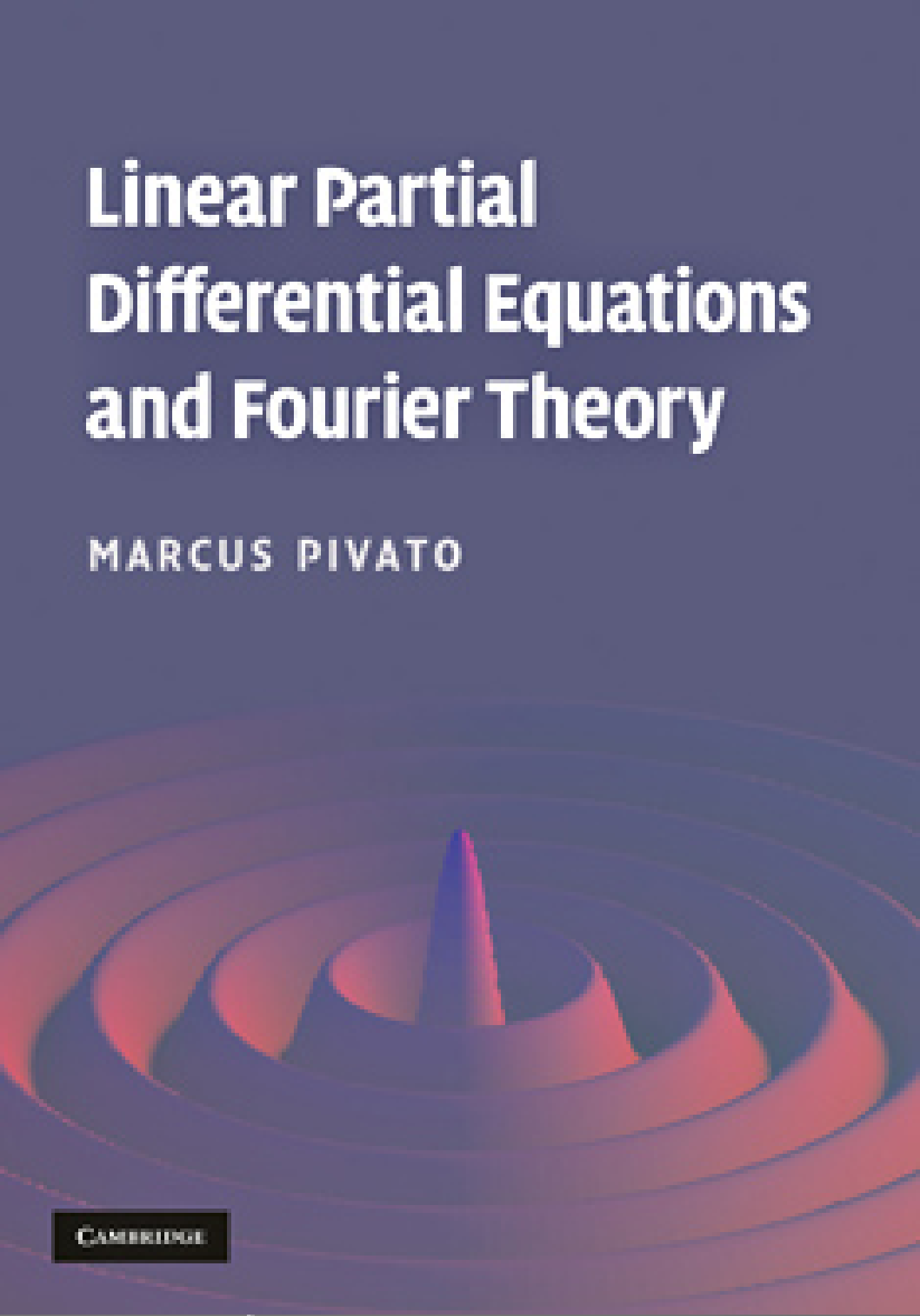
Linear Partial Differential Equations and Fourier Theory

MARCUS PIVATO



CAMBRIDGE

CAMBRIDGE www.cambridge.org/9780521199704

This page intentionally left blank

LINEAR PARTIAL DIFFERENTIAL EQUATIONS AND FOURIER THEORY

Do you want a rigorous book that remembers where PDEs come from and what they look like? This highly visual introduction to linear PDEs and initial/boundary value problems connects the theory to physical reality, all the time providing a rigorous mathematical foundation for all solution methods.

Readers are gradually introduced to abstraction – the most powerful tool for solving problems – rather than simply drilled in the practice of imitating solutions to given examples. The book is therefore ideal for students in mathematics and physics who require a more theoretical treatment than is given in most introductory texts.

Also designed with lecturers in mind, the fully modular presentation is easily adapted to a course of one-hour lectures, and a suggested 12-week syllabus is included to aid planning. Downloadable files for the hundreds of figures, hundreds of challenging exercises, and practice problems that appear in the book are available online, as are solutions.

MARCUS PIVATO is Associate Professor in the Department of Mathematics at Trent University in Peterborough, Ontario.

LINEAR PARTIAL DIFFERENTIAL EQUATIONS AND FOURIER THEORY

MARCUS PIVATO
Trent University

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore,
São Paulo, Delhi, Dubai, Tokyo

Cambridge University Press
The Edinburgh Building, Cambridge CB2 8RU, UK

Published in the United States of America by Cambridge University Press, New York

www.cambridge.org
Information on this title: www.cambridge.org/9780521199704

First published in print format 2010

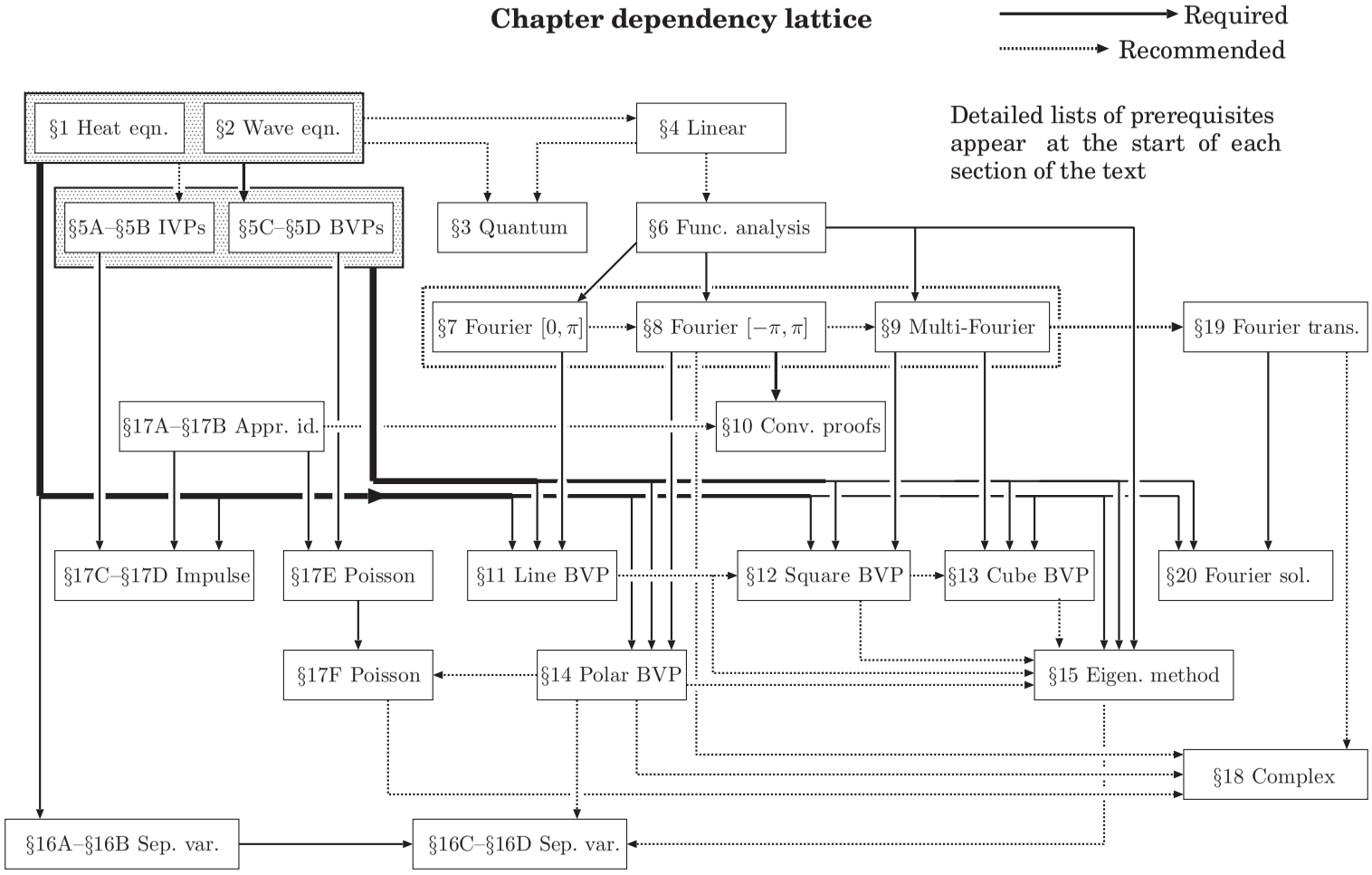
ISBN-13 978-0-511-76377-9 eBook (Adobe Reader)

ISBN-13 978-0-521-19970-4 Hardback

ISBN-13 978-0-521-13659-4 Paperback

To Joseph and Emma Pivato
for their support, encouragement,
and inspiring example.

Contents


Chapter dependency lattice
Required
Recommended
Detailed lists of prerequisites appear at the start of each section of the text
§1 Heat eqn.
§2 Wave eqn.
§4 Linear
§5A–§5B IVPs
§5C–§5D BVPs
§3 Quantum
§6 Func. analysis
§7 Fourier $[0, \pi]$
§8 Fourier $[-\pi, \pi]$
§9 Multi-Fourier
§19 Fourier trans.
§17A–§17B Appr. id.
§10 Conv. proofs
§17C–§17D Impulse
§17E Poisson
§11 Line BVP
§12 Square BVP
§13 Cube BVP
§20 Fourier sol.
§17F Poisson
§14 Polar BVP
§15 Eigen. method
§18 Complex
§16A–§16B Sep. var.
§16C–§16D Sep. var.

Preface

This is a textbook for an introductory course on linear partial differential equations (PDEs) and initial/boundary value problems (I/BVPs). It also provides a mathematically rigorous introduction to Fourier analysis (Chapters 7, 8, 9, 10, and 19), which is the main tool used to solve linear PDEs in Cartesian coordinates. Finally, it introduces basic functional analysis (Chapter 6) and complex analysis (Chapter 18). The first is necessary to characterize rigorously the convergence of Fourier series, and also to discuss eigenfunctions for linear differential operators. The second provides powerful techniques to transform domains and compute integrals, and also offers additional insight into Fourier series.

This book is not intended to be comprehensive or encyclopaedic. It is designed for a one-semester course (i.e. 36–40 hours of lectures), and it is therefore strictly limited in scope. First, it deals mainly with *linear* PDEs with constant coefficients. Thus, there is no discussion of characteristics, conservation laws, shocks, variational techniques, or perturbation methods, which would be germane to other types of PDEs. Second, the book focuses mainly on concrete solution methods to specific PDEs (e.g. the Laplace, Poisson, heat, wave, and Schrödinger equations) on specific domains (e.g. line segments, boxes, disks, annuli, spheres), and spends rather little time on qualitative results about entire classes of PDEs (e.g. elliptic, parabolic, hyperbolic) on general domains. Only after a thorough exposition of these special cases does the book sketch the general theory; experience shows that this is far more pedagogically effective than presenting the general theory first. Finally, the book does not deal at all with numerical solutions or Galerkin methods.

One slightly unusual feature of this book is that, from the very beginning, it emphasizes the central role of eigenfunctions (of the Laplacian) in the solution methods for linear PDEs. Fourier series and Fourier–Bessel expansions are introduced as the orthogonal eigenfunction expansions which are most suitable in certain domains or coordinate systems. Separation of variables appears relatively late in the exposition (Chapter 16) as a convenient device to obtain such eigenfunctions.

The only techniques in the book which are not either implicitly or explicitly based on eigenfunction expansions are impulse-response functions and Green's functions (Chapter 17) and complex-analytic methods (Chapter 18).

Prerequisites and intended audience

This book is written for third-year undergraduate students in mathematics, physics, engineering, and other mathematical sciences. The only prererequisites are (1) *multivariate calculus* (i.e. partial derivatives, multivariate integration, changes of coordinate system) and (2) *linear algebra* (i.e. linear operators and their eigenvectors).

It might also be helpful for students to be familiar with the following: (1) the basic theory of ordinary differential equations (specifically, Laplace transforms, Frobenius method); (2) some elementary vector calculus (specifically, divergence and gradient operators); and (3) elementary physics (to understand the physical motivation behind many of the problems). However, none of these three things are really required.

In addition to this background knowledge, the book requires some ability at abstract mathematical reasoning. Unlike some 'applied math' texts, we do not suppress or handwave the mathematical theory behind the solution methods. At suitable moments, the exposition introduces concepts such as 'orthogonal basis', 'uniform convergence' vs. '$\mathbf{L}_2$-convergence', 'eigenfunction expansion', and 'self-adjoint operator'; thus, students must be intellectually capable of understanding abstract mathematical concepts of this nature. Likewise, the exposition is mainly organized in a 'definition → theorem → proof → example' format, rather than a 'problem → solution' format. Students must be able to understand abstract descriptions of general solution techniques, rather than simply learn by imitating worked solutions to special cases.

Conventions in the text

- ∗ in the title of a chapter or section indicates 'optional' material which is not part of the core syllabus.
- (Optional) in the margin indicates that a particular theorem or statement is 'optional' in the sense that it is not required later in the text.
- Ⓔ in the margin indicates the location of an exercise. (Shorter exercises are sometimes embedded within the exposition.)
- ⧫ indicates the ends of more lengthy exercises.
- □ ends the proof of a theorem.
- ◇ indicates the end of an example.
- ⋄ ends the proof of a 'claim' within the proof of a theorem.
- △ ends the proof of a 'subclaim' within the proof of a claim.

Acknowledgements

I would like to thank Xiaorang Li of Trent University, who read through an early draft of this book and made many helpful suggestions and corrections, and who also provided Problems 5.18, 5.19, and 6.8. I also thank Peter Nalitolela, who proofread a penultimate draft and spotted many mistakes. I especially thank Irene Pizzie of Cambridge University Press, whose very detailed and thorough copy-editing of the final manuscript resulted in many dozens of corrections and clarifications. I would like to thank several anonymous reviewers who made many useful suggestions, and I would also like to thank Peter Thompson of Cambridge University Press for recruiting these referees. I also thank Diana Gillooly of Cambridge University Press, who was very supportive and helpful throughout the entire publication process, especially concerning my desire to provide a free online version of the book, and to release the figures and problem sets under a Creative Commons License. I also thank the many students who used the early versions of this book, especially those who found mistakes or made good suggestions. Finally, I thank George Peschke of the University of Alberta, for being an inspiring example of good mathematical pedagogy.

None of these people are responsible for any remaining errors, omissions, or other flaws in the book (of which there are no doubt many). If you find an error or some other deficiency in the book, please contact me at marcuspivato@trentu.ca.

This book would not have been possible without open source software. The book was prepared entirely on the LINUX operating system (initially REDHAT,[1] and later UBUNTU[2]). All the text is written in Leslie Lamport's LaTeX2e typesetting language,[3] and was authored using Richard Stallman's EMACS editor.[4] The illustrations were hand-drawn using William Chia-Wei Cheng's excellent TGIF object-oriented drawing program.[5] Additional image manipulation and post-processing was done with GNU IMAGE MANIPULATION PROGRAM (GIMP).[6] Many of the plots were created using GNUPLOT.[7,8] I would like to take this opportunity to thank the many people in the open source community who have developed this software.

Finally, and most importantly, I would like to thank my beloved wife and partner, Reem Yassawi, and our wonderful children, Leila and Aziza, for their support and for their patience with my many long absences.

[1] http://www.redhat.com. [2] http://www.ubuntu.com. [3] http://www.latex-project.org.
[4] http://www.gnu.org/software/emacs/emacs.html. [5] http://bourbon.usc.edu:8001/tgif.
[6] http://www.gimp.org. [7] http://www.gnuplot.info.
[8] Many other plots were generated using Waterloo MAPLE (http://www.maplesoft.com), which unfortunately is *not* open-source.

What's good about this book?

This text has many advantages over most other introductions to partial differential equations.

Illustrations

PDEs are physically motivated and geometrical objects; they describe curves, surfaces, and scalar fields with special geometric properties, and the way these entities evolve over time under endogenous dynamics. To understand PDEs and their solutions, it is necessary to visualize them. Algebraic formulae are just a language used to communicate such visual ideas in lieu of pictures, and they generally make a poor substitute. This book has over 300 high-quality illustrations, many of which are rendered in three dimensions. In the online version of the book, most of these illustrations appear in full colour. Also, the website contains many animations which do not appear in the printed book.

Most importantly, on the book website, all illustrations are *freely available* under a Creative Commons Attribution Noncommercial Share-Alike License.[1] This means that you are free to download, modify, and utilize the illustrations to prepare your own course materials (e.g. printed lecture notes or beamer presentations), as long as you attribute the original author. Please visit <http://xaravve.trentu.ca/pde>.

Physical motivation

Connecting the math to physical reality is critical: it keeps students motivated, and helps them interpret the mathematical formalism in terms of their physical intuitions about diffusion, vibration, electrostatics, etc. Chapter 1 of this book discusses the

[1] See http://creativecommons.org/licenses/by-nc-sa/3.0.

physics behind the heat, Laplace, and Poisson equations, and Chapter 2 discusses the wave equation. An unusual addition to this text is Chapter 3, which discusses quantum mechanics and the Schrödinger Equation (one of the major applications of PDE theory in modern physics).

Detailed syllabus

Difficult choices must be made when turning a 600+ page textbook into a feasible 12-week lesson plan. It is easy to run out of time or inadvertently miss something important. To facilitate this task, this book provides a lecture-by-lecture breakdown of how the author covers the material (see p. xxv). Of course, each instructor can diverge from this syllabus to suit the interests/background of their students, a longer/shorter teaching semester, or their personal taste. But the prefabricated syllabus provides a base to work from, and will save most instructors a lot of time and aggravation.

Explicit prerequisites for each chapter and section

To save time, an instructor might want to skip a certain chapter or section, but worries that it may end up being important later. We resolve this problem in two ways. First, p. xiv provides a '*chapter dependency lattice*', which summarizes the large-scale structure of logical dependencies between the chapters of the book. Second, every section of every chapter begins with an explicit list of 'required' and 'recommended' prerequisite sections; this provides more detailed information about the small-scale structure of logical dependencies between sections. By tracing backward through this 'lattice of dependencies', you can figure out exactly what background material you must cover to reach a particular goal. This makes the book especially suitable for self-study.

Flat dependency lattice

There are many 'paths' through the 20-chapter dependency lattice on p. xiv every one of which is only *seven* chapters or less in length. Thus, an instructor (or an autodidact) can design many possible syllabi, depending on their interests, and can quickly move to advanced material. The 'Suggested 12-week syllabus' on p. xxv describes a gentle progression through the material, covering most of the 'core' topics in a 12-week semester, emphasizing concrete examples and gradually escalating the abstraction level. The Chapter Dependency Lattice suggests some other possibilities for 'accelerated' syllabi focusing on different themes.

- *Solving PDEs with impulse response functions.* Chapters 1, 2, 5, and 17 only.
- *Solving PDEs with Fourier transforms.* Chapters 1, 2, 5, 19, and 20 only. (Pedagogically speaking, Chapters 8 and 9 will help the student understand Chapter 19, and Chapters 11–13 will help the student understand Chapter 20. Also, it is interesting to see how the 'impulse response' methods of Chapter 17 yield the same solutions as the 'Fourier methods' of Chapter 20, using a totally different approach. However, strictly speaking, none of Chapters 8–13 or 17 is logically necessary.)
- *Solving PDEs with separation of variables.* Chapters 1, 2, and, 16 only. (However, without at least Chapters 12, and 14, the ideas of Chapter 16 will seem somewhat artificial and pointless.)
- *Solving I/BVPs using eigenfunction expansions.* Chapters 1, 2, 4, 5, 6, and 15. (It would be pedagogically better to also cover Chapters 9 and 12, and probably Chapter 14. But, strictly speaking, none of these is logically necessary.)
- *Tools for quantum mechanics.* Section 1B, then Chapters 3, 4, 6, 9, 13, 19, and 20 (skipping material on Laplace, Poisson, and wave equations in Chapters 13 and 20, and adapting the solutions to the heat equation into solutions to the Schrödinger Equation).
- *Fourier theory.* Section 4A, then Chapters 6, 7, 8, 9, 10, and 19. Finally, Sections 18A, 18C, 18E, and 18F provide a 'complex' perspective. (Section 18H also contains some useful computational tools.)
- *Crash course in complex analysis.* Chapter 18 is logically independent of the rest of the book, and rigorously develops the main ideas in complex analysis from first principles. (However, the emphasis is on applications to PDEs and Fourier theory, so some of the material may seem esoteric or unmotivated if read in isolation from other chapters.)

Highly structured exposition, with clear motivation up front

The exposition is broken into small, semi-independent logical units, each of which is clearly labelled, and which has a clear purpose or meaning which is made immediately apparent. This simplifies the instructor's task; it is not necessary to spend time restructuring and summarizing the text material because it is already structured in a manner which self-summarizes. Instead, instructors can focus more on explanation, motivation, and clarification.

Many 'practice problems' (with complete solutions and source code available online)

Frequent evaluation is critical to reinforce material taught in class. This book provides an extensive supply of (generally simple) computational 'practice problems' at the end of each chapter. Completely worked solutions to virtually all of these problems are available on the book website. Also on the book website, the LaTeX source code for all problems and solutions is *freely available* under a Creative

Commons Attribution Noncommercial Share-Alike License.[2] Thus, an instructor can download and edit this source code, and easily create quizzes, assignments, and matching solutions for students.

Challenging exercises without solutions

Complex theoretical concepts cannot really be tested in quizzes, and do not lend themselves to canned 'practice problems'. For a more theoretical course with more mathematically sophisticated students, the instructor will want to assign some proof-related exercises for homework. This book has more than 420 such exercises scattered throughout the exposition; these are flagged by an 'Ⓔ' symbol in the margin, as shown here. Many of these exercises ask the student to prove a major result from the text (or a component thereof). This is the best kind of exercise, because it reinforces the material taught in class, and gives students a sense of ownership of the mathematics. Also, students find it more fun and exciting to prove important theorems, rather than solving esoteric make-work problems.

Ⓔ

Appropriate rigour

The solutions of PDEs unfortunately involve many technicalities (e.g. different forms of convergence; derivatives of infinite function series, etc.). It is tempting to handwave and gloss over these technicalities, to avoid confusing students. But this kind of pedagogical dishonesty actually makes students *more* confused; they know something is fishy, but they can't tell quite what. Smarter students know they are being misled, and may lose respect for the book, the instructor, or even the whole subject.

In contrast, this book provides a rigorous mathematical foundation for all its solution methods. For example, Chapter 6 contains a careful explanation of L^2-spaces, the various forms of convergence for Fourier series, and the differences between them – including the 'pathologies' which can arise when one is careless about these issues. I adopt a 'triage' approach to proofs: the simplest proofs are left as exercises for the motivated student (often with a step-by-step breakdown of the best strategy). The most complex proofs I have omitted, but I provide multiple references to other recent texts. In between are those proofs which are challenging but still accessible; I provide detailed expositions of these proofs. Often, when the text contains several variants of the same theorem, I prove one variant in detail, and leave the other proofs as exercises.

[2] See http://creativecommons.org/licenses/by-nc-sa/3.0.

Appropriate abstraction

It is tempting to avoid abstractions (e.g. linear differential operators, eigenfunctions), and simply present *ad hoc* solutions to special cases. This cheats the student. The right abstractions provide simple, yet powerful, tools that help students understand a myriad of seemingly disparate special cases within a single unifying framework. This book provides students with the opportunity to learn an abstract perspective once they are ready for it. Some abstractions are introduced in the main exposition, others are in optional sections, or in the philosophical preambles which begin each major part of the book.

Gradual abstraction

Learning proceeds from the concrete to the abstract. Thus, the book begins each topic with a specific example or a low-dimensional formulation, and only later proceeds to a more general/abstract idea. This introduces a lot of 'redundancy' into the text, in the sense that later formulations subsume the earlier ones. So the exposition is not as 'efficient' as it could be. This is a good thing. Efficiency makes for good reference books, but lousy texts.

For example, when introducing the heat equation, Laplace equation, and wave equation in Chapters 1 and 2, I first derive and explain the one-dimensional version of each equation, then the two-dimensional version, and, finally, the general, D-dimensional version. Likewise, when developing the solution methods for BVPs in Cartesian coordinates (Chapters 11–13), I confine the exposition to the interval $[0, \pi]$, the square $[0, \pi]^2$, and the cube $[0, \pi]^3$, and assume all the coefficients in the differential equations are unity. Then the exercises ask the student to state and prove the appropriate generalization of each solution method for an interval/rectangle/box of arbitrary dimensions, and for equations with arbitrary coefficients. The general method for solving I/BVPs using eigenfunction expansions only appears in Chapter 15, after many special cases of this method have been thoroughly exposited in Cartesian and polar coordinates (Chapters 11–14).

Likewise, the development of Fourier theory proceeds in gradually escalating levels of abstraction. First we encounter Fourier (co)sine series on the interval $[0, \pi]$ (§7A); then on the interval $[0, L]$ for arbitrary $L > 0$ (§7B). Then Chapter 8 introduces 'real' Fourier series (i.e. with both sine and cosine terms), and then complex Fourier series (§8D). Then, Chapter 9 introduces two-dimensional (co)sine series and, finally, D-dimensional (co)sine series.

Expositional clarity

Computer scientists have long known that it is easy to write software that *works*, but it is much more difficult (and important) to write working software that *other people*

can understand. Similarly, it is relatively easy to write formally correct mathematics; the real challenge is to make the mathematics easy to read. To achieve this, I use several techniques. I divide proofs into semi-independent modules ('claims'), each of which performs a simple, clearly defined task. I integrate these modules together in an explicit hierarchical structure (with 'subclaims' inside of 'claims'), so that their functional interdependence is clear from visual inspection. I also explain formal steps with parenthetical heuristic remarks. For example, in a long string of (in)equalities, I often attach footnotes to each step, as follows:

'$A \underset{(*)}{=} B \underset{(\dagger)}{\leq} C \underset{(\ddagger)}{<} D$. Here, $(*)$ is because [. . .]; $(\dagger)$ follows from [. . .], and $(\ddagger)$ is because [. . .].'

Finally, I use letters from the same 'lexicographical family' to denote objects which 'belong' together. For example: If $\mathcal{S}$ and $\mathcal{T}$ are sets, then elements of $\mathcal{S}$ should be $s_1, s_2, s_3, \ldots$, while elements of $\mathcal{T}$ are $t_1, t_2, t_3, \ldots$. If $\mathbf{v}$ is a vector, then its entries should be $v_1, \ldots, v_N$. If $\mathbf{A}$ is a matrix, then its entries should be $a_{11}, \ldots, a_{NM}$. I reserve upper-case letters (e.g. $J, K, L, M, N, \ldots$) for the bounds of intervals or indexing sets, and then use the corresponding lower-case letters (e.g. $j, k, l, m, n, \ldots$) as indexes. For example, $\forall n \in \{1, 2, \ldots, N\}$, $A_n := \sum_{j=1}^{J} \sum_{k=1}^{K} a_{jk}^{n}$.

Clear and explicit statements of solution techniques

Many PDE texts contain very few theorems; instead they try to develop the theory through a long sequence of worked examples, hoping that students will 'learn by imitation', and somehow absorb the important ideas 'by osmosis'. However, less gifted students often just imitate these worked examples in a slavish and uncomprehending way. Meanwhile, the more gifted students do not want to learn 'by osmosis'; they want clear and precise statements of the main ideas.

The problem is that most solution methods in PDEs, if stated as theorems in full generality, are incomprehensible to many students (especially the non-math majors). My solution is this: I provide explicit and precise statements of the solution method for almost every possible combination of (1) several major PDEs, (2) several kinds of boundary conditions, and (3) several different domains. I state these solutions as *theorems*, not as 'worked examples'. I follow each of these theorems with several completely worked examples. Some theorems I prove, but most of the proofs are left as exercises (often with step-by-step hints).

Of course, this approach is highly redundant, because I end up stating more than 20 theorems, which really are all special cases of three or four general results (for example, the general method for solving the heat equation on a compact domain with Dirichlet boundary conditions, using an eigenfunction expansion). However, this sort of redundancy is *good* in an elementary exposition. Highly 'efficient'

expositions are pleasing to our aesthetic sensibilities, but they are dreadful for pedagogical purposes.

However, one must not leave the students with the impression that the theory of PDEs is a disjointed collection of special cases. To link together all the 'homogeneous Dirichlet heat equation' theorems, for example, I explicitly point out that they all utilize the same underlying strategy. Also, when a proof of one variant is left as an exercise, I encourage students to imitate the (provided) proofs of previous variants. When the students understand the underlying similarity between the various special cases, *then* it is appropriate to state the general solution. The students will almost feel they have figured it out for themselves, which is the best way to learn something.

Suggested 12-week syllabus

Week 1: *Heat and diffusion-related PDEs*

Lecture 1: Appendix A–Appendix E *Review of multivariate calculus; introduction to complex numbers.*

Lecture 2: §1A–§1B *Fourier's law; the heat equation.*

Lecture 3: §1C–§1D *Laplace equation; Poisson equation.*

Week 2: *Wave-related PDEs; quantum mechanics*

Lecture 1: §1E; §2A *Properties of harmonic functions; spherical means.*

Lecture 2: §2B–§2C *Wave equation; telegraph equation.*

Lecture 3: Chapter 3 *The Schrödinger equation in quantum mechanics.*

Week 3: *General theory*

Lecture 1: §4A–§4C *Linear PDEs: homogeneous vs. nonhomogeneous.*

Lecture 2: §5A; §5B *Evolution equations and initial value problems.*

Lecture 3: §5C *Boundary conditions and boundary value problems.*

Week 4: *Background to Fourier theory*

Lecture 1: §5D *Uniqueness of solutions to BVPs*; §6A *inner products.*

Lecture 2: §6B–§6D *L^2-space; orthogonality.*

Lecture 3: §6E(i)–(iii) *L^2-convergence; pointwise convergence; uniform convergence.*

Week 5: *One-dimensional Fourier series*

Lecture 1: §6E(iv) *Infinite series*; §6F *orthogonal bases*; §7A *Fourier (co/sine) series: definition and examples.*

Lecture 2: §7C(i)–(v) *Computing Fourier series of polynomials, piecewise linear and step functions.*

Lecture 3: §11A–§11C *Solution to heat equation and Poisson equation on a line segment.*

Week 6: *Fourier solutions for BVPs in one and two dimensions*

Lecture 1: §11B–§12A *Wave equation on line segment and Laplace equation on a square.*

Lecture 2: §9A–§9B *Multidimensional Fourier series.*

Lecture 3: §12B–§12C(i) *Solution to heat equation and Poisson equation on a square.*

Week 7: *Fourier solutions for two-dimensional BVPs in Cartesian and polar coordinates*

Lecture 1: §12C(ii), §12D *Solution to Poisson equation and wave equation on a square.*

Lecture 2: §5C(iv); §8A–§8B *Periodic boundary conditions; real Fourier series.*

Lecture 3: §14A; §14B(i)–(iv) *Laplacian in polar coordinates; Laplace equation on (co)disk.*

Week 8: *BVPs in polar coordinates; Bessel functions*

Lecture 1: §14C *Bessel functions.*

Lecture 2: §14D–§14F *Heat, Poisson, and wave equations in polar coordinates.*

Lecture 3: §14G *Solving Bessel's equation with the method of Frobenius.*

Week 9: *Eigenbases; separation of variables*

Lecture 1: §15A–§15B *Eigenfunction solutions to BVPs.*

Lecture 2: §15B; §16A–§16B *Harmonic bases; separation of variables in Cartesian coordinates.*

Lecture 3: §16C–§16D *Separation of variables in polar and spherical coordinates; Legendre polynomials.*

Week 10: *Impulse response methods*

Lecture 1: §17A–§17C *Impulse response functions; convolution; approximations of identity; Gaussian convolution solution for heat equation.*

Lecture 2: §17C–§17F *Gaussian convolutions continued; Poisson's solutions to Dirichlet problem on a half-plane and a disk.*

Lecture 3: §14B(v); §17D *Poisson solution on disk via polar coordinates; d'Alembert solution to wave equation.*

Week 11: *Fourier transforms*

Lecture 1: §19A *One-dimensional Fourier transforms.*

Lecture 2: §19B *Properties of one-dimensional Fourier transform.*

Lecture 3: §20A; §20C *Fourier transform solution to heat equation; Dirchlet problem on half-plane.*

Week 12: *Fourier transform solutions to PDEs*

Lecture 1: §19D, §20B(i) *Multidimensional Fourier transforms; solution to wave equation.*

Lecture 2: §20B(ii); §20E *Poisson's spherical mean solution; Huygen's principle; the general method.*

Lecture 3: (Time permitting) §19G or §19H *Heisenberg uncertainty or Laplace transforms*.

In a longer semester or a faster paced course, one could also cover parts of Chapter 10 (*Proofs of Fourier convergence*) and/or Chapter 18 (*Applications of complex analysis*).

Part I

Motivating examples and major applications

A *dynamical system* is a mathematical model of a system evolving in time. Most models in mathematical physics are dynamical systems. If the system has only a finite number of 'state variables', then its dynamics can be encoded in an *ordinary differential equation* (ODE), which expresses the *time derivative* of each state variable (i.e. its rate of change over time) as a function of the other state variables. For example, *celestial mechanics* concerns the evolution of a system of gravitationally interacting objects (e.g. stars and planets). In this case, the 'state variables' are vectors encoding the position and momentum of each object, and the ODE describes how the objects move and accelerate as they interact gravitationally.

However, if the system has a very large number of state variables, then it is no longer feasible to represent it with an ODE. For example, consider the flow of heat or the propagation of compression waves through a steel bar containing 10^{24} iron atoms. We *could* model this using a 10^{24}-dimensional ODE, where we explicitly track the vibrational motion of each iron atom. However, such a 'microscopic' model would be totally intractable. Furthermore, it is not necessary. The iron atoms are (mostly) immobile, and interact only with their immediate neighbours. Furthermore, nearby atoms generally have roughly the same temperature, and move in synchrony. Thus, it suffices to consider the macroscopic *temperature distribution* of the steel bar, or to study the fluctuation of a macroscopic *density field*. This temperature distribution or density field can be mathematically represented as a smooth, real-valued function over some three-dimensional domain. The flow of heat or the propagation of sound can then be described as the *evolution* of this function over time.

We now have a dynamical system where the 'state variable' is not a finite system of vectors (as in celestial mechanics), but is instead a multivariate *function*. The evolution of this function is determined by its spatial geometry – e.g. the local 'steepness' and variation of the temperature gradients between warmer and cooler regions in the bar. In other words, the *time derivative* of the function (its rate

of change over time) is determined by its *spatial derivatives* (which describe its slope and curvature at each point in space). An equation that relates the different derivatives of a multivariate function in this way is a *partial differential equation* (PDE). In particular, a PDE which describes a dynamical system is called an *evolution equation*. For example, the evolution equation which describes the flow of heat through a solid is called the *heat equation*. The equation which describes compression waves is the *wave equation*.

An *equilibrium* of a dynamical system is a state which is unchanging over time; mathematically, this means that the time-derivative is equal to zero. An equlibrium of an N-dimensional evolution equation satisfies an $(N-1)$-dimensional PDE called an *equilibrium equation*. For example, the equilibrium equations corresponding to the heat equation are the *Laplace equation* and the *Poisson equation* (depending on whether or not the system is subjected to external heat input).

PDEs are thus of central importance in the thermodynamics and acoustics of continuous media (e.g. steel bars). The heat equation also describes chemical diffusion in fluids, and also the evolving probability distribution of a particle performing a random walk called *Brownian motion*. It thus finds applications everywhere from mathematical biology to mathematical finance. When diffusion or Brownian motion is combined with deterministic drift (e.g. due to prevailing wind or ocean currents) it becomes a PDE called the *Fokker–Planck equation*.

The Laplace and Poisson equations describe the equilibria of such diffusion processes. They also arise in electrostatics, where they describe the shape of an electric field in a vacuum. Finally, solutions of the two-dimensional Laplace equation are good approximations of surfaces trying to minimize their elastic potential energy – that is, soap films.

The wave equation describes the resonance of a musical instrument, the spread of ripples on a pond, seismic waves propagating through the earth's crust, and shockwaves in solar plasma. (The motion of fluids themselves is described by yet another PDE, the *Navier–Stokes equation*.) A version of the wave equation arises as a special case of Maxwell's equations of electrodynamics; this led to Maxwell's prediction of *electromagnetic waves*, which include radio, microwaves, X-rays, and visible light. When combined with a 'diffusion' term reminiscent of the heat equation, the wave equation becomes the *telegraph equation*, which describes the propagation and degradation of electrical signals travelling through a wire.

Finally, an odd-looking 'complex' version of the heat equation induces wave-like evolution in the complex-valued probability fields which describe the position and momentum of subatomic particles. This *Schrödinger equation* is the starting point of quantum mechanics, one of the two most revolutionary developments in physics in the twentieth century. The other revolutionary development was relativity theory. General relativity represents spacetime as a four-dimensional manifold,

whose curvature interacts with the spatiotemporal flow of mass/energy through yet another PDE: the *Einstein equation*.

Except for the Einstein and Navier–Stokes equations, all the equations we have mentioned are *linear* PDEs. This means that a sum of two or more solutions to the PDE will also be a solution. This allows us to solve linear PDEs through the *method of superposition*: we build complex solutions by adding together many simple solutions. A particularly convenient class of simple solutions are *eigenfunctions*. Thus, an enormously powerful and general method for linear PDEs is to represent the solutions using *eigenfunction expansions*. The most natural eigenfunction expansion (in Cartesian coordinates) is the *Fourier series*.

1

Heat and diffusion

> The differential equations of the propagation of heat express the most general conditions, and reduce the physical questions to problems of pure analysis, and this is the proper object of theory.
>
> *Jean Joseph Fourier*

1A Fourier's law

Prerequisites: Appendix A. **Recommended:** Appendix E.

1A(i) . . . in one dimension

Figure 1A.1 depicts a material diffusing through a one-dimensional domain $\mathbb{X}$ (for example, $\mathbb{X} = \mathbb{R}$ or $\mathbb{X} = [0, L]$). Let $u(x, t)$ be the density of the material at the point $x \in \mathbb{X}$ at time $t > 0$. Intuitively, we expect the material to flow from regions of *greater* to *lesser* concentration. In other words, we expect the *flow* of the material at any point in space to be proportional to the *slope* of the curve $u(x, t)$ at that point. Thus, if $F(x, t)$ is the flow at the point x at time t, then we expect the following:

$$F(x, t) = -\kappa \cdot \partial_x u(x, t),$$

where $\kappa > 0$ is a constant measuring the rate of diffusion. This is an example of *Fourier's law.*

1A(ii) . . . in many dimensions

Prerequisites: Appendix E.

Figure 1A.2 depicts a material diffusing through a two-dimensional domain $\mathbb{X} \subset \mathbb{R}^2$ (e.g. heat spreading through a region, ink diffusing in a bucket of water, etc.) We

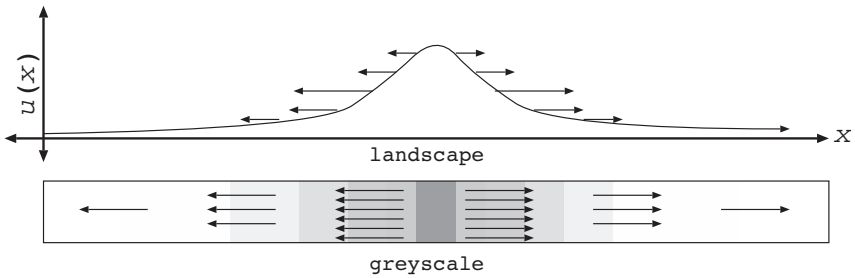


Figure 1A.1. Fourier's law of heat flow in one dimension.

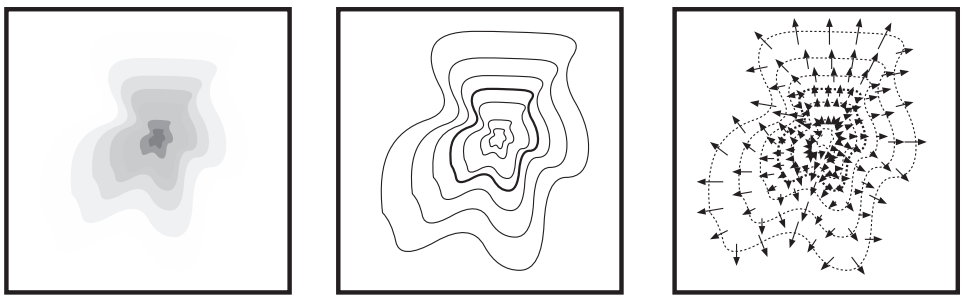


Figure 1A.2. Fourier's law of heat flow in two dimensions.

could just as easily suppose that $\mathbb{X} \subset \mathbb{R}^D$ is a D-dimensional domain. If $\mathbf{x} \in \mathbb{X}$ is a point in space, and $t \geq 0$ is a moment in time, let $u(\mathbf{x}, t)$ denote the concentration at $\mathbf{x}$ at time t. (This determines a function $u : \mathbb{X} \times \mathbb{R}_{\neq} \longrightarrow \mathbb{R}$, called a *time-varying scalar field*.)

Now let $\vec{\mathbf{F}}(\mathbf{x}, t)$ be a D-dimensional vector describing the *flow* of the material at the point $\mathbf{x} \in \mathbb{X}$. (This determines a *time-varying vector field* $\vec{\mathbf{F}} : \mathbb{R}^D \times \mathbb{R}_{\neq} \longrightarrow \mathbb{R}^D$.)

Again, we expect the material to flow from regions of high concentration to low concentration. In other words, material should flow *down the concentration gradient*. This is expressed by *Fourier's law of heat flow*:

$$\vec{\mathbf{F}} = -\kappa \cdot \nabla u,$$

where $\kappa > 0$ is a constant measuring the rate of diffusion.

One can imagine u as describing a distribution of highly antisocial people; each person is always fleeing everyone around them and moving in the direction with the fewest people. The constant κ measures the average walking speed of these misanthropes.

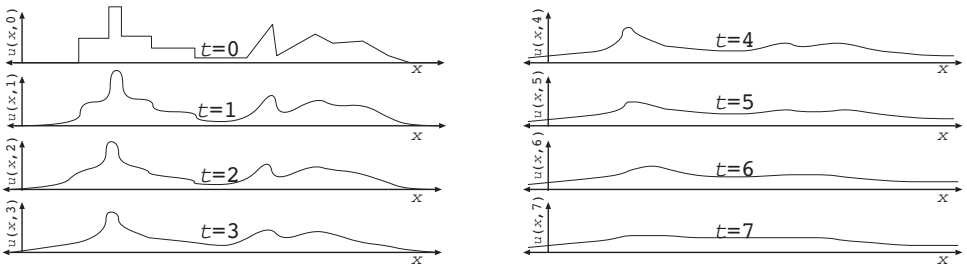


Figure 1B.1. The heat equation as 'erosion'.

1B The heat equation

Recommended: §1A.

1B(i) . . . in one dimension

Prerequisites: §1A(i).

Consider a material diffusing through a one-dimensional domain $\mathbb{X}$ (for example, $\mathbb{X} = \mathbb{R}$ or $\mathbb{X} = [0, L]$). Let $u(x, t)$ be the density of the material at the location $x \in \mathbb{X}$ at time $t \in \mathbb{R}_{+}$, and let $F(x, t)$ be the flux of the material at the location x and time t. Consider the derivative $\partial_x F(x, t)$. If $\partial_x F(x, t) > 0$, this means that the flow is *diverging*[1] at this point in space, so the material there is spreading farther apart. Hence, we expect the concentration at this point to *decrease*. Conversely, if $\partial_x F(x, t) < 0$, then the flow is *converging* at this point in space, so the material there is crowding closer together, and we expect the concentration to *increase*. To be succinct: the concentration of material will *increase* in regions where F converges and *decrease* in regions where F diverges. The equation describing this is given by

$$\partial_t u(x, t) = -\partial_x F(x, t).$$

If we combine this with Fourier's law, however, we get:

$$\partial_t u(x, t) = \kappa \cdot \partial_x \partial_x u(x, t),$$

which yields the *one-dimensional heat equation*:

$$\boxed{\partial_t u(x, t) = \kappa \cdot \partial_x^2 u(x, t).}$$

Heuristically speaking, if we imagine $u(x, t)$ as the height of some one-dimensional 'landscape', then the heat equation causes this landscape to be 'eroded', as if it were subjected to thousands of years of wind and rain (see Figure 1B.1).

[1] See Appendix E(ii), p. 562, for an explanation of why we say the flow is 'diverging' here.

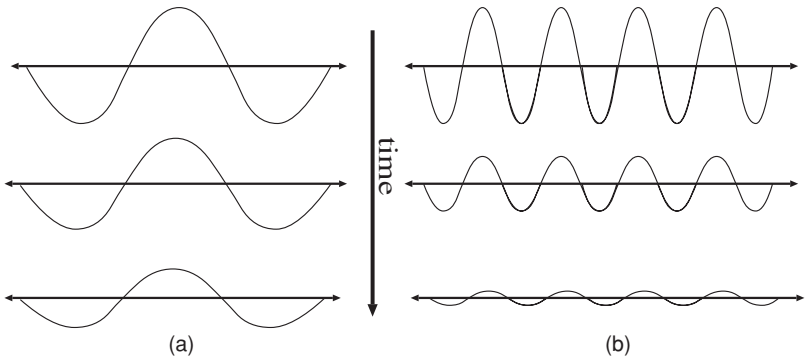


Figure 1B.2. Under the heat equation, the exponential decay of a periodic function is proportional to the square of its frequency. (a) Low frequency, slow decay; (b) high frequency, fast decay.

Example 1B.1 For simplicity we suppose $\kappa = 1$.

(a) Let $u(x,t) = \mathrm{e}^{-9t} \cdot \sin(3x)$. Thus, u describes a spatially sinusoidal function (with spatial frequency 3) whose magnitude decays exponentially over time.

(b) The dissipating wave. More generally, let $u(x,t) = \mathrm{e}^{-\omega^2 \cdot t} \cdot \sin(\omega \cdot x)$. Then u is a solution to the one-dimensional heat equation, and it looks like a standing wave whose amplitude decays exponentially over time (see Figure 1B.2). Note that the decay rate of the function u is proportional to the square of its frequency.

(c) The (one-dimensional) Gauss–Weierstrass kernel. Let

$$\mathcal{G}(x;t) := \frac{1}{2\sqrt{\pi t}} \exp\left(\frac{-x^2}{4t}\right).$$

Then $\mathcal{G}$ is a solution to the one-dimensional heat equation, and looks like a 'bell curve', which starts out tall and narrow, and, over time, becomes broader and flatter (Figure 1B.3). ◇

Ⓔ **Exercise 1B.1** Verify that all the functions in Examples 1B.1(a)–(c) satisfy the heat equation. ⧫

All three functions in Example 1B.1 start out very tall, narrow, and pointy, and gradually become shorter, broader, and flatter. This is generally what the heat equation does; it tends to flatten things out. If u describes a physical landscape, then the heat equation describes 'erosion'.

1B(ii) . . . in many dimensions

Prerequisites: §1A(ii).

More generally, if $u : \mathbb{R}^D \times \mathbb{R}_{\not\tau} \longrightarrow \mathbb{R}$ is the time-varying density of some material, and $\vec{\mathbf{F}} : \mathbb{R}^D \times \mathbb{R}_{\not\tau} \longrightarrow \mathbb{R}$ is the flux of this material, then we would expect the

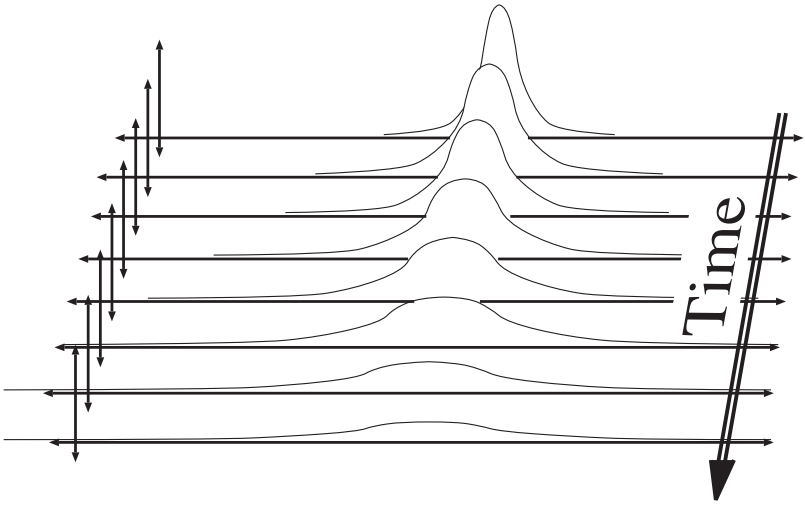


Figure 1B.3. The Gauss–Weierstrass kernel under the heat equation.

material to *increase* in regions where $\vec{\mathbf{F}}$ converges and to *decrease* in regions where $\vec{\mathbf{F}}$ diverges.[2] In other words, we have

$$\partial_t u = -\text{div}\,\vec{\mathbf{F}}.$$

If u is the density of some diffusing material (or heat), then $\vec{\mathbf{F}}$ is determined by Fourier's law, so we get the heat equation

$$\partial_t u = \kappa \cdot \text{div}\,\nabla u = \kappa\,\triangle u.$$

Here, $\triangle$ is the *Laplacian* operator,[3] defined as follows:

$$\boxed{\triangle u = \partial_1^2 u + \partial_2^2 u + \cdots + \partial_D^2 u}$$

Exercise 1B.2 (a) If $D = 1$ and $u : \mathbb{R} \longrightarrow \mathbb{R}$, verify that $\text{div}\,\nabla u(x) = u''(x) = \triangle u(x)$, for all $x \in \mathbb{R}$. Ⓔ

(b) If $D = 2$ and $u : \mathbb{R}^2 \longrightarrow \mathbb{R}$, verify that $\text{div}\,\nabla u(x, y) = \partial_x^2 u(x, y) + \partial_y^2 u(x, y) = \triangle u(x, y)$, for all $(x, y) \in \mathbb{R}^2$.

(c) For any $D \geq 2$ and $u : \mathbb{R}^D \longrightarrow \mathbb{R}$, verify that $\text{div}\,\nabla u(\mathbf{x}) = \triangle u(\mathbf{x})$, for all $\mathbf{x} \in \mathbb{R}^D$. ♦

By changing to the appropriate time units, we can assume $\kappa = 1$, so the heat equation becomes

$$\boxed{\partial_t u = \triangle u.}$$

[2] See Appendix E(ii), p. 562, for a review of the 'divergence' of a vector field.
[3] Sometimes the Laplacian is written as ∇^2.

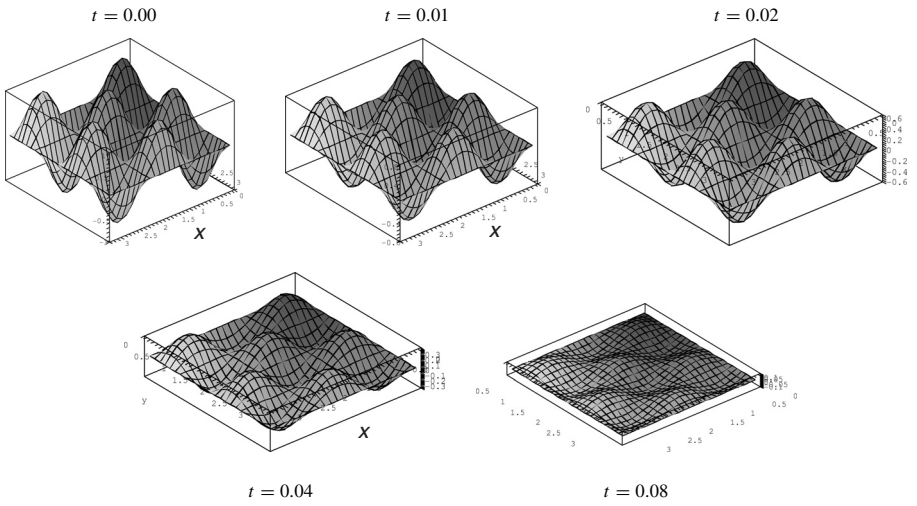


Figure 1B.4. Five snapshots of the function $u(x, y; t) = \mathrm{e}^{-25\,t} \cdot \sin(3x)\sin(4y)$ from Example 1B.2.

For example,

- if $\mathbb{X} \subset \mathbb{R}$, and $x \in \mathbb{X}$, then $\triangle u(x; t) = \partial_x^2\, u(x; t)$;
- if $\mathbb{X} \subset \mathbb{R}^2$, and $(x, y) \in \mathbb{X}$, then $\triangle u(x, y; t) = \partial_x^2\, u(x, y; t) + \partial_y^2\, u(x, y; t)$.

Thus, as we have already seen, the one-dimensional heat equation is given by

$$\partial_t\, u = \partial_x^2\, u,$$

and the the two-dimensional heat equation is given by

$$\boxed{\partial_t\, u(x, y; t) = \partial_x^2\, u(x, y; t) + \partial_y^2\, u(x, y; t).}$$

Example 1B.2

(a) Let $u(x, y; t) = e^{-25\,t} \cdot \sin(3x)\sin(4y)$. Then u is a solution to the two-dimensional heat equation, and looks like a two-dimensional 'grid' of sinusoidal hills and valleys with horizontal spacing $1/3$ and vertical spacing $1/4$. As shown in Figure 1B.4, these hills rapidly subside into a gently undulating meadow and then gradually sink into a perfectly flat landscape.

(b) The (two-dimensional) Gauss–Weierstrass kernel. Let

$$\mathcal{G}(x, y; t) := \frac{1}{4\pi t} \exp\left(\frac{-x^2 - y^2}{4t}\right).$$

Then $\mathcal{G}$ is a solution to the two-dimensional heat equation, and looks like a mountain, which begins steep and pointy and gradually 'erodes' into a broad, flat, hill.

(c) The D-dimensional Gauss–Weierstrass kernel is the function $\mathcal{G} : \mathbb{R}^D \times \mathbb{R}_+ \longrightarrow \mathbb{R}$ defined by

$$\boxed{\mathcal{G}(\mathbf{x}; t) = \frac{1}{(4\pi t)^{D/2}} \exp\left(\frac{-\|\mathbf{x}\|^2}{4t}\right).}$$

Technically speaking, $\mathcal{G}(\mathbf{x}; t)$ is a D-dimensional *symmetric normal probability distribution* with variance $\sigma = 2t$. ◇

Exercise 1B.3 Verify that all the functions in Examples 1B.2(a)–(c) satisfy the heat equation. Ⓔ ♦

Exercise 1B.4 Prove the *Leibniz rule* for Laplacians: if $f, g : \mathbb{R}^D \longrightarrow \mathbb{R}$ are two scalar fields, and $(f \cdot g) : \mathbb{R}^D \longrightarrow \mathbb{R}$ is their product, then, for all $\mathbf{x} \in \mathbb{R}^D$, Ⓔ

$$\triangle(f \cdot g)(\mathbf{x}) = g(\mathbf{x}) \cdot \Big(\triangle f(\mathbf{x}) \Big) + 2 \Big(\nabla f(\mathbf{x}) \Big) \bullet \Big(\nabla g(\mathbf{x}) \Big) + f(\mathbf{x}) \cdot \Big(\triangle g(\mathbf{x}) \Big).$$

Hint: Combine the Leibniz rules for gradients and divergences (see Propositions E.1(b) and E.2(b) in Appendix E, pp. 562 and 564). ♦

1C The Laplace equation

Prerequisites: §1B.

If the heat equation describes the erosion/diffusion of some system, then an *equilibrium* or *steady-state* of the heat equation is a scalar field $h : \mathbb{R}^D \longrightarrow \mathbb{R}$ satisfying the Laplace equation:

$$\boxed{\triangle h \equiv 0.}$$

A scalar field satisfying the Laplace equation is called a *harmonic function*.

Example 1C.1

(a) If $D = 1$, then $\triangle h(x) = \partial_x^2 h(x) = h''(x)$; thus, the *one-dimensional Laplace equation* is just

$$\boxed{h''(x) = 0.}$$

Suppose $h(x) = 3x + 4$. Then $h'(x) = 3$ and $h''(x) = 0$, so h is harmonic. More generally, the one-dimensional harmonic functions are just the *linear* functions of the form $h(x) = ax + b$ for some constants $a, b \in \mathbb{R}$.

(b) If $D = 2$, then $\triangle h(x, y) = \partial_x^2 h(x, y) + \partial_y^2 h(x, y)$, so the two-dimensional Laplace equation is given by

$$\boxed{\partial_x^2 h + \partial_y^2 h = 0,}$$

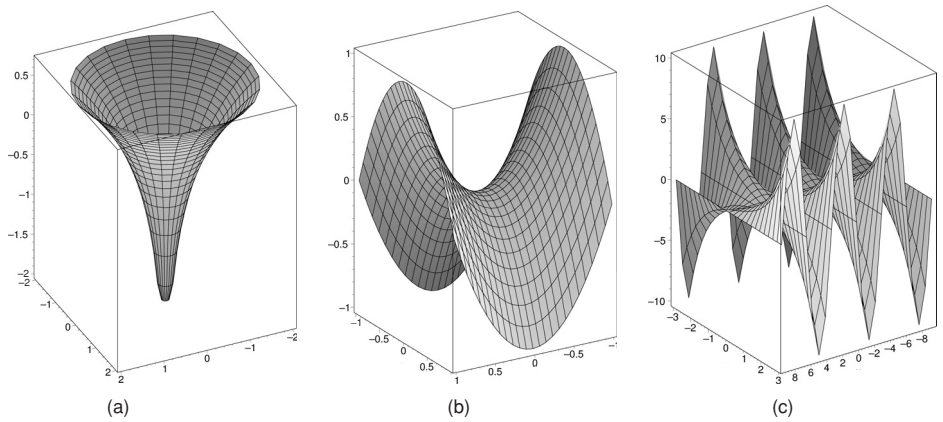


Figure 1C.1. Three harmonic functions. (a) $h(x, y) = \log(x^2 + y^2)$; (b) $h(x, y) = x^2 - y^2$; (c) $h(x, y) = \sin(x) \cdot \sinh(y)$. In all cases, note the telltale 'saddle' shape.

or, equivalently, $\partial_x^2 h = -\partial_y^2 h$. For example,

- Figure 1C.1(b) shows the harmonic function $h(x, y) = x^2 - y^2$;
- Figure 1C.1(c) shows the harmonic function $h(x, y) = \sin(x) \cdot \sinh(y)$. ◇

Ⓔ **Exercise 1C.1** Verify that these two functions are harmonic. ⧫

The surfaces in Figure 1C.1 have a 'saddle' shape, and this is typical of harmonic functions; in a sense, a harmonic function is one which is 'saddle-shaped' at every point in space. In particular, note that $h(x, y)$ has no maxima or minima anywhere; this is a universal property of harmonic functions (see Corollary 1E.2). The next example seems to contradict this assertion, but in fact it does not.

Example 1C.2 Figure 1C.1(a) shows the harmonic function $h(x, y) = \log(x^2 + y^2)$ for all $(x, y) \neq (0, 0)$. This function is well-defined everywhere except at $(0, 0)$; hence, contrary to appearances, $(0, 0)$ is *not* an extremal point. (Verification that h is harmonic is Problem 1.3, p. 23.) ◇

When $D \geq 3$, harmonic functions no longer define nice saddle-shaped *surfaces*, but they still have similar mathematical properties.

Example 1C.3

(a) If $D = 3$, then $\triangle h(x, y, z) = \partial_x^2 h(x, y, z) + \partial_y^2 h(x, y, z) + \partial_z^2 h(x, y, z)$. Thus, the three-dimensional Laplace equation reads as follows:

$$\boxed{\partial_x^2 h + \partial_y^2 h + \partial_z^2 h = 0.}$$

For example, let

$$h(x, y, z) = \frac{1}{\|(x, y, z)\|} = \frac{1}{\sqrt{x^2 + y^2 + z^2}}$$

for all $(x, y, z) \neq (0, 0, 0)$. Then h is harmonic everywhere except at $(0, 0, 0)$. (Verification that h is harmonic is Problem 1.4, p. 23.)

(b) For any $D \geq 3$, the D-dimensional Laplace equation reads as follows:

$$\boxed{\partial_1^2 h + \cdots + \partial_D^2 h = 0.}$$

For example, let

$$h(\mathbf{x}) = \frac{1}{\|\mathbf{x}\|^{D-2}} = \frac{1}{\left(x_1^2 + \cdots + x_D^2\right)^{\frac{D-2}{2}}}$$

for all $\mathbf{x} \neq \mathbf{0}$. Then h is harmonic everywhere in $\mathbb{R}^D \setminus \{\mathbf{0}\}$. ◇

Exercise 1C.2 Verify that h is harmonic on $\mathbb{R}^D \setminus \{\mathbf{0}\}$. ♦ Ⓔ

Harmonic functions have the convenient property that we can multiply together two lower-dimensional harmonic functions to get a higher dimensional one. For example,

- $h(x, y) = x \cdot y$ is a two-dimensional harmonic function. (**Exercise 1C.3** Verify this.) Ⓔ
- $h(x, y, z) = x \cdot (y^2 - z^2)$ is a three-dimensional harmonic function. (**Exercise 1C.4** Verify this.) Ⓔ

In general, we have the following:

Proposition 1C.4 *Suppose $u : \mathbb{R}^n \longrightarrow \mathbb{R}$ is harmonic and $v : \mathbb{R}^m \longrightarrow \mathbb{R}$ is harmonic, and define $w : \mathbb{R}^{n+m} \longrightarrow \mathbb{R}$ by $w(\mathbf{x}, \mathbf{y}) = u(\mathbf{x}) \cdot v(\mathbf{y})$ for $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{y} \in \mathbb{R}^m$. Then w is also harmonic.*

Proof **Exercise 1C.5** *Hint:* First prove that w obeys a kind of Leibniz rule: $\triangle w(\mathbf{x}, \mathbf{y}) = v(\mathbf{y}) \cdot \triangle u(\mathbf{x}) + u(\mathbf{x}) \cdot \triangle v(\mathbf{y})$. Ⓔ □

The function $w(\mathbf{x}, \mathbf{y}) = u(\mathbf{x}) \cdot v(\mathbf{y})$ is called a *separated solution*, and this proposition illustrates a technique called *separation of variables*. Exercise 1C.6 also explores separation of variables. A full exposition of this technique appears in Chapter 16.

Exercise 1C.6 Ⓔ

(a) Let $\mu, \nu \in \mathbb{R}$ be constants, and let $f(x, y) = \mathrm{e}^{\mu x} \cdot \mathrm{e}^{\nu y}$. Suppose f is harmonic; what can you conclude about the relationship between μ and ν? (Justify your assertion.)

(b) Suppose $f(x, y) = X(x) \cdot Y(y)$, where $X : \mathbb{R} \longrightarrow \mathbb{R}$ and $Y : \mathbb{R} \longrightarrow \mathbb{R}$ are two smooth functions. Suppose $f(x, y)$ is harmonic.

(i) Prove that $X''(x)/X(x) = -Y''(y)/Y(y)$ for all $x, y \in \mathbb{R}$.

(ii) Conclude that the function $X''(x)/X(x)$ must equal a constant c independent of x. Hence $X(x)$ satisfies the ordinary differential equation $X''(x) = c \cdot X(x)$.

Likewise, the function $Y''(y)/Y(y)$ must equal $-c$, independent of y. Hence $Y(y)$ satisfies the ordinary differential equation $Y''(y) = -c \cdot Y(y)$.

(iii) Using this information, deduce the general form for the functions $X(x)$ and $Y(y)$, and use this to obtain a general form for $f(x, y)$. ♦

1D The Poisson equation

Prerequisites: §1C.

Imagine $p(\mathbf{x})$ is the concentration of a chemical at the point $\mathbf{x}$ in space. Suppose this chemical is being *generated* (or *depleted*) at different rates at different regions in space. Thus, in the absence of diffusion, we would have the *generation equation*:

$$\partial_t \, p(\mathbf{x}, t) = q(\mathbf{x}),$$

where $q(\mathbf{x})$ is the rate at which the chemical is being created/destroyed at $\mathbf{x}$ (we assume that q is constant in time).

If we now include the effects of diffusion, we get the *generation–diffusion equation*:

$$\partial_t \, p = \kappa \triangle p + q.$$

A *steady state* of this equation is a scalar field p satisfying the *Poisson equation*:

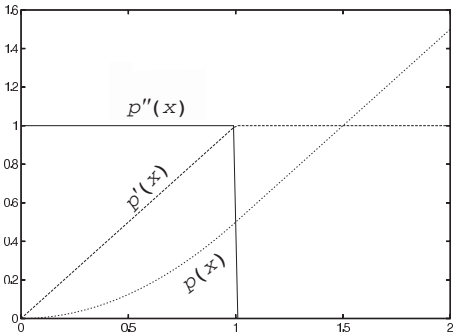
$$\boxed{\triangle p = Q,}$$

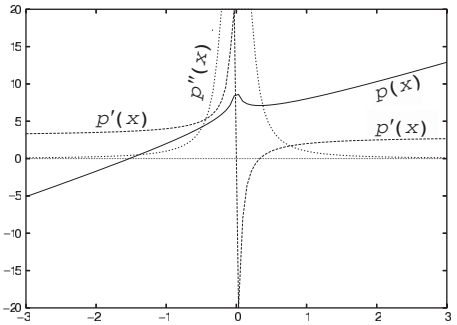
where $Q(\mathbf{x}) = -q(\mathbf{x})/\kappa$.

Example 1D.1 One-dimensional Poisson equation
If $D = 1$, then $\triangle p(x) = \partial_x^2 \, p(x) = p''(x)$; thus, the *one-dimensional Poisson equation* is just

$$\boxed{p''(x) = Q(x).}$$

We can solve this equation by twice-integrating the function $Q(x)$. If $p(x) = \int \int Q(x)$ is some double-antiderivative of G, then p clearly satisfies the Poisson equation. For example:





$$p''(x) = Q(x) = \begin{cases} 1 & \text{if } 0 < x < 1; \\ 0 & \text{otherwise;} \end{cases} \qquad p''(x) = Q(x) = 1/x^2;$$

$$p'(x) = \begin{cases} x & \text{if } 0 < x < 1; \\ 1 & \text{otherwise;} \end{cases} \qquad p'(x) = -1/x + 3;$$

$$p(x) = \begin{cases} x^2/2 & \text{if } 0 < x < 1; \\ x - \frac{1}{2} & \text{otherwise.} \end{cases} \qquad p(x) = -\log|x| + 3x + 5.$$

(a) (b)

Figure 1D.1. Two one-dimensional potentials.

(a) Suppose

$$Q(x) = \begin{cases} 1 & \text{if } 0 < x < 1; \\ 0 & \text{otherwise.} \end{cases}$$

Then define

$$p(x) = \int_0^x \int_0^y q(z)\mathrm{d}z\,\mathrm{d}y = \begin{cases} 0 & \text{if } x < 0; \\ x^2/2 & \text{if } 0 < x < 1; \\ x - \frac{1}{2} & \text{if } 1 < x \end{cases}$$

(Figure 1D.1(a));

(b) If $Q(x) = 1/x^2$ (for $x \neq 0$), then $p(x) = \int\int Q(x) = -\log|x| + ax + b$ (for $x \neq 0$), where $a, b \in \mathbb{R}$ are arbitrary constants (see Figure 1D.1(b)). ◇

Exercise 1D.1 Verify that the functions $p(x)$ in Examples 1D.1 (a) and (b) are both solutions to their respective Poisson equations. Ⓔ ♦

Example 1D.2 Electrical/gravitational fields

The Poisson equation also arises in classical field theory.[4] Suppose, for any point $\mathbf{x} = (x_1, x_2, x_3)$ in three-dimensional space, that $q(\mathbf{x})$ is the charge density at $\mathbf{x}$, and that $p(\mathbf{x})$ is the electric potential field at $\mathbf{x}$. Then we have the

[4] For a quick yet lucid introduction to electrostatics, see Stevens (1995), chap. 3.

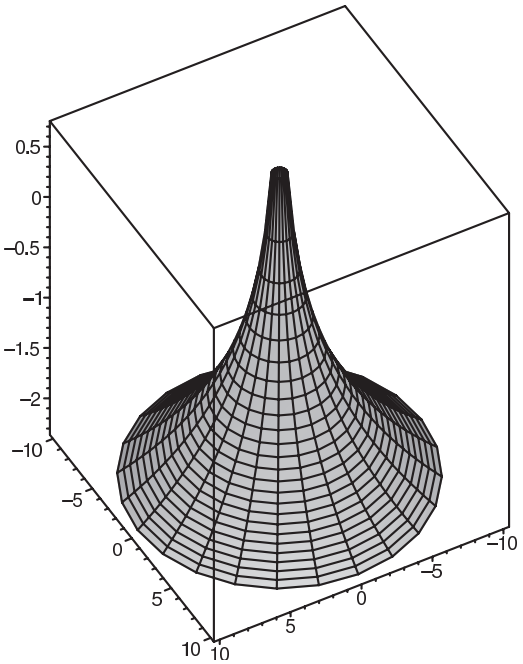


Figure 1D.2. The two-dimensional potential field generated by a concentration of charge at the origin.

following:

$$\triangle\, p(\mathbf{x}) = \kappa\, q(\mathbf{x}) \qquad (\kappa \text{ some constant}). \tag{1D.1}$$

If $q(\mathbf{x})$ were the *mass* density at $\mathbf{x}$ and $p(\mathbf{x})$ were the *gravitational* potential energy, then we would get the same equation. (See Figure 1D.2 for an example of such a potential in two dimensions.)

In particular, in a region where there is no charge/mass (i.e. where $q \equiv 0$), equation (1D.1) reduces to the Laplace equation $\triangle p \equiv 0$. Because of this, solutions to the Poisson equation (and especially the Laplace equation) are sometimes called *potentials*. ◇

Example 1D.3 The Coulomb potential
Let $D = 3$, and let

$$p(x, y, z) = \frac{1}{\|(x, y, z)\|} = \frac{1}{\sqrt{x^2 + y^2 + z^2}}.$$

In Example 1C.3(a), we asserted that $p(x, y, z)$ was harmonic everywhere except at $(0, 0, 0)$, where it is not well-defined. For physical reasons, it is 'reasonable' to

write the equation

$$\triangle p(0, 0, 0) = \delta_0, \tag{1D.2}$$

where δ_0 is the 'Dirac delta function' (representing an infinite concentration of charge at zero).[5] Then $p(x, y, z)$ describes the electric potential generated by a *point charge.* ◇

Exercise 1D.2 Check that Ⓔ

$$\nabla\, p(x, y, z) = \frac{-(x, y, z)}{\|(x, y, z)\|^3}.$$

This is the electric field generated by a point charge, as given by *Coulomb's law* from classical electrostatics. ⧫

Exercise 1D.3 Ⓔ

(a) Let $q : \mathbb{R}^3 \longrightarrow \mathbb{R}$ be a scalar field describing a charge density distribution. If $\vec{\mathbf{E}} : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ is the electric field generated by q, then *Gauss's law* says $\operatorname{div} \vec{\mathbf{E}} = \kappa\, q$, where κ is a constant. If $p : \mathbb{R}^3 \longrightarrow \mathbb{R}$ is the electric potential field associated with $\vec{\mathbf{E}}$, then, by definition, $\vec{\mathbf{E}} = \nabla p$. Use these facts to derive equation (1D.1).

(b) Suppose q is independent of the x_3 coordinate; that is, $q(x_1, x_2, x_3) = Q(x_1, x_2)$ for some function $Q : \mathbb{R}^2 \longrightarrow \mathbb{R}$. Show that p is also independent of the x_3 coordinate; that is, $p(x_1, x_2, x_3) = P(x_1, x_2)$ for some function $P : \mathbb{R}^2 \longrightarrow \mathbb{R}$. Show that P and Q satisfy the *two*-dimensional version of the Poisson equation, that is that $\triangle P = \kappa\, Q$.

(This is significant because many physical problems have (approximate) translational symmetry along one dimension (e.g. an electric field generated by a long, uniformly charged wire or plate). Thus, we can reduce the problem to two dimensions, where powerful methods from complex analysis can be applied; see Section 18B, p. 428.) ⧫

Note that the electric/gravitational potential field is *not uniquely defined* by equation (1D.1). If $p(\mathbf{x})$ solves the Poisson equation (1D.1), then so does $\widetilde{p}(\mathbf{x}) = p(\mathbf{x}) + a$ for any constant $a \in \mathbb{R}$. Thus we say that the potential field is well-defined *up to addition of a constant*; this is similar to the way in which the antiderivative

[5] Equation (1D.2) seems mathematically nonsensical but it *can* be made mathematically meaningful using *distribution theory*. However, this is far beyond the scope of this book, so, for our purposes, we will interpret equation (1D.2) as purely metaphorical.

$\int Q(x)$ of a function is only well-defined up to some constant.[6] This is an example of a more general phenomenon.

Proposition 1D.4 *Let $\mathbb{X} \subset \mathbb{R}^D$ be some domain, and let $p : \mathbb{X} \longrightarrow \mathbb{R}$ and $h : \mathbb{X} \longrightarrow \mathbb{R}$ be two functions on $\mathbb{X}$. Let $\widetilde{p}(\mathbf{x}) := p(\mathbf{x}) + h(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{X}$. Suppose that h is harmonic, i.e. $\triangle h \equiv 0$. If p satisfies the Poisson equation $\triangle p \equiv q$, then $\widetilde{p}$ also satisfies this Poisson equation.*

Ⓔ *Proof* **Exercise 1D.4.** *Hint:* Note that $\triangle \widetilde{p}(\mathbf{x}) = \triangle p(\mathbf{x}) + \triangle h(\mathbf{x})$. □

For example, if $Q(x) = 1/x^2$, as in Example 1D.1(b), then $p(x) = -\log(x)$ is a solution to the Poisson equation $p''(x) = 1/x^2$. If $h(x)$ is a one-dimensional harmonic function, then $h(x) = ax + b$ for some constants a and b (see Example 1C.1(a)). Thus $\widetilde{p}(x) = -\log(x) + ax + b$, and we have already seen that these are also valid solutions to this Poisson equation.

1E Properties of harmonic functions

Prerequisites: §1C, Appendix H(ii). **Prerequisites** (for proofs): §2A, §17G, Appendix E(iii).

Recall that a function $h : \mathbb{R}^D \longrightarrow \mathbb{R}$ is *harmonic* if $\triangle h \equiv 0$. Harmonic functions have nice geometric properties, which can be loosely summarized as 'smooth and gently curving'.

Theorem 1E.1 Mean value theorem
Let $f : \mathbb{R}^D \longrightarrow \mathbb{R}$ be a scalar field. Then f is harmonic if and only if f is integrable, and

$$\text{for any } \mathbf{x} \in \mathbb{R}^D, \text{ and any } R > 0, f(\mathbf{x}) = \frac{1}{A(R)} \int_{\mathbb{S}(\mathbf{x};R)} f(\mathbf{s}) d\mathbf{s}. \qquad (1E.1)$$

Here, $\mathbb{S}(\mathbf{x}; R) := \left\{\mathbf{s} \in \mathbb{R}^D;\ \|\mathbf{s} - \mathbf{x}\| = R\right\}$ is the $(D-1)$-dimensional sphere of radius R around $\mathbf{x}$, and $A(R)$ is the $(D-1)$-dimensional surface area of $\mathbb{S}(\mathbf{x}; R)$.

Ⓔ *Proof* **Exercise 1E.1.**
(a) Suppose f is integrable and statement (1E.1) is true. Use the *spherical means formula* for the Laplacian (Theorem 2A.1) to show that f is harmonic.

[6] For the purposes of the physical theory, this constant *does not matter*, because the field p is physically interpreted only by computing the *potential difference* between two points, and the constant a will always cancel out in this computation. Thus, the two potential fields $p(\mathbf{x})$ and $\widetilde{p}(\mathbf{x}) = p(\mathbf{x}) + a$ will generate identical physical predictions.

(b) Now suppose f is harmonic. Define $\phi : \mathbb{R}_{\neq} \longrightarrow \mathbb{R}$ by

$$\phi(R) := \frac{1}{A(R)} \int_{\mathbb{S}(\mathbf{x};R)} f(\mathbf{s}) \mathrm{d}\mathbf{s}.$$

Show that

$$\phi'(R) = \frac{K}{A(R)} \int_{\mathbb{S}(\mathbf{x};R)} \partial_{\perp} f(\mathbf{s}) \mathrm{d}\mathbf{s},$$

for some constant $K > 0$. Here, $\partial_{\perp} f(\mathbf{s})$ is the *outward normal derivative* of f at the point $\mathbf{s}$ on the sphere (see p. 567 for an abstract definition; see §5C(ii), p. 80, for more information).

(c) Let $\mathbb{B}(\mathbf{x}; R) := \left\{ \mathbf{b} \in \mathbb{R}^D ; \ \|\mathbf{b} - \mathbf{x}\| \leq R \right\}$ be the *ball* of radius R around $\mathbf{x}$. Apply *Green's Formula* (Theorem E.5(a), p. 567) to show that

$$\phi'(R) = \frac{K}{A(R)} \int_{\mathbb{B}(\mathbf{x};R)} \triangle f(\mathbf{b}) \mathrm{d}\mathbf{b}.$$

(d) Deduce that, if f is harmonic, then ϕ must be constant.

(e) Use the fact that f is continuous to show that $\lim_{r \to 0} \phi(r) = f(\mathbf{x})$. Deduce that $\phi(r) = f(\mathbf{x})$ for all $r \geq 0$. Conclude that, if f is harmonic, then statement (1E.1) must be true. □

Corollary 1E.2 Maximum principle for harmonic functions
Let $\mathbb{X} \subset \mathbb{R}^D$ be a domain, and let $u : \mathbb{X} \longrightarrow \mathbb{R}$ be a nonconstant harmonic function. Then u has no local maximal or minimal points anywhere in the interior of $\mathbb{X}$.

If $\mathbb{X}$ is bounded (hence compact)*, then u does obtain a maximum and minimum, but only on the boundary of $\mathbb{X}$.*

Proof (By contradiction.) Suppose $\mathbf{x}$ *was* a local maximum of u somewhere in the interior of $\mathbb{X}$. Let $R > 0$ be small enough that $\mathbb{S}(\mathbf{x}; R) \subset \mathbb{X}$, and such that

$$u(\mathbf{x}) \geq u(\mathbf{s}) \quad \text{for all } \mathbf{s} \in \mathbb{S}(\mathbf{x}; R), \tag{1E.2}$$

where this inequality is strict for at least one $\mathbf{s}_0 \in \mathbb{S}(\mathbf{x}; R)$.

Claim 1 *There is a nonempty open subset $\mathbb{Y} \subset \mathbb{S}(\mathbf{x}; R)$ such that $u(\mathbf{x}) > u(\mathbf{y})$ for all $\mathbf{y}$ in $\mathbb{Y}$.*

Proof We know that $u(\mathbf{x}) > u(\mathbf{s}_0)$. But u is continuous, so there must be some open neighbourhood $\mathbb{Y}$ around $\mathbf{s}_0$ such that $u(\mathbf{x}) > u(\mathbf{y})$ for all $\mathbf{y}$ in $\mathbb{Y}$. $\diamond_{\text{Claim 1}}$

Equation (1E.2) and Claim 1 imply that

$$f(\mathbf{x}) > \frac{1}{A(R)} \int_{\mathbb{S}(\mathbf{x};R)} f(\mathbf{s}) \mathrm{d}\mathbf{s}.$$

But this contradicts the mean value theorem. By contradiction, $\mathbf{x}$ cannot be a local maximum. (The proof for local minima is analogous.) □

A function $F : \mathbb{R}^D \longrightarrow \mathbb{R}$ is *spherically symmetric* if $F(\mathbf{x})$ depends only on the norm $\|\mathbf{x}\|$ (i.e. $F(\mathbf{x}) = f(\|\mathbf{x}\|)$ for some function $f : \mathbb{R}_{\not=} \longrightarrow \mathbb{R}$). For example, the function $F(\mathbf{x}) := \exp(-\|\mathbf{x}\|^2)$ is spherically symmetric.

If $h, F : \mathbb{R}^D \longrightarrow \mathbb{R}$ are two integrable functions, then their *convolution* is the function $h * F : \mathbb{R}^D \longrightarrow \mathbb{R}$ defined by

$$h * F(\mathbf{x}) := \int_{\mathbb{R}^D} h(\mathbf{y}) \cdot F(\mathbf{x} - \mathbf{y}) d\mathbf{y}, \quad \text{for all } \mathbf{x} \in \mathbb{R}^D$$

(if this integral converges). We will encounter convolutions in §10D(ii), p. 215 (where they will be used to prove the L^2-convergence of a Fourier series), and again in Chapter 17 (where they will be used to construct 'impulse-response' solutions for PDEs). For now, we state the following simple consequence of the mean value theorem.

Lemma 1E.3 *If $h : \mathbb{R}^D \longrightarrow \mathbb{R}$ is harmonic and $F : \mathbb{R}^D \longrightarrow \mathbb{R}$ is integrable and spherically symmetric, then $h * F = K \cdot h$, where $K \in \mathbb{R}$ is some constant.*

Ⓔ *Proof* **Exercise 1E.2.** □

Proposition 1E.4 Smoothness of harmonic functions
If $h : \mathbb{R}^D \longrightarrow \mathbb{R}$ is a harmonic function, then h is infinitely differentiable.

Proof Let $F : \mathbb{R}^D \longrightarrow \mathbb{R}$ be some infinitely differentiable, spherically symmetric, integrable function. For example, we could take $F(\mathbf{x}) := \exp(-\|\mathbf{x}\|^2)$. Then Proposition 17G.2(f), p. 415, says that $h * F$ is infinitely differentiable. But Lemma 1E.3 implies that $h * F = Kh$ for some constant $K \in \mathbb{R}$; thus, h is also infinitely differentiable.

(For another proof, see Evans (1991), §2.2, Theorem 6.) □

Actually, we can go even further than this. A function $h : \mathbb{X} \longrightarrow \mathbb{R}$ is *analytic* if, for every $\mathbf{x} \in \mathbb{X}$, there is a multivariate Taylor series expansion for h around $\mathbf{x}$ with a nonzero radius of convergence.[7]

Proposition 1E.5 Harmonic functions are analytic
Let $\mathbb{X} \subseteq \mathbb{R}^D$ be an open set. If $h : \mathbb{X} \longrightarrow \mathbb{R}$ is a harmonic function, then h is analytic on $\mathbb{X}$.

Proof For the case $D = 2$, see Corollary 18D.2, p. 456. For the general case $D \geq 2$, see Evans (1991), §2.2, Theorem 10. □

[7] See Appendices H(ii) and H(v), pp. 574, 579.

1F* Transport and diffusion

Prerequisites: §1B, §6A.

If $u : \mathbb{R}^D \longrightarrow \mathbb{R}$ is a 'mountain', then recall that $\nabla u(\mathbf{x})$ points in the direction of *most rapid ascent* at $\mathbf{x}$. If $\vec{\mathbf{v}} \in \mathbb{R}^D$ is a vector, then the dot product $\vec{\mathbf{v}} \bullet \nabla u(\mathbf{x})$ measures how rapidly you would be ascending if you walked in direction $\vec{\mathbf{v}}$.

Suppose $u : \mathbb{R}^D \longrightarrow \mathbb{R}$ describes a pile of leafs, and there is a steady wind blowing in the direction $\vec{\mathbf{v}} \in \mathbb{R}^D$. We would expect the pile to move slowly in the direction $\vec{\mathbf{v}}$. Suppose you were an observer fixed at location $\mathbf{x}$. The pile is moving past you in direction $\vec{\mathbf{v}}$, which is the same as you walking along the pile in direction $-\vec{\mathbf{v}}$; thus, you would expect the height of the pile at your location to increase/decrease at a rate $-\vec{\mathbf{v}} \bullet \nabla u(\mathbf{x})$. The pile thus satisfies the *transport equation*:

$$\partial_t u = -\vec{\mathbf{v}} \bullet \nabla u.$$

Now suppose that the wind does not blow in a *constant* direction, but instead has some complex spatial pattern. The wind velocity is therefore determined by a *vector field* $\vec{\mathbf{V}} : \mathbb{R}^D \longrightarrow \mathbb{R}^D$. As the wind picks up leaves and carries them around, the *flux* of leaves at a point $\mathbf{x} \in \mathbb{X}$ is then given by the vector $\vec{\mathbf{F}}(\mathbf{x}) = u(\mathbf{x}) \cdot \vec{\mathbf{V}}(\mathbf{x})$. But the rate at which leaves are piling up at each location is the *divergence* of the flux. This results in *Liouville's equation*:

$$\partial_t u = -\operatorname{div} \vec{\mathbf{F}} = -\operatorname{div}(u \cdot \vec{\mathbf{V}}) \underset{(*)}{=} -\vec{\mathbf{V}} \bullet \nabla u - u \cdot \operatorname{div} \vec{\mathbf{V}}.$$

Here, $(*)$ is by the Leibniz rule for divergence (Proposition E.2(b), p. 564).

Liouville's equation describes the rate at which u-material accumulates when it is being pushed around by the $\vec{\mathbf{V}}$-vector field. Another example: $\vec{\mathbf{V}}(\mathbf{x})$ describes the flow of water at $\mathbf{x}$, and $u(\mathbf{x})$ is the buildup of some sediment at $\mathbf{x}$.

Now suppose that, in addition to the deterministic force $\vec{\mathbf{V}}$ acting on the leaves, there is also a 'random' component. In other words, while being blown around by the wind, the leaves are also subject to some random diffusion. To describe this, we combine *Liouville's equation* with the *heat equation*, to obtain the *Fokker–Planck* equation:

$$\partial_t u = \kappa \, \triangle u - \vec{\mathbf{V}} \bullet \nabla u - u \cdot \operatorname{div} \vec{\mathbf{V}}.$$

1G* Reaction and diffusion

Prerequisites: §1B.

Suppose A, B and C are three chemicals, satisfying the following chemical reaction:

$$2A + B \Longrightarrow C.$$

As this reaction proceeds, the A and B species are consumed and C is produced. Thus, if a, b, and c are the concentrations of the three chemicals, we have

$$\partial_t c = R(t) = -\partial_t b = -\frac{1}{2}\partial_t a,$$

where $R(t)$ is the rate of the reaction at time t. The rate $R(t)$ is determined by the concentrations of A and B, and by a rate constant ρ. Each chemical reaction requires two molecules of A and one of B; thus, the reaction rate is given by

$$R(t) = \rho \cdot a(t)^2 \cdot b(t).$$

Hence, we get three ordinary differential equations, called the *reaction kinetic equations* of the system:

$$\left.\begin{aligned} \partial_t a(t) &= -2\rho \cdot a(t)^2 \cdot b(t), \\ \partial_t b(t) &= -\rho \cdot a(t)^2 \cdot b(t), \\ \partial_t c(t) &= \rho \cdot a(t)^2 \cdot b(t). \end{aligned}\right\} \tag{1G.1}$$

Now suppose that the chemicals A, B, and C are in solution, but are not uniformly mixed. At any location $\mathbf{x} \in \mathbb{X}$ and time $t > 0$, let $a(\mathbf{x}, t)$ be the concentration of chemical A at location $\mathbf{x}$ at time t; likewise, let $b(\mathbf{x}, t)$ be the concentration of B and $c(\mathbf{x}, t)$ be the concentration of C. (This determines three *time-varying scalar fields*, $a, b, c : \mathbb{R}^3 \times \mathbb{R} \longrightarrow \mathbb{R}$.) As the chemicals react, their concentrations at each point in space may change. Thus, the functions a, b, c will obey equations (1G.1) at each point in space. That is, for every $\mathbf{x} \in \mathbb{R}^3$ and $t \in \mathbb{R}$, we have

$$\partial_t a(\mathbf{x}; t) \approx -2\rho \cdot a(\mathbf{x}; t)^2 \cdot b(\mathbf{x}; t),$$

etc. However, the dissolved chemicals are also subject to *diffusion* forces. In other words, each of the functions a, b, and c will also be obeying the heat equation. Thus, we obtain the following system:

$$\begin{aligned} \partial_t a &= \kappa_a \cdot \triangle a(\mathbf{x}; t) - 2\rho \cdot a(\mathbf{x}; t)^2 \cdot b(\mathbf{x}; t), \\ \partial_t b &= \kappa_b \cdot \triangle b(\mathbf{x}; t) - \rho \cdot a(\mathbf{x}; t)^2 \cdot b(\mathbf{x}; t), \\ \partial_t c &= \kappa_c \cdot \triangle c(\mathbf{x}; t) + \rho \cdot a(\mathbf{x}; t)^2 \cdot b(\mathbf{x}; t), \end{aligned}$$

where $\kappa_a, \kappa_b, \kappa_c > 0$ are three different diffusivity constants.

This is an example of a *reaction–diffusion system*. In general, in a reaction–diffusion system involving N distinct chemicals, the concentrations of the different species are described by a *concentration vector field* $\mathbf{u} : \mathbb{R}^3 \times \mathbb{R} \longrightarrow \mathbb{R}^N$, and the chemical reaction is described by a *rate function* $F : \mathbb{R}^N \longrightarrow \mathbb{R}^N$. For example, in the previous example, $\mathbf{u}(\mathbf{x}, t) = (a(\mathbf{x}, t), b(\mathbf{x}, t), c(\mathbf{x}, t))$, and

$$F(a, b, c) = \big[-2\rho a^2 b, -\rho a^2 b, \rho a^2 b\big].$$

The *reaction–diffusion equations* for the system then take the form

$$\partial_t u_n = \kappa_n \triangle u_n + F_n(\mathbf{u}),$$

for $n = 1, \ldots, N$.

1H Further reading

An analogy of the Laplacian can be defined on any Riemannian manifold, where it is sometimes called the *Laplace–Beltrami operator*. The study of harmonic functions on manifolds yields important geometric insights (Warner (1983), Chavel (1993)).

The reaction–diffusion systems from §1G play an important role in modern mathematical biology (Murray (1993)).

The heat equation also arises frequently in the theory of Brownian motion and other Gaussian stochastic processes on $\mathbb{R}^D$ (Strook (1993)).

1I Practice problems

1.1 Let $f : \mathbb{R}^4 \longrightarrow \mathbb{R}$ be a differentiable scalar field. Show that $\operatorname{div} \nabla f(x_1, x_2, x_3, x_4) = \triangle f(x_1, x_2, x_3, x_4)$.

1.2 Let $f(x, y; t) = \exp(-34t) \cdot \sin(3x + 5y)$. Show that $f(x, y; t)$ satisfies the two-dimensional heat equation $\partial_t f(x, y; t) = \triangle f(x, y; t)$.

1.3 Let $u(x, y) = \log(x^2 + y^2)$. Show that $u(x, y)$ satisfies the (two-dimensional) Laplace equation, everywhere except at $(x, y) = (0, 0)$. *Remark:* If $(x, y) \in \mathbb{R}^2$, recall that $\|(x, y)\| := \sqrt{x^2 + y^2}$. Thus, $\log(x^2 + y^2) = 2 \log \|(x, y)\|$. This function is sometimes called the *logarithmic potential*.

1.4 If $(x, y, z) \in \mathbb{R}^3$, recall that $\|(x, y, z)\| := \sqrt{x^2 + y^2 + z^2}$. Define

$$u(x, y, z) = \frac{1}{\|(x, y, z)\|} = \frac{1}{\sqrt{x^2 + y^2 + z^2}}.$$

Show that u satisfies the (three-dimensional) Laplace equation, everywhere except at $(x, y, z) = (0, 0, 0)$. *Remark:* Observe that

$$\nabla u(x, y, z) = \frac{-(x, y, z)}{\|(x, y, z)\|^3}.$$

What force field does this remind you of? *Hint:* $u(x, y, z)$ is sometimes called the *Coulomb potential*.

1.5 Let

$$u(x, y; t) = \frac{1}{4\pi t} \exp\left(\frac{-\|(x, y)\|^2}{4t}\right) = \frac{1}{4\pi t} \exp\left(\frac{-x^2 - y^2}{4t}\right)$$

be the (two-dimensional) *Gauss–Weierstrass kernel*. Show that u satisfies the (two-dimensional) heat equation $\partial_t u = \triangle u$.

1.6 Let α and β be real numbers, and let $h(x, y) = \sinh(\alpha x) \cdot \sin(\beta y)$.

(a) Compute $\triangle h(x, y)$.
(b) Suppose h is *harmonic*. Write an equation describing the relationship between α and β.

2

Waves and signals

There is geometry in the humming of the strings.
Pythagoras

2A The Laplacian and spherical means

Prerequisites: Appendix A, Appendix B, Appendix H(v). **Recommended:** §1B.

Let $u : \mathbb{R}^D \longrightarrow \mathbb{R}$ be a function of D variables. Recall that the Laplacian of u is defined as follows:

$$\boxed{\triangle u = \partial_1^2 u + \partial_2^2 u + \cdots + \partial_D^2 u.}$$

In this section, we will show that $\triangle u(\mathbf{x})$ measures the discrepancy between $u(\mathbf{x})$ and the 'average' of u in a small neighbourhood around $\mathbf{x}$.

Let $\mathbb{S}(\epsilon)$ be the D-dimensional 'sphere' of radius ϵ around 0.

- If $D = 1$, then $\mathbb{S}(\epsilon)$ is just a set with two points: $\mathbb{S}(\epsilon) = \{-\epsilon, +\epsilon\}$.
- If $D = 2$, then $\mathbb{S}(\epsilon)$ is the *circle* of radius ϵ: $\mathbb{S}(\epsilon) = \{(x, y) \in \mathbb{R}^2; x^2 + y^2 = \epsilon^2\}$.
- If $D = 3$, then $\mathbb{S}(\epsilon)$ is the three-dimensional spherical shell of radius ϵ:

$$\mathbb{S}(\epsilon) = \left\{(x, y, z) \in \mathbb{R}^3;\ x^2 + y^2 + z^2 = \epsilon^2\right\}.$$

- More generally, for any dimension D,

$$\mathbb{S}(\epsilon) = \left\{(x_1, x_2, \ldots, x_D) \in \mathbb{R}^D;\ x_1^2 + x_2^2 + \cdots + x_D^2 = \epsilon^2\right\}.$$

Let A_ϵ be the 'surface area' of the sphere.

- If $D = 1$, then $\mathbb{S}(\epsilon) = \{-\epsilon, +\epsilon\}$ is a finite set, with two points, so we say $A_\epsilon = 2$.
- If $D = 2$, then $\mathbb{S}(\epsilon)$ is the circle of radius ϵ; the *perimeter* of this circle is $2\pi\epsilon$, so we say $A_\epsilon = 2\pi\epsilon$.
- If $D = 3$, then $\mathbb{S}(\epsilon)$ is the sphere of radius ϵ, which has *surface area* $4\pi\epsilon^2$.

Let

$$\mathbf{M}_\epsilon\, f(0) := \frac{1}{A_\epsilon} \int_{\mathbb{S}(\epsilon)} f(\mathbf{s}) \mathrm{d}\mathbf{s};$$

then $\mathbf{M}_\epsilon\, f(0)$ is the *average value* of $f(\mathbf{s})$ over all $\mathbf{s}$ on the surface of the ϵ-radius sphere around 0, which is called the *spherical mean* of f at 0. The interpretation of the integral sign '$\int$' depends on the dimension D of the space. For example,'$\int$' represents a *surface integral* if $D = 3$, a *line integral* if $D = 2$, and a simple two-point sum if $D = 1$. Thus we have the following.

- If $D = 1$, then $\mathbb{S}(\epsilon) = \{-\epsilon, +\epsilon\}$, so that $\int_{\mathbb{S}(\epsilon)} f(\mathbf{s})\mathrm{d}\mathbf{s} = f(\epsilon) + f(-\epsilon)$; thus,

$$\mathbf{M}_\epsilon\, f \;= \frac{f(\epsilon) + f(-\epsilon)}{2}.$$

- If $D = 2$, then any point on the circle has the form $(\epsilon \cos(\theta), \epsilon \sin(\theta))$ for some angle $\theta \in [0, 2\pi)$. Thus,

$$\int_{\mathbb{S}(\epsilon)} f(\mathbf{s})\mathrm{d}\mathbf{s} = \int_0^{2\pi} f\Big(\epsilon \cos(\theta), \epsilon \sin(\theta)\Big)\, \epsilon\, \mathrm{d}\theta,$$

  so that

$$\mathbf{M}_\epsilon\, f = \frac{1}{2\pi\epsilon} \int_0^{2\pi} f\Big(\epsilon \cos(\theta), \epsilon \sin(\theta)\Big)\, \epsilon \mathrm{d}\theta = \frac{1}{2\pi} \int_0^{2\pi} f\Big(\epsilon \cos(\theta), \epsilon \sin(\theta)\Big)\, \mathrm{d}\theta.$$

Likewise, for any $\mathbf{x} \in \mathbb{R}^D$, we define

$$\mathbf{M}_\epsilon\, f(\mathbf{x}) := \frac{1}{A_\epsilon} \int_{\mathbb{S}(\epsilon)} f(\mathbf{x} + \mathbf{s})\mathrm{d}\mathbf{s}$$

to be the average value of f over an ϵ-radius sphere around $\mathbf{x}$. Suppose $f : \mathbb{R}^D \longrightarrow \mathbb{R}$ is a smooth scalar field, and $\mathbf{x} \in \mathbb{R}^D$. One interpretation of the Laplacian is as follows: $\triangle f(\mathbf{x})$ measures the disparity between $f(\mathbf{x})$ and the *average value* of f in the immediate vicinity of $\mathbf{x}$. This is the meaning of the following theorem.

Theorem 2A.1

(a) *If $f : \mathbb{R} \longrightarrow \mathbb{R}$ is a smooth scalar field, then (as shown in Figure 2A.1), for any $x \in \mathbb{R}$,*

$$\triangle f(x) = \lim_{\epsilon \to 0} \frac{2}{\epsilon^2} \Big[\mathbf{M}_\epsilon\, f(\mathbf{x}) - f(\mathbf{x})\Big] = \lim_{\epsilon \to 0} \frac{2}{\epsilon^2} \left[\frac{f(x-\epsilon) + f(x+\epsilon)}{2} \;-\; f(x)\right].$$

(b)[1] *If $f : \mathbb{R}^D \longrightarrow \mathbb{R}$ is a smooth scalar field, then, for any $\mathbf{x} \in \mathbb{R}^D$,*

$$\triangle f(\mathbf{x}) = \lim_{\epsilon \to 0} \frac{C}{\epsilon^2} \Big[\mathbf{M}_\epsilon\, f(\mathbf{x}) - f(\mathbf{x})\Big] = \lim_{\epsilon \to 0} \frac{C}{\epsilon^2} \left[\frac{1}{A_\epsilon} \int_{\mathbb{S}(\epsilon)} f(\mathbf{x} + \mathbf{s})\mathrm{d}\mathbf{s} - f(\mathbf{x})\right].$$

(Here C is a constant determined by the dimension D.)

[1] Part (b) of Theorem 2A.1 is not necessary for the physical derivation of the wave equation which appears later in this chapter. However, part (b) *is* required to prove the Mean Value Theorem for harmonic functions (Theorem 1E.1, p. 18).

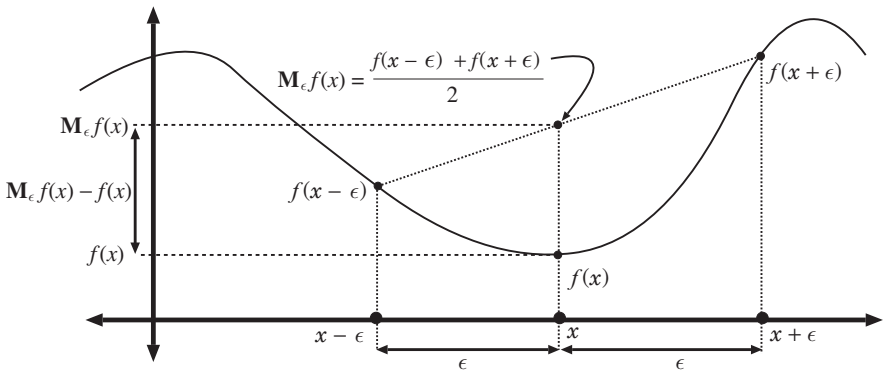


Figure 2A.1. Local averages: $f(x)$ vs. $\mathbf{M}_\epsilon\, f(x) := \frac{f(x-\epsilon)+f(x+\epsilon)}{2}$.

Proof (a) Using *Taylor's theorem* (see Appendix H(i), p. 573), we have

$$f(x+\epsilon) = f(x) + \epsilon f'(x) + \frac{\epsilon^2}{2} f''(x) + \mathcal{O}(\epsilon^3)$$

and

$$f(x-\epsilon) = f(x) - \epsilon f'(x) + \frac{\epsilon^2}{2} f''(x) + \mathcal{O}(\epsilon^3).$$

Here, $f'(x) = \partial_x\, f(x)$ and $f''(x) = \partial_x^2\, f(x)$. The expression '$\mathcal{O}(\epsilon)$' means 'some function (we don't care which one) such that $\lim_{\epsilon\to 0} \mathcal{O}(\epsilon) = 0$'.[2] Likewise, '$\mathcal{O}(\epsilon^3)$' means 'some function (we don't care which one) such that $\lim_{\epsilon\to 0} \frac{\mathcal{O}(\epsilon^3)}{\epsilon^2} = 0$.' Summing these two equations, we obtain

$$f(x+\epsilon) + f(x-\epsilon) = 2f(x) + \epsilon^2 \cdot f''(x) + \mathcal{O}(\epsilon^3).$$

Thus,

$$\frac{f(x-\epsilon) - 2f(x) + f(x+\epsilon)}{\epsilon^2} = f''(x) + \mathcal{O}(\epsilon)$$

(because $\mathcal{O}(\epsilon^3)/\epsilon^2 = \mathcal{O}(\epsilon)$). Now take the limit as $\epsilon \to 0$ to obtain

$$\lim_{\epsilon\to 0} \frac{f(x-\epsilon) - 2f(x) + f(x+\epsilon)}{\epsilon^2} = \lim_{\epsilon\to 0}\ f''(x) + \mathcal{O}(\epsilon) = f''(x) = \triangle f(x),$$

as desired.

[2] Actually, '$\mathcal{O}(\epsilon)$' means slightly more than this – see Appendix H(i). However, for our purposes, this will be sufficient.

(b) Define the *Hessian derivative matrix* of f at $\mathbf{x}$:

$$\mathsf{D}^2 f(\mathbf{x}) = \begin{bmatrix} \partial_1^2 f & \partial_1\partial_2 f & \dots & \partial_1\partial_D f \\ \partial_2\partial_1 f & \partial_2^2 f & \dots & \partial_2\partial_D f \\ \vdots & \vdots & \ddots & \vdots \\ \partial_D\partial_1 f & \partial_D\partial_2 f & \dots & \partial_D^2 f \end{bmatrix}.$$

Then, for any $\mathbf{s} \in \mathbb{S}(\epsilon)$, the *multivariate Taylor theorem* (see Appendix H(v), p. 579) says:

$$f(\mathbf{x}+\mathbf{s}) = f(\mathbf{x}) + \mathbf{s} \bullet \nabla f(\mathbf{x}) + \frac{1}{2}\mathbf{s} \bullet \mathsf{D}^2 f(\mathbf{x}) \cdot \mathbf{s} + \mathcal{O}(\epsilon^3).$$

Now, if $\mathbf{s} = (s_1, s_2, \dots, s_D)$, then $\mathbf{s} \bullet \mathsf{D}^2 f(\mathbf{x}) \cdot \mathbf{s} = \sum_{c,d=1}^{D} s_c \cdot s_d \cdot \partial_c\partial_d\, f(\mathbf{x})$. Thus, for any $\epsilon > 0$, we have

$$\begin{aligned}
A_\epsilon \cdot \mathbf{M}_\epsilon\, f(\mathbf{x}) &= \int_{\mathbb{S}(\epsilon)} f(\mathbf{x}+\mathbf{s})\mathrm{d}\mathbf{s} \\
&= \int_{\mathbb{S}(\epsilon)} f(\mathbf{x})\mathrm{d}\mathbf{s} + \int_{\mathbb{S}(\epsilon)} \mathbf{s} \bullet \nabla f(\mathbf{x})\mathrm{d}\mathbf{s} \\
&\qquad + \frac{1}{2}\int_{\mathbb{S}(\epsilon)} \mathbf{s} \bullet \mathsf{D}^2 f(\mathbf{x}) \cdot \mathbf{s}\,\mathrm{d}\mathbf{s} + \int_{\mathbb{S}(\epsilon)} \mathcal{O}(\epsilon^3)\mathrm{d}\mathbf{s} \\
&= A_\epsilon f(\mathbf{x}) + \nabla f(\mathbf{x}) \bullet \int_{\mathbb{S}(\epsilon)} \mathbf{s}\,\mathrm{d}\mathbf{s} \\
&\qquad + \frac{1}{2}\int_{\mathbb{S}(\epsilon)} \left(\sum_{c,d=1}^{D} s_c s_d \cdot \partial_c\partial_d\, f(\mathbf{x}) \right) \mathrm{d}\mathbf{s} + \mathcal{O}(\epsilon^{D+2}) \\
&= A_\epsilon f(\mathbf{x}) + \underbrace{\nabla f(\mathbf{x}) \bullet \mathbf{0}}_{(*)} \\
&\qquad + \frac{1}{2}\sum_{c,d=1}^{D} \left(\partial_c\, \partial_d\, f(\mathbf{x}) \cdot \left(\int_{\mathbb{S}(\epsilon)} s_c s_d\, \mathrm{d}\mathbf{s} \right) \right) + \mathcal{O}(\epsilon^{D+2}) \\
&= A_\epsilon f(\mathbf{x}) + \frac{1}{2}\underbrace{\sum_{d=1}^{D} \left(\partial_d^2\, f(\mathbf{x}) \cdot \left(\int_{\mathbb{S}(\epsilon)} s_d^2\, \mathrm{d}\mathbf{s} \right) \right)}_{(\dagger)} + \mathcal{O}(\epsilon^{D+2}) \\
&= A_\epsilon f(\mathbf{x}) + \frac{1}{2} \triangle f(\mathbf{x}) \cdot \epsilon^{D+1} K + \mathcal{O}(\epsilon^{D+2}),
\end{aligned}$$

where $K := \int_{\mathbb{S}(1)} s_1^2\, \mathrm{d}\mathbf{s}$. Here, $(*)$ is because $\int_{\mathbb{S}(\epsilon)} \mathbf{s}\,\mathrm{d}\mathbf{s} = \mathbf{0}$, because the centre-of-mass of a sphere is at its centre, namely $\mathbf{0}$; $(\dagger)$ is because, if $c, d \in [1 \dots D]$, and
Ⓔ $c \neq d$, then $\int_{\mathbb{S}(\epsilon)} s_c s_d \mathrm{d}\mathbf{s} = 0$ (**Exercise 2A.1** *Hint:* Use symmetry.)

Thus,

$$A_\epsilon \cdot \mathbf{M}_\epsilon f(\mathbf{x}) - A_\epsilon f(\mathbf{x}) = \frac{\epsilon^{D+1} K}{2} \triangle f(\mathbf{x}) + \mathcal{O}(\epsilon^{D+2}),$$

so

$$\begin{aligned}\mathbf{M}_\epsilon f(\mathbf{x}) - f(\mathbf{x}) &= \frac{\epsilon^{D+1} K}{2A_\epsilon} \triangle f(\mathbf{x}) + \frac{1}{A_\epsilon}\mathcal{O}(\epsilon^{D+2}) \\ &\underset{(*)}{=} \frac{\epsilon^{D+1} K}{2A_1 \cdot \epsilon^{D-1}} \triangle f(\mathbf{x}) + \mathcal{O}\left(\frac{\epsilon^{D+2}}{\epsilon^{D-1}}\right) \\ &= \frac{\epsilon^2 K}{2A_1} \triangle f(\mathbf{x}) + \mathcal{O}(\epsilon^3),\end{aligned}$$

where $(*)$ is because $A_\epsilon = A_1 \cdot \epsilon^{D-1}$. Thus,

$$\frac{2A_1}{K\,\epsilon^2}\Big(\mathbf{M}_\epsilon f(\mathbf{x}) - f(\mathbf{x})\Big) = \triangle f(\mathbf{x}) + \mathcal{O}(\epsilon).$$

Now take the limit as $\epsilon \to 0$ and set $C := 2A_1/K$ to prove part (b). □

Exercise 2A.2 Let $f : \mathbb{R}^D \longrightarrow \mathbb{R}$ be a smooth scalar field, such that $\mathbf{M}_\epsilon f(\mathbf{x}) = f(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^D$. Show that f is harmonic. Ⓔ ♦

2B The wave equation

Prerequisites: §2A.

2B(i) . . . in one dimension: the string

We want to describe mathematically vibrations propagating through a stretched elastic cord. We will represent the cord with a one-dimensional domain $\mathbb{X}$; either $\mathbb{X} = [0, L]$ or $\mathbb{X} = \mathbb{R}$. We will make several simplifying assumptions as follows.

(W1) The cord has uniform thickness and density. Thus, there is a constant *linear density* $\rho > 0$, so that a cord segment of length ℓ has mass $\rho\ell$.

(W2) The cord is *perfectly elastic*, meaning that it is infinitely flexible and does not resist bending in any way. Likewise, there is no air friction to resist the motion of the cord.

(W3) The only force acting on the cord is *tension*, which is a force of magnitude T pulling the cord to the right, balanced by an equal but opposite force of magnitude $-T$ pulling the cord to the left. These two forces are in balance, so the cord exhibits no horizontal motion. The tension T must be constant along the whole length of the cord. Thus, the

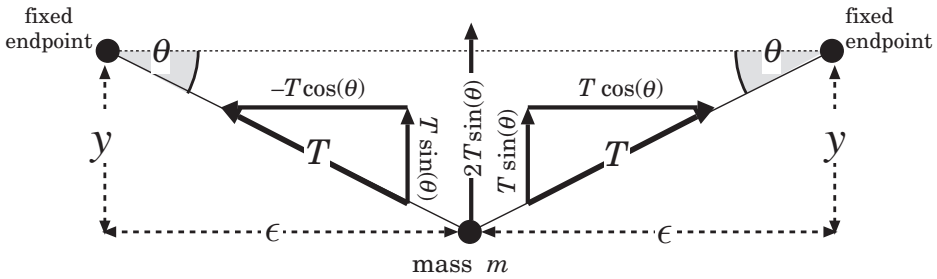


Figure 2B.1. Bead on a string.

equilibrium state for the cord is to be perfectly straight. Vibrations are deviations from this straight position.[3]

(W4) The vibrational motion of the cord is entirely *vertical*; there is no horizontal component to the vibration. Thus, we can describe the motion using a scalar-valued function $u(x, t)$, where $u(x, t)$ is the vertical displacement of the cord (from its flat equilibrium) at point x at time t. We assume that $u(x, t)$ is relatively small relative to the length of the cord, so that the cord is not significantly stretched by the vibrations.[4]

For simplicity, imagine first a single bead of mass m suspended at the midpoint of a (massless) elastic cord of length 2ϵ, stretched between two endpoints. Suppose we displace the bead by a distance y from its equilibrium, as shown in Figure 2B.1. The tension force T now pulls the bead diagonally towards each endpoint with force T. The horizontal components of the two tension forces are equal and opposite, so they cancel and the bead experiences no net horizontal force. Suppose the cord makes an angle θ with the horizontal; then the vertical component of each tension force is $T\sin(\theta)$, so the total vertical force acting on the bead is $2T\sin(\theta)$. But $\theta = \arctan(\epsilon/y)$ by the geometry of the triangles in Figure 2B.1, so $\sin(\theta) = \frac{y}{\sqrt{y^2+\epsilon^2}}$. Thus, the vertical force acting on the bead is given by

$$F = 2T\sin(\theta) = \frac{2Ty}{\sqrt{y^2+\epsilon^2}}. \tag{2B.1}$$

Now we return to our original problem of the vibrating string. Imagine that we replace the string with a 'necklace' made up of small beads of mass m separated by massless elastic strings of length ϵ as shown in Figure 2B.2. Each of these beads, in isolation, behaves like the 'bead on a string' in Figure 2B.1. However, now

[3] We could also incorporate the force of gravity as a constant downward force. In this case, the equilibrium position for the cord is to sag downwards in a 'catenary' curve. Vibrations are then deviations from this curve. This does not change the mathematics of this derivation, so we will assume for simplicity that gravity is absent and the cord is straight.

[4] If $u(x, t)$ were large, then the vibrations would stretch the cord and a *restoring force* would act against this stretching, as described by *Hooke's law*. By assuming that the vibrations are small, we are assuming we can neglect Hooke's law.

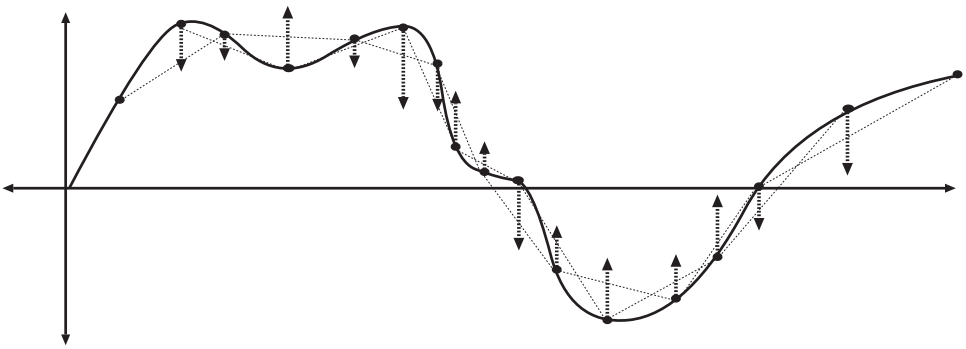


Figure 2B.2. Each bead feels a negative force proportional to its deviation from the local average.

the vertical displacement of each bead is not computed relative to the horizontal, but instead relative to the *average height* of the two neighbouring beads. Thus, in equation (2B.1), we set $y := u(x) - \mathbf{M}_\epsilon\, u(x)$, where $u(x)$ is the height of bead x, and where $\mathbf{M}_\epsilon\, u := \frac{1}{2}[u(x-\epsilon) + u(x+\epsilon)]$ is the average of its neighbours. Substituting this into equation (2B.1), we obtain

$$F_\epsilon(x) = \frac{2T[u(x) - \mathbf{M}_\epsilon\, u(x)]}{\sqrt{[u(x) - \mathbf{M}_\epsilon\, u(x)]^2 + \epsilon^2}}. \tag{2B.2}$$

(The 'ϵ' subscript in 'F_ϵ' is to remind us that this is just an ϵ-bead approximation of the real string.) Each bead represents a length-ϵ segment of the original string, so if the string has linear density ρ, then each bead must have mass $m_\epsilon := \rho\epsilon$. Thus, by Newton's law, the vertical acceleration of bead x must be as follows:

$$\begin{aligned} a_\epsilon(x) = \frac{F_\epsilon(x)}{m_\epsilon} &= \frac{2T[u(x) - \mathbf{M}_\epsilon\, u(x)]}{\rho\,\epsilon\,\sqrt{[u(x) - \mathbf{M}_\epsilon\, u(x)]^2 + \epsilon^2}} \\ &= \frac{2T[u(x) - \mathbf{M}_\epsilon\, u(x)]}{\rho\,\epsilon^2\,\sqrt{[u(x) - \mathbf{M}_\epsilon\, u(x)]^2/\epsilon^2 + 1}}. \end{aligned} \tag{2B.3}$$

Now we take the limit as $\epsilon \to 0$ to calculate the vertical acceleration of the string at x:

$$\begin{aligned} a(x) = \lim_{\epsilon\to 0} a_\epsilon(x) &= \frac{T}{\rho} \lim_{\epsilon\to 0} \frac{2}{\epsilon^2}\Big[u(x) - \mathbf{M}_\epsilon\, u(x)\Big] \cdot \lim_{\epsilon\to 0} \frac{1}{\sqrt{[u(x) - \mathbf{M}_\epsilon\, u(x)]^2/\epsilon^2 + 1}} \\ &\underset{(*)}{=} \frac{T}{\rho}\,\partial_x^2\, u(x) \frac{1}{\lim_{\epsilon\to 0}\sqrt{\epsilon^2 \cdot \partial_x^2\, u(x)^2 + 1}} \underset{(\dagger)}{=} \frac{T}{\rho}\,\partial_x^2\, u(x). \end{aligned} \tag{2B.4}$$

Here, $(*)$ is because Theorem 2A.1(a), p. 26, says that $\lim_{\epsilon\to 0} \frac{2}{\epsilon^2}[u(x) - \mathbf{M}_\epsilon\, u(x)] = \partial_x^2\, u(x)$. Finally, $(\dagger)$ is because, for any value of $u'' \in \mathbb{R}$, we have

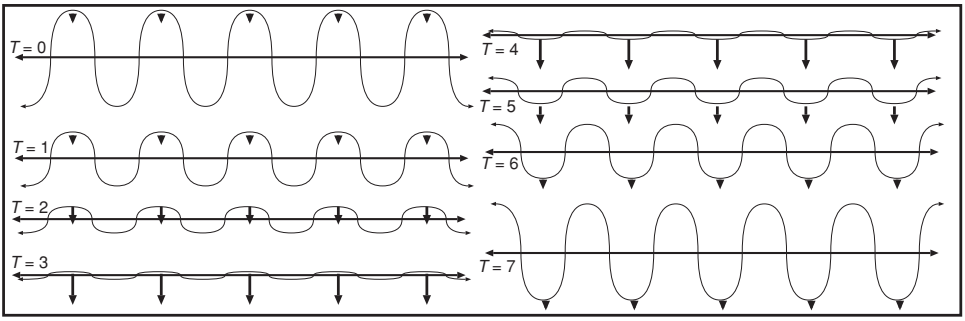


Figure 2B.3. One-dimensional standing wave.

$\lim_{\epsilon\to 0} \sqrt{\epsilon^2 u'' + 1} = 1$. We conclude that

$$a(x) = \frac{T}{\rho}\,\partial_x^2\, u(x) = \lambda^2\,\partial_x^2\, u(x),$$

where $\lambda := \sqrt{T/\rho}$. Now, the position (and hence velocity and acceleration) of the cord is changing in time. Thus, a and u are functions of t as well as x. And of course the acceleration $a(x,t)$ is nothing more than the second derivative of u with respect to t. Hence we have the *one-dimensional* wave equation:

$$\boxed{\partial_t^2\, u(x,t) = \lambda^2 \cdot \partial_x^2\, u(x,t).}$$

This equation describes the propagation of a transverse wave along an idealized string, or electrical pulses propagating in a wire.

Example 2B.1 Standing waves

(a) Suppose $\lambda^2 = 4$, and let $u(x;t) = \sin(3x)\cdot\cos(6t)$. Then u satisfies the wave equation and describes a *standing wave* with a *temporal frequency* of 6 and a *wave number* (or *spatial frequency*) of 3 (see Figure 2B.3).

(b) More generally, fix $\omega > 0$; if $u(x;t) = \sin(\omega\cdot x)\cdot\cos(\lambda\cdot\omega\cdot t)$, then u satisfies the wave equation and describes a *standing wave* of *temporal frequency* $\lambda\cdot\omega$ and *wave number* ω. ◇

Ⓔ **Exercise 2B.1** Verify Examples 2B.1(a) and (b). ♦

Example 2B.2 Travelling waves

(a) Suppose $\lambda^2 = 4$, and let $u(x;t) = \sin(3x - 6t)$. Then u satisfies the wave equation and describes a *sinusoidal travelling wave* with *temporal frequency* 6 and *wave number* 3. The wave crests move rightwards along the cord with velocity 2 (see Figure 2B.4(a)).

(b) More generally, fix $\omega \in \mathbb{R}$ and let $u(x;t) = \sin(\omega\cdot x - \lambda\cdot\omega\cdot t)$. Then u satisfies the wave equation and describes a *sinusoidal travelling wave* of *wave number* ω. The wave crests move rightwards along the cord with velocity λ.

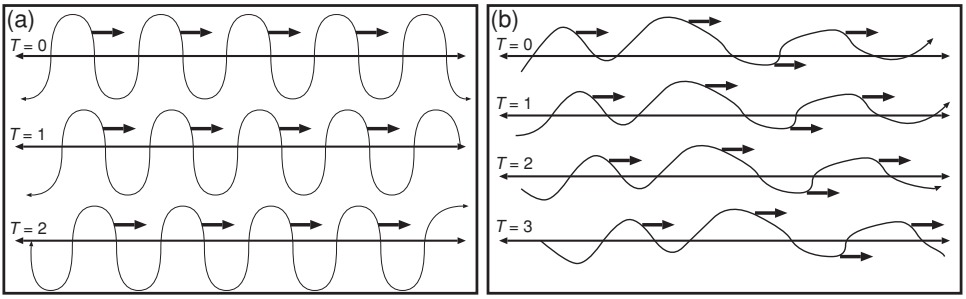


Figure 2B.4. (a) One-dimensional sinusoidal travelling wave. (b) General one-dimensional travelling wave.

(c) More generally, suppose that f is any function of one variable, and define $u(x;t) = f(x - \lambda \cdot t)$. Then u satisfies the wave equation and describes a *travelling wave*, whose shape is given by f, and which moves rightwards along the cord with velocity λ (see Figure 2B.4(b)). ◇

Exercise 2B.2 Verify Examples 2B.2(a)–(c). ⧫ Ⓔ

Exercise 2B.3 According to Example 2B.2(c), you can turn any function into a travelling wave. Can you turn any function into a standing wave? Why or why not? Ⓔ ⧫

2B(ii) . . . in two dimensions: the drum

Now suppose $D = 2$, and imagine a two-dimensional 'rubber sheet'. Suppose $u(x, y; t)$ is the vertical displacement of the rubber sheet at the point $(x, y) \in \mathbb{R}^2$ at time t. To derive the two-dimensional wave equation, we approximate this rubber sheet as a two-dimensional 'mesh' of tiny beads connected by massless, tense elastic strings of length ϵ. Each bead (x, y) feels a net vertical force $F = F_x + F_y$, where F_x is the vertical force arising from the tension in the x direction and F_y is the vertical force from the tension in the y direction. Both of these are expressed by a formula similar to equation (2B.2). Thus, if bead (x, y) has mass m_ϵ, then it experiences an acceleration $a = F/m_\epsilon = F_x/m_\epsilon + F_y/m_\epsilon = a_x + a_y$, where $a_x := F_x/m_\epsilon$ and $a_y := F_y/m_\epsilon$, and each of these is expressed by a formula similar to equation (2B.3). Taking the limit as $\epsilon \to 0$, as in equation (2B.4), we deduce that

$$a(x, y) = \lim_{\epsilon \to 0} a_{x,\epsilon}(x, y) + \lim_{\epsilon \to 0} a_{y,\epsilon}(x, y) = \lambda^2 \, \partial_x^2 \, u(x, y) + \lambda^2 \, \partial_y^2 \, u(x, y),$$

where λ is a constant determined by the density and tension of the rubber membrane. Again, we recall that u and a are also functions of time, and that

$a(x, y; t) = \partial_t^2 u(x, y; t)$. Thus, we have the *two-dimensional* wave equation:

$$\boxed{\partial_t^2 u(x, y; t) = \lambda^2 \cdot \partial_x^2 u(x, y; t) + \lambda^2 \cdot \partial_y^2 u(x, y; t),} \tag{2B.5}$$

or, more abstractly,

$$\boxed{\partial_t^2 u = \lambda^2 \cdot \triangle u.}$$

This equation describes the propagation of wave energy through any medium with a linear restoring force. For example:

- transverse waves on an idealized rubber sheet;
- ripples on the surface of a pool of water;
- acoustic vibrations on a drumskin.

Example 2B.3 Two-dimensional standing waves

(a) Suppose $\lambda^2 = 9$, and let $u(x, y; t) = \sin(3x) \cdot \sin(4y) \cdot \cos(15t)$. This describes a two-dimensional standing wave with temporal frequency 15.

(b) More generally, fix $\boldsymbol{\omega} = (\omega_1, \omega_2) \in \mathbb{R}^2$ and let $\Omega = \|\boldsymbol{\omega}\|_2 = \sqrt{\omega_1^2 + \omega_2^2}$. Then the function

$$u(\mathbf{x}; t) := \sin(\omega_1 x) \cdot \sin(\omega_2 y) \cdot \cos(\lambda \cdot \Omega t)$$

satisfies the two-dimensional wave equation and describes a standing wave with temporal frequency $\lambda \cdot \Omega$.

(c) Even more generally, fix $\boldsymbol{\omega} = (\omega_1, \omega_2) \in \mathbb{R}^2$ and let $\Omega = \|\boldsymbol{\omega}\|_2 = \sqrt{\omega_1^2 + \omega_2^2}$, as before. Let

$$SC_1(x) = \text{either } \sin(x) \text{ or } \cos(x);$$

let

$$SC_2(y) = \text{either } \sin(y) \text{ or } \cos(y);$$

and let

$$SC_t(t) = \text{either } \sin(t) \text{ or } \cos(t).$$

Then

$$u(\mathbf{x}; t) = SC_1(\omega_1 x) \cdot SC_2(\omega_2 y) \cdot SC_t(\lambda \cdot \Omega t)$$

satisfies the two-dimensional wave equation and describes a standing wave with temporal frequency $\lambda \cdot \Omega$. ◇

Ⓔ **Exercise 2B.4** Check Examples 2B.3(a)–(c). ♦