些不连续函数完全由连续函数的无穷级数构造而成，这似乎呈现出一个概念上的悖论。

虽然我们不会将函数论作为一个独立的主题来研究，但本节将介绍两种在应用中经常出现的重要函数：（1）阶跃函数（或 Heaviside 函数）；（2）狄拉克δ函数。第一种函数已在傅里叶级数一章中作过简要讨论。第二种函数，即狄拉克δ函数，从严格意义上讲并不是一个函数；更准确地说，它应称为广义函数或分布。然而，狄拉克*δ*函数这一术语在科学与工程学领域已根深蒂固，因此我们仍将采用这一（略有不准确的）术语。

你以前可能接触过狄拉克δ函数这个概念。狄拉克δ函数是一个很好的例子，说明某些数学概念的理论依据深受工程学和物理学领域的启发。事实上，为纪念物理学家保罗·狄拉克（1902年8月8日—1984年10月20日），人们常将这个函数称为狄拉克δ函数。保罗·狄拉克曾在描述量子力学时将狄拉克δ函数作为重要组成部分。狄拉克δ函数的另一个常见应用是在静电学中表示点电荷。在工程力学中，狄拉克δ函数用于表示集中载荷，

如上文所述，力作用于梁的情形就是一例。如前所述，狄拉克δ函数早已被广泛应用，远在其数学理论得到理解之前。20世纪50年代，名为Laurent Schwartz的数学家终于建立了更一般的广义函数理论（其中包括狄拉克δ函数）。虽然这一理论将函数的概念扩展到新的数学结构，但这里我们不打算深入探讨这一方向。相反，我们将专门讨论狄拉克δ函数及其相关的阶跃函数。讲解主要运用微积分工具展开。

5.1 术语

• **广义函数**。一种扩展了函数定义的概念。具体而言，广义函数是一种数学对象，它可能不符合普通（经典）函数的定义，但可以通过积分来描述它对其他函数的作用。所有普通函数都是广义函数，但有些广义函数并非普通函数。狄拉克δ函数是最广为人知的广义函数示例。

• **狄拉克δ函数**。一种广义函数，在物理上表示集中于一个极小的（相对于问题中其他尺度而言）时间或空间区间内的集中源。在数学上，狄拉克δ函数曾带来许多困难；最终，这些不同寻常的数学对象有了一个严谨的数学框架，通常称为广义函数理论。之所以采用“广义”一词，是因为这些新对象具有类似函数的用途，但从任何数学上的传统意义而言，它们都不是函数。尽管狄拉克δ函数的行为与任何已知的经典函数都不同（例如，它仅在一个点处非零，但其积分仍等于 1），它确实源于应用中的物理考量。例如，静电学理论中的点电荷可以用狄拉克δ函数表示。同样，材料力学中的点力也可以用狄拉克δ函数表示。

• **δ序列**。一个以整数n为索引的函数f，使得函数序列（$f1$，$f2$，$f3$，.....）随着$n \to \infty$的增大而变得越来越尖锐、狭窄。一个例子是常规的高斯函数，其形式写为

$$G(x;a,n) = \frac{1}{\sqrt{\pi(a/n)^2}} \exp\left(-\frac{x^2}{(a/n)^2}\right)$$

这个函数在图5.1中展示了n的几个不同取值。当n增加时，从图中可以很容易地看出，该函数变得越来越尖锐、越来越窄。在极限情况下，这个函数序列变成了一个狄拉克δ函数（因此得名）。

• **阶跃函数或海维赛德函数**。阶跃函数是在单个点 $x0$ 处不连续的函数。对于$x < x0$，函数值为零。在不连续点处，阶跃函数从 0 跃变至 1，并在$x > x0$ 时保持为 1。在$x = x0$ 处，函数在技术上未定义；不过，更好的理解方式是，函数在$x0$ 处可以取 0 到 1 之间的任意值 c。因此，阶跃函数实际上始终是一个完整的等价类 （参见 §2.2.3）函数，其中在$x0$ 处为函数选取的值决定了该等价类中的具体函数。

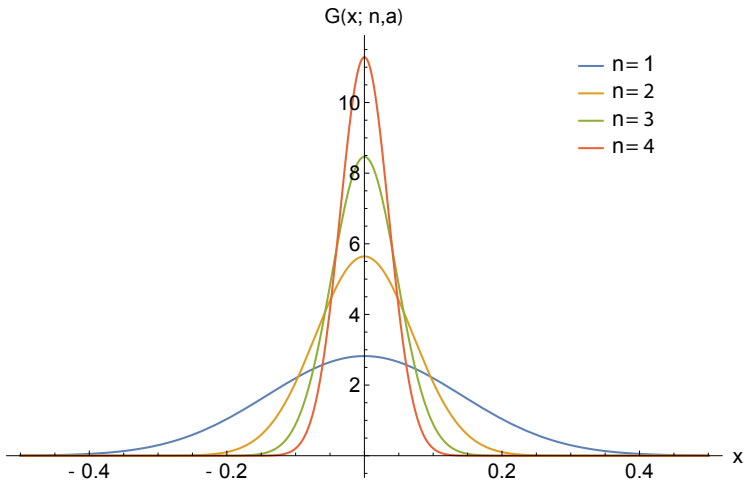


Fig. 5.1 A delta sequence. The Gaussian $G(x;\frac{1}{5},n)$ with standard deviation $\sigma=\frac{1}{5}n,\ n\in\mathbb{N}$ is an example of a *delta sequence*. As n increases, the function becomes increasingly peaked and narrow, even though the area stays constant at 1.

• **广义导数**。广义函数的导数。这一概念将通过分部积分来定义。其基本思想是赋予某种具体的含义，以说明例如对阶跃函数求导究竟意味着什么。从经典意义上讲，阶跃函数在不连续点处没有定义导数（即函数在长度为零的区间上跃变了一个有限距离，因此其经典导数为无限大）。广义函数理论赋予了阶跃函数的“导数”以明确的意义。虽然我们不会深入探讨广义函数理论本身（*per se*），但仍将采用一种方法，以直观（但在数学上严格）的方式理解阶跃函数的导数。

• **Function compact support.** A *function of compact support* is (for a single independent variable) a function that is nonzero only on some closed interval. While it is hard to conjure up images of such function, there are many examples that are even C^∞ (differentiable any number of times). One routinely adopted example is

$$f(x)=\exp(1/x)\exp[1/(1-x)]$$

This function is plotted in Fig. 5.2. Functions of compact support that are also C^∞ are sometimes called *bump functions* or *mollifying functions*.

• **伽马函数，** Γ。伽马函数是定义在实数轴上的解析函数。对伽马函数的一种理解是，它将阶乘的概念扩展到了实数。伽马函数的历史堪称一份古典数学界的“名人录”，许多数学家为函数的应用和性质作出了重要贡献，其中包括莱昂哈德·欧拉（1707–1783）、阿德里安-马里·勒让德（1752–1833）、卡尔·弗里德里希·高斯（1777–1855）、约瑟夫·刘维尔（1809–1882）、卡尔·魏尔斯特拉斯（1815–1897）和夏尔·埃尔米特（1822–1901）。

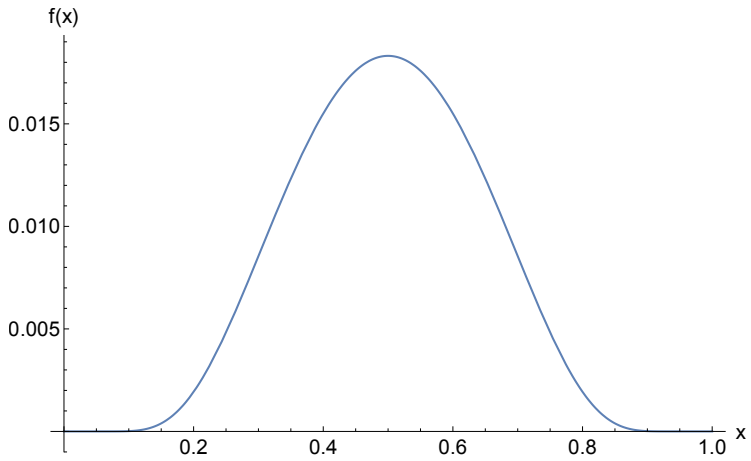


图 5.2 一个紧支撑函数。该函数仅在 $0 \leq x \leq 1$ 时非零，并且属于 C^{∞}。像这样的函数兼具紧支撑性和 C^{∞}，构成了一类重要的函数，称为隆起函数。

5.2 The Step and Delta Functions: The Basic Idea

本节旨在建立关于狄拉克δ函数和阶跃函数的基本直觉。狄拉克δ函数的直觉实际上相当容易形象化。从物理角度来说，狄拉克δ函数表示某种量（例如力、浓度、热能），它被理想化 为集中在单个点上。采用这种物理系统模型是有其目的的。例如，考虑尖楔压在刚性梁上所产生的力。图 5.3 展示了在两种分辨率下作用于梁上的力。从远处看，将力表示为仿佛作用在单个点上似乎是合理的。然而，我们必须考虑，把力施加到单个点上究竟意味着什么。从物理上讲，这是不可能的；所有物理接触都必须通过有限面积的相互作用发生。从物理上讲，点 并不存在；点是严格意义上的数学抽象。因此，如果我们放大力的作用区域，就会清楚地看到它并非发生在单个点上；事实上，力的施加发生在一个小面积上，而且一般来说，这类相互作用会以复杂的方式发生，我们无法详细了解其具体情况。

5.2.1 Mathematical Modeling of a Point Source

While it is true that *physically* we cannot apply a force to a single point, there are instances where *mathematically* we would like to model a quantity that is applied to a small area (relative to some other dimension; in this case, the length of the beam) as if it were applied to a point. The reason that we might want to make such a model are as follows.

1. Generally, we would not know the *actual* distribution of forces concentrated over a small region. Thus, we would not know how to represent the distribution of forces exactly.
2. Even if we could represent the distribution of forces exactly, there is the sense that, because of the difference in scales, $h \ll L$, it seems unlikely that the particular distribution of forces over h will be relevant.

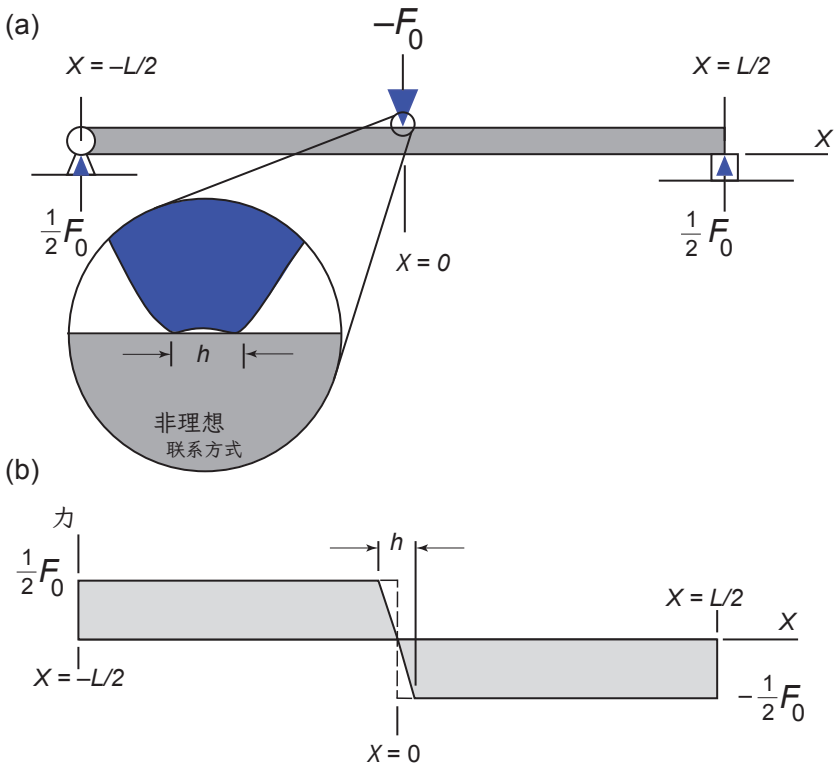


图 5.3：**非理想系统的理想化**。数学建模总是需要在一定程度上抽象“现实”。建模的诀窍在于，所作的抽象既要很好地近似实际问题，又要能得出数学上易于处理的结果。对许多我们感兴趣的物理量而言，使用狄拉克δ函数就是一种抽象（即一种并不精确、却仍能产生正确行为的表征）。在本图（a）部分的示例中，施加在梁上的实际力并不理想。这意味着它并非作用于单个点，而是关于x的复杂函数，通常我们并不知道其明确表达式（见放大图）。在本图（b）部分中，我们展示了力图（也称剪力图）的样子，其中的力分布在大小为h的小区域上。这里，我们近似认为该力在h的距离上均匀分布，尽管实际分布可能更复杂。

3. There is some hope that a mathematical model that represents the force as if it were applied to a single point would offer mathematical simplifications than would a more detailed model.

At this juncture, then, we need to address the question “*is it possible to generate a sensible model of the force as if it applied to a single point?*” This is our goal for this chapter.

To start our thinking about the problem, consider the set of functions illustrated in Fig. 5.4. Each of the rectangles shown has the same area– an area of unity (i.e., *area* = 1). Clearly, for each integer increase in n, the width of the rectangle decreases by a factor of 2, and the height increases by a factor of 2. Thus, as n increases, the rectangle becomes increasingly peaked and narrow. This function is given explicitly by

$$F_0\delta_n(x) = \begin{cases} F_0/(h/n) & \text{for } -\frac{h}{2n} < x < \frac{h}{2n} \\ 0 & otherwise \end{cases} \tag{5.1}$$

Some additional interpretation is helpful here. Consider the case $n = 1$ and $n = 2$. These case represents the force of the beam, were it to be spread uniformly over the interval of size h and $h/2$ respectively.

$$F_0\delta_1(x) = \begin{cases} F_0\frac{1}{h} & \text{for } -\frac{h}{2} < x < \frac{h}{2} \\ 0 & otherwise \end{cases} \tag{5.2}$$

$$F_0\delta_2(x) = \begin{cases} F_0\frac{2}{h} & \text{for } -\frac{h}{4} < x < \frac{h}{4} \\ 0 & otherwise \end{cases} \tag{5.3}$$

增加 n 的值会以显而易见的方式影响力密度。此外请注意，常数 F_0 不同于狄拉克δ函数 δ_n；它的作用是改变狄拉克δ函数的模。

作为这些函数的一个具体示例，假设我们要在区间 $-h/2 \leq x \leq h/2$上对 $F_0\delta_1$进行积分。我们将得到

$$\int_{\xi=-h/2}^{\xi=h/2} F_0\frac{1}{h}\delta_1(\xi)\,d\xi = F_0\int_{\xi=-h/2}^{\xi=h/2}\frac{1}{h}\delta_1(\xi)\,d\xi = F_0\frac{1}{h}h = F_0 \tag{5.4}$$

现在，我们得到了一个模型，其中相互作用的面积可以根据需要调节得任意大或小，同时施加在梁上的总力保持不变。利用这一表征，我们可以如下计算梁在任意位置x 处所受的力

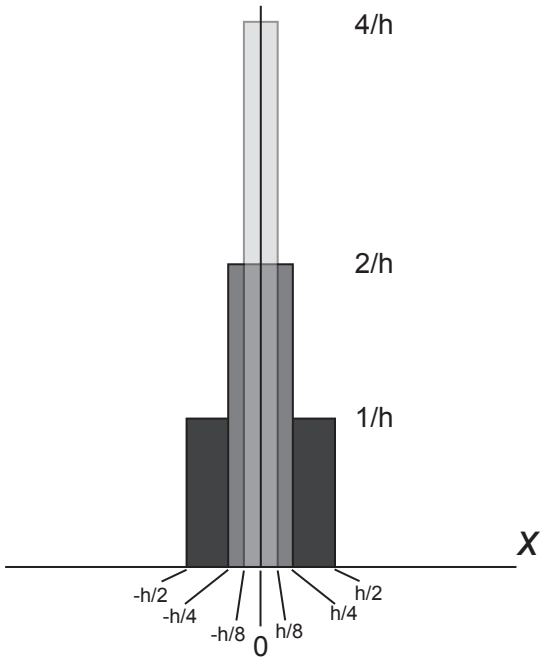
$$F(x) = \frac{F_0}{2} - \int_{\xi=-L/2}^{\xi=x} F_0\delta_n(\xi)\,d\xi \tag{5.5}$$

这一结果产生了如图 5.3 所示的梁中剪切力分布。我们可以如下理解描述该剪切力分布的方程（5.5）。在左边，支座提供大小为 $F_0/2$的向上力。当我们接近$x = 0$附近的区域时，单位宽度上的力在小距离h上均匀分布（按照我们的模型）。方程 5.5 给出的力表达式通过在该区域内随着x 增加而增大总向下力来表示这一点，使得在宽度 $-h/(2n) < x < h/(2n)$ 上对力密度积分后，可以恢复向下的力 F_0。最后，当x 大于$h/(2n)$时，所施加的总向下力就是F_0。只要稍加思考系统的物理学特性，应该就能与您的直觉相符。

5.2.2 The Path Forward: Limiting Behavior

现在，尽管我们的力分布模型比实际分布可能呈现的情况简单得多，但它仍未实现将力施加于单个点这一概念。事实上，对于那些在物理学课程或静力学课程中学习过梁的人来说，图 5.3 所示的力图看起来与他们以往见过的并不一样。通常，$x = 0$ 处的变化被视为仿佛力实际上施加在单个点上。在这种情况下，$x = 0$ 附近的力图会更像虚线所表示的曲线，而不是实线；实线显示了力在一个可观察到的有限区间内逐渐增大。有趣的是，

图 5.4 狄拉克δ函数的近似。
一列 矩形函数，以 n 为索引。矩形下方的面积固定为 1。随着n 增大，矩形变得更窄、更高，但面积始终保持为 1。



只要减小h 的值（或者根据方程 (5.1)，增大n ），我们就能使力的表征任意接近这条理想化虚线。这让我们开始对狄拉克δ函数的本质有所了解。看来，我们可以把这个函数看作前文所述函数的极限，因为该函数的宽度趋于零。

现在的问题是，如何从数学上严谨地阐释一个由物理模型启发的想法。数学与科学经常相互作用、彼此促进。狄拉克δ函数便是如此：物理学家和工程师早在它有坚实的数学理论作为基础之前，就已经使用它了。由于狄拉克δ函数可以直观地理解为一个具有合理解释的过程的极限，科学家和工程师广泛采用了它。相应地，数学家意识到，狄拉克δ函数并非通常意义上的函数。当时没有理论能够描述这类函数的数学性质，其应用也引发了不少数学难题。然而，如前所述，这个问题的数学理论最终在20世纪50年代得到了全面发展。如今，这类函数已在自然科学和数学领域得到广泛应用，人们可以放心使用。

5.3 A Construction for the Step and the Delta Functions

上述描述对狄拉克δ函数进行了较为直观的介绍。现在，我们将通过说明如何从适当的序列中构造狄拉克δ函数，进一步明确相关概念。一旦了解了这一构造过程及其适用的条件，我们就可以直接采用狄拉克δ函数。

作为一种更复杂过程的“简写”符号表示。幸运的是，无需在每个问题中重复描述这一过程的复杂性；一旦理解了狄拉克δ函数，就可以采用一组直观易懂的规则来使用它。

为了开始探讨狄拉克δ函数和阶跃函数，我们将修改之前的定义，以使用在某个区间上具有紧支撑且可微的函数。考虑以下函数及其积分。函数 δ_n 在 $x \in (-1/2, 1/2)$ 上严格来说具有紧支撑。指数 n 是大于零的整数。下面的函数按其面积 $a(n)$ 进行归一化，因此该函数的积分始终为 1（即，归一化函数的曲线下面积为 $area = 1$），与 n 的取值无关。该函数还是 $n+1$ 次可微的。

$$\delta_n(x) = \begin{cases} 0, & x < -\frac{1}{2} \\ \frac{1}{a(n)}(1-x^2)^n, & -\frac{1}{2} \le x \le \frac{1}{2} \\ 0, & x > \frac{1}{2} \end{cases} \tag{5.6}$$

$$H_n(x) = \int_{\xi=-1/2}^{\xi=x} (1-\xi^2)^n \, d\xi \tag{5.7}$$

$$a(n) = H_n(1/2) = \int_{\xi=-1/2}^{\xi=1/2} (1-\xi^2)^n \, d\xi \tag{5.8}$$

$H_n(x)$ 的表达式是将 $\delta_n(x)$ 积分到 x 的值；注意，这意味着 $H_n(1/2)$ 是函数 $\delta_n(x)$ 曲线下面积。这意味着 $\delta_n(x)$ 已经归一化，因此曲线下面积始终为 1。H_n 的数学表达式相当复杂，对于目前的讨论而言并非必需。为了帮助直观展示该函数及其积分，图 5.5 和图 5.6 绘出了 n 的各种取值下的情况 n。

A few things are apparent from these plots. First, the plot of the function $\delta_n(x)$ shows that $\delta_n(x)$ is a compact function (i.e., its domain is a finite one, $x \in [-1/2, 1/2]$) that becomes increasingly narrow and tall as n becomes large enough. It appears that, as we increase n, the maximum of $\delta_n(x)$ increases in value, and the values of the function not immediately near the origin become arbitrarily small, comparatively. In fact, we can determine the *second spatial moment about the vertical axis* (the variance) of the function $\delta_n(x)$ by integration by parts. This is a very messy computation, so the details will not be shown. However, the result is the following.

$$\begin{aligned} \sigma(n) &= \frac{1}{F_\Phi(1/2, n)} \int_{x=-1/2}^{x=1/2} x^2 (1-x^2)^n \, dx \\ &= \frac{1}{4(2n+3)} \end{aligned} \tag{5.9}$$

知道这个函数的特征宽度后，我们可以看出，随着n 无限增大，会得到一个有些奇怪的结果

$$\lim_{n \to \infty} \sigma(n) = 0 \tag{5.10}$$

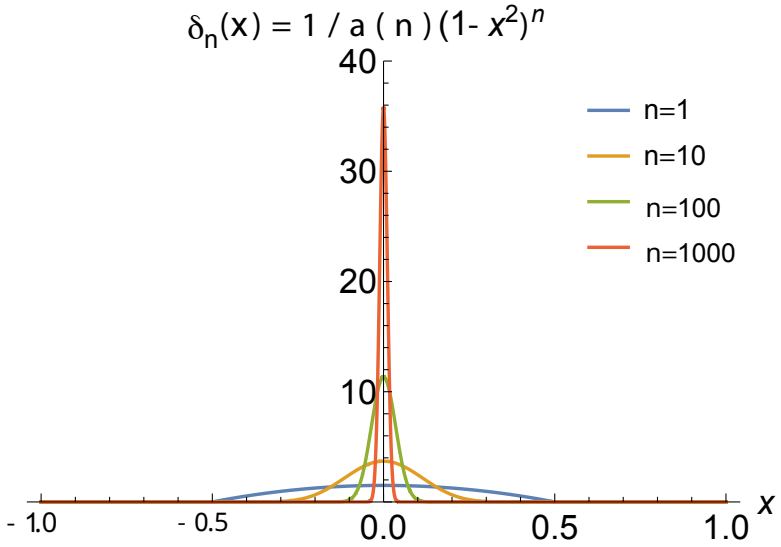


图5.5 狄拉克δ函数的近似。一列在极限下趋近于狄拉克δ函数的函数。这些序列使用函数 $\delta_n(x) = 1/a(n)(1-x^2)^n$ 构造，并通过 $a(n)$ 归一化，使其始终具有单位面积。随着n 的增大，该函数会迅速变得更窄、更尖峰。

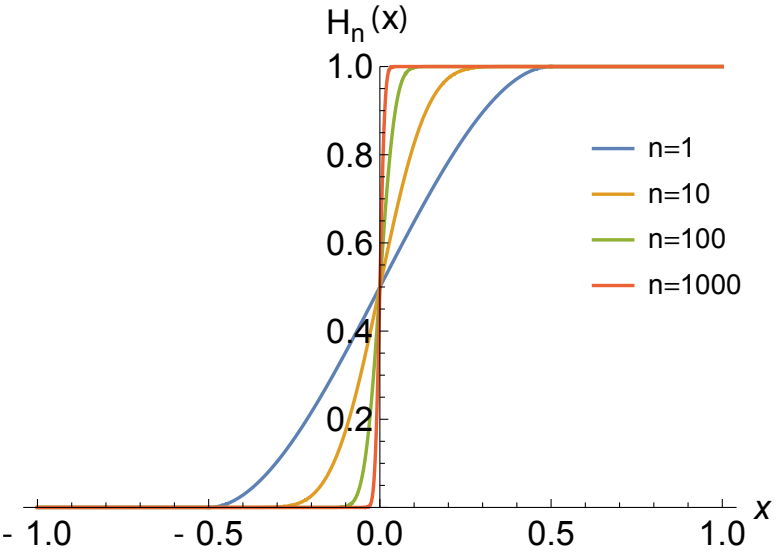


图5.6 阶跃函数的近似。一列在极限下趋近于狄拉克δ函数积分的函数。这些序列使用函数 赫维赛德阶跃函数 $1/\int 1-\xi^{2\,n}\,\xi\,(x)$ $a(n)_{x}{}^{-1/2}(\,)d$ 构造。随着n 的增大，该函数会迅速变得陡峭，呈阶跃状。阶跃函数也称为海维赛德函数，以推广其应用的电气工程师 Oliver Heaviside 命名。

再多做一些计算，我们就能求出 $\delta_n(x)$ 的曲线下面积（也称为零阶空间矩）

$$a(n) = \int_{\xi=-1/2}^{\xi=1/2} (1-\xi^2)^n \, d\xi = \frac{\sqrt{\pi}\Gamma(n+1)}{2\Gamma\left(n+\frac{3}{2}\right)} \tag{5.11}$$

用符号 Γ 表示的函数称为伽马函数；有关它的更详细介绍，请参阅例5.1（伽马函数）的灰色文本框。

例*5.1*（伽马函数）。

我们都熟悉阶乘的概念，n适用于任意整数。图 5.7 绘出了前几个整数对应的函数值。人们可能会想，是否存在一种类似于阶乘的概念，但适用于任意实数。事实证明确实存在，这个函数称为伽马函数。伽马函数是应用数学中最有趣、最重要的函数之一，最初由 Daniel Bernoulli 推导而来。不过，就我们的目的而言，可以把它看作是阶乘概念的推广。我们只给出伽马函数的定义，并开始使用它。之后，我们或许会回到这个定义，将其应用扩展到其他类型的问题。伽马函数定义为

$$\Gamma(n) = \int_{x=0}^{x=\infty} x^{n-1} e^{-x}\, dx \text{ , where} n \in \mathbb{R} \tag{5.12}$$

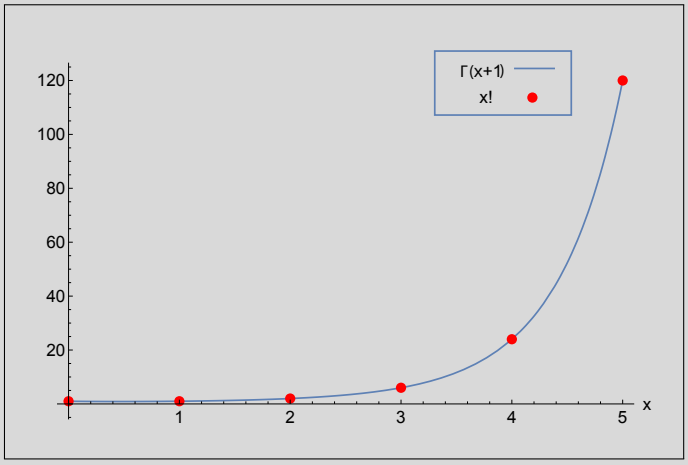


图 5.7：伽马函数在连续x范围内的曲线；作为比较，红色显示了x 取整数值时计算得到的阶乘。

这里，我们用n表示任意正实数，而不只是整数。虽然通常不会用n 表示实数，但这样有助于我们记住，当n 取整数值时，它就等于阶乘。

Now that we have defined the gamma function, we can turn to the area under the curve for the function $f(x) = (1-x^2)^n$. It turns out that

$$a(n) = \int_{x=-\frac{1}{2}}^{x=\frac{1}{2}} (1-x^2)^n\, dx = \frac{\sqrt{\pi}\Gamma(n+1)}{2\Gamma\left(n+\frac{3}{2}\right)} \tag{5.13}$$

which gives us the necessary result for the area under the curve, $a(n)$.

Some additional features of the gamma function

We have the following identity (which will not be proven here)

$$\Gamma(n) = (n-1)! \tag{5.14}$$

There are many interesting properties for the gamma function. Possibly the most relevant one is as follows

$$\begin{aligned} \gamma(1) &= 1 \\ \gamma(n+1) &= n\gamma(n) \\ n! &= \gamma(n+1) \end{aligned}$$

最后一条说明的是，对于整数值，伽马函数等价于阶乘。利用上述恒等式，并从 $n = 1$ 开始逐步推导，即可得到这一结论。

现在，让我们看看当n 变大时，这个函数会发生什么变化。图 5.5 绘出了n取不同值时的函数图像。当$n = 1$ 时，函数起初非常平坦；随着指数n 增大，函数迅速变得非常尖峭。为便于讨论，定义 $\delta_n(x) = {}^{1}\!/_{a(n)} f_1(x)$，使得

$$f_1(x) = (1-x^2)^n \tag{5.15}$$

还要注意以下内容

$$\lim_{n\to\infty} a(n) = \lim_{n\to\infty} \frac{\sqrt{\pi}\Gamma(n+1)}{2\Gamma\left(n+\frac{3}{2}\right)} = 0 \tag{5.16}$$

因此，随着n 增大，用于归一化曲线的面积也趋于零。

现在，注意以下内容。函数$f_1(x) = (1-x^2)^n$ 的最大值位于$x = 0$处，其值始终为$f_1(0) = 1$。不过，函数在其他所有点的取值都小于 1。回想一下，对于小于 1 的正数m_0，有

$$m_0 > m_0^2 > m_0^3 > m_0^4 \ldots \tag{5.17}$$

This means that as n grows, all of the values of the function become smaller, except the value at $x = 0$, which remains 1 regardless of how large n is. Thus, once normalized by $a(n)$, the only option for the function is to become more and more peaked as n becomes arbitrarily large, with the value at $x = 0$ being equal to $1/a(n)$.

Recalling that $a(n)$ is tending toward zero as n becomes arbitrarily large, and we find that in the limit our function f_δ behaves quite unusually in the limit. It apparently has zero width (as measured by the variance), and infinite height. And, because of its construction, it also has a total area of 1; in other words

$$\lim_{n\to\infty} \left(\int_{-\infty}^{\infty} f_\delta(x,n)\, dx \right) = 1$$

While it may seem like a peaked function with zero width could not somehow also have a finite area, note that the width (variance) times height (roughly, the area) of this function is, in the limit, a $0 \cdot \infty$ form. Recall from Chapter 1, we had the example of the function $g(x) = \frac{1}{x}\sin(x)$. In the limit, as $x \to 0$, we can show using L'Hôpital's rule that

$$\lim_{n\to\infty}\frac{\sin(x)}{x}=\lim_{n\to\infty}\frac{\cos(x)}{1}=1 \tag{5.18}$$

Therefore, it is not that unusual that a $0\cdot\infty$ form leads to a result that is neither 0 nor ∞. This is essentially the case that we have for our function f_δ.

观察函数 $\delta_n(x)$ 的积分 随n 增大而发生的变化也很有意思。虽然该函数在数学上由公式（5.7）所给出的表达式定义，但直接观察图5.6或许更有助于了解这个函数的行为。当n 的值很小时，函数 $\delta_n(x)$ 有些宽且光滑。因此，当我们积分到值域 $-1/2 < x < 1/2$内的x 值时，结果是一个光滑的“S”形函数，其最大值为 1。函数 F_Φ 的最大值为 1，因为函数 f_δ 经过归一化，其面积始终为 1。

对于F_Φ，随着n 的值变得任意大，该函数会越来越像一个阶跃 。这与我们对函数f_δ的了解一致：我们知道，随着n 变得任意大，函数f_δ 会变得越来越窄。因此，对该函数积分贡献最大的部分会集中在$x = 0$附近；于是，随着n 增大，函数F_Φ 在$x = 0$ 处呈现出从零陡然跃升至 1 的阶跃。

5.4 Delta Sequences

In the previous section, we developed a *sequence* of functions indexed by the integers, n that became increasingly peaked and narrow as n increased. These functions were defined by

$$f_1(x)=(1-x^2)^n$$

It turns out that there are any number of sequences of functions that one can make that, in the appropriate limit, the sequence defines a delta function. For example, the function

$$f(x)=\begin{cases}0 & \text{for } x<-\frac{1}{2}\\ \frac{\sqrt{\pi}\Gamma\left(\frac{n}{2}+1\right)}{\Gamma\left(\frac{n+1}{2}\right)}\left[\cos(n\pi x)\right]^n & \text{for } -\frac{1}{2}\le x\le\frac{1}{2}\\ 0 & \text{for } x>\frac{1}{2}\end{cases} \tag{5.19}$$

is also a delta sequence (this is plotted in Fig. 5.8).

The fact that there are multiple kinds of delta sequences that can lead to the same result indicates that the delta function is not a single entity; it is actually an equivalence class of sequences, any of which can be used equally well to define the (generalized) delta function. A delta sequence, then, can be defined as follows.

Definition 5.1 (delta convergent sequence.). A delta (convergent) sequence, $\delta_n(x)$, is any sequence of functions such that

$$\lim_{n\to\infty}\left(\int_{-\infty}^{\infty}\delta_n(x)f(x)\,dx\right)=f(0) \tag{5.20}$$

There are an infinite number of such functions. For example, if g(x) is any non-negative function such that

$$\int_{-\infty}^{\infty}g(x)\,dx=1 \tag{5.21}$$

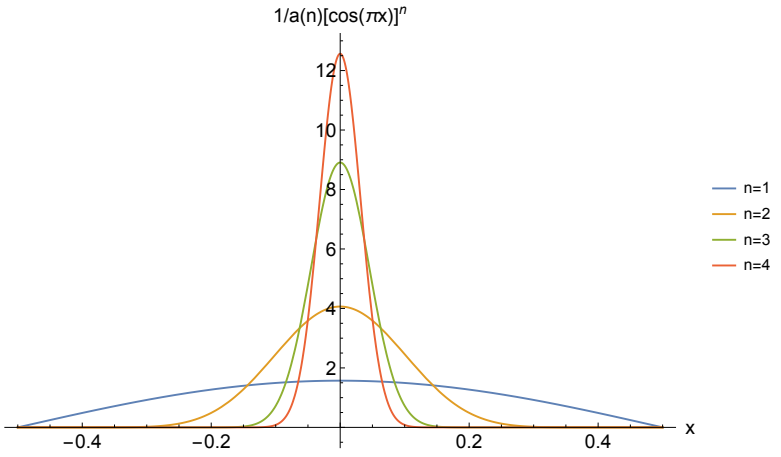


Fig. 5.8 Another delta sequence. The cosine function raised to the power n and normalized by its area (as a function of n) is a delta sequence.

then rescaling this function in the form

$$\delta_n(x) = ng(nx), \quad n = 1,2,3,\ldots \tag{5.22}$$

is a delta convergent sequence.

5.5 Properties of the Delta Function

Ultimately, the purpose for defining the delta function is because of its unique analytical properties, which make the representation of point sources in physical system very convenient. There are two primary analytical properties of the delta function that make it convenient in applications. Let I be either a compact ($\{I : x \in [a,b]\}$) or non-compact ($\{I : x \in [-\infty,\infty]\}$) interval containing the point x_0. Furthermore, we must assume that *the function is at least continuous* at x_0 (i.e., there cannot be a jump discontinuity there). Then, the following two properties are true for the delta function.

Property 1: Unit integral.

$$\int_{x\in I} \delta(x)\,dx = 1 \tag{5.23}$$

Property 2: The sifting property.

$$\int_{x\in I} f(x)\delta(x-x_0)\,dx = f(x_0) \tag{5.24}$$

The first property arises by construction; the delta sequences are always structured such that they have unit area, so the limiting function must also have unit area. The second of these properties is the so-called *sifting* property of the delta function. It is just an application of the definition given by Eq. 5.20, but one where the delta function is shifted by the amount x_0. The simplicity of the integration of the delta function is one of its

most attractive features.

Example 5.2 (The sifting property).

The sifting property of the delta function is called this because it offers a method to extract (i.e., to *sift out*）通过积分得到函数在单个点处的值。考虑以下示例。

$$\int_3^5 y^3 \ln(y-2)\sin(y^2)e^{-y}\delta(y-4)\,dy$$

Ordinarily, this integral would be impossible to compute analytically (without the delta function, a symbolic integration software cannot resolve the integrand of this integral). However, because of the sifting property of the delta function, this integral is simple. Essentially, the integral is zero everywhere, except at the single point where the independent variable of integration is equal to 4. At that point, we have

$$\delta(y-4) = \delta(4-4) = \delta(0)$$

Thus, the single point that contributes to the integral is the one where the argument of the delta function is zero, that is, the point $y = 4$. Thus, the result is

$$\int_3^5 y^3 \ln(y-2)\sin(y^2)e^{-y}\delta(y-4)\,dy = 4^3 \ln(4-2)\sin(4^2)e^{-4}$$

这只不过是将被积函数在 $y = 4$ 处取值。

作为第二个更抽象的例子，考虑下面的积分。与上述情况一样，假设 x_0 是积分定义域内定义的任意一点。那么，即使是高斯函数，与狄拉克δ函数相乘后进行积分，也可以很容易地求值。

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2a^2\pi}} \exp\left(-\tfrac{1}{2}\frac{x^2}{a^2}\right)\delta(x-x_0)\,dx = \frac{1}{\sqrt{2a^2\pi}} \exp\left(-\tfrac{1}{2}\frac{x_0^2}{a^2}\right)$$

We will see examples similar to this one in the course of studying the solution to partial differential equations. There, the Gaussian is known as the heat kernel, and the delta function would represent the distribution of heat specified as an initial condition.

5.6 An Explanation of the Sifting Property

For piecewise smooth functions, understanding how and why the delta function behaves the way it does is reasonably straightforwards. We need only use the ideas of limits, and the expansion defined by the Taylor series.

在接下来的推导中，我们假设函数 f 在零附近是解析的。为便于说明，我们将考察以下函数：f

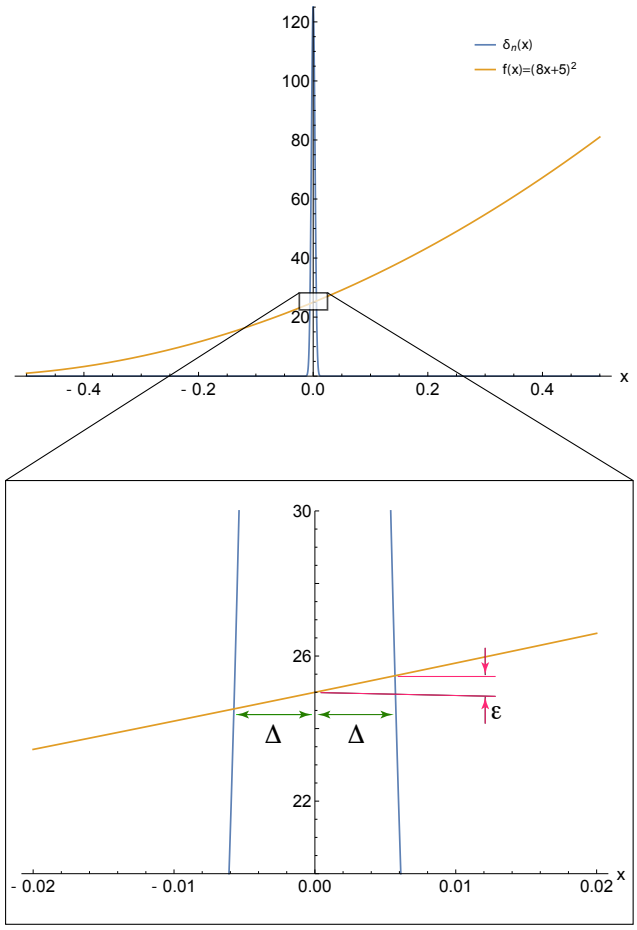


Fig. 5.9 A function, f in the vicinity of a delta sequence function. The function changes by less than or equal to the amount 2ε over the interval given by $\{I : -\Delta < x < \Delta\}$. The value of *epsilon* can be made as small as we like by making Δ smaller. By definition, there is always a delta sequence with width less than 2Δ around zero; we need only make the integer n large enough.

$$f(x) = (8x+5)^2 \tag{5.25}$$

图 5.9 展示了该函数的图像，并绘出 δ序列函数中的一个成员以供参考。

For the interval around zero, we can write the following Taylor series expansion for f because f is assumed to be analytic there.

$$f(\varepsilon) = f(0) + \Delta f'(0) + \frac{\Delta^2}{2!} f''(0) + \ldots \tag{5.26}$$

Note that we can take Δ to be as small as we like because we can always find a delta sequence that has width less than 2Δ simply by making n large enough. From chapter 1, we know that none of the derivatives of f must be finite in the region around zero (this is a fact arising from the fact that f is analytic). Therefore, we can always find a *Delta* sufficiently small such that

$$f(0) \gg \Delta f'(0)$$
$$f(0) \gg \tfrac{1}{2!}\Delta f''(0)$$
$$\cdots$$

This completes the construction. What we have shown is that we can always find a member of the delta sequence of functions such that the width of the function is small enough such that we can estimate f by its value at $x = 0$, with an error that can be made as small as we like. This validates the use of the delta function as having the property, in the limit, of

$$\int_{x\in I} f(x)\delta(x)\,dx = f(0) \tag{5.27}$$

将这一推导伸展至狄拉克δ函数在其他点的应用，只需对狄拉克δ函数进行适当平移（仿射变换）。

5.7 *δ序列极限的一个可计算示例

使用 δ序列的一个难点在于，大多数结果（至少在一般情形下）都是通过抽象构造来理解的。换句话说，人们不常能找到一个可计算示例，来展示 δ序列如何按照我们一直讨论的意义收敛。

有一些实例经过一定努力是可以计算的。在下面的实例中，给出了一个δ序列的实例。该实例的目的是说明，当函数序列与连续函数 $f(x)$ 进行积分时，狄拉克δ函数的筛选性质仍然成立。这一过程分为两个步骤。第一步，将δ序列与 $f(x)$ 进行积分。第二步，在积分完成后求 $n \to \infty$ 的极限，以表明所得结果确实符合狄拉克δ函数的筛选性质。需要注意的是，求极限和积分的过程不一定可交换！事实上，在这种情况下，两者显然不可交换。如果我们首先求δ函数的极限，然后再尝试进行积分，就会发现δ函数仅在一个点上具有支座；无论采用何种积分方式，这样的积分必然为零！

Example 5.3 (A computable example of the sifting property).

There are many functions, both compact (i.e., defined on a closed interval) and non-compact (defined for the entire real number line). For those who might be interested, compilations of such functions are available in the literature (e.g., Dang and Ehrhardt (2012)). The following function is one that is interesting, and also one that can be handled analytically.

$$\delta_n(x) = \begin{cases} \frac{1}{a_n}(1-4x^2)^n & \text{for } \frac{1}{2} \le x \le \frac{1}{2} \\ 0 & \text{otherwise} \end{cases}$$

其中

$$a_n = \frac{\sqrt{\pi}\Gamma(n+1)}{2\Gamma\left(n+\frac{3}{2}\right)}$$

This function is, in fact, the one provided at the start of this chapter. Examples of the shape of this function with increasing n are given in Fig. 5.5.

One of the nice properties of this function is that it can be integrated analytically. Thus, we can consider the following.

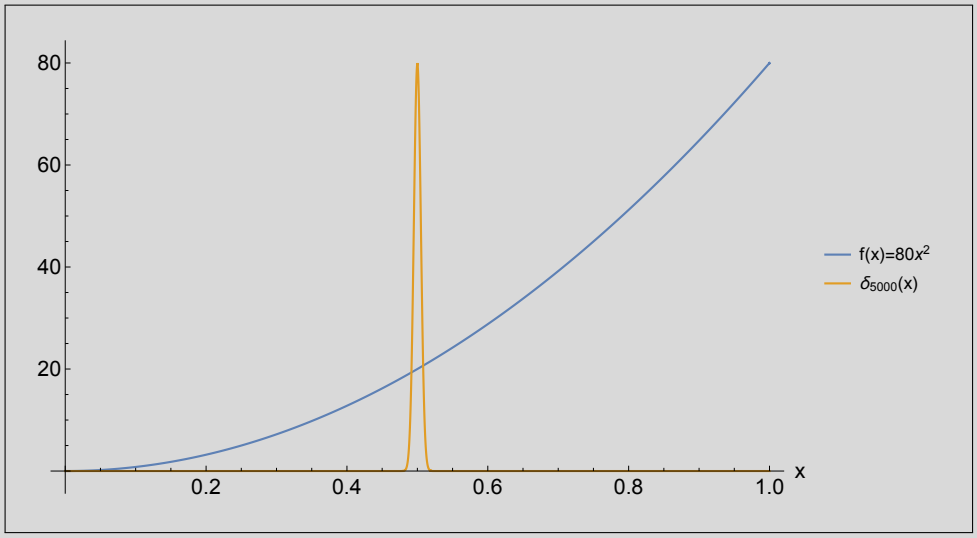


Fig. 5.10: The function $f(x) = 80x^2$ plotted against the delta function $\delta_n(x-\frac{1}{2})$ defined above, with $n = 5000$.

If we integrate our pre-delta function against an explicit function, say $f(x) = 80x^2$, then it may be possible to evaluate this integral. In this example, we will shift the pre-delta function by one-half; an example of the pre-delta function and the function $f(x)$ are plotted in Fig. 5.10. Mathematically, the problme is represented as follows

$$\int_{x=0}^{x=1} f(x)\delta_n(x-\tfrac{1}{2})\,dx = \int_{x=0}^{x=1} \underbrace{80x^2}_{f(x)} \underbrace{\frac{1}{a_n}(1-4(x-\tfrac{1}{2})^2)^n}_{\delta_n(x-\frac{1}{2})}\,dx$$

It turns out that this definite integral can be computed for any value of n. The result is

$$\int_{x=0}^{x=1} f(x)\delta_n(x-\tfrac{1}{2})\,dx = 80\frac{2+n}{2(3+2n)}$$

This is the result of the integration of any of the infinity of pre-delta functions, $\delta_n(x-\frac{1}{2})$, against the function $f(x) = 80x^2$ for any value of n. To find the final result, we need only take the limit of $n \to \infty$

$$\lim_{n\to\infty} 80\frac{2+n}{2(3+2n)} = 80\cdot\frac{1}{4} = 20$$

This means that we have for our particular choice of pre-delta function, we have the result

$$\lim_{n\to\infty}\left(\int_{x=0}^{x=1} 80x^2\left[\frac{1}{a_n}(1-4(x-\tfrac{1}{2})^2)^n\right]dx\right)=20$$

而这个结果正是$f(\frac{1}{2})=80(\frac{1}{2})^2$的结果。- 因此，这至少提供了一个明确的例子，说明狄拉克δ函数的极限恰好给出了狄拉克δ函数筛选性质所指定的结果。

5.8 阶跃函数及其（广义）导数

本章将狄拉克δ函数和阶跃函数放在一起讲解是有原因的。事实证明，两者之间存在一种有趣的关系。考虑下面的表征。

$$H_n(x-x_0)=\int_{y=-\infty}^{y=x}\delta_n(y-x_0)\,dy \tag{5.28}$$

其中，$H_n(x-x_0)$是一个“预”阶跃函数；也就是说，它类似于图5.11中展示的S形函数之一（图中展示的是$x_0=0$的情况）。随着n 的值增大，会发生两件事。首先，函数 δ_n 变得更加尖锐、狭窄（这对我们来说已经很熟悉了）；其次，S形函数变得更陡，更接近阶跃函数。由式（5.28）可知，H_n 是一个光滑且可微的函数。其导数为

$$\frac{d}{dx}H_n(x-x_0)=\delta_n(x-x_0) \tag{5.29}$$

因此，我们可以将狄拉克δ函数视为阶跃函数的导数。通常，只有当狄拉克δ函数出现在积分号下时，我们才应尝试对其进行解释；换言之，它不是普通函数，而是一种广义函数，从严格意义上讲，它在积分之外并未定义。然而，在积分之外将狄拉克δ函数用作符号表示，是一种常见甚至约定俗成的符号滥用，尤其是在物理学、工程学和偏微分方程数学中。然而，我们必须始终牢记，其含义 在于：狄拉克δ函数表示一个函数序列，其极限只能在积分过程中进行解释。

5.9 狄拉克δ函数是否具有傅里叶级数表示?

这个问题同时深刻而又简单。狄拉克δ函数根本不是函数，因此询问它是否具有傅里叶级数表征似乎有些奇怪。然而，“前”狄拉克δ函数δ_n 实际上是普通函数。根据所选的δ序列，这些前狄拉克函数甚至可能是C^∞ 光滑的。

In the problems from the previous chapter, the Fourier series for a Gaussian function was examined. It turns out that we can adopt this function to be a delta sequence as follows. First, we define the sequences

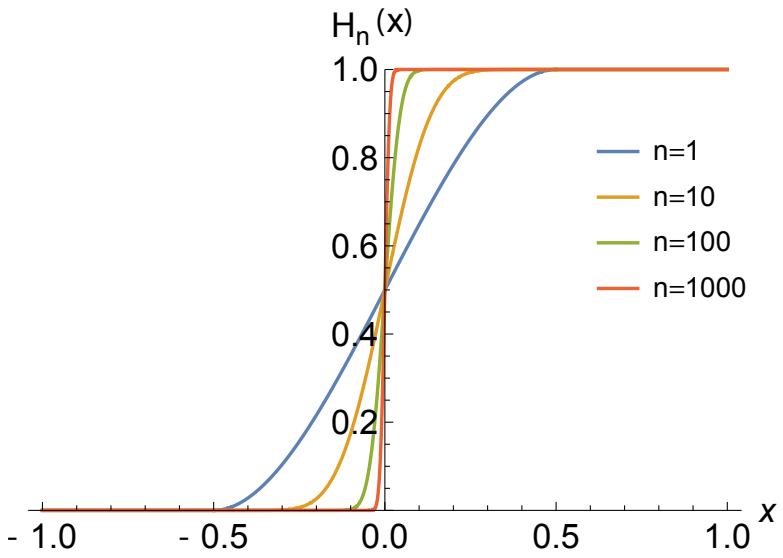


Fig. 5.11 Step function sequences. A *sequence* of functions that approaches the *integral* of the delta function in the limit. This function converges to the step function as $n \to \infty$.

as a function of m (here, we have switched from n to m to prevent confusion when examining the Fourier series) as follows.

$$\delta_m(x) = \frac{\alpha m}{\sqrt{\pi}\,\mathrm{erf}(\alpha m)} \exp\left(-\alpha^2 m^2 x^2\right) \tag{5.30}$$

$$A_0' = \frac{1}{2}\int_{-1}^{1} \delta_m(x)\,dx = \frac{1}{2}$$

$$A_n = \frac{1}{1}\int_{-1}^{1} \delta_m(x)\cos(n\pi x)\,dx = \exp\left(-\frac{\pi^2 n^2}{4\alpha^2 m^2}\right)$$

现在注意，当$m \to \infty$取极限时，我们得到的是狄拉克δ函数，而不是δ序列。傅里叶系数取值为

$$A_0' = \frac{1}{2}\int_{-1}^{1} \delta_m(x)\,dx = \frac{1}{2}$$

$$A_n = 1$$

这就得到一个级数，其中所有振幅都相同——它们全都等于 1（A_0'除外）。这说明，狄拉克δ函数包含所有可能的频率，并且振幅值通常不会减小。这是一个非常不寻常的结果！对于更普通的函数，可以证明傅里叶系数的振幅必然会逐渐减小（这与帕塞瓦尔定理有关），事实上，在求和时甚至会得到一个有限值。然而，对于狄拉克δ函数，我们不存在这种关系；振幅之和趋于无穷。虽然狄拉克δ函数的傅里叶级数确实违反了帕塞瓦尔定理，但我们不应对此感到过于惊讶。狄拉克δ函数不是普通函数，因此，证明帕塞瓦尔定理所需的条件并不适用于狄拉克δ函数。

通过对狄拉克δ函数的傅里叶级数求和，可以得到该函数的预测图像，观察这幅图像很有意思；该图像见图5.12。狄拉克δ函数的傅里叶变换在应用中得到了广泛使用，

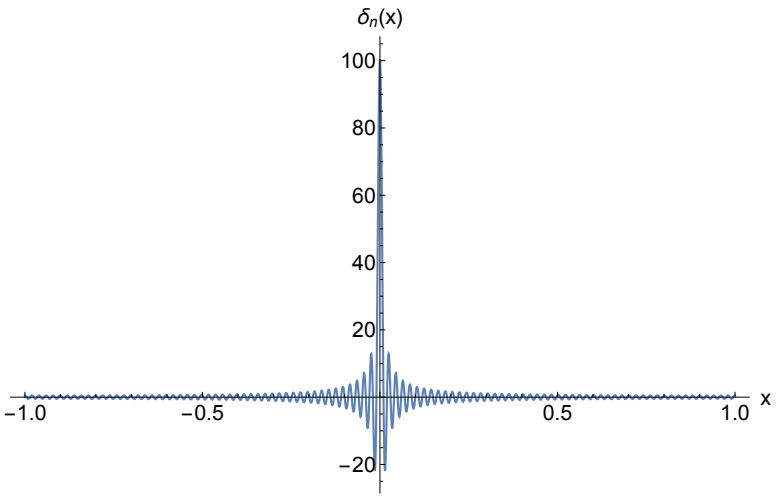


Fig. 5.12 The Fourier series sum for the delta function. The delta function has an unusual Fourier series expansion. In this figure, the series is approximated by the first 101 terms in the sum. Note that the resulting sum does actually become peaked and narrow at $x = 0$, as expected for approximations to the delta function.

such as in signal processing or modeling of point sources. We will see more about Fourier analysis of the delta function when we study the Fourier transform later on.

5.10 Some Identities for the Step and Delta Functions

There are a few identities that can be useful when working with step and delta functions. Some of these have been presented above, but are repeated here for reference.

$$\delta(-t) = \delta(t) \tag{5.31}$$

$$\delta(t-a) = \delta(a-t) \tag{5.32}$$

$$\int_0^b \delta(x-a) f(x)\,dx = f(a), \quad b > a \tag{5.33}$$

$$H(x-a) = [1 - H(a-x)] = [1 - H(-(x-a))] \tag{5.34}$$

赫维赛德阶跃函数(x) − 赫维赛德阶跃函数$(-x)$ = sgn(x), sgn(x) is the called the *sign function*, because it returns the sign of the argument (5.35)

$$B(x;a,b) = [H(x-a) - H(x-b)], \quad b > a \qquad \text{The } \textit{boxcar} \text{ function} \tag{5.36}$$

$$\int_0^b \frac{d}{dx} H(x)\,dx = \int \delta(x)\,dx = 1, \quad b > 0 \tag{5.37}$$

$$\int_0^b \frac{d}{dx} H(x-a)\,dx = \int \delta(x-a)\,dx = 1, \quad b > a \tag{5.38}$$

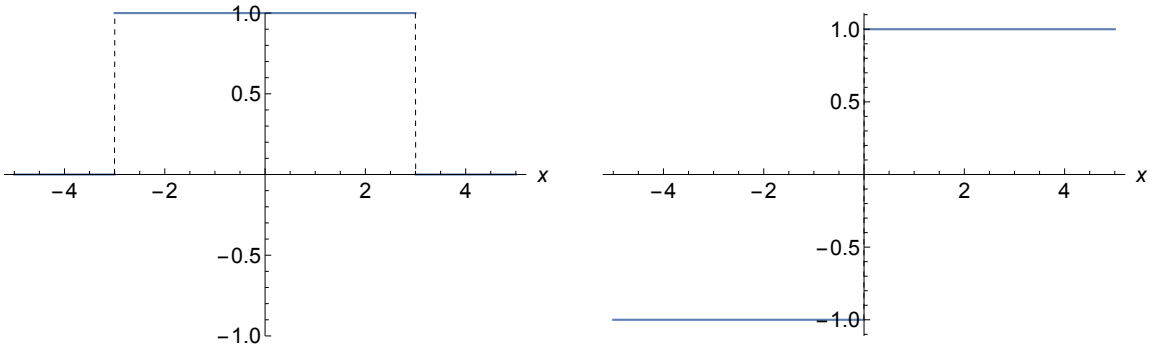


Fig. 5.13: (Left) The boxcar function with $a = -3$ and $b = 3$. (Right) The sign function.

这里介绍了两个新函数。第一个是棚车函数，因其形状类似于铁路棚车而得名。第二个称为符号函数，它具有一个有用的性质，即返回函数辐角的符号（± 1）。

习题

练习题

For the following problems, evaluate the integral.

1. $$\int_{-1}^{1} \delta(x)\,dx$$
2. $$\int_{-\infty}^{\infty} f(y)\delta(y)\,dy$$
3. $$\int_{-\infty}^{\infty} y^3\delta(y)\,dy$$
4. $$\int_{-\infty}^{\infty} y^3\delta(x-y)\,dy$$
5. $$\int_{z=-1}^{z=1} \delta(z)f(x-z)\,dz,$$ where $0 < x < 1$
6. $$\int_{z=0}^{z=1} \delta(x-z)f(z)\,dz,$$ where $0 < x < 1$
7. $$\int_{-\infty}^{\infty} \frac{\sin(y)}{\sin(x)}\delta(x-y)\,dy$$
8. $$\int_{-\infty}^{\infty} \frac{y^3-x^3}{x^2}\delta(x-y)\,dy$$
9. $$\int_{a}^{b} \sin(e^x)\delta(x-x_0)\,dx,$$ where $a < x_0 < b$
10. There are sequences that are the complement of delta sequences. These are the zero sequences. A delta sequence an extreme transformations of a function; when integrating a function against a delta sequence, the only information that passes through comes from a single point. The zero sequences act in a complementary way; although its integral is always unity, when integrating a function against a zero sequence, in the limit the result is zero. As an example, the following is an zero sequence on the real line.

$$\iota_n(x) = \begin{cases} \frac{1}{2n} & \text{for } -n \le x \le n,\ \ n \in \mathbb{N} \\ 0 & \text{otherwise} \end{cases}$$

11. Show that the function

$$(an)Z_n\left(\frac{x}{an}\right)$$

Is an zero function as $n \to \infty$

12. Show that every delta sequence can be transformed into an zero sequence by the substitution $n \to 1/n$.

Applied and More Challenging Problems

1. Heating of a surface at point can be represented by using the delta function. For this problem, suppose you are interested in finding the steady state temperature distribution for an insulated wire (so it does not loose heat to the environment) of length L and constant cross-sectional area, A. The wire is heated at its center point, $x = L/2$, by a powerful laser (where a small amount of insulation has been removed so that the wire can be heated). Suppose also that the left end of the wire is held in liquid nitrogen at $T_0 = 77$ Kelvin (K) and the right side is held at $T_1 = 194.5$ K in liquid CO_2. Assume $L = 1$ m.

$$\begin{aligned} K_T \frac{d^2T}{dx^2} &= -q_0 \delta(x - L/2) \\ T(0) &= T_0 \\ T(L) &= T_1 \end{aligned}$$

   Note! The delta function (curiously) must have units of $1/Length$ if it is to have an integral of unity (without units). Thus, the equation above is dimensionally correct. For this problem, assume the heat flux, q_0 is $q_0 = 200{,}000$ W m^{-2}, and the thermal conductivity is K= 385 W m^{-1}K^{-1}. Please plot the final solution over the whole domain.

2. One curious feature about delta functions is that they allow discrete data to be given a functional representation. Such features are part of why delta functions are not true "functions". To expand on this comment, first we just need to realize that the area under a point is identically zero when "measured" by an integral. As an example, think about the function $f(x), x \in [0,1]$ containing a single point, as follows

$$f_1(x) = \begin{cases} 2 & x = \frac{1}{2} \\ 0 & \text{otherwise} \end{cases}$$

   To integrate this, we might write

$$I_1(x) = \int_0^1 f_1(x)\, dx$$

   But, noting that $f(x)$ is zero at every point *except* at $x = a$, we must have the equivalent result

$$I_1(x) = \int_a^a f_1(x)\, dx = 0$$

   In other words, by a classical analysis, the area under a single point is zero. Now suppose we *allow* the delta representation, which you may recall is only technically defined when integrated. However, because of this understanding, we frequently write the delta (generalized) function as if it were a regular function. Thus, the function $f(x)$ can be written as follows

$$f_2(x) = 2\delta\left(x - \frac{1}{2}\right) \tag{5.39}$$

   So that upon integration we would find

$$I_2(x) = \int_0^1 \delta(x - \tfrac{1}{2})\, dx = 2$$

In other words, the discrete data point $(1/2, 2)$ has an integral interpretation if we consider the delta function.

含义在$f_1(x)$和$f_2(x)$之间有所不同。不过，狄拉克δ函数有一个有趣的用途：它能提供一种方法，以普通函数相同的形式表示离散数据。为明确说明这一点，我们来考虑以下离散数据集X，它由(x,y)数据对$X = \{(1,1),(2,2.1),(3,2.9),(4,4.2),(5,5.05)\}$给出。根据上述讨论，我们可以将这个数据集写成适合积分的形式，如下所示。为避免端点取值带来的问题，假设该函数的定义域为$x \in [0,6]$

$$g(x) = 1\delta(x-1) + 2.1\delta(x-2) + 2.9\delta(x-3) + 4.2\delta(x-4) + 5.05\delta(x-5) \tag{5.40}$$

With this definition, please try the following.

a. 使用常规求和后除以数据点的数量，计算X 的平均值。 b. 连续函数（例如$f(x)$，$x \in [0,L]$）的平均值通常表示为

$$A = \frac{1}{L}\int_0^1 f(x),\, dx$$

Where L measures the magnitude of the domain. One way of thinking about L is that it is also defined by an integral that measures distance

$$L = \int_0^L 1\, dx = x|_0^L = L$$

虽然这有些同义反复，但有助于建立以下对应关系。在处理狄拉克δ函数时，我们以不同的方式定义定义域的大小。例如，考虑 $g(x)$，$x \in [0,6]$

$$L_\delta = \int_{x=0}^{x=6} \delta(x-1) + \delta(x-2) + \delta(x-3) + \delta(x-4) + \delta(x-5)\, dx = 1+1+1+1+1 = 5$$

当涉及狄拉克δ函数时，我们希望定义域的测度为所涉及的狄拉克δ函数的数量。但这一点（至少从概念上讲）源于与连续距离的模拟物。考虑到这一点，使用适当的积分计算$g(x)$的平均值。

c. The *center of mass* of a discrete data set with N data points of the form (x_i, y_i), $i = 1 \ldots N$ is given by

$$\bar{y} = \frac{\sum_{i=1}^N x_i y_i}{\sum_{i=1}^N y_i}$$

Compute this value for the data set X. Then, illustrate how the same value can be computed using integral quantities of the function $g(x)$.

"Science is a partial differential equation. Religion is a boundary condition."

Alan Turing, genius developer of modern computing theory.

6

偏微分方程、守恒定律与本构方程导论

6.1 Introduction

Partial differential equations (PDEs) are the mathematical workhorses of continuum systems. The term *partial* differential equation indicates that the dependent variable depends on two or more independent variables. This is in contrast to an ordinary differential equation (Chapter 2), where there is always one dependent and one independent variable. For this chapter, we will use the symbol u to indicate the independent variable, and the conventional symbols x and t to represent the independent variables corresponding to one spatial dimensions and one time dimension. Examples of the three canonical types of PDEs (hyperbolic, parabolic, and elliptic) will be presented.

6.2 Terminology

- **Extensive variable.** An extensive variable (or property) is one that depends upon the size of the domain of interest. As an example, the mass of material in a domain is an extensive property of that domain.

- **强度变量**。 强度变量（或性质）是指其数值独立于研究域大小的变量。质量是研究域的一种广延 性质，而密度是一种强度 性质，因为它完全不依赖于定义域的总体大小。

- **Flux.** A flux is the amount of some extensive property passing through a unit area per unit time. Technically, a flux may be more generally defined in terms of vector quantities; under these circumstances the flux is defined as the amount of the extensive property passing through a unit area per unit time, where

the unit area is defined with respect to a specific unit vector indicating the direction normal to the surface; both normal and tangential fluxes can be defined in multiple dimensions. In one space dimension these additional distinctions are not material.

• **偏导数**。当函数定义中包含多个自变量时，就会定义偏导数。更具体地说，考虑一个关于一维空间变量x和一维时间变量t的函数。空间偏导数的形式与空间中的普通导数相同，其中函数的时间分量保持不变。

• **偏微分方程（PDE）**。包含多个自变量的微分方程。这些方程本身由函数与偏导数相乘后形成的代数组合构成。通常，此类方程源于对连续体系统的物理考量，尽管它们也可能以各种方式源自众多学科。对于此类方程，不一定每个自变量都必须出现偏导数。

• **偏微分方程的阶**。偏微分方程的阶由方程中出现的最高阶导数决定。

• **瞬态偏微分方程**。假设我们有一个包含一个空间变量和一个时间变量的偏微分方程。瞬态偏微分方程是指初始条件会随着时间变化，在空间维度上演化为新的状态。

- **Steady state conditions.** The steady state for a PDE is the condition where the time partial derivative is identically zero for all other combinations of the remaining independent variables. For example, in one space and one time dimension with domain $x \in [a,b]$, the steady state is defined by $\partial u(x,t)/\partial t = 0$ for all points in $[a,b]$; note that if this occurs at some finite time, t_1 then it must also be true that $\partial u(x,t)/\partial t = 0$ for all $t > t_1$. In other words, once the system achieves steady state, it remains in that state (unless the system is perturbed by external forces).

- **Ancillary conditions.** Like for ODEs, ancillary conditions are the additional information required to determine a *particular* solution from the set of all possible solutions to a PDE. Unlike ODEs, because more than one independent variable is involved, the ancillary conditions technically specify *whole functions* rather than just constants (although constant functions are very common!)

• **抛物型偏微分方程**。在一个时间维度和一个空间维度中，抛物型偏微分方程的特点是具有一个随时间推移而光滑演化的初始配置。抛物型偏微分方程的经典例子是热量方程或扩散方程。对于该方程，初始分布会随时间推移而扩散，逐渐变得更加光滑和均匀。

- **Hyperbolic PDE.** In one time and one space dimension, a hyperbolic PDE is one that represents wave-like behavior. An initial perturbation is translated through the domain either without changing shape, or with shape changes (which can result either from self-sharpening behavior in nonlinear wave equations, to dissipating behavior in equations that represent energy loss from waveforms as they translate). The first-order wave equation is the primary hyperbolic example presented in this chapter.

- **Elliptic PDE.** There is no exact analogue for an elliptic PDE in one time and one space dimension. However, in one space dimension alone, and elliptic PDE can be considered to be the *steady state* version of a corresponding parabolic PDE. As an example, a steady-state diffusion-reaction problem can generate a solution in which the solution is not uniform in space, but at each spatial point the solution does not change in time. Physically, this would indicate that the processes of diffusion and reaction balanced one another.

- **适定偏微分方程问题**。在偏微分方程的语境中，适定问题意味着以下几点

  1. 解存在。换言之，辅助条件提供了足够的信息，使问题可以得到求解。2. 解唯一。该要求表明，不仅必须找到解，还必须有足够的信息来确定问题的特定解。换一种说法，解不应代表一类函数，而应是一个具体函数，给定所有相关参数后即可计算得出。3. 解必须连续依赖于初始数据。这一概念不如前两点直观，但简言之，它意味着问题的解不能表现出混沌行为。这里，表现得不混沌意味着初始条件的微小变化（或扰动）不会导致结果发生巨大差异。这是适定线性方程的一个特征：系统受到微小扰动，只会使解产生微小变化。

6.3 偏导数与偏微分方程的类型

Many processes in science and engineering are governed by partial differential equations. Essentially, most systems that can be approximated as being a *continuum* and where one needs to understand the time-space behavior of the system (as opposed to the integrated qualities of the system as appears in material balances) are best described by partial differential equations. Before continuing with examples, however, it is useful to recall the definition of a *partial derivative*.

6.3.1 Partial derivatives

偏导数描述的是一个关于多个自变量的函数所对应曲线的斜率。例如，假设函数 u 取决于空间变量 x 和 y。那么，沿 x–和y–方向的导数分别为

$$\lim_{\Delta x\to 0}\frac{\partial u(x,y)}{\partial x}\equiv\frac{[u(x+\Delta x,y)-u(x,y)]}{\Delta x} \tag{6.1}$$

$$\lim_{\Delta y\to 0}\frac{\partial u(x,y)}{\partial y}\equiv\frac{[u(x,y+\Delta y)-u(x,y)]}{\Delta y} \tag{6.2}$$

$$\tag{6.3}$$

Here it is understood that Δx is a positive value. Similarly, the second derivative with respect to x and y can be specified by

$$\lim_{\Delta x\to 0}\frac{\partial^2 u}{\partial x^2}\equiv\frac{[\frac{\partial u}{\partial x}(x+\Delta x,y)-u(x,y)]}{\Delta x} \tag{6.4}$$

$$\lim_{\Delta y\to 0}\frac{\partial u(x,y)}{\partial y}\equiv\frac{[u(x,y+\Delta y)-u(x,y)]}{\Delta y} \tag{6.5}$$

这些表达式可以反复迭代，以推导出更高阶的导数。最后补充一点，在偏微分方程中表示偏导数有几种常用方法。另一种传统方法是使用下标表示求导所针对的自变量。下标重复n 次，表示n^{th} 阶导数。因此，以下各式等价：

$$\frac{\partial u}{\partial x}\ \Leftrightarrow\ u_x \qquad\qquad \frac{\partial u}{\partial t}\ \Leftrightarrow\ u_t \tag{6.6}$$

$$\frac{\partial^2 u}{\partial x^2}\ \Leftrightarrow\ u_{xx} \qquad\qquad \frac{\partial^2 u}{\partial t^2}\ \Leftrightarrow\ u_{tt} \tag{6.7}$$

6.3.2 线性偏微分方程类型的判定

偏微分方程主要可根据三个主要特征进行分类。

1. **方程阶数**。偏微分方程的特征之一是其最高阶导数（对任一自变量求导）的阶数。这称为偏微分方程的阶。科学与工程中遇到的绝大多数偏微分方程都是一阶或二阶方程，不过有时也会遇到四阶方程。四阶以上的方程很少见，但偶尔也会用于描述某些物理过程。所有含两个自变量的一阶方程都描述波状行为；这些方程列于表 6.1。表 6.2 给出了一些二阶方程的实例。表 6.3 给出了相关的二阶以上线性方程实例。

2. **Equation type.** There are several well-studied behaviors seen in linear partial differential equations. In particular, for linear second-order partial differential equations *in two independent variables* of the form

$$a\frac{\partial^2 u}{\partial t^2}+2b\frac{\partial^2 u}{\partial t\partial x}+c\frac{\partial^2 u}{\partial x^2}+d\frac{\partial u}{\partial x}+ed\frac{\partial u}{\partial t}+fu+g=0 \tag{6.8}$$

every such can be described as exhibiting three types of behavior. Note that more than one type of behavior may be exhibited by the same equation if the coefficients are functions; that is, the equation type can change as a function of time or space. For equations with constant coefficients, only one equation type is defined, and it is does not depend on any of the independent variables.

With the appropriate change of variables, every equation of the form of Eq. (6.8) can be put in its *canonical form*. While determining the equation type is a useful exercise, it does not necessarily tell us all that we need to know about a particular problem. For example, the fact that the general equation

given by Eq. (6.8) can always be put into its canonical form by a coordinate transformation completely ignores the fact that if such an transformation may make any associated boundary and initial conditions so complex as to render the problem nearly unsolvable. Thus, the idea of the canonical form must be taken with the proverbial grain of salt. While it is useful information, in more practical problems this information may be of minor importance.

a. **Parabolic.** The canonical form for the parabolic equation with one time and one space variable is

$$\boxed{\frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2} = 0}$$

These equations have generally smooth, dissipative solutions. The heat and diffusion conservation equations are examples.

b. **Hyperbolic.** The canonical form for the parabolic equation with one time and one space variable is

$$\boxed{\frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial x^2} = 0}$$

Hyperbolic equations are characterized by representing wave-like behavior. These waves can be smooth and periodic (like ocean waves), or sharp, like the wave front of a compressible gas that creates a sonic boom.

c. **Elliptic.** The canonical form for the parabolic equation with two space variables is

$$\boxed{\frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial x^2} = 0}$$

While mathematically, there can be time-dependent elliptic equations, most physical applications come from steady-state descriptions of problems that are determined by their boundary conditions. The steady, 2-dimensional distribution of heat in a plate would be an example. Steady groundwater flow is also an example of an elliptic equation.

There are some problems with formal classification systems for PDEs that prevent them from being universally useful. For example, there is a classification system for all linear second-order PDEs with two (or more) independent variables. While such classifications can provide some insight, they still have weaknesses. For example, in these classification systems, the role of derivatives less than order 2 are not considered at all. This can create some difficulties, especially when these first-order terms dominate the problem. As an example, the classical convection-diffusion equation in one spatial dimensions can be given by

$$\frac{\partial u}{\partial t} = -v_0 \frac{\partial u}{\partial x} + D \frac{\partial^2 u}{\partial x^2}$$

While the traditional classification schemes presented in most texts on partial differential equations would qualify this as a *parabolic* equation, this is not an entirely satisfactory answer. In particular, when the convective term (the first-order term) is much larger than the diffusive term (the second-order term), then this equation behaves as a *hyperbolic* equation. This is of some significance, because, for example, numerical methods for solving parabolic (smooth) and hyperbolic (sometimes abrupt) equations can be quite different. As a second example, consider the following elliptic equaiton in two spatial variables

$$K_x \frac{\partial^2 u}{\partial x^2} + K_y \frac{\partial^2 u}{\partial y^2} = 0$$

while this is a perfectly acceptable elliptic PDE, we encounter a problem in the limit $K_y \to 0$. In that limit, our equation is suddenly no longer elliptic by the conventional criterion. While it is true that it can be then considered to be an ordinary differential equation (since there is only one independent variable), it does raise some problems in categorization.

因此，本文虽然会使用抛物型、双曲型和椭圆型这些术语进行描述，但不会特别致力于介绍通过坐标变换来形式化表征线性偏微分方程的方法。对于高阶偏微分方程，严格来说并没有通用的分类体系；不过，如果高阶方程的解具有与二阶方程类似的行为特征，偶尔也会使用上述某一种分类（抛物型、双曲型、椭圆型）。

3. **线性或非线性**。线性方程描述了工程学和科学中的许多现象，通常是对更一般的非线性方程行为的近似。偏微分方程的线性可采用与确定常微分方程线性相同的方法来判定。对于某些现象（例如纳维–斯托克斯方程中的湍流），非线性是所描述现象不可或缺的部分。表6.4列出了一些非线性问题的实例。尽管非线性问题既相关又重要，但其解通常需要专门的方法，超出了本入门教材的范围。

Table 6.1: Examples of first-order linear PDEs with two independent variables. All first order PDEs represent waves of one form or another.

Type	自变量 in 2 Dimensions	Name	示例应用	方程示例
One-dimensional waves	空间，x time, t	对流方程	溶质输运	$\frac{\partial \mathbf{u}}{\partial x} = -v_0 \frac{\partial u}{\partial x}, \quad v_0 > 0$
一维波	space, x time, t	The Maxwell equations	Electrodynamics	$\frac{\partial E}{\partial x} = \frac{1}{c}\frac{\partial B}{\partial t}$ $\frac{\partial B}{\partial x} = \frac{1}{c}\frac{\partial E}{\partial t} + J$

6.4 The Origins of Partial Differential Equations in Science and Engineering: The Axioms of Conservation

There is a nearly inexhaustible supply of applications of PDEs in engineering and science. Essentially, any systems that a subject to the laws of continuum mechanics are usually expressed as partial differential equa-

Table 6.2: Examples of second-order linear PDEs with two independent variables

Type	自变量 in 2 Dimensions	Name	示例应用	方程示例
抛物型	空间, x time, t	热方程	Heat or mass transport	$\frac{\partial u}{\partial t} = K\frac{\partial^2 u}{\partial x^2}$
椭圆型	空间, x 空间, y	The Laplace equation	Steady groundwater flow	$K\frac{\partial^2 u}{\partial x^2} + K\frac{\partial^2 u}{\partial y^2} = 0$
双曲型	空间, x time, t	波动方程	深层运动 水波	$\frac{\partial^2 u}{\partial t^2} = K\frac{\partial^2 u}{\partial x^2}$
（无）	空间, x 时间, t	The linear Korteweg–de Vries equation	浅层运动 水波	$\frac{\partial u}{\partial t} = K\frac{\partial^3 u}{\partial x^3}$
（无）	空间, x 时间, t	The vibration equation	Vibrations in beams or thin plates	$\frac{\partial^2 u}{\partial t^2} = K\frac{\partial^4 u}{\partial x^4}$

Table 6.3: Examples of a few linear, higher-order PDEs with two independent variables

Name	自变量 in 2 Dimensions	应用	方程
The linear Korteweg–de Vries equation	空间, x 时间, t	浅层运动 水波	$\frac{\partial u}{\partial t} = K\frac{\partial^3 u}{\partial x^3}$
振动方程	空间, x 时间, t	Vibrations in beams 或薄板	$\frac{\partial^2 u}{\partial t^2} = K\frac{\partial^4 u}{\partial x^4}$

Table 6.4: Examples of some important nonlinear PDEs in two independent variables.

Name	自变量 in 2 Dimensions	应用	方程
非线性扩散	空间, x time, t	Solute transport at high concentrations	$\frac{\partial u}{\partial t} = D(u)\frac{\partial^2 u}{\partial x^2}$
Burgers' equation	space, x time, t	一维类比 to Navier-Stokes	$\frac{\partial u}{\partial t} = u\frac{\partial u}{\partial x} + \nu\frac{\partial^2 u}{\partial x^2}$
The nonlinear Korteweg–de Vries equation	空间, x 时间, t	shallow water waves with nonlinear restoring function	$\frac{\partial u}{\partial t} = \frac{\partial^3 u}{\partial x^3} - 6u\frac{\partial u}{\partial x}$

方程。通常，这些偏微分方程是通过考虑守恒原理得出的。守恒原理是公理化陈述（或定律），无法从更基本的思想推导出来，但与现有的所有其他信息和测量值一致。连续介质力学中关于守恒原理的主要公理化陈述可列举如下。

1. **质量守恒**。这体现了一个近似的理念：物质既不能被创造，也不能被毁灭。实际上，有些物质可以转化为能量（反之亦然），但这在连续介质力学中并不常见，因此可以假定质量守恒成立。流体力学中遇到的连续性方程，就是质量守恒的一种表达形式。

2. **动量守恒**。在连续体系统中，动量守恒与离散物体所遵循的牛顿定律类似。简言之，连续体系统的动量守恒指出，牛顿定律不仅适用于离散物体，也适用于人们可以想象从离散物体中截取出的任何部分。描述流体流动（纳维–斯托克斯方程）或受压力作用的固体容器中力分布的重要方程，正是由动量守恒得出的。严格来说，动量守恒公理包含两部分：线动量守恒和角动量守恒。本书的大部分内容将只关注线动量问题。

3. **能量守恒**。 能量守恒是质量守恒的一条补充公理，通常表述为能量既不能被创造，也不能被毁灭。然而，正如上文所述，我们知道，在某些极端条件下（例如物体运动速度极快、尺寸极小，或涉及量子力学相互作用时），质量和能量可以相互转换。一般而言，连续介质力学并不考虑此类系统，因此可以放心地将能量守恒视为一个真实可靠的公理化陈述。

作为一项相关说明，上述每个量都代表一个热力学广延量；也就是说，计算得到的质量、动量或能量的数值取决于系统的物理尺寸。相反，也可以考虑这三个量中任意一个单位体积所对应的数值（即质量密度、动量密度或能量密度），这样得到的密度就是强度量。在连续介质力学中，大多数情况下，最好用密度而非广延量来讨论因变量。为了便于后续使用，我们将用小写字母（特别是字母u）表示强度变量，用大写字母（例如U）表示广延变量。

通常，偏微分方程的建立源于将这些守恒律应用于可视为连续介质的系统。连续介质是指充满某种材料的任意空间区域，在该区域内，可以假设这种材料的性质从一点到另一点平滑变化。我们知道，从根本上说，这种假设并不可能完全成立，因为所有物质都由原子组成，而在分子尺度上，原子显然并非连续的。然而，当我们将大量原子放在一起考虑时，通常可以将其行为很好地近似为一个连续系统。事实上，这正是连续介质力学这一领域背后的基本思想。在连续介质力学中，材料由代表性体积内的原子来表示，这些原子的数量足够多，因此其行为可以视为连续介质。将理想气体定律应用于惰性气体，就是一个关于代表性体积的恰当例子。在接近标准温度和压力的条件下，惰性气体往往相当严格地遵循理想气体定律。显然，气体由大量独立的原子或分子组成，但如果其数量足够多，那么它们的行为就可以视为能够用单一的整体介质来描述。因此，在接近地球表面温度和压力的条件下，只要惰性气体具有足够大的摩尔浓度，就可以用理想气体定律对其进行准确描述。

然而，如果能够（至少在概念上）在例如 1 升的体积中，将少量气体原子（比如 $n = 10$ 个原子）保持在室温下，那么理想气体定律将无法准确预测压力。

accurately. In fact, the very definition of pressure (force per unit volume) would be challenging for a such a small number of atoms. The problem is that the amount of momentum transfer between a small number of atoms and the walls of the container would be a rapidly varying function of time. Because we think about the molecules as having more-or-less random thermal motion (i.e., they have kinetic energy that is proportional to the temperature), with only 10 atoms in a 1 liter container, the atoms collide with the walls in a way that appears somewhat random.

为了让这一点更加清楚，假设我们固定一个时间区间 Δt，在此期间，平均有 5 个分子与容器壁发生碰撞。当然，由于涉及的分子很少，有时这个数量会更多（比如在 Δt 内有 7 或 8 个分子与容器壁碰撞），有时则会更少（比如在 Δt 内有 2 或 3 个分子与容器壁碰撞）。从这个角度来看，与容器壁发生碰撞的原子数是一个随机变量。因此，测得的压力可能会剧烈变化，从接近零（没有分子与容器壁碰撞时）到最大值（在时间区间 Δt 内，所有十个分子都与曲面发生碰撞时）。显然，这种情况下，压力会随时间变化，并且与在该时间区间内撞击曲面的分子数成正比。因此，理想气体定律只在平均意义上成立。若假设分子彼此独立地运动，那么与容器壁碰撞次数的方差将与涉及的分子数成反比，这与您可能在统计学中学过的大数定律一致。因此，只要增加分子数，就总能降低示例中气体行为的方差。分子数非常大时，方差会迅速变得极小。在这些条件下，可以将气体视为连续介质，理想气体定律也会以极高的保真度成立。

6.5 微分守恒方程的推导：连续介质力学导论

假设一个系统可以视为连续介质，那么就可以根据所关注量的微分衡算方程建立守恒方程。本节将着重推导适用于质量、线动量或能量的一般平衡方程，而不论具体关注的是哪种量。这些量都遵循同一条基本守恒公理。根据上述讨论，这条公理可表述如下。

公理 1（公理：质量、动量或能量守恒——形式 1）。 对于任何孤立系统，系统内所含的质量、动量或能量总量为常数。

这里需要定义“孤立”一词。理想系统是指能够通过完美边界与宇宙其余部分隔离的系统，这些边界可以阻止质量、动量和能量的传输。因此，当用于描述系统时，孤立与理想是同义词。虽然在许多情况下，这是一种合理的建模近似，但应注意，现实中不存在理想系统。

While the axiom of conservation of mass, momentum, and energy is probably familiar as presented, it is not necessarily the most useful form of the axiomatic statement. The problem is that we are often interested in sub-volumes of an isolated system. Fortunately, there are other statements of the axiom of conservation of mass, momentum, and energy that are equivalent to the version given above, but are more useful for applications to sub-volumes of isolated systems.

公理2（质量、动量或能量守恒公理——形式2）。 对于从孤立系统中截取的任意子体积，均应按照质量守恒定律维持以下平衡，

momentum, and energy. For any volume, V cut out of a continuum, the following balance law is valid for the extensive variable U (where U represents the mass, momentum, or energy in the volume V)

$$\underbrace{\left\{\begin{array}{c}rate\ of\\ accumulation\\ of\ U\ in\ V\end{array}\right\}}_{①} = \underbrace{\left\{\begin{array}{c}rate\ of\ U;\\ entering\ the\\ 体积\ V\end{array}\right\}}_{②} - \underbrace{\left\{\begin{array}{c}rate\ of\ U\\ 留下的\\ 体积\ V\end{array}\right\}}_{③} + \underbrace{\left\{\begin{array}{c}rate\ that\ sources\ or\ reactions\\ 更改数量\\ 体积\ V\ 中\ U\ 的\end{array}\right\}}_{④} \tag{6.9}$$

这种守恒公理的形式与第一种形式等价，但它更适合用于推导微分衡算。该公理为我们提供了建立微分衡算的直接方法；我们只需将文字给出的表述转换为等价的数学语言表述。

6.5.1 守恒公理文字表述的推导

首先，如图 6.1 所示，我们将孤立系统的一个子体积定义如下；请注意，在 $\Delta x \to 0$趋于极限时，该体积将成为微分体积。在本节对衡算计算的介绍中，我们将施加一些约束条件，使问题可以视为只沿一个空间方向（x–方向）变化。因此，我们作出以下相当强的假设。

一维过程几何的假设。

所有关注的强度性质仅沿x–方向变化；因此，关注的强度性质在垂直于x–方向的任意横截面积上均为常数。对于一维系统，我们还要求横截面积A为常数。尽管所有真实系统都存在于三个空间维度中，但其中涉及的过程 通常可以理想化为涉及较少维度。为了更具体地说明这一点，假设图 6.1 所示的圆柱体是一根绝缘导线，且我们关注系统中的热传输。如果导线两端的边界条件均匀，那么我们可以通过合理的对称性论证，认为热量主要沿轴向传输（即沿x轴方向）。对于这类系统，我们可以根据假设忽略 热量沿y轴和z轴方向的传输（即在垂直于x轴的平面内传输）。因此，尽管系统的物理几何 是三维的，但就实际应用而言，该系统热传输的过程几何 是一维的。

To start defining the terms for the mathematical statement of the balance, we note the following

1. The intensive variable of interest (mass density (concentration), momentum density, or energy density) is denoted by the dependent variable $u(x,t)$. The variable u depends only on the independent variables x and t.
2. The *flux*, $j(x,t)$, is defined as the amount of mass, momentum, or energy passing through a unit area in a unit time. Thus, the fluxes described here must have units of mass per unit area per unit time, momentum per unit area per unit time, or energy per unit area per unit time. Because the property of interest has

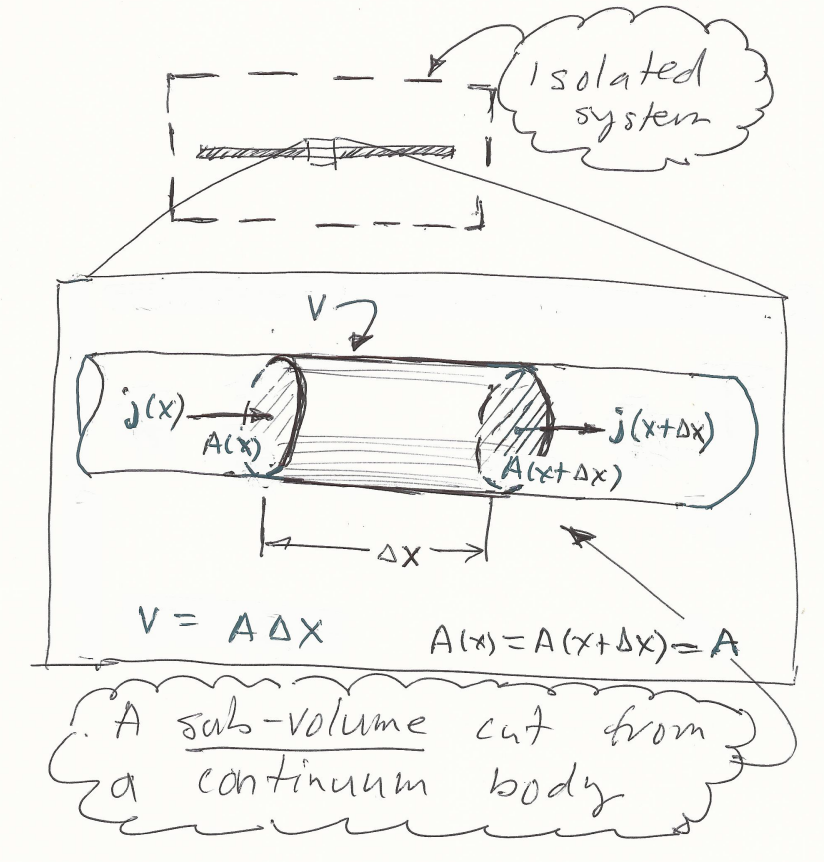


Fig. 6.1 A representative volume extracted from a continuum body. For our purposes, we will assume various symmetries that allow us to consider only one spatial dimension. In particular, the system shown has a constant cross-sectional area, A, and variables of the system change only in the $x-$direction (thus, they are constant on every cross-sectional area.)

a 1-dimensional process geometry, then there are non-zero fluxes only in the x-direction. The general expression above can be re-written for the 1-dimensional process geometry as follows

$$\left\{ \begin{matrix} rate\ of \\ U \\ entering\ V \end{matrix} \right\} = \int_{y,z \in A(x)} j_x(x,t)\, dy dz \tag{6.10}$$

$$= A j_x(x) \tag{6.11}$$

一个类似的表达式可以很容易地推导出来，用于表示离开体积 U 的速率。

$$\left\{ \begin{matrix} rate\ of \\ U \\ leaving\ V \end{matrix} \right\} = \int_{y,z \in A(x+\Delta x)} j_x(x+\Delta x,t)\, dy dz \tag{6.12}$$

$$= A j_x(x+\Delta x) \tag{6.13}$$

3. 体积V 内广延性质的总量由其密度在该体积上的积分给出。密度是质量的强度 对应量。因此，对控制体积V 内的密度进行积分，即可得到该控制体积内的质量。换句话说

$$\left\{\begin{array}{c} total \\ amount\ of \\ U \\ in\ V \end{array}\right\} = \int_{x,y,z\in V(x)} u(x,y,z,t)dx\,dy\,dz$$
$$= A\int_{x}^{x+\Delta x} u(x,t)\,dx \tag{6.14}$$

We are slightly abusing notation here; the variable of integration within the integral should, for absolute clarity, be different than the variable used for the bounds. However, for this development, the abuse is not likely to confuse and is likely to improve understanding, so we will not stand on propriety and instead opt for clarity.

With these definitions in place, we are in a position to evaluate each of the terms denoted ① – ④ in Eq. (6.9). Each term is taken in sequence in the material following.

Item ① : The accumulation term.

The accumulation term is defined as the "rate of accumulation of U in V." Given that we have an expression for the *total* amount of U in V, rate can be found by taking the time derivative. Note that because we have *two independent variables* (x and t), the derivatives must be defined as *partial derivatives*. We have as a result

$$\left\{\begin{array}{c} rate\ of \\ accumulation \\ of\ U\ in\ V \end{array}\right\} = A\frac{\partial}{\partial t}\int_{x}^{x+\Delta x} u(x,t)\,dx \tag{6.15}$$

请注意，尽管我们的文字形式涉及 U 在V中的总量，但在数学表征中，这已转换为u 在其中的强度量。根据上述步骤，应该可以清楚地看出其原因!

第 **2** 项： U 进入 V的速率

The groundwork done above makes this somewhat simple to evaluate. Recalling that we have a 1-dimensional process geometry, the rate of flux entering the volume V is given by

$$\left\{\begin{array}{c} rate\ of\ \mathrm{U}; \\ entering\ the \\ volume\ \mathrm{V} \end{array}\right\} = A j_x(x) \tag{6.16}$$

Item ③ : The rate of U leaving V

同样，我们可以得出，质量离开体积的速率V为

$$\left\{\begin{array}{c} rate\ of\ \mathrm{U}; \\ entering\ the \\ volume\ \mathrm{V} \end{array}\right\} = A j_x(x+\Delta x) \tag{6.17}$$

Item ④ : Source or Sink Term in V

Finally, we consider the amount of internal reaction or other source or sink within the volume to be given by a function s. Recalling that we have assumed that the process geometry is 1-dimensional, then this means that the source/sink term can depend only on x; the total amount of reaction or source or sink is then given by

$$\begin{aligned} \left\{\begin{array}{c} rate\ that\ sources\ or\ reactions \\ change\ the\ amount \\ of\ u\ in\ \ the\ volume\ V \end{array}\right\} &= \int_{x,y,z\in V(x)} s(x,y,z,t)dxdydz \\ &= A\int_x^{x+\Delta x} s(x,t)\,dx \end{aligned} \tag{6.18}$$

现在，我们可以考虑将上述文字表述转换为数学表述。将前面推导出的各文字量的数学表达式代入后，式（6.9）给出的公理化陈述即可用数学形式表示为。

$$A\frac{\partial}{\partial t}\int_x^{x+\Delta x} u(x,t)\,dx = A j(x) - A j(x+\Delta x) + A\int_x^{x+\Delta x} s(x,t)\,dx \tag{6.19}$$

Clearly, the constant A plays no role in this expression, so we can simplify this to

$$\frac{\partial}{\partial t}\int_x^{x+\Delta x} u(x,t)\,dx = j(x) - j(x+\Delta x) + \int_x^{x+\Delta x} s(x,t)\,dx \tag{6.20}$$

此时，这看起来还不太像一个微分平衡方程。不过，我们还没有使用一条重要的信息。由于我们将系统视为连续介质，因此可以让体积V任意小。更具体地说，对于一维表示，我们可以令 $\Delta x \to 0$。这将大大简化我们的表达式。首先注意，积分变量 x 满足 $x_0 < x < x+\Delta x$。因此，当 $\Delta x \to 0$ 时，我们还必须有 $|x - x_0| \to 0$。现在考虑函数 $u(x,t)$ 和 $s(x,t)$ 在x_0点附近的以下两个泰勒级数展开。这些展开式为

$$u(x,t) = u(x_0,t) + (x-x_0)\left.\frac{\partial u}{\partial x}\right|_{(x_0,t)} + \ldots \tag{6.21}$$

$$s(x,t) = s(x_0,t) + (x-x_0)\left.\frac{\partial s}{\partial x}\right|_{(x_0,t)} + \ldots \tag{6.22}$$

回顾一下，我们处于 $|x - x_0| \to 0$ 的情形。在第 1 章中，我们指出，解析函数的所有导数在其定义域内都必须保持有界。这意味着，我们总能找到 足够小的 Δ，使得上述每个泰勒级数中的第二项任意小。当 Δx 趋近于零（但不为零）时，泰勒级数给出如下近似

$$u(x,t) = [u(x_0,t) + \alpha_1 \Delta x] \tag{6.23}$$

$$s(x,t) = [s(x_0,t) + \alpha_2 \Delta x] \tag{6.24}$$

where the final term represents a measure of the error involved, with α_1 and α_2 being two finite constants. This means the two integrals can be simplified as follows

$$\frac{\partial}{\partial t}\int_x^{x+\Delta x} [u(x_0,t) + \alpha_1 \Delta x]\, d\xi = j(x,t) - j(x+\Delta x,t) + \int_x^{x+\Delta x} [s(x_0,t) + \alpha_2 \Delta x]\, d\xi$$

$$\frac{\partial}{\partial t}[u(x_0,t) + \alpha_1 \Delta x]\int_x^{x+\Delta x} d\xi = j(x,t) - j(x+\Delta x,t) + [s(x_0,t) + \alpha_2 \Delta x]\int_x^{x+\Delta x} d\xi$$

$$\frac{\partial}{\partial t}[u(x_0,t) + \alpha_1 \Delta x]\Delta x = j(x,t) - j(x+\Delta x,t) + s(x_0,t) + \alpha_2 \Delta x]\Delta x$$

$$\frac{\partial}{\partial t}[u(x_0,t) + \alpha_1 \Delta x] = -\frac{[j(x+\Delta x,t) - j(x,t)]}{\Delta x} + [s(x_0,t) + \alpha_2 \Delta x] \tag{6.25}$$

Note in the final line of this expression, we have rearranged the two flux terms so that the resulting expression is in the conventional form for the definition of the derivative. Now, if we take the limit as $\Delta x \to 0$, note that $u(x_0,t) \to u(x,t)$. We find the result known as the general conservation equation

$$\boxed{\frac{\partial u(x,t)}{\partial t} = -\frac{\partial j(x,t)}{\partial x} + s(x,t)} \tag{6.26}$$

6.6 Constitutive Equations and Flux Laws

It is not an understatement to say that that Eq. (6.26) is the most important equation in continuum mechanics. In one very simple equation, a relatively universal statement about the conservation of mass, momentum, and energy is expressed. Plus the statement is mathematically interpretable in terms of physical intuition:

accumulation in a control volume is equal to the difference between mass in and mass out (the flux term) plus any generation or reaction within the volume.

守恒方程并不会用因变量u来表示u 的守恒。相反，要解守恒方程，我们需要同时知道通量j和源项s。源项通常比较简单：它们要么由一些较为熟悉的项给出，代表动力学反应（或能量、动量的等效过程），要么是非齐次项（使方程本身成为非齐次方程）。然而，通量项则有所不同。根据问题的物理学原理，我们知道u 的通量应当以某种方式涉及变量u 本身。将通量j 与因变量u 的函数联系起来的表达式称为本构方程。它们在非平衡系统热力学和连续介质力学中都发挥着核心作用（事实上，这两个领域有很大的重叠）。本构方程理论中有一些深刻的结果，而且这一领域仍然是活跃的研究方向。下文将从连续体系统数学建模的角度，列出一些较为重要的本构关系；但除了少数旁注外，我们不会过多探讨它们的由来。许多本构方程最初是通过经验确定的（在发现时，由于它们在其他方面属于公理化陈述，往往被称为定律），但其中许多方程后来通过统计力学的表述获得了更为基础的解释。

It can be useful to think of PDEs as having characteristics of the three fundamental "types" that we have discussed (parabolic, hyperbolic, and elliptic). In the material that follows, we will combine the general conservation equation derived above (Eq. 6.26) with various flux laws to obtain representative equations of each of these three types. As an example of a *hyperbolic equation*, we will examine the first-order "pure convection" equation; such an equation would describe, for example, how a solute plume might move along a river. For *parabolic equations*, we will illustrate examples where fluxes are given by gradient laws. Representatives illustrating the conservation of mass, momentum, and energy will be presented. Parabolic PDEs are one of the most useful representations for describing the continuum mechanical behavior of quantities that are specified by gradient flux laws. Finally, for *elliptic equations* we will examine a reaction-diffusion problem at steady state.

There is an important note to be made at this juncture. The second-order wave equation (also a hyperbolic equation) is not one that arises naturally in the context of combining conservation and flux laws. While it arises from considerations of *conservation of momentum*, its derivation has some key differences, the foremost of these being that there is no simple flux law for the momentum in that case. Instead, we will adopt the following approach. The *first-order wave equation* (a hyperbolic equation) is one that does, in fact, arise from the combination of Eq. 6.26 and a flux law; in this chapter we focus on deriving it. A derivation of the second-order wave equation, using the more-or-less traditional approach, is given in the Appendix to this chapter. Later, during the study of the Fourier transform, we will have some new tools that allow us to re-visit this problem. In particular, we will be able to make a link between the first-order and second-order wave equations.

对于下文出现的每个实例，目前我们将讨论范围限制在一维空间以及表示时间的一维。空间方面的限制将在本文后续内容中放宽，届时我们将考虑多维空间中的解。

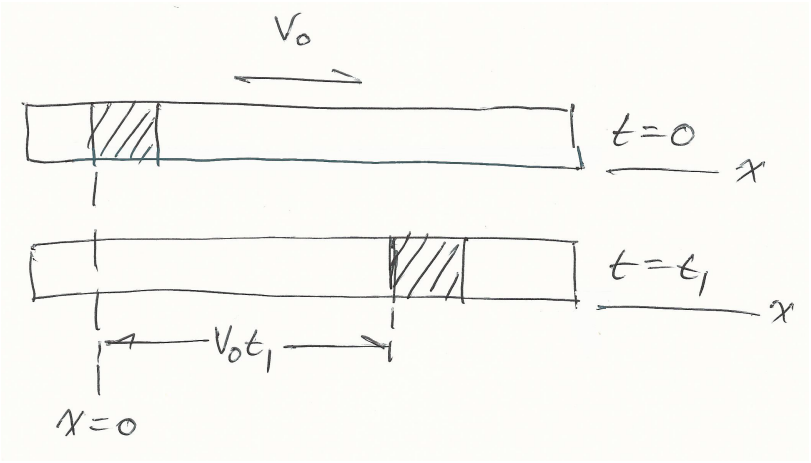


图 6.2 对流通量。均匀速度分布的理想平推流反应器中纯对流通量的作用。

6.6.1 通过对流通量描述的一阶双曲型偏微分方程

The most fundamental kind of flux is referred to by either the term *convective flux* or *advective flux*. There is some disagreement in the literature (and, occasionally, even debates) about the terminology, usually with strong statements that one or the other of these two terms is the "correct" terminology followed by some historical fact presented as proof. Regardless of these details, meaning of language is created via usage, and both terms are used to describe the same phenomenon widely enough that one should not attempt to distinguish between the two words unless the author has carefully defined the meaning. From here forward, only the term *convective flux* will be used.

对流通量是指借助整个介质的运动来传输强度量 u 的通量。举一个具体的例子：如果在河面上放出 1,000 只橡皮鸭，这些鸭子的运动就可以视为鸭子的对流通量（假设按照横截面积的适当定义进行了归一化）。另一个熟悉的例子是理想平推流反应器中化学物种脉冲的运动，如图 6.2 所示。其原理是，u 的运动遵循其所处介质的主体速度；显然，如果该介质是流体（气体或液体），这种说法最为合理。不过，有些物质并不明显属于其中任何一种，或是不同类型物质的混合物，但仍会表现出整体运动。岩浆、沙流和等离子体都是能够呈现类似液体的对流通量的物质。用数学关系表示，对流通量为

$$j_c(x,t) = u(x,t)v(x,t) \tag{6.27}$$

where u is the dependent variable (concentration, momentum per unit volume, or specific energy), and v is the velocity field that describes the motion of the material in which u is embedded.

Example 6.1 (Mass conservation: The continuity equation. A hyperbolic equation.).

A simple and important example of convective fluxes combined with the conservation equation is the *continuity0* equation. For now, assume that we interpret u as the density of a gas, $u \Leftrightarrow \rho$ in a steady flow field in one dimension, with no source or sink terms ($s = 0$). This system is composed of only

one chemical species – the fluid. The convective flux is then $j_c(x,t) = u(x,t)v(x)$, and the conservation equation gives

$$\frac{\partial u(x,t)}{\partial t} = -\frac{\partial (uv)}{\partial x} \tag{6.28}$$

例 *6.2*（质量守恒：一阶波动方程。双曲型方程。）

一阶波动方程会出现在一些条件下，例如图 6.2 所示的理想平推流反应器。对于这个例子，我们可以将u 视为稳态流动的均匀流场中，溶解在流体（液体或气体）里的某种物种的浓度。因此，该系统至少由两种化学物种组成：一种是流体，另一种是溶解在流体中的化学物种。假设不存在反应或其他源项，因此$s = 0$。通量由与上文相同的关系给出：$j_c(x,t) = u(x,t)v(x)$。守恒方程为

$$\frac{\partial u(x,t)}{\partial t} = -\frac{\partial (uv)}{\partial x} \tag{6.29}$$

对于许多液体，不可压缩的概念是适用的。尽管所有物质在一定程度上都可压缩，但通常可以很好地将液体近似为不可压缩。简而言之，这意味着进入控制体积的液体量不会超过流出该体积的液体量（即体积内不会有液体累积）。换一种说法，就是液体的密度为常数。由上面的连续性方程可立即得出，对于不可压缩流体，$\partial v/\partial x = 0$。这也意味着流体速度必须为常数，$v(x) = v_0$。在这些条件下，守恒方程可写成以下形式

$$\frac{\partial u(x,t)}{\partial t} = -v_0 \frac{\partial u}{\partial x} \tag{6.30}$$

Which is a first-order wave equation. Although we are not technically *solving* PDEs at this juncture, a little thought will indicate that in 1D, if the initial condition is simply being translated through space at a constant velocity, v_0. If the initial condition is given by the function $f(x)$, then the solution to the PDE is given by the translated version of the initial condition for time t, i.e.,

$$u(x,t) = f(x - v_0 t) \tag{6.31}$$

提醒一下，二阶双曲方程并非通常由守恒定律与通量定律结合而来。这并不是说这类方程不重要，只是它们的结构有别于本章讨论的其他方程（本章其他方程均由式（6.26）与通量定律结合而来）。稍后在讨论傅里叶变换时，我们会重新探讨二阶波动（双曲型）偏微分方程。届时，我们将能够展示一阶和二阶双曲方程之间一种有用的对应关系。

6.6.2 Second-Order (Parabolic) PDEs via Gradient Flux Laws

式 6.26 给出的守恒一般方程可以生成各种有趣的偏微分方程，用来描述许多重要的物理过程。这个守恒方程与通量定律之间存在一种有趣的对称性。前文中，我们看到对流通量定律 $j = u(x, t)v(x, t)$ 生成了一阶波动方程，例如，该方程描述了溶质在平推流反应器中的运动。本节将考察第二类通量定律——线性梯度定律。这类通量定律适用于在相应条件下描述质量、动量和能量的弥散传输。

A gradient law is any constitutive expression that describes the flux as being proportional to the spatial gradient of u. In one-dimension, the gradient is represented simply by the derivative, $\partial u/\partial x$. As mentioned above, gradient laws have a long history in the field of continuum mechanics. A few are given in the following.

$$\text{Fick's Law (mass):} \qquad j_g = -D\frac{\partial u}{\partial x} \qquad u = \text{concentration} \tag{6.32}$$

$$\text{Fourier's Law (energy):} \qquad j_g = -K\frac{\partial u}{\partial x} \qquad u = \text{temperature} \tag{6.33}$$

$$\text{Newton's Law of Viscosity (momentum):} \qquad j_g = -\mu\frac{\partial u}{\partial x} \qquad u = \text{velocity} \tag{6.34}$$

虽然梯度定律最初确实是通过实验确定的，但 1905 年，阿尔伯特·爱因斯坦展示了从后来被称为统计力学的理论中推导梯度定律的首批成果之一。爱因斯坦证明，如果可以将大分子视为由大量彼此不相互作用的粒子组成，这些粒子随机运动，并且每经过一个固定的时间间隔就改变方向，那么这种运动自然会导出菲克定律，作为对连续介质运动的近似。按照现代标准来看，他的推导或许略显粗略，但我们也必须记住，1905 年科学界尚未完全接受分子存在这一观点。爱因斯坦不仅从更基础的理论（统计力学）推导出了首批本构方程之一，他的工作也为物质的分子理论确实成立提供了最终证据之一。几年后（1908 年），杰出的法国实验物理学家让-巴蒂斯特·佩兰对爱因斯坦的理论进行了检验。佩兰的研究发表后，几乎再也没有任何有效论据能够有力反驳物质具有分子性质这一观点。

Example 6.3 (Mass conservation: Fick's second law. A parabolic equation.).

Sometimes Fick's gradient law combined with the conservation equation is called "Fick's second law". Given the information above, the result is straightforward to develop. Substituting the gradient law for mass for the flux, j, the conservation equation gives (again, assuming that $s = 0$, so there are no reactions or sources in this formulation)

$$\frac{\partial u}{\partial t} = -\frac{\partial}{\partial x}\left(-D\frac{\partial u}{\partial x}\right)$$

or

$$\frac{\partial u}{\partial t} = D\frac{\partial^2 u}{\partial x^2} \tag{6.35}$$

Clearly, the use of Fourier's law in the conservation equation for energy will yield analogous results. Thus, Eq. (6.35) is known both as the *diffusion* equation and the *heat* equation.

Example 6.4 (Momentum conservation: Burgers' Equation. Mixed parabolic and hyperbolic.).

Burgers' equation is a one-dimensional analogue to the well-known Navier-Stokes equations describing fluid motion. While Burgers' equation was initially derived as a tool for studying the Navier-Stokes equations (and, thus, one without particular physical relevance), it was later understood to also describe certain kinds of waves in fluids. The flux for Burgers' equation consists of the combination of a convective flux for momentum, and Newton's gradient law for viscosity. For conservation of momentum, the variable of interest is $q = \rho u$; in other words, it is the momentum per unit volume. Thus, the flux takes the form

$$j = q - \mu\frac{\partial q}{\partial x} \tag{6.36}$$

Substituting this result into the general conservation equation (and assuming that $s = 0$) yields

$$\frac{\partial q}{\partial t} = -\frac{\partial}{\partial x}\left(q - \mu\frac{\partial q}{\partial x}\right)$$

或者，在代入 $q = \rho u$ 并假设密度为常数后

$$\frac{\partial u}{\partial t} = -u\frac{\partial u}{\partial x} + \frac{\mu}{\rho}\frac{\partial^2 u}{\partial x^2} \tag{6.37}$$

This is the first example of a conservation equation that contains both *convective* and *gradient* fluxes.

Example 6.5 (Mass conservation: Convection-diffusion-reaction. Mixed parabolic and hyperbolic.).

The convection-diffusion-reaction equation is a familiar one science and engineering, and it has been used to describe everything from the motion of dissolved contaminants in groundwater to the spread of infectious diseases. Technically, systems that have convective motion experience *dispersion* which is the combination of spreading due to molecular diffusion plus spreading due to fluctuations (or velocity gradients) in the velocity field. Regardless, both diffusion and dispersion have the same mathematical form (assuming that neither is a strong function of space), and mathematically they are identical problems. The reaction term is expressed in the source or sink term, s, appearing in the conservation equation. This is the first example to contain a non-zero source term. For this effort,

consider u to be the concentration of a reacting chemical species, v_0 to be a spatially constant (uniform) one-dimensional velocity field, and the reaction rate to be first-order in u, that is $s = -ku$. The resulting expression is very similar to the previous result for Burgers' equation.

$$j = uv_0 - D\frac{\partial u}{\partial x} \tag{6.38}$$

$$s = -ku \tag{6.39}$$

Substituting this result into the general conservation equation yields

$$\frac{\partial u}{\partial t} = -\frac{\partial}{\partial x}\left(uv_0 - D\frac{\partial u}{\partial x}\right) - ku$$

or, upon using the fact that v_0 is a constant

$$\frac{\partial u}{\partial t} = -v_0\frac{\partial u}{\partial x} + D\frac{\partial^2 u}{\partial x^2} - ku \tag{6.40}$$

This is the first example of a conservation equation that contains both *convective* and *gradient* fluxes and a source term.

Example 6.6 (Mass conservation: The Groundwater Flow Equation in 1-dimension. Elliptic.).

Examples of *elliptic* problems with two spatial variables present the challenges noted above (specifically, the fluxes become vectors, which is a topic that will be covered in a separate chapter). Presenting examples of elliptic equations with one space dimension is also somewhat challenging. However, the following is a reasonable example. To start, recall the one-dimensional conservation equation on $x \in [0, L]$ of the form

$$\frac{\partial u(x,t)}{\partial t} = -\frac{\partial j(x,t)}{\partial x} + s(x,t)$$

In groundwater flow, Darcy's law is the gradient constitutive equation that relates hydraulic head, h to the rate of groundwater flow per unit area j. In one-dimension, the expression takes the form

$$j = -K\frac{\partial u}{\partial x}$$

where here u =hydraulic head in meters, K is the hydraulic conductivity in m/s. We are ignoring a amplification term known as the storativity for ease in the presentation. Substituting the flux into the general conservation equation yields

$$\frac{\partial u(x,t)}{\partial t} = K\frac{\partial^2 h(x,t)}{\partial x^2} + s(x,t)$$

While this result is technically a parabolic equation, we consider specifically the case where the groundwater flow is at steady state. Steady state is defined by the state in which, for all locations x in the domain, $\partial u/\partial t = 0$. Thus, the balance equation is given by

$$0 = K\frac{\partial^2 h(x)}{\partial x^2} + s(x)$$

严格来说，这个方程是椭圆型的。更常见的形式是，该方程包含两个空间变量，写作泊松方程。

$$K_x\frac{\partial^2 u}{\partial x^2} + K_y\frac{\partial^2 u}{\partial y^2} = -s(x)$$

如果在x –y定义域的边界上满足对称条件，我们可以设想这样一种条件：沿y方向的流量为零。在这些条件下，我们会得到一个与上面推导出的一维形式完全相同的方程（尽管该问题适用于二维定义域）。

希望此时你已经看出，守恒方程的简单推导确实是一种强有力的工具。结合通量的本构表达式以及源项的表征，可以推导出各种有意义的连续介质力学表达式。

6.7 辅助条件

确定偏微分方程需要多少个附加条件的问题，与常微分方程中类似问题非常相似。为了便于具体讨论，我们将继续讨论至多包含两个自变量的问题，其中一个是空间自变量，另一个是时间自变量。空间中包含两个自变量的情况则会带来一个复杂因素，即通量具有向量量的性质。因此，本章前面讨论过的椭圆型问题的更全面分析，将推迟到介绍多维守恒方程的章节中进行。

6.7.1 偏微分方程的阶数与确定必要的辅助条件

对于一维空间中的问题，确定所需辅助条件的数量相对直接。为清楚起见，我们考虑如下形式的扩散/热方程

$$\frac{\partial u}{\partial t} = \alpha^2\frac{\partial^2 u}{\partial x^2} \tag{6.41}$$

where α^2 is a parameter that is real and (necessarily) positive, and represents either the diffusion coefficient or the thermal conductivity. The way to think about these problem is to consider how many *integrations* are needed to eliminate the derivatives that appear. Here we are using the word “integrations” somewhat

informally (although, when we explore separation of variables, we will find that the solutions are indeed found by integration in the conventional sense that we have used the term in the study of ODEs).

In this example, we have the following derivatives to resolve

1. One time derivative of first order.
2. One space derivative of second order.

Thus, to resolve the problem, we must conduct one "integration" in time, and two "integrations" in space. This generates a set of three unknown constants, one from each integration. In reality, the unknowns may be whole *functions* because we are dealing with two dimensions; while the unknown values are fixed in one of the two dimensions, they remain potentially free in the other dimension. This will be more clear when we consider an example.

6.7.2 Types of Ancillary Conditions

关于辅助条件，还有一个细节需要注意。一般来说，我们应当将辅助条件视为一个条件层级，通常从指定函数的值开始，并且可能（取决于方程的阶）还包括指定函数的导数。换句话说，假设我们考虑上文所给扩散方程在空间中所需的两个辅助条件。在继续之前，我们将指出由一个空间变量和一个时间变量构成的二阶偏微分方程的三种基本辅助条件类型。我们假设这些方程由因变量u以及自变量x和t表示（因此，辅助条件所指定的函数为$u(x,t)$）。

1. **给定值条件**。在偏微分方程中，在解的时空边界上指定其值，是最基本的一类附加条件。如下文所述，要使问题适定，通常必须针对每个自变量，在边界的某一部分给定值条件。在偏微分方程的研究中，通常会对空间变量和时间变量区别对待（事实上，它们在物理意义上也截然不同——即使在一维空间中，我们也可以设想向前或向后移动，但时间却只能向前！）。因此，两种不同类型的给定值条件分别有各自的名称。

   a. **Specified value in the space dimension**. These conditions are called *Dirichlet* or *first-type* boundary conditions. The term "boundary" is used to emphasize that they are spatial conditions. Thus, these conditions specify the value of the dependent variable, u, at one or more spatial locations for all values of the independent variable, t, defining the time domain. This boundary condition is named after the German mathematician Johann Peter Gustav Lejeune Dirichlet (1805–1859).

   b. **Specified value in the time dimension**. When applied to the time dimension, the specified value condition is called an *initial condition*. Thus, and initial condition specifies the value of the dependent variable, u, at $t = 0$ (or some other time considered to be the start of the clock associated with the problem) for all values of the independent variable, x, defining the spatial domain.

2. **Specified derivative conditions.** Specified derivative conditions are, in a sense, "weaker" than specifying the value of the dependent variable. Thus, specifying the derivative of the dependent variable gives the problem more freedom to adapt to the requirements imposed by the PDE itself. In other words, one can meet a specified derivative condition at a boundary, but the value of the function at that boundary is otherwise unconstrained by the boundary condition itself. Because of this, some care has to be taken

在应用指定导数条件时，可以构造出仅施加指定导数条件却没有唯一解的问题。与给定值条件类似，在偏微分方程的研究中，惯例是使用不同的术语来描述这类空间条件和时间条件。

a. **在空间维度中指定导数**。 这些条件称为诺伊曼条件或第二类边界条件。这些条件规定了因变量u在一个或多个空间位置处、对定义时间域的自变量t的所有取值的导数。该条件以德国数学家卡尔·戈特弗里德·诺伊曼（1832–1925）的名字命名。

b. **时间维度上的指定导数**。 应用于时间维度时，指定导数条件仍称为初始条件；除此之外没有特殊原因，这只是约定俗成。因此，此初始条件指定了因变量u在$t = 0$时（或问题所关联时钟的其他起始时间）对于定义空间域的自变量x所有取值的导数。在更数学化的语境中，指定值加指定导数的初始条件有时也称为柯西条件，以纪念法国数学家奥古斯丁-路易·柯西男爵（1789–1857）。

3. **Linear combinations of specified value and specified derivative conditions.** There are instances where a linear combination of the two, more fundamental, types of ancillary conditions is useful and contains and often necessary physical content. These kinds of conditions will not be brought up in any detail for the remainder of this chapter, but an example of a useful linear combination of ancillary conditions as a model for a physical boundary condition will be discussed in the material related to *solving* PDEs. When applied to spatial boundaries, the linear combination condition is called a *Robin* condition after the French mathematician Victor Gustave Robin (1855–1897).
4. **Periodic conditions.** There are some applications where the boundary conditions are also required to behave periodically on the domain. This is not actually a separate boundary condition *per se*, so much as a constraint on the kinds of solutions that can be attained. The use of conditions to amend or constrain the boundary conditions is a topic that is more advanced than the material being covered in this text, so we will not discuss the in detail.

6.7.3 Applications of Ancillary Conditions with One Space and One Time Dimension

Let's return to the diffusion problem introduced above.

$$\frac{\partial u}{\partial t} = \alpha^2 \frac{\partial^2 u}{\partial x^2} \tag{6.42}$$

我们知道，总共需要三个辅助条件，以处理由积分方程产生的未确定信息（函数）。对于扩散问题，所涉及的积分必然要求指定两个边界 条件和一个初始 条件。

We could specify these conditions by any of the following options

选项 1（两个狄利克雷条件，一个初始条件）

$$u(0,t) = u_0(t) \qquad u(L,t) = u_L(t) \qquad u(x,0) = f(x)$$

Option 2 (one Dirichlet, one Neumann condition, one initial condition)

$$\frac{\partial u(0,t)}{\partial x} = g_0(t) \qquad u(L,t) = u_L(t) \qquad u(x,0) = f(x)$$

Option 3 (one Dirichlet, one Neumann condition, one initial condition)

$$u(0,t) = u_0(t) \qquad \frac{\partial u(L,t)}{\partial x} = g_L(t) \qquad u(x,0) = f(x)$$

In other words, we should generally work up through the sequence of functions in the order $u, u_x, u_{xx}, \cdots$ when specifying the ancillary conditions. Another way of stating this is that, for problems in one space and one time dimension, both space and time should have at least one Dirichlet (specified value) condition specified at some point in the problem.

Missing from the list above is the possible combination

Option 4 (two Neumann conditions)

$$\frac{\partial u(0,t)}{\partial x} = g_0(t) \qquad \frac{\partial u(L,t)}{\partial x} = g_L(t) \qquad u(x,0) = f(x)$$

The reason that this problem is not listed is that it requires special handling. This is best shown by using a direct example.

例 *6.7*（两个诺伊曼条件的问题。）

Consider the elliptic problem above for groundwater flow. Suppose we have the following PDE and two boundary conditions on the interval $x \in [0, L]$

$$\begin{aligned} K_x \frac{\partial^2 u}{\partial x^2} &= 0 \\ \frac{\partial u(0,t)}{\partial x} &= g_0 \\ \frac{\partial u(L,t)}{\partial x} &= g_L \end{aligned}$$

这个 pr oblem is actually an easy one to solve. It requires just two direct integrations of the first equation. An outline of the process is as follows

$$\frac{\partial}{\partial x}\left(\frac{\partial u}{\partial x}\right) = 0 \quad \text{Divide by } K_x \neq 0\text{; rewrite derivatives in equivalent form.}$$
$$\frac{\partial u}{\partial x} = C_1 \quad \text{Integrate once, as indefinite integration}$$
$$u(x) = C_1 x + C_2 \quad \text{Integrate a second time, as indefinite integration}$$

Now, note that the derivative of the solution implies

$$\frac{\partial u}{\partial x} = C_1$$

这表明导数处处为常数。因此，为了使原陈述有意义，我们必须有 $g_0 = g_L = G$，其中 G 为常数。换句话说，该解不允许定义域两端的通量彼此不同。此时的解是

$$u(x) = Gx + C_2$$

遗憾的是，我们已经没有足够的独立边界条件来求解C_2了！如果能够为该系统引入其他约束条件（通常通过考虑某种物理守恒原理），以弥补缺少第二个独立边界条件的问题，那么这种情况或许可以得到解决。例如，如果已知平均速度，即，

$$\frac{1}{L}\int_0^L u(x)\,dx = U_0$$

Then, this would be sufficient to replace the lack of information that we have by having a repeated Neumann condition. Specifically, note

$$\frac{1}{L}\int_0^L u(x)\,dx = U_0$$
$$\frac{1}{L}\int_0^L (Gx + C_2)\,dx = U_0$$
$$\frac{1}{L}\left(\frac{GL^2}{2} + C_2 L\right) = U_0$$
$$\implies C_2 = U_0 - \frac{GL}{2}$$

This is an indication that the application of two Neumann conditions must be considered very carefully in general to assure that the following are met:

1. The problem must make *physical* sense, including in the limits as time becomes arbitrarily large.
2. The problem must have sufficient content so that the problem is *well posed*. This just means that the solution exists, the solution is unique, and the solution does not exhibit "chaotic" behavior.

An problem combining all of the ideas in this section is given in the following example illustrating a set of boundary and initial conditions for a reaction-diffusion equation that is *well posed*.

例 *6.8*（一个适定的反应扩散方程。）。

A 个适定 问题是指满足以下条件的问题。

1. 该问题有解。这意味着偏微分方程本身的构造在内部是一致的（即在数学或物理上有意义），并且具有足够的辅助信息，因此可以求得解。

2. 该问题具有唯一性。这意味着，我们不仅能够找到解，还能够找到代表我们希望求解的特定条件的特解。这意味着，我们找到的并非一个解的等价类（其中每个解都是有效的，但任意两个解都不能表示为彼此的线性组合），而是一个由适当的辅助条件完全确定的单一解。

3. 该问题的行为不会呈现混沌性。通常，这被表述为要求问题的行为连续地依赖于初始数据，尽管这并非最直观的表述方式。简而言之，这一要求表明，如果对问题的辅助条件施加微小扰动，那么问题的解也只会发生微小变化。虽然可以从数学上对这一表述进行更加严格的定义，但对于我们的目的而言，理解其核心思想就足够了。

Consider the parabolic problem for a diffusion-reaction system on the interval $x \in [0, L]$.

$$\frac{\partial u}{\partial t} = D\frac{\partial^2 u}{\partial x^2} - ku$$

This problem has one spatial derivative of order 2, and one time derivative of order 1. Thus, we need two ancillary conditions for the spatial derivatives, and one for the time derivative. Now, consider the following statement

$$\begin{aligned}
& & \frac{\partial u}{\partial t} &= D\frac{\partial^2 u}{\partial x^2} - ku \\
B.C.\ 1 & & u(0,t) &= u_0 \\
B.C.\ 2 & & \frac{\partial u(L,t)}{\partial x} &= 0 \\
I.C.\ 1 & & u(x,0) &= f(x)
\end{aligned}$$

根据上述准则，边界条件包括一个狄利克雷条件和另一个条件（本例中为诺伊曼条件；如果第二个狄利克雷条件符合问题的物理背景，也同样可以接受）。因此，空间辅助条件似乎满足我们的要求。还要注意，辅助条件严格来说是函数。尽管它们是常函数，但两个边界条件都为每一个可能的值 $u(0,\ t)$ 或 $\partial u(L,\ t)/\partial x$ 指定了相应边界上的一个特定值；因此，这两个表达式

表示时间函数。稍加思考就能明白为何必然如此；如果这些函数没有为所有相关时刻定义，那么就会存在某些时刻缺少适当辅助条件，问题也就不再适定。关于初始条件也可以作类似的说明。初始条件指出系统在$t = 0$时的状态（原则上，也可以是指定为问题“起点”的任意时刻，t_0,）。当然，这个初始条件必须说明定义域内每个位置x发生的情况，否则就会有某些位置缺少辅助空间数据，从而无法求出随时间变化的解（问题同样会变得不适定）。

6.8 获得偏微分方程解的变换

偏微分方程的变换是这样一种思想的应用：“如果你有一个不知道如何解决的问题，就试着把它转换成一个你知道如何解决的问题。”尽管，原则上这是一个很好的想法，但在偏微分方程中普遍使用变换所面临的问题之一是，很难制定明确的规则来找出这样的变换，尤其是对于非线性问题。

6.8.1 对称性与变换

通过理解问题中所谓的微分对称性（即与系统的无穷小几何相关的对称性），可以掌握一些用于线性和非线性问题的变换。尽管这一主题远远超出了本文的范围，但这种方法的基本结果是由德国数学家埃米·诺特（1882–1935）发现的。埃米·诺特不仅是一位女性数学家，这一点本身就很值得注意，因为在那个时代女性数学家非常少；而且，她的工作还促成了数学领域一些深刻理论的发展，甚至催生了在她之前并不存在的完整学科。简而言之，诺特定理宣称，每一条守恒定律都有一个连续的对称性，反之亦然——对于每一个连续对称性，都存在一条与之对应的守恒定律。

What is meant by a symmetry here can be roughly thought about in terms of our conventional notion of the word. Take, for example, a square drawn of a piece of paper. If the square is rotated by any integer multiple of $\frac{\pi}{2}$ radians, then it appears to be unchanged. Thus, it has a rotational symmetry with respect to rotations of $\frac{\pi}{2}$ radians. Similarly, a square is unchanged by reflections across its horizontal or vertical axes; thus, it has two kinds of reflection symmetry. Many of the familiar and unfamiliar kinds of symmetries are discussed in a text by the famous physicist Herman Weyl (Weyl, 1952). While most familiar examples represent discrete symmetries (i.e., the rotational symmetry for a square happens for only certain angles of rotation), continuous symmetries are a bit more complex. While the details are not important here, one can think of continuous symmetries as symmetries that exist for a sequence of infinitesimal actions (e.g., rotations, translations).

Regardless, the importance of symmetries and of Emmy Noether's theorem have been critical to the study of PDEs, and to the study of various physical pheneomena (often expressed through PDEs). A few deep results in quantum mechanics have even been discovered by looking for certain kinds of physical symmetries in the mathematical laws that describe it.

6.8.2 An Example Transformation

For our purposes, we might best understand how transformations can lead to a simplified problem by an explicit example. In the following example, we propose a continuous symmetry in time; that is, we find that if we exponentially re-scale the dependent variable in time, we capitalize on a particular kind of symmetry (a dilation-of-time symmetry) that puts the problem in a new, much simpler form.

Example 6.9 (Transformation of a reaction-diffusion equation.).

Consider a parabolic problem similar to the one described in the previous example for a diffusion-reaction system on the interval $x \in [0,1]$. This problem meets our (somewhat limited) definition of being well-posed.

$$\frac{\partial u}{\partial t} = D\frac{\partial^2 u}{\partial x^2} - ku$$

$$B.C.\ 1 \qquad u(0,t) = 0$$

$$B.C.\ 2 \qquad u(1,t) = 0$$

$$I.C.\ 1 \qquad u(x,0) = f(x)$$

In the next chapter, we will examine finding solutions to this problem using separation of variables. Actually, the solution to the pure diffusion problem is well known, and is reasonably straightforward. While the solution to the reaction-diffusion problem is possible using separation of variables, it does contain a number of complications that make the solution unwieldy.

因此，我们考虑 Danckwerts（1951）提出的以下著名变换（Danckwerts 是一位化学工程师），以便将该问题转化为一个纯扩散问题。通过定义以下新的因变量w，即可实现这一变换。

$$u(x,t) = w(x,t)\exp(-kt)$$

To adopt this transformation, we simply substitute the right-hand side for u in both the PDE and ancillary conditions. The following results

$$\frac{\partial}{\partial t}(w(x,t)e^{-kt}) = D\frac{\partial^2}{\partial x^2}(w(x,t)e^{-kt}) - k(we^{-kt})$$

$$B.C.\ 1 \qquad w(0,t)e^{-kt} = 0$$

$$B.C.\ 2 \qquad w(1,t)e^{-kt} = 0$$

$$I.C.\ 1 \qquad w(x,0)e^{0} = f(x)$$

Applying the product rule for derivatives and extracting the exponential from the space derivative yields.

$$\frac{\partial w(x,t)}{\partial t}e^{-kt} - kw(x,t)e^{-kt} = e^{-kt}D\frac{\partial^2}{\partial x^2}w(x,t) - kw(x,t)e^{-kt}$$

边界条件 1 $$w(0,t)e^{-kt} = 0$$

边界条件 2 $$w(1,t)e^{-kt} = 0$$

I.C. 1 $$w(x,0)e^{0} = f(x)$$

Noting that the exponentials are not zero, and that for the initial condition the exponential is identically 1, we have the final result

$$\frac{\partial w(x,t)}{\partial t} = D\frac{\partial^2 w(x,t)}{\partial x^2}$$

边界条件 1 $$w(0,t) = 0$$

边界条件 2 $$w(1,t) = 0$$

I.C. 1 $$w(x,0) = f(x)$$

这是一个类扩散方程。

6.9 作为抛物型偏微分方程稳态出现的椭圆型偏微分方程

到目前为止，我们讨论了双曲型方程（一阶波动方程）和抛物型方程（热量/扩散方程），它们都是通过应用各种通量定律推导出来的。我们尚未讨论椭圆型方程。在这里，我们将讨论此类方程如何经常作为抛物型方程的稳态形式出现。由于本章将讨论范围限制在一个空间维度和一个时间维度，因此得到的方程是一个相对简单的椭圆型问题（最终得到的是常微分方程，而非偏微分方程）。不过，这里要阐明的要点是：（1）椭圆型方程如何经常作为抛物型方程的稳态形式出现；（2）对椭圆型问题在物理上的表现形成一些认识。

瞬态问题的稳态（即状态随时间变化，且解从初始配置开始随时间演变）就是时间偏导数趋近于零时的状态。这里使用趋近一词，是因为对于瞬态二阶偏微分方程，稳态通常在时间变得任意大时出现（即当$t \to \infty$时）；因此，它代表一种极限行为。

When transient problems are composed of one space and one time dimension, then the steady state solutions are identical in form to ordinary differential equations. There are some philosophically delicate points here to be made about the use of partial versus total derivatives for such problems (technically, as $t \to \infty$, they are no longer functions of time, thus there is only one independent variable, so one can certainly argue that they are formally, in fact, ODEs in the limiting case), but we will not dwell on these. Instead, we will treat such problems as being *de facto* ODEs, while continuing to keep in mind that the equation itself came from a more general form. An example of handling the determination of a steady state solution is given below.

Example 6.10 (A steady-state reaction-diffusion equation.). Let's start with the reaction-diffusion problem given above. Specifically, we have

$$\frac{\partial u}{\partial t} = D\frac{\partial^2 u}{\partial x^2} - ku$$

$$B.C.\ 1 \qquad u(0,t) = u_0$$

$$B.C.\ 2 \qquad \frac{\partial u(L,t)}{\partial x} = 0$$

$$I.C.\ 1 \qquad u(x,0) = f(x)$$

Now, as time becomes arbitrarily large, we expect this equation to reach some steady state. To determine if this is, in fact, true, we can examine what happens if the time derivative is set to zero. If a solution exists, then the problem has a steady state. Specifically, we have the problem (after dividing the balance equation through by D).

$$\frac{d^2 u}{dx^2} - \frac{k}{D}u = 0$$

$$B.C.\ 1 \qquad u(0,t) = u_0$$

$$B.C.\ 2 \qquad \frac{du(L,t)}{dx} = 0$$

Note that we no longer have need for the initial condition, because there is no integration in time required; the partial derivatives have also been replaced by ordinary derivatives, since time is no longer an independent variable in this equation.

Solving this equation is the same as described in Chapter 2 on ODEs. It is a homogeneous equation, and solving it requires only that we first determine which case it represents by computing the roots of the characteristic equation. Noting $a = 1$, $b = 0$, and $c = -k/D$, we have the roots given by

$$s_1 = +\sqrt{k/D} \qquad s_2 = -\sqrt{k/D}$$

For ease in notation, set $\beta = |\sqrt{k/D}|$. The solution is now

$$u(x) = C_1 \exp(\beta x) + C_2 \exp(-\beta x)$$

$$du(x)/dx = C_1\beta \exp(\beta x) - C_2\beta \exp(-\beta x)$$

The second boundary condition gives us

$$0 = C_1\beta \exp(\beta L) - C_2\beta \exp(-\beta L)$$

and the first boundary condition

$$u_0 = C_1 + C_2 \implies -C_2 = C_1 - u_0$$

Substituting this last relation into the second boundary condition expression gives

$$C_1 \left[\exp(\beta L) + \exp(-\beta L)\right] = u_0 \exp(-\beta L)$$

or, recalling the definition of the hyperbolic cosine $\cosh(x) = 1/2(e^x + e^{-x})$, we have

$$C_1 = \frac{u_0 \exp(-\beta L)}{2\cosh(\beta L)}$$

$$C_2 = u_0 - \frac{u_0 \exp(-\beta L)}{2\cosh(\beta L)}$$

The final solution is

$$u(x) = \frac{u_0 \exp(-\beta L)}{2\cosh(\beta L)} \exp(\beta x) + \left(u_0 - \frac{u_0 \exp(-\beta L)}{2\cosh(\beta L)}\right) \exp(-\beta x)$$

For reference, the solution to this problem is given in Fig. 6.3.

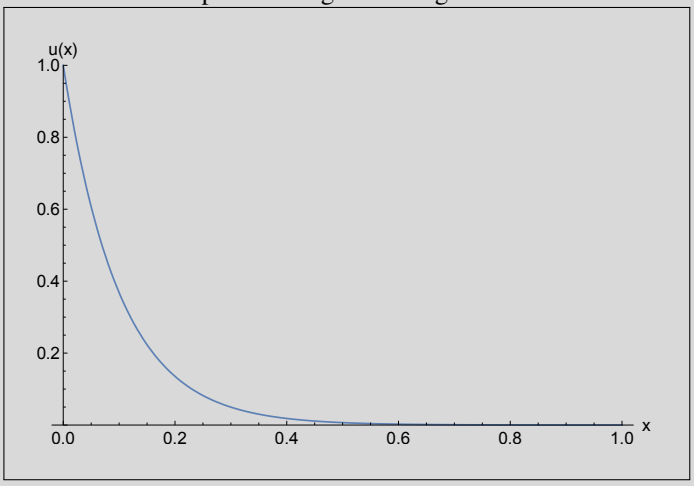


Fig. 6.3: The solution to the steady-state diffusion-reaction problem.

*附录：波动方程的推导

There are a number of different *kinds* of waves, but for our purposes it is sufficient to consider the following types

1. **Longitudinal waves**. For example, the first-order wave equation describing a purely translating signal; or the compressive waves that comprise sound.

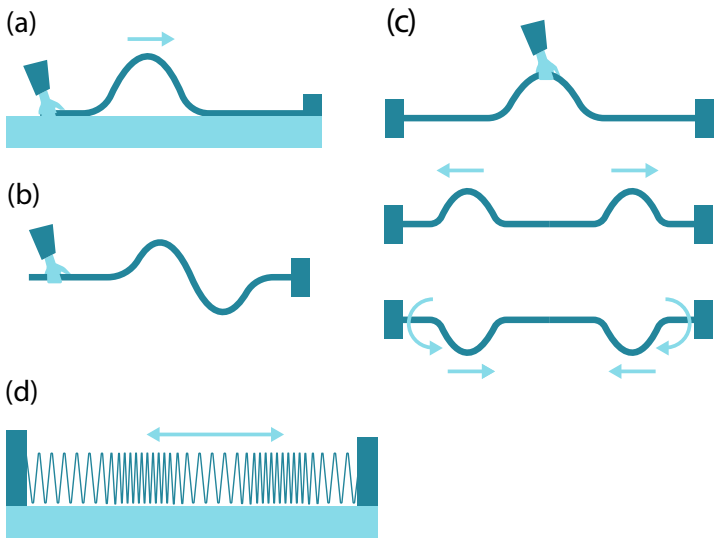


图 6.4 在弹簧玩具中可以观察到的波形。(a) 受到扰动的支座弹簧玩具形成横波，可用一阶波动方程描述。(b) 受到扰动的悬挂弹簧玩具形成横波，需要用二阶波动方程描述。(c) 悬挂的弹簧玩具因中心受到扰动而具有初始条件；形成两个沿相反方向传播的波。(d) 挤压支座弹簧玩具的中心形成的纵波；这也需要用二阶波动方程描述。

2. **Transverse waves**. An example here might be waves on a body of water, where the vertical displacement from the equilibrium position defines the wave.

3. **纵波和横波的组合**。这种运动的一个例子可以在机械弹簧中观察到。如果你曾经玩过弹簧玩具（Gluck，2010），很可能观察到过纵波（压缩波）和横波。当弹簧玩具的一端受到周期性驱动时，玩具弹簧会形成正弦函数状形状，同时也会出现压缩和伸展区域。因此，两种波都存在。

在图 6.4 中，展示了几种波的情况。第一种情况（图 6.4(a)）中，由地板支撑的弹簧玩具只有一个位移方向，也就是说，它只能在地板上方沿垂直方向产生位移。弹簧玩具一端的扰动将沿着弹簧玩具传播，而不会改变形状。这种波主要是一种横波，可以用一阶波动方程来描述。作为第二种波，考虑一个悬挂起来并固定在墙上的弹簧玩具。悬挂的弹簧玩具可以沿正方向和负方向发生位移，如图 6.4(b) 所示。作为另一种情况，假设弹簧玩具悬挂在两面墙之间（图 6.4(c)）。如果在弹簧玩具中部施加正弦初始位移（即初始条件），就会形成两个波，每个波的高度都是初始位移的一半，并且两个波沿相反方向传播。当波接触墙壁时，会发生反射，且其位移方向反转。这种运动由二阶波动方程描述。作为波动的最后一种模式，我们回到弹簧玩具由地板支撑的情况（图 6.4(d)）。通过压缩弹簧玩具的一部分然后释放，可以产生纵波。随后，纵波通过压缩和伸展弹簧玩具沿着平行于其中心轴的方向传播，但不会产生垂直位移。在弹簧玩具中部进行这一操作时，同样会形成两个波，这些纵波将沿两个方向传播，也可以用二阶波动方程来描述。实际弹簧玩具的许多运动都是横波和纵波的组合。

虽然弹簧玩具是一种玩具，但它是理解波动的一种非常有用的方式。它也为我们开始思考这些波提供了一种很好的方法。首先，回顾一下无限介质中的一阶波动方程，其表达式为

$$\frac{\partial u}{\partial t} = -c\frac{\partial u}{\partial x} \quad x \in [0,\infty) \tag{6.43}$$

$$u(x,0) = f(x) \tag{6.44}$$

Where here u represents the transverse displacement, and a represents the width of an initial *single* disturbance near the left boundary. As an example, one might use the function $f(x) = \sin(\pi x)[H(x) - H(x-a)]$ to represent the initial condition of a *single sine pulse* that travels along the system. A solution to this problem can be computed by a number of methods, although we have not yet learned about these. Here, we take a somewhat inverted approach by providing the solution, and then checking to validate it. The solution to the problem above is

$$u(x,t) = f(x-ct) \tag{6.45}$$

无论 o f what function $f(x)$ is chosen. In fact, we can codify this as a theorem, with following proof.

定理 6.1（具有初始条件的一阶波动方程的解）。具有平移速度 c 和初始条件 $u(x,0) = f(x)$ 的一阶波动方程的解为 $u(x,t) = f(x - ct)$。

证明.
The proof is not too difficult, but does require a change of variables, and appropriate differentiation of those variables. Let

$$y(x,t) = x - ct \tag{6.46}$$

with this definition, note the following.

$$\frac{\partial y}{\partial t} = -c \tag{6.47}$$

$$\frac{\partial y}{\partial x} = 1 \tag{6.48}$$

Now, technically y is a function. However, using the normal rules of differentiation (the chain rule), we can compute derivatives with respect to y. These derivatives can be computed as follows.

$$\begin{aligned} \frac{\partial u}{\partial t} &= \frac{\partial u}{\partial y}\frac{\partial y}{\partial t} \\ &= -c\frac{\partial u}{\partial y} \end{aligned} \tag{6.49}$$

$$\begin{aligned} \frac{\partial u}{\partial x} &= \frac{\partial u}{\partial y}\frac{\partial y}{\partial x} \\ &= \frac{\partial u}{\partial y} \end{aligned} \tag{6.50}$$

现在，将这些结果代入方程（6.43），我们发现等式两边确实相等，这表明$f(x - ct)$ 确实是一个解。由于该问题是一个一阶线性问题，我们还知道该解-

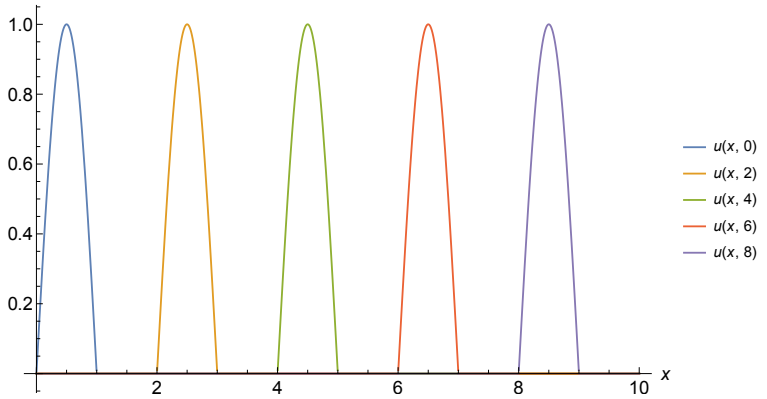


图6.5 一阶波动方程的解，其中 $c = 1$ 米/秒，$a = 1$ 米。假设横向位移为无量纲量，它们是函数 $f(x)$ =正弦(πx)[赫维赛德阶跃函数(x) – 赫维赛德阶跃函数$(x - 1)$] 沿距离 ct 的平移。时间分别为$t = 0,2,4,6,$ 和 8 秒，沿x轴的距离单位为米。

tion is unique. ■

一阶波动方程描述了波以速度 c 沿 $+x$ 方向平移运动且形状保持不变的情况。然而，这种描述存在一些不足。首先，它无法描述当扰动发生在介质中心而非一端时所产生的物理现象。对于这种情况（例如，图 6.4(b)），平移会从初始扰动处向两个方向传播。其次，由于一阶波动方程仅包含关于时间的一阶导数，因此只需要一个初始条件（初始位置）来确定任何未知的积分常数。然而，在某些物理情形下，我们可能希望同时指定空间中的初始配置（形状）以及 初速度。使用一阶波动方程无法做到这一点。

一种解决方法是运用已有知识，思考当初始扰动发生在波介质中部时，合适的解可能是什么样子。目前，我们只考虑初速度恒为零，且初始位移 $f(x)$ 关于$x = 0$对称的解。我们预计，波会从扰动处向两个方向（$+x$ 和 $-x$）传播。根据能量守恒，初始扰动的势能必须转化为两个波的动能。因此，每个波的形状都与初始扰动 $f(x)$ 相同，但高度只有其一半（此处，高度可视为质量的替代量）。根据我们对一阶波动方程的了解，可以设想如下形式的解

$$u(x,t) = \tfrac{1}{2}[f(x-ct)] + \tfrac{1}{2}[f(x+ct)] \tag{6.51}$$

这个解是基于对一阶波动方程行为的定性认识，以及由能量守恒得出的一个非常粗略的约束条件而提出的假设（我们认为高度与能量成正比，但并未详细证明这一点！）。现在问题来了——“如果这表示一个解，那么它是哪个偏微分方程的解？”我们的设想是：“这是二阶波动方程的一个解。”现在要做的就是证明这一点。同样，我们采用一种有些反常的做法：先给出解，再进行检验以验证它。下面给出一个定理及其证明，证明过程大体遵循一阶波动方程的证明思路。

Theorem 6.2 (The Second-Order Wave Equation). *Suppose we have a wave-like solution in an 1-dimensional infinite medium with the form* $u(x,t) = \frac{1}{2}[f(x-ct)] + \frac{1}{2}[f(x+ct)]$. *Here,* $f(x)$ *is the initial condition where* $f(x)$ *is a symmetric function around* $x = 0$, *and it is assumed that the initial velocity is identically zero. Then, the associated PDE describing this motion is given by*

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2} \tag{6.52}$$

证明。

The proof that the PDE given above (the classical "wave equation") is indeed solved by $u(x,t) = \frac{1}{2}[f(x-ct)] + \frac{1}{2}[f(x+ct)]$ follows very closely the proof associated with the first-order wave equation. Here, we have two translation variables, so we need two definitions as follows.

$$y(x,t) = x - ct \tag{6.53}$$
$$z(x,t) = x + ct \tag{6.54}$$

with this definition, note the following.

$$\frac{\partial y}{\partial t} = -c, \qquad \frac{\partial y}{\partial x} = 1 \tag{6.55}$$
$$\frac{\partial z}{\partial t} = c, \qquad \frac{\partial z}{\partial x} = 1 \tag{6.56}$$

Here, both y and z are a functions. Again, we can use the normal rules of differentiation (the chain rule), to compute derivatives with respect to y and z. These derivatives can be computed as follows.

$$\frac{\partial u}{\partial t} = \frac{\partial u}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial u}{\partial z}\frac{\partial z}{\partial t} = \underline{-c\frac{\partial u}{\partial y} + c\frac{\partial u}{\partial z}} \tag{6.57}$$
$$\frac{\partial u}{\partial x} = \frac{\partial u}{\partial y}\frac{\partial y}{\partial x} + \frac{\partial u}{\partial z}\frac{\partial z}{\partial x} = \underline{\frac{\partial u}{\partial y} + \frac{\partial u}{\partial z}} \tag{6.58}$$
$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial}{\partial t}\left(-c\frac{\partial u}{\partial y} + c\frac{\partial u}{\partial z}\right) = -c\frac{\partial}{\partial y}\left(\frac{\partial u}{\partial t}\right) = -c\frac{\partial}{\partial y}\left(\frac{\partial u}{\partial t}\right) + c\frac{\partial}{\partial z}\left(\frac{\partial u}{\partial t}\right) = \underline{c^2\frac{\partial^2 u}{\partial y^2} + c^2\frac{\partial^2 u}{\partial z^2}} \tag{6.59}$$
$$\frac{\partial^2 u}{\partial x^2} = \frac{\partial}{\partial t}\left(\frac{\partial u}{\partial y} + c\frac{\partial u}{\partial z}\right) = \frac{\partial}{\partial y}\left(\frac{\partial u}{\partial t}\right) = \frac{\partial}{\partial y}\left(\frac{\partial u}{\partial t}\right) + \frac{\partial}{\partial z}\left(\frac{\partial u}{\partial t}\right) = \underline{\frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}} \tag{6.60}$$

最后，将下划线标出的结果代入式 (6.52) 中的 $\partial^2 u / \partial t^2$ 和 $\partial^2 u / \partial x^2$，我们发现，二阶波动方程确实与所提出的解一致。■

这一推导过程本身并不是在推导波动方程。实际上，它做的是：（1）提出一个解；然后（2）说明该解与二阶（经典）波动方程的可能解一致。请回想，这个解是基于若干要求提出的（对称性、初速度为零）。事实证明，二阶波动方程允许比所提出的解更一般的解。由于时间导数有两个，初始条件和初始速度都可以作为辅助条件，用来确定时间积分中的两个未知积分常数。初始波形也不必是对称的；二阶波动方程同样允许非对称的初始条件。

二阶波动方程的传统推导（直接从动量守恒出发）相对容易找到——大多数偏微分方程入门教材都会以某种形式给出推导。然而，这些推导即便算得上妥当，论证也往往不够充分，而且常常建立在许多不正确或含糊不清的假设之上。文献中有基于可靠物理学的波动方程严谨推导（Antman，1980；Yong，2006）。对这方面感兴趣的读者，可以参考这两篇文献，它们是可靠推导的良好资料来源。二阶波动方程本身的推导并非特别困难，但它涉及的概念与本章其他部分所介绍的概念迥异，因此会大幅偏离本文的主要内容。故此，我们有意省略从基本原理推导波动方程的内容。不过，在学习傅里叶变换时，我们还会进一步探讨这个问题。届时，我们可以进一步了解二阶波动方程的性质。

有必要对可作为一阶或二阶波动方程解的函数种类f所受的限制作一些说明。到目前为止，我们尚未具体说明是否存在限制。由偏微分方程本身的性质可推测，对于一阶波动方程，函数必须至少具有一阶连续导数（关于x和t），而对于二阶波动方程，则必须具有二阶导数（关于x和t）（即函数f必须属于可微性类C^1或C^2）。尽管这确实是波动方程解的一个充分条件，但并非必要条件。虽然我们无法在本文中探讨其细节，但仍值得注意的是，波动方程的解可以按照此前第5章提到的广义函数来解释。这样看来，函数f甚至可以是不连续函数（例如由$B(x) = H(x-a) - H(x-b)$，$b > a$定义的箱形或阶跃函数）。

习题

Applied and More Challenging Problems

1. 常规扩散方程的形式为（如前文所推导，本章）

$$\frac{\partial u}{\partial t} = D\frac{\partial^2 u}{\partial x^2} \tag{6.61}$$

This derivation assumed that the diffusion coefficient was a constant, and did not depend on space. Assume now that we have the case that the diffusion constant depends upon space; that is, it is given by a function of x in the form $D(x)$. Starting by re-stating Fick's law (i.e., the diffusive mass flux constitutive equation) for this situation. Then, substitute this result into the general conservation equation. How does the result for the case of a spatially-dependent $D(x)$ change from the result given by Eq. (6.61)? Note, this last question is asking for more of a qualitative description rather that an extensive discussion.

2. Find the steady-state solution to the following first-order hyperbolic problem on the domain $x \in [0, L]$.

$$\frac{\partial u}{\partial t} = -v_0\frac{\partial u}{\partial x} - ku$$

$$B.C.\ 1 \qquad u(0,t) = u_0$$

$$I.C.\ 1 \qquad u(x,0) = 0$$

是否 th e solution to this problem match your intuition? This last question is a qualitative one, so a 定性 ve answer is all that is expected.

3. 本文的一个主题是：面对一个不知道如何求解的问题时，将其转化为一个可以求解的问题。不过，虽然这句话说起来容易，但要知道如何变换这类问题却并非总是那么容易！

由著名化学工程师 Peter Danckwerts（1916–1984）提出的一种有趣的变换，可以将一个扩散-反应问题转化为一个纯扩散问题，从而大大简化求解过程。首先，考虑下面这个具有诺伊曼边界条件的问题。

$$\frac{\partial u}{\partial t} = D\frac{\partial^2 u}{\partial x^2} - ku \tag{6.62}$$

$$B.C.1 \qquad \left.\frac{\partial u}{\partial x}\right|_{(0,t)} = 0 \tag{6.63}$$

$$B.C.2 \qquad \left.\frac{\partial u}{\partial x}\right|_{(L,t)} = 0 \tag{6.64}$$

$$I.C. \qquad u(x,0) = f(x) \tag{6.65}$$

Now, consider the following transformation of variables

$$u(x,t) = \exp(-kt)w(x,t)$$

Show that substituting this transformation into the diffusion-reaction problem given by Eqs. (6.62)-(6.62) will transform that set of equations into the following equations in terms of the variable w

$$\frac{\partial w}{\partial t} = D\frac{\partial^2 w}{\partial x^2} \tag{6.66}$$

$$B.C.1 \qquad \left.\frac{\partial w}{\partial x}\right|_{(0,t)} = 0 \tag{6.67}$$

$$B.C.2 \qquad \left.\frac{\partial w}{\partial x}\right|_{(L,t)} = 0 \tag{6.68}$$

$$I.C. \qquad w(x,0) = f(x) \tag{6.69}$$

4. Burgers 方程是一种非线性方程，可作为一维纳维–斯托克斯方程的原型，并应用于耗散波传输领域。该方程是非线性的，其形式为

$$u_t + uu_x - \nu u_{xx} = 0 \tag{6.70}$$

非线性方程通常很难通过解析方法求解（如果它们能够求解的话！）。对于这个问题，有一种著名的两步变换方法，可以将该方程转化为线性方程。这种变换称为 Cole-Hopf 变换，其步骤如下。首先，我们定义变量之间的以下关系。

$$u = w_x$$

Now, substituting this into the original equation, we find an equation in w of the form

$$w_{xt} + w_x w_{xx} - \nu w_{xxx} = 0$$

It is not clear at this juncture that anything has been made better, but we will press on. First note the following identity

$$\frac{\partial}{\partial x}\left(\frac{1}{2}w_x^2\right) = w_x w_{xx}$$

请注意，这里我们混合使用了不同的导数符号表示法，只要确保始终保持一致，这样做完全可以接受。通过简单应用链式法则，很容易验证这一结果。此外，值得回顾一下，导数的复合运算可以按以下方式进行组合。

$$\frac{\partial}{\partial x}w_x = \frac{\partial}{\partial x}\left(\frac{\partial w}{\partial x}\right) = \frac{\partial^2 w}{\partial x^2} = w_{xx}$$

Now, this means that the transformed Burgers' equation can be rewritten as follows

$$w_{xt} + \frac{\partial}{\partial x}\left(\frac{1}{2}w_x^2\right) - \nu w_{xxx} = 0$$

All of this work has paid off a bit, because now we can extract one partial derivative with respect to x outside the expression as a whole, giving

$$\frac{\partial}{\partial x}\left(w_t + \frac{1}{2}w_x^2 - \nu w_{xx}\right) = 0$$

Both 两边可以关于x 进行一次不定积分。结果 is

$$w_t + \frac{1}{2}w_x^2 - \nu w_{xx} = C_1$$

虽然这个方程似乎可能比原方程更好，但由于项$1/2w_x^2$的存在，它仍然是非线性的。现在，对于作为本题基础的第二次变换，证明由以下变量变换给出的是

$$w(x,t) = -2\nu \ln[z(x,t)]$$

将导致一个线性二阶偏微分方程，其形式为扩散方程，即，

$$z_t = \nu z_{xx}$$

其中我们假设$C_1 = 0$。

5. Cahn–Hilliard（CH）方程是一个平衡方程，用于预测流体中的两相从高度混合的初始配置开始如何分离。在这种情况下，u 实际上是一个因变量，用于编码两相几何结构的内部配置。变量 u 最初被视为取值为$u \in [0,1]$的变量。因此，u的值表示两个相中的一个。值为 1 表示第一相，值为 0 表示第二相。在连续介质力学研究中，取值为 1 或 0 以表示各相的变量有时称为相指示变量。

相指示变量u 的非线性平衡方程具有如下形式

$$\frac{\partial u}{\partial t} = \frac{\partial^2}{\partial x^2}\left(Du^3 - Du - \gamma D\frac{\partial^2 u}{\partial x^2}\right) \tag{6.71}$$

If you are interested in the topic, an example of the application of the CH equation in one dimension is reported in the following paper.

Argentina, M., Clerc, M., Rojas, R., Tirapegui, E.: 《一维Cahn–Hilliard模型的粗化动力学》。Physical Review E 71(4)，046210（2005）

该方程的行为十分有趣。从代表完全混合流体的（相对）均匀初始条件出发，CH方程可以预测流体将如何分离。作为一种模拟物，可以想想油水乳液如何随着时间推移形成大油滴：小油滴逐渐聚合在一起（你或许在再平常不过的油醋沙拉酱中观察到过这种现象.....）虽然CH方程在二维和三维系统中的应用最为有趣，但

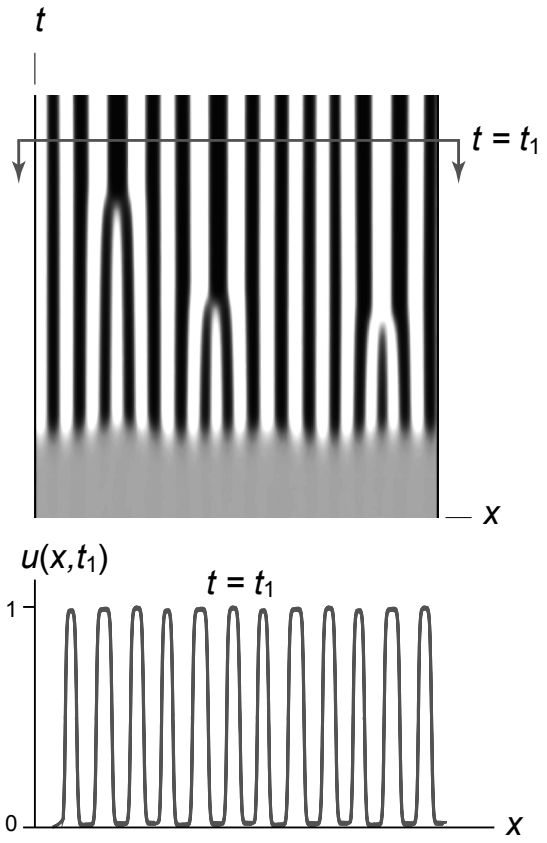


图 6.6 Cahn-Hilliard方程的解。该方程将相混合物的初始状态（由时间原点附近的灰色区域表示）变换为两相彼此分离的状态。请注意，这是一个展示复杂行为的案例，这种行为称为"涌现行为"（见第1章）。上图显示了$u(x,t)$随x和t变化的取值；定义域中的数值以强度图表示（白色=0、黑色=1，灰色表示介于两者之间的值），对应每个时空位置。下图展示了固定时间值下$u(x,t_1)$的空间变化；它对应于上图中标记为"$t = t_1$"的水平线。这个非线性方程竟能产生如此极端的行为，着实令人惊叹。图表依据Argentina等人（2005）提供的数据绘制。

are 1-dimensional solutions also. As an example, Fig. 6.6 provides the time-space value of the indicator function for one such solution.

Recall, the general conservation equation (with no source) for 1-dimension takes the form

$$\frac{\partial u(x,t)}{\partial t} = -\frac{\partial}{\partial x} j(x,t) \tag{6.72}$$

Comparing Eq. (6.72) with the expression given by Eq. (6.71), determine what the flux, $j(x,t)$ must be as a function of u. To be clear– you need only compare these two equations to determine what the the flux must be interpreted as in Eq. (6.71); this involves a bit of algebra and application of the derivative operators, but nothing else. To start, think about how you can pull out the operator $-\partial/\partial x$ from the right-hand side of Eq. (6.71). Once you do that, what this operator acts upon will then be the flux, $j(x,t)$.

6. 考虑一个抛物型问题，该问题与示例 6.9 中描述的问题类似，定义在区间 $x \in [0,1]$ 上。该问题与示例略有不同，因为反应速率系数是时间的函数。其平衡方程、边界条件和初始条件为

$$\frac{\partial u}{\partial t} = D\frac{\partial^2 u}{\partial x^2} - k(t)u$$

边界条件 1边界条件 2*I.C.* 1

$$u(0,t) = 0$$
$$u(1,t) = 0$$
$$u(x,0) = f(x)$$

For this problem, show that the transformation

$$u(x,t) = \exp\left(-\int_{\tau=0}^{\tau=t} k(\tau)\,d\tau\right)$$

can be used to transform the reaction-diffusion problem into a purely diffusion problem.

7. On the interval $x \in (-\infty,\infty)$ the heat/diffusion equation

$$\frac{\partial u}{\partial t} = D\frac{\partial^2 u}{\partial x^2}$$
$$u(x,0) = \delta(x)$$

has a solution known as the *fundamental solution* of the heat/diffusion equation. This solution is

$$u(x,t) = \frac{1}{\sqrt{4\pi Dt}}\exp\left(-\frac{x^2}{4Dt}\right)$$

Taking derivatives and substituting into the heat/diffusion PDE, show that this is indeed a solution to the PDE.

8. We have discussed first-order wave equations. In the Appendix we have shown that for the first-order wave equation

$$\frac{\partial u}{\partial t} = -c\frac{\partial u}{\partial x}$$
$$u(x,0) = f(x)$$

has the general solution $u(x,t) = f(x-ct)$. That derivation involved a change of variables, and was slightly complex. For this problem, show that the for the specific initial condition $f(x) = \exp(-x^2)$, that the solution $u(x,t) = \exp[-(x-ct)^2]$ is a solution to the first-order wave equation, and that it meets the proposed initial condition. Do this *directly* by taking derivatives of the given form for $u(x,t)$.

9. We have discussed first-order wave equations, and, in the Appendix, the relationship between first- and second-order wave equations. In that derivation, we showed that the second-order wave equation

$$\frac{\partial^2 u}{\partial t^2} = -c^2\frac{\partial^2 u}{\partial x^2}$$
$$u(x,0) = f(x)$$

has the general solution $u(x,t) = \frac{1}{2}f(x-ct) + \frac{1}{2}f(x+ct)$. That derivation involved a change of variables, and was moderately complex. For this problem, show that the for the specific initial condition $f(x) = \exp(-x^2)$, that the solution $u(x,t) = \frac{1}{2}\exp[-(x-ct)^2] + \frac{1}{2}\exp[-(x+ct)^2]$ is a solution to the seconds-order wave equation, and that it meets the proposed initial condition. Do this *directly* by taking derivatives of the given form for $u(x,t)$.

“李群、特殊函数与分离变量法之间的基本关系近来已得到阐明……线性偏微分方程的对称群与方程能够得到可分离解的坐标系之间的主要思想，通过实例最容易理解。”

Willard Miller，数学家，在将分离变量法发展为一种通用技术方面取得了进展。

7 分离变量法（SOV）

7.1 引言

分离变量法（SOV）已有数百年的研究历史，已知最早的成果由 Liouville 于 1846 年提出。尽管这一思想看似简单，但直至今日，人们仍在深入研究它，以便更好地理解该方法能够提供解的条件（Bluman 等人，2010；Kalnins 等人，2018）。事实证明，分离变量法与底层偏微分方程及其边界条件的结构（更具体地说，是微分几何）密切相关。要说明分离变量法何时以及为何有效，需要理解从拓扑学到称为群的代数结构等广泛的数学概念。将这些思想应用于偏微分方程的研究，推动了应用数学领域一些最受瞩目的成果。感兴趣的读者可以阅读 Miller（1977）的著作，以初步了解这一主题；该书还介绍了这一主题的部分历史。

All of this may make it sound like the method is difficult. In fact, the method itself poses no particular challenges in applying. In fact, it was originally developed by mathematicians such as Liouville (and later, John Bernoulli) who used physical motivation and observations from what they knew of existing solutions to PDEs to propose the method empirically. In other words, it was easy to see that, for many interesting cases, the method of separation of variables could successfully generate solutions. It is only when attempting to understand fully as to what conditions the method works (and why it does when it is successful) that one encounters the more advanced notions of symmetries and differential geometry.

For our purposes, we can take a more empirical tactic with the method. We will *propose* that the method be tried, and then *check* to see if this is indeed the case. There is nothing inherently unscientific or wrong with such an empirical approach! In fact, it is used very frequently in the study of mathematics. Recall, in our study of second order ODEs, we proposed that useful solutions to a homogeneous ODE of second order with constant coefficients could be represented by $u(x) = \exp(-sx)$; this proposition, of course, led ultimately to developing all possible solutions for such second order ODEs. Thus, this will be the approach

here, and, as we will ultimately determine, the method works for a number of PDEs on domains that are sufficiently regular in some sense.

正如初学新内容时常见的情况一样，理解一种方法的最佳方式是从实例入手。事实上，我们将首先通过两个变量下的三个典型方程（抛物型扩散方程、双曲型波动方程和椭圆型拉普拉斯方程）来演示求解方法。看到这种方法至少适用于本例后，我们就可以进一步加深对该方法的理解。

继续之前，请注意，即使对于线性偏微分方程，也很难给出可分离的一般条件。然而，对于三个典型实例（抛物型扩散方程、双曲型波动方程和椭圆型拉普拉斯方程），总是可以将变量分离法用作求解方法，前提是

1. 方程具有常系数 2. 方程是齐次的 3. 方程不含混合偏导数 4. 边界条件是线性齐次的，并且与坐标轴对齐 5. 初始条件由一个函数定义，且假定该函数具有有效的傅里叶级数表示

on

请注意，这些条件并非使用分离变量法的必要条件。然而，就我们的目的而言，它们是充分条件，也就是说，如果满足这些条件，就可以采用本章所述的方法使用分离变量法。这也意味着（根据必要和充分的定义），有些解可以通过分离变量法求得，而这些解并不满足上述四个条件中的全部条件（事实上，确实如此！）。

在本章余下的部分中，我们将探讨三种典型方程（抛物型热传导/扩散方程、双曲型波动方程和椭圆型拉普拉斯类方程）的通解。对于这三类方程，我们将重点研究仅包含两个自变量的系统（对于抛物型和双曲型方程，分别为一个空间变量和一个时间变量；对于拉普拉斯类方程，则为两个空间变量），并考虑齐次指定值（狄利克雷）边界条件。本章分析的一个重要组成部分，是继续培养我们在所研究问题的建模 方面的能力。具体而言，我们将展示如何利用基于物理 （而非数学）背景的各种约束条件，帮助我们求得能够反映所关注特定物理量的解。在掌握了三种典型方程的分离变量法（SOV）基本思想后，我们将把注意力转向涉及其他类型边界条件的情形。边界条件是数学建模中不可或缺的一部分，值得我们特别关注，以便更全面地理解各种类型的边界条件如何影响我们所建立的模型。

7.2 Terminology

- **分离变量法**。一种用于求解某些线性偏微分方程的过程。对于具有n 个自变量的偏微分方程，$\mathbf{x} = x_i$, $i = 1..n$，假设其解可表示为$u(\mathbf{x}) = f1(x1)\, f2(x2) \cdots f3(x3)$。

- **Transient.** The word *transient* in PDEs is used to indicate an process that changes in time. As an example, consider a drop of dye added to a transparent container of water. Starting from this initial state (the drop), the dye continues to spread in time; in more formal terms, if $u(\mathbf{x},t)$ is the dye concentration for

all locations at all possible times, then initially $\frac{\partial u}{\partial t} \neq 0$.

• **稳态**。瞬态的反义词。处于稳态的过程不会随时间发生变化。如果 $u(\mathbf{x},t)$ 表示所关注的因变量，那么在所有位置均有 $\partial_u \partial_t = 0$。瞬态过程有可能达到 稳态。在我们的大多数模型中，这需要满足 $t \to \infty$，但也有一些由偏微分方程描述的过程能够在有限时间内达到稳态。例如，对于具有零级反应的扩散-反应问题（即反应速率与浓度无关），该过程会在某个有限时间内完成。随着时间任意增大，许多瞬态问题会趋向于稳态；这取决于所涉及的偏微分方程的具体形式。例如，许多热传导/扩散方程都具有明确定义的稳态。相反，按照我们所定义的稳态，许多波动型方程并不存在稳态。

• **热传导/扩散方程**。这是通常在偏微分方程入门课程中学习的三个经典方程之一（热传导/扩散方程、波动方程、拉普拉斯方程）。它是一个抛物型方程，用于描述具有扩散特性的现象。在一维空间和一维时间中，(x,t)它的形式为

$$\frac{\partial u}{\partial t} = \alpha^2 \frac{\partial^2 u}{\partial x^2} \tag{7.1}$$

The fact that the coefficient is α^2 is to emphasize that it *must* be a positive quantity. However, this is usually understood, and this formalism is often not maintained in practice. Unlike the wave equation (below), α^2 has physical significance (it is the heat or mass diffusion constant), but α (not squared) does not have a direct physical interpretation.

- **The wave equation**. Another of the classical equations studied in introductory PDEs. It is a hyperbolic equation, and represents phenomena that have *wave like* behavior. There are actually two wave equations studied in introductory PDEs. The first is the first-order wave equation. In one space and one time dimension, (x,t) it takes the form

$$\frac{\partial u}{\partial t} = -v_0 \frac{\partial u}{\partial x} \tag{7.2}$$

This equation describes the translation of an initial disturbance (initial condition) rectilinearly at velocity v_0. The initial condition translates without changing shape.
The second form of the wave equation is one that is more commonly thought of as *the* wave equation. In one space and one time dimension, (x,t) it takes the form

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2} \tag{7.3}$$

系数为c^2 这一事实是为了强调它必须 是一个正值。奇怪的是，与热传导/扩散方程不同，在实际应用中通常会保留平方项。这在一定程度上是因为量c具有物理意义，它可以解释为波速的一种测度。

• **拉普拉斯方程**。在入门课程中学习的经典偏微分方程中，拉普拉斯方程是涉及两个或更多自变量（空间维度）的椭圆型方程的原型。热量/扩散类型方程的稳态通常会得到一个拉普拉斯方程。拉普拉斯方程在二维空间中的标准形式为

$$K\frac{\partial^2 u}{\partial x^2}+K\frac{\partial^2 u}{\partial x^2}=0 \tag{7.4}$$

拉普拉斯方程以皮埃尔-西蒙·拉普拉斯侯爵（1749年3月23日–1827年3月5日）命名。他是一位著名的法国学者，在包括工程学、数学、统计学、物理学、天文学和哲学在内的众多领域都有贡献。如果拉普拉斯方程的右端由自变量的某个函数指定，那么该方程就称为泊松方程。该方程以西梅翁·丹尼斯·泊松男爵（1781年6月21日–1840年4月25日）命名。事实证明，泊松是拉普拉斯的热忱学生，而拉普拉斯认为泊松实际上就像自己的儿子一样。

• **调和函数**。拉普拉斯方程的解称为调和函数。调和函数具有一些独特的性质，值得我们专门加以研究。

• **叠加原理**。线性问题在数学中具有特殊地位——由于线性，一个问题的两个解可以相加，所得结果仍是该问题的一个解。这一点特别有用，例如，可以将一个复杂问题分解为两个较简单的问题，而这两个简单问题的解之和等价于原来的复杂问题。然后，可以将这两个较简单问题的解线性叠加，从而得到一个对复杂问题有效的解。本章末尾将对此进行更明确的讨论。

7.3 热方程/扩散方程

首先，我们将以无处不在的热传导/扩散方程为例。抛物型热传导/扩散方程可能是工程学和物理学中最有用的方程之一，可用于描述种类繁多的现象，从质量的扩散（显然如此！）到流体中由黏性剪切引起的动量扩散（牛顿黏性定律与动量守恒方程相结合）。在一维空间中，仅由（线性）扩散引起的质量平衡，$u(x,t)$，其形式为

$$\frac{\partial u}{\partial t}=\alpha^2\frac{\partial^2 u}{\partial x^2},\ x\in[0,\ell] \tag{7.5}$$

Here, we have used α^2 as the coefficient to remind us that this equation is a heat/diffusion equation only for cases where the coefficient is *positive*. While one can formulate an "backward" heat/diffusion equation, such an equation has some technical challenges, not the least of which is that it violates the second law of thermodynamics! (It is interesting to note, however, that the backward heat equation does have uses in inverse modeling of data). The domain length is ℓ, although later we will use $\ell=1$ to simplify the initial examples.

如上一章所述，这是一个二阶方程，且有两个自变量。因此，要解决这个问题，我们需要在空间中进行两次"积分"，并在时间上进行一次"积分"。这也意味着，若要确定唯一解，就需要给出两个边界（空间固定）条件和一个初始（时间固定）条件。对于此例，我们将研究 $x\in[0,1]$ 和 $t\geq 0$ 上的解，其中时间最初从 $t=0$ 开始计量。辅助条件为定义域两端的齐次狄利克雷条件，并指定初始条件的一般函数。从物理角度考虑，我们对该问题施加两个额外约束：(1)

$u(x,t)$ 的解必须在所有时间内保持有界，并且（2）所有解必须表示总质量为有限且常数的量（我们通过要求初始条件的傅里叶级数存在这一稍强条件来施加该要求）。于是，整个问题，包括物理约束，可以写成

$$\frac{\partial u}{\partial t} = \alpha^2 \frac{\partial^2 u}{\partial x^2}, \text{ for } x \in [0,\ell] \tag{7.6}$$

B.C. 1	$u(0,t) = 0$	(7.7)
B.C. 2	$u(\ell,t) = 0$	(7.8)
I.C. 1	$u(0,x) = \varphi(x)$, for $x \in [0,1]$	(7.9)
约束条件 1	Solutions stay bounded as $t \to \infty$	(7.10)
约束条件 2	初始条件具有有效的傅里叶级数表示	(7.11)

这里，“I.C.”代表边界条件，“I.C.”代表初始条件。请注意，最后一个条件表示问题物理学所施加的约束条件。我们不认为扩散问题会无限增长；事实上，我们预期的是相反的结果。扩散类问题往往会使质量在空间中扩散开来，从而形成尽可能均匀的质量分布，并在长时间极限下仍满足所施加的边界条件。虽然这一约束条件的必要性可能并不明显，但它确实符合我们对所涉及物理过程的理解。稍后我们将发现，这一约束条件使我们能够从所有数学上可能的解的集合中提取出具有物理意义的解。数学上可能的解并不总是代表物理上相关的解！

在继续介绍分离变量法之前，我们先简要岔开一下，确定热方程的一个物理性质，该性质稍后会很有用。这一性质适用于具有齐次边界条件的热方程。

7.3.1 A Property of the Heat/Diffusion Equation: Decreasing Variance

In thermodynamics, we learn that the entropy of a sequence of states for isolated system must always increase (or not change). We can consider a 1-space dimension heat problem to be isolated in the sense that zero Dirichlet or Neumann conditions represent the appropriate conditions. In the case of the Dirichlet conditions, the assumption is that the system has heat reservoirs that are contain substantial more heat energy than does the medium of interest (e.g., typically for heat problems in 1-dimension, one considers the transport of heat along an insulated rod between two reservoirs fixed at constant temperature.) Now, consider the case where we have homogeneous Dirichlet or Neumann boundary conditions (i.e., regardless of which condition we choose, the value is zero at the boundaries $x = 0$ and $x = \ell$). Now, consider multiplying the heat/diffusion equation by the independent variable u. This yields

$$u\frac{\partial u}{\partial t} = \alpha^2 u \frac{\partial^2 u}{\partial x^2}, \text{ for } x \in [0,\ell] \tag{7.12}$$

With a little manipulation, this can easily be put in the form

$$\frac{1}{2}\frac{\partial u^2}{\partial t} = \alpha^2 \frac{\partial}{\partial x}\left(u\frac{\partial u}{\partial x}\right) - \alpha^2 \frac{\partial u}{\partial x}\frac{\partial u}{\partial x}, \text{ for } x \in [0,\ell] \tag{7.13}$$

(You can confirm this by expanding the derivatives on the left and right sides of the expression). Integrating this result over $x \in [0,1]$ gives

$$\int_{x=0}^{x=\ell} \frac{1}{2}\frac{\partial u^2}{\partial t}\,dx = \alpha^2 \int_{x=0}^{x=\ell} \frac{\partial}{\partial x}\left(u\frac{\partial u}{\partial x}\right)dx - \alpha^2 \int_{x=0}^{x=\ell} \frac{\partial u}{\partial x}\frac{\partial u}{\partial x}, \text{ for } x \in [0,\ell] \tag{7.14}$$

注意，右侧第一项很容易积分，得到

$$\alpha^2 \left(u\frac{\partial u}{\partial x}\right)\Bigg|_{x=0}^{x=\ell} \tag{7.15}$$

由于假设边界条件是齐次的，因此该项恒等于零。由此得到

$$\frac{\partial}{\partial t}\int_{x=0}^{x=\ell} u^2\,dx = -2\alpha^2 \int_{x=0}^{x=\ell} \frac{\partial u}{\partial x}\frac{\partial u}{\partial x}, \text{ for } x \in [0,\ell] \tag{7.16}$$

这是一个有趣的表达式；它表明，量u^2 随时间的变化率始终为负（即不断减小）。在对热方程/扩散方程进行更深入的数学探讨时，这类分析是证明最大值原理 的第一步。该原理指出，因变量u的最大值总是出现在以下位置之一：（1）定义域内部的初始条件处，或（2）边界上。假设系统的平均温度为$u = 0$。在这种情况下，u^2 表示系统中温度（或浓度）的方差。因此，理解方程 7.16 的一种方式是：它表明系统的方差 始终在减小；换句话说，随着时间推移，系统正变得更加均匀。这与系统随时间从一种状态演变到另一种状态时熵不断增加的概念完全一致。这也符合我们对热量或质量传输现象的观测和直觉。我们知道，热量往往会在物体中扩散，化学物质也往往会扩散，直到其空间梯度消散。总的来说，我们上面概述的简短证明说明，热量和质量倾向于在空间中自发扩散，而不是自发地集中在空间中。这是热方程/扩散方程的一项有用性质。它不仅让我们了解该方程的整体行为（并且这种行为与我们对宇宙的观测相符），还能作为一项有用的约束条件，帮助确定问题的正确数学解。我们将在下一节看到这一点。

7.3.2 热方程/扩散方程的分离变量法

要开始使用分离变量法（SOV），我们首先假设解具有以下形式。为简化计算，我们取 $\ell = 1$；根据第 3 章（傅里叶级数）中确立的原则，转换为其他定义域长度也很简单。

$$u(x,t) = X(x)T(t) \tag{7.17}$$

Viewing this empirically, we will simply *try* this solution to see if it leads to a useful result. Toward that goal, we substitute Eq. (7.17) into the governing differential equation to get

$$\frac{\partial(X(x)T(t))}{\partial t} = \alpha^2 \frac{\partial^2(X(x)T(t))}{\partial x^2} \tag{7.18}$$

由于$X(x)$和$T(t)$都只是单个变量的函数，我们可以采用以下$X' \Leftrightarrow dX/dx$、$X'' \Leftrightarrow d^2X/dx^2, T' \Leftrightarrow dT/dt$来简化符号表示。这样，原偏微分方程可写成

$$X(x)T'(t) = \alpha^2 X''(x)T(t) \tag{7.19}$$

两边除以 $\alpha^2 X(x)T(t)$，得到

$$\underbrace{\frac{T'(t)}{\alpha^2 T(t)}}_{\substack{\text{Function of}\\ \text{only } t\\ G(t)}} = \underbrace{\frac{X''(x)}{X(x)}}_{\substack{\text{Function of}\\ \text{only } x\\ F(x)}} \tag{7.20}$$

请注意，这里出现了一种不寻常的情况：对于这种形式的方程，左边显然只是t 的函数（我们可以称之为$G(t)$），而右端只是x 的函数（我们可以称之为$F(x)$）。简而言之，这意味着对于$x \in (0,1)$和$t > 0$的所有可能组合，都满足以下条件：

$$G(t) = F(x), \text{ for any possible combinations of } x \in (0,1) \text{ and } t > 0 \tag{7.21}$$

乍看之下，这或许并不特别惊人，但实际上施加了一个非常强的条件。无论函数$X(x)$和$T(t)$是什么，它们都必须使方程（7.19）两边相等。由于左边只取决于t，右端只取决于x，而且对于定义域（x, t）中任意选取的一对值，这个方程都必须成立，因此*only*可能的选择是

$$G(t) = F(x) = Constant \tag{7.22}$$

在下文中，我们将这个常数称为 $\pm\lambda^2$。再次注意，我们将常数值平方，以此规定该常数本身为正。在此处，我们尚未为该常数指定符号，因此将它与 $\pm$算子一同列出。

稍加思考便可看出，在这一阶段我们得到以下条件。

$$\frac{T'(t)}{\alpha^2 T(t)} = \frac{X''(x)}{X(x)} = \pm\lambda^2 \tag{7.23}$$

因此，实际上发生的是，这个问题中的时间和空间分量已经被分离为两个常微分方程，它们分别用彼此正交的变量表示。为了更清楚地理解这一点，我们可以将上面的表达式改写为

$$\frac{T'(t)}{\alpha^2 T(t)} = \pm\lambda^2 \qquad\qquad \frac{X''(x)}{X(x)} = \pm\lambda^2 \tag{7.24}$$

或者等价地，写成以下两个独立的常微分方程

$$\frac{dT}{dt} = \pm\lambda^2\alpha^2 T(t) \qquad\qquad \frac{d^2X}{dx^2} = \pm\lambda^2 \tag{7.25}$$

这分别是两个一阶和二阶齐次常微分方程。我们在第2章中学习过这两种情况，它们的解相对简单。这里有必要对符号表示作一些说明。虽然原方程涉及偏导数，但请注意，此处我们采用了常导数的惯用符号。这在一定程度上源于这一惯例的悠久历史，有时也会引起一些混淆。不过，稍微思考一下这些导数的定义，就能帮助澄清这一点。根据构造，$X(x)$只依赖于自变量x，因此不存在其他可针对该函数求导的变量。所以，严格来说，这里定义的是常导数。对时间变量中的常微分方程也可以作类似的解释。这样，我们就把一个含有x和t偏导数的偏微分方程，表示成了两个常微分方程的组合，其中时间和空间的影响彼此独立地演变。好消息是，我们已经熟悉常微分方程。

为了得到偏微分方程的一个解，我们现在需要求解这两个常微分方程。让我们先考虑其中较简单的一个——关于时间的一阶常微分方程。这个问题可以通过简单的分离变量法求解。

$$\frac{dT}{dt} = \pm\lambda^2\alpha^2 T(t) \tag{7.26}$$

or,

$$\frac{dT}{T} = \pm\lambda^2\alpha^2\, dt \tag{7.27}$$

Integrating both sides over $t \in [0,t] \Rightarrow T \in [T(0), T(t)]$ gives

$$\int_{T(0)}^{T(t)} \frac{dT}{T} = \int_0^t \pm\lambda^2\alpha^2\, dt \tag{7.28}$$

The solution is the conventional exponential one of the form

$$\ln\left(\frac{T(t)}{T(0)}\right) = \pm\lambda^2\alpha^2 t \tag{7.29}$$

Or, solving for $T(t)$

$$T(t) = T(0)\exp[\pm\lambda^2\alpha^2 t] \tag{7.30}$$

Technically, both the positive and negative components of this equation represent *solutions* to the original ODE. However, we know from physical reasoning that the concentrations in a pure diffusion problem have their maximum somewhere in the initial state of the system ($t = 0$) (including the boundaries, assuming that the boundaries are steady). In fact, this is exactly what we proved to ourselves in Section 7.3.1. This is important, because it indicates that the positive exponential root of the time component of the ODE is not a valid one because it leads to temperatures or concentrations that grow arbitrarily large as time increases, which is not consistent with the proof that the variance must decrease with increasing time. Thus, we have explicitly that only the negative exponential is *physically* relevant

$$T(t) = T(0)\exp[-\lambda^2\alpha^2 t] \tag{7.31}$$

This represents the solution to one of the two ODEs we have generated. The second is given by

$$\frac{d^2X}{dx^2} + \lambda^2 = 0 \tag{7.32}$$

where we have adopted the sign imposed by the solution to the ODE in time. This second-order ODE has the characteristic function

$$s^2 + \lambda^2 = 0 \tag{7.33}$$

替换 tuting this into the quadratic formula immediately gives the two possible roots for the equation

$$s_1 = -\lambda\, i \qquad\qquad s_2 = +\lambda\, i \tag{7.34}$$

提醒一下，这表示常系数齐次二阶常微分方程解的情形3；因此，两个根是共轭复数。在第2章中，我们发现，这种情形的解由以下公式给出。

$$X(x) = C\cos(\lambda x) + D\sin(\lambda x) \tag{7.35}$$

Where C and D must be determined by two ancillary conditions. Before proceeding, note that the general solution (before any anciallry conditions are imposed) is given by

$$u(x,t) = T(t)X(x) = T(0)\exp[-\lambda^2\alpha^2 t]\,[C\cos(\lambda x) + D\sin(\lambda x)] \tag{7.36}$$

$T(0)$、C和 D均为待定常数；由于 $T(0)$ 同时乘以 C和 D，显然这三个常数并非相互独立。事实上，实际上只有两个相互独立的常数。因此，解可以写成

$$u(x,t) = T(t)X(x) = \exp[-\lambda^2\alpha^2 t]\,[A\cos(\lambda x) + B\sin(\lambda x)] \tag{7.37}$$

where $A = C \times T(0)$ and $B = D \times T(0)$. In the future, we will use this form as the general solution for the heat/diffusion PDE. Note that this kind of reduction in the number of constants will appear in applications of the SOV method to other kinds of PDEs. For convenience, we will often reduce the equations in the future by *specifying* one of the constants to be unity, which leaves the second constant to be set via the ancillary conditions. Thus, for the discussion here, we can make the statement $T(0) = 1$, and this is equivalent to removing the extra (unnecessary) constant.

Continuing forward, we note that for this particular case the two ancillary conditions that we need are given by Eqs. (7.7)-(7.8), which we write here as

$$B.C.\ 1 \qquad T(t)X(0) = 0 \tag{7.38}$$
$$B.C.\ 2 \qquad T(t)X(1) = 0 \tag{7.39}$$

Because these boundary conditions hold for all time, and because $T(t)$ is not, in general, identically zero, we must then have the following two equivalent conditions on the spatial boundaries

$$B.C.\ 1 \qquad X(0) = 0 \tag{7.40}$$
$$B.C.\ 2 \qquad X(1) = 0 \tag{7.41}$$

Imposing the first of these on the solution given by Eqs. (7.35) gives

$$A\cos(0) + B\sin(0) = 0 \tag{7.42}$$

or

$$A = 0 \tag{7.43}$$

Thus, we at least know now that, whatever the solution to the problem, it involves a decaying exponential in space, and a sine function. To eliminate the second unknown, we impose the second of the two boundary conditions

$$B\sin(\lambda) = 0 \tag{7.44}$$

尽管该问题的解可能是 $B = 0$，但我们注意到，这将得到 $A = B = 0 \Leftrightarrow u(x,t) \equiv 0$。这个解被称为平凡 或零 解。尽管它确实代表该问题的一个有效解，但它仅 对应于初始条件为零的问题。换句话说，它是初始条件和边界条件全为零的问题的解，因此不存在空间导数。换句话说，初始条件和边界条件为零，于是在所有时间内永远什么也不会发生。由此可以看出，为什么它可能被称为平凡解。

因此，一般来说，我们不可能让A 和B 同时等于零。由此可知，必须有$B \neq 0$，因此我们可以将式 (7.44) 的两边除以B，得到

$$\sin(\lambda) = 0 \tag{7.45}$$

奇怪的是，虽然乍看之下这似乎没有太大改进，但稍加思考便会发现，它有一个非平凡解。事实上，它有无限多个解！我们必须有这个解

$$\lambda = 0,\ \pi,\ 2\pi,\ 3\pi \ldots \tag{7.46}$$

这表明 λ 可以是 π 的任意整数倍。注意：由于正弦函数 $\sin(-\lambda) = -\sin(\lambda)$ 具有奇对称性，我们尚未考虑 λ 的负值。这里出现的负号可以

be adsorbed ultimately into the constant B, which remains undetermined at this juncture. A similar argument can be made for solutions where the cosine function was selected, so we will not repeat this argument in the future. Thus, the application of the two boundary conditions gives us not just a single solution, but actually an *infinite number* of possible solutions, one for each value of n. Because there exists an infinite number of solutions, then there must also be a corresponding infinite number of multiplicative constants, B_n. This is easy to see by noting that if $\sin(n\pi x)$ is a solution, then $B_n \sin(n\pi x)$ is also a solution.

At this juncture, we have the following:

1. 我们已将解分解为 u(x,t) = T (t)X(x)。2. 我们已证明，可以将这一分解代回控制方程偏微分方程，得到两个涉及常数 $-\lambda^2$ 的常微分方程。3. 可以求出两个解 T (t) 和 X(x)，它们分别表示为 T (t) = exp[$-\lambda^2\alpha^2$t] 和 X(x) =B sin(λx)，其中必须满足 λ = nπ，且 n = 0，1，2，3.....。函数 X(x) 是利用两个空间辅助（边界）条件确定的。初始条件尚未使用。4. 关于 X(x) 的问题解表明，我们有无限多个可能的解，每个 n 值对应一个解；每个这样的解都（可能）对应一个不同的剩余常数值 B_n。

Using this information, apparently our solution at this point in the analysis is given by the infinite set of possible independent solutions, $u_n(x,t)$

$$\begin{aligned} u_n(x,t) &= T(t)X(x) \\ &= \exp[-n^2\pi^2\alpha^2 t]B_n \sin(n\pi x) \end{aligned} \tag{7.47}$$

Clearly, if the u_n represent independent solutions, then so does its sum. Thus, the most general solution we can write at this point in our analysis is as follows

$$u_n(x,t) = \sum_{n=0}^{\infty} \exp[-n^2\pi^2\alpha^2 t]B_n \sin(n\pi x) \tag{7.48}$$

We have an unusual seeming situation here. We have a single remaining condition (the initial condition), but we have an infinite number of constants, B_n, to determine. While initially this may seem intractable, we should note that, unlike the constant value boundary conditions, the initial condition contains in general a *function*. Thus, in some sense, the initial condition contains more information than the two boundary conditions. To make more sense of this, we can apply the initial condition to see what happens. Recall, the initial condition states

$$I.C.\ 1 \qquad\qquad u(0,x) = \varphi(x) \tag{7.49}$$

Thus, at the initial time, $t = 0$, we must have the following

$$u(x,0) = \sum_{n=0}^{\infty} \exp[-n^2\pi^2\alpha^2(0)]B_n \sin(n\pi x) \tag{7.50}$$

or,

$$\varphi(x) = \sum_{n=0}^{\infty} B_n \sin(n\pi x) \tag{7.51}$$

We assumed at the start of the analysis that $\phi(x)$ is a function that meets the requirement that its Fourier series exists (see Chapter 3); thus, this result indicates that the B_n are, for this particular case, *exactly the amplitude constants for the Fourier sine series representation of* $\phi(x)$! In some ways this result should not be surprising. However, it is interesting to note that the SOV method automatically generates the Fourier sine series for the initial condition. In fact, this analysis is close to how Fourier originally found the Fourier series in the first place (he was interested in finding solutions to the heat equation). We have already studied Fourier series in Chapter 3, and using that information, we can simply write down the value for B_n by

$$B_n = 2\int_{x=0}^{x=1} \varphi(x)\sin(n\pi x)\,dx \tag{7.52}$$

This, then, formally concludes the solution to the problem. Note that for $n = 0$, the resulting contribution to the sum in Eq. (7.50) is zero. For this reason, it is customary to begin the numbering for the sine series at $n = 1$. The final result for the heat/diffusion problem with two homogeneous Dirichlet conditions on the interval $x \in [0,1]$ is

$$\boxed{u_n(x,t) = \sum_{n=1}^{\infty} \exp[-n^2\pi^2\alpha^2 t] B_n \sin(n\pi x)} \tag{7.53}$$

其中

$$\boxed{B_n = 2\int_{x=0}^{x=1} \varphi(x)\sin(n\pi x)\,dx} \tag{7.54}$$

Now that we have seen how the method works, it will be useful to see an example worked out in detail. Note that the general method for SOV outlined above also works for the linear wave and Laplace equations in two variables, with very minor modifications. We will explore those solutions after an example.

Example 7.1 (Separation of Variables: The Heat/Diffusion Equation with Dirichlet Boundary Conditions).

Consider the following equation on the interval $x \in [0,1]$.

	$\frac{\partial u}{\partial t} = \alpha^2 \frac{\partial^2 u}{\partial x^2},\ x \in [0,1]$
B.C. 1	$u(0,t) = 0$
B.C. 2	$u(1,t) = 0$
I.C. 1	$u(0,x) = \delta(x - \frac{1}{2})$
约束条件	Solutions stay bounded as $t \to \infty$
约束条件	The initial condition has a valid Fourier series representation

Rather than reproducing all of the details above, the SOV method will be conducted here in a summary fashion. This is done to help build familiarity with the method; eventually, once the details are well understood, one can "jump" to later stages of the solution with confidence, and then complete the solution from there. To get us started, we recall that the substitution of $u(x,t) = T(t)X(x)$ yields two, separated equations of the form

$$T(t) = 1 \cdot \exp(-\lambda^2 \alpha^2 t)$$
$$X(x) = A\cos(\lambda x) + B\sin(\lambda x)$$

Where, here, recall that we are arbitrarily setting the constant in the first ODE to unity because it is essentially a redundant constant. Thus, the general solution, before applying any other conditions, is

$$u(x,t) = \exp(-\lambda^2 \alpha^2 t)[A\cos(\lambda x) + B\sin(\lambda x)]$$

Application of the first boundary condition at $x = 0$ rapidly shows us that we must have

$$0 = \exp(-\lambda^2 \alpha^2 t)[A\cos(0) + B\sin(0)]$$

Whatever the value of the exponential is, we know it is not zero for any finite time. Thus, we can divide both sides by the exponential, and also evaluate $\cos(0)$ and $\sin(0)$. This gives

$$0 = A + 0 \cdot B \to A = 0$$

The cosine functions do not contribute to the solutions of this problem. Thus, the solutions must be of the form

$$u(x,t) = \exp(-\lambda^2 \alpha^2 t) B\sin(\lambda x)$$

Applying the second boundary condition gives

$$0 = B\sin(\lambda)$$

The constant B cannot be zero unless the initial condition is also zero (which it is not). Thus we must have

$$0 = \sin(\lambda)$$

or,

$$\lambda = \pi,\ 2\pi,\ 3\pi \ldots = n\pi, \text{ for } n \in \mathbb{N}$$

The solution is now (recognizing that the most general solution must be a linear combination of all possible independent solutions)

$$u(x,t) = \sum_{n=1}^{\infty} \exp(-n^2\pi^2\alpha^2 t) B_n \sin(n\pi x)$$

利用初始条件，我们得到

$$\delta(x - \tfrac{1}{2}) = \sum_{n=1}^{\infty} B_n \sin(n\pi x)$$

这表明，系数 B_n 正是狄拉克δ函数的傅里叶级数中对应的振幅。回想一下，虽然狄拉克δ函数并不是一个<escape></*escape*>真正的<*escape*></escape>函数，但我们可以将其视为傅里叶级数有意义的对象来处理（这一点已在第4章中讲过）。狄拉克δ函数的系数特别容易确定。

$$\begin{aligned} B_n &= 2\int_{x=0}^{x=1} \delta(x - \tfrac{1}{2}) \sin(n\pi x)\, dx \\ &= 2\sin\left(\frac{n\pi}{2}\right) \end{aligned}$$

因此，我们的最终解形式如下。

$$u(x,t) = \sum_{n=1}^{\infty} \exp(-n^2\pi^2\alpha^2 t) \left[2\sin\left(\frac{n\pi}{2}\right)\right] \sin(n\pi x)$$

这里揭示了解的一个重要特征。n 的每个值都对应正弦函数的一个特定频率。每个这样的函数都由衰减指数函数项调制，该项是n^2的函数。这表明，解中的高频分量比低频分量衰减得快得多。从物理角度来看，这确实说得通。$x \in [0,1]$ 中的高频波，是指波峰很多且彼此相距很近的函数。只需少量扩散，这些波峰就会充分混合，并从解中基本消失。对于解中的低频分量，在相同区间内波峰少得多，而且相对而言，每个波峰都离相邻波峰较远。因此，扩散过程需要更长时间才能使这些波峰混合，因为它们扩散所需的特征距离要大得多！我们将在另一个示例中进一步考察这种行为。

供参考，图 7.1 展示了解在不同时刻的图像。

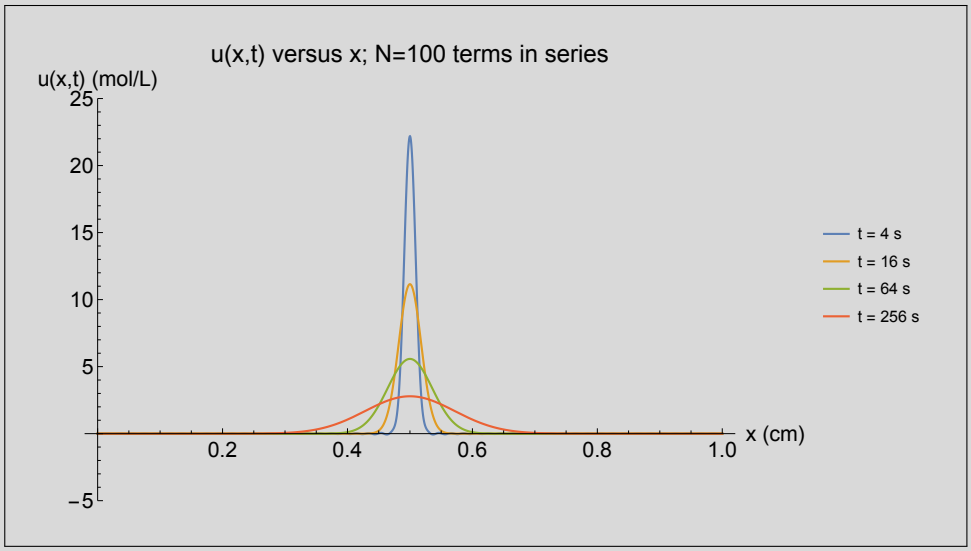


Fig. 7.1: The solution to the diffusion problem for several selected times. For this plot, $x \in [0,1]$ cm, $0 < t < 256$ s, $\alpha^2 = 1 \times 10^{-5}$ cm^2/s. For this figure, the number of terms in the series was truncated at $N = 100$.

7.4 The Wave Equation

如上一章所述，波动方程描述了多种现象，包括水波（例如海洋中的重力波）、振动和压缩波（例如声音传播）。回想一下，一维空间中的波动方程需要两个空间附加条件和两个时间附加条件。这里给出的狄利克雷条件示例，例如，可以描述两端固定的拨弦所产生的波动。同样，为了简洁起见，我们假设区间为$x \in [0,1]$，且时间从$t = 0$开始计量。具体而言，我们可以将此例视为u表示两端固定的弦在时间和空间上的位移。其他现象的波动方程（例如，对于压缩波，u表示可压缩介质的密度）中，u也有同样容易理解但物理意义不同的含义。请注意，在我们所考虑的形式中，没有阻尼项，因此经典波动方程（后面将会看到）不允许系统发生能量损失。这是一种近似，我们将在后文中进一步详细讨论。

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2}, \text{ for } x \in [0,1] \tag{7.55}$$

B.C. 1	$u(0,t) = 0$	(7.56)
B.C. 2	$u(1,t) = 0$	(7.57)
I.C. 1	$u(0,x) = \varphi_0(x), \text{ for } x \in [0,1]$	(7.58)
I.C. 2	$\frac{\partial u(0,x)}{\partial t} = \varphi_1(x), \text{ for } x \in [0,1]$	(7.59)
约束条件 1	The initial condition has a valid Fourier series representation	(7.60)
Constraint 2	The initial condition is at least C^1 continuous in time	(7.61)

This represents our starting point for our analysis of the wave equation using SOV. Note that, for this case, u represents the displacement from zero for the string. Thus the two initial conditions (required because of the second-order derivative in time) represent the initial displacement and the initial velocity for the string.

7.4.1 波动方程的一个性质：能量守恒

Before continuing, it is helpful to consider the physical context of this equation. If we think of u as being the *displacement* of a vibrating string fixed at both ends (such as a guitar string), then this equation describes how the displacements of such a string evolve in time. If one imagines, for example, using a strobe light to capture instantaneous images at times t_i, $i = 1, 2, 3, 4, 5$ of the vibrating string, the solutions $u(x,t_i)$ would represent the state of the displacement field (from zero displacement) at the the five indexed times.

Our derivation of the wave equation did not account for any kinds of energy loss; thus, we would expect the resulting model to be an approximation, valid for early times, that did not include any energy dissipation mechanism (since we did not model one!) Thus, the total energy in the string remains constant at all times. If u is the displacement field, and the one-dimensional density is $\rho = M/\ell$ (where ℓ is the length of the string, equal to unity for this example), then it is not hard to show that the following equation defines the energy per unit length in the string for all (x,t) (this is not shown here, but is available as a problem to work out)

$$e(x,t) = \rho \left(\frac{\partial u}{\partial t}\right)^2 + \rho c^2 \left(\frac{\partial u}{\partial x}\right)^2 \tag{7.62}$$

相应地，弦的总能量，E_0，由下式给出

$$E_0 = \rho \int_{x=0}^{x=\ell} \left(\frac{\partial u}{\partial t}\right)^2 + c^2 \left(\frac{\partial u}{\partial x}\right)^2 dx \tag{7.63}$$

By definition, we must have that E_0 is a constant. Thus, the energy in the string is conserved for all times. This provides a third, important constraint for our problem. Whatever the solution of the problem is, it must at least obey the following conservation of energy principle (imposed independently by the physical interpretation of the system)

$$\textit{Constraint } 3 \qquad \begin{array}{c}\text{The energy } E_0 \text{ remains} \\ \text{constant for all time}\end{array} \tag{7.64}$$

With this property established, we can continue the analysis. We will find that this bit of *physical* analysis constructed ahead of time will make the solution to the problem substantially less complex! Note that all real physical systems have some kind of loss of energy (in accordance with the second law of thermodynamics), so the wave equation must represent an appropriate *idealization* of systems that represents an approximation.

无阻尼波动方程的能量守恒对其解具有一些重要的物理意义。特别是，它表明波动方程不存在稳态解（平凡解除外，即初始条件处处为零，并且永远保持为零）。虽然不存在稳态解，但波动方程确实存在周期性解。这意味着，如果在定义域中的某个点固定观察位置，并观测因变量的一个值（例如波高），那么该值将再次出现，而且实际上会出现无限多次！周期性是一种有趣的行为特征，广泛存在于许多物理系统中，从摆的运动到潮汐运动皆是如此。从这一角度来看，波动方程符合我们的物理直觉（至少在一定程度上如此！）

In all real systems, there is a gradual loss of energy (damping) in media that support waves, so ultimately there can be a steady-state solution for systems where energy is lost.

7.4.2 Separation of Variables for the Wave Equation

As with the case for the heat/diffusion equation, we start by positing that a solution may exist of the form

$$u(x,t) = T(t)X(x) \tag{7.65}$$

and then substitute this form into the governing PDE to see if solutions of this form are, in fact, possible (and useful!). Following the example of the heat/diffusion equation, we can very quickly get to the following expression

$$\frac{T''(t)}{c^2T(t)} = \frac{X''(x)}{X(x)} \tag{7.66}$$

As previously, one now makes the argument that independent functions of t and x can only be equal to one another for all possible pairs, (x,t), if they are equal to the same constant. By convention (and for distinction among the three basic problems that we will solve by SOV), we set the contant equal to ω^2. Thus, we have

$$\frac{T''(t)}{c^2T(t)} = \frac{X''(x)}{X(x)} = \pm\omega^2 \tag{7.67}$$

Or, equivalently

$$T''(t) = \pm c^2\omega^2 T(t) \tag{7.68}$$

$$X''(x) = \pm\omega^2 X(x) \tag{7.69}$$

As with the case of the heat/diffusion equation, we can start our analysis with the time variable. Because we have not yet assigned a sign to the separation constant, ω, we have three possible solutions. Thus, we need

以考察分离常数分别等于 $-\omega^2$、 $\omega^2=0$ 和 $+\omega^2$ 的情况。换句话说，我们需要考察以下常微分方程。为了使这些方程更易于求解，我们已在右侧列中计算出特征方程的根。

$$(a) \qquad -\omega^2 \Rightarrow \qquad T''(t)+c^2\omega^2 T(t)=0 \qquad s_1=c\omega,\ \ s_2=-c\omega \tag{7.70}$$

$$(b) \qquad 0 \ \Rightarrow \qquad T''(t)=0 \qquad s_1=s_2=0 \tag{7.71}$$

$$(c) \qquad +\omega^2 \Rightarrow \qquad T''(t)-c^2\omega^2 T(t)=0 \qquad s_1=ic\omega,\ \ s_2=-ic\omega \tag{7.72}$$

因此，相应的解如下

$$(a) \qquad -\omega^2 \Rightarrow \qquad T(t)=\alpha\exp(c\omega t)+\beta\exp(-c\omega t) \tag{7.73}$$

$$(b) \qquad 0 \ \Rightarrow \qquad T(t)=\alpha+\beta t \tag{7.74}$$

$$(c) \qquad +\omega^2 \Rightarrow \qquad T(t)=\alpha\cos(c\omega t)+\beta\sin(c\omega t) \tag{7.75}$$

尽管这些都是波动方程在数学上可能的解，但只有其中一个符合我们希望解具有常数且有限的能量这一要求。解 (a)（式 (7.73)）可以排除，因为无论我们为 α 和 β 选取什么参数（除非二者均为零，此时得到平凡解），所得解都会随着 $t\to\infty$呈指数增长或指数衰减。因此，能量无法保持恒定。对于解 (b)（式 7.74），当 $\beta\neq 0$ 时，式 (7.63) 给出的总能量会随时间线性增长（或衰减）。然而，如果 $\beta=0$，那么解将完全不依赖时间；对于给定的边界条件，这同样只对应于平凡解。因此，只剩下式 (7.75）可作为可能的解。虽然该解随时间振荡，但无论时间增长得多么久，它至少都是有界的。因此，此时我们采用该解作为唯一可行的解，并将注意力转向空间中的第二个常微分方程。请注意，这意味着我们采用分离常数由 $+\omega^2$ 明确给出的情形。

对于空间解，分离常数现在已有一个固定值。因此，解的形式只有一种可能

$$X''(x)=+\omega^2 X(x) \tag{7.76}$$

or,

$$X''(x)-\omega^2 X(x)=0 \tag{7.77}$$

根据上述分析，我们可以立即写出解为

$$X(x)=A\cos(\omega x)+B\sin(\omega x) \tag{7.78}$$

The solution as a whole to this stage is then given by

$$u\ (x,t)\ \ =T\ (t)\ X\ (x) \tag{7.79}$$

$$u(x,t)=[\alpha\cos(c\omega t)+\beta\sin(c\omega t)][A\cos(\omega x)+B\sin(\omega x)] \tag{7.80}$$

This solution at least seems consistent with the solution explored previously for the heat equation, except now instead of having spatial waves that decay in time, we have waves in both time and space. To make further sense out of this result, we can, as before, begin applying the ancillary conditions. To start, we can apply the two Dirichlet boundary conditions. The first of these gives

$$0 = [\alpha\cos(c\omega t) + \beta\sin(c\omega t)][A\cos(0) + B\sin(0)] \tag{7.81}$$

We have already determined that both α and β cannot be zero. Also, while the remainder of the time component of the solution may be zero for some times (since it is the sum of two oscillating functions that are each sometimes zero), they are not zero for *all* times, which is what is required by the boundary condition. This means that we must have

$$[A\cos(0) + B\sin(0)] = 0 \tag{7.82}$$

or,

$$A = 0 \tag{7.83}$$

The complete solution at this juncture is now

$$u(x,t) = [\alpha\cos(c\omega t) + \beta\sin(c\omega t)][B\sin(\omega x)] \tag{7.84}$$

indicating that the spatial component of the solution will ultimately be represented by only sine functions. Applying the second boundary condition (and applying the same logic regarding the time-component of the solution) gives the result

$$[B\sin(\omega)] = 0 \tag{7.85}$$

This is identical to the problem for the heat equation, and we find that the only possibility is that ω take on the values

$$\omega = \pi,\ 1\pi,\ 3\pi\ldots = n\pi,\ \ n = 1,\ 2,\ 3,\ldots \tag{7.86}$$

这里请记住，4 这一情形 $n = 0$ 确实能够得到该问题的一个解，但它对应的是零解（正弦$(0\pi x) = 0$），因此按照惯例，我们不将其纳入最终结果。

We have made some significant progress here. If we substitute the possible values for ω into our solution now, we have

$$u_n(x,t) = [\alpha_n\cos(cn\pi t) + \beta_n\sin(cn\pi t)][B_n\sin(n\pi x)],\ \ n = 1,\ 2,\ 3,\ldots \tag{7.87}$$

Note that we have included a subscript n to the solution variable, u_n, because we actually have an infinite number of them now, one for each value of n. Note that we have again a situation where the time and space constants are multiplied, and it should be clear that we can set $B_n = 1$ with no loss of generality (because $\alpha_n B_n$ forms a single new constant, and the same is true for $\beta_n B_n$; rather than renaming constants, we can just set $B_n = 1$, and still have two undetermined constants in the end.) As with the heat equation, we also

conclude that the most general possible solution at this point is just the linear superposition of all of the possible solutions

$$u(x,t) = \sum_{n=1}^{\infty} [\alpha_n \cos(cn\pi t) + \beta_n \sin(cn\pi t)] \sin(n\pi x) \tag{7.88}$$

From our previous experience with heat/diffusion problem, it is possible to guess at how things will evolve from here. We have *two* ancillary initial conditions for this problem, and imposing those will generate series that allow us to reproduce the two initial conditions at the time $t = 0$. Imposing the second initial condition first, we find

$$u(x,0) = \sum_{n=1}^{\infty} \alpha_n \sin(n\pi x) \tag{7.89}$$

or,

$$\varphi_0(x) = \sum_{n=1}^{\infty} \alpha_n \sin(n\pi x) \tag{7.90}$$

Without doing any additional work, we note that this is exactly the sine series expansion of the function $\varphi_0(x)$. The constants α_n are thus known, and are given explicitly by

$$\alpha_n = 2 \int_{x=0}^{x=1} \varphi_0(x) \sin(n\pi x)\, dx \tag{7.91}$$

对于 ap plying the second boundary condition, we first need the time derivative of our proposed solution

$$\frac{\partial u}{\partial t} = \sum_{n=1}^{\infty} c\pi n[-\alpha_n \sin(cn\pi t) + \beta_n \cos(cn\pi t)] \sin(n\pi x) \tag{7.92}$$

And, for $t = 0$ we must have

$$\varphi_1(x) = \sum_{n=1}^{\infty} c\pi n \beta_n \sin(n\pi x) \tag{7.93}$$

如果令 $\gamma_n = c\pi n\beta_n$，我们得到

$$\varphi_1(x) = \sum_{n=1}^{\infty} \gamma_n \sin(n\pi x) \tag{7.94}$$

因此，该结果就是 $\varphi_1(x)$ 的傅里叶级数，其中

$$\gamma_n = 2 \int_{x=0}^{x=1} \varphi_1(x) \sin(n\pi x)\, dx \tag{7.95}$$

or, converting back to β_n

$$\beta_n = \frac{2}{c\pi n}\int_{x=0}^{x=1} \varphi_1(x)\sin(n\pi x)\,dx \tag{7.96}$$

因此，该问题的最终解为

$$\boxed{u(x,t) = \sum_{n=1}^{\infty}[\alpha_n \cos(cn\pi t) + \beta_n \sin(cn\pi t)]\sin(n\pi x)} \tag{7.97}$$

$$\boxed{\alpha_n = 2\int_{x=0}^{x=1} \varphi_0(x)\sin(n\pi x)\,dx} \tag{7.98}$$

$$\boxed{\beta_n = \frac{2}{c\pi n}\int_{x=0}^{x=1} \varphi_1(x)\sin(n\pi x)\,dx} \tag{7.99}$$

例 7.2（分离变量法：具有狄利克雷边界条件的波动方程 *s).*

吉他至少是大多数人都熟悉的乐器（尽管与理想情况下所希望的结果相比，吉他在工程学问题中的直接应用确实较少），因此它是理解波动方程背后直觉的良好载体。吉他两座琴桥之间的长度（有效弦长）约在 62.9 至 65.1 厘米之间，但目前制造的大多数原声吉他和电吉他采用 $L = 64$ 厘米会是一个很好的代表值。事实证明，按标准调弦（EADGBE）调音时，空弦吉他的弦振动频率相当明确，因为音阶与所使用的乐器无关！一份使用特定弦组的参考资料测得标准吉他调弦的频率 f 如下（所有数值单位为 1/s=赫兹=Hz）：E—82.41，A—110.00，D—146.83，G—196.00，B—246.94，E（高音）—329.63。对于本例，我们将考察 G 弦的运动，其频率为 196.00 Hz。回顾波长、 λ 与长度 L 之间的关系，即

$$\lambda = 2L \tag{7.100}$$

This is because the fundamental solution to the problem is represented by half of a full sine series (i.e., the sine function goes from zero, to a maximum, and back to zero in a distance L; this is only *half* of the full sine function, which has a negative component that has to be traversed before one full period of the sine function is expressed.) To compute the value of c, then, we need only recall the expression

$$c = f\lambda \tag{7.101}$$

$$= f2L = 196\ s^{-1} \times 2 \times 0.64\ m = 250.88\ m/s \tag{7.102}$$

	$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2},\ x \in [0,L]$
B.C. 1	$u(0,t) = 0$
B.C. 2	$u(L,t) = 0$
I.C. 1	$u(0,x) = 4\delta_0 \frac{x}{L}(1 - \frac{x}{L})$
I.C. 2	$\frac{\partial u}{\partial t} = 0$

其中 here δ_0 is the initial maximum displacement (with units of length) during a plucking of the string.

Rather than reproducing all of the details above, the SOV method will be conducted here in a summary fashion, as we did for the heat/diffusion problem.

$$T(t) = \alpha\cos(c\omega t) + \beta\sin(c\omega t)$$
$$X(x) = A\cos(\omega x) + B\sin(\omega x)$$

Where, here, recall that we are arbitrarily setting the constant in the first ODE to unity because it is essentially a redundant constant. Thus, the general solution, before applying any other conditions, is

$$u(x,t) = [\alpha\cos(c\omega t) + \beta\sin(c\omega t)][A\cos(\omega x) + B\sin(\omega x)]$$

Application of the first boundary condition at $x = 0$ rapidly shows us that we must have

$$0 = A\cos(0) + B\sin(0)$$

由此得到结果 $A = 0$。余弦函数对这一特定问题的解没有贡献。但请注意，对于其他边界条件，情况可能会发生变化！到目前为止，解必须具有以下形式。

$$u(x,t) = [\alpha\cos(c\omega t) + \beta\sin(c\omega t)]\sin(\omega x)$$

Applying the second boundary condition (again, noting that the time component of the solution is generally non-zero, thus we can focus on the space component as the portion that must be zero) gives

$$0 = B\sin(\omega L)$$

The constant B cannot be zero unless the initial condition is also zero (which it is not). Thus we must have

$$0 = \sin(\omega L)$$

or,

$$\omega L = \pi,\ 2\pi,\ 3\pi \ldots = n\pi, \text{ for } n \in \mathbb{N}$$
$$\omega = \pi/L,\ 2\pi/L,\ 3\pi/L \ldots = \frac{n\pi}{L}, \text{ for } n \in \mathbb{N}$$

Recognizing that the most general solution must be a linear combination of all possible independent solutions, we form the sum to represent this. Recall, we have the multiplication of B with α and β; this generates two new constants. Rather than rename them, we can use the original constant names, and think of $B = 1$ as a condition we can impose without eliminating possible solutions.

$$u(x,t) = \sum_{n=1}^{\infty} [\alpha_n \cos(cn\pi/Lt) + \beta_n \sin(cn\pi/Lt)] \sin(n\pi x/L)$$

Using the initial condition, we have

$$4\delta_0 \tfrac{x}{L}(1 - \tfrac{x}{L}) = \sum_{n=1}^{\infty} \alpha_n \sin(n\pi x/L)$$

This indicates that the coefficients α_n are just the corresponding amplitudes for the Fourier series for the initial condition. This is given by

$$\begin{aligned}\alpha_n &= \frac{2}{L}\int_{x=0}^{x=L} 4\delta_0 \tfrac{x}{L}(1 - \tfrac{x}{L}) \sin(n\pi x/L)\, dx \\ &= 16\delta_0 \frac{1-(-1)^n}{\pi^3 n^3}\end{aligned}$$

The second boundary condition indicates that the initial velocity of the string is zero everywhere. Thus, we need the Fourier series for the zero function, which is just zero (in other words, for this case, the initial velocity does not impact the solution because it is zero, and we have $\beta_n = 0$ for all values of n). Therefore, our final solution takes the form

$$u(x,t) = \sum_{n=1}^{\infty} \cos(cn\pi t) \left[16\delta_0 \frac{1-(-1)^n}{\pi^3 n^3}\right] \sin(n\pi x/L)$$

There is some interesting interpretation that can be done here as we did with the heat/diffusion equation.

For reference, the plot of the solution for various times is given in Fig. 7.1. For this plot, we have taken $c = 250.88$ m/s.

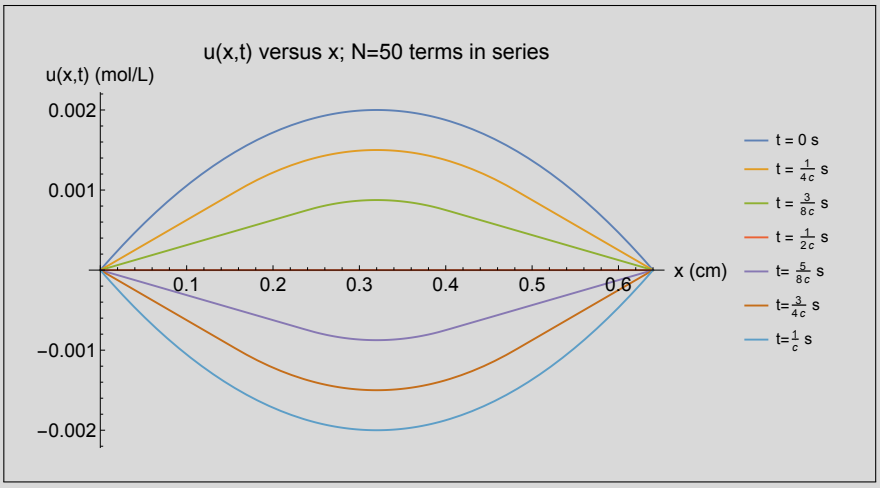


Fig. 7.2: The solution to the wave problem of a plucked guitar string; spatial profiles are given for for several selected times. For this plot, $x \in [0, 64]$ cm, $0 < t < 1/c = 0.004$ s. For this figure, the number of terms in the series was truncated at $N = 50$.

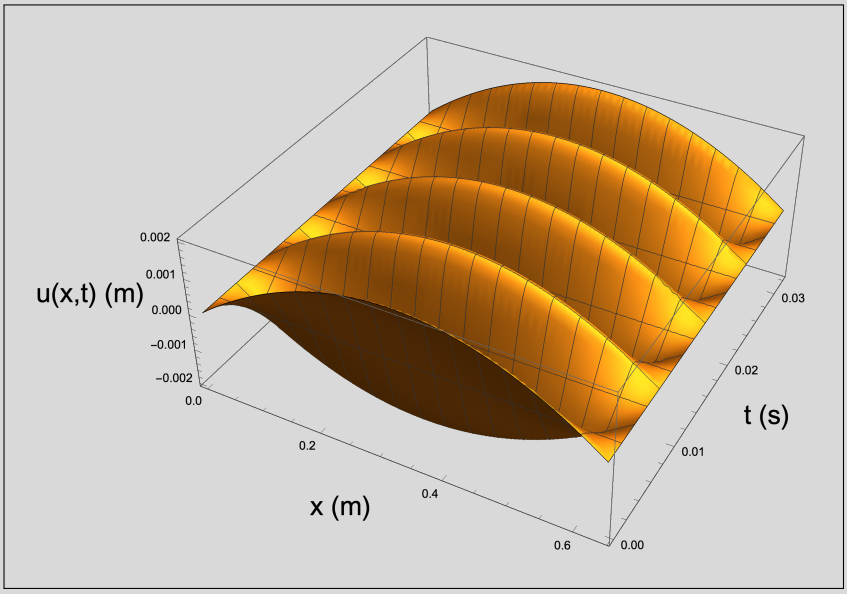


Fig. 7.3: The solution to the wave problem of a plucked guitar string. Here, the full time-space solution is shown as a surface. Parameters are the same as for the previous figure. Note that a total of four full periods in time are illustrated. This solution assumes that energy is not lost to heat, which is an approximation that does not hold for long times (otherwise guitar strings would never stop vibrating after being plucked!)

7.5 拉普拉斯方程

拉普拉斯方程是空间中的一种椭圆型 方程；它是应用数学中如此基础的一部分，以至于堪称偏微分方程中的真正原型。实际上，拉普拉斯方程常常作为热传导/扩散方程的稳态 形式出现；也就是说，多维热方程的稳态形式会导出一个拉普拉斯方程。从技术上讲，我们已经在第 6.8 节研究过一个一维的类拉普拉斯方程，当时考察了一个稳态扩散-反应问题。关于多维空间中偏微分方程的更一般分析，将在另一章中讨论。不过，要说明椭圆型方程的特征，二维拉普拉斯方程是最简单的研究对象。二维拉普拉斯方程的形式为

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 \tag{7.103}$$

其中x 和y 表示二维空间中的常规正交坐标轴。当该方程的右端是x 和y的函数时，该方程称为泊松方程。本章讨论分离变量法时，我们不会研究泊松方程。

To more thoroughly understand this equation, first note the following heat/diffusion equation in two space dimensions.

$$heat/diffusion \qquad \frac{\partial u}{\partial t} = K\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial x^2}\right) \tag{7.104}$$

Now note, the two *steady state* versions of these problems are identical, and both are of the form of the Laplace equation.

$$heat/diffusion\ (steady) \qquad 0 = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \tag{7.105}$$

Recall that the heat/diffusion equation is a parabolic equation that changes in time. In the long-time limit of this equation, the result is mathematically an elliptic equation. This is interesting in its own right. Elliptic equations apply to a wide variety of physical processes, and many of them can be thought of as being the steady-state of some more complex *transient* problem.

7.5.1 A Property of the Laplace Equation: The Maximum Principle

As mentioned above, the 2-dimensional Laplace equation can be thought of as the steady-state solution to the heat/diffusion equation in two spatial dimensions (and one time dimension!) We have already illustrated that the heat equation with homogeneous boundary conditions is characterized by having a variance for u that decreases in time. Thus, one can think of the Laplace equation as being one that has the smallest possible variance of all possible solutions to the heat equation. It turns out that this smallest possible variance happens as time tends toward infinity, and these solutions are equivalent to the solutions of the Laplace equation with the same boundary conditions.

A little thought will indicate that for homogeneous boundary conditions, the only steady-state solution to the heat equation is given by the solution

$$u(x,t\to\infty)=C_1 \tag{7.106}$$

where C_1 is a constant that is possibly zero. So, in general The steady state solutions to the heat/diffusion equation are not tremendously interesting at long times when the boundary conditions are homogeneous. We will handle nonhomogeneous for the heat/diffusion equation (and wave equation) later. For now, we can consider the Laplace equation in two space dimensions by thinking about the $y-$ spatial direction as being "time like" (for the purposes of pondering the qualities of the problem and its solution), and the $x-$direction as being space like. Under those circumstances, we might suspect that our overall scheme of separation of variables will still work, just by analogy with other equations in two independent variables that we have examined. In the material below, we will see that SOV does in fact work for the Laplace equation considered in this way. However, before continuing on, it is useful to establish a few other properties for solutions to the Laplace equation.

由于历史原因，拉普拉斯方程的解被称为调和函数。这种术语主要有助于理解，因为数学和工程学文献中到处都能看到"调和函数"这一术语。调和函数被赋予了一个专门名称，这也表明它们必定具有有趣的性质（我们很少给无趣的事物起专门的名字！）。调和函数有许多值得注意的有趣性质。下面介绍两个关于二维空间中调和函数的定理，它们很容易证明（Olver，2014年，第4章），但此处不作证明。

定理 7.1（调和函数的平均值）。设 $u(x,y)$ 是二维区域 $D(x,y)$内拉普拉斯方程（通常具有非齐次边界条件）的解。那么，对于定义域 D 内的每个(x_0,y_0) 点，以下结论成立

$$u(x_0,y_0)=\frac{1}{2\pi}\int_{\theta=0}^{\theta=2\pi}u(x_0+a\cos\theta,y_0+a\sin\theta)\,d\theta \tag{7.107}$$

其中 a 是任意一个不与边界 D 相交的圆的半径（但可以与边界相切）。换句话说，定义域内每个圆的圆心处的函数 u 值，都是函数 u 在圆内的平均值。

这是一个非常有趣的结论，它使我们能够确定有关调和函数的许多其他性质。结合下一个定理（同样不附证明），这些性质几乎告诉了我们所有可能希望了解的关于调和函数的信息。

Theorem 7.2 (Harmonic Functions are Analytic). *Recall that a function is analytic in a domain, $D(x,y)$, when it has a convergent Taylor series for each point (x,y) contained in the domain. All harmonic functions are analytic at each point in the interior of the domain D (this emphasizes that they may not be analytic on the boundary of the domain).*

Using the first of these two theorems, we can prove one additional thing about harmonic functions. While the proof is not difficult, we report only the result here.

Theorem 7.3 (Maximum-Minimum Theorem for Harmonic Functions). *The solution to the Laplace equation has both its maximum and its minimum values of u on the boundary of the domain D. This means that if the boundary conditions are specified by functions of u that are finite, then the solution inside the domain is also finite.*

7.5.2 SOV for the Laplace Equation

To be specific, consider the following steady-state heat transport problem on a unit square (where $x \in [0,1]$ and $y \in [0,1]$). For this problem, think of the dependent variable, u as being the temperature in a 1 meter by 1 meter square with a unit value for the thermal conductivity. The problem statement is

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 \tag{7.108}$$

B.C. 1	$u(x,0) = \sin(\pi x)$	(7.109)
B.C. 2	$u(x,1) = 0$	(7.110)
B.C. 3	$u(0,y) = 0$	(7.111)
B.C. 4	$u(1,y) = 0$	(7.112)
约束条件	解及其导数保持有界。	

这些约束条件中的最后一项再次源于物理条件。回想一下，我们将这类问题视为某种抛物型（热量/扩散）方程的稳态形式，那么我们此前假定适用于这类方程的物理约束也必须适用（因此，解必须有界）。约束条件的第二部分可以看作一项要求所有通量均为有限值的条件。请记住，通量通常由线性通量关系给出（即自变量的一阶导数乘以一个常数）；因此，要求导数保持有界，就意味着不允许任何通量趋于正无穷或负无穷。这显然与问题的物理性质一致！

现在，如果至少从概念上把y–方向看作类似“时间”的方向，那么从某种意义上说，这个问题似乎是这样一种问题：前两个边界条件提供的信息会向定义域内部传播，就像初始条件一样。实际上，问题的物理性质代表的是稳态条件，但以这种方式思考问题多少会有所帮助。至少，这表明我们常用的分离变量法或许能够奏效，因为它已经成功应用于结构类似的问题。

To check to see if our SOV approach will work for this problem, we can start with the conventional supposition: We want to determine if a solution of the form

$$u(x,y) = X(x)Y(y) \tag{7.113}$$

can be found. Substituting directly into the Laplace equation, we find

$$X''Y + XY'' = 0 \tag{7.114}$$

and, after dividing both sides by XY we have

$$\frac{X''}{X} = -\frac{Y''}{Y} \tag{7.115}$$

与之前热/扩散方程和波动方程的解一样，要使这一结果成立，唯一的可能性是两边（彼此独立）都等于同一个常数。对于这种情况，我们仍需确定该常数的符号，尽管显然最终解中分离常数会有一个正值和一个负值。回想一下，对于热/扩散

问题，我们已经确定，要在空间中得到满足齐次狄利克雷条件的解，合适的选择是令该常数为负。这里也有同类问题。在x方向上，有两个齐次边界条件。因此，类比可知，分离常数的正确符号应使分离后的常微分方程为 $X''+\lambda X = 0$。这种情况下，分离常数必须为负。请注意，如果我们交换 x 和y 的作用（也就是说，如果采用非齐次边界条件$y(0,y) = \sin(\pi x)$），我们就会得出相反的结论！

Choosing constant to be "-λ^2", then we have the two ODEs

$$X'' + \lambda^2 X = 0 \tag{7.116}$$

$$Y'' - \lambda^2 Y = 0 \tag{7.117}$$

As before, we can take these equations one at a time to determine whether or not our proposed solution is actually a reasonable one. Because the $x-$variable has two homogeneous Dirichlet conditions [$u(x=0,y) = 0$ and $u(x=1,y) = 0$], we propose starting with that one. The solution is the familiar one that we also saw in the analysis of the heat/diffusion equation.

$$X(x) = A\cos(\lambda x) + B\sin(\lambda x) \tag{7.118}$$

The two homogeneous boundary conditions demand that, whatever the solution for $X(x)$, it must be zero at the two boundaries in the $x-$direction. This gives again a familiar result

$$\begin{aligned} X(0) &= 0 \\ 0 &= A\cos(0) + B\sin(0) \end{aligned}$$

and, as we have seen previously, this implies

$$A = 0 \tag{7.119}$$

Thus, in the $x-$direction, the functions we use to "build" our solution will be the sine functions. Applying the second boundary condition gives

$$\begin{aligned} X(1) &= 0 \\ 0 &= B\sin(\lambda) \\ \lambda &= n\pi, \;\; n = 1,\, 2,\, 3\ldots \quad \text{(or, equivalently } n \in \mathbb{N}) \end{aligned}$$

This implies then

$$X_n(x) = B_n \sin(n\pi x), \;\; n \in \mathbb{N} \tag{7.120}$$

As with our previous solutions, we find that the solution demands that there be an infinite number of possible sine functions involved, each with a different frequency as modulated by the quantity $n\pi$.

Turning our attention to the second ODE, we have

$$Y'' - \lambda^2 Y = 0 \tag{7.121}$$

The characteristic equation for this problem is (noting $a = 1$, $b = 0$, and $c = -\lambda^2$)

$$s^2 - \lambda^2 = 0 \tag{7.122}$$

and this has roots

$$s = \frac{0 \pm \sqrt{4\lambda^2}}{2} \tag{7.123}$$

or,

$$s_1 = \lambda, \;\; s_2 = -\lambda \tag{7.124}$$

The solutions that correspond to these roots are the exponential ones.

$$Y(y) = C_1 \exp(\lambda y) + C_2 \exp(-\lambda y) \tag{7.125}$$

The two remaining boundary conditions will help us work out the final solution for the problem. At this juncture, we have

$$\begin{aligned} u_n(x,y) &= X_n(x)Y_n(y) \\ u_n(x,y) &= B_n \sin(n\pi x)[C_1 \exp(n\pi y) + C_2 \exp(-n\pi y)] \end{aligned} \tag{7.126}$$

一个经验法则是始终<escape>*first*</escape>应用齐次边界条件；应用其中一个自变量为零时的边界条件也很有帮助。为此，我们注意到，“边界条件2”是齐次的，并且在$y = 0$处取值。

$$\begin{aligned} u_n(x,y) &= 0 \\ B_n \sin(n\pi x)[C_1 \exp(\lambda y) + C_2 \exp(-\lambda y)] &= 0 \end{aligned} \tag{7.127}$$

Solving this, we find

$$C_1 = -C_2 \exp(-2\lambda) \tag{7.128}$$

Substituting this into the solution gives us

$$u(x,y) = B_n \sin(n\pi x) C_2 [-\exp(-2\lambda)\exp(\lambda y) + \exp(-\lambda y)] \tag{7.129}$$

正如我们在前面的实例中所看到的，我们可以令 $B_n C_2 \Rightarrow B_n$ （即，我们可以取$C_2 = 1$），因为两个待定常数的乘积在数学上表示一个单独的待定常数。因此，我们有

$$u_n(x,y) = B_n \sin(n\pi x)[-\exp(-2\lambda)\exp(\lambda y) + \exp(-\lambda y)] \tag{7.130}$$

Our final step is to apply the remaining boundary condition at $y = 0$. Before proceeding, we note that the most general condition for the solution is a linear combination of all possible solutions to this point. While we could have imposed that condition at several points in the previous analysis, it is most convenient to impose it it now. This gives the solution in the form (where we have also substituted $\lambda = n\pi$)

$$u(x,y) = \sum_{n=1}^{\infty} B_n \sin(n\pi x)[-\exp(-2\pi)\exp(-n\pi y) + \exp(-n\pi y)] \tag{7.131}$$

Now, imposing the initial condition, we find

$$\sin(\pi x) = \sum_{n=1}^{\infty} B_n \sin(n\pi x)[1 - \exp(-2n\pi)] \tag{7.132}$$

The term on the right-hand side is just a Fourier series for $\sin(\pi x)$, but one that has a rescaled value for B_n (note that the rescaling function, $[1 - \exp(-2n\pi)$, does not depend on x). We can multiply both sides by $\sin(m\pi x)$, integrate, and use the conventional properties of orthogonality to complete the problem.

$$\int_{x=0}^{x=1} \sin(\pi x)\sin(m\pi x)\,dx = \sum_{n=1}^{\infty} B_n \int_{x=0}^{x=1} \sin(n\pi x)\sin(m\pi x)\,dx[1 - \exp(-2n\pi)] \tag{7.133}$$

On the left-hand side of this expression, we must have that $m = 1$; all of the other possible integer values for m lead to an integral that is zero by orthogonality. On the right hand side, because $m = 1$, we must also have only a single term in the sum corresponding to $n = 1$. Putting this all together, we find

$$\frac{1}{2} = B_1 \int_{x=0}^{x=1} \sin(\pi x)\sin(\pi x)\,dx[1 - \exp(-2\pi)] \tag{7.134}$$

Or, we find that

$$B_1 = \frac{1}{1 - \exp(-2\pi)}, \;\; B_n = 0 \;\; \text{for } n > 1 \tag{7.135}$$

The final solution to our problem is

$$u(x,y) = \frac{1}{1 - \exp(-2\pi)} \sin(\pi x)[-\exp(-2\pi)\exp(-\pi y) + \exp(-\pi y)] \tag{7.136}$$

Upon additional algebra, the result is the somewhat simple looking result

$$u(x,y) = \sin(\pi x)\exp(-\pi y) \tag{7.137}$$

Solutions to the Laplace equation are interesting from a number of perspectives. In many applications, they are the surfaces that *minimize* the energy associated with the surface. This would include both the steady-state versions of the heat/diffusion equation, and the steady-state version of the wave equation. In both cases, we have discussed the notion of the energy associated with the solution u. Even though the underlying equations have different physical interpretations, it is interesting to note that they are both governed by the same steady-state solution. In both cases, this particular steady state solution is the *minimal surface*, which is just a general term used to describe a surface that minimizes some ancillary function (in this case, the energy).

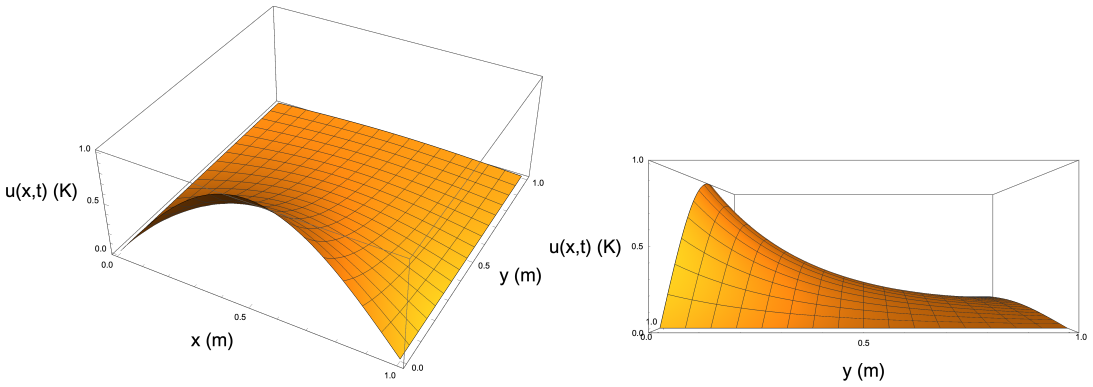


Fig. 7.4: **Solutions to Laplace Equation.** The Laplace equation solved via the SOV method. This approach works when the boundaries in one direction are homogeneous, and in the other direction there is one homogeneous boundary, and one (non-zero) specified value boundary.

7.6 Linear Superposition

线性是一种非常强大的性质，只要它存在，就能发挥重要作用。虽然事实上，许多（甚至大多数）物理学和工程学中使用的微分方程通常都是非线性的，但几乎所有这些方程都存在线性区间（或近似），而这些区间（或近似）可能非常有用。

在线性微分方程理论中，最有用的概念之一是叠加概念。这个概念并不晦涩或复杂；它只是线性这一性质的自然结果。在许多情况下，我们可能希望将一个复杂的线性问题分解为两个或更多个较简单的线性问题。这里的关键特征是，较简单问题的和必须在某种意义上等价于原始的复杂问题。

定义 7.1（线性叠加）。对于任意一个线性齐次微分方程的两个解，这两个解的线性组合仍然是该微分方程的一个解。这有时称为叠加原理，它同样适用于常微分方程和偏微分方程。

这段话st way to understand the principle of linear superposition is, as is normal, through an example.

Example 7.3 (Separation of Variables: The use of the linear superposition principle).

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 \tag{7.138}$$

B.C. 1	$u(x,0) = \sin(\pi x)$	(7.139)
B.C. 2	$u(x,1) = \sin(\pi x)$	(7.140)
B.C. 3	$u(0,y) = 0$	(7.141)
B.C. 4	$u(1,y) = 0$	(7.142)
Constraint	The solution and its derivatives remain bounded.	

Example 7.4 (Separation of Variables: The Laplace equation as a model for soap film geometry).

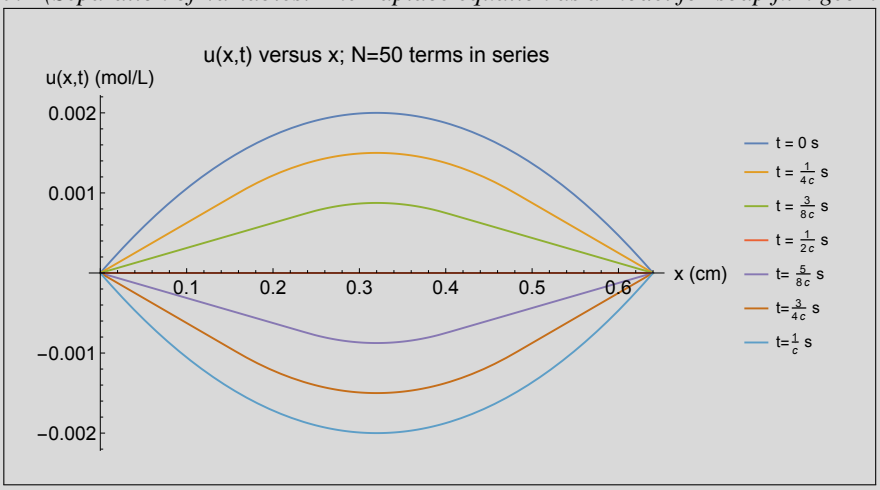


Fig. 7.5: The solution to the wave problem of a plucked guitar string; spatial profiles are given for for several selected times. For this plot, $x \in [0,64]$ cm, $0 < t < 1/c = 0.004$ s. For this figure, the number of terms in the series was truncated at $N = 50$.

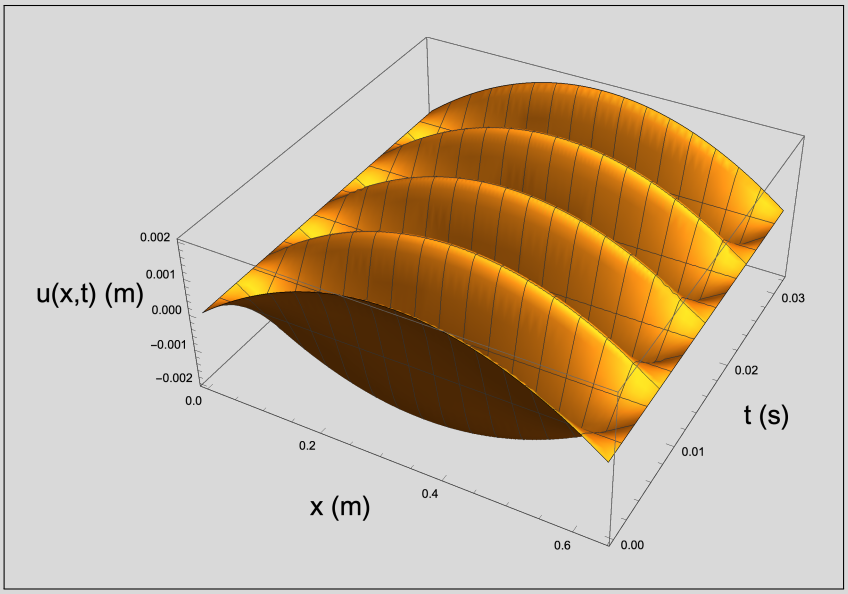


Fig. 7.6: The solution to the wave problem of a plucked guitar string. Here, the full time-space solution is shown as a surface. Parameters are the same as for the previous figure. Note that a total of four full periods in time are illustrated. This solution assumes that energy is not lost to heat, which is an approximation that does not hold for long times (otherwise guitar strings would never stop vibrating after being plucked!)

“粗略地说，傅里叶所创立的是一种数学方法，能够将任何模式——无论多么复杂——转换成由简单波组成的语言。他还展示了如何将这些波形转换回原始模式。换句话说，正如电视摄像机将图像转换为电磁频率，而电视机再将这些频率转换回原始图像一样，傅里叶展示了如何以数学方式实现类似的过程。他所推导出的、用于将图像转换为波形并再转换回来的方程，称为傅里叶变换。”

Michael Talbot，《全息宇宙》，1991年

8 傅里叶变换

8.1 引言

在本书迄今介绍的内容中，我们已经看到，傅里叶级数（以及基于斯图姆-刘维尔理论所定义函数的相关级数）用途广泛。傅里叶级数应用于众多学科，从计算机科学（其中它们可以成为各种压缩算法的组成部分）到自然现象的时间序列分析。毫不夸张地说，傅里叶级数的提出是现代历史上最重要的数学成果之一。

一类无法通过傅里叶级数分析处理的问题，是定义在整个实数轴上的函数问题。对于任意一个有限区间上的（具有足够良好性质的）函数，通过对函数进行平移并进行周期延拓，可以将该函数表示为傅里叶级数。然而，对于定义在实数轴上的函数（即 $\lim_{L\to\infty} x \in (-L,L)$），则不存在与之对应的傅里叶级数定义。

实际上，可以将傅里叶级数的基本概念扩展到整个实数轴。简而言之，这种转化将引出连续傅里叶变换的定义。尽管傅里叶变换并不特别复杂，但其中有许多细节常常妨碍人们轻松理解这一变换。此外，还有一种称为有限傅里叶变换的方法，适用于周期系统。事实上，最终会发现，后一种变换只不过是第 4 章所述傅里叶级数方法的一种应用。不过，将这种方法视为一种变换 而非用于表示函数的无限三角级数，意味着视角上的转变，而且是一个有益的转变。

本章旨在将对傅里叶变换的理解分解为更易于消化的部分。虽然傅里叶变换有许多用途，但我们将采用这种方法继续研究偏微分方程的求解方法。

8.2 Terminology

- **Integral Transform.** A method where a function is integrated against some *kernel* over its entire domain; the result is that its independent variable is replaced by a new independent variable contained in the kernel function. The purposes is to map the use the properties of integration to (1) bring out specific physical properties of the function (e.g., as Fourier transforms do by converting to wave space), or (2) to simplify some of the operations of the original equations (e.g., derivatives often become algebraic when integral transformed). A key component of the transforms is that they are both linear and invertible. In other words, a transformed function can be recovered by computing the appropriate inverse transform.

- **Fourier Transform.** The Fourier transform is a generalization of the Fourier series to the infinite real line. It is an integral transform. If x is the original independent variable, and ξ is the kernel variable, then the kernel function is $\exp(-i\xi x)$. Upon transforming a function of x a new function is recovered that is a function of ξ and no longer a function of x. This is sometimes represented by $f(x) \xrightarrow{\mathscr{F}[f]} \hat{f}\xi$.
- **Finite Fourier Transform**. A representation of the conventional Fourier series as a transform method. Even though a finite Fourier transform occurs on a bounded interval $x \in [a,b]$, it can still be cast as an integral transform. In this case, the inverse transform is the classical infinite sum that defines the Fourier series.
- **Euler's Identity.** We have seen this identity previously, but as a reminder it is the identity linking sine, cosine, and the complex exponential. Two frequently encountered forms of the identity are

$$\exp(ix) = \cos(x) + i\sin(x)$$
$$\exp(i\pi x) = \cos(\pi x) + i\sin(\pi x)$$

  The second of these is useful because the two trigonometric functions go through a full period as x goes between 0 and 1.
- **The Spectrum**. For Fourier series and transforms, the transformation of the independent variable, x, is to a new variable, ξ, that measures the frequency of the component of the sine and cosine functions used to represent it. The relationship between the *amplitude* of the trigonometric functions versus the corresponding frequency variable, ξ.
- **卷积**。一种积分运算，通过连续移动两个函数中的一个，使一个函数与另一个函数相乘并进行积分。对于两个函数 $f(x)$ 和 $g(x)$，卷积在区间 $x \in [a,b]$ 上定义为。

$$(f * g)(x) = \int_a^b f(x)g(x-z)\,dz$$

  ˇ 虽然有限区间上的卷积确实存在（Ljubarskii，1977），但更常见的是在半无限或无限直线上出现。对于傅里叶级数，卷积定义在无限实数轴上，我们有以下定义。

$$(f * g)(x) = \int_{-\infty}^{\infty} f(z)g(x-z)\,dz = \int_{-\infty}^{\infty} f(x-z)g(z)\,dz$$

- **Square-Integrable.** A function is said to be square-integrable on $x \in [a,b]$ if

$$F(x) = \int_a^b f^2(x)\,dx < \infty$$

Functions that are square-integrable belong to a space called the Hilbert space, L^2. Square-integrable functions are an important class of functions in many areas of analysis.

- **Tempered Distribution.** A version of distribution theory that allows the Fourier transform to be generalized to more kinds of functions. While the conventional requirement for a Fourier transform to exist is that the function be *square integrable* on $x \in (-\infty, \infty)$, the theory of tempered distributions allows the extension of Fourier Transforms to (1) generalized functions (such as the delta function), and (2) functions that do not grow *too fast* (and this means only that they must generally grow slower than an exponential function). Although the theory of tempered distributions is not covered in this text, their existence is needed to explain the use of several functions that arise naturally in applications of the Fourier transform.

8.3 回归傅里叶级数：复傅里叶级数

Recall that every function defined over a symmetric interval around zero, $x \in [-L, L]$, can be expressed as the sum of an *even* plus an *odd* function

$$f(x) = f_{even}(x) + f_{odd}(x) \tag{8.1}$$

Also recall, that over the same symmetric interval, the Fourier sine series is able to represent only odd functions, whereas the Fourier cosine series is able to represent only even functions. Thus, in general, we need both the Fourier sine and cosine series to represent a function (one to represent f_{even} and one to represent f_{odd}). The expressions developed previously were of the form

$$f(x) = A_0' + \sum_{n=1}^{n=\infty} A_n \cos\left(n\pi \tfrac{x}{L}\right) + \sum_{n=1}^{n=\infty} B_n \sin\left(n\pi \tfrac{x}{L}\right) \tag{8.2a}$$

$$A_0' = \frac{1}{2L} \int_{x=-L}^{x=L} f(x)\, dx \ \text{ (the average)} \tag{8.2b}$$

$$A_n = \frac{1}{L} \int_{x=-L}^{x=L} f(x) \cos\left(n\pi \tfrac{x}{L}\right)\, dx \tag{8.2c}$$

$$B_n = \frac{1}{L} \int_{x=-L}^{x=L} f(x) \sin\left(n\pi \tfrac{x}{L}\right)\, dx \tag{8.2d}$$

These results summarize what we have developed for the general Fourier series on a symmetric interval around $x = 0$.

Although generally there is no need to do so, we can also express these results entirely using the complex exponentials defined by Euler's equation. Recall, Euler's equation stated

$$e^{in\pi x/L} = \cos\left(n\pi \tfrac{x}{L}\right) + i \sin\left(n\pi \tfrac{x}{L}\right) \tag{8.3}$$

Making linear combinations of this result allows us to define the following relationships

$$\cos\left(n\pi\tfrac{x}{L}\right) = \frac{e^{in\pi x/L} + e^{-in\pi x/L}}{2} \tag{8.4}$$

$$\sin\left(n\pi\tfrac{x}{L}\right) = \frac{e^{in\pi x/L} - e^{-in\pi x/L}}{2i} \tag{8.5}$$

Substituting these into Eqs. (8.2a)-(8.2d) is algebraically messy, but otherwise not difficult. The result is

$$f(x) = A_0' + \sum_{n=1}^{n=\infty} A_n \frac{e^{in\pi x/L} + e^{-in\pi x/L}}{2} + \sum_{n=1}^{n=\infty} B_n \frac{e^{in\pi x/L} - e^{-in\pi x/L}}{2i} \tag{8.6a}$$

$$A_0' = \frac{1}{2L}\int_{x=-L}^{x=L} f(x)\,dx \ \text{(the average)} \tag{8.6b}$$

$$A_n = \frac{1}{L}\int_{x=-L}^{x=L} f(x)\frac{e^{in\pi x/L} + e^{-in\pi x/L}}{2}\,dx \tag{8.6c}$$

$$B_n = \frac{1}{L}\int_{x=-L}^{x=L} f(x)\frac{e^{in\pi x/L} - e^{-in\pi x/L}}{2i}\,dx \tag{8.6d}$$

Or, collecting like exponential terms

$$f(x) = A_0' + \sum_{n=1}^{n=\infty} \frac{A_n - iB_n}{2} e^{in\pi x/L} + \sum_{n=1}^{n=\infty} \frac{A_n + iB_n}{2} e^{-in\pi x/L} \tag{8.6e}$$

$$A_0' = \frac{1}{2L}\int_{x=-L}^{x=L} f(x)\,dx \ \text{(the average)} \tag{8.6f}$$

$$\frac{A_n - iB_n}{2} = \frac{1}{L}\int_{x=-L}^{x=L} f(x)e^{-in\pi x/L}\,dx \tag{8.6g}$$

$$\frac{A_n + iB_n}{2} = \frac{1}{L}\int_{x=-L}^{x=L} f(x)e^{in\pi x/L}\,dx \tag{8.6h}$$

最后，请注意，将所使用的系数重新命名如下会更加方便。

$$f(x) = C_0 + \sum_{n=1}^{n=\infty} C_n e^{in\pi x/L} + \sum_{n=1}^{n=\infty} D_n e^{-in\pi x/L} \tag{8.7}$$

$$C_0 = \frac{1}{2L}\int_{x=-L}^{x=L} f(x)\,dx \ \text{(the average)} \tag{8.8}$$

$$C_n = \frac{1}{2L}\int_{x=-L}^{x=L} f(x)e^{-in\pi x/L}\,dx \tag{8.9}$$

$$D_n = \frac{1}{2L}\int_{x=-L}^{x=L} f(x)e^{in\pi x/L}\,dx \tag{8.10}$$

最后一步，注意到C_n 和 D_n 的定义方式相同，唯一的区别在于它们的指数符号相反。然而，如果我们将C_n 的表达式从 $n = -1$ 求和到 $n = -\infty$，便可得到以下一个简便的关系。

$$\sum_{n=1}^{n=\infty} D_n e^{-in\pi x/L} = \sum_{n=-1}^{n=-\infty} C_n e^{in\pi x/L} \tag{8.11}$$

Combining Eqs. (8.8), (8.9), (8.10), and (8.11), the result is a single sum expressing the entire series

$$f(x) = \sum_{n=-\infty}^{n=-1} C_n e^{in\pi x/L} + C_0 + \sum_{n=1}^{n=\infty} C_n e^{in\pi x/L} \tag{8.12}$$

$$C_0 = \frac{1}{2L}\int_{x=-L}^{x=L} f(x)\,dx \tag{8.13}$$

$$C_n = \frac{1}{2L}\int_{x=-L}^{x=L} f(x)e^{-in\pi x/L}\,dx \tag{8.14}$$

注意：与傅里叶级数一样，当 $n = 0$ 时，C_n 的表达式并不总是会计算出正确的量。不过，如果确实如此，我们就可以将级数写成下面这种更加紧凑的形式

$$f(x) = \sum_{n=-\infty}^{n=\infty} C_n e^{in\pi x/L} \tag{8.15a}$$

$$C_n = \frac{1}{2L}\int_{x=-L}^{x=L} f(x)e^{-in\pi x/L}\,dx \;\; n = \ldots -2,-1,0,1,2\ldots \tag{8.15b}$$

提醒一下，上述结果只是对零附近对称区间 $x \in [-L,L]$，上的一般傅里叶级数所得的结果，其中我们用复指数表示了正弦函数和余弦函数。虽然这些表达式的形式看起来更简单，但其含义与用正弦和余弦级数来解释时完全相同。求和后，用这种形式表示的傅里叶级数结果，应与直接使用正弦和余弦级数计算出的结果完全一致。这里尤其需要注意的是，所得级数中不应包含复数。尽管中间步骤中会出现复数，但最终表达式中只应出现实数（假设原函数一开始就是实函数！）。与第 4 章中研究的传统傅里叶级数一样，这些函数必须满足所概述的狄利克雷条件

Note that because of the change of bounds for the summation, we still have to compute exactly the same number of coefficients for any finite approximation to the series (that number being $2N+1$, where N is the number of terms that we decide to compute, and the plus one comes from the C_0 constant). In addition, some find this form of the expression to be less intuitive than the one involving sine and cosine functions (and not involving any complex numbers).

Example 8.1 (A Fourier Series Using the Complex Series Form). For the function $f(x) = e^x$ on $[-1,1]$, the conventional Fourier series is found from

$$A_0' = \frac{1}{2}\int_{x=-1}^{x=1} e^x\,dx = \sinh[1] \tag{8.16}$$

$$A_n = \frac{1}{2}\int_{x=-1}^{x=1} e^x \sin[n\pi x]\,dx = \frac{2(-1)^n \sinh(1)}{1+\pi^2 n^2} \tag{8.17}$$

$$B_n = \frac{1}{2}\int_{x=-1}^{x=1} e^x \cos[n\pi x]\,dx - \frac{2\pi(-1)^n n \sinh(1)}{1+\pi^2 n^2} \tag{8.18}$$

with

$$e^x = A_0' + \sum_{n=1}^{n=\infty} A_n \cos(n\pi x) + \sum_{n=1}^{n=\infty} B_n \sin(n\pi x) \tag{8.19}$$

The series solution for this is plotted in Fig. 8.1. The corresponding complex exponential series can be found as follows.

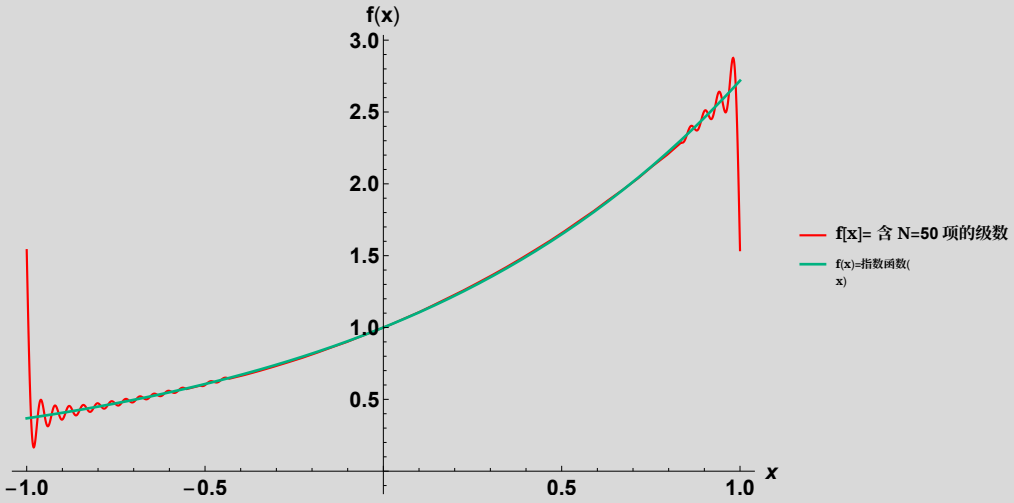


图 8.1：$f(x) = e^x$ 含 $N = 50$ 项的傅里叶级数。

$$\begin{aligned} C_n &= \frac{1}{2}\int_{x=-1}^{x=1} e^x e^{-in\pi x}\,dx \quad n = \ldots -2,-1,0,1,2\ldots \\ &= \frac{(-1)^n (L+i\pi n)\sinh(L)}{L^2+\pi^2 n^2} \end{aligned}$$

因此，得到的级数为

$$e^x = \sum_{n=-\infty}^{n=\infty} \frac{(-1)^n (L+i\pi n)\sinh(L)}{L^2+\pi^2 n^2} e^{i\pi x}$$

For plotting purposes, we can re-express this result as

$$e^x = \sum_{n=-\infty}^{n=\infty} \frac{(-1)^n \sinh(L)}{L^2+\pi^2 n^2}(L+i\pi n)\left[\cos(n\pi x)+i\sin(n\pi x)\right]$$

请注意，这个求和式中的最后一个产物涉及两个复数的乘法。进行这种乘法时，会涉及许多相互抵消的项。经过大量代数运算，可以证明所有虚部都完全抵消（这是必然的）。结果是

$$\begin{aligned}
e^x &= \sum_{n=-\infty}^{n=\infty} \frac{(-1)^n \sinh(L)}{L^2+\pi^2 n^2}(L+i\pi n)\left[\cos(n\pi x)+i\sin(n\pi x)\right] \\
&= \sinh(1) + \frac{2\pi\sinh(1)\sin(\pi x)}{1+\pi^2} - \frac{4\pi\sinh(1)\sin(2\pi x)}{1+4\pi^2} + \frac{6\pi\sinh(1)\sin(3\pi x)}{1+9\pi^2} - \frac{8\pi\sinh(1)\sin(4\pi x)}{1+16\pi^2} \\
&\quad + \frac{10\pi\sinh(1)\sin(5\pi x)}{1+25\pi^2} \\
&\quad - \frac{2\sinh(1)\cos(\pi x)}{1+\pi^2} + \frac{2\sinh(1)\cos(2\pi x)}{1+4\pi^2} - \frac{2\sinh(1)\cos(3\pi x)}{1+9\pi^2} + \frac{2\sinh(1)\cos(4\pi x)}{1+16\pi^2} \\
&\quad - \frac{2\sinh(1)\cos(5\pi x)}{1+25\pi^2} + \cdots
\end{aligned}$$

该级数的图形与之前图 8.1 所示的图形完全相同。总而言之，虽然傅里叶级数的复指数表示始终 等价于用三角函数表示的级数，但对于所有用途而言，它不一定更加方便。特别是，如果希望实际对该级数求和，复指数形式通常并不更加方便。然而，在将傅里叶级数伸展到实数轴上（从而生成傅里叶变换）的过程中，复数形式将具有一定的优势。

8.4 The Fourier Series as a Transform

在实数轴上建立傅里叶变换之前，我们先从变换的角度重新审视傅里叶级数的定义，这将有助于理解相关概念。在这一背景下，傅里叶级数有时也被称为有限傅里叶变换 （参见 Churchill（1972））。这种重新命名仅代表对相关运算的视角发生了变化；从函数角度来看，即使将傅里叶级数视为作用于有限区间上的一种变换，其具体细节也不会发生任何改变。首先，我们来复习线性变换的概念。

8.4.1 Linear transforms

Definition. In the context of function spaces, a linear transformation, $\mathscr{L}(\cdot)$, takes a function as an argument (its domain), and generates a new function as its output (its range). Explicitly, this is defined by

$$\mathscr{L}[f] = g \tag{8.20}$$

请注意，与f和g相关联的自变量可能不同。

Example 8.2 (Linear Transformations on Functions). Consider the following two functions

$$f(x) = 4x$$
$$g(y) = 6y$$

Then the linear operation $f(g)$ is defined by

$$f[g(y)] = 4(6y) = 24y$$

这可以被视为从一个函数空间到另一个函数空间的线性变换。

As a second example, consider the operation of integration of a function. The integration process can be thought of as a linear transform (or, equivalently, linear operator). Note the following

$$\int (\alpha f(x) + \beta g(x))\, dx = \alpha \int f(x)\, dx + \beta \int g(x)\, dx$$

如上所示，变换保持加法和乘法运算不变的这一性质，定义了线性变换。因此，积分可以被视为一种线性变换。下面给出一个具体的例子。

$$f(x) = 4x^2 + 2$$
$$\mathscr{L}(f) = \int_{x=0}^{x=1} f(x)\, dx$$

因此

$$\begin{aligned}\mathscr{L}(4x^2+2) &= \int_{x=0}^{x=1} (4x^2+2)\, dx \\ &= 4\int_{x=0}^{x=1} x^2 + dx + 2\int_{x=0}^{x=1} 1\, dx \\ &= \frac{10}{3}\end{aligned}$$

将积分作为一种运算来应用的过程显然是线性的。

8.4.2 The Finite Fourier Transform

Now that we have had some reminders about the concept of a linear transform, we can revisit the Fourier series, but viewing it as a transform of one function (in its function space) to another function (in a different function space). Because this mapping between functions spaces is reversible, we will be examining the

Fourier series as a set of two operations: (1) a forward transform that take a function in a real variable and sends it to a new function defined by the sequence of amplitudes and frequencies, and (2) an inverse function that takes the transformed function of amplitudes and frequencies, and sends it back to being a conventional function in the original independent variable.

回顾一下，前面给出的复傅里叶级数定义是

$$f(x) = \sum_{n=-\infty}^{n=\infty} C_n e^{in\pi x/L} \tag{8.21a}$$

$$C_n = \frac{1}{2L}\int_{x=-L}^{x=L} f(x)e^{-in\pi x/L}dx \;\; n = \ldots -2,-1,0,1,2\ldots \tag{8.21b}$$

实际上，我们可以将其写成一个单独的表达式，如下所示。

$$f(x) = \sum_{n=-\infty}^{n=\infty}\left(\frac{1}{2L}\int_{x=-L}^{x=L} f(x)e^{-in\pi x/L}dx\right)e^{in\pi x/L} \tag{8.22}$$

以这种方式书写，我们可以将傅里叶级数解释为两个运算组成的序列。

1. 首先，有一种运算，它接受一个函数，并生成该函数的变换版本。注意到，该函数可以看作有序对 $(x, f(x))$，那么，求傅里叶振幅的过程会生成一个新的有序对 (n, C_n)，它表示原函数，但处于变换空间中（定义域为 n，值域为 C_n）。用符号表示，假设我们用 $\mathscr{F}_{FS}(\cdot)$ 表示这种变换。那么，我们可以用以下任一种方式来表示变换过程

$$(x, f(x)) \xrightarrow{\mathscr{F}_{FS}} (n, C_n)$$
$$\mathscr{F}_{FS}[f(x)] = C_n(n)$$

按照惯例，在将傅里叶级数描述为线性变换时，采用以下符号表示（或许多类似符号表示中的任意一种）：$C_n(n) \Leftrightarrow \hat{f}(n)$。这样，变换函数的形式 $\hat{f}(n)$ 就能提醒我们，它最初来自函数 $f(x)$。因此，上述两个变换按传统方式写作：

$$(x, f(x)) \xrightarrow{\mathscr{F}_{FS}} (n, \hat{f}(n))$$
or
$$\mathscr{F}_{FS}[f(x)] = \hat{f}(n)$$

With these definitions in place, we can think of the finite Fourier transform as being the inner integral in Eq. (8.22). This yields the definition of the forward transform

$$\boxed{\mathscr{F}_{FS}[f(x)] = \hat{f}(n) = \frac{1}{2L}\int_{x=-L}^{x=L} f(x)e^{-in\pi x/L}dx} \tag{8.23}$$

2. Second, there is an *inverse* operation that returns a function, $\hat{f}(n)$, in Fourier space back to a real function of the original variable, $f(x)$.

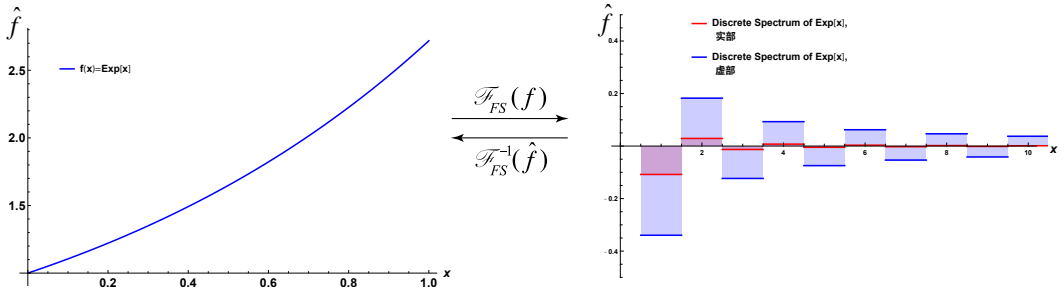


Fig. 8.2: The function $f(x) = \exp(x)$ (left) and its finite Fourier transform spectrum (right).

$$(n, \hat{f}(n)) \xrightarrow{\mathscr{F}_{FS}^{-1}} (x, f(x))$$

or

$$\mathscr{F}_{FS}^{-1}\left[\hat{f}(n)\right] = f(x)$$

Again, in reference to Eq. (8.22), note that if we adopt the notation given directly above for $\hat{f}$, then we have a definition for the *inverse transform* $\mathscr{F}_{FS}^{-1}$

$$\boxed{\mathscr{F}_{FS}^{-1}\left[\hat{f}(n)\right] = f(x) = \sum_{n=-\infty}^{n=\infty} \hat{f}(n) e^{in\pi x/L}} \tag{8.24}$$

An example showing the function $f(x) = e^x$ and its Finite Fourier transform spectrum is given in Fig. 8.2.

8.5 The Fourier Transform

We are now in a position where we can more carefully examine the idea of how to define the Fourier series on the real line, i.e., for the open interval $I = \lim_{L\to\infty}[-L < x < L]$. As mentioned above, the Fourier transform is an extension of the idea of the Fourier series, but extended to the entire real line. The Fourier transform was an original invention of Fourier; before him, the notion of integral transforms was not developed. While the Fourier series opened entirely new concepts regarding the very definition of the function, the Fourier integral extended these ideas, and had enormous impact across mathematics ranging from integration theory to the solutions to partial differential equations.

In the following, the Fourier transform is derived from the Fourier series. This derivation proceeds in a way that reasonably matches our intuition (i.e., if the Fourier transform is the extension of the Fourier series to the entire real line, then we should be able to develop it from the Fourier series by letting $L \to \infty$). However, there are many technical issues in the *formal* derivation of the Fourier transform that we will not examine in detail. Therefore, the derivation below should be considered the outline to a more careful, detailed proof. Regardless, the development has substantial intuitive value, and the presentation is frequently offered as the de facto derivation of the Fourier transform

8.5.1 Derivation of the Fourier Integral

To start, note that by definition of the complex Fourier series (or, equivalently, the finite Fourier transform) for any admissible function f on $x \in [-L, L]$ is given by

$$f(x) = \sum_{n=-\infty}^{n=\infty} C_n e^{in\pi x/L} \tag{8.25}$$

$$C_n = \frac{1}{2L} \int_{x=-L}^{x=L} f(x) e^{-in\pi x/L} dx \tag{8.26}$$

Combining these into a single expression (as we did for the finite Fourier transform), we find

$$f(x) = \sum_{n=-\infty}^{n=\infty} \left[\frac{1}{2L} \int_{x=-L}^{x=L} f(x) e^{-in\pi x/L} dx \right] e^{in\pi x/L} \tag{8.27}$$

Now, let

$$\begin{aligned} \xi_n &= \frac{n\pi}{L} \;\Leftrightarrow\; n = \frac{\xi_n L}{\pi} \\ \Delta\xi_n &= \left(\frac{(n+1)\pi}{L} - \frac{n\pi}{L} \right) = \frac{\pi}{L} \end{aligned} \tag{8.28}$$

Substituting these into Eq. (8.27) gives the following sum

$$f(x) = \sum_{n=-\infty}^{n=\infty} \left[\frac{1}{2\pi} \int_{x=-L}^{x=L} f(x) e^{-i\xi x} dx \right] e^{i\xi x} \Delta\xi_n \tag{8.29}$$

This sum is now in the form of a conventional Riemann sum (i.e., the sum that was covered in introductory calculus and led to the definition of the integral). The integral in this case is obtained in the limit as $L \to \infty$ (or, equivalently, as $\Delta\xi_n \to 0$)

$$f(x) = \lim_{\substack{\Delta\xi \to 0 \\ L \to \infty}} \left(\sum_{\xi L/\pi=-\infty}^{\xi L/\pi=\infty} \left[\frac{1}{2\pi} \int_{x=-L}^{x=L} f(x) e^{-i\xi x} dx \right] e^{i\xi x} \Delta\xi_n \right) \tag{8.30}$$

The final result is

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{\xi=-\infty}^{\xi=\infty} \left[\frac{1}{\sqrt{2\pi}} \int_{x=-\infty}^{x=\infty} f(x) e^{-i\xi x} dx \right] e^{i\xi x} d\xi \tag{8.31}$$

$$\tag{8.32}$$

Note that we have distributed the factor $1/(2\pi)$ in two parts, so that the forward and reverse transforms have symmetry in the constants in front of them. While the derivation above is compelling, there is at least one detail that was not described fully. We have apparently let the index, n, be transformed from an index set over the integers to a real number defined on the entirety of the real line. This was somewhat hidden by the

transformation of variables from n to ξ_n, but this transformation did occur. While we did not justify this step rigorously, such justifications are available (see Lanczos (1966)).

与有限傅里叶变换一样，我们定义如下的正变换和逆变换。

$$\boxed{\hat{f}(\xi) = \mathscr{F}[f(x)] = \frac{1}{\sqrt{2\pi}} \int_{x=-\infty}^{x=\infty} f(x) e^{-i\xi x}\, dx} \tag{8.33}$$

$$\boxed{f(x) = \mathscr{F}^{-1}[f(x)] = \frac{1}{\sqrt{2\pi}} \int_{\xi=-\infty}^{\xi=\infty} \hat{f}(\xi) e^{i\xi x}\, d\xi} \tag{8.34}$$

虽然该变换与有限傅里叶变换直接对应，但还有一些需要考虑的事项。下面简要介绍这些事项。

1. 要使傅里叶变换存在，函数f必须在$x \in (-\infty, \infty)$上平方可积。这仅意味着

$$\boxed{\int_{x=-\infty}^{x=\infty} |f(x)|^2\, dx < \infty} \tag{8.35}$$

并且这个有限数是唯一的。虽然我们可能不需要直接使用这一点（因为大多数傅里叶变换都能在表格中查到），但应当明白，只有在此意义下为有限的函数才能进行变换。本章稍后将说明，如何在少数特殊情形下放宽这一要求，例如狄拉克δ函数、海维赛德函数和常函数。

2. 所得变换的光滑性由被变换函数的光滑性决定，其方式与傅里叶级数的情形完全类似。换言之，表征

$$f(x) = \mathscr{F}^{-1}[f(x)] = \frac{1}{\sqrt{2\pi}} \int_{\xi=-\infty}^{\xi=\infty} \hat{f}(\xi) e^{i\xi x}\, d\xi \tag{8.36}$$

对于解析函数，该级数将一致收敛。对于连续函数，傅里叶变换也将逐点收敛。对于具有跳跃间断点的函数，傅里叶变换将以与傅里叶级数相同的方式收敛；也就是说，它收敛到该值

$$f(x) = \frac{f(x^+) + f(x^-)}{2} \tag{8.37}$$

3. Because the delta function has a representation as a Fourier series, it also has a representation in the transformed space. In other words, the transform of generalized functions like the delta function can be given meaning. In this sense, the generalized functions act the same as any other function in the transformed space.

8.5.2 傅里叶变换的记号与限制

在整个应用数学领域中，恐怕没有哪个领域像傅里叶变换这样，因采用多种不同的符号表示而造成如此多的混淆。傅里叶变换表的使用者需要了解若干与符号表示有关的事项，具体如下。

1. 首先，对于我们所采用的定义，傅里叶变换及其逆变换积分前的系数分别为 $\mathscr{F}$ 和 $\mathscr{F}^{-1}$，取对称值 $1/\sqrt{2\pi}$（即变换及其逆变换前出现相同的常数）。有些文献倾向于在定义$f(x)]$,的变换前保留因子 $1/(2\pi)$，而在逆变换前使用系数 1。2. 在我们的定义中，正变换由指数函数集合$e^{-i\xi x}$定义，逆变换则由$e^{i\xi x}$定义。有些文献会调换由基函数$e^{-i\xi x}$定义的变换。3. 傅里叶变量 ξ有多种常用表示方式。这个变量通常称为波数空间变量 （多维情况下也称为波矢 ）。符号k、 α和 ω 也常用来表示我们记作 ξ的变量。4. 表示函数傅里叶变换的记号有多种。最常见的记号是在变换上加“帽子”（即 $\hat{f} = \mathscr{F}[f]$），但也经常使用其他记号。另一种广泛使用的约定是用小写字母及其对应的大写字母表示变换对（即$F = \mathscr{F}[f]$）。还有一些记号会使用其他符号，例如在字母上方加横线（即 $\bar{f} = \mathscr{F}[f]$；我们将保留此记号专门表示拉普拉斯变换），以指示变换。

使用不同形式表示傅里叶变换时，主要困难在于需要通过表格查找函数的适当变换。简而言之，任何表格中列出的结果都取决于变换本身的定义，以及与变换相关的符号表示。对于基本变换定义相同的情况，主要问题只是变量名称的变化（例如，表格可能使用 α 而不是 ξ）。然而，对于采用不同变换—逆变换对定义的表格，差异就更加显著。最重要的一点是：如果使用表格（或 Mathematica 等软件）来确定傅里叶变换，那么变换对的定义必须与表格中所采用的定义相一致。

傅里叶变换的限制。傅里叶变换的限制很少，至少没有多少是我们需要关注的。傅里叶变换存在的传统要求是，被变换的函数必须是平方可积的。因此，为使 $f(x)$ 的变换存在，我们需要确保

$$I = \int_{-\infty}^{\infty} f^2(x)\,dx < \infty \tag{8.38}$$

在应用中经常出现的一些函数（例如海维赛德函数）并非平方可积。对于这类函数，傅里叶变换理论有一种扩展，称为温和分布理论。我们不会详细讨论这一理论，但至少需要指出，这种理论框架确实存在。注意到这一框架后，我们便可以对一些原本无法进行变换的函数（例如海维赛德函数）作出合理的解释。

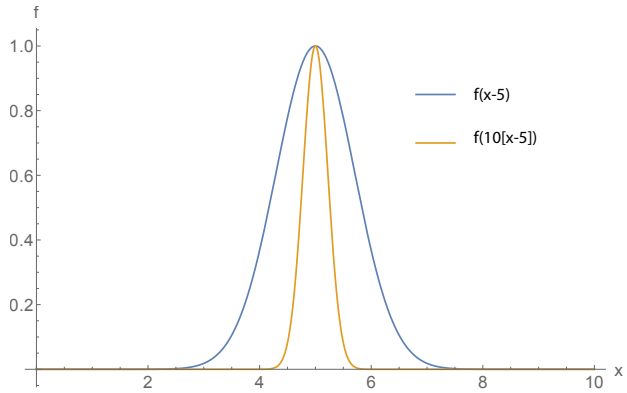


Fig. 8.3 The function f upon both shifting and scaling so that it takes the form $f(a[x-c])$. Here, the constant a controls the scale of the function (the larger the value of a, the narrower the function), and c controls the shift (subtracting a constant shifts functions to the right; adding a constant shifts functions to the left).

8.6 Fourier Transform Identities

傅里叶变换在实数空间与变换空间之间具有一些有趣的性质（更准确地说，是恒等式）。正是这些性质在一定程度上使傅里叶变换在应用中十分有用。总之，对于以下函数形式，傅里叶变换在变换空间中有简单的表征：（1）自变量平移一个常数的函数；（2）自变量乘以一个常数缩放因子的函数；（3）由求导运算定义的函数。下面将简要讨论这几种情况。

8.6.1 Fourier Transform of Shifted Functions

回顾一下，当从自变量中减去一个常数c，从而形成$f(x-c)$时，函数$f(x)$会向右平移（见图 8.3）。当对形式为$f(x-a)$的函数进行变换时，以下恒等式成立。

$$\mathscr{F}[f(x-a)] = e^{-ia\xi}\hat{f}(\xi) \tag{8.39}$$

This identity is easily proven by making the variable transform $z = x - c$, and substituting this into the integral definition of the Fourier transform.

8.6.2 Fourier Transform of Scaled Functions

Scaling the independent variable of a function by a constant expands or compresses the functional relationship to the domain. In Fig. 8.3, the effect of multiplying the independent variable by a constant a is illustrated. Intuitively, one might expect the compression of the domain that such a scaling (with $a > 1$) creates to affect the Fourier representation by increasing the high-frequency components (i.e., the narrower function requires higher frequency components to resolve). This is essentially what happens. The following relationship holds for Fourier transforms of a scaled function

$$\mathscr{F}[f(ax)] = \frac{1}{a}\hat{f}\left(\frac{x}{|a|}\right) \tag{8.40}$$

This identity can be proved by the variable transformation $z = ax$.and substituting this into the integral definition of the Fourier transform.

8.6.3 Fourier Transform of Derivatives

从求解偏微分方程的角度来看，傅里叶变换具有一个主要的实用特性：它将导数（变换后的自变量的导数）转化为代数运算。这一点很容易通过分部积分来证明；这一推导留作练习。假设我们要求满足以下两个条件

$$B.C.1 \qquad u(x,t) \to 0 \ \ as \ |x| \to \infty$$

$$B.C.2 \qquad \left.\frac{\partial u}{\partial x}\right|_{(x,t)} \to 0 \ \ as \ |x| \to \infty$$

Then, we have the result

$$\begin{aligned}\mathscr{F}\left[\frac{\partial u}{\partial x}\right] &= \frac{1}{\sqrt{2\pi}}\int_{x=-\infty}^{x=\infty} \left.\frac{\partial u}{\partial x}\right|_{(x,t)} e^{-i\xi x}\,dx \\ &= i\xi\hat{u}\end{aligned} \tag{8.41}$$

This result can be iterated for higher-order derivatives (assuming that for each such higher derivative, the derivative is constrained so that it goes to zero as $|x| \to \infty$)

$$\mathscr{F}\left[\frac{\partial^2 u}{\partial x^2}\right] = (i\xi)^2\hat{u} = -\xi^2\hat{u} \tag{8.42}$$

Note that the following action of the transform on derivatives that are not with respect to a transformed variable

$$\begin{aligned}\mathscr{F}\left[\frac{\partial u}{\partial t}\right] &= \frac{1}{\sqrt{2\pi}}\int_{x=-\infty}^{x=\infty} \left.\frac{\partial u}{\partial t}\right|_{(x,t)} e^{-i\xi x}\,dx \\ &= \frac{\partial}{\partial t}\frac{1}{\sqrt{2\pi}}\int_{x=-\infty}^{x=\infty} u(x,t)e^{-i\xi x}\,dx \\ &= \frac{\partial \hat{u}}{\partial t}\end{aligned} \tag{8.43}$$

While these results may not necessarily seem significant, they are actually is exceptionally useful. The following examples helps illustrate the transform of the derivative.

Example 8.3 (Solution to the first-order convection equation using the Fourier transform). Consider the following problem, representing the pure convection of a pulse of solute at velocity c with concentration given by u.

$$\frac{\partial u}{\partial t} = -c\frac{\partial u}{\partial x} \qquad -\infty < x < \infty$$

$B.C.1$ $\quad u(x,t)$ remains bounded for all x and t

$I.C.$ $\quad u(x,0) = \exp(-x^2)$

This problem represents the translation in the $+x$ direction of the initial pulse, unchanged in shape, at velocity c. From what we know about this problem from our study of separation of variables, we expect the solution to be of the form $u(x,t) = \exp[-(x-ct)^2]$. However, our goal here is to *derive* this result. Toward that end, we take the Fourier transform of both sides of the PDE and of the initial condition. This gives

$$\frac{d\hat{u}}{dt} = -i\xi c\hat{u} \qquad -\infty < x < \infty$$

$B.C.1$ $\quad \hat{u}(\xi,t)$ remains bounded for all ξ and t

$I.C.$ $\quad u(\xi,0) = \frac{1}{\sqrt{2}}\exp(-\xi^2/4)$

The solution to this problem reduces to a first order, integrable ODE. The solution is

$$\hat{u}(\xi,t) = \frac{1}{\sqrt{2}}\exp(\frac{-\xi^2}{4})\exp[-ict\xi]$$

To complete the problem, we need to invert the transform. Taking the inverse of both sides yields

$$\mathscr{F}^{-1}[\hat{u}(\xi,t)] = \mathscr{F}^{-1}\left[\exp(-ict\xi)\frac{1}{\sqrt{2}}\exp(\frac{-\xi^2}{4})\right]$$

Here, the easiest way to proceed is to recognize that the first exponential is one that shifts the following transformed function. Thus, the way forward is to note that the second function, $\frac{1}{\sqrt{2}}\exp(\frac{-\xi^2}{4})$, has inverse transform $\exp(-x^2)$. The presence of the first exponential indicates that this function should be shifted by the amount $-ct$ in the independent variable x. Thus, the result is

$$u(x,t) = \exp\left[-(x-ct)^2\right]$$

Example 8.4 (Solution to the diffusion equation using the Fourier transform). Consider the following diffusion / heat transport problem

$$\frac{\partial u}{\partial t} = D\frac{\partial^2 u}{\partial x^2} \qquad -\infty < x < \infty$$

$$B.C.1 \qquad u(x,t) \to 0 \ \ as\ |x| \to \infty$$

$$B.C.2 \qquad \left.\frac{\partial u}{\partial x}\right|_{(x,t)} \to 0 \ \ as\ |x| \to \infty$$

$$I.C. \qquad u(x,0) = \delta(x)$$

This problem represents the diffusion of a "point" concentration or initial distribution of heat at the center of an infinite domain. Typically in Fourier transform problems, the two boundary conditions are assumed to be true even when they are not stated; in short, this just indicates that the behavior of most physical systems as one gets arbitrarily distant from the initial condition. Although, technically speaking, periodic initial conditions might also be possible to *solve*, they often do not make physical sense (e.g., it would require an infinite amount of energy to heat up an infinitely large system!).

Although this problem may look somewhat daunting, the Fourier transform makes it pretty simple. Transforming both sides of the equation and the initial condition, we have

$$\frac{\partial \hat{u}}{\partial t} = -\xi^2 D\hat{u} \qquad -\infty < x < \infty$$

$$I.C. \qquad \hat{u}(x,0) = \mathscr{F}[\delta(x)]$$

Note that we do not need the two boundary conditions at this juncture, since they were accounted for in the integration by parts that lead to the transformation of the derivative.

The Fourier transform of the initial condition is listed in tables; however, for this particular transform, the definition is also sufficient

$$\begin{aligned}\mathscr{F}[\delta(x)] &= \frac{1}{\sqrt{2\pi}}\int_{x=-\infty}^{x=\infty} \delta(x)e^{-i\xi x}\,dx \\ &= \frac{1}{\sqrt{2\pi}}\end{aligned}$$

或者，也可以通过查阅变换表来求得结果。

$$\mathscr{F}[\delta(x-a)] = \frac{1}{\sqrt{2\pi}}e^{-ia\xi} \tag{8.44}$$

which, evaluated at $a = 0$, generates the same result for the transform. Now, we have a simple first-order ODE of the form

$$\frac{d\hat{u}}{dt} = -\xi^2 D\hat{u} \qquad -\infty < x < \infty$$

$$I.C. \qquad \hat{u}(x,0) = \frac{1}{\sqrt{2\pi}}$$

This is easily solved using separation of variables to compute the integral

$$\frac{d\hat{u}}{\hat{u}} = -\xi^2 D dt \qquad -\infty < x < \infty$$

$$I.C. \qquad \hat{u}(x,0) = \frac{1}{\sqrt{2\pi}}$$

This can be recognized to give the conventional exponential solution

$$\hat{u}(\xi,t) = \frac{1}{\sqrt{2\pi}} \exp\left(-\xi^2 Dt\right) \tag{8.45}$$

这为该问题提供了一个解，但这个解仍然处于变换空间中。换句话说，我们得到了 $\hat{u}(\xi,t)$ 的解，但我们想要的是 $u(x,t)$ 的解。为了得到这个解，我们只需对最后一个方程的两边使用逆变换。

$$\mathscr{F}^{-1}\left[\hat{u}(\xi,t)\right] = \mathscr{F}^{-1}\left[\frac{1}{\sqrt{2\pi}} \exp\left(-\xi^2 Dt\right)\right] \tag{8.46}$$

根据定义，求出该方程左边的逆变换很容易：$\mathscr{F}^{-1}\left[\hat{u}(\xi,t)\right] = u(x,t)$。对于右端，我们至少可以提出常数（换句话说，总是可以从积分中提出常数），从而得到

$$u(x,t) = \frac{1}{\sqrt{2\pi}} \mathscr{F}^{-1}\left[\exp\left(-\xi^2 Dt\right)\right] \tag{8.47}$$

此时，我们可以尝试根据定义对表达式进行逆变换。然而，在大多数情况下，这会非常困难。傅里叶变换相关的大部分工作都要借助表格（或 Mathematica 之类的软件）计算傅里叶逆变换。使用表格时，做法是查看表格中变换一侧的内容（这些表格总是将函数及其变换分别列在两列中），找到一个变换（忽略常数后）与我们的结果形式相同。为此，大多数表格都会列出以下变换对

$$e^{-a^2x^2} \Leftrightarrow \frac{1}{a\sqrt{2}} e^{-\xi^2/(4a^2)}$$

We just need to make our transformed result match this transform (and thereby figure out what the associated constants are). To do so, we first note that we can operate on transform pairs with constants (or variables that are not part of the transform) as if they were equations. In other words, if we multiply a function by a constant, its transform will also be multiplied by the same constant. To see how this works, we can work out the example above in detail. We start with the exponential, and note that the two transformed functions (our function and the one listed in the table) will be identical for the case that

$$-\xi^2 Dt = \frac{-\xi^2}{4a^2}$$

Or, solving this for a^2 and a

$$a^2 = \frac{1}{4Dt}$$
$$a = \frac{1}{\sqrt{Dt/4}}$$

Substituting this back into our transform pair, this gives us

$$e^{-\frac{x^2}{4Dt}} \Leftrightarrow \sqrt{2Dt}e^{-\xi^2 Dt}$$

Recall, we are trying to find the inverse transform of just $\mathscr{F}^{-1}\left[\exp\left(-\xi^2 Dt\right)\right]$ that appears in Eq. (8.47). We are very close with the transform pair listed above. Now, we consider multiplying both sides by $1/\sqrt{2Dt}$. This gives the result

$$\frac{1}{\sqrt{2Dt}}e^{-\frac{x^2}{4Dt}} \Leftrightarrow e^{-\xi^2 Dt}$$

其中 is exactly the transform result that we need. Now, we can invert the transform given in Eq. (8.47)

$$\begin{aligned} u(x,t) &= \frac{1}{\sqrt{2\pi}}\mathscr{F}^{-1}\left[\exp\left(-\xi^2 Dt\right)\right] \\ &= \frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{2Dt}}e^{-\frac{x^2}{4Dt}} \\ &= \frac{1}{\sqrt{\pi}\sqrt{4Dt}}e^{-\frac{x^2}{4Dt}} \end{aligned}$$

This result was not too difficult to obtain. The result is an important one; the function that we derived is called the *heat kernel*, and it appears widely in solutions to the heat (or diffusion) equation.

8.7 Convolutions

本节将介绍卷积的一般概念。虽然本材料的最终目标是将其应用于傅里叶变换（以及稍后的拉普拉斯变换），但这里将讨论并说明卷积的一些一般性质。在探讨卷积在傅里叶变换中的应用之前，现在正是讨论卷积的作用的好时机。要直观地了解卷积如何“工作”，最直接的物理方式之一就是考察其在平滑操作中的应用。当然，并非所有卷积都一定是平滑操作；但使用卷积进行平滑是一个很好的例子，能够从函数和几何两个角度帮助我们理解卷积的作用。

8.7.1 Convolution Definition and Examples

卷积是一般性的数学过程，应用广泛。例如，它们可用于平滑含噪函数，而其离散版本还是称为卷积神经网络（CNNs）的人工神经网络的核心组成部分。在信号处理中，卷积用于创建滤波器，以突出或减弱信号的特定特征。如果你曾在电子表格程序中对一组数据使用过方框滤波器，那么这实际上就是一种离散卷积。该过程通常如下进行。首先，根据方框滤波器将覆盖的单元格数量，选择其宽度，例如选择电子表格中的一行。假设我们选取的数量为 3；这意味着，对于任意单元格 i，滤波器将包含其上方的单元格（单元格 $i-1$）和其下方的单元格（单元格 $i+1$）。现在，假设对这三个单元格取平均值（即将三个单元格之和乘以 $1/3$）。单元格i 中的值现在被该平均值替换，然后继续处理该列中的下一个单元格，并使用相同的相对方法计算平均值。最终，我们得到移动方框平均，或简称为该列数据经过方框滤波后的结果。在新的数据集中，每个数据点前后的值都会被纳入考虑，因此数值变得更加平滑。这就是卷积的基本概念，并且可以扩展到连续函数。

For continuous functions f and g, on the real line, the convolution is defined by either of the following two expressions

$$(f*g)(x) = \gamma \int_{-\infty}^{\infty} f(x-\xi) g(\xi)\, d\xi \tag{8.48}$$

$$(g*f)(x) = \gamma \int_{-\infty}^{\infty} g(x-\xi) f(\xi)\, d\xi \tag{8.49}$$

$$\text{Note:} (f*g)(x) = (g*f)(x)$$

The constant γ is a rescaling value that is usually set to $\gamma = 1$. Describing what is going on during a convolution is a bit tricky, and here some images will help. To start, lets define some compact function that is also a density function (so that its area is 1). It is not at all necessary that this function be compact or a density, but it does help with the visualization and interpretation. Here is the function we will use

$$w(x;\alpha) = \frac{\pi a}{4}\left[H\left(x+\tfrac{1}{\alpha}\right) - H\left(x-\tfrac{1}{\alpha}\right)\right] \cos\left(\frac{\pi\alpha x}{2}\right) \tag{8.50}$$

where here α controls the width of the density function (larger α corresponds to narrower w.

As a second function, we define the *boxcar* function (presumably because it can take the form of a railroad boxcar under the correct conditions). This function is the difference of two Heaviside functions, and is defined by

$$B(x;a;a+\Delta) = [H(x-a) - H(x-(a+\Delta))] \tag{8.51}$$

These two functions w and B are plotted in Fig. 8.4. Note that the density function w is both symmetric around $x = 0$. This is not necessary for convolutions in general; it is, however, necessary for smoothing operations. The symmetry of the density function around $x = 0$ creates a result where the boxcar function is smoothed, but its center of mass does not change. The fact that w is a *density* function with unit area assures that the total area of the smoothed boxcar function is the same as the original. When conducting smoothing operations using convolutions, these are often desirable properties.

现在，要计算卷积，我们只需计算方程 (8.48) 或 (8.48) 中的任意一个。两者的结果完全相同，因此，在解中选择哪个函数来表示 f 或 g 并无区别。

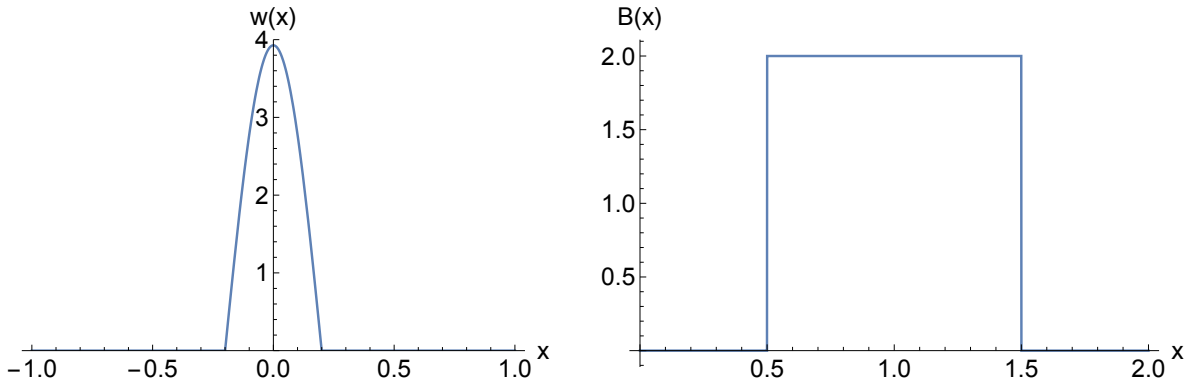


图 8.4: （左）由式 (8.50) 给出的密度函数，其中 $a = 5$。（右）矩形函数，$2B$ 其中$a = \frac{1}{2}$且$a + \Delta = \frac{3}{2}$。

evaluating these integrals. Note however, the choice of which function contains the displacement $(x-\xi)$ can influence the difficulty of the resulting integral. Suppose we set up our convolution as follows. Note that for B we have $a = \frac{1}{2}$ and $a+\Delta = \frac{3}{2}$, and for w we have taken $\alpha = 5$

$$\begin{aligned} B'(x;\tfrac{1}{2},\tfrac{3}{2}) &= \int_{-\infty}^{\infty} w(x-\xi)B(\xi)\,d\xi \\ &= \int_{-\infty}^{\infty} \left[\frac{5\pi}{4}\left[H\left((x-\xi)+\tfrac{1}{5}\right)-H\left((x-\xi)-\tfrac{1}{5}\right)\right]\cos\left(\frac{5\pi(x-\xi)}{2}\right)\right] \\ &\quad\times\left[H(\xi-\tfrac{1}{2})-H(\xi-\tfrac{3}{2})\right]\,d\xi \end{aligned} \tag{8.52}$$

While this expression is not easy to evaluate, it is possible to do so. Later on, we will see how Fourier transforms can be used to *evaluate* this integral. The result is the following somewhat lengthy result

$$\begin{aligned} B'(x;\tfrac{1}{2},\tfrac{3}{2}) = \frac{1}{2\sqrt{2}}&\operatorname{sgn}\left(\frac{17}{10}-x\right)\left[-\sin\left(\frac{5\pi x}{2}\right)-\cos\left(\frac{5\pi x}{2}\right)+\sqrt{2}\right] \\ &-\operatorname{sgn}\left(\frac{3}{10}-x\right)\left(-\sin\left(\frac{5\pi x}{2}\right)+\cos\left(\frac{5\pi x}{2}\right)+\sqrt{2}\right) \\ &-\operatorname{sgn}\left(\frac{7}{10}-x\right)\left[\sin\left(\frac{5\pi x}{2}\right)-\cos\left(\frac{5\pi x}{2}\right)+\sqrt{2}\right] \\ &+\operatorname{sgn}\left(\frac{13}{10}-x\right)\left[\sin\left(\frac{5\pi x}{2}\right)+\cos\left(\frac{5\pi x}{2}\right)+\sqrt{2}\right] \end{aligned} \tag{8.53}$$

This result is plotted in Fig. 8.5.

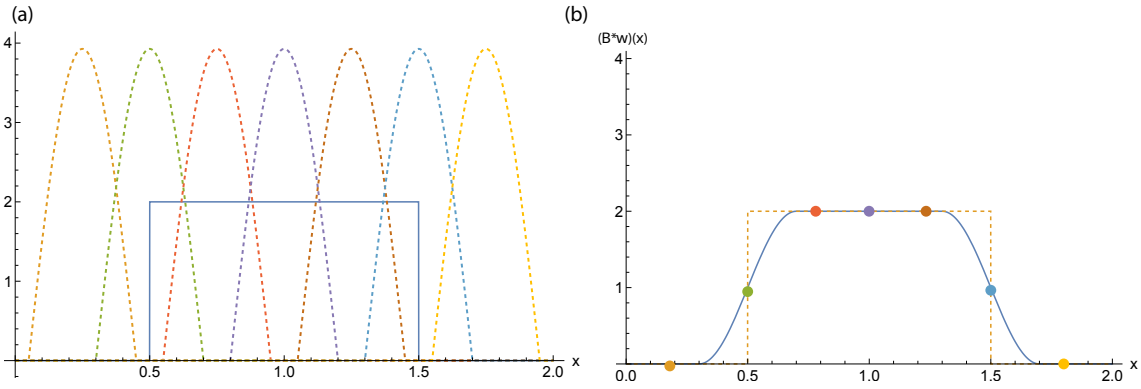


图 8.5：（a）式 (8.52) 中卷积的思想是，将密度函数w的平移副本与矩形函数相乘后进行积分。我们可以将每次此类积分视为在x的固定值下进行。然后，对 ξ 变量的每次积分都会得到 (b) 中所示新曲线上的一个点。（b）卷积得到的函数。曲线上的每个彩色圆都对应于w 的平移副本与B的乘积的积分；颜色对应于 (a) 中所示的曲线。当w 是以$x = 0$为中心的对称密度函数时，质心和B的面积都保持其原始值。

Here, the sign function $sgn(x)$ is defined by

$$\operatorname{sgn}(x) = \begin{cases} -1 & x < 0 \\ 0 & x = 0 \\ 1 & x > 0 \end{cases} \tag{8.54}$$

在图 8.6 中，展示了卷积过程的另一种表示方式。这里，我们可以将卷积理解为发生在由坐标 x 和 ξ 定义的二维空间中。回顾一下，卷积中的被积函数表示为 $w(x-\xi) \times B(\xi)$。因此，当在 $(x\ \xi)$ 空间中绘制时，函数$B(\xi)$ 与 x 无关，如图 8.6(a) 所示（蓝色曲线）。然而，权重函数是 x 和 ξ 的函数。回顾一下，权重函数关于 $x = 0$ 对称。因此，$w(x) = w(-x)$。如果愿意，我们可以将函数 w 重新写为 $w(-(\xi - x))$，这与 $w(\xi - x)$ 完全相同。因此，虽然函数B 在 $(x\ \xi)$ 空间中保持静止，但我们可以将函数 w 视为在 $(x\ \xi)$ 空间中持续平移 x 这一距离的函数。这就得到了图 8.6(a) 中所示的函数（橙色曲线）。

最后，卷积本身涉及对这两条曲线的乘积进行积分，其中 ξ 从$-\infty$到 ∞。由于w和B都是紧支撑的，实际积分区间为 $\xi \in [\frac{3}{10}, \frac{17}{10}]$（请注意，这两个分数出现在式 8.53 给出的解中）。该区域绘于图 8.6(c) 中。图中展示了三个取不同固定x 值的代表性平面。任意此类平面与函数$w(x-\xi) \times B(\xi)$（图 8.6(c) 中的橙色曲面）相交所形成的曲线，都会在卷积过程中进行积分。每条此类曲线的面积给出了解曲线上的一个点，如图 8.6(b) 所示。

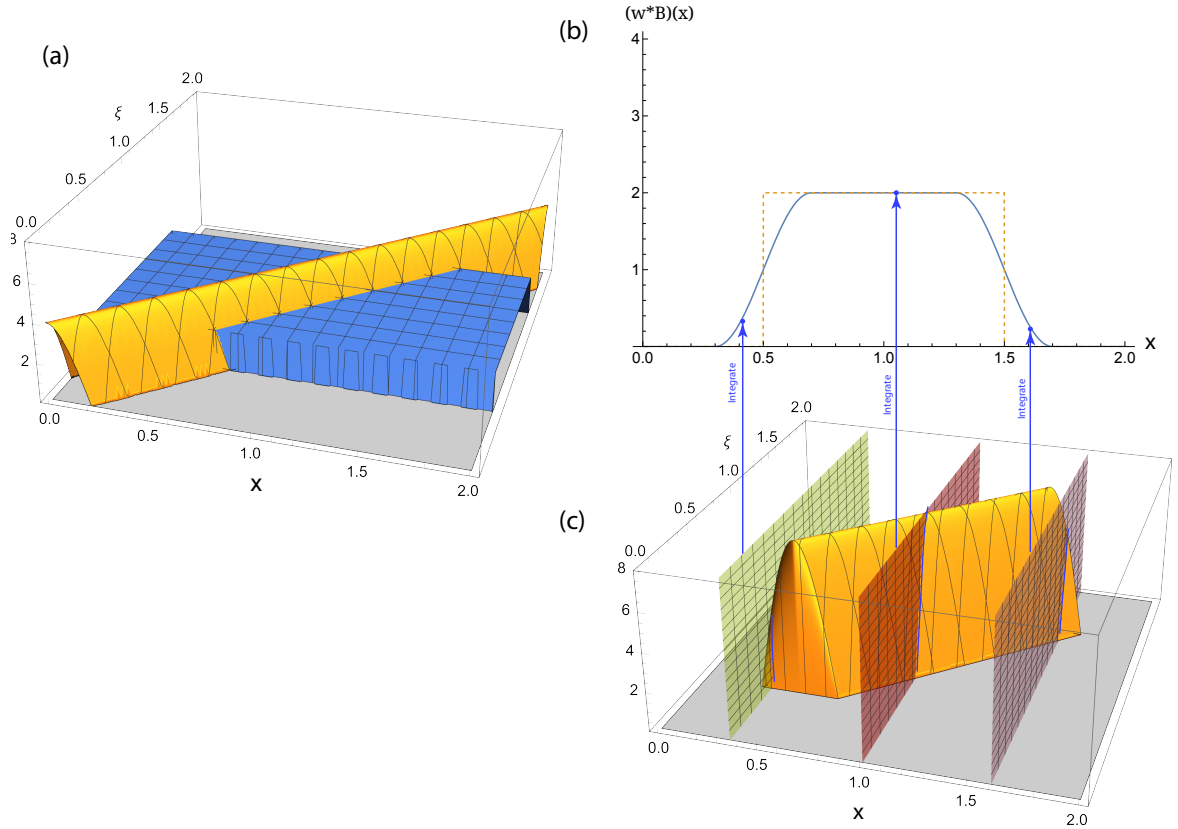


图 8.6：（左）卷积 $w * B = \int_{-\infty}^{\infty} w(x-\xi)B(\xi)d\xi$ 的图形化解释。(a) 函数 $B(\xi)$ 和 $w(x-\xi)$ 分别表示为自变量 x 和 ξ 的函数。(b) 卷积的结果是原函数 B 的平滑版本。请注意，此处虚线表示 B(x)，而蓝色实线表示卷积。(c) 卷积 $(w*B)(x)$ 通过在 x 固定时沿 ξ 方向积分求得；卷积上的每个点都对应这样一次积分。橙色曲线表示 $w(x-\xi)\times B(\xi)$。蓝线表示函数 $w(x-\xi)\times B(\xi)$ 与垂直于 x-ξ 坐标平面的垂直平面（x 为常数）相交形成的交线。

8.8 Convolutions in Fourier Transforms

有时，傅里叶变换会得到一个没有已知反函数的函数，但它可以表示为两个已知其逆的函数的乘积。乍看之下，这似乎是不常见的情况，但实际上，由于大多数初边值问题的结构，这种情况相当常见。

The following is a fairly remarkable feature of Fourier transforms, further illustrating the utility of the method. When one would like to inverse transform the product of two Fourier transformed functions, the result can be expressed

$$\mathscr{F}^{-1}[\hat{g}(\xi)\cdot\hat{f}(\xi)] = \frac{1}{\sqrt{2\pi}}\int_{\xi=-\infty}^{\xi=\infty} f(x-\xi)g(\xi)\,d\xi \tag{8.55}$$

$$= \frac{1}{\sqrt{2\pi}}\int_{\xi=-\infty}^{\xi=\infty} f(\xi)g(x-\xi)\,d\xi \tag{8.56}$$

Thus, the result states that the inverse transform of the product of two Fourier transformed functions $\hat{f}$ and $\hat{g}$ is just the convolution of the functions f and g in their *non-transformed* state. Note that for this definition, the scaling factor γ is not zero. This scaling factor shows up in the convolution directly from the definition of the Fourier transform.

We can also define the inverse of the convolution above, where one is taking the Fourier transform of the convolution of functions f and g.

$$\mathscr{F}[g(x)\cdot f(x)] = \frac{1}{\sqrt{2\pi}}\int_{x=-\infty}^{x=\infty} \hat{f}(x-\xi)\hat{g}(\xi)\,dx \tag{8.57}$$

$$= \frac{1}{\sqrt{2\pi}}\int_{x=-\infty}^{x=\infty} \hat{f}(\xi)\hat{g}(x-\xi)\,dx \tag{8.58}$$

虽然此时可能还不明显这种关系有何用处，但通过一个具体例子便可轻松看出。在这个例子中，我们考虑之前分析过的问题，不过初始条件更复杂（或者至少原则上如此）。

As an example of the power of the Fourier transform of convolutions, the following example is provided.

Example 8.5 (Computing a convolution using the Fourier transform). One way that the convolution identities can be used is to actually compute the value of a convolution using the Fourier transform. Suppose we take the following two functions

$$f(x) = 2\sqrt{2}\exp(-x^2) \tag{8.59}$$

$$g(x) = H(x+a) - H(x-a) \tag{8.60}$$

The convolution of these two quantities is (using the weighting factor $\gamma = 1/\sqrt{2\pi}$)

$$(f*g)(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} 2\sqrt{2}\exp[-(x-\xi)^2][H(x+a)-H(x-a)]\,d\xi \tag{8.61}$$

直接计算这个积分将面临相当大的困难。然而，利用卷积的傅里叶变换的性质，我们知道，其变换结果仅为f与 $\hat{}g$的产物。注意，傅里叶变换表（参见本章附录）提供了以下内容。

$$\hat{f} = 2\exp\left(-\frac{\xi^2}{4}\right) \qquad \text{(entry 15 in the Fourier transform table)} \tag{8.62}$$

$$\hat{g} = \frac{1}{\xi}\sqrt{\frac{2}{\pi}}\sin(a\xi) \qquad \text{(entry 32 in the Fourier transform table)} \tag{8.63}$$

Thus, we have

$$\mathscr{F}[(f*g)(x)] = 2\exp\left(-\frac{\xi^2}{4}\right)\frac{1}{\xi}\sqrt{\frac{2}{\pi}}\sin(a\xi) \tag{8.64}$$

现在，再次查阅傅里叶变换表，我们发现方程 8.64 列在表中的第 29 项。因此，对方程 8.64 两边取逆变换，得到结果

$$(f*g)(x) = \text{erf}(a+x) - \text{erf}(a-x) \tag{8.65}$$

这个结果并不费太大力气就得到了，但需要仔细运用卷积的傅里叶变换性质。

一般来说，确定卷积的值并不是傅里叶变换的主要用途。在实数轴上，傅里叶变换可以为求解线性偏微分方程提供一种非常有效的方法。当然，这要求方程中的所有项都有已知的傅里叶变换。其过程是通过傅里叶变换将偏微分方程转化为更容易求解的常微分方程。问题求解后，可以使用逆变换，将函数还原为原始（实空间）变量中的形式。使用傅里叶变换及变换表时有一些细微之处；要熟悉这一过程，需要经过一些练习。虽然傅里叶变换并非总是最佳解法，但它仍是求解偏微分方程时可用工具中的一种。下面的例子将以大家熟悉的实数轴上的热传导/扩散方程来说明这一过程。

Example 8.6 (A solution to the heat/diffusion equation using the Fourier transform- General initial conditions).

Consider the following diffusion / heat transport problem

$$\frac{\partial u}{\partial t} = D\frac{\partial^2 u}{\partial x^2} \qquad -\infty < x < \infty$$

$$B.C.1 \qquad u(x,t) \to 0 \ \ as\ |x| \to \infty$$

$$B.C.2 \qquad \left.\frac{\partial u}{\partial x}\right|_{(x,t)} \to 0 \ \ as\ |x| \to \infty$$

$$I.C. \qquad u(x,0) = \varphi(x)$$

这个问题描述了无限域中一般初始条件的扩散。这是对我们先前研究过的较简单初始条件（δ条件）的伸展。

对方程两边及初始条件进行变换，我们得到

$$\frac{\partial \hat{u}}{\partial t} = -\xi^2 D\hat{u} \qquad -\infty < x < \infty$$

$$I.C. \qquad \hat{u}(x,0) = \mathscr{F}[\varphi(x)]$$

对偏微分方程进行变换后，我们得到一个形式简单的一阶常微分方程

$$\frac{d\hat{u}}{dt} = -\xi^2 D\hat{u} \qquad -\infty < x < \infty$$
$$I.C. \qquad \hat{u}(x,0) = \hat{\varphi}(x)$$

This is easily solved using separation of variables to compute the integral

$$\frac{d\hat{u}}{\hat{u}} = -\xi^2 D dt \qquad -\infty < x < \infty$$
$$I.C. \qquad \hat{u}(x,0) = \hat{\varphi}(x)$$

This can be recognized to give the conventional exponential solution

$$\hat{u}(\xi,t) = \hat{\varphi}(x)\exp\left(-\xi^2 Dt\right) \tag{8.66}$$

This provides a *solution* to the problem, but the solution is still in the transformed space. In other words, we have the solution for $\hat{u}(\xi,t)$, but we want the solution for $u(x,t)$. To get this solution, we need only use the inverse transform on both sides of the last equation

$$\mathscr{F}^{-1}\left[\hat{u}(\xi,t)\right] = \mathscr{F}^{-1}\left[\hat{\varphi}(x)\exp\left(-\xi^2 Dt\right)\right] \tag{8.67}$$

Obtaining the inverse transform of the left-hand side of this equation is easy; by definition $\mathscr{F}^{-1}\left[\hat{u}(\xi,t)\right] = u(x,t)$. For the right-hand side, we are a bit stuck, since we have not yet specified what $\hat{\varphi}$ is. However, the convolution theorem can come to the rescue here.

$$u(x,t) = \mathscr{F}^{-1}\left[\hat{\varphi}(x)\exp\left(-\xi^2 Dt\right)\right] \tag{8.68}$$

which, using the convolution theorem, is

$$u(x,t) = \frac{1}{\sqrt{2\pi}}\int_{x=-\infty}^{x=\infty} \varphi(x-\xi)\frac{1}{\sqrt{2Dt}}e^{-\frac{x^2}{4Dt}} \tag{8.69}$$

or, equivalently

$$u(x,t) = \frac{1}{\sqrt{2\pi}}\int_{x=-\infty}^{x=\infty} \varphi(\xi)\frac{1}{\sqrt{2Dt}}e^{-\frac{(x-\xi)^2}{4Dt}} \tag{8.70}$$

最后这个结果尤其重要。项 $1/\sqrt{2Dt}\exp(-(x-\xi)^2/(4Dt))$ 是一个颇为著名的表达式，通常称为热核。这个方程是一个通用方程，通过与热核进行适当的卷积，来定义任意初始条件 φ 对应的解。如果能够计算出所得积分，就能得到这个原本极难问题的显式解。即使无法解析地求出解，至少也能将结果表示为积分。这类积分通常可以使用专门的数值方法进行高精度计算，这些方法旨在求出无限区域上有限积分的值。

8.9 傅里叶变换的伸展

有许多函数在应用中经常出现，但却无法进行傅里叶变换，至少按照我们所介绍的约定是如此。回想一下，我们指出傅里叶变换存在的首要条件是：被变换的函数必须具有唯一且有限的平方可积积分（即该函数平方可积）。然而，许多常见且有用的函数并不平方可积。例如，$\sin(x)$ 和 $\cos(x)$ 都不是平方可积的，因为它们的平方积分趋于无穷。类似地，狄拉克δ函数也不是平方可积的，因为严格来说，δ是一个分布，而不是函数；事实证明，δ函数的平方并无一般意义。这使我们处境颇为棘手，因为这些函数都相当基础，而且在应用中也经常出现。

幸运的是，有一个更一般的框架，可以对这类函数的傅里叶变换作出合理解释。这一理论涉及第5章简要介绍过的分布理论，并与傅里叶变换的定义相结合，从而形成一个更一般的框架，称为缓增分布的傅里叶变换。粗略地说，缓增分布就是“增长不会太快”的函数。这里的“增长不会太快”是指函数以某种有限的多项式速率增长（而不是以某种指数速率增长）。我们不会专门研究缓增分布，但重要的是要知道，前述理论发展有可靠的基础。在接下来的内容中，我们将为一些不满足最初平方可积要求、因而无法进行变换的函数推导傅里叶变换。不过，它们都属于缓增分布这一类函数。由于这一更复杂的理论能够确定这类函数的变换，因此，如果我们能求出其变换（及其逆变换），就可以在某种意义上“获准”扩充可变换函数库。以下内容只关注以下函数的变换：（1）赫维赛德阶跃函数和狄拉克δ函数；（2）常函数；（3）正弦函数和余弦函数。另有一些函数（如自然对数）也适用类似的概念，但这里不详细介绍相关结果（不过，它们的变换列在傅里叶变换表中）。

8.9.1 Fourier Transform of the Delta Function

As discussed in Chapter 5, the delta is not a true *function* (even though we use the name *delta function*); rather, it is a mathematical construct that represents a impulse (with infinite magnitude) applied to a single point. The delta function can be given clear mathematical meaning when it is integrated against other (appropriately constrained) functions.

因此，用傅里叶变换来表示狄拉克δ函数，从直觉上看是有道理的；傅里叶变换只是将输入函数f与积分核$e^{-i\xi x}$相乘后进行积分，以确定变换函数。利用狄拉克δ函数的常规性质和傅里叶变换的平移性质，我们可以轻松推导出以下结果。

$$\begin{aligned}\mathscr{F}[\delta(x)] &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\delta(x)e^{-i\xi x}\,dx \\ &= \frac{1}{\sqrt{2\pi}}\end{aligned} \tag{8.71}$$

$$\begin{aligned}\mathscr{F}[\delta(x-a)] &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\delta(x-a)e^{-i\xi x}\,dx \\ &= \frac{1}{\sqrt{2\pi}}e^{-ia\xi}\end{aligned} \tag{8.72}$$

Note that this last identity implies the following

$$\mathscr{F}[\delta(x+a)] = \frac{1}{\sqrt{2\pi}}e^{ia\xi} \tag{8.73}$$

and using the Fourier inversion formula gives

$$\mathscr{F}[e^{-ia\xi}] = \sqrt{2\pi}\delta(x) \tag{8.74}$$

8.9.2 海维赛德函数的傅里叶变换

To develop the Fourier transform of the Heaviside function, we take a bit of an indirect path. Recall, the Heaviside function (or the *unit step function*) is given by

$$H(x) = \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases} \tag{8.75}$$

根据我们在 §8.5 中对傅里叶变换的可容许函数的讨论，该函数并不符合要求，因为它不是平方可积的。尽管如此，我们仍然可以建立一种合理的常函数的傅里叶变换概念。首先，考虑以下变换函数。

$$\hat{f}(\xi) = \frac{1}{\sqrt{2\pi}}\frac{1}{i\xi} \tag{8.76}$$

While this proposed transform has simply proposed without explanation, we can proceed to see where the proposition leads us. We can consider the inverse Fourier transform for $\hat{f}$ by computing the following integral.

$$\begin{aligned}\mathscr{F}^{-1}\left[\frac{1}{\sqrt{2\pi}}\frac{1}{i\xi}\right] &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}\frac{1}{i\xi}e^{i\xi x}d\xi \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{e^{i\xi x}}{i\xi}d\xi \\ &= \frac{1}{2\pi i}\int_{-\infty}^{\infty}\frac{\cos(\xi x)+i\sin(\xi x)}{\xi}d\xi\end{aligned}$$

现在，注意到 $\sin(\xi)/\xi$ 是一个偶函数，而 $\cos(\xi)/\xi$ 是一个奇函数，显然余弦项的计算结果为零。此外，正弦项关于$x = 0$ 对称，因此导致了

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{\sin(\xi x)}{\xi}d\xi$$
$$= \frac{1}{\pi}\int_{0}^{\infty}\frac{\sin(\xi x)}{\xi}d\xi$$

As a final step, we will express this integral as follows

$$\frac{1}{\pi}\lim_{L\to\infty}\left[\int_0^L \frac{\sin(\xi x)}{\xi}d\xi\right] \tag{8.77}$$

This integral in the square brackets is called the sine integral. Many symbolic mathematical packages can evaluate this integral; in particular, as $L \to \infty$, we have

For $x > 0$

$$\frac{1}{\pi}\lim_{L\to\infty}\left[\int_0^L \frac{\sin(\xi x)}{\xi}d\xi\right] = \frac{1}{2} \tag{8.78}$$

and for $x < 0$

$$\frac{1}{\pi}\lim_{L\to\infty}\left[\int_0^L \frac{\sin(\xi x)}{\xi}d\xi\right] = -\frac{1}{2} \tag{8.79}$$

At this point it is helpful to develop a new type of step-like function which we might call the zero-average Heaviside function. Suppose that we define the zero mean Heaviside, H_0 by

$$H_0(x) = \begin{cases} -\frac{1}{2} & x < 0 \\ \frac{1}{2} & x > 0 \end{cases} \tag{8.80}$$

With this definition and the result above, we have the following

$$\mathscr{F}[H_0(x)] = \frac{1}{\sqrt{2\pi}}\frac{1}{i\xi} \tag{8.81}$$

With a little thought, we can relate the standard Heaviside function to the zero-mean Heaviside function simply by shifting $H(x)$ vertically by $1/2$.

$$H_0(x) = H(x) - \frac{1}{2} \tag{8.82}$$

因此，我们可以将式（8.80）写成如下形式。

$$\hat{f} = \mathscr{F}\left[H(x) - \frac{1}{2}\right] = \frac{1}{\sqrt{2\pi}}\frac{1}{i\xi} \tag{8.83}$$

So, we have the Fourier transform for the function $H(x) - \frac{1}{2}$, but not the one for $H(x)$. Note that this is in line with our initial requirement that the functions amenable to Fourier transform must be integrable; $H(x) - 1/2$ has, by symmetry properties, a zero integral and thus qualifies (N.B., one must be careful with such arguments on infinite domains– sometimes the infinite does things that are quite counterintuitive).

我们可以利用傅里叶变换的线性性继续推导。

$$\mathscr{F}\left[H(x)-\frac{1}{2}\right]=\frac{1}{\sqrt{2\pi}}\frac{1}{i\xi} \tag{8.84}$$

$$\mathscr{F}\left[H(x)\right]-\mathscr{F}\left[\frac{1}{2}\right]=\frac{1}{\sqrt{2\pi}}\frac{1}{i\xi} \tag{8.85}$$

根据有关狄拉克δ函数的傅里叶变换的材料，我们有

$$\mathscr{F}\left[H(x)\right]-\sqrt{\pi}\delta(\xi)=\frac{1}{\sqrt{2\pi}}\frac{1}{i\xi} \tag{8.86}$$

在两边同时加上 $\sqrt{\pi}\delta(\xi)$，即可得到最终结果。

$$\mathscr{F}\left[H(x)\right]=\sqrt{\frac{\pi}{2}}\delta(\xi)+\frac{1}{\sqrt{2\pi}}\frac{1}{i\xi} \tag{8.87}$$

上述推导过程可以重复进行，以定义负海维赛德函数。

$$H_{-}(x)=\begin{cases}1 & x<0\\ 0 & x\geq 0\end{cases} \tag{8.88}$$

以便

$$\mathscr{F}\left[H(x)\right]=\sqrt{\frac{\pi}{2}}\delta(\xi)-\frac{1}{\sqrt{2\pi}}\frac{1}{i\xi} \tag{8.89}$$

8.9.3 The Fourier Transform of the Constant Function

Once we have defined the Fourier transform of the delta function and the unit step function, some somewhat odd things happen. For example, we have found the Fourier transform for the unit step function, which is a decidedly non-integrable function on the real line. The *reason* that these odd things happen is because we have allowed the delta function into our world of acceptable functions. While we can make sense of the Fourier transform of the delta function, we must remember that the delta function is actually a *distribution*, and not a function at all! While we have managed to illustrate the properties of the Fourier transform of the delta and Heaviside functions, it has used some arguments that are not necessarily obvious ones.

We have another "trick up our sleeves", so to speak, to develop the Fourier transform of a tempered distribution but using our existing knowledge of the Fourier transform. The constant function is a tempered distribution (it is the polynomial defined by cx^0, and grows at rate zero). To achieve this, we can use the results for the Heaviside and negative Heaviside functions above. Note the following:

$$c(H(x)+H_{-}(x))=c,\ \ x\in(-\infty,\infty) \tag{8.90}$$

However, we already have the Fourier transform for these two functions from the discussion in the previous section. Thus

$$\mathscr{F}[c(H(x)+H_{-}(x))] = \mathscr{F}[c] \tag{8.91}$$

and from the Fourier transforms for H and H_{-} we have

$$c\left(\sqrt{\frac{\pi}{2}}\delta(\xi)+\frac{1}{\sqrt{2\pi}}\frac{1}{i\xi}\right)+c\left(\sqrt{\frac{\pi}{2}}\delta(\xi)-\frac{1}{\sqrt{2\pi}}\frac{1}{i\xi}\right)=\mathscr{F}[c] \tag{8.92}$$

Upon simplifying and rearranging, we have

$$\mathscr{F}[c] = c\sqrt{\frac{\pi}{2}}\delta(\xi)+c\sqrt{\frac{\pi}{2}}\delta(\xi) \tag{8.93}$$

Recall that $\mathscr{F}[c]$ is just the integral

$$\mathscr{F}[c] = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} c\exp(-ix\xi)\,dx \tag{8.94}$$

Substituting this on the left-hand side gives

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} c\exp(-ix\xi)\,dx = c\sqrt{\frac{\pi}{2}}\delta(\xi)+c\sqrt{\frac{\pi}{2}}\delta(\xi) \tag{8.95}$$

And simplifying this result gives us

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} 1\times\exp(-ix\xi)\,dx = \sqrt{2\pi}\delta(\xi) \tag{8.96}$$

Rewriting the left-hand side as $\mathscr{F}[1]$ gives us a final result of

$$\mathscr{F}[1] = \sqrt{2\pi}\delta(\xi) \tag{8.97}$$

or, equivalently

$$\mathscr{F}\left[\frac{1}{\sqrt{2\pi}}\right] = \delta(\xi) \tag{8.98}$$

8.9.4 The Fourier Transform of $\sin(ax)$ *and* $\cos(ax)$

The two functions $\sin(ax)$ and $\cos(ax)$ are interesting functions to consider transforming. Again, they do not meet the requirements we set out for functions that have a Fourier transform, because they are not square integrable. However, clearly we have that $sin(x)^2 \leq 1$ and $cos(x)^2 \leq 1$ for all values of x, and we now know that the function $f(x) = 1$ has a Fourier transform. Thus, it is reasonable to wonder if both the sine and cosine functions can transformed in some way that makes sense to us without resorting to the theory of tempered distributions. We can use the following construction based on our existing notions of the Fourier transform, but again relaxing the assumption that we can relax the condition that functions must be square-integrable.

First, we note the following identities

$$\cos(ax) = \frac{1}{2}(\exp[iax] + \exp[-iax]) \tag{8.99}$$

$$\sin(ax) = -i\frac{1}{2}(\exp[iax] - \exp[-iax]) \tag{8.100}$$

考察 g the table of Fourier transforms, note that transform number 9 gives the needed transform, i.e.,

$$\mathscr{F}\left[\sqrt{\frac{1}{2\pi}}\exp(-iax)\right] = \delta(\xi - a) \tag{8.101}$$

我们可以相当直接地利用这些信息来证明

$$\mathscr{F}[\cos(ax)] = \sqrt{\frac{\pi}{2}}[\delta(\xi + a) + \delta(\xi - a)] \tag{8.102}$$

$$\mathscr{F}[\sin(ax)] = i\sqrt{\frac{\pi}{2}}[\delta(\xi + a) - \delta(\xi - a)] \tag{8.103}$$

注意，对于 $a = 0$，这分别化为常函数 $f(x) = 1$ 和零函数 $f(x) = 0$ 的相应傅里叶变换。

例 *8.7*（利用傅里叶变换求解泊松方程）。在关于分离变量法的章节中，我们探讨了齐次拉普拉斯方程的一些解。拉普拉斯方程仍然只涉及两个自变量，但在这种情况下，这两个变量都是空间变量。我们已经证明，对于这种情况，分离变量法的运作方式与处理其他含两个自变量的问题基本相同。

我们可以利用傅里叶变换求解无限平面上含两个空间变量的拉普拉斯方程，但我们将看到，这需要分别对每个变量使用变换。事实上，在这个例子中，我们还能展示傅里叶变换的另一项优势：使用变换方法时，可以很容易地推导出非齐次问题的解。

To start, recall that when the Laplace equation has a source term it is usually called the Poisson equation. Consider the following Poisson equation on the infinite 2-dimensional plane.

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = -\delta(x)\delta(y) \qquad -\infty < x < \infty,\ -\infty < x < \infty$$

*B.C.*1 $$u(x,y) \to 0 \ \ as\ x^2 + y^2 \to \infty$$

*B.C.*2*a* $$\left.\frac{\partial u}{\partial x}\right|_{(x,y)} \to 0 \ \ as\ x^2 + y^2 \to \infty$$

*B.C.*2*b* $$\left.\frac{\partial u}{\partial y}\right|_{(x,y)} \to 0 \ \ as\ x^2 + y^2 \to \infty$$

Here, we have used $d^2 = x^2 + y^2$ as the square of the distance between any point and the origin as the appropriate measure on the plane.

现在，考虑对方程两边关于变量 x 进行变换。回顾一下，傅里叶变换定义为

$$\mathscr{F}[f(x)] = \hat{f}(\xi) = \int_{-\infty}^{\infty} f(x)\, e^{-ix\xi}\, dx$$

对x–变量进行变换，我们得到结果。

$$-\xi^2 \hat{u}(\xi, y) + \frac{\partial^2 \hat{u}}{\partial y^2} = \sqrt{\frac{1}{2\pi}}\delta(y)$$

请注意，在这里我们仅通过对变量 x 和 ξ应用傅里叶变换，变换了变量 x，这与目前为止的惯例一致。此时，我们借助变换消去了两个导数项中的一个。不过，我们仍然面对二阶非齐次方程。你可能还记得，在我们复习常微分方程时，这类方程可以用参数变易法求解，但求得解仍需付出一些努力。

作为另一种方法，可以考虑进行第二次傅里叶变换，但这次是在变量 y 上进行。注意，我们有 $-\infty < y < \infty$，而这些恰好是傅里叶变换适用的条件。我们完全可以使用两个独立的变换，分别对 x 和 y 中的量进行变换，但必须额外注意跟踪所涉及的变量。假设我们使用变换对（x ξ）表示关于 x 的自变量傅里叶变换对，并使用（y η）表示关于 y 的自变量变换对。那么，y 坐标的变换可表示为

$$\mathscr{F}_y[f(y)] = \bar{f}(\eta) = \int_{-\infty}^{\infty} f(y)\, e^{-iy\eta}\, dy$$

Now, with this definition in place, we can conduct a second transform of our Poisson equation. This gives

$$\mathscr{F}_y\left[-\xi^2 \hat{u}(\xi, y)\right] + \mathscr{F}_y\left[\frac{\partial^2 \hat{u}}{\partial y^2}\right] = \mathscr{F}_y\left[-\sqrt{\frac{1}{2\pi}}\delta(y)\right]$$

Recalling that constants pass through integrals, and that ξ is independent of y, then we have the result

$$-\xi^2 \bar{\hat{u}}(\xi, \eta) - \eta^2 \bar{\hat{u}}(\xi, \eta) = -\sqrt{\frac{1}{2\pi}}\sqrt{\frac{1}{2\pi}}$$

where $\bar{\hat{u}}$ represents the function u Fourier transformed in both the x and y variables. Note that we now have a function of ξ and η, indicating that both spatial independent variables have been transformed. We are also left with a purely algebraic equation, which is very convenient. Solving this for $\bar{\hat{u}}$ gives

$$\begin{aligned}\bar{\hat{u}} &= -\sqrt{\frac{1}{2\pi}}\sqrt{\frac{1}{2\pi}}\frac{1}{\xi^2+\eta^2} \\ &= -\frac{1}{2}\sqrt{\frac{1}{2\pi}}\left\{\sqrt{\frac{2}{\pi}}\frac{1}{\xi^2+\eta^2}\right\}\end{aligned}$$

Inverting this requires only that we keep careful track of which variables we are working on, and the fact that ξ and η are independent. First, we can find the inverse transform with respect to η. This is given symbolically by

$$\mathscr{F}_y^{-1}\left[\bar{\hat{u}}\right] = -\frac{1}{2}\sqrt{\frac{1}{2\pi}}\mathscr{F}_y^{-1}\left[\left\{\sqrt{\frac{2}{\pi}}\frac{1}{\xi^2+\eta^2}\right\}\right]$$

Note that the term on the right-hand side in braces matches entry 17 from the table of Fourier transforms. To be clear, here we set $a = \xi$, and note that the requirement that $a > 0$ implies $\xi \Rightarrow |\xi|$. This yields the inverse transform (where $\eta \to y$) as follows

$$\hat{u} = -\frac{1}{2}\sqrt{\frac{1}{2\pi}}\left\{\frac{1}{|\xi|}\exp(-\xi|y|)\right\}$$

Now we have only the ξ variable to invert. Looking at the expression, we can see that entry number 20 is very close to our needed inverse. To assist in inverting, we can rearrange the result as follows

$$\hat{u}(\xi,y) = -\frac{1}{2}\sqrt{\frac{1}{2\pi}}\frac{1}{\sqrt{2\pi}}\left\{\frac{\sqrt{2\pi}}{\xi}\exp(-\xi|y|)\right\}$$

Note that now the shorthand notation for the transform in the y-variable is given by an overbar so that we can distinguish between transforms in the two variables.

In the expression above, the term in braces matches the table exactly, so we find the final inverse transform to give us the result

$$u(x,y) = -\frac{1}{4\pi}\ln(x^2+y^2)$$

This is a well-known result, and is often called the *fundamental* solution for the Poisson problem on the infinite plane. Noting that we have defined $d(x,y) = \sqrt{x^2+y^2}$, we can use the properties of logarithms to express this result by

$$
\begin{aligned}
u(x,y) &= -\frac{1}{4\pi}\ln(x^2+y^2) \\
&= -\frac{1}{2\pi}\ln\left[(x^2+y^2)^{\frac{1}{2}}\right] \\
&= -\frac{1}{4\pi}\ln[d(x,y)]
\end{aligned}
$$

最后需要说明的是，如果我们在该问题中加上任意 常数，那么上面的解仍然是一个解。由于我们放宽了函数必须平方可积这一要求，因此对于无限平面，我们没有额外的约束条件来确定这个常数。因此，我们必须将我们的解视为一个解的等价类，其中每个解仅相差某个常数。这也表明，$\ln(a^2+x^2)$ 的傅里叶变换在某些情况下可能包含第二部分，即含有狄拉克δ函数$c\delta(\xi)$ 的部分，其中 c 是某个尚未确定的常数。这个额外的狄拉克 δ 函数在进行逆变换时，会生成常数项；从技术上说，该常数项属于我们的潜在解集合。下面给出了这个外观颇为异常的函数图像。

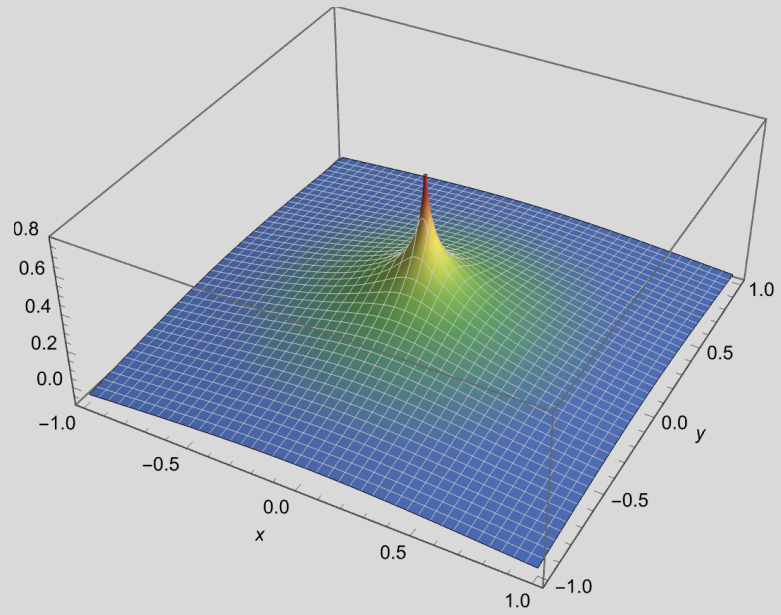


Fig. 8.7: The solution for the 2-dimensional Poisson equation. Note that the lower boundary of the vertical axis is set to $-1/10$.

8.10 *双曲型方程再探

如第6章所述，考察应用于连续介质的动量守恒，可以得到二阶双曲方程。然而，它们并不像热方程/扩散方程这类抛物型方程那样，可以通过简单的守恒定律与通量定律相结合来推导。借助傅里叶变换，我们可以从一阶描述推导出二阶双曲方程。这种做法并不常见；在教科书中，这两个方程很少以这种方式联系起来。不过，采用这种方法处理问题能提供另一种视角，或许也能揭示比常规推导更多的物理内涵。

首先， 让我们再次看看一阶波动方程的傅里叶变换。回想一下，我们有

$$\frac{\partial u}{\partial t} = -c\frac{\partial u}{\partial x} \tag{8.104}$$

$$I.C. \qquad u(x,0) = f(x) \tag{8.105}$$

通常要求函数u和f平方可积（或者，回想一下扩展后的情形，它们至少是缓增分布）。该表达式的傅里叶变换为

$$\frac{d\hat{u}}{dt} = -i\xi c \tag{8.106}$$

$$\hat{u}(\xi,0) = \hat{f}(\xi) \tag{8.107}$$

这个问题的解相当直接，可以得到如下形式的指数函数

$$\hat{u}(\xi,t) = \underline{\hat{f}(\xi)\exp(-i\xi ct)} \tag{8.108}$$

现在，对于二阶方程，回想一下我们有

$$\frac{\partial^2 u}{\partial t^2} = c^2\frac{\partial^2 u}{\partial x^2} \tag{8.109}$$

$$I.C.1 \qquad u(x,0) = f(x) \tag{8.110}$$

$$I.C.2 \qquad \left.\frac{\partial u}{\partial t}\right|_{(\xi,0)} = 0 \tag{8.111}$$

而变换为

$$\frac{d^2\hat{u}}{dt^2} - c^2\xi^2\hat{u} = 0 \tag{8.112}$$

$$I.C.1 \qquad \hat{u}(\xi,0) = \hat{f}(\xi) \tag{8.113}$$

$$I.C.2 \qquad -i\xi\hat{u}(\xi,0) = 0 \tag{8.114}$$

这个二阶齐次常微分方程的解相当直接。它是

$$u(\xi,t) = \tfrac{1}{2}\underline{\hat{f}(\xi)\exp(-ic\xi t)} + \tfrac{1}{2}\hat{f}(\xi)\exp(ic\xi t) \tag{8.115}$$

At this point, it is possible to see a close relationship between the solution to the first-order wave equation and the second-order wave equation in Fourier space. Note that the first order wave equation has a solution with the transform of the initial conditions modified by a single exponential. The solution to the second-order wave equation has a similar solution, except the initial condition is modified by two different exponentials. You may recall that exponentials of the form $\exp(-i\xi ct)$ or $\exp(i\xi ct)$ simply *shift* their inverse functions back in real space by the amount $\pm ct$, depending on the sign of the argument of the exponential. Note also, that the two underlined terms in Eqs. (8.108) and (8.115) are identical!

在傅里叶空间中，很容易看出这两个解彼此相关。一阶解由初始条件向前平移得到，平移距离为ct。二阶方程包含两项，两项都是初始条件的缩放版本（缩放因子为 12），并且都平移ct，但方向相反。

注意到指数函数的作用仅仅是平移反演后的函数，因此这两个方程的反演相对容易。由此，我们得到以下两个解。

一阶 $$u(x,t) = f(x-ct) \tag{8.116}$$

二阶 $$u(x,t) = \tfrac{1}{2}f(x-ct) + \tfrac{1}{2}f(x+ct) \tag{8.117}$$

利用傅里叶变换分析无限直线上的一阶和二阶波动方程，可以较为容易地看出这两个方程之间的对应关系。两个表达式都会导致空间平移相同的距离 ct。如前所述，一阶波动方程适用于只有一个波沿单一方向传播的问题（例如，溶质脉冲随流体在平推流反应器中移动）。二阶波动方程对应两个波，每个波的高度都是原来的一半，并且沿相反方向移动。这里可以设想一条狭窄灌渠中的扰动（例如，假设有人把一块大石头扔进灌溉渠）；波从初始扰动的位置开始，但会向两个方向传播。

All of this was predicated on the requirement that the initial velocity $\partial u/\partial t$ was identically zero in our analysis. One of the additional feature that make the second-order wave equation useful is that it contains more unknown constants of integration (thus, it requires one additional condition on $\partial u/\partial t$ in order to determine an unique solutions). When $\partial u/\partial t \neq 0$, other interesting phenomena are manifest, and the similarity between the first-order and second-order wave equations begins to be less obvious. However, the examples given above should provide some substantial evidence that, fundamentally, the two equations are indeed related! The second-order wave equation can be thought of, quite accurately and correctly, as a generalization of the first-order wave equation.

Appendix. Fourier Transform Table

Table 8.1: Fourier transform pairs.

	$f(x)$	$\hat{f}(\xi)$	约束条件
1.	$\frac{\partial f(x)}{\partial x}$	$i\xi\hat{f}(\xi)$	
2.	$\frac{\partial^2 f(x)}{\partial x^2}$	$-\xi^2\hat{f}(\xi)$	
3.	$f(ax)$	$\frac{1}{a}\hat{f}(\xi/a)$	$a>0$
4.	$f(x-a)$	$\exp(-ia\xi)\hat{f}(\xi)$	
5.	$\sqrt{\frac{1}{2\pi}}$	$\delta(\xi)$	
6.	0	0	
7.	1	$\sqrt{2\pi}\delta(\xi)$	
8.	$\delta(x)$	$\sqrt{\frac{1}{2\pi}}$	
9.	$\delta(x-a)$	$\sqrt{\frac{1}{2\pi}}\exp(-ia\xi)$	
10.	$\sqrt{\frac{1}{2\pi}}\exp(-iax)$	$\delta(\xi-a)$	
11.	$H_0(x)$	$\sqrt{\frac{\pi}{2}}\delta(\xi)$	
12.	$H(x)$	$\sqrt{\frac{\pi}{2}}\delta(\xi)+\frac{1}{\sqrt{2\pi}}\frac{1}{i\xi}$	

	$f(x)$	$\hat{f}(\xi)$	约束条件
13.	赫维赛德阶跃函数$(x-a)$	$1-$赫维赛德阶跃函数$(a-x)$	
	$\sqrt{\frac{\pi}{2}}\delta(\xi)+\frac{1}{\sqrt{2\pi}}$指数函数$(-ia\xi)$	$i\xi$	
14.	赫维赛德阶跃函数$(a-x)$	$1-$赫维赛德阶跃函数$(x-a)$	
	$\sqrt{\frac{\pi}{2}}\delta(\xi)-\frac{1}{\sqrt{2\pi}}$指数函数$(-ia\xi)$	$i\xi$	
15.	$e^{-a^2x^2}$	$\frac{1}{a\sqrt{2}}e^{-\frac{\xi^2}{4a^2}}$	$a>0$
16.	$e^{-a\lvert x\rvert}$	$\sqrt{\frac{2}{\pi}}\frac{a}{a^2+\xi^2}$	$a>0$
17.	1	$a\,e$	
18.	$xe^{-a\lvert x\rvert}$	$\sqrt{\frac{2}{\pi}}\frac{1}{a^2+\xi^2}$	$a>0$
19.	$-2\sqrt{\frac{2}{\pi}}ia\xi(\xi^2+a^2)^2$	$a>0$	
20.	$\frac{1}{\sqrt{2}}a$ 指数函数$\left(-\frac{x^2}{4}a\right)$	指数函数$(-a\xi^2)$	$a>0$
21.	自然对数(x^2+a^2)	$\sqrt{2\pi}$指数函数$(-a\lvert\xi\rvert)$	$a>0$
22.	xe^{-ax^2}	$-i2\sqrt{2}\,\frac{\xi}{a^{3/2}}$指数函数$\left(-\frac{\xi^2}{4a}\right)$	$a>0$
23.	$x^2e^{-ax^2}$	$-4\sqrt{2}\,\frac{2-\xi^2}{a^{5/2}}$指数函数$\left(-\frac{\xi^2}{4a}\right)$	$a>0$
24.	正弦(ax)	$i\sqrt{\frac{\pi}{2}}[\delta(\xi+a)-\delta(\xi-a)]$	
	余弦(ax)	$\sqrt{\frac{\pi}{2}}[\delta(\xi+a)+\delta(\xi-a)]$	

	$f(x)$	$\hat{f}(\xi)$	约束条件
25.	$\arctan(x)$	$-\frac{i}{\xi}\sqrt{\frac{\pi}{2}}\exp(-\lvert\xi\rvert)$	
26.	$\cos(\pi x)[H(x+1)-H(x-1)]$	$\sqrt{\frac{2}{\pi}}\frac{\xi\sin(\xi)}{\pi^2-\xi^2}$	
27.	$\sin(\pi x)[H(x+1)-H(x-1)]$	$-\sqrt{2\pi}\frac{i\sin(\xi)}{\pi^2-\xi^2}$	
28.	$\operatorname{erf}(x)$	$-\frac{1}{\xi}\sqrt{\frac{2}{\pi}}i\exp\left(-\frac{\xi^2}{4}\right)$	
29.	$\operatorname{erf}(a+x)-\operatorname{erf}(a-x)$	$\frac{2}{\xi}\sqrt{\frac{2}{\pi}}\exp\left(-\frac{\xi^2}{4}\right)\sin(a\xi)$	
30.	$\operatorname{erf}(x)-\operatorname{erf}(x-a)$	$\frac{1}{\xi}\sqrt{\frac{2}{\pi}}i\exp\left[-\frac{1}{4}\xi(\xi+4ia)\right]$ $\times[1-\exp(ia\xi)]$	
31.	$T(x;-a,a)=$ $\begin{cases}(1+\frac{1}{a}x), & -a<x<0\\ (1-\frac{1}{a}x), & 0<x<a\end{cases}$	$-\sqrt{\frac{2}{\pi}}\frac{(\cos(a\xi)-1)}{a\xi^2}$	$a>0$
32.	$B(x;\ -a,a)=$赫维赛德阶跃函数$(x+a)-$赫维赛德阶跃函数$(x-a)$	$\sqrt{\frac{2}{\pi}}$正弦$(a\xi)\xi a$	$a>0$

	$f(x)$	$\hat{f}(\xi)$	约束条件
33.	$B(x+b_0;-a,a)$ $=H(x+a+b_0)-H(x-a+b_0)$	$\sqrt{\frac{2}{\pi}}\frac{\sin(a\xi)}{\xi}\exp(ib_0\xi)$	$a>0$ $b_0>0$
34.	$B(x-b_0;-a,a)$ $=H(x+a-b_0)-H(x-a-b_0)$	$\sqrt{\frac{2}{\pi}}\frac{\sin(a\xi)}{\xi}\exp(-ib_0\xi)$	$a>0$ $b_0>0$
35.	$\frac{1}{a^2+x^2}$	$\frac{1}{a}\sqrt{\frac{\pi}{2}}\exp(-a\lvert\xi\rvert)$	
36.	$\frac{a}{x^2+a^2}$	$\sqrt{\frac{\pi}{2}}\exp(-a\lvert\xi\rvert)$	$a>0$
37.	$\frac{x}{a^2+x^2}$	$-\mathrm{sgn}(\xi)ia\sqrt{\frac{\pi}{2}}\exp(-a\lvert\xi\rvert)$	$a>0$
38.	$\frac{a^4}{a^4+x^4}$	$\left(\frac{1}{4}+\frac{i}{4}\right)\sqrt{\pi}a\exp\left(-\frac{(1+i)a\xi}{\sqrt{2}}\right)\times$ $\left[e^{\sqrt{2}a\xi}\left(e^{i\sqrt{2}a\xi}-i\right)H(-\xi)\right.$ $\left.-ie^{i\sqrt{2}a\xi}H(\xi)+H(\xi)\right]$	$a>0$

注意。如果您使用 Mathematica 执行傅里叶变换，则需要将以下内容设置为 FourierTransform 和 InverseFourierTransform 命令的一个选项：傅里叶参数 -> {0, -1}。这样可以确保傅里叶变换的定义 与本文采用的定义完全一致。有关更多详细信息，请参阅 Mathematica 帮助页面。函数注释：

erf 误差函数，$\mathrm{erf(x)}=\frac{2}{\sqrt{\pi}}\int \mathrm{x}\ 0\ \exp(-\eta^2)\mathrm{d}\eta$.erfc 互补误差函数 $1-\mathrm{erf(t)}$

δ 狄拉克δ函数 赫维赛德阶跃函数 B(x; –a,a) x 上的矩形函数 $\in[-\mathrm{a,a}]$T(t; –a,a) x 上的对称三角函数 $\in[-\mathrm{a,a}]$, 最大值为 1，位于 $\mathrm{t}_1=0$sgn(x) 符号函数；$\mathrm{sgn(x)}=1,\ \mathrm{x}>0,$ $\mathrm{sgn(x)}=-1,\ \mathrm{x}<0,\ \mathrm{sgn(x)}=0,\ \mathrm{x}=0$$\mathrm{H_0(x)}=\begin{cases}-^2\ \mathrm{x}\ \frac{1}{2} & >0\ \mathrm{x}\end{cases}$ 零均值赫维赛德函数

习题

练习题

1. 证明：$\mathscr{F}[f(x-a)] = e^{-iax}\hat{f}(\xi)$。
2. 证明：$\mathscr{F}[f(ax)] = \frac{1}{|a|}\hat{f}\left(\frac{\xi}{a}\right)$

3. Prove that $(f*g)(x) = (g*f)(x)$. Hint: this can be done using a transform of variables of the form $\eta = (x-\xi)$.
4. Work out the following transforms in detail.

$$\mathscr{F}[\cos(x)] = \sqrt{\frac{\pi}{2}}[\delta(\xi+1)+\delta(\xi-1)] \tag{8.118}$$

$$\mathscr{F}[\sin(x)] = i\sqrt{\frac{\pi}{2}}[\delta(\xi+1)-\delta(\xi-1)] \tag{8.119}$$

   Begin by using the Euler identity to transform the cosine and sine functions to equivalent functions expressed in terms of the complex exponential. Then, directly integrate the resulting expressions using the definition of the Fourier transform given by Eq. 8.33. Finally, you can use the definition of the transform given in the table under entry number 9 to identify the delta function in the result.

5. Note the following two transforms (entries number 9 and 10 in the table of Fourier transforms).

$$\mathscr{F}[\delta(x-a)] = \sqrt{\frac{1}{2\pi}}\exp(-ia\xi)$$

$$\mathscr{F}\left[\sqrt{\frac{1}{2\pi}}\exp(-iax)\right] = \delta(\xi-a)$$

   There appears to be some symmetry between the forward and inverse transforms. To see why this is true, start with the following inverse transform

$$\left[\sqrt{\frac{1}{2\pi}}\exp(-iax)\right] = \mathscr{F}^{-1}[\delta(\xi-a)]$$

   Note that using the definition of the inverse, this is

$$\sqrt{\frac{1}{2\pi}}\exp(-iax) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\delta(\xi-a)\exp[i\xi x]\,d\xi$$

   Propose the change of variables $\xi = -\eta$ to transform this integral. Take extra care in transforming the bounds of integration! After the transformation is complete, switch back to the original variable ξ by simply renaming $\eta \to \xi$.
   Finally, note that $\delta(x-a) = \delta(a-x)$. This symmetry relationship allows us to develop a result that shows that the Fourier transform of the delta function can be used in reverse. That is, if you know the transform of the delta function, then you also have the transform of the function $1/\sqrt{2\pi}\exp[-iax]$ by

反转函数及其变换的角色，并交换变量的位置。为了明确起见，我们希望根据上述推导证明，由 9 可推出由 10 给出的变换同样成立。

关于函数 $\delta(x-a)$，有一个关键特征使得这种推广成为可能——那是什么？你能否对函数 $f(x)$ 的条件作出一般性陈述，使得傅里叶变换能够以这种“逆向”方式使用？

Applied and More Challenging Problems

6. Solve the following heat/diffusion equation problem using the Fourier transform.

$$\frac{\partial u}{\partial t} = D\frac{\partial^2 u}{\partial x^2}, \quad x \in (-\infty, \infty) \tag{8.120}$$

$$\textit{condition } 1 \qquad u(x,t) \to 0 \text{ as } |x| \to \infty \text{ for all } x \text{ and } t \tag{8.121}$$

$$\textit{condition } 2 \qquad \frac{\partial u}{\partial x} \to 0 \text{ as } |x| \to \infty \text{ for all } x \text{ and } t \tag{8.122}$$

$$I.C. \qquad u(x,0) = \exp(-x^2) \quad x \in (-\infty, \infty) \tag{8.123}$$

7. Solve the following heat/diffusion equation problem using the Fourier transform.

$$\frac{\partial u}{\partial t} = D\frac{\partial^2 u}{\partial x^2}, \quad x \in (-\infty, \infty) \tag{8.124}$$

$$\textit{condition } 1 \qquad u(x,t) \to 0 \text{ as } |x| \to \infty \text{ for all } x \text{ and } t \tag{8.125}$$

$$\textit{condition } 2 \qquad \frac{\partial u}{\partial x} \to 0 \text{ as } |x| \to \infty \text{ for all } x \text{ and } t \tag{8.126}$$

$$I.C. \qquad u(x,0) = \delta(x) \tag{8.127}$$

8. Solve the following heat/diffusion equation problem using the Fourier transform. Solve this problem two ways. (1) First, try the transformation $u(x,t) = w(x,t)\exp(-kt)$ to develop a new problem in w that does not contain the reaction term, and then solve that using the Fourier transform. (2) Second, try solving the problem directly using the Fourier transform. Which method seemed easier to you?

$$\frac{\partial u}{\partial t} = D\frac{\partial^2 u}{\partial x^2} - ku, \quad x \in (-\infty, \infty) \tag{8.128}$$

$$\textit{condition } 1 \qquad u(x,t) \to 0 \text{ as } |x| \to \infty \text{ for all } x \text{ and } t \tag{8.129}$$

$$\textit{condition } 2 \qquad \frac{\partial u}{\partial x} \to 0 \text{ as } |x| \to \infty \text{ for all } x \text{ and } t \tag{8.130}$$

$$I.C. \qquad u(x,0) = \delta(x) \tag{8.131}$$

9. **A problem of pure convection.** In some instances, such as in packed beds or in rivers, it is permissible to make the approximation that the transport is primarily *convective* (or sometimes the word *advective* is used), meaning that dissolved species in the fluid (gas or liquid) phase move with the fluid with very little spreading (or dispersion). If one has, for example, a first-order reaction representing the conversion

of a chemical compound in a river to some end product, one possible model for the reactant given by the following first-order PDE

$$\frac{\partial u}{\partial t} = -v_0 \frac{\partial u}{\partial x}, \quad x \in (-\infty, \infty) \tag{8.132}$$

condition 1 $$u(x,t) \to 0 \text{ as } |x| \to \infty \text{ for all } x \text{ and } t \tag{8.133}$$

I.C. $$u(x,0) = \beta \exp\left(-\alpha x^2\right) \tag{8.134}$$

该模型假设初始条件是浓度的高斯函数分布（沿河流路径随距离 x 变化，并且在垂直于x的平面上保持常数）。这假设河流可以近似为具有一维过程几何。

使用傅里叶变换求解上述问题。然后，以 $(1/b)u(x,t)$ 的形式绘制该问题的解，其中 $x \in [-10, 125]$ m 且 $u \in [0,1]$。使用 $v_0 = 10\text{km} \cdot \text{h}^{-1}$, $\alpha = 0.1$ 和 $\beta = 1$。绘制时间 $t = 0.1$,1,5 和 10 小时时的函数。注意，我们没有指定 β 的单位，但它们必须与u的单位相同；因此u/b 是无量纲的。

10. **纯对流和反应问题**。假设对于上述给出的河流纯对流问题，现在还存在一阶变换反应。初始浓度为高斯函数时，其浓度将由以下一阶偏微分方程规定

$$\frac{\partial u}{\partial t} = -v_0 \frac{\partial u}{\partial x} - \gamma u, \quad x \in (-\infty, \infty) \tag{8.135}$$

condition 1 $$u(x,t) \to 0 \text{ as } |x| \to \infty \text{ for all } x \text{ and } t \tag{8.136}$$

I.C. $$u(x,0) = \beta \exp\left(-\alpha x^2\right) \tag{8.137}$$

This model assumes that the initial condition is a Gaussian distribution of concentration (changing in distance, x along the river path, and constant in planes perpendicular to x). This assumes that the river can be approximated as having a 1-dimensional process geometry. Here, γ is a first-order kinetic reaction rate; set $\gamma = 0.1\ \text{h}^{-1}$.
Solve the problem above using Fourier transforms. Then, plot the solution to the problem in the form $(1/b)u(x,t)$ for $x \in [-10, 125]$ and $u \in [0,1]$. Use $v_0 = 10\ \text{km} \cdot \text{h}^{-1}$, $\alpha = 0.1$, and $\beta = 1$. Plot the function at times $t = 0.1, 1, 5$ and 10 hours. Note that we have not specified the *units* for β, but they must be the same as those for u; thus u/b is dimensionless.

11. **A convection-diffusion problem.** Suppose we would like to solve the following convection-diffusion problem using the Fourier transform.

$$\frac{\partial u}{\partial t} = D\frac{\partial^2 u}{\partial x^2} - v_0 \frac{\partial u}{\partial x}, \quad x \in (-\infty, \infty) \tag{8.138}$$

condition 1 $$u(x,t) \to 0 \text{ as } |x| \to \infty \text{ for all } x \text{ and } t \tag{8.139}$$

条件 2 $$\frac{\partial u}{\partial x} \to 0 \text{ as } |x| \to \infty \text{ for all } x \text{ and } t \tag{8.140}$$

I.C. $$u(x,0) = \beta \exp(-x^2) \tag{8.141}$$

Note: you will want to use the transform given by #4 in the table of transforms.

12. **一个对流-扩散-反应问题**。假设我们希望使用傅里叶变换求解以下对流-扩散问题。

$$\frac{\partial u}{\partial t} = D\frac{\partial^2 u}{\partial x^2} - v_0\frac{\partial u}{\partial x} - ku, \quad x \in (-\infty, \infty) \tag{8.142}$$

$$\textit{condition } 1 \qquad u(x,t) \to 0 \text{ as } |x| \to \infty \text{ for all } x \text{ and } t \tag{8.143}$$

$$\textit{condition } 2 \qquad \frac{\partial u}{\partial x} \to 0 \text{ as } |x| \to \infty \text{ for all } x \text{ and } t \tag{8.144}$$

$$I.C. \qquad u(x,0) = \beta\delta(x - x_0) \tag{8.145}$$

注意：你会需要使用变换表中的第 4 个变换。

13. **无限介质中的波动方程**。波动方程的行为与热传导/扩散方程有很大不同。在目前为止我们所建立的形式中，波在传播过程中不会发生阻尼。因此，在无限介质中（即在数学上，对于一维情况而言，实数轴上），位于$x = 0$ 附近、$t = 0$ 时刻的初始扰动，将从扰动位置向外朝无穷远处传播，对所有$t > 0$均如此。如果波动方程是二阶方程（将在下文中具体写出），那么随着时间t的增加，初始扰动将沿 $+x$ 和 $-x$ 两个方向向无穷远处传播。实际上，我们可以使用傅里叶变换来求解无限介质中的这一方程，就像求解热传导/扩散方程一样。下面的步骤将引导我们完成这一求解过程。

a. Begin by transforming the wave equation *and* its initial condition by conducting a Fourier transform with respect to the variable x. For the wave equation, assume the following balance equation and ancillary conditions

$$\begin{aligned} \frac{\partial^2 u}{\partial t^2} &= c^2\frac{\partial^2 u}{\partial u^2} \\ u(x,0) &= H(x+1) - H(x-1) \\ \left.\frac{\partial u}{\partial t}\right|_{(x,0)} &= 0 \\ u(x,t) \text{ and } \partial u/\partial x &\to 0 \text{ as } |x| \to \infty \end{aligned}$$

Note that here we have both and initial condition and an initial derivative for u with respect to time. As we have discussed, we need one ancillary condition for each derivative in time and in space. Because we have a second-order derivative in time for this equation, we need *two* conditions to eliminate the unknown constants of integration that will result. Also recall that the conditions on u in space as $|x| \to \infty$ serve as the two necessary ancillary conditions in space.

b. The result from your previous step should take the form

$$\frac{\partial^2 \hat{u}}{\partial t^2} + c^2 \xi^2 \hat{u} = 0$$
$$\left.\frac{d\hat{u}}{dt}\right|_{(\xi,0)} = 0$$
$$\hat{u}(\xi,0) = \frac{1}{\xi}\sqrt{\frac{2}{\pi}}\sin(\xi)$$

This equation may initially look challenging, but it is in fact just a linear second-order ODE. We have studied the solutions to such equations in Chap. 3. Recall that the solution for an equation like this is found by first developing the characteristic equation, and then solving for the roots of that equation. The form of the roots tell you which of three cases the solution belongs to. Do not forget to use the two ancillary conditions to eliminate the two constants of integration!

c. Your result for the solution is not on the Fourier transform table as such. However, we can generate the solution as follows. First, obtain the transforms given on the table as entries 33 and 34. Note that each of those involve an imaginary exponential function that can be converted to trigonometric functions by using the Euler identity: $\exp(ib_0\xi) = \cos(b0\xi) + i\sin(b_0\xi)$ and $\exp(-ib_0\xi) = \cos(b0\xi) - i\sin(b_0\xi)$. So, make this substitution to these to transform pairs. Then, average the two results (i.e., take $\frac{1}{2}$ of the sum of the two transforms). You should now have a result that matches your solution in Fourier space, and this will allow you to compute the inverse transform. Your final result should be as follows

$$u(x,t) = \underbrace{\frac{1}{2}[H(x+1+ct) - H(x-1+ct)]}_{\text{half-height wave traveling in negative direction}} + \underbrace{\frac{1}{2}[H(x+1-ct) - H(x-1-ct)]}_{\text{half-height wave traveling in positive direction}} \tag{8.146}$$

So the initial single perturbation (the boxcar function) is split into two boxcar functions of half the height, and traveling with velocity c in opposite directions.

d. Suppose we think of the area under the wave times a unit thickness as being proportional to mass. At $t = 0$ we can think of the initial wave as as having two parts moving at velocity c (for the positive half), and $-c$ (for the negative half); this is what leads to its eventual splitting for $t > 0$. We have discussed previously that this wave equation conserves energy. If we use the kinetic energy alone to represent the total, then each wave has an energy equal to Ac^2. Compute the energy in the initial condition, and show that this amount is conserved for all $t > 0$.

14. 无限介质中的泊松方程。使用傅里叶变换求解以下问题，其中 $a > 0$，$b > 0$。你的解应为例题中给出的泊松方程解的平移版本。

$$\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}=-\delta(x-a)\delta(y-b) \qquad -\infty<x<\infty,\ -\infty<x<\infty$$

$$B.C.1 \qquad u(x,y)\to 0 \ \ as\ x^2+y^2\to\infty$$

$$B.C.2a \qquad \left.\frac{\partial u}{\partial x}\right|_{(x,y)}\to 0 \ \ as\ x^2+y^2\to\infty$$

$$B.C.2b \qquad \left.\frac{\partial u}{\partial y}\right|_{(x,y)}\to 0 \ \ as\ x^2+y^2\to\infty$$

Some problems involving hypergeometric functions. The hypergeometric functions are series solutions to differential equations. There existence has been known about since at least the mid-1600s. The famous mathematicians Leonhard Euler and Carl Friedrich Gauss studied the properties of these mathematical functions in Late 1700s and early 1800s. The *generalized* hypergeometric functions appear to have been introduced by Clausen (1828) in 1828.

超几何函数有多种表示方式。这里不作详细介绍，仅列举其中几种并给出其定义。以下两个广义超几何函数与后续内容相关。

$${}_0F_1(x;a)=\sum_{n=0}^{\infty}\frac{1}{n!}\frac{\Gamma(a)}{\Gamma(a+n)}x^n \tag{8.147}$$

$${}_0F_2(x;a;b)=\sum_{n=0}^{\infty}\frac{1}{n!}\frac{\Gamma(a)}{\Gamma(a+n)}\frac{\Gamma(b)}{\Gamma(b+n)}x^n \tag{8.148}$$

Note that these are presented in a non-standard form in terms of the independent parameters a and b; this is done to make the presentation clearer. If you study generalized hypergeometric functions in the future, the notation will be slightly different from this. These functions have well-defined infinite series in general, and many computer mathematics packages (such as Mathematica) have definitions for these functions.

In addition to these, we introduce a function that is related to the hypergeometric functions called the Airy $Ai(x)$ function. This special function was discovered by British astronomer George Biddell Airy (1801–1892). It is defined in terms of the hypergeometric functions as follows.

$$Ai(x)=\frac{1}{3^{2/3}\Gamma\left(\frac{2}{3}\right)}\,{}_0F_1\left(\frac{x^3}{9},\frac{2}{3}\right)-x\frac{1}{3^{1/3}\Gamma\left(\frac{1}{3}\right)}\,{}_0F_1\left(\frac{x^3}{9},\frac{4}{3}\right) \tag{8.149}$$

There are many equivalent expressions for the Airy $Ai(x)$ function, but we this form has been chosen to help illustrate how the generalized hypergeometric functions provide the framework for many solutions to higher-order PDEs. Because the Airy function shows up in applications to physical systems (e.g., optics, quantum mechanics, waves in fluids), it has been given its own symbol, and is defined in many computer mathematics packages (such as Mathematica).

Now, the reason that these functions are important is that they allow us to describe the solutions to PDEs with higher-order derivatives in space. Toward that goal, we now note the following Fourier transform pairs.

	函数	变换
A.	$\frac{1}{\sqrt{2\pi}}$ 指数函数$(\alpha_3 ik^3 t)$	$\frac{1}{(3\alpha_3 t)^{1/2}}\mathrm{Ai}\left(\frac{x}{(3\alpha_3 t)^{1/2}}\right)$
B.	$\frac{1}{\sqrt{2\pi}}$ 指数函数$(-\alpha_4 k^4 t)$	$\frac{1}{(\alpha_4 t)^{3/4}}\left[\frac{1}{8\sqrt{\alpha_4 t}\,\Gamma\left(\frac{5}{4}\right)}{}_0F_2\left(x256\alpha_4 t\ 12\,,\frac{3}{4}\right)-x^2\Gamma\left(\frac{3}{4}\right){}_0F_2\left(x256\alpha_4 t\ 54\,,\frac{3}{2}\right)\right]$

表 8.2：由（空间）高阶偏微分方程产生的两个函数的傅里叶变换。

With this information available, it is possible for us to examine two higher-order (in space) PDEs. This is done in the following two questions.

15. **三阶偏微分方程**。如上所述，空间中的三阶导数出现在几个有趣的物理应用中，例如光学、水波的描述，或用于对晶体薄膜外延生长中的气相沉积进行建模（Lam 和 Vlachos，2001）。使用傅里叶变换求解以下带有狄拉克δ初始条件的三阶偏微分方程。

$$\frac{\partial u}{\partial t} = -\alpha_3 \frac{\partial^3 u}{\partial x^3},\ x \in (-\infty, \infty) \tag{8.150}$$

condition 1 $$u(x,t) \to 0 \text{ as } |x| \to \infty \text{ for all } x \text{ and } t \tag{8.151}$$

condition 2 $$\frac{\partial u}{\partial x} \to 0 \text{ as } |x| \to \infty \text{ for all } x \text{ and } t \tag{8.152}$$

condition 3 $$\frac{\partial^2 u}{\partial x^2} \to 0 \text{ as } |x| \to \infty \text{ for all } x \text{ and } t \tag{8.153}$$

I.C. $$u(x,0) = \beta\delta(x) \tag{8.154}$$

Note here that α_3 has the units of L^3/T (length cubed per unit time). After finding the solution, plot the solution for $\alpha_3 = 2$ cm^3/hr and $\beta = 10$. If this is done in Mathematica, note that the Airy $Ai(x)$ function is given by

$$Ai(x) = \texttt{AiryAi[x]}$$

对于绘图，使用 $x \in [-10, 10]$；对于值域，使用区间 $u \in [-10, 10]$。使用三个绘图时间，分别为 $t = 0.1$、1，和 10。

16. **Fourth-order PDE.** Fourth-order derivatives in space arise in a number of interesting problems. While they are not generally well known (even in those that study PDEs), they do have applications including (a) describing ice formation, (b) modeling liquid flows in the lungs, (c) and smoothing of raster images in computational image processing (Greer et al., 2006; You and Kaveh, 2000).
Find the solution to this problem using Fourier transforms.

$$\frac{\partial u}{\partial t} = -\alpha_4 \frac{\partial^4 u}{\partial x^4}, \quad x \in (-\infty, \infty) \tag{8.155}$$

condition 1 $$u(x,t) \to 0 \text{ as } |x| \to \infty \text{ for all } x \text{ and } t \tag{8.156}$$

condition 2 $$\frac{\partial u}{\partial x} \to 0 \text{ as } |x| \to \infty \text{ for all } x \text{ and } t \tag{8.157}$$

condition 3 $$\frac{\partial^2 u}{\partial x^2} \to 0 \text{ as } |x| \to \infty \text{ for all } x \text{ and } t \tag{8.158}$$

condition 4 $$\frac{\partial^3 u}{\partial x^3} \to 0 \text{ as } |x| \to \infty \text{ for all } x \text{ and } t \tag{8.159}$$

I.C. $$u(x,0) = \beta \delta(x) \tag{8.160}$$

Note here that α_3 has the units of L^4/T (length to the fourth per unit time).
求出解后，绘制 $\alpha_4 = 10$ 厘米^4/hr 和 $\beta = 10$ 的解图。如果使用 Mathematica 完成，请注意，超几何函数 ${}_0F_2$ 可通过以下命令调用

$${}_0F_2(x;a;b) = \texttt{ HypergeometricPFQ[\{\}, a, b, x]}$$

对于绘图，使用 $x \in [-10,10]$；对于值域，使用区间 $u \in [-2,5]$。使用三个绘图时间，分别为 $t = 0.1$、1，和 10。

17. **Fifth order PDE.** While it may or may not be a surprise at this juncture, it turns our that there are applications for fifth-order PDEs. In it simplest possible form, one can write

$$\frac{\partial u}{\partial t} = -\alpha_5 \frac{\partial^5 u}{\partial x^5}, \quad x \in (-\infty, \infty) \tag{8.161}$$

I.C. $$u(x,0) = \delta(x) \tag{8.162}$$

plus the conventional conditions on the function and its derivatives.

这类方程的研究背景是（通常为非线性的）五阶 Korteweg-De Vries（KDV）方程的线性化版本。KDV 方程最初广受关注，是因为其解包含称为孤子的行波。对这类解的研究始于19世纪中期，随后在20世纪50年代及之后又掀起了研究热潮。该方程的五阶版本已被用于描述多种现象，包括等离子体波或毛细-引力水波的运动（Bridges 等，2002）。这个问题的解形式颇为复杂，涉及含有七个参数的广义超几何函数。

$$u(x,t)=\frac{\sqrt{\frac{1}{2}\left(5+\sqrt{5}\right)}}{1171875000\,\pi\,(\alpha_5 t)^{9/5}}\left\{16500x^3\Gamma\left(-\frac{21}{5}\right)\left[x^5\,{}_0F_7\left(\{\};\frac{11}{10},\frac{6}{5},\frac{13}{10},\frac{3}{2},\frac{8}{5},\frac{17}{10},\frac{9}{5};\frac{x^{10}}{(50000\,\alpha_5 t)^2}\right)\right.\right.$$
$$\left.-8400\,(\alpha_5 t)\,{}_0F_7\left(\{\};\frac{1}{2},\frac{3}{5},\frac{7}{10},\frac{4}{5},\frac{11}{10},\frac{6}{5},\frac{13}{10};\frac{x^{10}}{(50000\,\alpha_5 t)^2}\right)\right]$$
$$+9375\left(\sqrt{5}-1\right)(\alpha_5 t)^{1/5}x^2\Gamma\left(-\frac{12}{5}\right)\left[4200(\alpha_5 t)\,{}_0F_7\left(\{\};\frac{2}{5},\frac{1}{2},\frac{3}{5},\frac{7}{10},\frac{9}{10},\frac{11}{10},\frac{6}{5};\frac{x^{10}}{(50000\,\alpha_5 t)^2}\right)\right.$$
$$\left.-x^5\,{}_0F_7\left(\{\};\frac{9}{10},\frac{11}{10},\frac{6}{5},\frac{7}{5},\frac{3}{2},\frac{8}{5},\frac{17}{10};\frac{x^{10}}{(50000\,\alpha_5 t)^2}\right)\right]$$
$$-31250\left(\sqrt{5}-1\right)(\alpha_5 t)^{2/5}x\Gamma\left(-\frac{8}{5}\right)\left[1800\,(\alpha_5 t)\,{}_0F_7\left(\{\};\frac{3}{10},\frac{2}{5},\frac{1}{2},\frac{3}{5},\frac{4}{5},\frac{9}{10},\frac{11}{10};\frac{x^{10}}{(50000\,\alpha_5 t)^2}\right)\right.$$
$$\left.-x^5\,{}_0F_7\left(\{\};\frac{4}{5},\frac{9}{10},\frac{11}{10},\frac{13}{10},\frac{7}{5},\frac{3}{2},\frac{8}{5};\frac{x^{10}}{(50000\,\alpha_5 t)^2}\right)\right]$$
$$+4418826048\,(\alpha_5 t)^{3/5}\Gamma\left(-\frac{39}{5}\right)\left[600\,(\alpha_5 t)\,{}_0F_7\left(\{\};\frac{1}{5},\frac{3}{10},\frac{2}{5},\frac{1}{2},\frac{7}{10},\frac{4}{5},\frac{9}{10};\frac{x^{10}}{(50000\,\alpha_5 t)^2}\right)\right.$$
$$\left.\left.-x^5\,{}_0F_7\left(\{\};\frac{7}{10},\frac{4}{5},\frac{9}{10},\frac{6}{5},\frac{13}{10},\frac{7}{5},\frac{3}{2};\frac{x^{10}}{(50000\,\alpha_5 t)^2}\right)\right]\right\} \tag{8.163}$$

使用 Mathematica（或你选择的符号数学程序）完成以下任务。(a) 说明该方程确实是上文给出的五阶偏微分方程的一个解。(b) 对于条件$\alpha_5 = 1\,\mathrm{m}^5/\mathrm{h}$，在 $x \in [-10,10]$ 上绘制解，其中 $u \in [-1,1]$。分别绘制时间 $t = 0.1$,1 和 10 小时对应的解。当 $t \to \infty$时，该解会发生什么变化？仅通过考察涉及时间的项（即超几何函数之外的那些项），你能否证明 当 $t \to \infty$时，该解趋近于什么？

“人追随的只是幻影。”拉普拉斯变换的发现者皮埃尔-西蒙·拉普拉斯

9 拉普拉斯变换

9.1 引言

拉普拉斯变换和傅里叶变换有许多相似之处。在一些教材中，先讨论拉普拉斯变换；在另一些教材中，则先研究傅里叶变换。由于傅里叶变换与傅里叶级数有关，而且其逆变换很容易理解，因此本书先探讨傅里叶变换。虽然拉普拉斯变换通常被认为不如傅里叶变换直观，但它仍包含许多契合问题物理特性的内容。在这里，我们将拉普拉斯变换视为傅里叶变换的一种直观伸展；这种伸展旨在扩大变换所适用的函数空间。

9.2 术语

- **拉普拉斯变换**。在区间 $[0,\infty)$ 上的一种积分变换，通常用于变换时间变量。拉普拉斯变换可以看作傅里叶变换的推广，其变换核函数增加了一个实值衰减指数函数。因此，核函数为 $\exp(-\alpha t)\exp(-i\xi t)$。通常，通过定义 $s=\alpha+i\xi$ 来表示，使得核函数为 $\exp(-st)$。在本章中，函数 f 的拉普拉斯变换定义为

$$F(s)=\int_{\tau=0}^{\tau=t} f(\tau)\exp(-s\tau)\,d\tau$$

  由于发生误解的可能性极小，在接下来的论述中，我们不再正式使用积分变量（τ），而是在积分中直接使用t。这种写法在技术上并不严谨，但与大多数教科书中的符号表示一致。

- **Complex plane.** Recall that complex numbers involve both a real and an imaginary component. Thus, $s = \alpha + i\xi$ is a complex number. Complex numbers can also be thought of as a pair, (α, ξ), where α gives the real component, and ξ the imaginary component. Such numbers can be plotted on the complex plane in the same way that one might plot a pair of real numbers on the plane. One can also define a *function* of complex numbers, $f(\alpha + i\xi) = f(\alpha, \xi)$. Such functions can be represented as a surface with height f and coordinates on the complex plane of (α, ξ).

- **拉普拉斯变换中的卷积**。卷积已经被多次讨论，并且在研究拉普拉斯变换时再次出现。由于拉普拉斯变换定义在半轴上，卷积同样受到限制。因此，对于两个函数 f 和 g，半轴卷积可表示为

$$(f * g)(t) = \int_0^\infty f(t-\tau)g(\tau)\,d\tau = \int_0^\infty f(t)g(t-\tau)\,d\tau$$

- **围道积分**。我们已经定义了复平面，以及可能定义在复平面上的函数。微积分中的线积分 有一种模拟物，它将积分推广到了复平面。虽然这类积分（及相关定理）并不难，但确实需要对复分析进行大量学习。本主题超出了本文预期的讨论范围。不过，尽管我们不描述如何计算围道积分的过程，定义围道积分仍然是有意义的。拉普拉斯逆变换通常被理解为一个围道积分过程。

9.3 拉普拉斯变换

拉普拉斯变换的一个动机是扩大能够通过该方法进行分析的函数类型。例如，回想一下，对于傅里叶变换，函数至少必须在$x \in (-\infty, \infty)$上具有有限的平方积分。因此，例如，函数$f(x) = x$这个看似简单的函数，却不存在傅里叶变换。

One way of thinking about the Laplace transform is that it extends the kinds of functions that can be transformed by including a conditioning function within the transform itself. The role of this conditioning function is to eradicate "fast growing" components of a function that prevent it from having a defined square integral. Transforms, like the Fourier or Laplace transforms, are qualified as *functionals* that map entire *functions* from one space to another. Recall, the Fourier transform of the function $f(x)$, assuming that it exists, would be given in the complex exponential form. Laplace transforms are often used to transform functions that are dependent on time. Thus, by convention we will adopt t as the independent variable for this chapter. For easy correspondence with the Laplace transform, we will write out the Fourier transform in the independent variable t. The result is as follows.

$$\hat{f}(\xi) = \mathscr{F}[f(t)] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) e^{-i\xi t}\,dt \tag{9.1}$$

In the case of Fourier transforms, the functional *uniquely* maps functions with independent variable t (here, noting our change in notation for this chapter only) to a new function, F, entirely in terms of the variable ξ, i.e.,

$$\mathscr{F}[f(t)] \to \hat{f}(\xi) \tag{9.2}$$

As a matter of vocabulary, the function $e^{-i\xi t}$ is called the *kernel* of the Fourier transform.

现在，这种特定的变换本身并没有什么神圣不可侵犯之处（除了它悠久的历史和相对直观的物理解释之外！）。事实上，如果我们稍微修改其定义，将其改为另一种定义良好的泛函（通过改变核函数），这种变换仍然可能具有价值。实际上，存在大量的积分变换，每一种都采用不同的核函数，以突出变换函数的某些特定特征。例如，在维基百科上可以查到，截至目前，其中定义的积分变换种类已接近 20 种。因此，只要能够证明所提出的核函数具有某种特定价值（例如，能够求解原本无法求解的问题），探索用于实施变换的不同核函数便在应用数学分析中占有一席之地。

在我们的讨论中，让我们考虑一种最简单的傅里叶变换改造方式，它有助于扩大可进行变换的函数空间。作为一个具体例子，我们考虑函数$f(t) = x$，其中 $-\infty < t < \infty$。显然，这个函数的傅里叶变换没有明确定义。傅里叶变换将给出

$$\mathscr{F}[f(t)] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} t \cdot e^{-i\xi t}\, dt \tag{9.3}$$

由于$e^{-i\xi t}$是一个振荡函数，因此当 $|t| \to \infty$时，这个积分无界。不过，假设我们决定不用$e^{-i\xi t}$，而改用核函数$e^{-\alpha t}e^{-i\xi t}$，其中 $\alpha > 0$ 是某个实数。这样的加法会对核函数的行为产生什么影响呢？首先，我们知道（至少对于正的t 值），$e^{-\alpha t}$ 是一个相对于例如任意 多项式都快速衰减的函数。在这里，快速 的意思是，对于t的任意多项式$P_n(t)$（其中n 是该多项式的次数或阶），我们总有

$$\lim_{t\to\infty} P_n(t)e^{-\alpha t} = 0 \tag{9.4}$$

对于任意 $\alpha > 0$。（要看出这一点，只需将指数函数展开为其幂级数；该级数处处收敛，且指数函数包含所有阶的多项式项。）这一加法带来的实际结果是，原先不可积的函数现在借助这个新核函数变得可积了。不过，还有一个细节需要处理。让我们再次回到傅里叶变换的定义，但采用这一修改。现在，我们把修改后的傅里叶变换称为 $\mathscr{F}'$ 变换。对于$f(t) = t$，结果为

$$\mathscr{F}'[f(t)] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} t \cdot e^{-\alpha t}e^{-i\xi t}\, dt \tag{9.5}$$

The exponential $e^{-\alpha t}$ accomplishes our goal for $t \geq 0$, but actually makes the problem worse for $t < 0$. In other words, for $t < 0$, the sign of the exponential is *positive*, and the exponential grows without bounds. There are some functions that still converge for this transform on the bounds $-\infty < t < \infty$; in this application, the transformation is known as the *bilateral* or *two-sided* Laplace transform. This is a somewhat specialized application that will not be further discussed in this text. Instead of insisting on convergence over $-\infty < t < \infty$, we will adopt the convention that our $\mathscr{F}'$ transform is used only for functions defined on $t \in [0, \infty)$. Under these circumstances the increasing exponential is no longer a problem. Restricting the domain of integration, we can now compute the integral

$$\mathscr{F}'[t] = \frac{1}{\sqrt{2Pi}} \int_{0}^{\infty} t \cdot e^{-\alpha t}e^{-i\xi t}\, dt \tag{9.6}$$

this integral, although an improper one, can be computed.

$$\mathscr{F}'[t] = \frac{1}{\sqrt{2\pi}} \int_0^\infty t \cdot e^{-\alpha t} e^{-i\xi t}\, dt = \frac{1}{\sqrt{2\pi}} \frac{1}{(a+i\xi)^2} \tag{9.7}$$

因此，我们这种新变换似乎有一个优点：对于诸如$f(t) = t$这样的函数，它能得到收敛（因而可计算）的积分，而这类函数此前无法进行变换。

这种新变换称为拉普拉斯变换。不过，在完成定义之前，我们必须指出拉普拉斯变换中另外两项约定俗成的改动。首先，正变换中不包含常数$1/\sqrt{2\pi}$；而是将它与逆变换中的同一常数合并。这不会改变变换对，只会改变常数 $1/(2\pi)$ 在正变换与逆变换之间的分配方式。其次，按照惯例，我们令 $s = \alpha + i\xi$。这两项改动构成了拉普拉斯变换的标准形式，记作 $\mathscr{L}$。

$$\mathscr{L}[f(t)] = \frac{1}{\sqrt{2\pi}} \int_0^\infty f(t) \cdot e^{-st}\, dt \tag{9.8}$$

现在请注意，对函数$f(t) = t$先前的分析作出这些改动后，我们得到

$$\mathscr{L}[f(t)] = \int_0^\infty f(t) \cdot e^{-st}\, dt = \frac{1}{s^2} \tag{9.9}$$

因此，这是一个相当重要的发现！使用傅里叶变换时，我们只能处理当 $|t|$（或者用上一章的符号表示，当 $|x|$）变大时迅速衰减的函数，才能确保傅里叶积分保持有限。现在，只要函数的增长速度低于指数函数，我们就大幅拓宽了可进行变换的函数范围（甚至包括指数级增长的函数，只要相应积分收敛；稍后还会进一步讨论）。为此我们付出了一点代价：现在我们的变换只适用于实数轴的正半轴。不过，考虑到它与初值问题的对应关系——实数轴的正半轴是唯一对应物理现实的部分——这表明拉普拉斯变换可能有助于变换时间域中的问题。

现在我们已经对拉普拉斯变换有了一个明确的定义。

定义 9.1（拉普拉斯变换）。拉普拉斯变换，$\mathscr{F}[f(t)] = F(s)$，是一种积分变换，主要用于对半实数轴上的函数进行变换，$t \in [0,\infty)$。其定义为

$$\boxed{\mathscr{L}[f(t)] = F(s) = \int_0^\infty f(t) \cdot e^{-st}\, dt}$$

where s is a complex number. We restrict our analysis to functions, $f(t)$, where the following must be true

1. f is piecewise continuous on $0 \le t \le T$, for all $T > 0$. This means only that f has a finite number of points of discontinuity.
2. We must have
$$\lim_{t\to\infty} f(t)e^{-st} = 0 \tag{9.10}$$
so that the transform is given by a finite integral.

3. Sometimes, a more practical version of this last statement is given by the following: There are *real numbers* M and c such that

$$|f(t)| \leq Me^{ct} \text{ for all values of } t \tag{9.11}$$

因此，当$Re(s) > c$时，拉普拉斯变换存在，也就是说，s 的实数集 部分（$s = \alpha + i\xi$中的实数α）大于c。因此，这意味着拉普拉斯变换中定义的衰减指数函数比增长指数函数e^{ct}衰减得更快；结果是一个有限积分。

例 *9.1*（拉普拉斯变换的计算。）

拉普拉斯变换的强大之处，首先可以从我们研究那些无法用傅里叶变换分析的函数时看出来。例如，考虑下面这个当 $t \to \infty$ 时趋于无穷的函数。

$$f(t) = at^5 \tag{9.12}$$

显然，随着t 变大，这个函数会迅速增长。不过，拉普拉斯变换仍然有定义，因为这个多项式的增长速度比我们在变换中加入的衰减指数函数慢。因此，结果是有限的。变换的形式为

$$\mathscr{L}[at] = \int_0^\infty at^5 e^{-st}\, dt \tag{9.13}$$

这个问题可以通过直接（但繁琐地）使用五次分部积分来求解，每次使用分部积分都会使t^5的指数阶降低一阶。此处不列出具体步骤；然而，结果是

$$\mathscr{F}[at^5] = a\left(\int_0^\infty t^5 e^{-st}\, dt\right) \tag{9.14}$$

$$\mathscr{F}[at^5] = \frac{120a}{s^6} \tag{9.15}$$

9.4 Some Notes About Laplace Transformed Functions

Unlike the Fourier transform, it is a bit harder to relate the *shape* of a Laplace transformed function to the properties of the original function. For example, the Fourier transform of a pure sine function is a delta function shifted to the appropriate wave number; intuitively, this makes sense, because the sine funciton contains a single wavelength, and this is represented in its transform. For a Laplace transform, however, the transform of a sine function is a hyperbolic function in the transform space – not a result that immediately appeals to an easy physical interpretation.

A second issue is that the transformed functions are a function of $s = \alpha + i\xi$– a complex number. Recall from Chp. 2 that a complex variable can be though of as a two variables, one representing the real-line component, and one representing the imaginary component. While it this is also true for the Fourier transform,

the transform variable for that case is given by $0+i\xi$ Thus, a Fourier-transformed function can be plotted as a function of only ξ. To properly interpret a Laplace-transformed function, however, we must plot the function as a complex number $F(s)$ $s=\alpha+i\xi=(\alpha,\xi)$; that is, we must plot it as if it were a function in 2-dimensions!

It can be useful to examine the shapes of a few functions and their transforms, which is done in the following example.

例 *9.2*（函数及其变换的几个实例。）

Recall from Chapter 2, a complex number $s=\alpha+i\xi$ can be represented in polar coordinates by specifying its magnitude, r (or absolute value), and its argument, θ. These are defined by

$$r=\sqrt{\alpha^2+\xi^2}$$
$$Arg(s)=\theta=\tan^{-1}\left(\frac{\xi}{\alpha}\right)$$

One complicating feature of the Laplace transform is that the *result*, $\mathscr{L}[f(t)]$ of the transform is the complex function F(s). Recall that a function of a complex number can be thought of as being a function in a two-dimensional plane, where the vertical axis represents the *imaginary* component of s, and the horizontal axis represents the *real* part of s. Further complicating things is that the output of the function is, in general, itself a complex number. Thus, in some very reasonable ways, the Laplace transform is a vector function, that returns a vector as its output, i.e.,

$$F(s)=F_r(\alpha+i\xi)+iF_I(\alpha+i\xi)$$

where here F_r represents the real component of $F(s)$ ($F_r=Re[F(s)]$), and F_I represents the imaginary part of $F(s)$ ($F_I=Im[F(s)]$). This complicates the matter because we now have a result with four dimensions (α,ξ,F_r,F_I)! There are a number of methods for representing such functions, however. Here, we will map the functions as follows. First, the horizontal and vertical components will represent the real and imaginary components of s as is customary. Second, the magnitude (or modulus) of $F(s)$ will then be plotted on the vertical axis. Finally, the argument (θ) will be overlaid on the surface as a color field to represent the fourth dimension of our plot. While this is a lot of information to take in, the plots of our transformed functions are still interesting to examine. A few examples appear below.

1. The function $f(t)=1$ for $t>0$ is has the transform $\mathscr{L}(1)=\frac{1}{s}$. Note that at $s=0+i0$, the transform function has singularity ($\frac{1}{s}$ tends toward infinity there) which in complex analysis is known as a *pole*. This imposes no particular problems in complex analysis. There is an analogue to the Stokes theorem (called the *residual theorem* in the complex plane that allows an integral over a complex domain to be replaced by a closed line integral. Converting to the complex coordinates $s=\alpha+i\xi$ (as we defined above)

we can plot the resulting transform in the complex plane. This plot is given in Fig. 9.1.

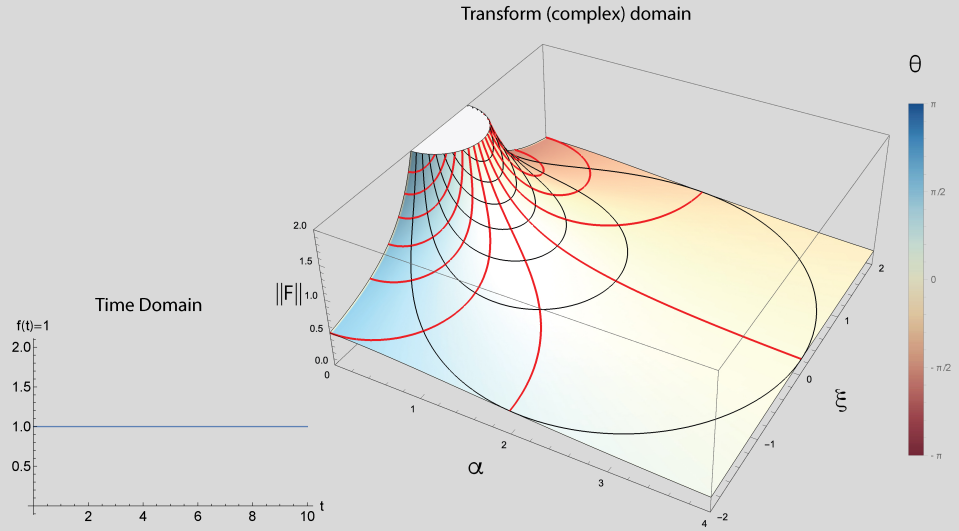


图 9.1：$f(t)$ 的拉普拉斯变换 $=1$。垂直分量表示函数的模（r），颜色表示辐角 θ。黑色和红色直线分别表示s的虚部（黑色）和实部（红色）为常数的直线。

2. 函数 $f(t)$ $=$正弦(t)，当 $t>0$时，其变换为 $\mathscr{L}[$正弦$(t)] = \frac{1}{1+秒^2}$。这里，当 s趋近于 $s=1+0i$或 $s=-1+0i$时，存在奇点。变换在复坐标 $s=\alpha+i\xi$ 中的图像见图 9.2。

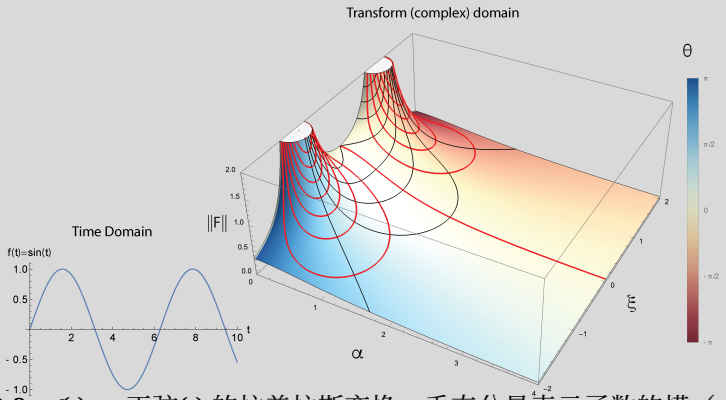


图 9.2：$f(t)$ $=$正弦(t) 的拉普拉斯变换。垂直分量表示函数的模（r），颜色表示辐角，θ。黑色和红色直线分别表示s的虚部（黑色）和实部（红色）的等值线。

3. The function $f(t)=e^{2t}$ for $t>0$, with transform $\mathscr{L}[e^{2t}]=\frac{1}{s-2}$. Note that for this function, the transform $\mathscr{L}[e^{2t}]$ is not defined for $Re(s)\leq 2$. Recall, this happens because the exponential e^{-st}

must decay faster than the exponential e^{2t} grows; thus we must have $Re(s) > 2$. This result means that there is a discontinuity in the function that spans the entire complex plane along the line $s = 2 + i\xi$ for all values of ξ. As mentioned above, this does not present significant problems in the inversion integral. In the complex plane, one can use the residual theorem (analogous to the Stokes theorem of vector calculus) to compute the inverse.

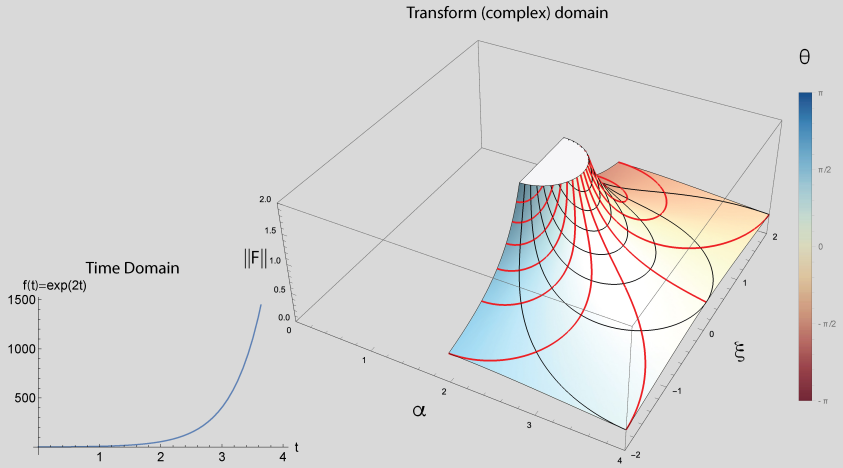


Fig. 9.3: Laplace transform of $f(t) = e^{2t}$, with transform $\mathscr{L}[e^{2t}] = \frac{1}{s-2}$. The vertical component gives the magnitude (r) of the function, and the color indicates the argument, θ. The black and red lines indicate lines of constant imaginary (black) and real (red) components of s

9.5 The Inverse Laplace Transform

拉普拉斯变换积分的定义中有一个细微之处，对逆变换有着重要影响。类比傅里叶变换，我们会认为逆拉普拉斯变换需要对变换函数关于变换变量s进行积分。乍看之下，这既熟悉（与傅里叶变换相比较），也合乎情理。然而，请记住，现在s是一个复变量。根据你先前学习复变函数分析的程度，你可能知道，复变量函数可以看作是关于两个自变量的函数——实变量和虚变量（第2章曾简要介绍过）。因此，对复变量进行积分可以看作是在复平面上积分（按照惯例，纵轴称为虚轴，横轴称为实轴。）

9.5.1 *布罗姆维奇积分

拉普拉斯逆变换， $\mathscr{L}^{-1}[F(s)]$ ，形式上可以用所谓的布罗维奇积分表示。虽然我们不再进一步探讨这一形式（它需要理解复变函数分析），但从概念上理解这个积分很有帮助，因为它对应于傅里叶变换的逆变换。用复布罗姆维奇积分表示的逆变换为

$$\mathscr{L}^{-1}[F(s)] = \lim_{\xi\to\infty} \frac{1}{2\pi i} \int_{s=b-i\xi}^{s=b+i\xi} F(s)e^{st}\,ds \tag{9.16}$$

这里添加“lim”是为了强调，在积分步骤中，只有变量的虚部是在无穷处取值的。还应注意，归一化常数与傅里叶变换中出现的常数类似，只出现在逆变换中。显然，最终结果是相同的，因此，这种选择仅仅是为了方便定义变换（同时使逆变换的定义略微失去对称性）。

9.5.2 Post's Inversion Theorem

The Bromwich integral allows one to compute the *inverse* Laplace transform, but it requires that one also have expertise in complex analysis, which is not a common area of study in more modern instruction in applied mathematics.

However, there are alternatives. In 1930, a mathematician named Emil Post (Post, 1930) presented the following Laplace inversion formula as a the limit of a sequence. Assume that we have the Laplace transform pair $f(t) \rightleftharpoons F(s)$.

Theorem 9.1 (Post's inversion theorem). *Suppose we have the Laplace transform pair $f(t) \rightleftharpoons F(s)$. Then, the inversion of $F(s)$ can be determined by computing the following limit*

$$f(t) = \lim_{k\to\infty} \frac{(-1)^k}{k!} \left(\frac{k}{t}\right)^{k+1} F^{(k)}\left(\frac{k}{t}\right) \tag{9.17}$$

where $F^{(k)}$ is the kth derivative of $F(s)$ with respect to s [N.B., the independent variable is changed after the differentiation process so that $F^{(k)}(s) \to F^{(k)}(k/t)$].

It is somewhat frequently stated that this inversion formula is not of practical value, but this statement should be viewed dubiously. There are several reasons to view this inversion formula as a useful tool, including the following.

1. Post逆变换公式完全用实变量表示，因此不涉及复分析 2. 目前已有符号软件包，可用于帮助求出导数的递推关系，并计算 Post反演定理中所指定的此类函数各项的极限。

例 *9.3*（*Post* 反演定理的应用）。

In the example above, we saw that $\mathscr{L}[at] = \frac{a}{s^2}$. Show that the inverse transform of $F(s) = 1/s^2$ is in fact, $f(t) = t$ using Post's inversion theorem.

To start, we need the derivatives of $F(s)$ with respect to s. A little work allows us to find the general form for this derivative.

$$\begin{aligned}
F^0(s) &= \frac{1}{s^2} \\
F^1(s) &= -2\frac{1}{s^3} \\
F^2(s) &= 6\frac{1}{s^4} \\
F^3(s) &= -24\frac{1}{5} \\
F^4(s) &= 120\frac{1}{s^6} \\
&\cdots \\
F^{(k)}(s) &= (-1)^k(k+1)!\frac{1}{s^{k+2}}
\end{aligned}$$

Now, making the change $s \to k/t$, and substituting into Post's inversion formula, we need to evaluate

$$f(t) = \lim_{k\to\infty} \frac{(-1)^k}{k!}\left(\frac{k}{t}\right)^{k+1}(k+1)!(-1)^k\frac{1}{\left(\frac{k}{t}\right)^{k+2}}$$

Noting that $(-1)^k(-1)^k = 1$ and that $(k+1)!/k! = (k+1)$, we have

$$f(t) = \lim_{k\to\infty} \frac{(k+1)!}{k!}\frac{k^{k+1}}{t^{k+1}}\frac{t^{k+2}}{k^{k+2}}$$

While this looks somewhat messy, we can simplify a little bit by making the substitution $m = k+1$. Then we have

$$f(t) = \lim_{k\to\infty} \frac{m!}{(m-1)!}\frac{(m+1)^m}{t^m}\frac{t^{m+1}}{(m+1)^{m+1}}$$

noting that $m!/(m-1)! = m$, then this reduces to

$$\begin{aligned}
f(t) &= \lim_{m\to\infty} \frac{m}{(m+1)}t \\
&= t
\end{aligned}$$

Fortunately, most of the Laplace transforms and their inverses have been computed and tabulated previously. In fact, this is the very thing that makes the transform methods so useful – much of the effort for

computing the transforms and inverse transforms is removed from the user because most of the "interesting" functions have been computed previously. In the appendix to this chapter, there is a somewhat lengthy table of transforms and inverse transforms. Regardless of the presence of such tables, however, it is important to understand the transforms themselves. This means understanding what the transform is doing when it is computed, why it works, and how one might compute its inverse. The material above provides one essentially all that one might need in this respect, and, in particular, the Post inversion theorem allows one to compute the inverse Laplace transform without having to have previously studied the theory of complex variable analysis.

9.6 Heaviside 阶跃函数与狄拉克δ函数的拉普拉斯变换

拉普拉斯变换比傅里叶变换应用得更广泛，部分原因是它们可用于信号处理。回想一下，拉普拉斯变换用于半轴 $t \in [0,\infty)$，因此变换中自然包含了零时刻。这与许多从初始条件开始演变的过程相对应。例如，拉普拉斯变换常用于预测含有电阻器、电感器和电容器的电路随时间的行为。鉴于这些是经典（晶体管出现前）电路的构成元件，这类问题的解是一项重要应用。另一个例子是电网分析。假设有人想了解，如果网格上突然出现一个"尖峰"负荷（类似狄拉克δ函数的负荷），或者在$t=a$ 时刻突然开始持续增加负荷（Heaviside负荷），电网上可能会发生什么。这些都对应着真实的物理现象，电网基础设施的规划和运营人员会关注这些现象。同样，拉普拉斯变换也可用于分析这些问题。

It turns our that the Heaviside and delta functions are used routinely in modeling many kinds of dynamic systems because of their obvious interpretations as spike or sudden constant sources or sinks. Unlike the Fourier transform, there is no real difficulty in determining their transforms (once one has adopted the formalisms associated with the delta function). The transform of the Heaviside function is given by the following

$$\begin{aligned}\mathscr{L}[H(t-a)] &= \int_0^\infty H(t-a)\exp(-ts)\,dt \\ &= \int_a^\infty 1\times\exp(-ts)\,dt \\ &= -\frac{1}{s}\exp(-ts)\Big|_a^\infty \\ &= \frac{1}{s}\exp(-as)\end{aligned} \tag{9.18}$$

And for the delta function

$$\begin{aligned}\mathscr{L}[\delta(t-a)] &= \int_0^\infty \delta(t-a)\exp(-ts)\,dt \\ &= \exp(-as)\end{aligned} \tag{9.19}$$

关于在不连续点处应用狄拉克δ函数的警示说明

It is necessary to make an important note about the Laplace transform of the delta function. The casual use of generalized functions without carefully understanding their deeper theory can sometimes lead to trouble. One problem that has arisen classically with the Laplace transform is the interpretation of a delta function placed at the origin. Why might this create a problem? To start, lets consider the following

$$\int_0^\infty \delta(t)\exp(-st)\,dt \tag{9.20}$$

现在，这个积分看起来似乎相当简单。不过，请回想一下，我们将狄拉克δ函数定义为一个由函数构成的序列，以n为索引，在某种意义下，当$n \to \infty$时，该序列“收敛”到狄拉克δ函数。我们提出了若干函数，而这些函数通常关于原点对称（例如，我们在狄拉克δ函数一章中使用的一个例子涉及余弦函数的幂，即$(cos[\pi t])^n$，它关于$t = 0$对称）。假设我们将狄拉克δ函数视为一个δ序列 $\delta_n(t)$ 的极限。进一步考虑这个问题，我们可以将上述积分写成如下形式。

$$\int_0^\infty \delta(t)\exp(-st)\,dt = \int_{-\infty}^\infty \left[\lim_{n\to\infty} \delta_n(t)H(t)\exp(-st)\right] dt \tag{9.21}$$

如果我们的狄拉克δ函数确实是一个关于零对称的函数序列，我们可能会预期这样的级数收敛后得到的积分值为 1/2。为什么会这样呢？对于 δ序列中的每一个成员（对所有 $n > 0$），恰好有一半的 δ序列位于（$-\infty$,0）上，另一半则位于（0,∞）上。当只取正半部分时（这由 Heaviside截断函数保证），结果应当是 1/2，而不是更常见的值 1。

事实上，这个问题多少是无法解决的，至少在传统分布理论的框架下是如此。根据所选的δ序列，式 (9.21) 给出的积分可以收敛到 0 到 1 之间的任意数。换句话说，狄拉克δ函数与海维赛德函数的乘法没有唯一结果，因此应当避免这种运算（至少我们应当避免！）。我们不会试图修补如何定义作用于 $t = 0$ 处的狄拉克δ函数这一概念，而是直接指出，在我们的拉普拉斯变换框架中，这不是一种允许的运算。凡是狄拉克δ函数与海维赛德函数可能重叠的地方，都适用这一结论——例如， $\delta(t-a)H(t-a)$ 的拉普拉斯变换同样没有良好定义。因此，我们可以使用狄拉克δ函数，但不能在定义域中已有不连续点的地方使用它们！这实际上不会给我们带来太多麻烦，因为物理上可能出现这种情况的场合很少（不过，信号处理理论以及物理学应用中都曾指出这一问题）。虽然分布理论的某些伸展可以对狄拉克δ函数与海维赛德函数的这类乘法作出合理解释，但对我们的目的而言，最好的做法就是指出，在我们的框架中，这些量没有定义。

9.7 不连续函数的拉普拉斯变换

海维赛德函数 赫维赛德阶跃函数(秒 $-a$) 是一个只有一个不连续点的函数，该不连续点位于 秒 $-a$。上面，我们能够计算这个不连续函数的拉普拉斯变换，这表明具有不连续点的更一般函数可能也可以进行变换。事实上，结果证明确实如此。为了开始思考这个问题，我们来看看下面的例子。考虑以下两个函数f和 g，它们定义为

$$f(t) = \cos(\pi t) \tag{9.22}$$
$$g(t) = \exp(-t) \tag{9.23}$$

我们可以用多种方法构造一个不连续函数，不妨尝试下面这种方法。在0 < 秒 < a 之间，函数由 f (t) 定义；在时间 t = a 之后，函数由 f (t) 与 g(t − a) 之和定义。请注意，后一个函数在与 f (t) 相加之前先进行了平移。假设我们令这个平移后的时间为 t′ = t − a。因此，函数 g(t′) 与 g(t) 完全相同，只是在时间上向前平移了 a（见图 9.4）。请注意，这种平移在时间上向前，是因为在 t = a 时，t′ = a − a = 0；因此，g(t = 0) 的值在新的坐标系中映射到时间 g(t′ = 0)，或者等价地，在原来的时间坐标中映射到 g(t − a)。由此，我们得到下面这个不连续函数，称为 w(t)，其定义为

$$w(t) = \begin{cases} \cos(\pi t) & t < a \\ \cos(\pi t) + \exp[-(t-a)] & t \geq a \end{cases} \tag{9.24}$$

这种形式对我们的推导来说并不是最有用的。我们可以利用海维赛德函数，将其改写为一个等价的单一方程，如下所示。

$$w(t) = \cos(\pi t) + \exp[-(t-a)]H(t-a), \quad \text{for all } t > 0 \tag{9.25}$$

需要非常明确的是，这类问题通常具有良好的定义。然而，请注意，根据上一节中的警告，我们不能考虑函数 $g(t) = \delta(t)$，因为这样一来，在不连续点处就会出现狄拉克δ函数与赫维赛德阶跃函数相乘的情况（而这是我们不允许的）。

此时，我们可以通过对表达式两边应用拉普拉斯变换，按形式化步骤继续推导。

$$W(s) = \mathscr{L}[\cos(\pi t)] + \mathscr{L}[\exp[-(t-a)]H(t-a)], \quad \text{for all } t > 0 \tag{9.26}$$

其中 $W(s)$ 是 $w(t)$ 的拉普拉斯变换。这两个函数中的第一个就是 $\cos(\pi t)$ 的拉普拉斯变换。可以利用欧拉公式将其转换为复指数（即 $\cos(\pi t) = \frac{1}{2}[\exp(i\pi t) - \exp(-i\pi t)]$），然后进行积分来计算该变换。不过，与傅里叶变换的情况很相似，我们通常不必亲自计算常见函数的变换，因为这些变换都是众所周知的。本章末尾的表格列出了大量此类变换；特别是，它给出了我们所需的变换。

$$\mathscr{L}[\cos(\pi t)] = \frac{s}{\pi^2 + s^2} \tag{9.27}$$

那么，对于 $\mathscr{L}[\exp[-(t-a)]$赫维赛德阶跃函数$(t-a)]$ 呢？嗯，严格来说，这个并不在表格所列的范围内，但它同样是一个指数函数，因此积分起来也并不困难。我们有

$$\begin{aligned} \mathscr{L}[\exp[-(t-a)]H(t-a)] &= \int_{t=0}^{\infty} \exp[-(t-a)]H(t-a)\exp(-ts)\,dt \\ &= \int_{t=a}^{\infty} \exp[-(t-a)]\exp(-ts)\,dt \end{aligned} \tag{9.28}$$

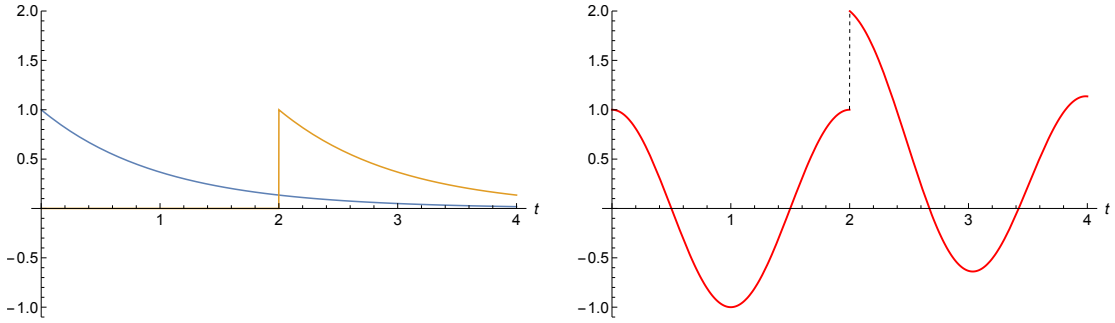


Fig. 9.4: (Left) The exponential function $\exp(-t)$ and the shifted and cutoff exponential function $\exp[-(t-2)]H(t-2)$. (Right) The function $w(t) = \cos(\pi t) + \exp[-(t-2)]H(t-2)$. The dashed line indicates the jump.

虽然可以直接计算这个积分，但为了适用于更广泛的应用，我们将按以下方式处理这个问题。首先，令 $\tau = t - a$，使得 $dt = d\tau$，且 $t = \tau + a$。现在注意，该积分可以转化为如下形式。

$$\begin{aligned}\mathscr{L}[\exp[-(t-a)]H(t-a)] &= \int_{\tau=0}^{\infty} \exp[-(\tau)]\exp(-(\tau+a)s)\,d\tau \\ &= \int_{\tau=0}^{\infty} \exp[-(\tau)]\exp(-\tau s)\exp(-as)\,d\tau \\ &= \exp(-as)\int_{\tau=0}^{\infty} \exp[-\tau]\exp(-\tau s)\,d\tau \end{aligned} \tag{9.29}$$

最后这个积分就是负指数项的拉普拉斯变换再乘以 exp(–作为)。请注意，我们用于积分变量的名称并不会影响函数的变换。查阅变换表，我们得到的是

$$\exp(-as)\mathscr{L}[\exp(-\tau)] = \exp(-as)\frac{1}{s+1} \tag{9.30}$$

这里我们只需注意，在积分中，τ 所起的作用与 t 完全相同（事实上，我们可以直接重新命名自变量 t，以免混淆）。现在可将最终结果表述如下。

$$\begin{aligned} W(s) &= \mathscr{L}[\cos(\pi t)] + \mathscr{L}[\exp[-(t-a)]H(t-a)] \\ &= \frac{s}{\pi^2 + s^2} + \exp(-as)\frac{1}{s+1} \end{aligned} \tag{9.31}$$

This result can be generalized as follows.

Theorem 9.2 (Laplace transform of a shifted and cut off function). *Suppose we have a function g, shifted forward in time by the amount a, whose value is identically zero for $t < a$ (i.e., cut off for $t < a$). This function can be represented by $w(t) = g(t-a)H(t-a)$. The Laplace transform of $w(t)$ is given by*

$$W(s) = \exp(-as)G(s)$$

*Wh*其中 $G(s) = \mathscr{L}[g(t)]$；也就是说，$G(s)$ 是未经过移位的函数的拉普拉斯变换 *g.*

理解这一结果的一种有用方式是：移位并截断后的函数$g(t-a)H(t-a)$，在拉普拉斯域中表示为$g(t)$的拉普拉斯变换乘以一个按指数衰减的权重 $\exp(-as)$。

例 *9.4*（跃变的拉普拉斯变换。）假设有一个电力系统，其负荷 $\ell(t)$ 等于常数值 ℓ_1 （单位适当）。在时间$t = a$ a 时，新增一项负荷，使总负荷达到 ℓ_2 单位。描述这一情况的函数可写如下

$$\ell(t) = \begin{cases} \ell_1(t) & t < a \\ \ell_1(t) + \ell_2(t-a) & t \geq a \end{cases} \tag{9.32}$$

求该函数的拉普拉斯变换。

解。我们可以利用 Heaviside 截断滤波器的形式化表示，将这个分段函数转换为相应形式。结果为

$$\ell(t) = \ell_1(t) + \ell_2(t-a)H(t-a)$$

The Laplace transform is given by

$$\begin{aligned} L(t) &= \mathscr{L}[\ell_1(t)] + \mathscr{L}[\ell_2(t-a)H(t-a)] \\ &= L_1(t) + \exp(-as)L_2(t) \end{aligned}$$

Note- even though we do not have explicit formulas for $\ell_1(t)$ and $\ell_2(t)$, we were still able to determine the Laplace transform symbolically. At this point, we could be given *any* admissible functions (i.e., that have a defined Laplace transform), and the result could be found by simply substituting the appropriate transformed functions.

9.8 Laplace Transform of the Derivative

One of the most familiar applications of the Laplace transform is for solving first- and second-order ODEs on the real half-line (i.e., on $t \in [0, \infty)$). However, they are also *very useful* for solving PDEs. In both cases, the real power of the Laplace transform arises because of what it does to derivatives. Much like the Fourier transform, the Laplace transform changes derivatives in real space into algebraic quantities in transformed space. However, unlike the Fourier transform, the Laplace transform comes with a tiny bit of overhead. Because the integration defining the transform has a lower bound of zero, the integration leads to capturing

the initial conditions of the problem when the integration is completed. Thus, we have the following identities for the Laplace transform of derivatives.

定理 9.3（导数的拉普拉斯变换）。对于任意函数 $f(t)$，若 $f(t)$、$f'(t)$ 以及 $f''(t)$ 连续，则以下变换成立。

$$\mathscr{L}\left[\frac{df(t)}{dt}\right] = \mathscr{L}[f'(t)] = sF(s) - f(0) \tag{9.33}$$

$$\mathscr{L}\left[\frac{d^2f(t)}{dt^2}\right] = \mathscr{L}[f''(t)] = s^2F(s) - sf(0) - f'(0) \tag{9.34}$$

其中，$f(0)$ 和 $f'(0)$ 表示f的已知辅助条件。该结论可通过简单应用分部积分来证明，这里留作练习。

这里有几点需要说明。首先，如果需要，可以类似地定义高阶导数（例如，参见 Spiegel（1965））。其次，要使n 阶导数的拉普拉斯变换存在，我们需要函数f可导n 次，并且所得导数是连续函数。不过，th阶导数只需分段连续（即n^{th} 阶导数函数可能存在不可微点）。最后，如有需要，甚至可以适当放宽可导这一条件。以下定理定义了不连续函数的拉普拉斯变换。

Theorem 9.4. *Suppose $w(t)$ is not continuous, but instead it contains a jump discontinuity at $t = a$, and is represented by $w(t) = f(t) + g(t-a)H(t-a)$. Then, the expression for the derivative is modified as follows*

$$\mathscr{L}\left[\frac{dw(t)}{dt}\right] = \mathscr{L}[w'(t)] = sF(s) - w(0) + s\exp(-sa)G(s). \tag{9.35}$$

Note that for the case where $a = 0$, this expression holds with $w(t) = f(t) + g(t)$ by definition.

证明。
By definition, we have

$$\mathscr{L}\left[\frac{dw(t)}{dt}\right] = \mathscr{L}\left[\frac{df(t)}{dt}\right] + \mathscr{L}\left[\frac{d}{dt}\left(g(t-a)H(t-a)\right)\right] \tag{9.36}$$

Here, $f(t)$ is a regular, continuous function ($w(t)$ is discontinuous). Thus, $\mathscr{L}[f] = sF(s) - f(0)$. Now the trick is to determine the second of the two transforms on the right-hand side; to be explicit, this is given by

$$\mathscr{L}\left[\frac{\partial}{\partial t}\left(g(t-a)H(t-a)\right)\right] = \int_0^\infty \frac{d}{dt}\left(g(t-a)H(t-a)\right)\exp(-st)\,dt \tag{9.37}$$

这里，我们可以使用分部积分法。假设我们选择.....进行分部积分。

$$u = \exp(-st) \qquad dv = \frac{d}{dt}[g(t-a)H(t-a)]$$
$$du = -s\exp(-st) \qquad v = g(t-a)H(t-a)$$

Integration by parts gives us

$$\begin{aligned}\mathscr{L}\left[\frac{\partial}{\partial t}(g(t-a)H(t-a))\right] &= \exp(-st)g(t-a)H(t-a)|_0^\infty - (-s)\int_0^\infty g(t-a)H(t-a)\exp(-st)\,dt \\ &= 0+s\int_0^\infty g(t-a)H(t-a)\exp(-st)\,dt \\ &= s\exp(-sa)\int_0^\infty g(t)H(t)\exp(-st)\,dt \quad \text{(using the shifting property)} \\ &= s\exp(-sa)\int_0^\infty g(t)\exp(-st)\,dt \\ &= s\exp(-sa)G(s) \end{aligned} \tag{9.38}$$

Putting this all together, we obtain

$$W(s) = sF(s) - f(0) + s\exp(-sa)G(s) \tag{9.39}$$

■

Example 9.5 (Laplace transform of derivatives.). Find the Laplace transform of the following

1. Find $\mathscr{L}[f'(t)]$ for $f(t) = t^2 + f_0$. The ancillary condition is $f(0) = f_0$, where f_0 is some constant.
2. Find $\mathscr{L}[f''(t)]$ for $f(t) = \sin(t)$ if the ancillary conditions are $f(0) = 5$, and $f'(0) = 2$.
3. Find Find $\mathscr{L}[w'(t)]$ for $w(t) = (t-5)^2H(t-5)$ if the ancillary condition is $w(0) = 0$.

解.

1. Recall, for this case we have the theorem $[f'(t)] = sF(s) - f(0)$; thus, we need the transform of f and the value of $f(0)$. The transform is given by $\mathscr{L}[t^2 + f_0] = 2/s^3 + f_0/s$. Note that from the expression for f, we have $f(0) = f_0$. According to the theorem, then, we have

$$\begin{aligned}\mathscr{L}[f'(t)] &= s(2/s^3 + f_0/s) - f_0 \\ &= 2/s^2 + f_0 - f_0 \\ &= 2/s^2\end{aligned}$$

2. According to the theorem above, the Laplace transform of the second derivative is given by $\mathscr{L}[f''(t)] = s^2F(s) - sf(0) - f'(0)$. Thus, we need the transform of f, and the values of $f(0)$ and $f'(0)$. Collecting these, we have

$$\begin{aligned}\mathscr{L}[f(t)] = \mathscr{L}[\sin(t)] &= \frac{1}{1+s^2} \\ f(0) = \sin(0) &= 0 \\ f'(0) = \cos(0) &= 1\end{aligned}$$

Using the theorem, the result is

$$\mathscr{L}[f''(t)] = s^2 \frac{1}{1+s^2} - 1$$

或者，经过简单的代数运算，可以将其写成以下形式

$$\mathscr{L}[f''(t)] = -\frac{1}{1+s^2}$$

3. 间断函数导数的变换定理要求我们确定$w(t) = f(t) + g(t-a)$赫维赛德阶跃函数$(t-a)$。在这里，很明显，$g(t-a)$赫维赛德阶跃函数$(t-a) = (t-a)^2$赫维赛德阶跃函数$(t-a)$，但是对于函数$f(t)$我们应该怎么办呢？一种解法是取$f(t) = 0$。然后，我们得到

$$\begin{aligned} W(s) &= s \times 0 - 0 + s\exp(-5s)\frac{2}{s^3} \\ &= \frac{2\exp(-5s)}{s^2} \end{aligned}$$

9.8.1 时间卷积

在傅里叶变换一章中，卷积（我们通常认为它存在于空间中）被定义为两个函数乘积的积分，其中一个函数相对于另一个函数发生了平移。已证明卷积具有以下实用性质：（1）两个函数的卷积的变换 等于这两个函数的变换的乘积；（2）两个变换函数的乘积的逆变换 等于这两个函数（未经变换的状态下）的卷积。拉普拉斯变换中的卷积也具有相同的性质。一个主要区别是，拉普拉斯变换中的卷积仅在半轴上定义。具体而言，两个函数f和g 的卷积定义为

$$(f * g)(t) = \int_0^t f(t-\tau)g(\tau)\,d\tau \tag{9.40}$$

为了观察变换的“单边性”对所得卷积的影响，我们可以重复一项与之前计算的傅里叶变换卷积示例类似的分析。在这里，我们再次考虑两个紧支撑函数。首先，我们有一个紧支撑 但连续的函数，其定义为

$$w(t) = \frac{a\pi}{4}\left[H(t) - H\left(t - \tfrac{2}{a}\right)\right]\cos\left(\tfrac{\pi}{2}(at-1)\right) \tag{9.41}$$

where a is any positive number. Recall that a compact function is nonzero on some finite interval, and zero everywhere else. For this function, it is compact on the interval $[0, 2/a]$. It is also a technically *probability density* function or normalized *weighting function* because the area under the curve is 1 for all values of a. The second function we will look at is the *boxcar* function, which is the difference between two Heaviside functions. It is defined as it was in the chapter on Fourier transforms.

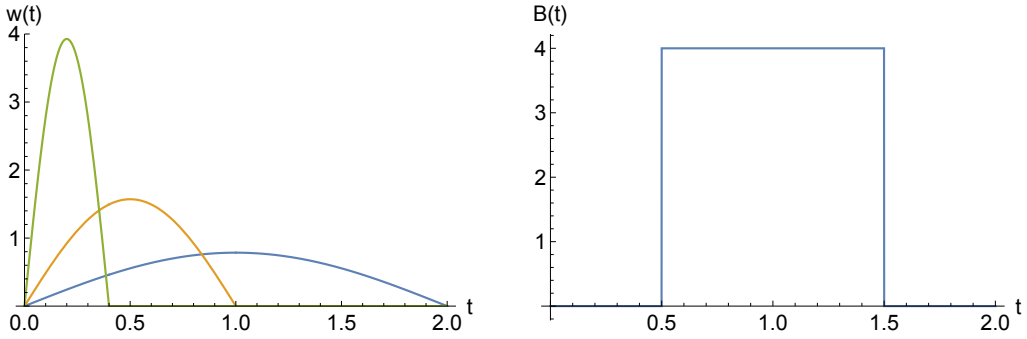


Fig. 9.5: (Left) The weighting function $w(t)$ with parameters $a = 1$ (blue), $a = 2$ (orange) and $a = 5$ (green). (Right) The boxcar function $B(t;t_1,t_2)$ with parameters $h_0 = 4$, $t_1 = \frac{1}{2}$, and $t_2 = \frac{3}{2}$.

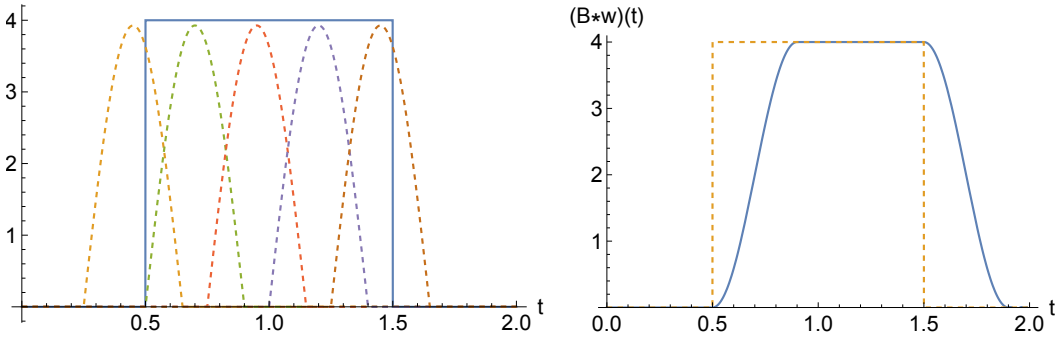


Fig. 9.6: (Left) The boxcar function with five copies of the weighting function superimposed at discrete positions. The convolution computes the integral of the multiplication of these two functions as they are *continuously* overlapped over the bounds of the integral. (Right) The boxcar function $B(t;\frac{1}{2},\frac{3}{2})$ with parameters $h_0 = 4$ after convolution with the weighting function $w(t)$ with $a = 5$. The result is the smoothed function shown in blue. The original function is shown in the dashed orange line for comparison.

$$B(t;t_1,t_2) = h_0[H(t-t_1) - H(t-t_2)] \tag{9.42}$$

where here t_1, t_2, and h_0 are parameters. The first two are the starting and ending points for the function interval (that is, the function is non-zero in $t \in [t_1,t_2]$), and h_0 sets the height of the function. Plots of $w(t)$ and $B(t)$ are given in Fig. 9.5.

The convolution defined on the half-line looks different than the convolution defined on the whole line (see Fig. 8.5 for comparison), even thought the two functions used in these two examples have the same shape. The difference is that, for the Fourier transform, the weighting function w was centered around zero. For the Laplace transform, this is not possible because the transform is defined for only $t > 0$. Thus, the weighting function must be shifted forward in time by one-half the width of $w(t)$ so that it is defined over a positive domain.

The practical result of this is that the convolution behaves qualitatively differently than for the zero-centered weighting function used in the previous example given in the Chapter on Fourier transforms

(Fig. 8.5). If one examines Fig. 9.6, it is possible to see this difference. Note that the *shape* of the resulting function is the same as for the previous example, but instead of the center of mass of the convolution being conserved (i.e., the same before and after the operation), it has moved to the right. This results exclusively because the weighting function has itself been moved to the right so that it is defined on the domain $t > 0$.

These differences are not merely an interesting side note. Convolutions in time are often used to represent the *response* of a system to some signal or process in time. Thus, in Fig. 9.6 we would not expect the convolution to allow a *response* to the weighting function at times earlier than $t = 1/2$ which is the time that the the boxcar function begins. Allowing the convolution to do so would, in essence, allow the weighting function to weight components of the boxcar function that had not arrived yet! This feature of time convolutions is sometimes known as *obeying causality* or as the convolution being *non-anticipative*. In short, it means simply that time convolutions do not allow their weighting functions to act on portions of the signal (in this example, the boxcar function) that have not yet arrived in time. No output can precede the input that led to it.

9.9 Convolutions and the Laplace Transform

与傅里叶变换类似，拉普拉斯变换在处理卷积积分时也具有良好的性质。卷积积分有两种产生方式。第一，在尝试求拉普拉斯变换的逆变换时，所得结果可能可以表示为时间变量的卷积。第二，可能会遇到包含卷积的积分微分方程；拉普拉斯变换可以帮助处理这类复杂问题。以下两个定理不加证明地给出。

定理 9.5（逆卷积定理）。假设 $\mathscr{L}[f(t)] = F(s)$ 且 $\mathscr{L}[g(t)] = G(s)$。令 $w(t) = \mathscr{L}^{-1}[F(s)G(s)]$。则有

$$w(t) = \int_0^t f(t-\tau)g(\tau)\,d\tau = \int_0^t f(\tau)g(t-\tau)\,d\tau \tag{9.43}$$

例 *9.6*（逆卷积示例）。假设我们正在用拉普拉斯变换处理一个问题，并得到了以下结果

$$W(s) = \frac{a}{s^2(a^2+s^2)}$$

查阅拉普拉斯变换表后，我们发现其中没有列出这样的变换。不过，若令 $F(s) = 1/s^2$ 且 $G(s) = a/(a^2+s^2)$，可以看到，这两个变换都列在表中。我们有

$$F(s) = \frac{1}{s^2} \to f(t) = t$$
$$G(s) = \frac{a}{a^2+s^2} \to g(t) = \sin(at)$$

因此，式（9.6）的逆变换为

$$w(t) = \int_0^t \tau \sin[a(t-\tau)]\, d\tau$$

Such convolutions are not always easy to solve. In this case, however, the integral can be done by parts. The result is

$$w(t) = \frac{1}{a^2}[at - \sin(at)]$$

Technically, we have discovered a new Laplace transform pair this way! We could put the following pair in the table if we were so inclined.

$$\frac{1}{a^2}[at - \sin(at)] \xrightarrow{\mathscr{L}} \frac{a}{s^2(a^2+s^2)}$$

The second convolution theorem allows us to convert convolutions in real space to multiplications in transform space.

Theorem 9.6 (Forward convolution theorem.). *Suppose $\mathscr{L}[f(t)] = F(s)$ and $\mathscr{L}[g(t)] = G(s)$. Let $w(t)$ be given by the convolution*

$$w(t) = \int_0^t f(t-\tau)g(\tau)\, d\tau = \int_0^t f(\tau)g(t-\tau)\, d\tau \tag{9.44}$$

Then

$$\mathscr{L}[w(t)] = W(s) = F(s)G(s) \tag{9.45}$$

例 *9.7*（卷积的拉普拉斯变换）。

Integro-differential equations are notoriously difficult to solve. The Laplace transform provides one 方法 d for potential solution. Take the following example

$$\begin{aligned} \frac{du}{dt} &= \int_0^t \exp(t-\tau)u(\tau)\, d\tau \\ u(0) &= 1 \end{aligned}$$

For this problem, if we take the Laplace transform of both sides, we find

$$sU(s) - 1 = \frac{1}{s-1}U(s)$$

Solving this for $U(s)$ gives

$$U(s) = \frac{s-1}{s(s-1)-1}$$

Using the table of transforms, we find the solution

$$u(t) = \tfrac{1}{5}e^{t/2}\left[5\cosh\left(\tfrac{\sqrt{5}t}{2}\right) - \sqrt{5}\sinh\left(\tfrac{\sqrt{5}t}{2}\right)\right]$$

We can easily check this solution. While it is tedious to work out, the derivative of the solution is

$$\frac{du}{dt} = \frac{2e^{t/2}\sinh\left(\tfrac{\sqrt{5}t}{2}\right)}{\sqrt{5}}$$

We can also compute, in this case, the integral

$$\begin{aligned}\frac{du}{dt} &= \int_0^t \tfrac{1}{5}e^{\tau/2}\left[5\cosh\left(\tfrac{\sqrt{5}\tau}{2}\right) - \sqrt{5}\sinh\left(\tfrac{\sqrt{5}\tau}{2}\right)\right]\exp(t-\tau)\,d\tau \\ &= \frac{2e^{t/2}\sinh\left(\tfrac{\sqrt{5}t}{2}\right)}{\sqrt{5}}\end{aligned}$$

So, the solution checks out. A plot of the solution is provided below.

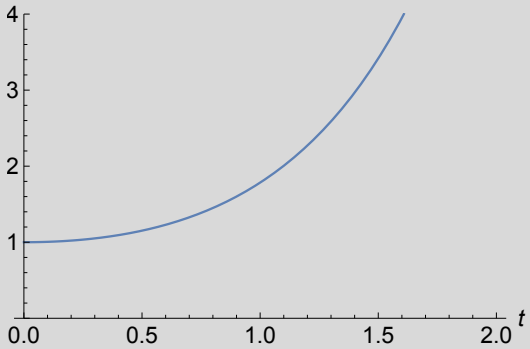


Fig. 9.7: The solution to the integro-differential equation specified in this example.

Note that this feature of Laplace transforms allows us to transform general integrals of time. The trick here is to realize that a generic integral in time is just a convolution with the function $\mu(t) = 1, t \in (-\infty, \infty)$. Note that *this function is distinct from the Heaviside function* because it contains no jumps anywhere in the domain. Additional note that all translations of this function are identical– they are all equal to for all values of t; thus $\mu(t) = \mu(t-a)$ for all values of a. Now, suppose $f(t)$ is a function that has a well-defined Laplace transform. Then, consider the following integral.

$$w(t) = \int_{\tau=0}^{\tau=t} f(\tau)\,d\tau \tag{9.46}$$

其中， τ 是积分变量，并且我们在积分上下限中更明确地标示了积分变量。注意到我们已定义单位函数 $\mu(t) = 1$，因此可以将此函数写成卷积

$$w(t) = \int_{\tau=0}^{\tau=t} 1 \times f(\tau)\, d\tau$$
$$= \int_{\tau=0}^{\tau=t} \mu(t-\tau) f(\tau)\, d\tau \tag{9.47}$$

Noting that $\mathscr{L}[\mu(t)] = \mathscr{L}[1] = \frac{1}{s}$, then the result for *any* integral of time is the transform

$$\mathscr{L}\left[\int_{\tau=0}^{\tau=t} f(\tau)\, d\tau\right] = \frac{1}{s} F(s) \tag{9.48}$$

9.10 使用拉普拉斯变换求常微分方程的解

拉普拉斯变换的主要用途之一是求解线性常微分方程。采用这种方法求解有一些优点。首先，非齐次方程或积分微分方程等复杂问题，都可以通过拉普拉斯变换以相当直接的方式处理。以这种方式求解常微分方程时，主要困难在于问题完成后确定适当的逆变换。

关于常微分方程的应用，我们没有太多其他内容需要学习。请注意，按照惯例，拉普拉斯变换用于变换时间 变量。然而，定义拉普拉斯变换的主要特征并不是自变量的名称（即时间，用t表示），而是变换定义在半轴上。因此，定义在半轴上的任何常微分方程都适合使用拉普拉斯变换处理，无论自变量是时间、空间，还是其他变量。下面的几个实例将有助于巩固使用拉普拉斯变换求解线性常微分方程的方法。

Example 9.8 (A steady-state diffusion-reaction problem). As mentioned above, the application of the Laplace transform does not require that the independent variable be time, but, rather, that the independent variable be defined on the half-line. The following is an example where we look at the solution to a steady-state reaction-diffusion problem defined on $x \in [0,\infty)$. Recall, the heat/diffusion equation with a first-order reaction at steady state is, by definition, no longer a function of time (i.e., by definition, $\partial u/\partial t \equiv 0$ for steady state to exist). Suppose we have a steady reaction-diffusion problem on the half real line as follows

$$D\frac{d^2u}{dx^2} - ku = 0, \;\; x \in [0,\infty)$$
$$u(x=0) = u_0$$
$$\left.\frac{du}{dx}\right|_{(x\to\infty)} = 0$$

For simplicity, let $\alpha = k/D$. Noting that the domain for space is $x \in [0,\infty)$, the Laplace transform is the correct one to use. We have one bit of trouble here, however. We are not given the derivative of the function at $x = 0$; instead, we are given a *requirement* that must be respected as x grows large. We can still proceed, however. Assume for now that we specify

$$\left.\frac{du}{dx}\right|_{(} x=0) = c_1$$

where c_1 is some currently unknown constant. Then, we can proceed as follows. Transforming the ODE yields

$$[s^2U(s) - su_0 - c_1] - \alpha U(s) = 0$$

Note that the Laplace transform *always* automatically incorporates the ancillary information into the solution. This is easy to solve, with the result being

$$U(s) = \frac{su_0 + c_1}{s^2 - \alpha}$$

Referring to the Laplace transform tables, we find that the inverse transform is given by

$$\begin{aligned}
\mathscr{L}^{-1}[U(s)] &= u_0\mathscr{L}^{-1}\left[\frac{s}{s^2-\alpha}\right] \\
&= u_0\cosh[\sqrt{\alpha}x] + c_1\mathscr{L}^{-1}\left[\frac{1}{s^2-\alpha}\right] \\
&= u_0\cosh[\sqrt{\alpha}x] + c_1\sinh[\sqrt{\alpha}x] \\
&= u_0\cosh\left[\sqrt{\frac{k}{D}}x\right] + c_1\sinh\left[\sqrt{\frac{k}{D}}x\right]
\end{aligned}$$

为了求出这个常数，我们可以先求导，然后令这些导数为零。或者，我们注意到 $\cosh(x)$/双曲正弦$(x) \to 1$ 当 $x \to \infty$ 时。因此，如果我们令 $c_1 = -1$，那么我们的

ancillary condition is met. If we set $k/D = 0.1$ and $u_0 = 1$, the the solution is given by the plot shown 下方。

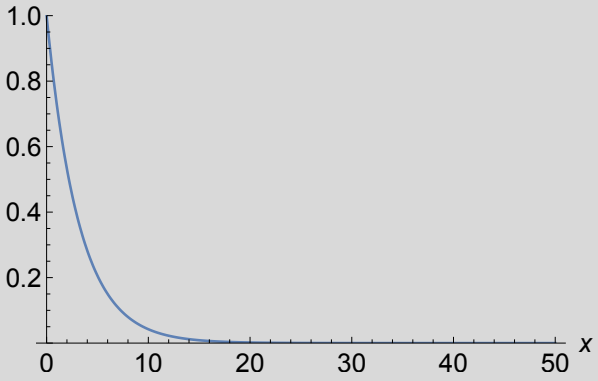


图 9.8：上述振荡质量问题的解。这里，纵轴表示相对于平衡（零）位置的位移。

例 *9.9*（使用拉普拉斯变换求解二阶常微分方程）。 在力学中分析带摩擦力的动态质量—弹簧系统的运动时，会遇到以下问题：摩擦力会抑制运动。在此表达式中，变量 $y(t)$ 表示相对于平衡位置的位移，且可以为正，也可以为负。

$$\underbrace{my''(t)}_{mass\times accelleration} = -\underbrace{\beta_0 y'(t)}_{friction\ damping} - \underbrace{k_0 y(t)}_{spring\ force} \tag{9.49}$$

如果质量块还受到一个驱动其运动的外力 $F(t)$（例如，可以设想用磁力驱动铁磁性质量块），那么就将此驱动力函数加到方程右端，从而得到一个非齐次常微分方程。假设该系统受到周期性驱动，驱动力为一个振荡且衰减的电磁力，$f(t) = \exp(-t)\sin(t)$。方程及其初始位置和速度可以表示为

$$\begin{aligned} y''(t) + \beta y'(t) + k y(t) &= \exp(-t)\sin(t) \\ y'(0) &= 0 \\ y(0) &= 0 \end{aligned}$$

where here, to simplify, $\beta = \beta_0/m$ and $k = k_0/m$. Ordinarily, this would pose a very difficult problem to solve. Using Laplace transforms will make this somewhat easier, but we will still encounter some challenges in inverting the solution. For this problem, assume that $\beta = 2$ and that $k = 2$. Proceeding, it is fairly direct to find the transform of the ODE.

$$\left[s^2Y(s)-sy(0)-y'(0)\right]+\beta\left[sY(s)-y(0)\right]+kY(s)=\frac{1}{1+(1+s)^2}$$

利用附加条件以及 β 和 k 的值，我们得到

$$(s^2+2s+2)Y(s)=\frac{1}{1+(1+s)^2} \tag{9.50}$$

注意，左边可以写成

$$[(s+1)^2+1]Y(s)=\frac{1}{1+(1+s)^2} \tag{9.51}$$

解出 $Y(s)$ 得到

$$Y(s)=\frac{1}{1+(1+s)^2}\frac{1}{1+(1+s)^2} \tag{9.52}$$

好消息是，我们已经得到了拉普拉斯域中的解。然而，我们仍需要将其逆变换回时间函数。在此，注意到我们可以从拉普拉斯变换表中查到 $\mathscr{L}^{-1}\left[1/(1+(1+s)^2)\right]=\exp(-t)\sin(t)$ 的逆变换。在本例中，我们将该变换与自身相乘。令 $F(s)=1/[1+(1+s)^2]$，且 $G(s)=1/[1+(1+s)^2]$,，根据卷积定理可得

$$\mathscr{L}^{-1}[Y(s)]=\mathscr{L}^{-1}\left[\frac{1}{1+(1+s)^2}\frac{1}{1+(1+s)^2}\right]$$

$$y(t)=\int_0^t \exp(-\tau)\sin(\tau)\exp[-(t-\tau)]\sin(t-\tau)\,d\tau$$

可以利用欧拉公式将正弦函数转换为指数函数，从而计算这个积分。然后即可计算该积分（它是指数函数的积分），并可反向利用欧拉公式，从结果中还原出正弦函数和余弦函数。计算过程在代数上较为复杂，但除此之外并不困难。结果为

$$y(t)=\frac{1}{2}\exp(-t)\left[-t\cos(t)+\sin(t)\right] \tag{9.53}$$

As always, it is a good idea to check our answers. Toward that goal, we have

$$y'(t)=\frac{1}{2}\exp(-t)t\sin(t)-\frac{1}{2}\exp(-t)(\sin(t)-t\cos(t))$$

$$y''(t)=\frac{1}{2}\exp(-t)\sin(t)-\exp(-t)t\sin(t)+\frac{1}{2}\exp(-t)t\cos(t)+\frac{1}{2}e^{-t}(\sin(t)-t\cos(t))$$

Computing the left-hand side sum $y''(t)+2y'(t)+2y(t)$ gives the result $\exp(-t)\sin(t)$. This matches the right-hand side of the original ODE, so we have verified that our solution is correct. A plot of the solution appears below.

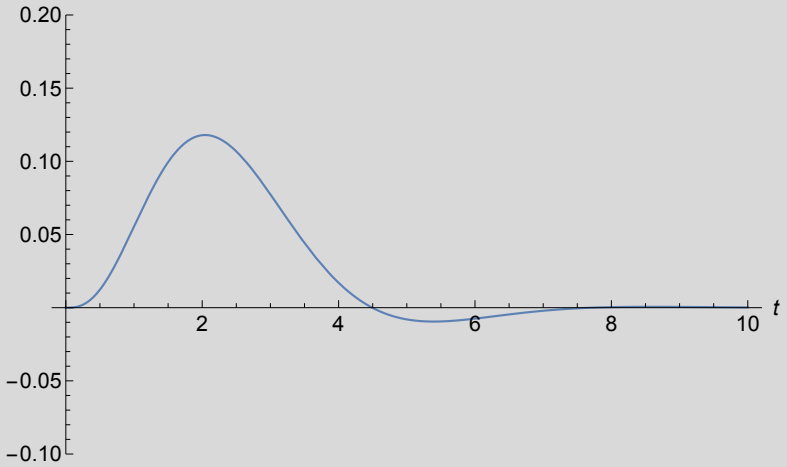


Fig. 9.9: The solution for the oscillating mass problem described above. Here, the vertical axis defines the displacement relative to the equilibrium (zero) position.

9.11 Solutions to PDEs Using Laplace Transforms

The solution to PDEs using Laplace transforms is very similar to the use of Fourier transforms for the same purpose. The primary benefit realized by using the Laplace transform is, as for the Fourier transform, the conversion of derivatives into algebraic quantities. As might be expected, the Laplace transform is a tool that to transform independent variables on the half-line.

在偏微分方程的拉普拉斯变换应用中，实际上并没有什么新的概念需要学习。然而，求解这类问题的过程有时确实需要一些创造性思维。为了说明这一点，下面给出一些使用拉普拉斯变换求解偏微分方程的实例。

Example 9.10 (All roads lead to Rome: Solution to the transient first-order wave equation in an infinite domain). In Chap. 6 we introduced the first-order wave equation. That equation was revisited in Chap. 8, where we used Fourier transforms to solve the problem on the real line in space for a specified initial condition. In particular, in Example 8.3, we found the solution to the following problem.

$$\frac{\partial u}{\partial t} = -c\frac{\partial u}{\partial x} \qquad -\infty < x < \infty$$

$B.C.1$ $\quad u(x,t)$ remains bounded for all x and t

$I.C.$ $\quad u(x,0) = \exp(-x^2)$

Using Fourier transforms, we found that the solution was

$$u(x,t) = \exp\left[-(x-ct)^2\right]$$

This solution should, mathematically speaking, be entirely independent of the method of solution (assuming that the method is a valid one!) Thus, we *should be able* to find this same solution via the Laplace transform. To do so, we can, as usual, take the transform of the PDE (which will automatically incorporate the initial condition when using the Laplace transform). The result of that step is as follows

$$\mathscr{L}\left[\frac{\partial u}{\partial t}\right] = \mathscr{L}\left[-c\frac{\partial u}{\partial x}\right] \qquad -\infty < x < \infty$$

$$I.C. \qquad u(x,0) = \exp(-x^2)$$

with the understanding that the solution must remain bounded. Taking the Laplace transform, we find (after some rearrangement) the following first-order nonhomogeneous ODE.

$$\frac{\partial U(s)}{\partial x} + \frac{s}{c}U(s) = \frac{1}{c}\exp(-x^2)$$

The solution to linear first-order ODEs in two variables is always possible (assuming that they are well-posed), at least as an integral solution. Recall, we need an *integrating factor* of $\exp(sx/c)$ to solve this particular problem. Multiplying both sides of the equation by the integrating factor yields

$$\exp\left(\frac{sx}{c}\right)\frac{dU(x,s)}{dx} + \exp\left(\frac{sx}{c}\right)\frac{s}{c}U(x,s) = \frac{1}{c}\exp\left(\frac{sx}{c}\right)\exp(-x^2)$$

The first two terms are the derivative of $\exp\left(\frac{sx}{c}\right)U(x,s)$, thus we have

$$\frac{d}{dx}\left[\exp\left(\frac{sx}{c}\right)U(x,s)\right] = \frac{1}{c}\exp\left(\frac{sx}{c}\right)\exp(-x^2)$$

Integrating both sides with respect to x gives us

$$\int_{z=-\infty}^{z=x}\frac{d}{dz}\left[\exp\left(\frac{sz}{c}\right)U(x,s)\right]dz = \frac{1}{c}\int_{z=-\infty}^{z=x}\exp\left(\frac{sz}{c}\right)\exp(-z^2)\,dz$$

where here we have been careful to use a variable of integration (z) because the independent variable x appears in the bounds of the integral. The left-hand side of this equation is easy to evaluate. The right-hand side is actually somewhat *difficult* to evaluate because it integrates to some kind of an error-function. Rather than evaluate the right-hand side, we will leave it in integral form for now. This is *not* usual. However, it will ultimately serve us well (and the author, of course, has already solved this problem, so this choice comes from knowledge of at least one method that leads to a solution!) Thus, we have

$$\left[\exp\left(\frac{sz}{c}\right)U(x,s)\right]\Big|_{z=-\infty}^{z=x} = \frac{1}{c}\int_{z=-\infty}^{z=x}\exp\left(\frac{sz}{c}\right)\exp(-z^2)\,dz$$

and evaluating, we have

$$\exp\left(\frac{sx}{c}\right)U(x,s) = \frac{1}{c}\int_{z=-\infty}^{z=x}\exp\left(\frac{sz}{c}\right)\exp(-z^2)\,dx$$

or, simplifying

$$U(x,s) = \frac{1}{c}\int_{z=-\infty}^{z=x}\exp\left(\frac{s(z-x)}{c}\right)\exp(-z^2)\,dz$$

Recall that the Laplace transform is an integration. You may recall from calculus that we can change the order of integration of two integrals if the functions involved have defined integrals, and the region that the double integral covers is *simple* (that is, the region formed on the $s-x$ plane is bounded by two functions). For improper integrals, the situation is a bit more complicated, and the general result is known as Fubini's theorem. In short, this theorem says that we can change the order of integration if the absolute value of the functions converge for each integral. This is true for our integrals, so changing the order of integration is allowed. (N.B., for the purposes of comparison of Fourier and Laplace transforms, this example goes beyond what are covering in this text; you would not be expected to know when one might apply Fubini's theorem!) In our case, we have the following result

$$\begin{aligned}\mathscr{L}^{-1}[U(x,s)] &= \mathscr{L}^{-1}\left[\frac{1}{c}\int_{z=-\infty}^{z=x}\exp\left(\frac{s(z-x)}{c}\right)\exp(-z^2)\,dz\right]\\ &= \frac{1}{c}\int_{z=-\infty}^{z=x}\mathscr{L}^{-1}\left[\exp\left(\frac{s(z-x)}{c}\right)\right]\exp(-z^2)\,dz\end{aligned}$$

Consulting the Laplace transform tables, this gives

$$\begin{aligned}u(x,t) &= \mathscr{L}^{-1}\left[\frac{1}{c}\int_{z=-\infty}^{z=x}\exp\left(\frac{s(z-x)}{c}\right)\exp(-z^2)\,dz\right]\\ &= \frac{1}{c}\int_{z=-\infty}^{z=x}\mathscr{L}^{-1}\left[\exp\left(\frac{s(z-x)}{c}\right)\right]\exp(-z^2)\,dz\\ &= \frac{1}{c}\int_{z=-\infty}^{z=x}\delta\left(\frac{tc-x+z}{c}\right)]\exp(-z^2)\,dz\end{aligned}$$

这里，狄拉克δ函数仅在$tc-x+z=0$或$z=x-ct$时非零。因此，积分得到结果

$$u(x,t) = \exp[-(x-ct)^2]$$

将其与通过傅里叶变换得到的结果进行比较，我们看到所得答案相同。不过，这种情况表明，拉普拉斯变换可能并不是最便于求解的途径。请记住一个很好的概念：虽然针对某个特定问题所应用的所有有效数学方法都应得到相同的解，但这并不意味着每种方法都需要付出相同的努力！在本例中，傅里叶变换仅用几行数学推导就解决了一阶波动方程问题。对于拉普拉斯变换，结果相同，但方法要复杂得多。一般而言并非总是如此——对于某些问题，拉普拉斯变换会比傅里叶变换更容易得到解。当两种方法都适用时，可以将两种方法都尝试一下。经过一些分析后，应该会比较清楚哪种方法能够更容易地得到结果。

Example 9.11 (Solution to the transient second-order wave equation in an infinite domain). Solutions to the wave equation are to many less intuitive than solutions to, for example, the heat/diffusion equation. This is a result of the second-order derivative in time, which does not generate the same intuitive interpretation as does the first derivative in time (which can be interpreted an accumulation at fixed location term)

拉普拉斯变换通过变换时间变量，使波动方程的求解成为可能；对于时间的二阶导数，这种方法尤其有用。让我们在一个半无限域中定义一个波动方程。在这里，我们将在$x = 0$处设置边界条件，以此来“驱动”该问题。

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2}$$

*B.C.*1	$u(0,t) = \sin(t)$
*B.C.*2	$u(x,\ t)$ 对所有 $x > 0$ 和 $t > 0$均保持有界。
*I.C.*1	$u(x,0) = 0$
*I.C.*2	$\left.\frac{\partial u}{\partial t}\right\|_{(x,0)} = 0$

Taking the Laplace transform of the PDE gives us the following result

$$\frac{d^2U}{dx^2} - s^2 U(s) = 0$$

*B.C.*1.	$U(x,s) = \frac{1}{s^2+1}$
*B.C.*2	$U(x,s)$ remains bounded for all $x > 0$ and $s > 0$

This is a homogeneous, second-order ODE in the variable x with a well-specified boundary condition at $x = 0$. Note that here the two boundary conditions have been transformed as well as the PDE itself. Also note, as it typical for Laplace transforms, the ancillary conditions associated with the variable

being transformed (in this example, the initial conditions, because the transformed variable is t) are incorporated directly into the transformation via the rules for transforms of derivatives.

To solve this problem, we use the standard methods for second-order ODEs with constant coefficients. The characteristic equation is $r^2 - s^2/c^2 r = 0$. Using the quadratic formula (noting $a = 1$, $b = 0$, and $c = -s^2$, this yields two real roots, $r_1 = -s/c$ and $r_2 = s/c$. The solution is

$$U(x,s) = c_1 \exp\left(-\frac{sx}{c}\right) + c_2 \exp\left(\frac{sx}{c}\right)$$

Noting that the solution must remain bounded for all $x > 0$ and $t > 0$, then we must have that $c_2 = 0$. The constant c_1 can be found using the boundary condition. At $x = 0$ we find that $c_1 = 1/(s^2+1)$. The solution is, then

$$U(x,s) = \frac{1}{s^2+1} \exp\left(-\frac{sx}{c}\right)$$

The final step is to convert this solution back into real time rather than the transform variable s. Taking the inverse transform of both sides of the solution gives us

$$\mathscr{L}^{-1}[U(x,s)] = \mathscr{L}^{-1}\left[\frac{1}{s^2+1} \exp\left(-\frac{x}{c}s\right)\right]$$

There are two ways that the inversion of the right-hand side can be done. First, we might note that $1/(s^2+1)$ and $\exp(-xs/c)$ are both transforms that are available on the table of Laplace transforms. Thus, the convolution method could be used here. A simpler approach is to recall the *shifting* property of the Laplace transform. The exponential term in the result can also be interpreted as a shifting operator, whose role is to shift the independent variable t by the amount $-x/c$ during the inversion of $1/(s^2+1)$. Taking this latter approach, we find the solution to be

$$u(x,t) = \sin\left(t - \tfrac{x}{c}\right) H\left(t - \tfrac{x}{c}\right)$$

下图给出了几个具有代表性的时间点对应的解。

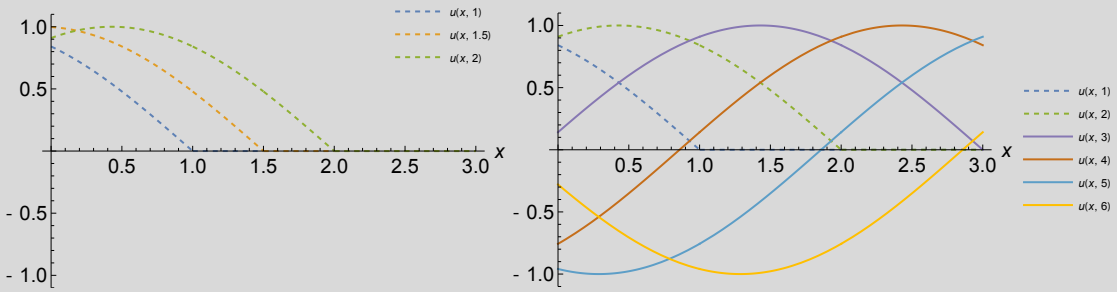


Fig. 9.10: The solution to the wave equation on a half-infinite domain. For this plot, the vertical axis is wave displacement (distance), x is measured in m, and $c = 1$ m/s. The times are between $t = 1$ and $t = 6$ s.

Example 9.12 (solution to the transient heat/diffusion equation on the half-line). The heating of an insulated rod by the application of a constant temperature at one end is a problem we have examined previously using separation of variables. If the rod is very long compared to the characteristic time scale of heat transport and only "early" times (several multiples of the characteristic time scale) are of interest, it can be useful to treat the rod as being infinitely long.

Our heat transport problem will take the following form. Here, assume that the temperature, u is given in non-dimensional form (i.e., it has been normalized such that the boundary temperature is 1 and the initial temperature is zero).

$$\frac{\partial u}{\partial t} = K\frac{\partial^2 u}{\partial x^2}$$

$$B.C.1 \qquad u(0,t) = u_1$$

$$B.C.2 \qquad u(x,t) < \infty$$

$$I.C. \qquad u(x,0) = 0$$

The interpretation here is that the rod is initially of uniform temperature, $u = 0$. At $t = 0$, the left end is suddenly put in contact with a large reservoir of heat at temperature $u(0,t) = u_1$. The second boundary condition is set to require that the temperatures are not allowed to grow without bound, which would be physically unrealistic.

First, note that in principle, we could conduct the Laplace transform on either t or on x (or both!), since both variables are defined on the interval $[0,\infty)$. To keep the problem from getting complicated, we will transform only the variable t (which is the conventional choice).

通过求出该偏微分方程（PDE）及两个边界条件的拉普拉斯变换，即可得到其解。假设边界条件依赖于时间，即使它们是时间的常函数也是如此（这一点在两个边界条件的自变量列表中也有明确体现）。其变换如下所示。

$$sU(s) - 0 = K\frac{\partial^2 U}{\partial x^2} \qquad 0 \le x < \infty,\ t > 0$$

$$B.C.1 \qquad U(0,s) = \frac{u_1}{s}$$

$$B.C.2 \qquad U(x,s) < \infty$$

$$I.C. \qquad u(x,0) = 0$$

Note that, because we are taking the transform of the time variable, the Laplace transform commutes with the spatial derivative. In other words, *assuming that t is the transformed variable*, then $\mathscr{L}[\partial u/\partial x] = \partial/\partial x\mathscr{L}[u] = \partial U/\partial x$. Obviously this argument can be repeated for higher-order derivatives. Now we have a second-order, non-homogeneous ODE with constant coefficients. The ODE takes the form

$$\frac{\partial^2 U}{\partial x^2} - \frac{s}{K}U(s) = 0$$

$$B.C.1 \qquad U(0,s) = \frac{u_1}{s}$$

$$B.C.2 \qquad U(x,s) < \infty$$

The solution to this ODE is straightforward. The characteristic equation is given by $r^2 - s/Kr = 0$, and this indicates that $a = 1$, $b = 0$, and $c = -s/K$. Thus the roots are $r_1 = +\sqrt{4s/K}$ and $r_2 = -\sqrt{4s/K}$. The solution must be of the general form

$$U(x,s) = c_1 \exp\left(-2x\sqrt{\frac{s}{K}}\right) + c_2 \exp\left(2x\sqrt{\frac{s}{K}}\right)$$

And, without much effort, we can immediately see that $c_2 = 0$ because of the boundedness imposed by the second boundary condition.

$$U(x,s) = c_1 \exp\left(-2x\sqrt{\frac{s}{K}}\right)$$

Using the first boundary condition, we find that $c_1 = \frac{u_1}{s}$, which gets us to the point (noting that the quantity $2x$ is moved to inside the root)

$$U(x,s) = \frac{u_1}{s} \exp\left(-\sqrt{\frac{4x^2}{K}}\sqrt{s}\right)$$

From the table of Laplace transforms, the entry number 27 is the one needed here. Taking $a = \sqrt{\frac{4x^2}{K}}$, then we have the Laplace inversion of the form

$$\begin{aligned} u(x,t) &= u_1 \operatorname{erfc}\left(\frac{\sqrt{\frac{4x^2}{K}}}{2\sqrt{t}}\right) \\ &= u_1 \operatorname{erfc}\left(\sqrt{\frac{x^2}{Kt}}\right) \end{aligned}$$

下图给出了几个具有代表性的时间点对应的解。

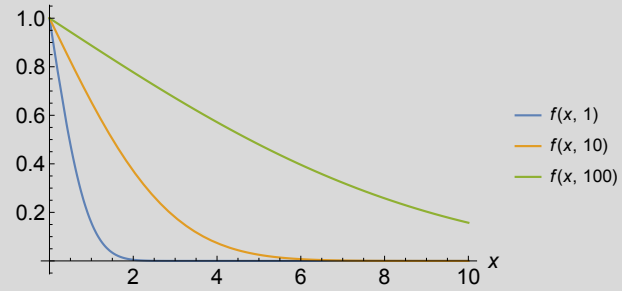


Fig. 9.11: The solution to the heat equation on a half-infinite domain. For this plot, the vertical axis is temperature (difference), the temperature, u has been non-dimensionalized such that $u_1 = 1$, x is measured in mm, and $K = 1$ mm^2/min. The times are taken to be $t = 1$, 10, and 100 minutes.

Appendix 2. Laplace Transform Table

Table 9.1: Laplace transform pairs.

It is assumed that a is a real number ($a \in \mathbb{R}$), and $Re(s) > 0$ unless other constraints are provided.
n 被假定为正整数（$n \in \mathbb{N}$）

	$f(t)$	$F(s)$	约束条件
1.	$\frac{\partial f(t)}{\partial t}$	$sF(s)-f(0)$	
2.	$\frac{\partial^2 f(t)}{\partial t^2}$	$s^2F(s)-sf(0)-f'(0)$	
3.	$\frac{\partial^n f(t)}{\partial t^n}$	$s^nF(s)-s^{(n-1)}f(0)-s^{(n-2)}f'(0)$ $-\ldots-f^{(n-1)}(0)$	
4.	$\exp(-at)$	$\frac{1}{s+a}$	$a>0$
5.	$\exp(at)$	$\frac{1}{s-a}$	$s>a$ and $a>0$
6.	$\exp(at)(1+at)$	$\frac{s}{(s-a)(s-a)}$	$s>a$ and $a>0$
7.	$\frac{\exp(-at)-\exp(-bt)}{a+b}$	$\frac{1}{(s+a)(s+b)}$	$a,b>0$
8.	$\frac{\exp(at)-\exp(bt)}{a-b}$	$\frac{1}{(s-a)(s-b)}$	秒 > a ,b和 a ,b > 0
9.	$\frac{\exp(at)-\exp(-bt)}{a+b}$	$\frac{1}{(s-a)(s+b)}$	秒 > a ,b和 a ,b > 0
10.	$\frac{te^{-at}}{(a-b)^2}+\frac{te^{-bt}}{(a-b)^2}+\frac{2e^{-at}}{(a-b)^3}-\frac{2e^{-bt}}{(a-b)^3}$	$\frac{1}{(s+a)^2(s+b)^2}$	$a,b\in\mathbb{R}$
11.	$\frac{1}{3}e^{-t/2}\left[\sqrt{3}\sin\left(\frac{\sqrt{3}t}{2}\right)+3\cos\left(\frac{\sqrt{3}t}{2}\right)\right]$	$\frac{s+1}{s(s+1)+1}$	

	$f(t)$	$F(s)$	Constraint
12.	$\frac{1}{5}e^{t/2}\left[5\cosh\left(\frac{\sqrt{5}t}{2}\right) - \sqrt{5}\sinh\left(\frac{\sqrt{5}t}{2}\right)\right]$	$\frac{s-1}{s(s-1)-1}$	$s>1$
13.	$t^0 = 1$	$\frac{1}{s}$	
14.	t	$\frac{1}{s^2}$	
15.	t^2	$\frac{2}{s^3}$	
16.	$t^{\frac{1}{2}}$	$\frac{\sqrt{\pi}}{2s}\sqrt{\frac{\pi}{s}}$	
17.	$t^{-\frac{1}{2}}$	$\sqrt{\frac{\pi}{s}}$	
18.	t^n	$\frac{n!}{s^{(n+1)}}$	$n \in \mathbb{N}$
19.	t^a	$s^{-(1+a)}\Gamma(1+a)$	$a>-1$
20.	t^{-a}	$s^{-(1-a)}\Gamma(1-a)$	$a \geq 0$
21.	$t^{\frac{1}{n}}$	$s^{-\left(1+\frac{1}{n}\right)}\Gamma\left(1+\frac{1}{n}\right)$	$n \in \mathbb{N}$
22.	$t^{-\frac{1}{n}}$	$s^{-(1-\frac{1}{n})}\Gamma\left(1-\frac{1}{n}\right)$	$n \in \mathbb{N}$

	$f(t)$	$F(s)$	约束条件
23.	$\frac{1}{a^2+t^2}$	$-\frac{1}{a}\mathrm{Ci}(as)\sin(as)$ $+\frac{1}{2a}[\pi-2\,\mathrm{Si}(as)]\cos(as)$	$a>0$
24.	$\frac{t}{a^2+t^2}$	$\mathrm{Ci}(as)\cos(as)$ $+\frac{1}{2}[\pi-2\,\mathrm{Si}(as)]\sin(as)$	$a>0$
25.	$J_0(at)$	$\frac{1}{\sqrt{s^2+a^2}}$	$a>0$
26.	$\mathrm{erf}\left(\frac{a}{2\sqrt{t}}\right)$	$\frac{1-\exp(-a\sqrt{s})}{s}$	$a>0$
27.	$\mathrm{erfc}\left(\frac{a}{2\sqrt{t}}\right)$	$\frac{\exp(-a\sqrt{s})}{s}$	$a>0$
28.	$\delta(t)$	1	
29.	$\delta(t-a)$	$\exp(-as)$	$a>0$
30.	$\begin{cases} H(t-a) \\ 1-H(a-t) \end{cases}$	$\frac{\exp(-as)}{s}$	$a>0$
31.	$\begin{cases} H(a-t) \\ 1-H(t-a) \end{cases}$	$\frac{1-\exp(-as)}{s}$	$a>0$
32.	$B(t;a,a+\Delta)=$ $[H(t-a)-H(t-(a+\Delta)]$	$\frac{\exp(-as)[1-\exp(-\Delta s)]}{s}$	$\Delta>0$ and $a>0$
33.	$T(t;a)=$ $\frac{t}{a}[H(t)-H(t-a)]$ $+\left(2-\frac{t}{a}\right)[H(t-a)-H(t-2a)]$	$\frac{\exp(-2as)[\exp(as)-1]^2}{s^2}$	$a>0$

	$f(t)$	$F(s)$	约束条件
34.	正弦(at)	$\frac{a}{s^2+a^2}$	$a \in \mathbb{R}$
35.	$\cos(at)$	$\frac{s}{s^2+a^2}$	$a \in \mathbb{R}$
36.	$\exp(-bt)\sin(at)$	$\frac{a}{a^2+(b+s)^2}$	$a \in \mathbb{R},\ b \in \mathbb{R}$
37.	$\exp(-bt)\cos(at)$	$\frac{b+s}{a^2+(b+s)^2}$	$a \in \mathbb{R},\ b \in \mathbb{R}$
38.	$\sin\left(\frac{\pi}{a}t\right)[H(t)-H(t-a)]$	$\frac{\pi a[1-\exp(-as)]}{a^2s^2+\pi^2}$	$a > 0$
39.	$\cos\left(\frac{\pi}{a}t\right)[H(t)-H(t-a)]$	$\frac{sa^2[1+\exp(-as)]}{a^2s^2+\pi^2}$	$a > 0$
40.	$\sin(at+b_0)$	$\frac{s\sin(b_0)+a\cos(b_0)}{s^2+a^2}$	$a < 0$ 或 $a = 0$ 或 $a > 0$
41.	$\cos(at+b_0)$	$\frac{秒\ 余弦(b_0)\ -a正弦(b_0)}{秒^2+a^2}$	$a < 0$ 或 $a = 0$ 或 $a > 0$
42.	$t\sin(at)$	$\frac{2as}{(s^2+a^2)^2}$	
43.	$t\cos(at)$	$\frac{秒^2-a^2}{(s^2+a^2)^2}$	

$f(t)$	$F(s)$	约束条件
44. $\sinh(at)$	$\frac{a}{s^2 - a^2}$	$a < 0$ 或 $a = 0$ 或 $a > 0$ and $s > \lvert a \rvert$
45. $\cosh(at)$	$\frac{s}{s^2 - a^2}$	$a < 0$ 或 $a = 0$ 或 $a > 0$ and $s > \lvert a \rvert$

Notes:

J_0	Bessel function of the first kind, zero order
erf	The error function, $\operatorname{erf}(t) = \frac{2}{\sqrt{\pi}} \int_0^t \exp(-\tau^2)\, d\tau$.
erfc	The complimentary error function $1 - \operatorname{erf}(t)$
δ	The Dirac delta function
H	The Heaviside function
$B(x; a, a+\Delta)$	The boxcar function on $t \in [a, a+\Delta]$
$T(t; a)$	The symmetric triangle function on $t \in [0, 2a]$, with maximum of 1 located at $t = a$
Ci	The cosine integral function, $\operatorname{Ci}(z) = -\int_{\tau=z}^{\infty} \frac{\cos(\tau)}{\tau}\, d\tau$
Si	The sine integral function, $\operatorname{Si}(z) = -\int_{\tau=0}^{\tau=z} \frac{\sin(\tau)}{\tau}\, d\tau$
Γ	The gamma function

习题

练习题

Solve the following problems using Laplace transforms.

1. u″(t)+2u′(t)+5u(t) =0u(0) =1u′(0) =12. u″(t) −2u′(t) +2u(t) =0u(0) =0u′(0) =13. u″(t)−2u′(t) +2 u(t) =0u(0) =5u′(0) =04. u″(t)−3u′(t) +2u(t) =0 u(0) =0u′(0) =−15. u″(t) −3u′(t) +2u(t) =10余弦(t) u(0) =0u′(0) =−16. u″(t)−2u′(t) +2u(t) =指数函数(− t) u(0) =0u′(0) =07. u″(t)−4u′(t) +4u(t) =余弦(t) u(0) =0u′(0) =08. u″(t)−u′(t) +4u(t) =余弦(t) u(0) =0u′(0) =09. u″(t)−4u′(t) =赫维赛德阶跃函数(t −5) u(0) =0u′(0) =010. u″(t)−u′(t) =3t^{2}赫维赛德阶跃函数 (t −1) u(0) =0u′(0) =1

应用与更具挑战性的问题

1. 证明 $\mathscr{L}\left[\frac{df}{dt}\right] = sF(秒) - f(0)$。使用分部积分证明这一结论。2. 求解

$$\frac{du}{dt} = -\int_0^t \exp[-(t-\tau)]u(\tau)\,d\tau$$
$$u(0) = 1$$

3. Redo the analysis of Example 9.12, but this time assume that the initial condition is $u(x,0) = u_0$, a constant. You will find that you generate a nonhomogeneous equation for U upon transformation. Although frequently nonhomogeneous problems create substantial extra work, sometimes when the nonhomogeneous term is a constant, we can propose a change of variables. Please try the change of variables $V = (U + K/su_0)$. You should end up with the ODE and boundary conditions

$$\frac{\partial^2 V}{\partial x^2} - \frac{s}{K}V(s) = 0$$
$$B.C.1 \qquad V(0,s) = \frac{u_1}{s} - \frac{K}{s}u_0$$
边界条件2 $$V(x,s) < \infty$$

Solve this problem, and then substitute back to the original variable U. You should be able to invert the final result.

4. Post反演定理并不特别方便使用，但它是少数几种可以避免在复平面上进行围道积分的方法之一（本书不会讲解围道积分）。不过，为了证明自己能够推导出某些逆变换，请尝试以下问题。使用Post反演定理，证明$F(s) = 1/s$的逆变换是函数$f(t) = 1$。提示：先尝试将k的所有幂化简为较简单的形式。例如

$$t^{k+1}t^{-(k+2)} = (t^k)t(t^{-k})t^{-2} = t^{-1}$$

这类代数运算将帮助你使用Post反演定理成功计算出相应的量。

5. 使用拉普拉斯变换求解以下问题。请注意，与例9.10不同，这个问题会产生一个积分（由积分因子产生），在尝试对结果进行逆变换之前，可以（也应该！）直接求出该积分。

$$\begin{aligned} & \frac{\partial u}{\partial t} = -c\frac{\partial u}{\partial x} \qquad -\infty < x < \infty \\ I.C. \qquad & u(x,0) = H(x) \end{aligned}$$

绘制$c = 1$米/秒和$t = 1$，2秒和3秒时的结果（全部绘制在同一张图上）。在定义域0米$< x <$5米上作图。

6. 使用拉普拉斯变换求解以下一阶波动方程。

$$\begin{aligned} & \frac{\partial u}{\partial t} = -c\frac{\partial u}{\partial x} \qquad -\infty < x < \infty \\ I.C. \qquad & u(x,0) = H(x)\exp(-x) \end{aligned}$$

变换后，这道题会得到一个非齐次一阶常微分方程，可以用积分因子求解。解题时，请务必明确区分变量s和x。以下几点提示或许能帮助你顺利解题。首先，请注意，x方向的积分上下限为$-\infty$和x。其次，在这种情况下，使用积分变量来表示空间积分会非常有帮助（变量z很合适，它能提醒你这是一个空间积分）。最后，请记住，积分海维赛德函数时，可以通过限制积分上下限来处理（即，当$z < 0$时，$H(z)$为零）。最后提示：使用积分因子、通过考虑海维赛德函数来简化积分，并将所有指数函数移至结果右端后，你应该能得到如下形式的常微分方程解

$$U(x,s) = \frac{1}{c}\int_{z=0}^{z=x} \exp\left[-\frac{s}{c}(x-z)\exp(-z)\right] dz$$

For this integral, note that $z > 0$ and that $x > z$. So, in all possible cases, this integral is composed of *two decaying exponentials in z*, so it must be convergent. To actually solve the integral, however, it is much more convenient to put it in the form

$$U(x,s) = \frac{1}{c}\exp\left[-\frac{s}{c}x\right]\left\{\int_{z=0}^{z=x}\exp\left[z(\frac{s}{c}-1)\right]\,dz\right\}$$

This result is offered as a way to *check* your results, of your analysis. Do make sure that you show all of your work to get to this point. Note that the integral in this result has the variable of integration z, so the term $(1-s/c)$ can be treated as essentially a constant for the purposes of integration. After integrating, be sure to separate all functions involving s to the extent possible. You should find a result that can be inverted using the table of Laplace transforms. Plot your results for $c = 1$ m/s and $t = 1, 2$ and 3 s (all on the same plot). Plot on the domain $0 \text{ m} < x < 5$ m.

7. Solve the following heat/diffusion problem using Laplace transforms.

$$\frac{\partial u}{\partial t} = \alpha^2\frac{\partial^2 u}{\partial x^2} \qquad x \in (-\infty,\infty)$$

$B.C.1$ $\quad u(x,t)$ remains bounded for all (x,t)

$B.C.2$ $\quad \left.\frac{\partial u}{\partial x}\right|_{(x,t)}$ remains bounded for all (x,t)

$I.C.$ $\quad u(x,0) = \sin(x)+1$

Upon transforming this problem, you will find that the result is an inhomogeneous second-order ODE in the variable x. You will need to use the method variation of parameters (Chp. 3.4.4) to complete the problem. Plot solutions for $\alpha^2 = 1\times10^5$ cm^{2}s and times $t = 1, 100, 1000$ s over the range $0 < x < 100$ cm.

8. 使用拉普拉斯变换求解以下热量/扩散问题。

$$\frac{\partial u}{\partial t} = \alpha^2\frac{\partial^2 u}{\partial x^2} \qquad x \in (0,\infty)$$

$B.C.1$ $\quad u(x,t) = \sin(t)+1$

$B.C.2$ $\quad \left.\frac{\partial u}{\partial x}\right|_{(x,t)}$ remains bounded for all (x,t)

$I.C.$ $\quad u(x,0) = 0$

Your solution for this problem will not be explicit. You should end up with a convolution integral as your final result.

9. 使用拉普拉斯变换求解下列热/扩散问题。

$$\frac{\partial^2 u}{\partial t^2} = c^2\frac{\partial^2 u}{\partial x^2} \qquad x \in (0,\infty)$$

$B.C.1$ $\quad u(0,t) = \sin(t)$

$B.C.2$ $\quad \left.\frac{\partial u}{\partial x}\right|_{(x,t)}$ remains bounded for all (x,t)

$I.C.$ $\quad u(x,0) = 0$

绘制 $\alpha^2 = 1$ 的解以及时间 $t = 1,\ 5,\ 10$。

“要对抗指数函数，用其他指数函数武装自己似乎合情合理。”摘自《深度架构的表达能力》，作者为 Bengio 和 Delalleau（2011）

10

前馈神经网络入门：一种解析方法

10.1 引言

人工神经网络（人工神经网络），或简称神经网络，已成为应用数学许多领域中经常使用的工具。对人工神经网络而言，一个简明但并非完全通用的定义是：人工神经网络由函数复合的加权和构成。这类网络通常用于以下两种任务之一：（1）将函数拟合到一组示例数据（回归）；或（2）将一组示例数据划分为不同类别（分类），其中类别的边界是网络预测的函数。无论用于回归还是分类，神经网络的基本思想都相同。为了让本入门介绍保持简明，本文主要关注如何通过前馈神经网络生成近似函数来拟合数据集。从现在起，我们假设所讨论的人工神经网络旨在生成近似函数以拟合数据集；不过，我们也应记住，本文讨论的人工神经网络基础知识同样适用于数据分类。深度学习的目标是重建输入与输出之间的关系。我们假设存在一个未知函数 $\hat{y}$，它能够近似已知的数据实例（这些实例将一组拟议的自变量与一组观测值（自变量）相关联）。每个这样的函数 $\hat{y}$都包含若干可调系数。为了进行这样的近似，我们提出一个度量指标，以衡量 $\hat{y}$ 与观测数据之间的误差，然后尝试优化可调系数，使误差尽可能小。利用数据集训练网络后，应得到一个深度神经网络 $\hat{y}$，按照所选度量指标，该网络的结果应接近观测数据。

For now it is enough to think of ANNs as a hierarchical network of compositions of transformation functions. In their simplest form (the only form we will consider here), the networks are arranged in a hierarchy called, suggestively, *layers*. Within each layer, there are one or more *nodes*. A node represents a linear or nonlinear operation; we will describe these operations in more detail later. A fully connected feedforward neural network (FNN) is one where each node within a layer receives input from all of the

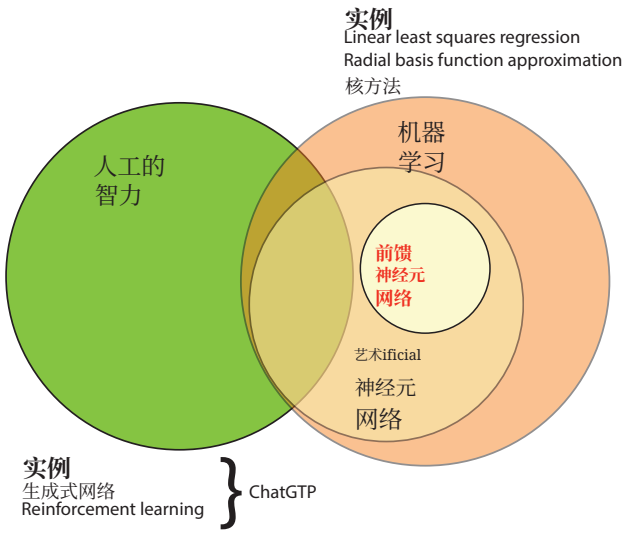


图 10.1 机器学习不一定是人工智能的子集。例如，很少有人会认为线性最小二乘拟合方法代表了“人工智能”。然而，它可以用一个简单的前馈网络来表示，因此是机器学习 方法的子集。

nodes the previous layer, and outputs information to all of the nodes in the subsequent layer. In this material of this chapter, our focus will be on FNNs exclusively.

前馈神经网络的深度 与输入和输出之间函数复合的层数 有关。过去，人们曾对网络深度的作用进行过大量讨论。如今人们已经认识到，随着网络宽度和深度的增加，向网络中添加层会大致以指数级提升其表达能力（本质上，即函数所要估计的自变量被划分成多少个区域）（Bengio 和 Delalleau，2011）。前馈神经网络不断加深，对于开发能够近似数千至数百万（甚至更多）个自变量行为的神经网络至关重要（截至 2023年，GTP-3 聊天机器人使用了超过 1750 亿个自变量）。

In computational sciences (and other disciplines) there is a concept known as the *curse of dimensionality* (Bellman, 1957) that relates to the problem of describing high-dimensional data using naive methods. In short, this concept recognizes that as the number of independent and dependent variables (the dimension) increases, the volume of the space needed to represent them increases. This volume expansion can occur so rapidly with increasing dimension that the available data become sparse (i.e., if one represented the data in a high-dimensional matrix in the independent variables, nearly all the entries would be zero). Thus, in order to obtain a reliable result, the amount of data needed would grow exponentially with the dimensionality. Among the properties that make neural networks useful, and in particular deep neural networks with many layers, is their ability to overcome the curse of dimensionality through function compositions.

10.2 Terminology

In the study of FNNs (and ANNs in general), terminology is an important component. It is probably not an overstatement to say that the differences in terminology between computational sciences and applied mathematics has hampered understanding of what FNNs are (in a mathematical sense), and how they operate. In the glossary of terms below, the terminology established is intended to be both clear and consistent with existing practice. The definitions presented are specific to fully-connected feedforward networks (sometimes

called multilayer perceptron networks). When it comes to the presentation of FNNs in graph form, there is not currently a single unified approach. In the definitions below, a particular effort has been made to present a systematic and reproducible method to both form and interpret the graph representations of feedforward networks. The glossary following defines many of the terms needed to describe FNNs. The glossary is organized hierarchically rather than alphabetically because this helps organizes the information in a more practical way.

• **人工智能**。对于人工智能一词，目前并没有公认的定义。我们不必试图梳理迄今所有可能的定义和观点；就本文而言，可以将其定义为：开发能够执行通常需要人类智能才能完成的任务的计算机系统。实例包括图像识别、针对语音输入生成合理且逻辑正确的响应，以及通过适当训练生成此前不存在的新数据（例如，根据文字提示生成图像）。虽然这个定义有些循环论证（每当新的计算算法实现这一目标后，“通常需要人类智能”这一概念就会被重新修订！），但它仍然很有用。

• **机器学习**。与人工智能一词类似，机器学习 也没有单一且明确无歧义的定义。在本文中，这个术语用于指任何能够利用对当前误差的测度，自动生成误差更小的新模型的算法（无论该算法是否被视为人工智能）。在人工神经网络的应用中，学习循环 会采用某种形式的梯度下降 （下文定义）来优化代价函数残差，并通过反向传播 更新权重，从而表示出更好的近似。虽然机器学习常被视为子集 ，属于人工智能，但本文不采用这种区分。例如，线性最小二乘拟合可以表示为一种具有已知算法的神经网络，并且可以成功用于以大量数据点作为输入实例的高维数据集。然而，很少有人会认为这类方法代表了“人工智能”。

- **Artificial neural network**. An artificial neural network (FNN) is a hierarchical network constructed from weighted compositions of functions. The networks are arranged in a hierarchy called *layers*. Each layer contains a number of *nodes*, where each node represents a linear or nonlinear function. The vector of independent variables for the network is known as the *input* or *feature set*. The result of the network is a vector or scalar (or other mathematical object) containing the dependent variable, and is called the *output* or *target*. The input data is processed through the network by making a sequence of transformations of the input. These transformations are done layer-wise, with the output from one layer forming the input to the next. The input to any node in an internal layer of the network is a weighted sum of the output from the nodes from the previous layer. The functions used to transform the variables are frequently called *activation functions*, but for the problem of function fitting, might better be thought of as *basis* functions.

• **（全连接）前馈神经网络（FNN）**（也称为**多层感知机网络或 MLP**）。这是一种人工神经网络，其层与层之间的连接只有一个方向：从输入指向输出。假设某一层中的每个节点都会接收来自前一层所有 节点的输入，并将信息输出到后一层所有 节点。前馈结构有别于包含反馈回路或并行回路的人工神经网络，或.....。此定义也不同于用于优化神经网络权重的任何算法过程。有时，简单的前馈网络也称为多层感知器，这一术语源自早期对线性网络（即变换为线性的网络）的研究。不过，如今多层感知机一词也涵盖采用非线性变换的简单前馈网-

络，因此在现代用法中，这两个术语可以互换。尽管前馈神经网络（多层感知器）常被归入人工智能，但不这样归类也有充分理由。虽然前馈神经网络采用非线性激活函数，但它们并不使用非线性网络结构（例如回路或强化反馈）来求解。

• **深度神经网络**。关于深度神经网络的定义，存在一些不同意见。就本研究而言，任何包含至少两个隐藏层（定义见下文）且能够对输入数据进行非线性变换的神经网络，都属于深度神经网络（参见 Chui 等人，1994）。之所以选择两个层，是因为非线性本身并不足够；原则上，即使早期的二元感知机模型也是非线性的，因为它们包含阈值处理（阶跃）激活函数。向神经网络中增加层（这会使系统规模呈线性增长）的一个关键优势在于，激活函数的非线性能够使问题的解变得复杂得多，且复杂程度呈指数级增长。若要使用单隐藏层网络实现同等效果，则需要使其宽度呈指数级增长。这是一个重要概念。深度神经网络可以被认为具有与浅层神经网络相同的表达能力，但所需的总节点数量要少得多。例如，已有研究表明，对于修正线性单元（ReLU）激活函数，单隐藏层网络若要达到与深度网络相同的表达能力（即逼近函数的能力），所需的节点数量将呈指数级增加（Bengio 和 Delalleau，2011；Montufar等人，2014年）。

• **浅层神经网络**。与许多关于神经网络的术语一样，浅层一词似乎没有统一的含义。在本文中，浅层指包含零个或一个隐藏层的人工神经网络。已知含有单个隐藏层的网络具有良好的连续函数再现能力（例如，可用偏差平方或 $L2$ 度量指标来衡量）。然而，与具有相同表达能力的深度神经网络（其所需单元数仅随深度呈多项式增长）相比，这类网络所需的单元数可能会呈指数级增长。

• **网络架构**。网络架构是对组合成网络的运算进行的规定，通常以图形方式表示。对于全连接前馈神经网络，网络架构（L，$\mathbf{p}$）由一个正整数 L 和向量$\mathbf{p} = (p1, \ldots, pL)$完全确定；前者表示隐藏层的数量，后者表示每个隐藏层的宽度。

• **节点**（有时也称为**单元**或**神经元**；在图论中有时称为**顶点**）。在上述前馈神经网络的定义中，引入了节点这一概念。第k 层的节点 $+1$ 就是一个算子（计算科学文献中通常使用函数一词），它对第k 层的输入进行变换（见图 10.2）。这些变换可以是线性或非线性的。然而，要使前馈神经网络在一般情况下能够成功地进行函数估计（或等价地，对数据进行分类），就需要在某些节点层实施非线性变换。在本文中，**网络中的每个节点都表示某种算子**，该算子与输入结合后会生成一个激活函数。节点可以是输入节点（位于自变量进入网络的边界处，以正方形表示）、输出节点（位于因变量预测值形成的边界处，以正方形表示），也可以是内部节点（以圆形表示）；根据上下文即可理解这些定义。

• **Node number**. In this work, *structure preserving* notation indicates a notation in which the network layer structure is embedded in the parameter and function representations. For this notation, each *node*

number is denoted by a pair indicating the layer, and the node number within that layer. Thus, $(1,2)$ indicates "layer 1, node number 2". In small networks, we will use *simplified notation*, where each node is given a unique integer. While this notation does not indicate connectivity well, it does make the presentation simpler, which is ideal for small networks.

- **Source node**. Nodes are connected by links (defined below). A node in layer k with a link whose output is directed to node n in layer $k+1$ is called a *source node* for node n. The node n in layer $k+1$ may have as many source nodes as there are total nodes in layer k.

- **Destination node**. A node in layer $k+1$ with a link whose input is directed from a source node in layer k.

- **Bias node**. Homogeneous functions are such that when each of their independent variables is zero, the returned dependent variable is also zero. To allow for shifts in the functions defined for an FNN, one adds *bias nodes* that represent the addition of a constant (bias). Thus the magnitude of the shift is determined by its weight, b. The weight variable is uniquely identified by a subscript using the convention described under the definition of *link*. See Fig. 10.2 for an example of the implementation a bias term in graph form.

• **权重**。如上所述，网络中的每条连接（或边）都有相应的权重。权重没有统一的符号表示。在本章中，采用保持结构的符号表示时，节点之间的连接以下标表示源节点、以上标表示目标节点；在有助于提高清晰度时，还会使用颜色。因此，$w_{2,2\ 1,2}$ 或 $w_{2,2\ 1,2}$ 表示连接节点 $(1,2)$（第1层的节点2）与节点 $(2,2)$（第2层的节点2）的权重。采用简化符号表示时，为每个节点指定一个唯一的整数索引。因此，节点j 和k 之间的权重可用更简单的符号表示为$w_{j,k}$。

• **连接**（在图论中有时称为**边**）。前馈神经网络中的每个节点都通过一条连接与至少一个其他节点相连。连接是一条有向直线，通常用箭头表示，用来指示信息流的方向。系统中的每个节点至少有一条输入连接（箭头指向该节点）和一条输出连接（箭头从该节点指向外部）。在前馈神经网络中，连接所表示的不仅仅是简单连接性。前馈神经网络中的每条连接都对应一个特定的**权重** w。这种组织方式隐含着一个概念：与第$k+1$层节点关联的输入连接定义了对k层源节点输出的线性组合。这个线性组合是各源节点输出分别乘以对应权重后求和。输入连接和输出连接将用虚线表示，以便与其他连接区分；可以认为它们的权重为 1。**简言之，每条连接都表示一个加权项，该项从源节点出发，终止于目标节点。在每个目标节点处，所有源连接的输出与相应权重构成的线性组合都是**隐含的。注意：在简化符号表示中，网络中的每个节点都被赋予一个唯一的整数标签，因此网络中的每条连接都可以用这两个整数表示。为连接编号时，约定先将源节点列为下标，再列目标节点。因此，$w_{1,2}$表示与一条连接对应的权重，该连接将节点1的输出指向节点2的输入。如果涉及的源节点是偏置节点，则约定使用符号b 代替w，但下标的约定与普通权重相同。

- **Layer**. As mentioned above, and FNN is composed of nodes arranged in layers. An example of a generic FNN is given in Fig. 10.2. The concept of layers is a convenient organizing structure for fully-connected

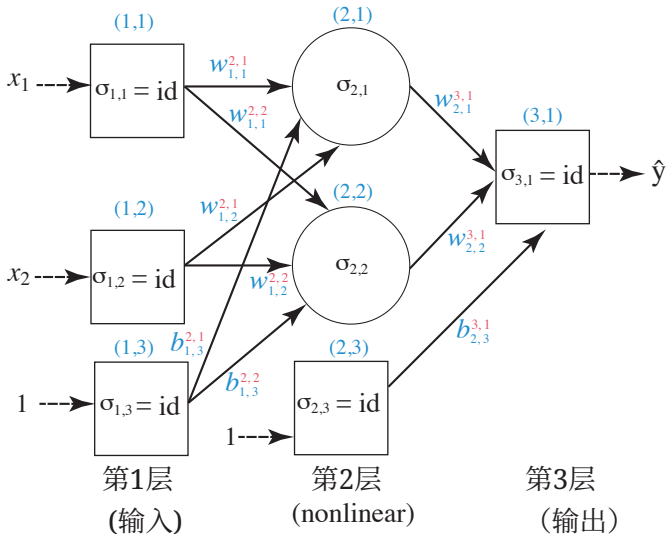


Fig. 10.2 A generic feedforward artificial neural network, illustrated in graph form with structure-preserving notation. In this example, there are two independent (feature) variables, x_1 and x_2, and one dependent (target) variable, y. The example contains three layers, and one of these is a *hidden layer* (i.e., it is neither the first nor the last layer of the network).

简单前馈网络，因为采用分层网络拓扑可以直观地展示节点连接性。对于未完全连接或包含反馈回路的网络，层的概念就不那么重要了，取而代之的是节点之间简单连接性的概念。

- **Hidden layer**. A hidden layer is, technically, any layer that is not an input or output layer. Thus, any network with three or more layers will contain at least one hidden layer. Note that hidden layers are generally nonlinear (although there may be instances of deep networks with many layers that impose a layer of linear transformations for specific purposes). All hidden layers are made up of *interior* nodes.

• **特征**。（也称为**输入**或**自变量**）。机器学习这一学科使用的术语在计算科学中也很常见。在神经网络中，自变量向量有时常被称为特征，以表明其作为描述变量的作用，这些描述变量使人们能够预测复杂函数的结果（注意：此处的“复杂”表示自由度，而非包含虚部）。例如，在离散时间t_i对电阻器的温度N T_i次观测，将生成N 个形如（t_i，T_i）的数对。N 个t_i 值构成数据集的特征。在机器学习中，特征集往往会变得相当庞大。神经网络（结合适当的优化和权重调整算法）能够为包含大量特征的数据集生成优化拟合，这是使用神经网络的主要动机之一。

• **目标**。（也称为**输出**或**因变量**）。因变量的集合。目标通常是一个标量值，由与自变量的某个特定向量相关联的实数指定。例如，在离散时间t_i对电阻器温度进行观测，得到N 个温度T_i，从而形成N 对形如（t_i，T_i）的数据。数值T_i 构成单个自变量（特征）t_i对应的目标值。原则上，目标也可以由向量或张量（矩阵）输出指定。按照惯例，神经网络预测的目标值会在输出变量上方加上扬抑符（帽子）标记，以表示所得函数是一个估计量。

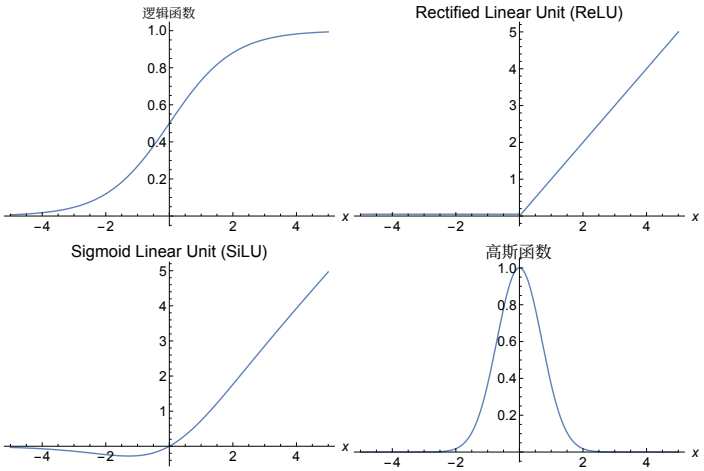


Fig. 10.3 Examples of four activation functions. Activation functions a selected on the basis of network purpose, and the depth of the network. Functions whose derivative approaches zero as $|x| \to \infty$ often do not perform well in deep networks, because the (numerical) propagation of information back through the network is proportional to the derivatives of the function.

• **激活函数**（或**基函数**）。神经网络中的每个节点都表示一个算子，用于变换输入。按照惯例，这些算子称为激活函数。算子和函数之间的技术差异不太可能引起混淆，因此我们通常会遵循这一惯例。这个函数可以是线性的，也可以是非线性的。由于线性函数的复合仍会映射为一个（新的）线性函数，现代神经网络的一个或多个隐藏层中的激活函数具有非线性这一特征至关重要。毫不夸张地说，使用此类非线性激活函数是前馈神经网络能够用于函数逼近的原因之一。虽然激活函数这一术语很常见，但在某些应用中，将这些函数视为基函数或许更合适；通过对基函数进行缩放、平移和线性组合，可以导出新的函数。

图 10.3 给出了若干激活函数实例。激活函数的选择取决于具体应用（例如，结果是否需要具有连续导数；网络是否非常深）。尽管修正线性单元（*relu*）激活函数的导数并不连续，但它已成为深度神经网络中应用最广泛的函数。这些激活函数是本章讨论的重点，后续材料还将对其进行更详细的介绍。

• **损失函数**。为了优化网络，需要使用某种可微的目标度量指标（回归中的拟合优度；分类中的适当数据分离度）。从解析角度看，有限维向量空间中所有拟合优度度量指标（误差的绝对值，即$L1$ 范数；或误差平方和，即$L2$ 范数）都是等价的。然而在实际应用中，采用某种误差度量而非另一种，可能会简化计算。出于本文的目的，我们将专门采用误差平方和作为范数。我们的一些实例涉及无限维（函数）向量空间，其中不同范数不一定等价。这不会造成值得注意的问题。

- **Identity operator**. There is occasionally use for an identity operator as one of the possible operators (functions), σ, that can be used at a node. For our purposes, the identity operator is defined by

$$\mathrm{id}(x) = x$$

That is, the operator leaves the input unchanged.

• **优化**。通过此过程，可以确定一组参数（视为自变量），使由指定误差度量指标构成的目标函数取得最小值。误差度量指标通常是模型与数据之间的绝对差（$L1$ 度量指标），或最小二乘差（$L2$ 度量指标），不过也会使用其他指标。优化可以通过解析或数值方法完成。对于有限维问题（例如，最小化一个基于有限数量数据的函数），原则上所有误差度量指标都应得出等效结果。实际上，由于数值计算方案中采用的近似方法，有时某种误差度量指标可能比另一种具有更好的性质。有时会通过指定“arg min”算子来表示优化过程。因此，可以写作 $\arg\min[f(x; w1, w2)]$，以求得使指定误差最小化的$w1$和$w2$的值。

• **多分辨率**。多分辨率一词在小波理论广为流行时开始普遍使用，并一度成为小波的同义词。不过，此后它的含义变得更加宽泛。在本章中，多分辨率 仅指可以对函数进行局部细化，而不影响函数的其余部分。这是分析人工神经网络时反复出现的主题，因此，多分辨率的这一广义定义很有用。

- **Backpropagation**. Backpropagation is the *numerical scheme* used to adjust the weights when optimizing an ANN in response to the current state of the error in the solution. This method relies on an application of the chain rule to assign adjustments to the weights. Backpropagation methods were discovered in the early 1970s, but it took some time for the algorithms to be assimilated by the machine learning community. Because this is a scheme used for solving ANNs numerically, it will not be a component of the material in this chapter. However, the scheme is so central to the development and solution methods for ANNs, it is necessary to be aware of the scheme and its purpose.

• **监督式机器学习（在神经网络语境下）**。假设我们有N 个数据实例，形式为（$\mathbf{x}_i$，$\mathbf{y}_i$），其中$\mathbf{x}_i$是一个向量，包含与观测值y_i相关的自变量。利用自变量和因变量已知的数据集，对前馈神经网络进行参数化以执行回归或分类，这一问题称为监督学习。这里，“监督”一词表示示例数据明确标注了相应的自变量。大型神经网络通常无法（或者说，实际中不可能）仅用解析方法求解，而是通过计算求解。通过计算而非解析方法求解简单前馈神经网络时，最主要的显著区别在于所采用的损失函数最小化方法。在前馈神经网络的机器学习中，优化通常采用若干种梯度下降方法之一，并结合反向传播方法来准确更新网络中的权重。由于此处仅介绍可解析求解的网络，因此不再进一步讨论梯度下降或反向传播方法。

10.3 前馈神经网络图形表示的解读

在进一步讨论人工神经网络之前，先概述其图形化描述方式会很有帮助；这样我们便掌握了恰当的“结构术语”，从而更好地理解实例。我们知道，神经网络的基本运算是接收输入，使用线性或非线性算子对输入进行变换，然后生成输出。这个过程在节点处进行，由按层排列的一系列非线性运算构成（如上文所定义）。在每一层之间，前一层节点的输出通过线性组合进行变换。为了使这一过程更加直观，这些运算通常以连通图的形式呈现。尽管相关文献中的表述似乎并不统一，本文仍采用以下规则来描述人工神经网络的图形结构。

1. 虚线表示输入或输出。相关箭头的方向可以帮助判断它表示的是输入还是输出。2. 在此处采用的符号表示中，所有节点都包含激活函数，并用符号 σ 表示，下标表示其节点编号。节点编号由有序对 (i, j) 指定，其中 i 表示层索引，j 表示该层内的节点索引。通常只使用一种形式的激活函数，此时无需为激活函数添加索引。激活函数可以是线性或非线性（具体形式将在后文讨论）。对于输入节点和输出节点这一特殊情况，算子为恒等算子 $\sigma(\cdot) = \mathrm{id}(\cdot)$，它将输入值原样作为输出。3. 正方形表示输入节点或输出节点。这些节点也称为边界节点，表示它们负责将外部输入传入网络，或将输出传出网络。至少始终存在一个输入层和一个输出层；输出层可能不会对数据执行任何非线性变换。4. 图中的圆形表示带有相应激活函数的节点；激活函数可以是线性的，也可以是非线性的。这些算子将输入其中的加权输入进行变换，生成一个输出值。所有用圆形表示的节点都是内部节点，说明它们不是图的输入或输出边界。5. 节点之间的实线（带有表示方向的箭头）表示加权输出。线尾节点的输出乘以一个权重值 $w^{k}_{,l}i, j$，该权重值标注在线上方。在这种符号表示中，i, j 表示第 $k-1$ 层的一个节点，而 k, l 表示第 k 层的一个节点。这称为保结构记法。6. 每个内部节点都会有一个或多个指向它的加权输入。每个内部节点的输入都是指向该节点的加权输入之和。7. 偏置项是可以出现在网络任意层的输入。这类节点以 1 为输入，通过恒等算子变换该输入，并输出数字“1”。然后，该输出乘以一个权重值 $w^{k}_{,l}i, j$，该权重值标注在线上方。通常会向每个内部节点添加偏置项，其编号方式与其他节点相同。因此，$b^{k}_{,l}$ i, j 表示节点 (i, j) 与节点 (k, l) 之间连接的偏置项。由于偏置节点可能使图变得复杂，因此通常只需注明每个内部节点都包含一个偏置项，而不必在图中明确表示它们。

To complete the connection between the graph and the mathematical representations, we consider the example of the network shown in Fig. 10.2, and repeated here as Fig. 10.4 for convenience. The output from any layer will is usually indicated by $\mathcal{N}$ (for “network”), subscripted by its layer index. Starting with the

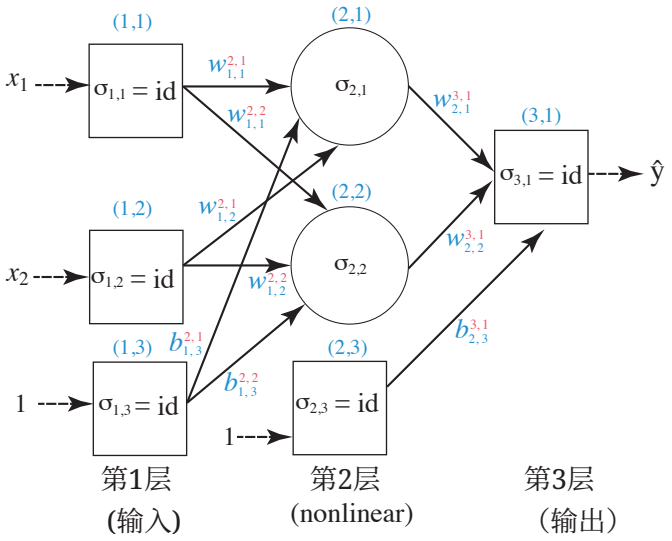


Fig. 10.4 The feedforward artificial neural network from Fig. 10.2, repeated here for convenience. This network is specified with structure-preserving notation.

first layer, we note that the linear composition rule indicated by the arrows connecting the first and second layer would provide the following *output* from layer 1

$$\begin{aligned}\mathscr{N}_1(x_1,x_2,1) &= \begin{pmatrix} w^{2,1}_{1,1} & w^{2,1}_{1,2} \\ w^{2,2}_{1,1} & w^{2,2}_{1,2} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} b^{2,1}_{1,3} \\ b^{2,2}_{1,3} \end{pmatrix} \\ &= \begin{pmatrix} w^{2,1}_{1,1}x_1 + w^{2,1}_{1,2}x_2 + b^{2,1}_{1,3} \\ w^{2,2}_{1,1}x_1 + w^{2,2}_{1,2}x_2 + b^{2,2}_{1,3} \end{pmatrix} \end{aligned} \tag{10.1}$$

The second layer takes the vector $\mathscr{N}_1$ as input, and forms the following output (as indicated by the arrows between layers 2 and 3)

$$\mathscr{N}_2(x_1,x_2,1) = \begin{pmatrix} \sigma_{2,1}\left(w^{2,1}_{1,1}x_1 + w^{2,1}_{1,2}x_2 + b^{2,1}_{1,3}\right) \\ \sigma_{2,2}\left(w^{2,2}_{1,1}x_1 + w^{2,2}_{1,2}x_2 + b^{2,2}_{1,3}\right) \end{pmatrix} \cdot \begin{pmatrix} w^{3,1}_{2,1} \\ w^{3,1}_{2,2} \end{pmatrix} + b^{3,1}_{2,3} \tag{10.2}$$

Note that this output is now scalar valued. Again, this is very typical of ANNs. A vector-valued input is transformed by a sequence of compositional vector operations. In the last step, a contraction reduces the dimensionality of the output (usually) to a scalar, although other output forms (e.g., vector or tensor outputs) are possible.

Finally, the third layer operates on the output $\mathscr{N}_2$ with the identity. The result is

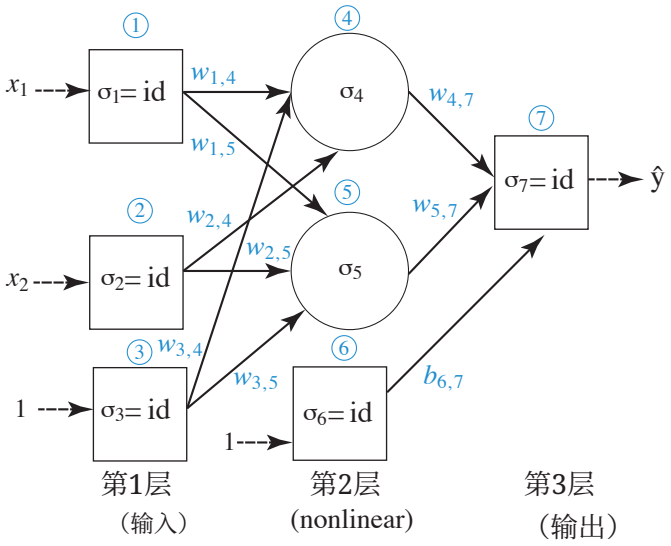


Fig. 10.5 The feedforward artificial neural network from Fig. 10.2, repeated here for convenience. This network is specified with simplified notation that does not preserve structure.

$$\begin{aligned}\hat{y}(x_1,x_2,1) &= \begin{pmatrix}\sigma_{2,1}\left(w^{2,1}_{1,1}x_1 + w^{2,1}_{1,2}x_2 + b^{2,1}_{1,3}\right)\\ \sigma_{2,2}\left(w^{2,2}_{1,1}x_1 + w^{2,2}_{1,2}x_2 + b^{2,2}_{1,3}\right)\end{pmatrix} \cdot \begin{pmatrix}w^{3,1}_{2,1}\\ w^{3,1}_{2,2}\end{pmatrix} + b^{3,1}_{2,3}\\ &= w^{3,1}_{2,1}\sigma_{2,1}\left(w^{2,1}_{1,1}x_1 + w^{2,1}_{1,2}x_2 + b^{2,1}_{1,3}\right) + w^{3,1}_{2,2}\sigma_{2,2}\left(w^{2,2}_{1,1}x_1 + w^{2,2}_{1,2}x_2 + b^{2,2}_{1,3}\right) + b^{3,1}_{2,3}\end{aligned} \tag{10.3}$$

While in this example the application of the identity in layer 3 is a bit tautological, it does serve a specific purpose. Viewed this way, the nodes of *every* layer conduct *some* operation. It is not uncommon to use a filtering function as the final step in the network, in which case the final activation function would not be the identity.

Even a small network can rapidly lead to long and complex compositional functions. The utility of network representations begins to become quite apparent for networks of any substantial size.

10.3.1 More on Notation

上述符号表示很有用，因为网络结构嵌入在符号表示中。因此，可以从图生成网络的数学形式；也可以轻松地从数学表征生成图，因为图的结构在符号表示中得到了完整表征。

While this kind of notation can be useful (e.g., in coding, or with networks that are not strictly feedforward), it is not ideal for examining small networks such as we will use. The alternative is to abandon the labeling of layers. This means that each node receives a single integer index, and each weight requires only two integers to indicate its connection. The downside is that the graph must accompany the mathematical representation or there can be a lack of clarity about the location of nodes (although with strictly feedforward networks this is less of a problem).

An example of the *simplified* notation is given in Fig. 10.5. For small networks, this notation is substantially easier to understand, and it will be the default option for the remainder of this chapter. For comparison, the analysis above can be repeated. The result is a bit easier to read, and is probably better for initial forays into understanding ANNs.

$$\begin{aligned}\hat{y}(x_1,x_2,1) &= \begin{pmatrix}\sigma_4\left(w_{1,4}x_1+w_{2,4}x_2+b_{3,4}\right)\\ \sigma_5\left(w_{1,5}x_1+w_{2,5}x_2+b_{3,5}\right)\end{pmatrix}\cdot\begin{pmatrix}w_{4,7}\\ w_{5,7}\end{pmatrix}+b_{6,7}\\ &= w_{4,7}\sigma_4\left(w_{1,4}x_1+w_{2,4}x_2+b_{3,4}\right)+w_{5,7}\sigma_5\left(w_{1,5}x_1+w_{2,5}x_2+b_{3,5}\right)+b_{6,7}\end{aligned} \tag{10.4}$$

10.4 人工神经网络的历史与目的

现代人工神经网络的发展提供了一个有趣的历史案例，展示了科学进步实际上是如何展开的。尽管类似人工神经网络的网络早已在数学文献中出现，Warren McCulloch 和 Walter Pitts（McCullough 和 Pitts，1943年）发表了被认为是首篇将认知概念与计算联系起来的论文。他们的模型本质上是一个二元模型：二进制输入的总和经过阈值（或赫维赛德阶跃）函数筛选；二元输出为 1 表示总和超过阈值（即 1 表示“真”），而 0 表示总和低于阈值（即 0 表示“假”）。Piccinini（2004）的论文对他们的工作进行了历史梳理和分析。尽管他们的工作如今已不再被视为认知的原型，但它仍开启了仿生计算方法的时代。McCulloch 和 Pitts 提出的理论代表了人们对后来发展为现代神经网络的领域的首次探索。然而，他们的模型并不包含学习步骤。换句话说，模型只是被设定在某个状态（即阈值被设为某个指定值），随后网络便对输入作出响应。

20世纪50年代末，一位名叫 Fran Rosenblatt 的心理学家通过提出一种带有反馈步骤的网络，重新承担起机器学习的使命；该反馈步骤根据输出误差调整权重。计算的基本单元被称为 感知机。经典感知机的图形表示见图 10.7。简而言之，感知机是一种二分类器，它接收输入值向量 $\mathbf{x}=(x_1, x2,\ldots,1)$，并构造一个线性组合。随后，该组合经过 Heaviside 函数处理，充当阈值单元。输出为 1 或 0，具体取决于输入函数。关键在于，当解存在时，有一种简单的算法能够以系统化的方式调整权重，使感知机收敛到该解。遗憾的是，如果分类是不可能的（具体而言，如果待分类的集合无法使用线性函数进行线性可分），那么这种学习方法将永远不会收敛。感知机的局限性在于：（1）它只能学习数据中的线性关系；（2）学习算法依赖于二进制误差。这要求分类必须没有误差；也就是说，数据必须能够由线性函数完全分离。即使只有一个阻碍线性分离的点，也足以使该方法永远不会收敛。

Rosenblatt 的工作继续推进，并成功构建了首台仿生计算机。Mark I 感知机是一种机电设备，由 Frank Rosenblatt、Charles Wightman 等人在 1957年和1958年间研制。Mark I 感知机的主要用途是对符号进行模式识别。权重编码在电流电位器中，学习过程中通过电动机自动调节电位器来更新权重。图10.6 展示了 Resenblatt 与 Mark I 的合影，以及该装置的网络图。

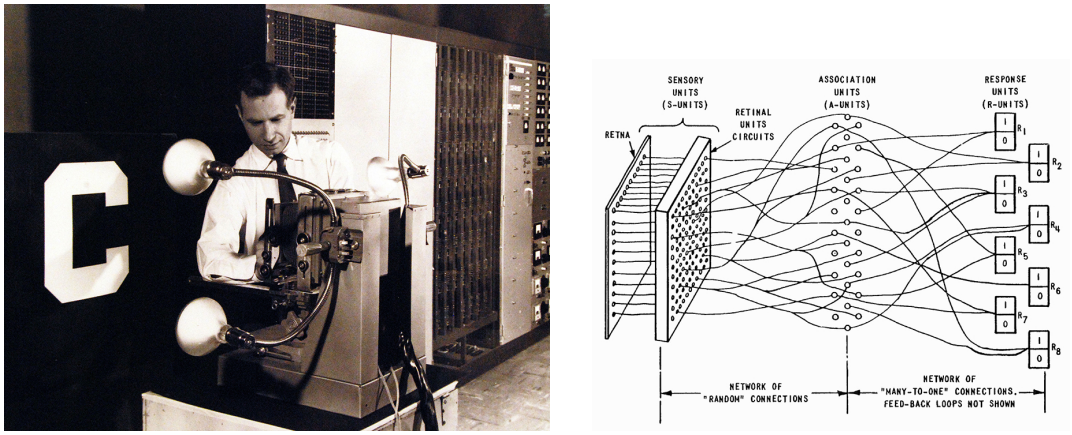


Fig. 10.6: (Left). Roseanblatt working with the Mark I Perceptron, circa 1960. (Right) A schematic of network comprising the Mark I Perceptron from the operator's manual (adapted from (Hay et al., 1960)). Photograph by the National Museum of the U.S. Navy, licensed under the Public Domain Mark 1.0.

学习算法理论的下一个重要进展由 Bernard Widrow 及其博士生 Ted Hoff 于 1960 年在斯坦福大学提出。他们开发了一种采用改进学习规则的网络；该方法被称为自适应线性元素（ADALINE）规则或 Widrow-Hoff 规则。这种新方法的主要区别在于反馈误差的定义方式，以及如何利用该误差调整两个感知机的权重和偏置。Rosenblatt 使用分类误差（即 1 或 0）作为学习度量指标，而 ADALINE 引入了损失函数（或目标函数）的概念。这使得非整数误差能够被计算出来，而这是确保权重调整幅度与误差度量指标成正比所必需的。Widrow 和 Hoff 还开发了一种早期形式的梯度下降算法，称为 *delta* 学习规则。有关该算法的详细信息可在其他资料中找到；然而，其重要特点在于，它是一种简单的算法，能够使权重调整幅度与非整数误差度量指标成正比。

尽管这一早期模型提出了许多不错的想法，但一系列因素恰好交织在一起，导致人们对继续研究人工神经网络的兴趣大幅下降。首先，媒体和相关科学家对最初的感知机大肆吹捧。尽管早期取得了一些成功，但有关“思考机器”新世界的种种夸张承诺终究未能实现。其次，或许部分是为了回应围绕感知机的热潮，麻省理工学院的 Marvin Minsky 和 Seymour Papert 发布了一份未发表的技术手稿（后来经修订，于 Minsky 和 Papert（1969，1988）合著的书中出版），说明这些线性单层感知机网络无法学习简单的非线性二元函数；“异或”函数便是常见的例子。Minsky 和 Papert 的本意确实推动了这一领域的发展，但最终也导致人们对人工神经网络整体的兴趣降温。此外，Minsky 和 Papert 对多层感知机的能力并不确定，并预言它们不太可能解决这一问题。与此同时，Rosenblatt 声称，通过增加多个层，可以显著提升感知机的泛化能力。然而，由于当时尚无已知的训练方法，这些技术难题使多层感知机无法得到直接应用。支持与贬低单层感知机（以及认为多层感知机毫无用处这一观点）的人之间的论战，清晰地分成了两派。这是一个例子，说明科学研究有时会因某些不太科学的原因而出现颇为反常的偏离，

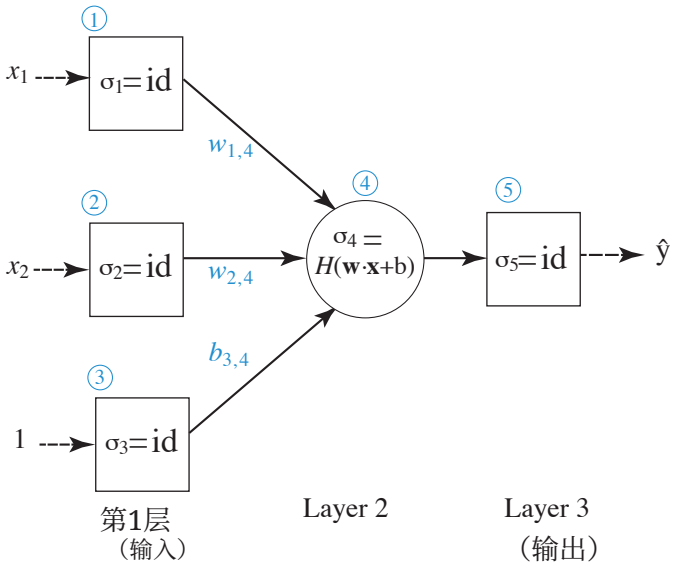


Fig. 10.7 The classical perceptron model of Rosenblatt.

这种情况可能会在科学领域发生。尽管科学研究中此类反常路径通常并不常见（也不可能无限期持续），但了解其发生的原因仍有助于防止它们在未来重演。Olazaran（1996）在一篇引人入胜的分析中，对这一事件作了相当详尽的考察。

从 20 世纪 70 年代到 80 年代中期，只有少数专注于此的研究人员在开展人工神经网络研究。然而，在此期间，人们提出了多种形式的反向传播算法；这一发明似乎由多位研究人员几乎同时完成（Linnainmaa，1970；Werbos，1974；LeCun，1985；Rumelhart 等，1986）。该算法是人工神经网络现代发展的关键一步。此外，尽管 Minsky 和 Papert 提出了相关担忧，一些研究人员仍继续探索使用多个隐藏层的神经网络。

20 世纪 80 年代初，Hopfield（1982、1984）发表了数篇论文，推动人工神经网络研究重获生机，并开启了后来大致被称为深度学习第二次革命的时期。最终，作为人工神经网络研究再度兴起的最后一股推动力，James McClelland 和 David Rumelhart 于 1986 年和 1987 年出版了一套教科书，总结了当时已知的大部分内容（McClelland 等，1986、1987）。特别是，这些新教材刊载了反向传播学习算法，使这一算法广为人知。这些进展促成了前馈神经网络（也称多层感知器，MLP）的诞生，并全面推动了人工神经网络研究的复兴。

10.5 Some Principles of Feedforward Neural Networks

Before delving into the construction and solution for ANNs, it is useful to outline why one should study them at all. While ANNs have been researched for many years, it has been only in more contemporary times that the computational algorithms and hardware that allow the solution to deep neural networks to become routine. Recall, deep neural networks (DNNs) are neural networks that have more than one hidden layer. The reason that the number of hidden layers is important is because the methods to optimize networks with more

than one hidden layer is a relatively recent innovation. However, once research on DNNs became practical, it was discovered that they were able to compute solutions *substantially* more efficiently than previously known methods.

人工神经网络通常如此擅长拟合或分类高维数据，其原因正是一个受到高度关注的研究领域；研究既从构建描述性理论的角度展开，也包括旨在提出高效算法的更多实证工作。尽管人们常常将人工神经网络描述为黑箱，但这种描述并不完全准确。尽管目前一般还不存在关于人工神经网络的完备理论，但人们已在为许多特定架构发展理论方面取得了实质性进展。关于这一主题的文献浩如烟海，所用的数学方法涉及函数论、拓扑学、经典逼近理论和统计学中的非常具体的技术方法。人们采用的方法种类繁多，或许正是这一主题引人入胜的最佳证明之一。尽管关于人工神经网络仍有许多有待了解之处，但过去几年里，人们在理解描述其行为的理论方面已取得重大进展，而且几乎没有证据表明这种进展不会持续下去。

以下内容概述了截至2023年，从相关文献中总结出的一般人工神经网络（ANNs），尤其是深度神经网络（DNNs）的一些基本原则。这里的“原则”一词并不意味着“公理”，而是采用该词的通俗含义（即，这些原则是对当前知识状况的总结性观测）。以下原则不分先后顺序，既基于迄今为止关于人工神经网络的定性知识，也基于定量知识。

1. **人工神经网络的效果好得令人难以置信**。文献中常能看到关于深度神经网络“效果好得令人难以置信”的评论。这里使用“令人难以置信”一词，是为了强调，直到最近，人们才知道存在与人工神经网络同样高效的方法。例如，利用分段线性样条近似非线性函数的理论已经相当成熟，但效率相对较低。人们直到最近才开始真正理解人工神经网络为何能够做到这一点。迄今为止，初步解释似乎涉及：(a) 激活函数的复合使用方式（在深度神经网络中）能够缩小参数空间；(b) 此类复合结构（以及可以推测，它们所表征的数据）具有自相似性；以及 (c) 存在能够捕捉数据内部“隐藏”结构的降维表示（Recanatesi等人，2019）。由于大多数数据集都包含某种结构（尽管可能尚未被识别），许多研究人员正在探索这样一种观点：复合函数特别适合对此类数据进行建模。实际结果是，如果某个基函数集适合表征某些类型的数据结构，那么它或许只需较少的可调参数就能完成表征。这里可以打个比方：考虑在$x \in [0,1]$上表征函数$f(x) =$ 正弦(πx)。泰勒级数可以用一组多项式函数来近似$f(x)$，但要将近似误差降低到理想水平，可能需要很多项。然而，傅里叶级数仅用一项就能精确收敛到这个函数！Daubechies 等人（2022）的一篇论文概述了当前关于深度神经网络为何如此有效的一些看法。

2. **存在通用逼近定理**。在20世纪90年代初，许多研究人员开始探索可由神经网络表征的函数类型这一问题。这个概念有时称为神经网络的表达能力。早期的大多数研究（例如 Cybenko（1989年））都采用经典逼近理论（在这里，我们可以将其视为一种更一般的理论，并将傅里叶级数作为其特例），来描述人工神经网络可能具备的能力。其他研究人员指出，Schwartz 在20世纪40年代的研究为理解表征问题提供了依据。近期的研究主要聚焦于两个主题：(1) 针对实际使用的函数类型（例如 ReLU 函数）得出结果；(2) 估算近似的准确度与网络规模之间的关系；以及 (3) 考虑独特的泛函结构

由深度网络诱导（例如，Petersen 和 Voigtlaender（2018））。这些最新结果表明，ReLU 前馈网络不仅能够近似（在实数集上稠密）勒贝格可积函数（采用平方误差积分（L_2）度量指标），而且与浅层网络相比，使用深度网络时收敛速率更快。不过，值得注意的是，人工神经网络的近似定理自 20 世纪 80 年代以来就已存在，并不取决于网络的深度。虽然近似非常一般的函数的能力显然很重要（如果人工神经网络无法近似广泛类别的函数，它们就毫无用处！），但它们不同寻常的有效性并非必然与其近似函数的能力相关。

3. **涉及多分辨率表示**。学习傅里叶级数时，我们了解到，傅里叶级数是在整个定义域上全局表征的一种方式。也就是说，级数中每个带权项都会以周期性的方式对解作出贡献，而且这种方式在整个定义域内保持一致。另一方面，小波是具有尖峰状非周期分量的函数（有时甚至是紧支撑函数），可以有选择地突出局部特征（即定义域的某个子集中的特征）。之后我们会看到一些实例，其中某些特定的激活函数可以同时表征函数的局部特征（例如峰值）和全局特征（例如容易用傅里叶级数表示的分量）。因此，从原理上说，人工神经网络似乎能够表现出多分辨率行为。这里的多分辨率仅指：可以对函数进行局部细化，而不影响函数的其余部分。文献中早在很久以前就已认识到这一点（Chui 等人，1994），但直到最近，它才成为研究人工神经网络的学界广泛关注的话题。人们或许会预期（或希望？），优化过程会根据哪种表征效率最高，选择能够表示相应局部或全局表征的激活函数。目前，人们尚不清楚在应用中，局部表征与全局表征是如何在内部解析的（Raghu 等人，2021）。此外，Daubechies 等人（2022）的论文总结了目前对人工神经网络所体现的多分辨率行为的大部分认识。

4. **优化过程中无须找到全局最小值**。通常，无法确定一个包含众多节点和层的复杂人工神经网络是否已针对某个指定损失函数进行优化，以找到全局最小值；通常，这类问题并非凸问题，因此可能存在许多局部极小值。训练方法转而通过对数据进行统计抽样来优化神经网络。人们希望这些方法能够找到某种优化结果（可能是局部最优），并在损失可接受的前提下提供足够好的近似。在这种情况下，找到的最优解是否为全局最优并不重要；相反，只要事先规定近似误差可接受即可。近期研究（Haeffele 和 Vidal，2017）表明，对于配置恰当的人工神经网络，局部极小值也是全局极小值（即存在许多局部极小值，但它们的值都相同，且各自都代表全局最小值）。

5. **过完备性是常态，而非例外**。尝试分解已知函数时，完备正交基可能非常高效。然而，对于逆问题——我们希望确定最适合一组数据的函数——使用完备正交基实际上可能会造成阻碍。问题在于，若要在参数空间中搜索唯一解（例如完备正交基所给出的解），这是一个困难的问题。相反，使用过完备的基函数集合，可以搜索任何可接受的解；采用过完备基时，这类解可能有很多（Lewicki 和 Sejnowski，2000）。在某种程度上，这与前一条关于全局与局部极小值的原则有关。

6. **正则化是一种实用的方法**。正则化是一个涵盖范围广泛的术语，指的是增加信息（通过对解空间施加约束，或在损失函数中添加用于惩罚某些类型结果的项）。人工神经网络之所以能够取得成功，其中一个原因是应用了此类方法，

以正则化得到的解。正则化的作用是缩小搜索范围内的参数空间，并减少生成可接受解所需的参数数量。因此，当有大量解都与数据一致时，某些类型的正则化（例如在损失函数中对参数数量进行惩罚）能够打破平局，倾向于选择参数数量最少的解。尽管这些好处看似显而易见，但正则化对深度神经网络究竟有多大效用，仍是一个研究课题（Zhang 等人，2021年）。Zhang 等人（2021年）的研究结果（以及其他研究）表明，正则化有助于高效地找到解，但并非总是必要的。

第一部分：线性前馈网络

10.6 用于线性回归的前馈网络

现在，关于人工神经网络和前馈神经网络的必要背景知识已经介绍完毕，是时候将其中一些理念付诸实践了。几乎每个人至少在概念上都熟悉线性回归的思想。事实证明，用尽可能“最佳”的解拟合一组数据这一问题，是初步探索前馈神经网络应用的理想切入点。

前馈神经网络常用于对一组数据进行简单回归或分类。因此，人们通常假设，这些数据（可能包含也可能不包含因数据收集过程中的不确定性而产生的噪声成分）实际上由某个未知函数描述。如果数据是类别型的，那么该函数描述的是所提出的自变量空间中各类别之间的边界。如果数据代表某种（假定的）连续介质过程，那么就假定回归（采用某种拟合优度指标）是适当的方法，可根据函数所依赖的自变量对该函数进行近似。前馈神经网络还有第三种用途，但文献中通常不会对此进行讨论。这种用途是有意对已知函数进行近似（例如，像傅里叶级数这样的做法）。

In this section, we will consider only linear regression problems on sets of data. Later, we will extend these ideas to nonlinear regression on data, and nonlinear approximation of specified functions.

10.6.1 An example of FNNs for linear regression of data

The use of examples an effective way to see how the concept of FNNs can be combined with ideas from other areas of applied mathematics. In the present example, a two-layer FNN will be constructed to find the best-fit solution to a set of data. Here, the assumption is that the data are well-represented by a *linear function*. The example being presented is somewhat trivial; we will fit a plane through three points. This is a linear problem that has a single unique solution for a specified loss function (and this solution can be found by more conventional methods); however, examining simple versions of otherwise complex problems can be a very effective means for bringing clarity to the methods used.

例如，假设我们有以下自变量（特征）数据集合x_1 和x_2, ，目标是用线性模型拟合这些数据。

Table 10.1 Data for a linear perceptron model. This data set is equivalently specified by the vectorized variable $\mathbf{R} = \{(1,1,5),(2,3,7),(2,1,8)\}$

.

示例编号	x_1	x_2	y
1	1	1	5
2	1	2	7
3	2	1	8

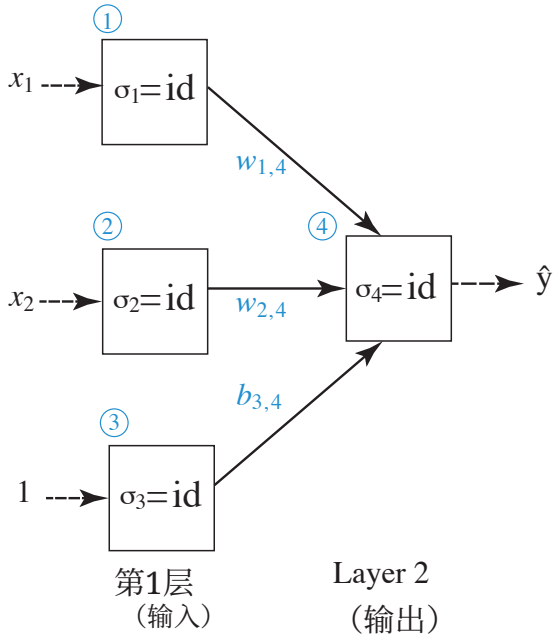


Fig. 10.8 A two layer network. This is a fully *linear network* with a single input and a single output layer; there are no hidden layers (hence, no interior nodes). While the original perceptrons represented binary functions, modern usage of the word frequently includes single-node networks with continuous activation functions (such as the linear one illustrated here).

对于高维数据集，自变量（及其关联数据）的符号表示有时颇具挑战。为帮助缓解这个问题，我们用以下符号表示数据的子集。符号 X 用于表示自变量数据集合。因此，对于上面的例子，表格中给出的有序数据对集合可以表示为 $X(x_1,x_2)$（即 $X(x_1,x_2) = \{(1, 1), (2, 3),(2,1)\}$）。因此，我们可以用变量 $X(x_1,x_2,y)$ 来指代整个自变量（特征）数据集；如果不需要明确列出自变量，则可简称为 X。顾名思义，X_i表示第 i^{th} 个自变量元组。描述因变量数据集合时也采用类似的符号表示。因此，我们有 $Y(y) = \{5,7,8\}$。最后，整个数据集是自变量与因变量之间的一种关系，记为 $R = \{XRY\}$。稍微宽泛地使用符号表示，在上面的例子中，我们将其记作 $R = \{(1,1, 5),(2,3,7),(2,1,8)\}$

The appropriate linear model is given by

$$\hat{y}(x_1,x_2) = w_{1,4}x_1 + w_{2,4}x_2 + b_{3,4} \tag{10.5}$$

假设我们要寻找的解确实是线性的，那么生成所需参数的网络可以由简单感知器构成。简单感知器是一种没有隐藏层的网络。从历史上看，感知器是第一个经实际实现（即在实验室中构建为机电系统）的二元分类网络，并且被证明具备机器学习能力；这一成果由 Warren McCulloch 和 Walter Pitts 于20世纪40年代取得（McCullough 和 Pitts，1943年）。最初的目标是

简单感知器旨在模拟生物神经元（因此，如今人们普遍使用“神经网络”这一术语）。现代对生物神经网络的认识表明，这种模型过于简化；但它仍为网络模型提供了一个有用的原型。人们有时认为简单感知器是严格的线性器件，但事实并非完全如此。原始感知器模型的输出涉及第三层，其中包含一个非线性阈值函数，用于生成二元输出。通常，将每个节点都执行恒等变换、并带有一个单独汇聚层（作为第2层）的双层模型称为感知器，尽管它们并非二分类器。简单的（线性）感知器模型为理解前馈神经网络的工作原理提供了良好基础。对于当前问题，简单感知器网络的结构如图10.8所示。

至此，我们已经具备用于构建并求解该神经网络所表示问题的全部组件。剩下的唯一一步是为损失函数建立显式形式，然后将其最小化。最小平方误差损失函数由观测目标值 y_j 与近似函数在与目标 y_j 所对应的自变量取值处预测的值 $\hat{}y_j$ 之间的平方差给出。该式可以简洁地写成

$$J(\mathbf{R};w_{1,4},w_{2,4},b_{3,4}) = \sum_{j\in\ell(1)} [y_j - \hat{y}(x_{1,j},x_{2,j})]^2 \tag{10.6}$$

其中，$\ell(1) = \{1,2,3\}$ 是数据点数量的索引集。i 下标用于索引示例数据中与 x_1 和 x_2 相关联的 x_1 和 x_2 值。有时会根据函数中数据的数量对损失函数进行归一化，但这主要是计算优化中的一种归一化处理。如果使用解析方法进行优化，归一化并不重要。

With such a simple network, the loss function is easily written out explicitly, and this will help to clarify the notation further. The loss function given above can be more verbosely written as

$$\begin{aligned} J(\mathbf{R};w_{1,4},w_{2,4},b_{3,4}) = {} & [5 - w_{1,4} - w_{2,4} - b_{3,4}]^2 + [7 - w_{1,4} - 2w_{2,4} - b_{3,4}]^2 \\ & + [8 - 2w_{1,4} - w_{2,4} - b_{3,4}]^2 \end{aligned} \tag{10.7}$$

It is necessary to tend to another notational detail. As with the list of dependent and independent variables comprising the data, the list of parameters that are being optimized can grow quite long. Thus, it is common to give them a single summary vector notation. Here, we define $\theta = \{w_{1,4}, w_{2,4}, b_{3,4}\}$. This vector of parameters can be indexed similarly to those for $\mathbf{X}$ (i.e., $\theta_{1,3} = w_{1,4}$, etc.) Now, the expression above can be more clearly written out by

$$\begin{aligned} J(\mathbf{R};\theta) = {} & [5 - w_{1,4} - w_{2,4} - b_{3,4}]^2 + [7 - w_{1,4} - 2w_{2,4} - b_{3,4}]^2 \\ & + [8 - 2w_{1,4} - w_{2,4} - b_{3,4}]^2 \end{aligned} \tag{10.8}$$

Now the goal is to optimize the weights of the loss function such that it reaches its minimum. To do so, we can take derivatives of the loss function with respect to each adjustable parameter, and set each such derivative equal to zero. If the resulting set of equations can be solved simultaneously, then we have a solution to the optimization problem. Computing the appropriate derivatives leads to the following set of equations

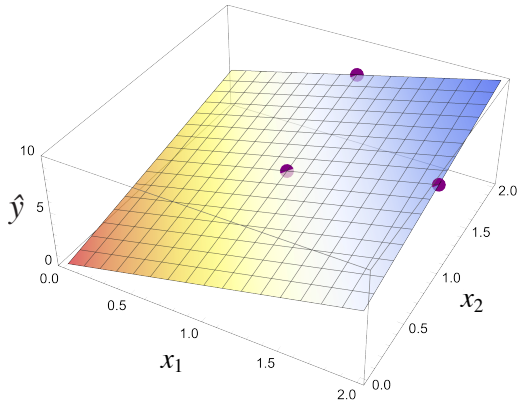


Fig. 10.9 The optimized solution to Eq. (10.5) with Eq. (10.6) as the loss function. Points plotted represent the data used to *train* the simple perceptron network.

$$12w_{14} + 10w_{24} + 8b_{34} - 56 = 0 \tag{10.9}$$
$$10w_{14} + 12w_{24} + 8b_{34} - 54 = 0 \tag{10.10}$$
$$8w_{14} + 8w_{24} + 6b_{34} - 40 = 0 \tag{10.11}$$

This linear system is easily solved to yield

$$\begin{aligned} w_{14} &= 3 \\ w_{24} &= 2 \\ b_{34} &= 0 \end{aligned}$$

And this yields the *approximating function*

$$\hat{y} = 3x_1 + 2x_2 \tag{10.12}$$

Because there is only a single unique plane between any three (unique) points in $\mathbb{R}^3$, the error is identically zero for this case; this is easily verified by computing the value of Eq. (10.6) with this approximating function. A plot of the fit to the data and the data used in fitting appear in Fig. 10.9.

Example 10.1 (Fitting a linear expression to a data set).

In the previous example, the system was somewhat of a special case because there is only one plane through three lines in $\mathbb{R}^3$. Thus, the residual error in this example was identically zero. However, the process changes very little if there are more data point (or, indeed, even if there are more independent variables). Suppose we had a fourth data point in the set of data given by the triplet $(2,2,13)$, so that $\mathbf{R} = \{(1,1,5),(2,3,7),(2,1,8),(2,2,13)\}$. The analysis would be identical, except that the new loss function would be

$$\begin{aligned} J(\mathbf{R}; w_{1,4}, w_{2,4}, b_{3,4}) = {} & [5 - w_{1,4} - w_{2,4} - b_{3,4}]^2 + [7 - w_{1,4} - 2w_{2,4} - b_{3,4}]^2 \\ & + [8 - 2w_{1,4} - w_{2,4} - b_{3,4}]^2 + [13 - 2w_{1,4} - 2w_{2,4} - b_{3,4}]^2 \end{aligned} \tag{10.13}$$

导数的值略有变化，从而得到以下方程组。

$$20w_{14} + 18w_{24} + 12b_{34} - 108 = 0 \tag{10.14}$$
$$2018w_{14} + 20w_{24} + 12b_{34} - 106 = 0 \tag{10.15}$$
$$2012w_{14} + 12w_{24} + 8b_{34} - 66 = 0 \tag{10.16}$$

The solutions is routine, and found to be

$$w_{14} = \frac{9}{2} \tag{10.17}$$
$$w_{24} = \frac{7}{2} \tag{10.18}$$
$$b_{34} = -\frac{15}{4} \tag{10.19}$$

损失函数不再为零。我们能做到的最好结果是损失函数的最小值，即由 $J(\textbf{实数集}; 9/2, 7/2, -15/4) = 9/4$ 给出的最小值。基于这一优化结果修订后的图如下所示。

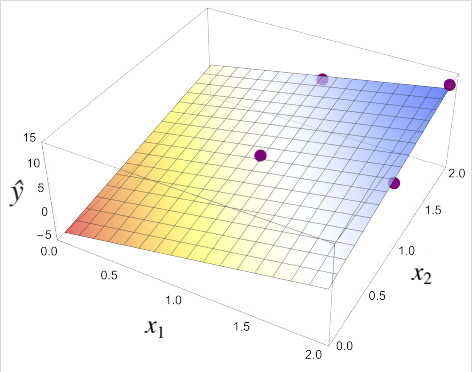


Fig. 10.10: Example of fitting a plane to four data points by minimizing the sum of the squared errors.

PART II: Nonlinear Feedforward Networks

前面的实例展示了由简单线性函数构成的神经网络，并说明传统的线性最小二乘拟合算法可以看作一个两层神经网络。虽然这是阐释神经网络概念的一个很好的例子，但它并未运用现代神经网络所具备的许多优势。具体而言，这种简单网络无法实现以下功能：（1）根据定义，线性映射无法很好地表示非线性函数；（2）这种简单网络无法实现深度神经网络固有的效率。

在接下来的内容中，我们会在解析分析可行的范围内放宽这两项限制；也就是说，我们将研究无需借助计算机即可进行优化的问题。有人可能会问，直接进行解析优化是否算作机器学习；参照术语表中的机器学习定义，答案似乎是否定的！事实确实如此。机器学习涉及的算法是数值方法，可用于迭代训练大型复杂的人工神经网络。然而，即使不使用符合机器学习定义的算法，人工神经网络依然存在，也可以对其进行研究而无需这些算法。有人或许会认为（这也是本章的观点），从解析角度理解人工神经网络的结构和行为，是在机器学习算法所支持的更广泛背景下研究人工神经网络的重要第一步。

本章第二部分的目标是扩展我们对全连接前馈神经网络（FNN）的理解。具体来说，我们将（1）采用非线性激活函数来引入非线性；（2）开始探讨深度的作用，考察具有一个或多个隐藏层的神经网络。我们的重点主要是只有一个隐藏层的网络，严格来说，这些网络并不算深度网络。然而，至少在解析的背景下，与非线性表征和最优参数确定相关的概念，对于深度网络而言并无实质差异。对于具有两个或更多层的神经网络，由于微分时需要使用链式法则（或复合法则），网络的微分会变得更加复杂；这也反映了神经网络数值优化早期面临的一个挑战，而反向传播算法最终解决了这一问题。本章第二部分以讨论一种称为修正线性单元（ReLU）函数的特定非线性激活函数开篇。

10.7 ReLU 函数

为了进行计算，我们必须选择一种用于网络的激活函数。归根结底，网络的深度和用途会在一定程度上影响激活函数的选择。

对于深度神经网络，激活函数的选择会显著影响网络的性能（即优化的难易程度）。就计算效率而言，修正线性单元（ReLU）函数在回归和分类任务中的表现都优于其他激活函数。它具有多项优势，因此得到了广泛应用。其主要原因在于实际应用方面：（1）计算该函数及其导数的开销很小；（2）当 $|x| \to \infty$ 时，ReLU 函数不会出现“梯度消失”问题。这一特性对于深度网络中通过反向传播在网络中传递权重信息十分重要。作为一个具体的反例，过去人们曾非常青睐S形函数（见图10.3）。然而，当 $|x| \to \infty$ 时，这些函数会出现梯度消失问题，并在一段时间内阻碍了深度网络的研究（Glorot 和 Bengio，2010）。

The ReLU function has a number of the other nice features. It is a nonlinear and non-polynomial function, which is essential for its use as a basis function Pinkus (1999). While it is nonlinear, it also has the property that any weighted sums of the ReLU function generate a piecewise-linear result. This is, interestingly, also true for compositions: the composition of a ReLU function with any weighted sum of ReLU functions generates a piecewise linear function. These qualities give the ReLU function an algebraic structure that makes analysis of large networks more tractable.

Among the various possible definitions, we adopt the following for the ReLU function defined on $x \in (-\infty, \infty)$

$$\mathrm{ReLU}(x)=\begin{cases}0 & x<0\\ 1 & x>0\\ 0 & x=0\end{cases} \tag{10.20}$$

虽然 ReLU 函数不具备连续导数，但它的其他特征足以弥补这一缺陷。这是一个很好的例子，说明某项技术的发展最初主要出于实际应用需求，而后来人们发现，它具有特定的理论结构，这使其表现得尤为有效。with

$$\frac{d}{dx}\mathrm{ReLU}(x)=\begin{cases}0 & x<0\\ 1 & x>0\\ \text{undefined} & x=0\end{cases} \tag{10.21}$$

在前馈神经网络的文献中，通常会在$x=0$处为 ReLU 函数的导数任意指定一个值（通常介于 0 和 1 之间）。然而，事实证明，$x=0$处导数的不连续性不会给回归问题带来解析或数值上的问题。下文有关 ReLU 函数微积分的章节将对此作进一步说明。

由于在后续内容中使用这一符号表示会很方便，我们将函数H_0定义为海维赛德函数

$$H_0(x)=\begin{cases}0 & x\le 0\\ 1 & x>0\end{cases} \tag{10.22}$$

因此，ReLU 函数的另一种表达式为

$$\mathrm{ReLU}(x)=xH_0(x) \tag{10.23}$$

10.8 ReLU 的代数性质

在应用中，通常重要的是了解如何对 ReLU 函数进行缩放、平移和相加。虽然 ReLU 函数是非线性的，但它们仍具有一些近似线性的性质（这在一定程度上也是其取得成功的原因）。为展开讨论，我们先关注单个自变量；随后再讨论推广到多个自变量的情形。下文将介绍 ReLU 函数的缩放（通过标量乘法实现）、平移和加法。注意，ReLU 函数是非线性的。ReLU 函数既不满足齐次性，也不满足可加性。换句话说，我们有

$$\mathrm{ReLU}(ax)\neq a\mathrm{ReLU}x \tag{10.24}$$

and

$$\mathrm{ReLU}(x+y)\neq\mathrm{ReLU}(x)+\mathrm{ReLU}(y) \tag{10.25}$$

The piecewise-linear form of the ReLU function does give it some advantages. The quasi-linearity of the function helps tremendously in analytic efforts for solutions and interpretation of FNNs. In addition, the ReLU function avoids some of the problems created by other candidates for activation functions that have been used in the past. For example, the logistic ("s" shaped) function is smooth, but because its derivatives rapidly tend toward zero as $|x|\to\infty$, such functions can create difficulties when using numerical methods

such as gradient decent and conventional backpropagation of information throughout the network (which both rely on some linearly proportional function of the *derivatives*).

单变量仿射函数的形式为 $f(x) = ax$ 加上它的所有平移（例如，$f(x) = a(x+\beta) = ax + a\beta$是线性函数加上它的所有平移，$\beta$。ReLU 函数是分段仿射的；由于已有大量关于这类函数的理论，这些理论有助于更好地理解使用 ReLU 作为激活函数的深度神经网络的函数。

10.8.1 Translation

以 ReLU 函数代数的平移操作定义作为讨论的起点，虽然有些不同寻常，但对于理解其在人工神经网络中的应用而言，这可能是最重要的概念。假设在一维空间中，我们定义 开始 T^β 为一种运算，它将所作用的函数平移β 的距离（该值可以是正数，也可以是负数），那么我们有

$$T_\beta(\mathrm{ReLU}(x)) = \mathrm{ReLU}(x+\beta) \tag{10.26}$$

The translation operator acts only on the variable x, so some care is needed in interpreting it correctly. For example, if a scalar multiplies x, we interpret the translation as follows.

$$T_\beta \mathrm{ReLU}(ax) = \mathrm{ReLU}(a(x+\beta)) \tag{10.27}$$

This operation is correct regardless of the sign of a. Sequential translations can be defined (although one does not usually encounter this)

$$\begin{aligned} T_\beta T_\gamma \mathrm{ReLU}(ax) &= T_\beta \mathrm{ReLU}(a(x+\gamma)) \\ &= \mathrm{ReLU}(a([x+\beta]+\gamma)) \\ &= \mathrm{ReLU}(a(x+\beta+\gamma)) \end{aligned} \tag{10.28}$$

Note that the translation operator is a linear one:

$$T_\beta[\mathrm{ReLU}(w_{1,3}x) + \mathrm{ReLU}(a_2 x)] = \mathrm{ReLU}(w_{1,3}(x+\beta)) + \mathrm{ReLU}(a_2(x+\beta)) \tag{10.29}$$

While β the the amount of the *translation* of the ReLU function, it is not in general equivalent to the *bias* term that is typically adopted in applications. To make this clear, define $b = a\beta$. Then

$$\mathrm{ReLU}(a(x+\beta)) = \mathrm{ReLU}(ax+b)$$

其中，b 是偏置项。从现在开始，符号b 始终表示偏置，而符号β 则用于表示平移。图 10.11 绘制了平移和缩放后的*relu* 函数，供参考。

10.8.2 Addition

The addition of ReLU functions is straightforward. Combining the operations above with that of addition gives

$$\begin{aligned} T_{\beta_1}(a_1\mathrm{ReLU}(x)) + T_{\beta_2}(a_2\mathrm{ReLU}(x)) &= \mathrm{ReLU}(a_1(x+\beta_1)) + \mathrm{ReLU}(a_2(x+\beta_2)) \\ &= \mathrm{ReLU}(a_1 x + b_{2,3}) + \mathrm{ReLU}(a_2 x + b_2) \end{aligned} \tag{10.30}$$

An example of the sum of *relu* functions appears in Fig. 10.12.

While this definition is straightforward, note that when $\beta_1 = \beta_2 = \beta$, the sum is a *degenerate* case; that is, sum of the two functions can be replaced by a single ReLU function.

$$\begin{aligned} &= \mathrm{ReLU}(a_1(x+\beta)) + \mathrm{ReLU}(a_2(x+\beta)) \\ &= T_\beta(a_1\mathrm{ReLU}(x)) + T_\beta(a_2\mathrm{ReLU}(x)) \\ &= T_\beta(\mathrm{ReLU}((a_1+a_2)x)) \end{aligned}$$

Setting $a_3 = a_1 + a_2$ gives the result

$$= \mathrm{ReLU}(a_3(x+\beta)) \tag{10.31}$$

In applications, degenerate cases are inefficient (by unnecessarily increasing the number of parameters involved). Depending upon the system being investigated using ReLU, one might also seek to avoid cases that are *nearly degenerate*, i.e., for two one-dimensional ReLU functions shifted by β_1 and β_2 respectively, then if ε_0 represents a minimum (scaled) separation or "resolution" between any two ReLU functions, a nearly degenerate case would be indicated by

$$\left|1 - \frac{\beta_1}{\beta_2}\right| < \varepsilon_0 \tag{10.32}$$

The problem created by nearly degenerate cases is a practical one. For such cases, the concept of nearly degenerate would be established *a priori*; it means only that the expressiveness gained by adding two ReLU

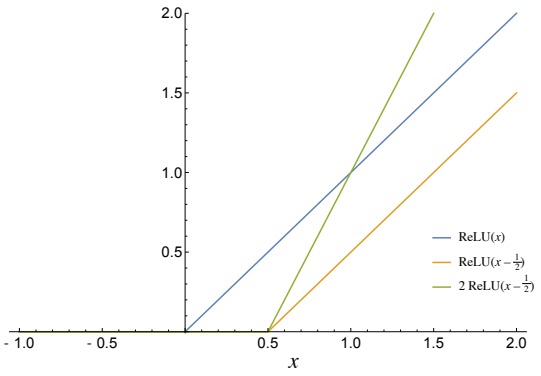


图 10.11 平移和缩放后的 ReLU 函数。

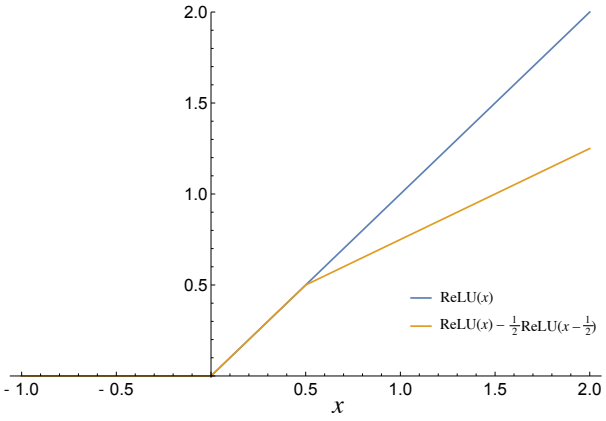


Fig. 10.12 Addition of non-degenerate ReLU functions.

functions that are sufficiently close together (in terms of the shift variable, β) is not warranted by the quality of the data or other external factors.

10.8.3 Multiplication by a scalar

Multiplication by a positive scalar is commutative with the ReLU function. Thus, we can write for any real number a

$$\text{ReLU}\left(a^2 x\right) = a^2\,\text{ReLU}\,(x) \tag{10.33}$$

Unfortunately, this kind of linear behavior does not extend to cases where the constant might be negative. Suppose a is any real number, then we have

$$\begin{aligned}\text{ReLU}\,(a(x+\beta)) &= \text{ReLU}\,(\text{sng}(a)|a|(x+\beta)) \\ &= |a|\,\text{ReLU}\,(\text{sng}(a)(x+\beta))\end{aligned} \tag{10.34}$$

其中回忆起 $\beta = b/a$，符号函数定义为

$$\text{sgn}(a) = \frac{a}{|a|} = \begin{cases} +1 & \text{for } a > 0 \\ 0 & \text{for } a = 0 \\ -1 & \text{for } a < 0 \end{cases} \tag{10.35}$$

Conversely, we also have

$$\begin{aligned}c\text{ReLU}\left(a'(x+\beta)\right) &= |c|\text{sng}(c)\text{ReLU}\,(a(x+\beta)) \\ &= \text{sng}(c)\,\text{ReLU}\left(|c|a'(x+\beta)\right)\end{aligned} \tag{10.36}$$

Note that we can now define $a = a'|c|$ to write

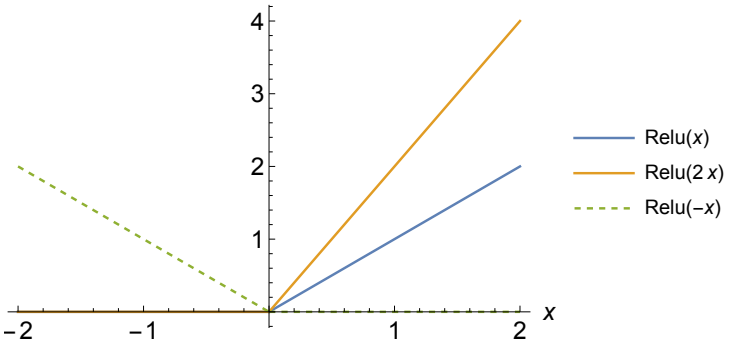


Fig. 10.13 Three ReLU functions, differing only by the constant, a.

$$c\text{ReLU}\left(a'(x+\beta)\right) = \text{sng}(c)\,\text{ReLU}\left(a(x+\beta)\right) \tag{10.37}$$

This last relation does allow some simplification of the operation of scalar multiplication with a ReLU function. In the last layer of a network, we often have forms that are simple weighted sums of ReLU functions. This last relation allows us to replace those weights by ± 1, where now the magnitude is fixed, but the sign must be determined. Later, we will encounter examples where this property allows us to make effective constraints that allow optimization to be done more efficiently.

10.8.4 Multiplication of ReLU *functions*

Although not commonly arising in applications, the product of ReLU functions is nonetheless well-defined. At the end of this chapter, we will provide an example of using a ReLU squared function as the activation. To define the product of ReLU functions, it is helpful to adopt the following form for the function

$$\text{ReLU}\left(ax+b\right) = \text{ReLU}\left(a(x+\beta)\right) = a(x+\beta)H_0(a(x+\beta)) \tag{10.38}$$

这里，赫维赛德阶跃函数$_0(x)$ 是前面定义的海维赛德函数。

假设我们将 ReLU($a_1(x-b_{2,3})$) 与 ReLU ($a_2(x-b_2)$) 相乘。利用式 (10.38)，很容易验证（使用海维赛德函数的性质）

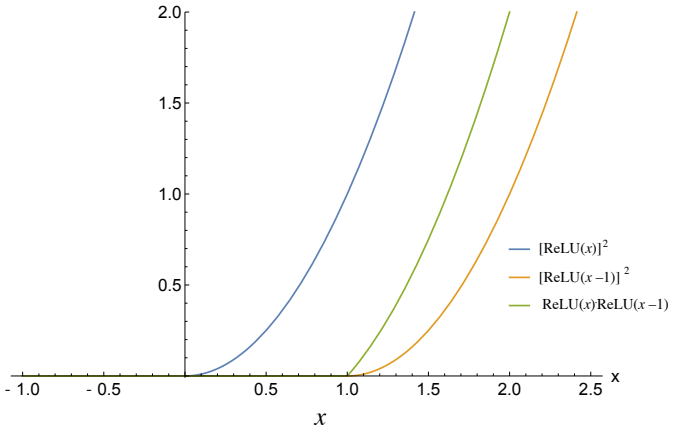
$$\text{ReLU}\left(a_1x-b_{2,3}\right)\cdot\text{ReLU}\left(a_2x-b_2\right) = (a_1x-b_{2,3})H_0\left[a_1x-b_{2,3}\right](a_2x-b_2)H_0\left[a_2x-b_2\right] \tag{10.39}$$

从这一点开始，关于 Heaviside 函数的配置有四种可能的情况。因此，不存在简单的 简化方法。ReLU^2 情况是在应用中使用的一种情况，并且具有简化形式。此情况可以按如下方式简化。

$$\begin{aligned}\text{ReLU}^2(a(x-b)) &= H(ax+b)\text{ReLU}\left(ax+b\right))H(ax+b)\text{ReLU}\left(ax+b\right)) \\ &= H(ax+b)(ax-b)^2\end{aligned} \tag{10.40}$$

Example plots of the multiplication of *relu* functions appears in Fig. 10.14.

图 10.14 ReLU函数的乘积。



10.8.5 Composition of ReLU *Functions*

Recall that the composition of two functions f and g with independent variable x is a given by

$$F(x;f,g) = f(g(x)) \tag{10.41}$$

ReLU 函数的简单复合具有这样的性质：它们会生成另一个 ReLU 函数。这里，简单 是指仅将一个 ReLU 函数与另一个 ReLU 函数进行复合。对于简单复合，我们有如下结论。

$$\begin{aligned}\text{ReLU}\left(\text{ReLU}\left(ax-b\right)\right) &= \text{ReLU}\left(a\text{ReLU}\left(x-b/a\right)\right)\\ &= \text{ReLU}\left(a\text{ReLU}\left(x\right)-b/a\right)\\ &= \text{ReLU}\left(x-b/a\right)\end{aligned} \tag{10.42}$$

While the composition of two ReLU functions is somewhat uninteresting, things change dramatically when one considers the composition to a ReLU functions with a linear combination of two others. Because of the non-additivity of ReLU functions, such compositions result in complicated relationships that cannot be easily expressed in simpler terms. In other words, expressions like $\text{ReLU}\left(\text{ReLU}\left(a_1x+b_{2,3}\right)+\text{ReLU}\left(a_2x+b_2\right)\right)$ cannot be more simply expressed in general. As a concrete examples, consider

$$f(x) = \text{ReLU}\left[\text{ReLU}\left(x-1\right)-\frac{10}{9}\text{ReLU}\left(x-2\right)\right] \tag{10.43}$$

$$f(x) = \frac{1}{10}\text{ReLU}\left[\text{ReLU}\left(x-5\right)+\frac{1}{2}\text{ReLU}\left(x-8\right)\right] \tag{10.44}$$

这些复合函数的图像见图 10.15。很容易证明，两个 ReLU 函数之和的所有此类复合都是正函数（其中包括零函数）。无需深入细节即可说明，ReLU 函数复合的性质是解释神经网络取得巨大成功的原因之一。

如上所述（即使对于宽度固定的神经网络，这一点也成立），随着深度增加，自变量的底层空间中由线性函数定义的区域数量呈指数增长（参见 Hanin（2019））。换言之，ReLU 函数生成了关于自变量的分段连续函数，而构成该函数的片段数量以（渐近的）指数阶增长。ReLU 函数组合的这一特征非常有效。简而言之，使用单层也可以实现相同的结果，但等效的单层网络所包含的参数数量将比多层网络多出指数级。由于这些结果是由估计值的渐近分析确定的，因此通常需要足够大的自变量维度和足够多的层数，才能在实践中实现近似指数级的扩展。然而，这正是许多深度神经网络在传统实践中的结构方式，因此函数复合的幂可以发挥非常有效的作用。

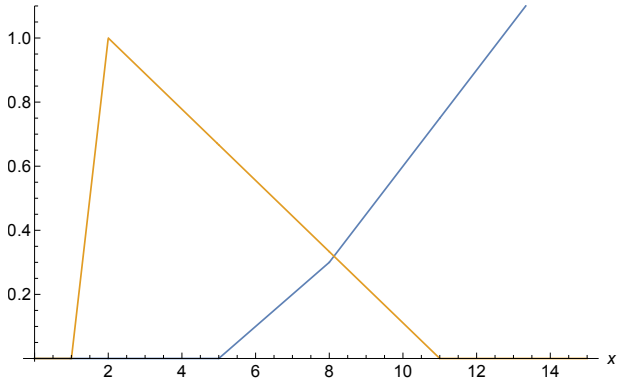


Fig. 10.15 The composition of two ReLU functions. Orange line: $f(x) = \mathrm{ReLU}\left[\mathrm{ReLU}(x-1) - \frac{10}{9}\mathrm{ReLU}(x-2)\right]$. Blue line: $f(x) = \frac{1}{10}\mathrm{ReLU}\left[\mathrm{ReLU}(x-5) + \frac{1}{2}\mathrm{ReLU}(x-8)\right]$

10.9 Calculus of ReLU Functions

10.9.1 Integration of the ReLU *function*

Using the alternative expression for the ReLU function can be found by using the Heaviside (step) function.

$$H_0(x) = \begin{cases} 0 & x \leq 0 \\ 1 & x > 0 \end{cases} \tag{10.45}$$

根据这个定义，我们可以得到 ReLU 函数的另一种表达式。

$$\mathrm{ReLU}(x) = xH_0(x) \tag{10.46}$$

$$\frac{d}{dx}\text{ReLU}\,(x) = H_0(x) \tag{10.47}$$

$$\text{ReLU}\,(x) = \int_{z=0}^{z=x} H_0(z)\,dz \tag{10.48}$$

$$\text{ReLU}\,(x) = xH_0(x) \tag{10.49}$$

of these, the last one provides a convenient way of integrating the ReLU function. Assuming $x > 0$, then

$$\begin{aligned}\int_{-\infty}^{z=x} \text{ReLU}\,(z)\,dz &= \int_{z=0}^{z=x} zH(z)\,dz \\ &= \int_{z=0}^{z=x} z\,dz \\ &= \frac{x^2}{2}\end{aligned} \tag{10.50}$$

10.9.2 Derivative of the ReLU *function*

在优化过程中，需要用到 ReLU 函数的导数。虽然许多参考文献都将不连续点处的导数指定为一个任意值，但实际上这并非必要。同样，将 ReLU 函数表示为如下形式。

$$f(x) = \text{ReLU}\,(ax+b)\,H_0(ax+b) \tag{10.51}$$

we can consider the derivative of the *relu* function in the distributional sense (cf. Chp. 5). Taking the derivative of both sides of Eq. (10.51), we obtain the distributional derivative

$$\begin{aligned}\frac{df}{dx} &= \frac{d}{dx}\left[\text{ReLU}\,(ax+b)\,H_0(ax+b)\right] \\ &= aH_0(ax+b) + \text{ReLU}\,(ax+b)\,\delta(ax+b)\end{aligned} \tag{10.52}$$

Now note that the second term in this expression involves the delta function. Recall that in a distributional sense, the delta function is only nonzero at the point where its argument is zero, i.e., when $ax+b=0$. But, it is multiplied on the outside by a function that will return zero when $ax+b=0$. Thus, this term is always zero. More correctly, we might establish this by integration of both sides of this result from $(-\infty, x)$ with $x > 0$

$$\begin{aligned}\int_{-\infty}^{x}\frac{df}{dx}\,dx &= \int_{-\infty}^{x} aH_0(ax+b)+\mathrm{ReLU}\,(ax+b)\,\delta(ax+b)\,dx\\ &= \int_{-\infty}^{x} aH_0(ax+b)\,dx+\mathrm{ReLU}\,(0)\,\delta(0)\\ &= \int_{-b/a}^{x} a\,dx\\ &= ax+b \ \text{ for } x>-b/a \qquad (10.53)\\ &= \mathrm{ReLU}\,(ax+b)\,H_0(ax+b) \qquad (10.54)\end{aligned}$$

showing that we recover the original function after integrating the derivative (as we must). The practical result of this analysis is that the derivative of the ReLU function can safely be set to zero at the point of discontinuity, and this result is at least consistent with the distributional interpretation of the derivative.

10.10 *ReLU* 函数在多维中的应用

ReLU 函数的一项有趣特性是，这些函数的简单复合也可能变得多么复杂。随着自变量（特征）数量增加，这一点会变得更加明显。这里讨论的主要目的是建立一些直觉，了解这类函数的行为和表现形式。我们将范围限制在二维，考虑依赖于变量对（x，y）的 ReLU 函数。具体而言，我们将为加权（斜率和偏置）参数选取特定值。假设我们定义以下函数

$$\sigma_1(x,y) = \mathrm{ReLU}\,(x+2y-4)+2 \qquad (10.55)$$
$$\sigma_2(x,y) = \mathrm{ReLU}\,(-x+1/2y-1) \qquad (10.56)$$

and

$$\sigma_3(x,y) = \mathrm{ReLU}\,(\sigma_1(x,y)-\sigma_2(x,y)) \qquad (10.57)$$

尽管两个 ReLU 函数之差的复合未必看起来复杂，但其图像展示了 ReLU 函数的复合能够生成越来越多区域。这种增长大致呈多项式（即以某个由多项式描述的因子增长）：随着自变量数量或网络宽度增加，区域数量也会增加；而随着层数增加，区域数量则呈指数增长（Pascanu 等，2014 年）。

As an example of this behavior, the functions σ_1, σ_2, and σ_3 are plotted in Fig. 10.16.

许多研究者研究了不同配置的前馈神经网络中，ReLU 函数在自变量空间形成的区域数量。Montufar等人（2014年）和Pascanu等人（2014年）详细给出了ReLU 前馈神经网络所形成区域数量的估计值（以及下界）。

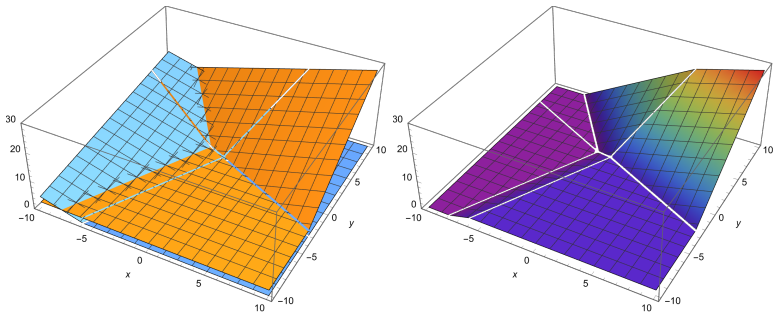


图10.16 （左）两个自变量中的 ReLU 函数，σ_1（橙色）和 σ_2（蓝色）。（右）函数 $\sigma_3 = \mathrm{ReLU}(\sigma_1 - \sigma_2)$。右侧图形的颜色色标表示 σ_3 的值，并添加该色标以改善函数可视化。

10.11 关于多维 ReLU 函数、网络与逼近理论的简要评述

逼近理论研究如何用更简单的函数尽可能好地逼近函数，以及这类近似对原函数的表征效果如何。从这个角度看，傅里叶级数也可视为逼近理论的一部分。然而，逼近理论所使用的函数可能与我们熟悉的光滑三角函数大不相同。逼近理论涉及的内容极为广泛，无法在此一一概述，但相关文献中有具体实例可供参考。例如，如径向基函数理论所述，可以将高斯函数组合起来，近似其他函数。小波可视为傅里叶变换的一种推广，它着重表示所表征函数的局部结构，也是用于构造其他函数近似的一类函数。讨论这些方法时，主要会遇到以下问题：（1）所提出的函数是否足够稠密（按某种测度而言），能够表征待近似的函数类型；（2）所研究的基函数是否正交（如傅里叶级数）。关于第一个问题，已有许多论文研究人工神经网络的逼近理论（相关文献综述见（Pinkus, 1999年））；普遍共识是，ReLU 函数确实足够稠密，能够表征定义在紧区间（或体积，或维度大于1的空间）上的L_2 函数。第二个问题也很有意思。人们早已知道，在非常宽泛的条件下，高斯函数乃至紧支撑凸起函数都能构成L_2 函数的稠密基。然而，构造这类基的方法一直缺乏，主要是因为相关方法的约束过少。

神经网络通过不试图构造必然唯一的解来避开这个问题。换句话说，采用数值方法求解时，目标通常只是找到某个满足所需误差条件的解。由于梯度下降（以及一种称为随机梯度下降的相关方法）具有独特的工作方式，即使潜在解的空间极其庞大，这些方法也能够找到解。

One recent method that has been adopted to examine the abilities of FNNs in particular is to compare them with existing methods for approximation. For example, one can construct triangle functions using a combination of three ReLU functions as follows

$$f[x] = 4\mathrm{ReLU}(x) - 8\mathrm{ReLU}(x-1/2) + 4\mathrm{ReLU}(x-1) \tag{10.58}$$

This function is plotted in Fig. 10.17.

One correspondence one can make is with the conventional finite-element constructions that have been well studied in the literature. Although many functions are used in practice, the simple triangle function is the most common. In terms of ReLU functions, this is

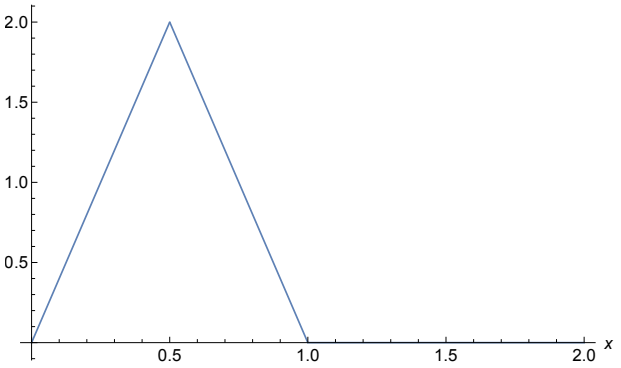


图 10.17 由 ReLU 函数$f[x] = 4\text{ReLU}(x) - 8\text{ReLU}(x - 1/2) + 4\text{ReLU}(x - 1)$之和生成的对称紧支撑函数。

This function is useful because it indicates that, in some senses, ReLU is equivalent to a compact basis function for linear finite elements. In fact, it has been shown that there are equivalences between finite element basis functions and ReLU functions (He et al., 2022). A paper by Telgarsky (2015) provides an analysis of error bounds for compositions of sawtooth functions that correspond to deep ReLU networks.

研究人员采用的另一种方法是，展示如何利用这类函数构造单位分解（即定义在$x \in [0,1]$上的有限个函数的集合，使得定义域内每个点处这些函数的和都为 1）。这类构造有强大的解析框架，能够展示 ReLU 函数组成的网络如何表示高维数据集（Yarotsky，2017；Poggio 等人，2017）。另一种已有研究的方法是利用 ReLU 函数构造类似傅里叶的正交函数，然后证明这些函数具有良好的表达能力（Daubechies 等人，2022）。例如，考虑由（10.58）给出的紧支撑函数$f(x)$。现在考虑

$$g1(x) = f(x) - f(x-1) \tag{10.59}$$

$$g2(x) = g1(2x) + g1(2(x-1)) \tag{10.60}$$

These functions are plotted in Fig. 10.18. It is clear that these functions are orthogonal. Functions constructed in this manner a generalization of Fourier series, and are known as Riesz bases. More information on such constructions can be found in the literature (Daubechies et al., 2022).

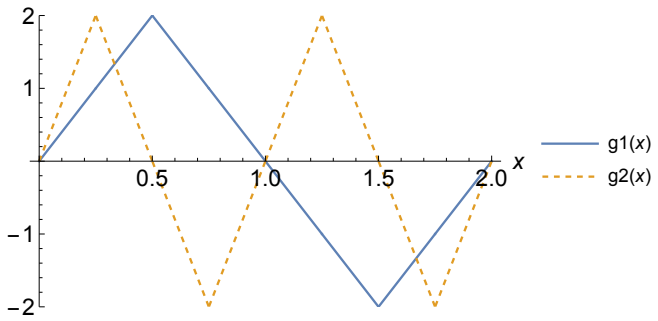


Fig. 10.18 Orthogonal functions constructed from *relu* functions.

While these examples are far from "proofs" that ReLU functions can effectively reproduce many continuous functions, it does provide some intuition about how such a goal might be pursued. There is still much

that remains unknown about the expressive power of deep ReLU networks, but this remains an active area of investigation. A paper by Daubechies et al. (2022) provides a particularly insightful analysis along the lines discussed above.

10.11.1 脊函数

在结束关于 ReLU 函数及其在近似中的作用的这一节之前，我们不妨先看看一类特殊的函数，它们可以由高维 ReLU 函数构造而成：这类函数称为脊函数。脊函数是函数的线性加权和，可形成一种高维结果，类似于一维三角函数。下文给出了两个二维脊函数的示例，其图像见图 10.19。

脊函数的显著特征是，在适当的坐标系中，它们所依赖的自变量比定义该空间的自变量更少。具体来说，请看图 10.19。如果我们将坐标系旋转 $\pi/2$ （根据情况取正号或负号），那么其中一个坐标轴将与$x_1 + x_2 - 1 = 0$ （对于$r_1(x-1,x_2)$）以及$x_1 - x_2 = 0$（对于$r_2(x_1,x_2)$）方向上的折痕对齐。这样一来，这些函数便不再依赖于与折痕对齐的变量，而只依赖于与折痕垂直的变量。在相关文献中，这有时被描述为找到一个嵌入高维空间中的低维流形 （这里的“流形”只是欧几里得空间中子集的一种推广）。换句话说，尽管函数嵌入在二维空间中，我们仍可将其表示为仅含一个变量的函数。目前关于人工神经网络为何表现如此出色的一些理论，涉及对这类子流形的识别。不过，我们展示这些函数并非特别为了讨论流形。后面我们会发现，这些函数在确定用于学习异或函数的神经网络时有具体用途。

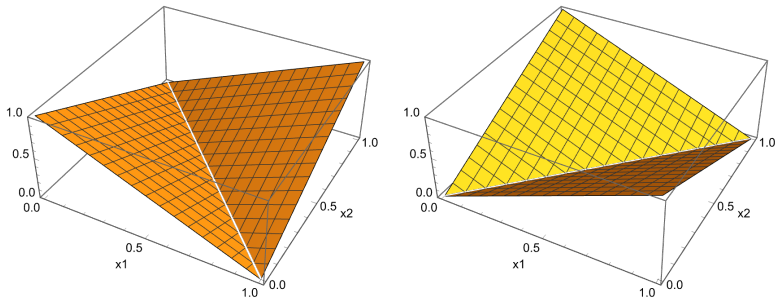


图 10.19 两个岭函数。这些函数的表达式见正文。函数$r_1(x_1,x_2)$位于左侧，函数$r_2(x_1,x_2)$位于右侧。

The expression for these two ridge functions is given as follows

$$r_1(x_1,x_2) = \mathrm{ReLU}(x_1 + x_2 - 1) + \mathrm{ReLU}(-x_1 - x_2 + 1) \tag{10.61}$$
$$r_2(x_1,x_2) = \mathrm{ReLU}(-x_1 + x_2) + \mathrm{ReLU}(x_1 - x_2) \tag{10.62}$$

These examples illustrate two shapes that can be useful in applications. Later, we will have an opportunity to construct a neural network for the exclusive or, XOR, function. This function has been offered as an example

that cannot be constructed with a one-hidden-layer neural network model; we will have the opportunity construct a one-layer model that illustrates that this is untrue.

10.12 用于非线性回归的前馈网络

现在我们已经概述了对 ReLU 函数的一种恰当描述，便可以考虑构建一些能够执行指定函数的示例网络。本节将考察应用于数据集的单隐层和双隐层前馈神经网络。

10.12.1 使用 ReLU 的单隐层前馈神经网络

单隐层前馈神经网络与自由节点线性样条法有许多相似之处（DeVore，1998）。事实上，对于单层 ReLU 前馈神经网络而言，两者本质上是相同的（Plonka 等，2022）。虽然单层 ReLU 网络的效率不及更深的网络，但由于其相对透明，是开展研究时一个有用的模型。换言之，我们可以轻松地解读单层 ReLU 网络的结果。首先，下面举例说明如何用单层 ReLU 网络确定二元与函数。

Example 10.2 (Learning the AND function).

The binary AND function is an interesting example to examine using a single-layer ReLU network. The *AND* function is a binary function whose functional relations are specified by the following table

x_1	x_2	y
0	0	0
0	1	0
1	0	0
1	1	1

Table 10.2: The binary AND function.

It is not difficult to determine that the solution $y(x_1,x_2) = \text{ReLU}(x_1 + x_2 - 1)$ is a continuous function that matches the binary AND function at the points $(1,1),(0,1),(1,0)$ and $(1,1)$. The question now is, can we determine this function by constructing an appropriate neural network and then optimizing the parameters. To do so, we first draw out an appropriate network: This is done in Fig. 10.2.

To start the process, we first convert our figure into the equivalent mathematical representation. It is not difficult to work out the representation for this simple network, it is

$$\hat{y}(x_1,x_2) = \text{ReLU}\left(w_{1,4}x_1 + w_{2,4}x_2 + b_{3,4}\right) \tag{10.63}$$

Now, we have a number of choices for the *loss* function for this problem. However, the L_2 loss function（最小二乘- squares) is convenient. The loss function can be represented by the difference between our *data*

and the predicted function, $\hat{y}$ at the same values of the independent variables for which we have data. Thus, our loss function is given by the sum

$$J(x_1,x_2,w_{1,4},w_{2,4},b_{3,4}) = [0-\hat{y}(0,0)]^2+[0-\hat{y}(1,0)]^2+[0-\hat{y}(0,1)]^2+[0-\hat{y}(1,1)]^2 \tag{10.64}$$

This problem is simple, and it can be solved uniquely without imposing any additional constraints. However, the ReLU function creates a number of branching options in the loss function without additional constraints. Or, in other words, the loss function is described by four piecewise function statements, each of which has a zero portion an a positive portion arising from the application of the ReLU function. To be concrete about this, the following results from the expansion of the loss function

$$\begin{aligned} J(x_1,x_2,w_{1,4},w_{2,4},b_{3,4}) = &\left[1-\left(\begin{cases}0 & b_{3,4}+w_{1,4}+w_{2,4}<0\\ b_{3,4}+w_{1,4}+w_{2,4} & b_{3,4}+w_{1,4}+w_{2,4}\geq 0\end{cases}\right)\right]^2 \\ &+\left(\begin{cases}0 & b_{3,4}+w_{1,4}<0\\ b_{3,4}+w_{1,4} & b_{3,4}+w_{1,4}\geq 0\end{cases}\right)^2+\left(\begin{cases}0 & b_{3,4}+w_{2,4}<0\\ b_{3,4}+w_{2,4} & b_{3,4}+w_{2,4}\geq 0\end{cases}\right)^2 \\ &+\left(\begin{cases}0 & b_{3,4}<0\\ b_{3,4} & b_{3,4}\geq 0\end{cases}\right)^2 \end{aligned} \tag{10.65}$$

这个方程有 4 个二次项，每个这样的项都涉及 2 个定义域；因此可能有 16 种不同的情况组合（包括任何冗余情况）需要检查！不过，恰当地施加约束条件可以将其基本减少到零。请考虑以下约束条件。首先，我们知道在点 (0，0) 处，对函数的唯一贡献是 b_{34}。由于函数在此点应为零，显然 $b3,_4 \leq 0$（这样对其应用 ReLU 函数时，结果就是零）。此外，与函数关于直线 $x_1 = x_2$, 对称，也就是说 $y(0，1) = y(1，0)$。定义该函数的网络也具有对称性。因此，我们必须有 $w1,_4 = w2,_4$（没有任何机制能够打破这种对称性，因此对称的输入数据必然会得到对称的解）。最后，我们知道 $y(1，1) = 1$，所以必须有 $w1,_4 = w2,_4 > 0$。

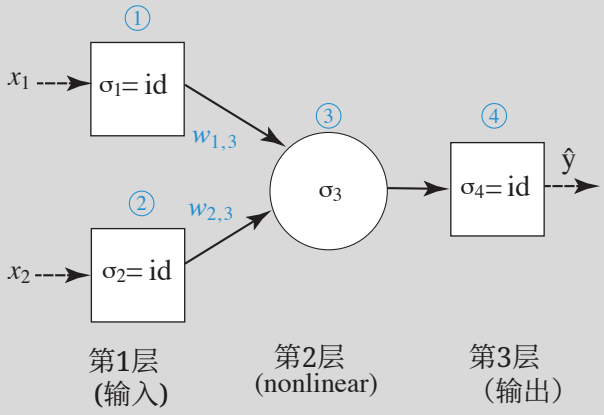


图 10.20：用于确定与函数的网络。

要优化该系统，我们只需找出参数空间中使每个自变量（$w_{1,4}$,$w_{2,4}$,$b_{3,4}$）对应的导数均为零的点。施加额外的约束条件后，每个导数都只有一个表达式。结果如下

$$a_1(w_{1,4},w_{2,4},b_{3,4}) = \frac{\partial J}{\partial w_{1,4}} = 2(-1+b34+w14+w24) \tag{10.66}$$

$$a_2(w_{1,4},w_{2,4},b_{3,4}) = \frac{\partial J}{\partial w_{2,4}} = -2+6b34+8w24 \tag{10.67}$$

$$a_3(w_{1,4},w_{2,4},b_{3,4}) = \frac{\partial J}{\partial w_{2,4}} = -2+6b34+8w24 \tag{10.68}$$

由于a_2 和a_3 相同，因此我们必须将约束条件$w1_{,4} - w2_{,4} = 0$作为第三个方程。这样便得到三个未知数的三个线性方程。求解过程很直接，此处不再详述。解为

$$w_{1,4} = 1 \tag{10.69}$$

$$w_{2,4} = 1 \tag{10.70}$$

$$b_{3,4} = -1 \tag{10.71}$$

与函数的最终解由以下表达式给出

$$\hat{y}(x_1,x_2) = \text{ReLU}\,(x_1+x_2-1) \tag{10.72}$$

很容易验证，对于（x_1,x_2）的四种可能输入值，此解都是正确的
。最后再次强调，这个问题无需施加额外约束也能求解。采用这种方法可以得到一个解，但每个导数函数都需要用四个独立区间来定义，从而构成分段导数。如果所得区间互不重叠，则最多可能需要检查 64 种情况。如果找到多个使损失函数达到最小值的解，我们就选择其中的全局最小值。因此，使用约束条件有助于我们以简单的方式求解这个问题；不过，蛮力法也同样可行！

As a second example, we illustrate the problem of fitting a set of nonlinear data that has no a priori *known* functional behavior other than the data that it contains. This example is somewhat of an artifice– three data points are fit with two ReLU functions, yielding a solution that simply connects the three points via straight lines. Obviously we do not need an FNN to accomplish this task. However, it again gives us a tractable example that helps further illustrate the process of generating and optimizing FNNs.

Example 10.3 (Interpolating data points linearly with ReLU *functions).*
假设我们有以下极其简单的数据集，**R**，其中一个自变量（特征）维度和一个因变量（目标）维度由有序对 $(x,\ y)$ 指定，如下所示

$$X = \{(0,0),(\tfrac{1}{2},1),(1,\tfrac{3}{4})\} \tag{10.73}$$

该数据集的定义域 为 $x \in [0,1]$。我们的任务是用两个 ReLU 函数拟合这些数据；这属于插值而非回归（即每两个相邻点之间用一条线段连接），因此应当能够实现零误差拟合。

给定任意两个点，都存在某个 ReLU 函数经过这两点。然而，由于存在一般的平移和垂直位移，这个解并不唯一。这一点可见图 10.21。因此，从一开始我们就应该意识到，如果要获得唯一解，可能需要添加一些约束条件。数值优化问题时，通常会将斜率和平移的初始状态设为随机数。因此，即使没有约束条件，解也可能收敛到某种能够使误差最小化的组合。它并非唯一解这一事实通常无关紧要。然而，在解析地处理这个问题时，我们就没有这样的便利了！

Given that any two points can be fit with a single ReLU function, then for our case (three points) we need two ReLU functions to exactly fit the data. Our first step is to write out an appropriate network that can be used to solve the problem. In Fig. 10.22 we have drawn out such a network. In drawing out this figure we have already imposed one constraint. Note that in layer 2 there is no addition of bias terms. If you take a moment to think about what this does, it forces the result to be a sum of two ReLU functions with no vertical translations; in other words, we have implicitly set b for the final linear composition at node 5 to be zero. This dispenses with the non-uniqueness problem that was illustrated in Fig. 10.21. As we will see, the resulting solution is now a unique one.

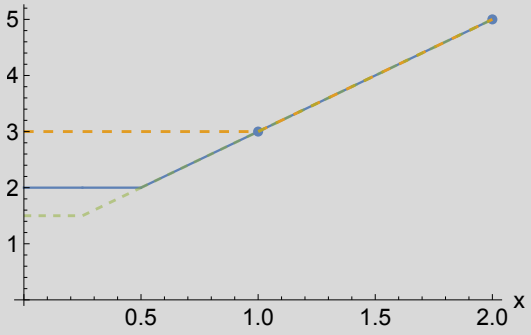


Fig. 10.21: Nonuniqueness in solutions of ReLU functions through two points. Additional constraints are required to obtain a unique solution.

将网络图转化为数学表达式，我们得到如下结果。

$$F(x;w_{1,3},w_{1,4},b_{2,3},b_{2,4},w_{3,5},w_{4,5}) = w_{3,5}\mathrm{ReLU}\,(w_{1,3}x - b_{2,3}) + w_{4,5}\mathrm{ReLU}\,(w_{1,4}x - b_{2,4}) \tag{10.74}$$

We can use the properties of ReLU functions and what we know about the data to impose an additional constraint that makes the system simpler. Recall that the linear combination of two ReLU function with unspecified coefficients can always be re-expressed with a weighting of unity, and and unknown sign function. In other words, we can write the function above as

$$F(x;w_{1,3},w_{1,4},b_{2,3},b_{2,4},\eta_1,\eta_2)=\eta_1\text{ReLU}\left(w_{1,3}x-b_{2,3}\right)+\eta_2\text{ReLU}\left(w_{1,4}x-b_{2,4}\right) \tag{10.75}$$

其中， η_1 和 η_2 要么是 $+1$ ，要么是 -1。我们还可以进一步说明。快速浏览数据后可以发现，所需的曲线不是单调的；也就是说，线段斜率必须在中点处变号。这要求 η_1 和 η_2 符号相反。两者分别取哪个值并不重要。如果施加约束$w^{1,3}$,$w1_{,4}>0$，那么我们可以将解改写为

$$F(x;w_{1,3},w_{1,4},b_{2,3},b_{2,4})=\text{ReLU}\left(w_{1,3}x-b_{2,3}\right)-\text{ReLU}\left(w_{1,4}x-b_{2,4}\right) \tag{10.76}$$

最后，对于这个简单问题，还必须施加最后一个约束。我们知道，两个*ReLU* 函数的原点必须分别平移 $\beta=0$ 和 $\beta=1/2$。这可以从数据本身推断出来。我们希望找到两条能够精确通过三个点的线段。根据数据本身，我们知道两个分段线性函数的定义域（分别记作PW_1 和PW_2）。我们从数据中得到以下信息。

$$PW_1(x):\{x:x\in[0,\tfrac{1}{2}]\} \tag{10.77}$$

$$PW_2(x):\{x:x\in[\tfrac{1}{2},1]\} \tag{10.78}$$

注意，中间点$x=\frac{1}{2}$ 同时属于两个定义域，这正是我们所期望的（这保证了连续性）。根据这些信息，我们通过观察可知，$\beta_{2,\ 3}=0$ 以及$b2_{,\ 4}=\frac{1}{2}$。将其代入我们试图参数化的函数F，可得

$$F(x;w_{1,3},w_{1,4})=\text{ReLU}\left(w_{1,3}x\right)-\text{ReLU}\left(w_{1,4}(x-\tfrac{1}{2})\right) \tag{10.79}$$

Take note that we have rewritten the second ReLU function in the form where the translation appears explicitly. Also recall, we have made two sets of constraints: (1) we have explicitly put in the translation terms (because our problem allowed no other choice), and (2) we have assumed $w_{1,3},w_{1,4}>0$ to get the form given by Eq. 10.79.

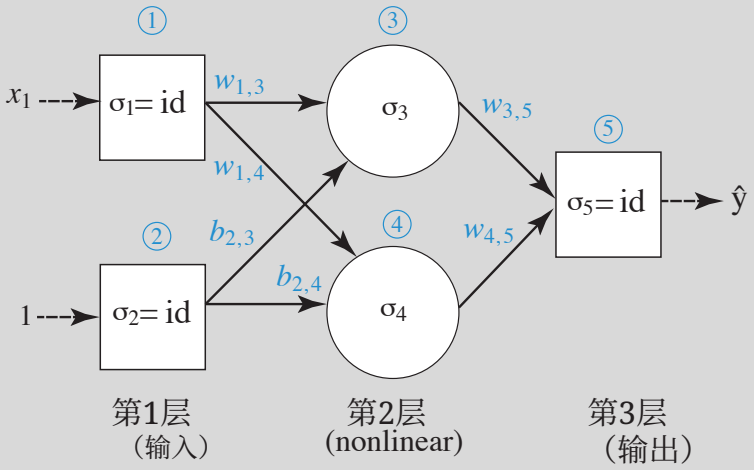


图 10.22：与式 (10.74) 对应的前馈神经网络。这里，$\mathrm{ReLU}_1 = \mathrm{ReLU}(w^{1,3}(x - b^{2,3}))$ 和 $\mathrm{ReLU}_2 = \mathrm{ReLU}(\alpha_2(x - b_2))$

To fit these data, as always, we need *some metric* for defining the error. In this particular case, the data form a C^0 (piecewise linear) function, and the ReLU basis functions are *complete*, and are able to represent the piecewise linear function exactly, provided we can determining the values of $w_{1,3}$ and $w_{1,4}$ explicitly.
Now, we adopt the ℓ_2 metric as a measure the sum of the errors involved in our estimate at the three points of interest, i.e., for N examples, the ℓ_2 error is

$$J(\mathbf{R}) = \sum_{i=1}^{i=N} [y_i - \hat{y}(\mathbf{X}_i; \theta]^2 \tag{10.80}$$

$$\begin{aligned} J = &(0 - [\mathrm{ReLU}\,(w_{1,3} \cdot 0) - \mathrm{ReLU}\,\left(w_{1,4} \cdot (0 - \tfrac{1}{2})\right)])^2 \\ &+ (1 - [\mathrm{ReLU}\,\left(w_{1,3} \cdot \tfrac{1}{2}\right) - \mathrm{ReLU}\,\left(w_{1,4} \cdot (\tfrac{1}{2} - \tfrac{1}{2})\right)])^2 \\ &+ (\tfrac{3}{4} - [\mathrm{ReLU}\,(w_{1,3} \cdot 1) - \mathrm{ReLU}\,\left(w_{1,4} \cdot (1 - \tfrac{1}{2})\right)])^2 \end{aligned} \tag{10.81}$$

利用 ReLU 的性质，可以将该表达式简化为

$$J = (1 - \tfrac{1}{2}w_{1,3})^2 + (\tfrac{3}{4} - w_{1,3} + \tfrac{1}{2}w_{1,4})^2 \tag{10.82}$$

To minimize the residual, we take the derivatives with respect to $w_{1,3}$ and $w_{1,4}$, and set them equal to zero. This gives

$$\frac{\partial J}{\partial w_{1,3}} = \tfrac{5}{2} w_{1,3} - w_{1,4} - \tfrac{5}{2} \tag{10.83}$$

$$\frac{\partial J}{\partial w_{1,4}} = w_{1,3} - \tfrac{1}{2} w_{1,4} - \tfrac{3}{4} \tag{10.84}$$

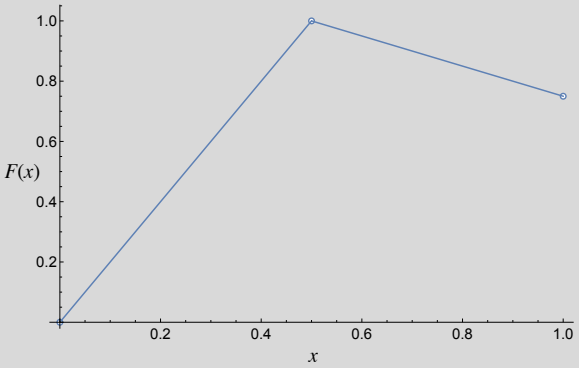


Fig. 10.23: An *extremely simple* one-dimensional data set containing three points.

此时，我们只需要求解以下线性方程组。

$$\tfrac{5}{2} w_{1,3} + w_{1,4} - \tfrac{5}{2} = 0 \tag{10.85}$$

$$w_{1,3} + \tfrac{1}{2} w_{1,4} - \tfrac{3}{4} = 0 \tag{10.86}$$

The solution to this set of equations is given by $w_{1,3} = 2$ and $w_{1,4} = \frac{5}{2}$. This gives the solution

$$\hat{y}(x) = 2\text{ReLU}\,(x) + \tfrac{5}{2}\text{ReLU}\left(x - \tfrac{1}{2}\right) \tag{10.87}$$

It is easy to verify that this is the has a residual error of $J = 0$; the results are plotted in Fig. 10.23. As a final note, the process outlined above is essentially the same if we have many *more* data points than we do ReLU functions. In that event, the optimization process results in the *best possible* fit as measured by the least-squares (L_2) error metric. However, this would be an example of *regression* rather than *interpolation*. So, while it would be a best fit (meaning that the derivatives of the loss function would be identically zero for the minimizing set of parameters), the value of the error itself, J would not be zero. For the case of interpolation, the error, J, is identically zero.

示例 *10.4*（通过反向分析实现异或函数。）

The *XOR* function is a classic example that is often used as an example for machine learning. This function is given by the following binary relation

x_1	x_2	y
0	0	0
0	1	1
0	0	1
1	1	0

Table 10.3: The binary AND function.

In the context of *classification* (which is usually how it is solved), a 2-hidden layer network is required in order to obtain a solution. Even then, the problem is quite under-constrained, and it requires careful analysis to arrive at a correct result. There appear to be existing analytical solutions to optimizing a network describing XOR.

在本例中，我们将寻找一种分几个步骤逐步求解的方法。我们将利用第 10.11.1 节中介绍的脊函数相关知识，为求解方法提供启发。此前，我们研究了以下脊函数。

$$r_2(x_1,x_2) = \text{ReLU}(x_2 - x_1) + \text{ReLU}(x_1 - x_2) \tag{10.88}$$

As a reminder, the plot of this function appears in Fig. 10.24. If one compares this function to the requirements of the XOR function, we find that this function represents *one possible solution* (out of an infinite number of potential solutions) that is composed of only ReLU functions. In fact, this solution is composed of exactly two ReLU functions, which is minimum possible number.

异或函数具有镜像对称性，这是我们在确定解时可以加以利用的一点。这种对称性仅意味着，我们总是可以在解中交换变量x_1和x_2，而结果保持不变。认识到这一点后，便可纳入大量约束。为说明这一点，我们先写出一个具有三个输入值（x_1,x_2,1）、隐藏层中包含两个 ReLU 函数且具有一个输出 $\hat{y}$ 的通用单隐藏层网络的表达式。该网络的一个通用版本如图 10.25 所示。

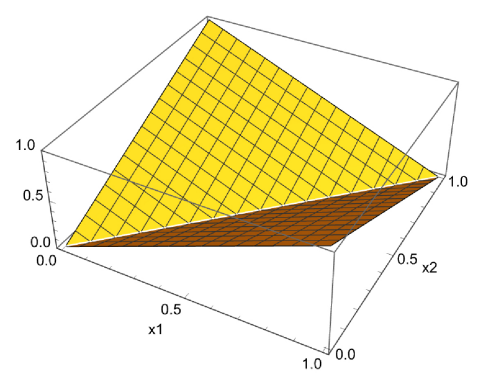


Fig. 10.24: A ridge function.

利用我们对异或函数的了解，可以立即确定$b6,_7 = 0$（这是来自点$x_1 = 0$、$x_2 = 0$，其解为$y = 0$。根据镜像对称性的要求，我们还得到条件$w1,_4 = w2,_5$和$w2,_4 = w2,_5$（这是为了使我们总能互换x_1和x_2,，以及要求$w4,_7 = w5,_7$。由于我们的解是正定的，并且利用我们总能用符号函数替换 ReLU 网络最后一层权重这一性质，可以表述为$w4,_7 = w5,_7 = 1$。仅凭对称性，我们就已经取得了很大帮助！
由此得到网络估计量 $\hat{}y$ 的以下表达式。

$$\hat{y}(x_1,x_2) = \text{ReLU}(w_{1,4}x_1 + w_{2,4}x_2 + b_{3,4}) + \text{ReLU}(w_{2,4}x_1 + w_{1,4}x_2 + b_{3,5}) \tag{10.89}$$

Again, using symmetry, and the solution for the point $x_1 = 0, x_2 = 0$ with $y = 0$, we have that the two bias terms must equal, the value must be negative or zero. To constrain these further, we can evaluate the function $\hat{y}$ at $(x_1,x_2) = (0,1)$, $(1,0)$ and $(1,1)$. We find the following options

$$\begin{aligned} &\text{either} && w_{1,4} + b_{3,4} > 0, \text{ or } w_{2,4} + b_{3,4} > 0 \\ &\text{and} && w_{1,4} + w_{2,4} + b_{3,4} = 0 \end{aligned}$$

Any solution that meets these constraints is valid. The constraints are met by $b_{3,4} = 0$, which is the simplest option.

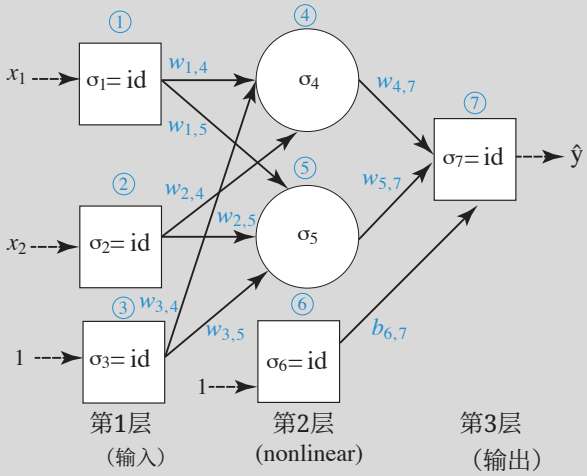


图 10.25：一个通用的单隐藏层网络。

Our estimator is now relatively simple.

$$\hat{y}(x_1,x_2) = \text{ReLU}(w_{1,4}x_1 + w_{2,4}x_2) + \text{ReLU}(w_{2,4}x_1 + w_{1,4}x_2) \tag{10.90}$$

which contains only two unknown parameters. The loss function is given by

$$J(w_{1,4}, w_{2,4}) = [0 - \hat{y}(0,0)]^2 + [1 - \hat{y}(1,0)]^2 + [1 - \hat{y}(0,1)]^2 + [0 - \hat{y}(1,1)]^2 \tag{10.91}$$

其中

$$\hat{y}(0,0) = 0 \tag{10.92}$$
$$\hat{y}(1,0) = \text{ReLU}\,(w_{1,4}) + \text{ReLU}\,(w_{2,4}) = 1 \tag{10.93}$$
$$\hat{y}(0,1) = \text{ReLU}\,(w_{2,4}) + \text{ReLU}\,(w_{1,4}) = 1 \tag{10.94}$$
$$\hat{y}(1,1) = \text{ReLU}\,(w_{1,4} + w_{2,4}) + \text{ReLU}\,(w_{2,4} + w_{1,4}) = 0 \tag{10.95}$$

In this case, we do not necessarily need to take derivatives of the objective function because the solution is so highly constrained (although that process *will* lead to the correct result!). The expression for $haty(1,1)$ clearly indicate that $w_{1,4} = -w_{2,4}$. In addition, either of $\hat{y}(1,0)$ or $\hat{y}(0,1)$ require that $w_{1,4} = 1$ and $w_{2,4} = -1$ (or $w_{1,4} = -1$ and $w_{2,4} = 1$, either choice is a solution). The resulting network can be drawn out as given in Fig. 10.26.

最后补充一点，我们本可以通过在输出之前添加一层，利用这个网络进行分类。这一层会包含一个海维赛德函数，以及一个额外的未知参数（偏置）。这样便会得到一个具有两个隐藏层的网络。不过，既然我们已经得到了一个解，稍加思考就会发现，用偏置为 $-1/2$的海维赛德函数对我们的解进行事后筛选，也能达到同样的效果。这一点也绘于图 10.26 中。

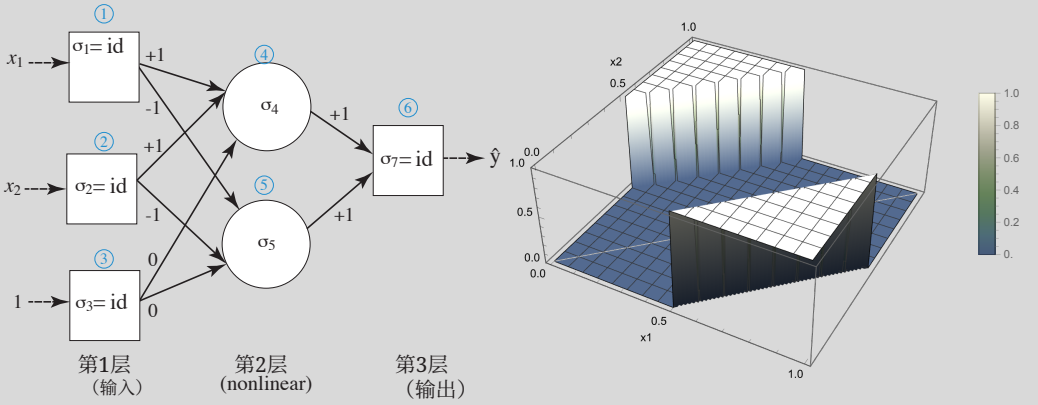


Fig. 10.26: (Left). The XOR network solution. (Right). The solution filtered through the Heaviside function $H(\hat{y} - 1/2)$.

10.13 使用非线性前馈神经网络逼近已知函数

使用简单函数的加权和来近似给定函数的过程属于近似理论。此前介绍的材料中曾简要提及这一点。这里，我们演示如何用分段连续 ReLU 网络近似连续函数。人们可能出于许多原因，想用近似函数来近似给定函数；例如，与原函数相比，对分段线性逼近进行某些变换通常更容易。在这个例子中，我们着重讨论如何用 ReLU 函数拟合多项式。具体来说，我们考察二次函数的拟合。

As the quadratic of interest, take the function

$$y_0(x) = x^2 + 1,\ x \in [0,1] \tag{10.96}$$

这个例子很有意思，因为它还包含一个全局偏置项，$B = 1$，该偏置项使函数沿垂直方向平移 1 个单位。该网络在图 10.28 中以网络图的形式进行了图示。

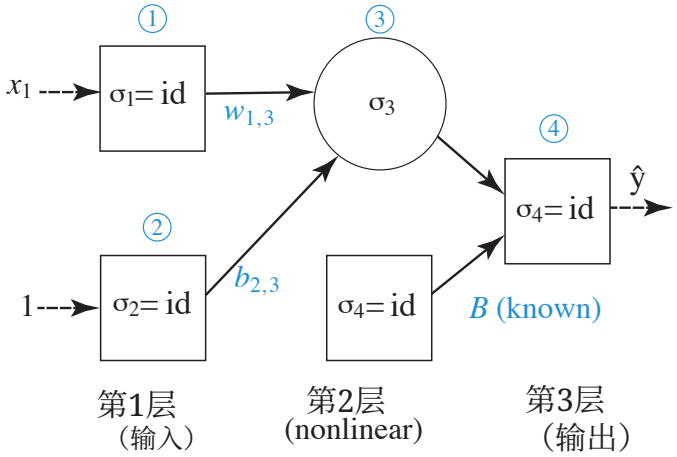


Fig. 10.27 A single-hidden-layer, one-node network for polynomial approximation.

To begin, note that for a single-hidden-layer ReLU network, the bias term can always be handled by adding a bias term to the output node. In other words, we set

$$B(x) = 1,\ x \in [0,1] \tag{10.97}$$

for the output node. Here, we have given this constant a special symbol so that it does not depend on network numbering schemes. From here forward, we *assume* that this can be done for any constants added to a polynomial.

The remaining problem is to fit the homogeneous polynomial

$$y(x) = x^2 \tag{10.98}$$

Using ReLU functions. To outline the process, consider the case where we use a single ReLU function to represent this polynomial. Thus, we are approximating $y(x) = x^2$ by $\hat{y}(x)$, where

$$\hat{y}(x) = \mathrm{ReLU}\,(w_{1,3}x + b_{2,3}) \tag{10.99}$$

Because we are now in the context of continuous functions rather than discrete data points, our error function (or residual), J is now an *integral* rather than a sum

$$J(w_{1,3}, b_{2,3}) = \int_0^1 \left[x^2 - \text{ReLU}\left(w_{1,3}x - b_{2,3}\right)\right]^2 dx \tag{10.100}$$

为帮助解决这个问题，我们注意到由定义域宽度施加的以下约束条件 .

$$b_{2,3} < 1 \tag{10.101}$$

这个约束条件可以防止 ReLU 函数的贡献恒等于零。利用 ReLU 函数的性质，残差J 可以写成以下形式。

$$J(w_{1,3}, b_{2,3}) = \int_0^1 x^4\, dx - \int_{b_{2,3}}^1 2w_{1,3}x^2(x - b_{2,3})\, dx + \int_{b_{2,3}}^1 w_{1,3}^2(x - b_{2,3})^2\, dx \tag{10.102}$$

这里，积分取决于积分域的下界$_{2,3}$。一如既往，在尝试优化之前，我们会对函数 $\hat{}y$ （因此也对损失函数）施加额外约束。再次强调，也可以 不预先添加约束而继续进行。然而，这样会面临多种可能的解，它们取决于参数的线性组合，而这类潜在解的数量可能很快就会变得难以应付。如果能提前施加合理的约束，就可以避免这个问题。不过，我们必须确保所提出的解不会无意中排除“最佳”解。通过检查问题，我们施加以下约束。首先，我们假设由于函数$y(x) = x^2$ 在$x \in [0, 1]$上严格为正，因此解必须为正。这要求$w1_{,3} > 0$。其次，我们假设偏置b_{23} 必须小于或等于零，否则得到的 $\hat{}y$ 预测将是线性的。现在问题表述为

$$\begin{gathered} J(w_{1,3}, b_{2,3}) = \int_0^1 x^4\, dx - \int_{b_{2,3}}^1 2w_{1,3}x^2(x - b_{2,3})\, dx + \int_{b_{2,3}}^1 w_{1,3}^2(x - b_{2,3})^2\, dx \\ \text{Constraints}: \; w_{1,3} > 0, \;\; b_{2,3} \leq 0 \end{gathered} \tag{10.103}$$

融入 g these two constraints, the expression for the loss function be integrated to give the following

$$J(w_{1,3}, b_{2,3}) = \frac{1}{5} - \frac{(b_{2,3} + w_{1,3})^2 \left(b_{2,3}^2 - 2b_{2,3}w_{1,3}(w_{1,3} + 1) + (3 - 2w_{1,3})w_{1,3}^2\right)}{6w_{1,3}^3} \tag{10.104}$$

This is a nice result; our constraints have given us a single expression for the error function rather than a list of cases, and this dramatically simplifies the analysis. Continuing forward, we need to determine the derivatives of the loss function.

$$\frac{\partial J}{\partial w_{1,3}} = \frac{1}{6}\left(\frac{3b_{2,3}^4}{w_{1,3}^4} - \frac{2b_{2,3}^3}{w_{1,3}^2} + 6b_{2,3} + 4w_{1,3} - 3\right) \tag{10.105}$$

$$\frac{\partial J}{\partial b_{2,3}} = -\frac{2b_{2,3}^3}{3w_{1,3}^3} + \frac{b_{2,3}^2}{w_{1,3}} + 2b_{2,3} + w_{1,3} - \frac{2}{3} \tag{10.106}$$

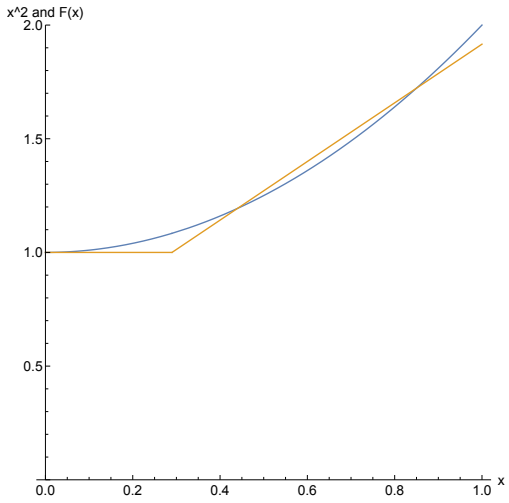


Fig. 10.28 The solution for a single ReLU unit fitting the expression $f(x) = x^2 + 1$. The solution represents the global minimum residual.

Setting the two derivatives equal to zero, and solving simultaneously gives a single solution that satisfies the constraints

$$\frac{3b_{2,3}^4}{w_{1,3}^4} - \frac{2b_{2,3}^3}{w_{1,3}^2} + 6b_{2,3} + 4w_{1,3} - 3 = 0 \tag{10.107}$$

$$-\frac{2b_{2,3}^3}{3w_{1,3}^3} + \frac{b_{2,3}^2}{w_{1,3}} + 2b_{2,3} + w_{1,3} - \frac{2}{3} = 0 \tag{10.108}$$

求解这样一组多项式方程并非易事。一种方法是使用数值方法对系统进行多变量求根。还有一些结构化方法，例如 Gröbner 基方法（Adams and Loustaunau，2022），可以简化此类分析。这里使用符号数学软件包 Mathematica 求得了解析解。这些解是

$$w_{1,3} = \frac{1}{5}(4+\sqrt{6}) \qquad b_{2,3} = \frac{1}{25}(2+3\sqrt{6}) \qquad J = \frac{73-28\sqrt{6}}{3125} \approx 0.00141 \tag{10.109}$$

Example 10.5 (Two-hidden layer approximation). The deeper that networks become, the more challenging it is to solve them analytically. Generally, as ReLU FNNs get deeper, the number of individual regions goes up significantly. Even for networks with a single independent variable and a constant number of nodes in the hidden layers, the rate of increase is is still a power of the number of hidden layers.

在此示例中，我们将使用一个具有 2 个隐藏层的网络来近似一个紧支撑函数。我们感兴趣的函数为

$$y(x) = \begin{cases} 0 & x < 0 \\ \sin[\pi x/2] & 0 \le x \le 2 \\ 0 & x > 2 \end{cases} \tag{10.110}$$

这是一个正函数，由正弦函数的一个波瓣定义，横跨区间 $x \in [0,2]$；不过请注意，我们并不要求解的定义域与之相同。最优近似可能出现在稍大的定义域上，近似的递增部分从稍微位于 $x = 0$ 左侧的位置开始（这就是模型中出现偏置 $b^{2,3}$ 的作用）。这个问题有趣的一点是，它关于位于 $x = 1$处的纵轴对称。和之前一样，提前设定合理的约束条件，有助于我们成功解决这个问题。首先，我们要求解具有对称性。另一个认识是，无论我们的解是什么，它都必定是某种三角形函数。一个 ReLU 函数将拟合 $x = 1$左侧函数的递增部分，另一个 ReLU 函数将拟合 $x = 1$右侧函数的递减部分。这样也就确定了两个 ReLU 函数中一个的平移位置——它必须设在 $x = 1$处。

With these constraints in place, we find that there is another constraint that is needed for consistency with those already defined. The slope of the approximation $\hat{y}(x)$ must be equal in magnitude, but with opposite signs (otherwise the approximator would not be symmetric). A little though will indicate that if the ReLU function approximates $\hat{y}(x)$ using a slope of w, then the second ReLU function must have a slope of $-w$. Finally, there is no bias in the final solution because the both the triangle function approximation and the original function are both compact functions that go to zero at large values of $|x|$.

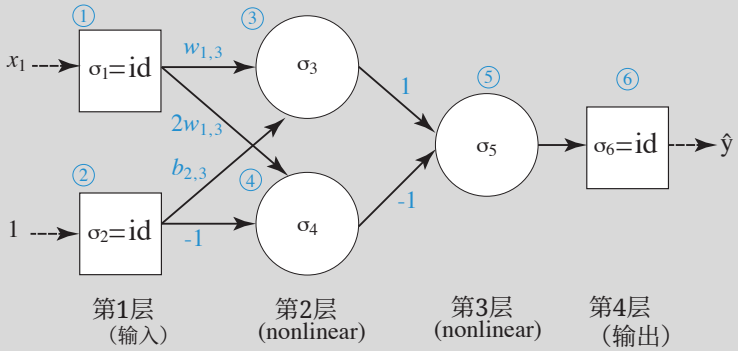


Fig. 10.29: A two-hidden-layer network graph. Constraints have helped fix several of the parameters, and those are displayed over the appropriate link.

The mathematical model associated with this graph is given by

$$\hat{y}(x) = \mathrm{ReLU}\left[\mathrm{ReLU}\left(w_{1,3}x + b_{2,3}\right) - \mathrm{ReLU}\left(2w_{1,3}(x-1)\right)\right] \tag{10.111}$$

Note that this expression assumes that $w_{1,3} > 0$ and that $b_{2,3} > 0$; the minus sign before the second term is to enforce the slope equal to $-w_{1,3}$ for the portion of $\hat{y}(x)$ where $x > 1$. The factor of 2 on the $(1,3)$ link is necessary to attain this result.

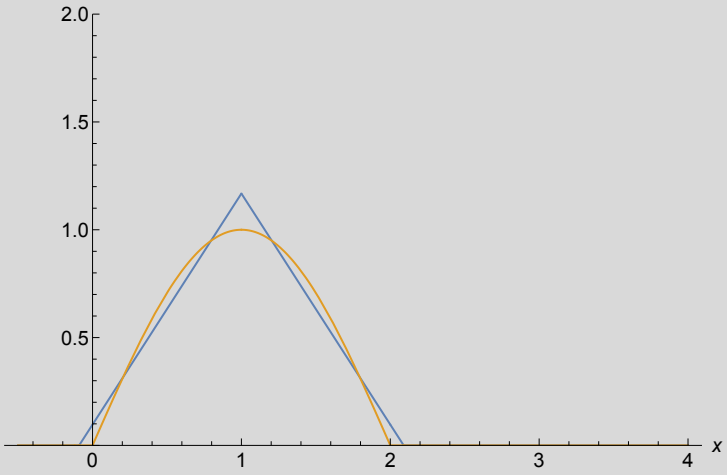


Fig. 10.30: The *relu* FNN fit to the sine function described in the example problem statement.

按照惯例，我们现在构建损失函数。其定义如下：

$$\begin{aligned} J(w_{1,3}, b_{2,3}) &= \int_{-\infty}^{\infty} [\sin(\pi x/2) - \hat{y}(x)]^2 \, dx \\ &= \int_{-\infty}^{\infty} \Big[\sin(\pi x/2) - \text{ReLU}\Big(\text{ReLU}\left(w_{1,3} x + b_{2,3}\right) - \text{ReLU}\left(2 w_{1,3}[x-1]\right) \Big) \Big]^2 \, dx \\ &\text{Constraints}: \quad w_{1,3} > 0, \; b_{2,3} > 0 \end{aligned} \tag{10.112}$$

The integral bounds are set to $x \in (-\infty, \infty)$ here because both the function and its approximation are compact, but we do not know the exact domain of the approximating function yet (this is part of the problem we are solving!) This domain accomplishes the integration without needing to know the exact bounds. Another tip–differentiation commutes with integration here. Thus, when finding the minimum, it is easier to differentiate the integral form and then integrate. Using that approach, we find the following result for the two derivatives.

$$\frac{\partial J}{\partial w_{1,3}} = \frac{4\pi^2 b_{2,3}^3 w_{1,2} + 12\pi^2 b_{2,3}^2 w_{1,2} + 12\pi^2 b_{2,3} w_{1,2} - 24\pi b_{2,3} + 4\pi^2 w_{1,2} - 48}{3\pi^2} \tag{10.113}$$

$$\frac{\partial J}{\partial b_{2,3}} = \frac{6\pi^2 b_{2,3}^2 w_{1,2}^2 + 12\pi^2 b_{2,3} w_{1,2}^2 + 6\pi^2 w_{1,2}^2 - 24\pi w_{1,2}}{3\pi^2} \tag{10.114}$$

As before, we solve this set of equations with the derivative equal to zero, and solve them analytically using symbolic mathematics software. The result is

$$w_{1,3} = \frac{4\pi}{36 - 36\pi + 9\pi^2} \tag{10.115}$$

$$b_{2,3} = \frac{-6 + 2\pi}{\pi} \tag{10.116}$$

$$\begin{aligned} J(w_{1,3}, b_{2,3}) = \frac{1}{3\pi^2}\Bigg(& 3\pi^2 + \frac{32\pi^4}{(36 - 36\pi + 9\pi^2)^2} + \frac{96\pi^3(2\pi - 6)}{(36 - 36\pi + 9\pi^2)^2} + \frac{96\pi^2(2\pi - 6)^2}{(36 - 36\pi + 9\pi^2)^2} \\ & + \frac{32\pi(2\pi - 6)^3}{(36 - 36\pi + 9\pi^2)^2} - \frac{192\pi}{36 - 36\pi + 9\pi^2} - \frac{96\pi(2\pi - 6)}{36 - 36\pi + 9\pi^2} \Bigg) \\ & \approx 0.00861 \end{aligned} \tag{10.117}$$

习题

1. Prove by constructing one representative example that the sum of three ReLU functions

$$f(x)=\sum_{i=1}^{3}\mathrm{ReLU}_i(x+b_i) \tag{10.118}$$

defined on $x\in[0,1]$ with biases of b_i defined by $-1<b_3<b_2<b_1<0$ yield a continuous function with three segments.

2. Prove by constructing a representative example that two ReLU functions in two independent variables of the form

$$f(x,y)=ReLU(a_1x+b_1)+ReLU(a_2y+b_2)$$

结果是一个连续函数。示例中定义了多少个互不相连的区域？在 $(x,\ y)$ 平面上

$$f(x,y)=ReLU(x-1)+ReLU(2y-2)$$

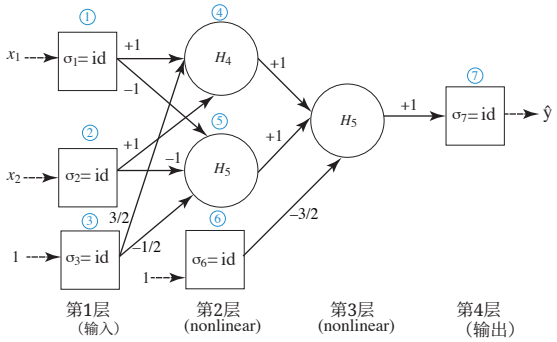
3. 定义（但不求解）一个单隐层 ReLU 前馈神经网络的示例，该网络可用于估计曲线 $f(x)=$正弦(πx)，其中 $x\in[0,1]$。假设隐藏层中只有一个节点。首先写出该网络的图形式，并使用正文中介绍的简化约定标注连接。完成图后，写出该图对应的数学表达式。

4. 定义（但不求解）一个单隐层 ReLU 前馈神经网络的示例，该网络可用于估计曲线 $f(x)=$正弦(πx)，其中 $x\in[0,1]$。假设隐藏层中有两个节点。首先写出该网络的图形式，并使用正文中介绍的简化约定标注连接。完成图后，写出该图对应的数学表达式。

5. Draw the graph for a ReLU FNN that contains 2 hidden layers with two independent variables (x_1,x_2) and two nodes in each hidden layer. Assume that the output is a single (weighted) scalar. Assume that there is a bias term that enters at the input, and at the first hidden layer. First write out the graph form of the network, and label the connections using the simplified conventions presented in the text. After the graph is complete, write out the appropriate mathematical expression for the graph.

6. For the network graph illustrated below (Fig. 10.31, write out the mathematical expression for the network (note that here, H indicates a *Heaviside* function rather than a ReLU function). Once you have that written out, verify that the network produces behavior equivalent to the XOR function.

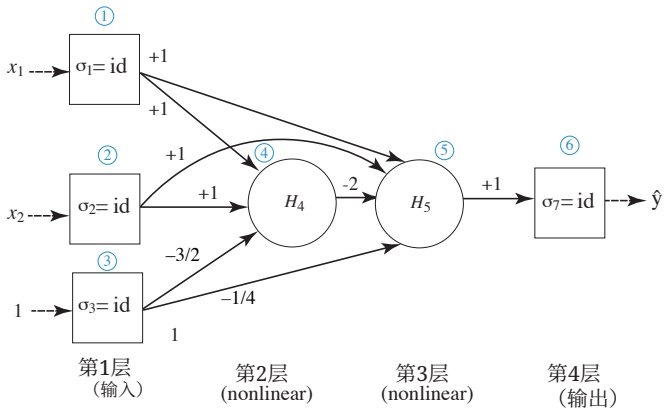
Fig. 10.31 Problem 6.



7. Skip layer networks are still feedforward networks, but they are not simply connected (this term indicates that every node is connected to every other node only in the previous and following layers) but *multiply connected*. In the literature, these are sometimes called skip layers. In applications, these kinds of networks are sometimes used to help avoid the "vanishing gradients" problem in numerical backpropagation.

图 10.32 所示的图是一个跳跃连接前馈神经网络的示例。对于图中所示的网络图，写出该网络的数学表达式（请注意，此处的H表示*Heaviside*函数，而不是 ReLU 函数）。写出表达式后，验证该网络产生的行为等同于异或函数。用一两句话解释你认为此处所示的网络与问题 6 中的网络为何表示相同的函数。

Fig. 10.32 Problem 7.



8. 如简介中所述，人工神经网络中可以使用许多不同的激活函数，尤其是在网络不是很深的情况下。逻辑函数是一个性质良好的函数（即它具有各阶导数），其行为类似于光滑版的海维赛德函数。一般而言，该函数表示为

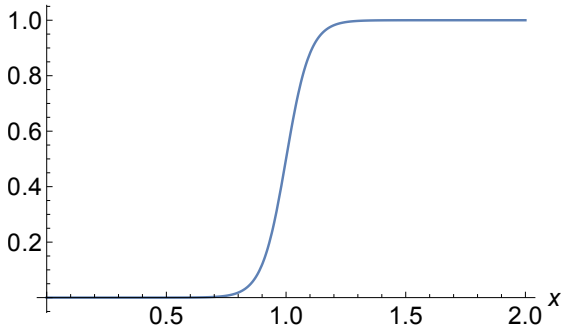
$$\mathrm{Lo}(x) = \frac{1}{1+\exp[-k(x+b)]} \tag{10.119}$$

where b translates the function, and the parameter k changes the slope of the function. To get a Heaviside like function, we can set k to a large value like $k = 20$.

$$\mathrm{Lo}(x) = \frac{1}{1+\exp[-20(x+b)]} \tag{10.120}$$

Lo(x) 的图像如图 10.33 所示。

图 10.33 问题 8 中使用的逻辑激活函数。



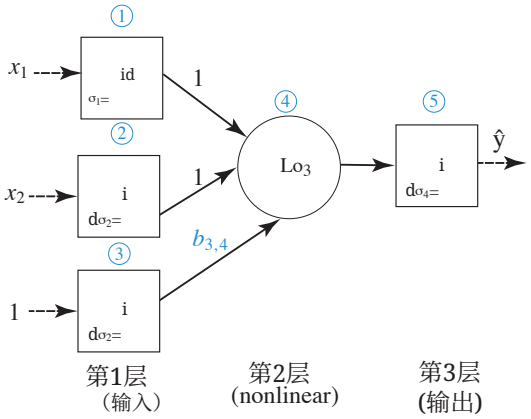
Using this function, write out the mathematical expression for $\hat{y}$ for the network illustrated in Fig. 10.34. Then, determine the L_2 loss function using the four discrete values associated with the AND function. As a reminder, the AND function is defined as given in Table 10.4.

x_1	x_2	y
0	0	0
0	1	0
1	0	0
1	1	1

Table 10.4: The binary AND function.

Determine the value of $b_{3,4}$ by optimizing the loss function for the variable $b_{3,4}$.

Fig. 10.34 Problem 8 A network for the AND function.



9. Following the analysis in Example 10.5, determine an approximation for the function

$$y(x) = \begin{cases} 0 & x < 0 \\ \sin[\pi x/2] & 0 \le x \le 2 \\ 0 & x > 2 \end{cases} \tag{10.121}$$

对问题施加以下附加约束：$b2,3 = 0$。回想一下，我们要求解保留原问题的对称性，因此$x = 1$左侧的斜率仍然是

并且在$x = 1$的右侧，大小相同，但符号相反。这将产生一个单个未知参数。你可以使用符号计算软件来完成这个问题，但也可以在不借助此类工具的情况下进行优化。

与例题中得到的解相比，这个解有何不同？误差项相差百分之多少？计算百分比时，请用两个误差的平均值作分母。

10. 在尝试求解复杂问题的解析解时，利用对称性可能是一个非常有效的方法。例如，考虑以下函数

$$y(x) = \begin{cases} 0 & x < 0 \\ \sin[\pi x/2] & 0 \le x \le 4 \\ 0 & x > 4 \end{cases} \tag{10.122}$$

至少需要三个 ReLU 单元才能表示这个函数，因此这个问题比较复杂。考虑利用对称性来减轻该网络的计算负担，请完成以下任务。

a. 不进行任何 计算，手动画出函数$y(x)$的草图。将函数按以下方式划分为三个定义域：(1) $x < 1$，(2) $1 \le x \le 3$，以及 (3) $x > 3$。注意，第一个和第二个区间都是无界的！现在，在这三个区间上画出该函数的分段线性连续近似，并使其保持关于$x = 2$的对称性。仔细确定斜率和偏置项的约束。画出一个能够近似此函数的前馈神经网络（FNN）的图。

b. 现在，重复这一过程，但这次求一个由四个区间组成的解，区间分别为：(1) $x < 1$，(2)$1 \le x \le 2$，(3) $2 \le x \le 3$，以及 (3) $x > 3$。

例10.5和问题10的解包含了我们所需的全部信息；若运用对称性的思想，就能构造出一个四单元 ReLU 网络的最优解。请仔细确定斜率和偏置项所受的约束。画出能够近似此函数的前馈神经网络图。

参考文献

Adams, W.W., Loustaunau, P.: 《Gr¨obner 基入门》，第3卷。美国数学学会（2022） Antman, S.S.: 《弦的大幅振动方程》。《美国数学月刊》87(5), 359–370（1980） Argentina, M., Clerc, M., Rojas, R., Tirapegui: 《一维 Cahn-Hilliard 模型的粗化动力学》。Physical Review E 71(4), 046210（2005） Bellman, R.E.: 《动态规划》。兰德公司研究报告，普林斯顿大学出版社（1957） Bengio, Y., Delalleau: 《《深度架构的表达能力》。载于：算法学习理论：第22届国际会议，ALT 2011，芬兰埃斯波，2011年10月5日至7日。论文集22。第18–36页。施普林格（2011） Bluman, G.W., Cheviakov, A.F., Anco, S.C.: 《对偏微分方程应用对称方法》，第168卷。施普林格（2010） Bord´e, C.J.: 《国际单位制的基本单位、基本常数与现代量子物理学》。《英国皇家学会哲学汇刊A辑：数学、物理与工程科学》363(1834), 2177–2201（2005） Bridges, T.J., Derks, G., Gottwald, G.: 《五阶 KdV 方程孤立波的稳定性与不稳定性：一个数值框架》。Physica D: Nonlinear Phenomena 172(1-4), 190–216（2002） Chui, C.K., Li, X., Mhaskar, H.N.: 《用于局部近似的神经网络》。《计算数学》63(208), 607–623（1994） Churchill, R.V.: 《运算数学》。麦格劳-希尔（1972） Clausen, T.: 《关于当形如$y = 1 + \frac{\alpha}{1} \cdot \frac{\beta}{\gamma} x + \frac{\alpha \cdot \alpha}{1 \cdot 2} \cdot \frac{\beta \cdot \beta}{\gamma \cdot \gamma + 1} x^2 +$等的级数是形如$1 + \frac{\alpha'}{1} \cdot \frac{\beta'}{\gamma'} \cdot \frac{\delta'}{\varepsilon'} x 1 + \frac{\alpha'}{1} \cdot \frac{\beta'}{\gamma'} \cdot \frac{\delta'}{\varepsilon'} x + \frac{\alpha \cdot \alpha}{1 \cdot 2} \cdot \frac{\beta \cdot \beta}{\gamma \cdot \gamma + 1} \cdot \frac{\delta' \cdot \delta' + 1}{\varepsilon' \cdot \varepsilon' + 1} x^2 +$等的平方时的情形》。纯粹与应用数学杂志3, 89–91（1828） Courant, R., Hilbert, D.: 《数学物理方法》，第一卷。John Wiley & Sons，纽约（1953） Courant, R., Robbins, H.: 《什么是数学？：思想与方法的入门》。I. Stewart 修订，牛津大学出版社（美国），第二版（1996） Craig, R.T.: 《现代数学原理》。普伦蒂斯霍尔，新泽西州恩格尔伍德克利夫斯（1969） Cybenko, G.: 《用 Sigmoid 函数叠加进行近似》。《控制、信号与系统数学》2(4), 303–314（1989） Danckwerts, P.: 《通过同时扩散和化学反应吸收至各种形状的粒子及下落液滴中》。《法拉第学会汇刊》47, 1014–1023（1951） Dang, Q.A., Ehrhardt, M.: 《关于狄拉克δ序列及其生成函数》。《应用数学快报》25(12), 2385–2390（2012） Daubechies, I., DeVore, R., Foucart, S., Hanin, B., Petrova, G.: 《非线性近似与（深度）ReLU网络》。《构造性近似》55(1), 127–172（2022）

De Winter, J., Wagemans, J.：《Attneave的睡猫苏醒：基于轮廓直线版本的日常物体识别》。《知觉》37(2), 245–270 (2008) DeVore, R.A.：非线性近似。《数值学报》7, 51–150 (1998) Ewald, W.：一阶逻辑的涌现。载于：Zalta, E.N.（编）《斯坦福哲学百科全书》（2019年春季版）。https://plato.stanford.edu/entries/logic-firstorder-e mergence/ (2022) Ferraro, G., 等：从19世纪20年代初以前的级数理论兴起与发展。施普林格 (2008) Ferreir´os, J.：集合论的早期发展。载于：Zalta, E.N.（编）《斯坦福哲学百科全书》（2022年春季版）。https://plato.stanford.edu/archives/spr2022/entr ies/settheory-early/ (2022) Ford, A., Peat, F.D.：语言在科学中的作用。《物理学基础》18(12), 1233–1242 (1988) Franklin, J.：猜测科学：帕斯卡之前的证据与概率。约翰·霍普金斯大学出版社 (2001) Glorot, X., Bengio, Y.：理解深度前馈神经网络训练的难点。载于：第十三届国际人工智能与统计学会议论文集，第249–256页。JMLR研讨会与会议论文集 (2010) Gluck, P.：关于软弹簧和弹簧玩具的项目。《物理教育》45(2), 178 (2010) Gowers, T., Barrow-Green, J., Leader, I.：《普林斯顿数学指南》。普林斯顿大学出版社 (2008) Greer, J.B., Bertozzi, A.L., Sapiro, G.：一般几何空间上的四阶偏微分方程。《计算物理学杂志》216(1), 216–246 (2006) Haeffele, B.D., Vidal, R.：神经网络训练中的全局最优性。载于：IEEE计算机视觉与模式识别会议论文集，第7331–7339页 (2017) Hald, A.：《1750年以前的概率与统计史及其应用》。John Wiley & Sons，霍博肯，新泽西州 (2003) Hanin, B.：利用有界宽度和ReLU激活函数的深度神经网络实现通用函数逼近。《数学》7(10), 992 (2019) Hardy, G.H., Rogosinski, W.W.：《傅里叶级数》。多佛出版社：米尼奥拉，纽约州 (1999) Hay, J.C., Lynch, B.E., Smith, D.R.：《Mark I 感知机操作手册》。技术报告，康奈尔航空实验室公司，布法罗，纽约州 (1960) He, J., Li, L., Xu, J.：从层次基视角看ReLU深度神经网络。《计算机与应用数学》120, 105–114 (2022) Hopfield, J.J.：具有涌现集体计算能力的神经网络与物理系统。《美国国家科学院院刊》79(8), 2554–2558 (1982) Hopfield, J.J.：具有分级响应的神经元具有与双态神经元类似的集体计算特性。《美国国家科学院院刊》81(10), 3088–3092 (1984) Jordan, C.：《傅里叶级数论述》。《法国科学院院报》，巴黎92, 228–230 (1881) Kalnins, E.G., Kress, J.M., Miller, W.：《变量分离与超可积性》。IOP出版社 (2018) K¨ohler, W., Held, R.：模式视觉的皮层相关物。《科学》110(2860), 414–419 (1949) Krause, A.L., Burton, A.M., Fadai, N.T., Van Gorder, R.A.：球面上反应–平流–扩散系统中的涌现结构。《物理评论E》97, 042215 (2018年4月)，URL https://link.aps .org/doi/10.1103/PhysRevE.97.042215 Krishnachandran, V.：微分方程：历史回顾。arXiv预印本 arXiv:2012.06938 (2020) Lam, R., Vlachos, D.：薄膜外延生长的多尺度模型：生长模式转变。《物理评论B》64(3), 035401 (2001) Lanczos, C.：《傅里叶级数论述》。奥利弗与博伊德出版社 (1966)

拉普拉斯（侯爵），P.S.：《概率哲学论》。Bachelier，巴黎 (1825)，第六版法文版的英文译本：Truscott FW, Emory FL，《概率哲学论》。John Wiley & Sons/Chapman and Hall，伦敦 (1902) LeCun, Y.：用于不对称阈值网络的学习程序。Cognitiva 85论文集，第599–604页 (1985) Lewicki, M.S., Sejnowski, T.J.：学习过完备表示。《神经计算》12(2), 337–365 (2000) Li, T.Y., Yorke, J.A.：周期三蕴含混沌。《美国数学月刊》82, 985–992 (1975) Lighthill, M.：《傅里叶级数》（第六版）。剑桥大学出版社：伦敦，英国 (1970) Linnainmaa, S.：将算法的累积舍入误差表示为局部舍入误差的泰勒展开。博士论文，硕士论文（芬兰语），赫尔辛基大学 (1970) Liouville, J.：关于物质点运动方程可积分的若干特殊情形。《纯粹与应用数学杂志》，第345–378页 (1846) ˇLjubarskii, J.I.：有限区间上的卷积算子。《苏联消息报数学辑》11(3), 583 (1977) McClelland, J.L., Rumelhart, D.E., PDP研究组：《并行分布式处理：认知微观结构探索：基础》。MIT出版社，剑桥，马萨诸塞州 (1986) McClelland, J.L., Rumelhart, D.E., PDP研究组：《并行分布式处理：认知微观结构探索》第2卷：心理与生物模型。MIT出版社，剑桥，马萨诸塞州 (1987) McCullough, E., Pitts, W.：神经活动中内在思想的逻辑演算。《数学生物物理学通报》5, 115–133 (1943) Mendie, J.G., Udofia, S.N.：语言与现实观念。《国际人文与创新杂志》（IJHI）2(4), 121–124 (2019) Miller, Jr, W.：《对称性与变量分离》。Addison-Wesley出版社，公司，雷丁，马萨诸塞州 (1977) Minsky, M.L., Papert, S.A.：《感知机：计算几何导论》(1969) Minsky, M.L., Papert, S.A.：《感知机：扩充版》(1988) Montufar, G.F., Pascanu, R., Cho, K., Bengio, Y.：深度神经网络线性区域的数量。《神经信息处理系统进展》27 (2014) Munkres, J.：《拓扑学》（第二版）。培生，埃塞克斯，英国 (2014) O'Connor, J., Robertson, E.：Colin Maclaurin——传记。圣安德鲁斯大学数学与统计学院，苏格兰，https://mathshistory.st-andrews.ac.uk/Biographies/M aclaurin/ (2017年5月)，访问日期：2021年10月15日 Olazaran, M.：关于感知机争论官方历史的社会学研究。《科学社会研究》26(3), 611–659 (1996) Olver, P.J.：《偏微分方程导论》。施普林格 (2014) Papin, L.：这不是一个宇宙：隐喻、语言与表征。《PMLA》107(5), 1253–1265 (1992) Pascanu, R., Montufar, G., Bengio, Y.：分段线性激活下深度前馈网络响应区域的数量。arXiv预印本 arXiv:1312.6098v5 (2014) Petersen, P., Voigtlaender, F.：利用深度ReLU神经网络对分段光滑函数进行最优近似。《神经网络》108, 296–330 (2018) Phelan, S.E.：复杂性科学究竟是什么？《涌现》3(1), 120–136 (2001) Piccinini, G.：第一个心智与大脑计算理论：细读McCullough和Pitts的"神经活动中内在思想的逻辑演算"。《Synthese》141, 175–215 (2004) Pinkus, A.：神经网络多层感知机模型的逼近理论。《数值学报》8, 143–195 (1999)

Playfair, J.：《几何元素：包含欧几里得《几何原本》前六卷，并附两卷立体几何，另增平面三角学与球面三角学元素》。Bell & Bradfute and G. G. & J. Robinson，伦敦（1795年）Plonka, G., Riebe, Y., Kolomoitsev, Y.：一维 ReLU 神经网络模型的样条表征与冗余。arXiv 预印本 arXiv:2207.14609（2022年）Poggio, T., Mhaskar, H., Rosasco, L., Miranda, B., Liao, Q.：为什么以及何时深层网络（而非浅层网络）能够避免维数灾难：综述。《国际自动化与计算杂志》14(5)，503–519（2017年）Post, E.L.：广义微分。《美国数学会汇刊》32(4)，723–781（1930年）Protter, M.H., Charles Jr, B., 等：实分析入门。施普林格科学与商业媒体（2012年）Raghu, M., Unterthiner, T., Kornblith, S., Zhang, C., Dosovitskiy, A.：视觉变换器的视觉方式是否与卷积神经网络相同？《神经信息处理系统进展》34，12116–12128（2021年）Recanatesi, S., Farrell, M., Advani, M., Moore, T., Lajoie, G., Shea-Brown, E.：深度神经网络中的维数压缩与扩展。arXiv 预印本 arXiv:1906.00443（2019年）Reynolds, O.：XXIX. 关于决定水流是呈直线流动还是呈曲线流动的条件，以及平行渠道中阻力定律的实验研究。《伦敦皇家学会哲学汇刊》174，935–982（1883年）Rosen, K.H.：《离散数学及其应用》。McGraw-Hill，纽约，第八版（2019年）Rumelhart, D.E., Hinton, G.E., Williams, R.J.：通过反向传播误差学习表征。《自然》323(6088)，533–536（1986年）Sasser, J.E.：常微分方程史：最初的一百年。《中西部数学史学会论文集》1（1992年）Snow, J.E.：关于实数与连续介质的观点。《现代逻辑评论》9(1-2)，95–113（2003年）Spiegel, M.R.：《拉普拉斯变换》。McGraw-Hill，纽约（1965年）Telgarsky, M.：深层前馈网络的表征优势。arXiv 预印本 arXiv:1509.08101（2015年）Torfs, K., Panis, S., Wagemans, J.：碎片化物体轮廓的识别：不同组成过程之间的动态交互。《视觉认知》18(8)，1133–1164（2010年）Werbos, P.：超越回归：行为科学中的预测与分析新工具。博士论文，哈佛大学应用数学委员会，马萨诸塞州剑桥（1974年）Weyl, H.：《对称性》。普林斯顿大学出版社（1952年）Yarotsky, D.：深度 ReLU 网络近似的误差界。《神经网络》94，103–114（2017年）Yong, D.：《弦、链与绳索》。《SIAM评论》48(4)，771–781（2006年）You, Y.L., Kaveh, M.：用于噪声去除的四阶偏微分方程。《IEEE图像处理汇刊》9(10)，1723–1730（2000年）Zhang, C., Bengio, S., Hardt, M., Recht, B., Vinyals, O.：理解深度学习（仍然）需要重新思考泛化。《美国计算机协会通讯》64(3)，107–115（2021年）Zygmund, A.：《三角级数》（1935年版重印）。多佛出版社：纽约州米尼奥拉（1955年）

索引

解析函数，47，82 辅助条件 柯西条件，257 狄利克雷边界条件，256 第一类，256 初始条件，256 诺伊曼边界条件，257 周期条件，257 罗宾条件，257 第二类，257 指定边界导数，257 指定边界值，256 第三类，257 架构，406 辐角，37 人工智能，405 人工神经网络，405 架构，406 前馈神经网络（FNN），405 历史，414 公理，30，242 基函数，170 比较，200

复杂性，12；其性质，12；守恒公理，243；能量守恒公理，242；质量守恒公理，242；动量守恒公理，242；文字表述，244；逐点收敛，87；序列收敛，75；级数收敛，77；一致收敛，88；卷积，312；卷积，329，333；狄拉克δ函数，212，217；δ序列，222；傅里叶级数，228；恒等式，230；性质，223；筛选性质，224；因变量，102；导数定义，102；广义导数，213；阶，102；偏导数，237；决定论，10；不连续点；可去间断点，50；定义域，101；涌现行为，13；图灵模型，14；等价类，49

猫 Attneave，18；Rosenblueth 和 Wiener，1；柯西序列，75；混沌系统，13；朱利亚集，14；复数模，37；辐角，37；布哂维奇积分，369；模，37；留数定理，366

前馈神经网络（FNN）；函数近似，447；近似理论，434；定义，405；线性回归，419；非线性，423；非线性回归，437；原则，416；通量，236；复傅里叶级数，313；傅里叶级数，155，164；正弦级数，168；基函数，170，179；区间变换，185；余弦级数，174；狄利克雷条件，180；偶函数和奇函数，188；周期延拓，197；谱，181；关于零对称的区间，188；傅里叶变换，320；卷积，329，333；推导，321；导数，325；有限傅里叶变换，317；傅里叶积分，321；热/扩散问题，326；恒等式，324；核函数，363；线性，317；符号表示，323；限制，323；缩放函数，324；平移函数，324；变换表，348；波动方程，346；傅里叶变换的伸展，337；一阶波动方程，326；
Fubini 定理，389；函数；解析，47，82；连续，43；等价类，49；偶函数和奇函数，188
  functional, 91
  gamma function, 213
  Heaviside, 217
  piecewise continuous, 45
  step, 217
  transcendental, 104
functional, 91
  kernel, 94

广义函数，212 广义导数，213 狄拉克δ函数，212、217 几何 一维，244 欧几里得，30 过程几何，244 热传导/扩散方程 方差递减性质，281 分离变量法，282 海维赛德函数，217 导数，228 高阶偏微分方程 五阶，359 四阶，358 三阶，358 超几何函数 广义，357 当且仅当（iff），75 自变量，101 积分核，94 勒贝格，70 黎曼，67 积分变换，312 积分，66 国际单位制（SI），9 区间，33 闭区间，33 半开区间，33 开区间，33 无界，33 孤立系统，243

朱利亚集，14

核函数，94、363 热量，
336 Korteweg-德弗里斯方程，
9
35 语言，2 形式语言，2 自
然语言，2 拉普拉斯方程 最
大值原理，301 分离变量法，
303 拉普拉斯变换，361
Post反演定理，369 布哥维奇
积分，369 定义，364 狄拉克
δ函数，371 不连续函数，
372 海维赛德函数，371 逆，
368 核函数，363 勒贝格积分，
70 连接，407 机器学习，
405 数学，3 作为一种语言，
3 模型，4、7 混沌的，13 定
义，4 涌现行为，13 大量变
量，12 规律性，12 相似性，
6 系统，4 建模，4、7 抽象，
7 算法思维，8 复杂性，12
计算，8 决定论，10 奥卡姆
剃刀，15 过拟合，16 简约性，
15 规律性，12 修订，9 简洁
性，15 系统，4 单位，9

modulus, 37

necessary and sufficient, 76
network
  artificial neural, 405
  shallow, 406
  architecture, 406
  backpropagation, 410
  deep, 406
  fully connected feedforward (FNN), 405
  graphs, 411
  link, 407
  loss function, 409
  node, 406
  optimization, 410
  simplified notation, 413
  structure-preserving notation, 411
  weight, 407
node, 406
  bias, 407
  destination, 407
  number, 406
  source, 407
numbers
  complex, 35
  integers, 28
  irrational, 28
  natural, 28
  rational, 28
  real, 28

Occam’s razor, 15
ordinary differential equation, 102
  homogeneous, 103
  boundary condition, 104
  first order, 105
  homogeneous solution, 141
  initial condition, 104
  integrable, 108
  linear, 103
  nonlinear, 103
  second order, 125
  second-order homogeneous, 131
  second-order nonhomogeneous, 141
  separable, 111
  variation of parameters, 141
overfitting, 16

简约性，15 最大限度的不作承诺，16 偏微分方程，附加条件，236 椭圆型，239 双曲型，239 阶，238 抛物型，239 适定，237 幂级数，79

值域，101 ReLU 函数，424 代数性质，425 微积分，431 复合，430 多维，433 脊函数，436 数乘，428 平移，426 脊函数，436 二阶偏微分方程 热传导/扩散方程，280 拉普拉斯方程，301 线性叠加，307 波动方程，291 分离变量法，277 应用条件，278 序列 柯西，75 柯西收敛准则，75 收敛，75 发散，75 有限，73 无限，73 函数序列，74 级数，76 比较判别法，78 收敛，77 定义，76 无限，76 积分判别法，77 莱布尼茨收敛判别法，79 极限判别法，78

部分和，76 部分和，76 逐点收敛，87 幂，79 比值判别法，78 根值判别法，78 泰勒，82 一致收敛，88 集合，26 笛卡尔积，27 补集，27 元素，26 子集，26 并集，27 集合 区间，32 无界区间，33 简洁性，15 平方可积，323 阶跃函数恒等式，230 阶跃函数，217 导数，228 叠加，307 线性，307 叠加，线性，307 系统，4 理想，243 孤立，243 泰勒级数，81，82 单位，9 基本单位，9 全集，26 变量 因变量，102 广延，235 自变量，101 强度，235, 244 波动方程 能量守恒，29 分离变量法，293 权重，407