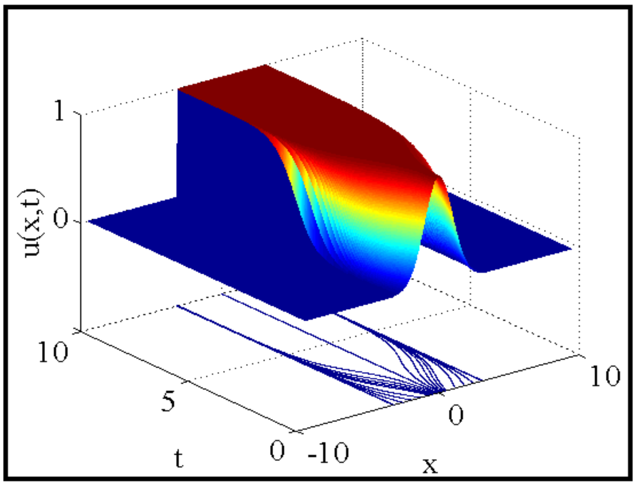
University College Dublin
An Coláiste Ollscoile, Baile Átha Cliath

School of Mathematical Sciences
Scoil na nEolaíochtaí Matamaitice
Partial Differential Equations (ACM30220)



Dr Lennon Ó Náraigh

Lecture notes in Partial Differential Equations, September 2012

Partial Differential Equations (ACM30220)

- Subject: Applied and Computational Maths
- School: Mathematical Sciences
- Module coordinator: Dr Lennon Ó Náraigh
- Credits: 5
- Level: 3
- Semester: First

In science, partial differential equations arise whenever spatial heterogeneity or evolutionary processes are important. Applications are found in fields as diverse as finance, engineering, physics, and the biosciences. This module is an elementary introduction to the theory of partial differential equations and emphasizes explicit solution techniques for linear and simple nonlinear equations. **Fourier series:** Finite-dimensional vector spaces, Scalar products, Fourier series, up to and including convergence and term-by term differentiation and integration; **One-dimensional diffusion equation:** Physical motivation, solution in terms of sine and cosine series, well-posedness of the solution, **Diffusion with forcing:** Diffusion with spatial and temporal forcing, breaking the solution into a homogeneous part and a particular integral, **One-dimensional wave equation:** Physical motivation, separation of variables solution, dAlemberts solution, causality, **Classification of linear first- and second-order PDEs:** Classification in terms of the elliptic/parabolic/hyperbolic scheme, classification in terms of the linear/nonlinear dichotomy, **Fourier transforms:** Definition and properties, the Dirac delta function, convolution, applications in solving boundary-value problems, **Greens functions:** Definition and motivating example for Poissons equation, general theory for infinite domains in terms of Fourier transforms, **One-dimensional linear advection equation:** Physical motivation, solution by the method of characteristics, **One-dimensional Burgers equation with no diffusion:** Method of characteristics, breaking waves, Riemann problems, **Burgers equation with diffusion:** Proof of regularity, the Cole-Hopf transformation **Further topics may include:** Greens functions for finite domains, further problems in non-linear first-order PDEs

What will I learn?

On completion of this module students should be able to

1. Compute Fourier series for a given periodic generating function;
2. Solve the diffusion equation for a standard set of boundary and initial conditions;
3. State and prove the properties of well-posedness for the diffusion problem, and for similar standard linear PDEs;
4. Recognize situations where decomposition into homogeneous and particular solutions is necessary, and solve such problems accordingly;
5. Classify PDEs as elliptic/parabolic/hyperbolic, and as linear/nonlinear;
6. Solve linear inhomogeneous problems using the Greens function technique;
7. Solve the one-dimensional wave equation using dAlemberts formula or separation of variables, depending on the context;
8. Understand the notion of causality as applied to the wave equation;
9. Solve Burgers equation using the method of characteristics; solve Riemann problems in the same way;
10. Obtain a bound on the gradient of the solution in the regularized Burgers problem, apply the Cole-Hopf transformation to the same.

Editions

First edition: September 2010

Second edition: September 2011

This edition: September 2012

Contents

Chapter 1

Introduction

1.1 Module summary

Here is the executive summary of the module:

After this module, you will be able to solve three standard **linear, constant-coefficient, second-order** partial differential equations (PDEs). You will also be able to solve certain **non-linear, first-order** PDEs too. These equations arise in numerous fields including engineering, finance, fluids, quantum mechanics, and electromagnetism.

In more detail, we will follow the following programme of work:

1. We review the theory of finite-dimensional vector spaces, with special emphasis on **orthonormal bases**. We then pass over to infinite-dimensional vector spaces, in the form of Fourier series;

2. We classify linear second-order PDEs;

3. We use Fourier series to solve these PDEs;

4. We study PDEs with **sources** intensively. We use the Green's function to solve these equations.

5. We will study a archetypal first-order nonlinear PDE called the **Burgers equation**. We will solve it using the **method of characteristics**. We will show that this method can fail, leading to **breaking waves**.

1.2 Learning and Assessment

Learning:

- Thirty six classes, three per week.
- In some classes, we will solve problems together or look at supplementary topics.
- To develop an ability to *solve problems autonomously*, you will be given homework exercises, and it is recommended that you do *independent study*.

Assessment:

- Three homework assignments, for a total of 20%;
- Three in-class tests, for a total of 20%;
- One end-of-semester exam, 60%

Textbooks

- Lecture notes will be put on the web. These are self-contained. They will be available *before* class. It is anticipated that you will print them and bring them with you to class. You can then annotate them and follow the proofs and calculations done on the board. Thus, you are still expected to attend class, and I will occasionally deviate from the content of the notes, give hints about solving the homework problems, or give a revision tips for the final exam.
- Here are some books for extra reading, if desired:
  - **For linear second-order PDEs:** *Mathematical methods for physicists*, G. B. Arfken, H. J. Weber, and F. Harris, Wiley, Fifth Edition (One copy of third edition in library, 510).
  - **For technical, analysis issues:** *Analysis: An Introduction*, R. Beals, Cambridge University Press, 2004 (Available for free on the internet).

Office hours

I do not keep specific office hours. If you have a question, you can visit me whenever you like – from 09:00-18:00 I am usually in my office if not lecturing. It is a bit hard to get to. The office number, building name, and location are indicated on a map at the back of this introductory chapter.

Otherwise, email me:

`lennon.onaraigh@ucd.ie`

1.3 Some pretty pictures from PDE theory

This section is given as means of motivation, and is not examinable.

The diffusion equation

Consider an ensemble of particles exercising Brownian motion. The particles are confined in an interval $[0,L]$, and a typical particle has the equation of motion

$$\frac{dx}{dt}=\xi(t),$$

where $\xi(t)$ is a random, normally-distributed variable with the correlation

$$\langle\xi(t)\xi(t')\rangle=2D\delta(t-t').$$

Technically, this equation is not correct, since we cannot differentiate a sequence of random numbers. However, an analytically sound solution does in fact exist, as the limit of the following difference equation:

$$x(t)=\lim_{\tau\rightarrow 0}x_\tau(N),\qquad t=N\tau,$$

where

$$x_\tau(n+1)=x_\tau(n)+\sqrt{2D\tau}w_n,$$

and where

$$w_n\sim_{\text{IID}}\mathcal{N}(0,1)\implies\langle w_n w_{n'}\rangle=\delta_{n,n'}.$$

A sample path is shown in Fig. 1.1.

Introducing the **probability distribution function**

$$C(x,t)=\langle\delta(x-x(t))\rangle,\qquad\text{average over all particles,}$$

we have

$$\begin{aligned}\frac{\partial C}{\partial t} &= -\left\langle\delta'(x-x(t))\frac{dx(t)}{dt}\right\rangle,\\ &= -\langle\delta'(x-x(t))\xi(t)\rangle,\\ &= -\frac{\partial}{\partial x}\langle\delta(x-x(t))\xi(t)\rangle.\end{aligned}$$

Using the Furutsu–Novikov theorem, it can be shown that

$$\langle\delta(x-x(t))\xi(t)\rangle=-D\frac{\partial}{\partial x}\langle\delta(x-x(t))\rangle,$$

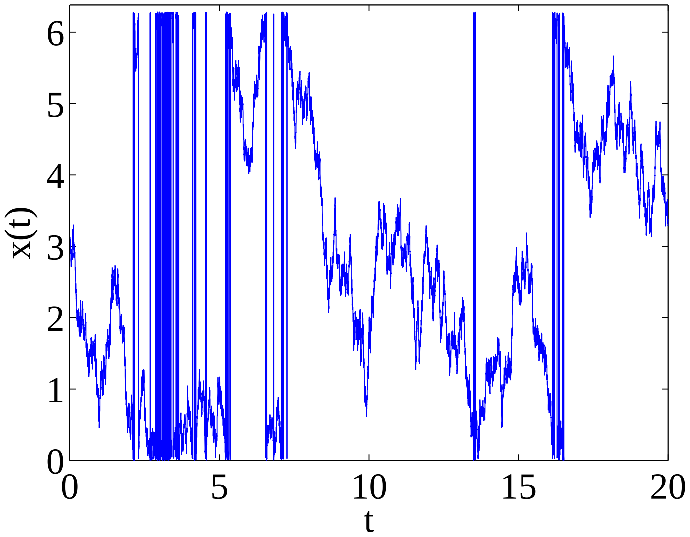


Figure 1.1: Sample path on the periodic domain, $D = 1$, $\tau = 0.001$.

hence

$$\frac{\partial C}{\partial t} = D\frac{\partial C}{\partial x^2}. \tag{1.1}$$

The function $C(x,t)$ therefore has the interpretation of the **concentration** of particles at position x and time t, and Eq. (1.1) is the **diffusion equation**. If we assume that the equation has **periodic boundary conditions**, so that a particle leaking out at $x = L$ re-enters the box at $x = 0$, then this equation has a very simple analytical solution, as we will find out:

$$C(x,t) = \sum_{n=-\infty}^{\infty} C_n \mathrm{e}^{-D(2\pi/L)^2 n^2 t} \mathrm{e}^{\mathrm{i}(2\pi/L)nx}, \tag{1.2}$$

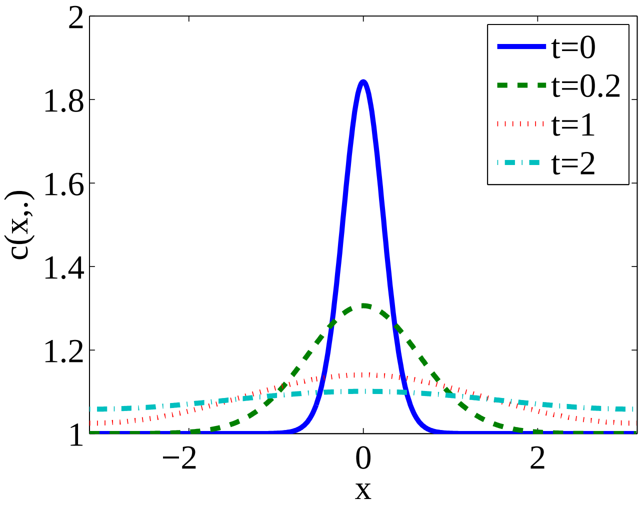
where the C_n's are chosen to match the prescribed **initial conditions** of the concentration:

$$\begin{aligned} C(x,0) &= f(x), \qquad \text{initial conditions}, \\ C(x,0) &= \sum_{n=-\infty}^{\infty} C_n \mathrm{e}^{\mathrm{i}(2\pi/L)nx}. \end{aligned}$$

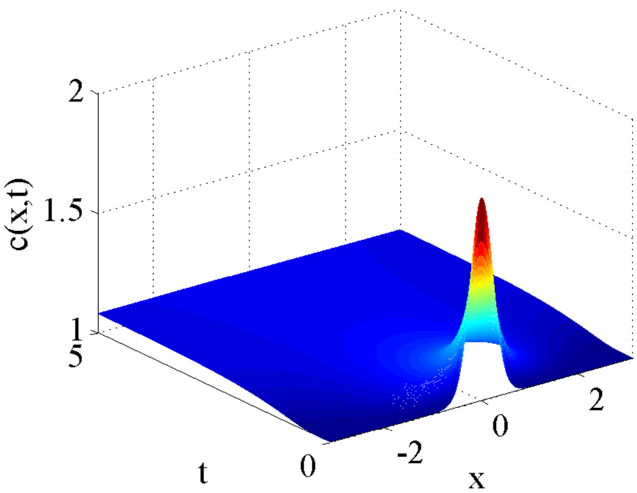
Let's look at a typical solution of this equation for the initial data

$$f(x) = \mathrm{e}^{-x^2/(2\sigma^2)} + 1.$$

I have implemented the numerical solution (1.2) in MATLAB, and the results are shown in Fig. 1.2 (a),



(a)



(b)

Figure 1.2: Solutions of the diffusion equation

for $D = 1$ and $\sigma = 0.2$. The initially 'peaked' distribution of concentrations, representing clustering near the origin $x = 0$ is flattened very quickly by the diffusion. After two time units, the concentration is nearly constant. Indeed, **diffusion forces any initial data to a constant, average value** – diffusion destroys concentration gradients and creates homogeneity. This is especially visible in the **space-time plot** (b).

Equation (1.2) is an example of a **spectral method**:

1. $FT[C](k,t) = \int_0^L C(x,t)\mathrm{e}^{-\mathrm{i}kx}$ is the **Fourier transform** of the concentration,

$$\widehat{C}_k(t) := FT[C](k,t) = \int_0^L C(x,t)\mathrm{e}^{-\mathrm{i}kx}\,\mathrm{d}x.$$

2. This is substituted back into the original diffusion equation to give

$$\frac{d\widehat{C}_k}{dt} = -Dk^2\widehat{C}_k$$

   Thus, the PDE is reduced to an ODE.

3. Solve the ODE (numerically if necessary):

4. Transform back into 'real space':

$$C(x,t) = IFT[\widehat{C}_k(t)] = \sum_{\substack{n=-\infty \\ k=(2\pi/L)n}}^{\infty} \mathrm{e}^{\mathrm{i}kx}\widehat{C}_k(t).$$

Numerically, this is a great recipe to implement because of the existence of **fast Fourier transforms**. I have used it many times to solve the celebrated **Cahn–Hilliard equation**:

$$\frac{\partial\phi}{\partial t} = D\nabla^2\left(\phi^3 - \phi - \gamma\nabla^2\phi\right).$$

This equation describes the separation of a binary alloy into its component parts:

$$\phi_{\text{separated}} = \begin{cases} +1, & \text{Alloy component 1,} \\ -1 & \text{Alloy component 2.} \end{cases}$$

I solve the CH equation numerically in Fig. 1.3. I have the following **initial conditions** and **boundary conditions**:

$$\begin{aligned} \boldsymbol{x} &\in [0, L = 2\pi]^2, \\ \phi(x+L, y, t) &= \phi(x, y, t), \\ \phi(x, y+L, t) &= \phi(x, y, t), \\ \phi(x, y, t=0) &= \text{random number between } -1 \text{and } +1. \end{aligned}$$

The numerical solution yields all kinds of insights into domain formation. For example, we can appropriately define the length scale of one of these domains (extended regions where $\psi = \pm 1$) – call it ℓ. We can demonstrate numerically that

$$\ell \sim t^{1/3}.$$

This is called the rate of domain **coarsening**. Understanding domain coarsening sheds light on many different subject areas:

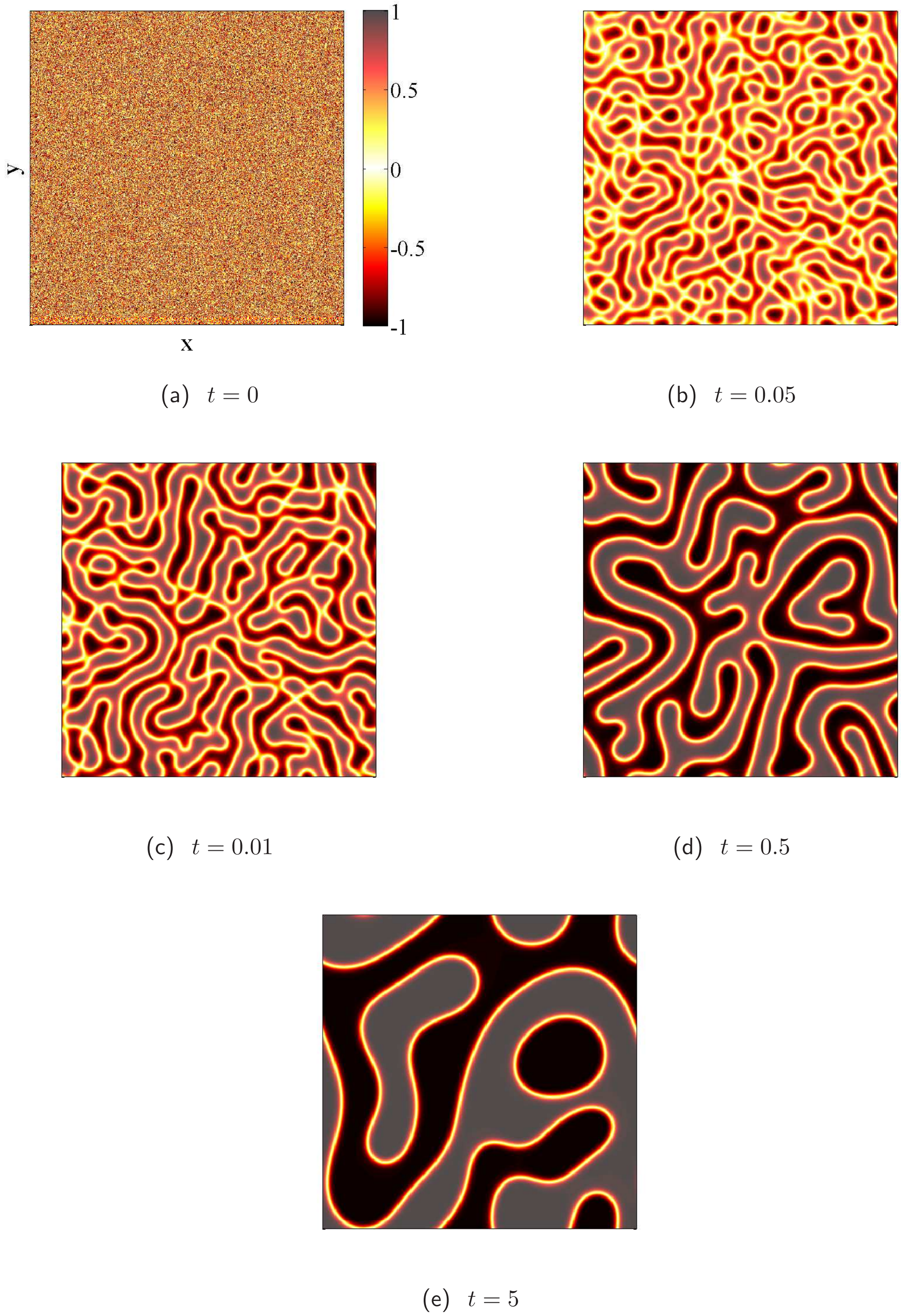
- droplet formation and droplet breakup,
- interfacial flows,
- polymer dynamics,
- emulsification and phase separation

This module will open up these areas, and others, for you to study.

Note on numerical methods

This module is a purely analytic one ('pen and paper'). However, from time to time, there will be pictures and solutions of PDEs. These will, more often than not, be generated using MATLAB. The codes used for these studies will be available on Blackboard and are cross-referenced in the notes. For example:

- The trajectory of a Brownian particle – brownian.m
- The one-dimensional diffusion equation – diffusion__one__d.m
- The two-dimensional Cahn–Hilliard equation – cahnhilliard__solve.m



(a) $t = 0$

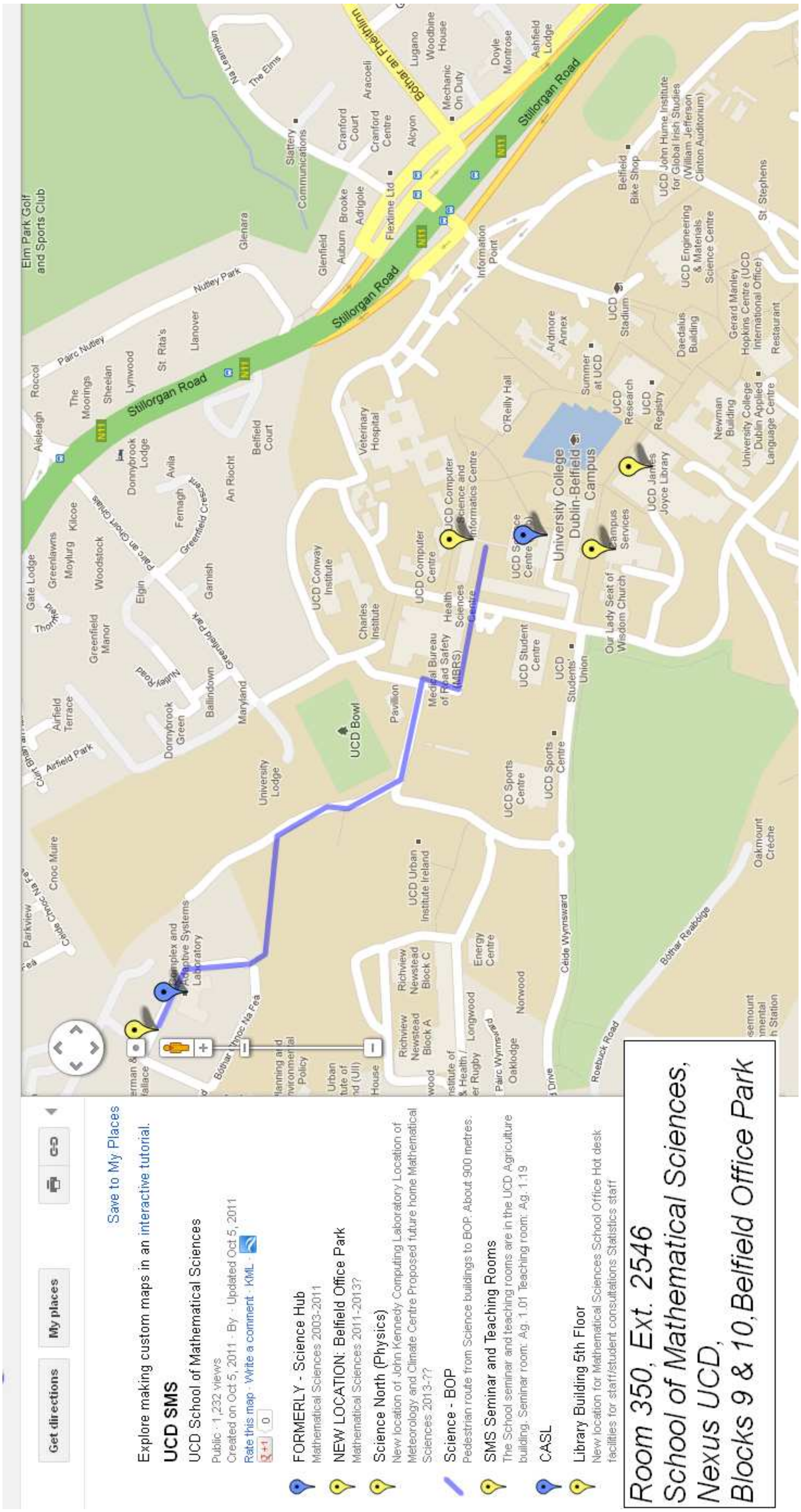
(b) $t = 0.05$

(c) $t = 0.01$

(d) $t = 0.5$

(e) $t = 5$

Figure 1.3: Solutions of the Cahn–Hilliard equation in two dimensions. Here $D = 1$ and $\gamma = 0.0024$.


Get directions
My places
Save to My Places
Explore making custom maps in an interactive tutorial.
UCD SMS
UCD School of Mathematical Sciences
Public · 1,232 views
Created on Oct 5, 2011 · By · Updated Oct 5, 2011
Rate this map · Write a comment · KML
FORMERLY - Science Hub
Mathematical Sciences 2003-2011
NEW LOCATION: Belfield Office Park
Mathematical Sciences 2011-2013?
Science North (Physics)
New location of John Kennedy Computing Laboratory Location of Meteorology and Climate Centre Proposed future home Mathematical Sciences 2013-??
Science - BOP
Pedestrian route from Science buildings to BOP. About 900 metres.
SMS Seminar and Teaching Rooms
The School seminar and teaching rooms are in the UCD Agriculture building. Seminar room: Ag 1.01 Teaching room: Ag 1.19
CASL
Library Building 5th Floor
New location for Mathematical Sciences School Office Hot desk facilities for staff/student consultations Statistics staff
Room 350, Ext. 2546
School of Mathematical Sciences,
Nexus UCD,
Blocks 9 & 10,Belfield Office Park
Stillorgan Road
University College Dublin-Belfield Campus
UCD Bowl
Elm Park Golf and Sports Club

Chapter 2

Finite-dimensional vector spaces

Overview

Recall that in previous modules, in particular "20150–Vector Integral and Differential Calculus" you studied the notion of a vector space. In that module, the vector spaces was, invariably, $\mathbb{R}^2$ or $\mathbb{R}^3$. In this chapter we review the theory of *linear vector spaces* and discuss a more general case, where the vector space is an abstract object.

2.1 Definitions and examples

Definition 2.1 *A set V is called a* **real vector space** *if the following properties hold:*

1. *An operation*

$$\begin{aligned} V \times V &\rightarrow V, \\ (\boldsymbol{x}, \boldsymbol{y}) &\rightarrow \boldsymbol{x} + \boldsymbol{y} \end{aligned}$$

   is given, called **addition of vectors**, *such that*

   (a) *The addition is associative: $(\boldsymbol{x} + \boldsymbol{y}) + \boldsymbol{z} = \boldsymbol{x} + (\boldsymbol{y} + \boldsymbol{z})$;*

   (b) *The addition is commutative: $\boldsymbol{x} + \boldsymbol{y} = \boldsymbol{y} + \boldsymbol{x}$;*

   (c) *There is an additive identity: $\boldsymbol{x} + 0 = \boldsymbol{x}$;*

   (d) *There are inverses: $\boldsymbol{x} + (-\boldsymbol{x}) = 0$.*

   These properties make the vector space into an **abelian group**.

2. *An operation*

$$\begin{aligned}\mathbb{R} \times V &\rightarrow V, \\ (\lambda, \boldsymbol{x}) &\rightarrow \lambda \boldsymbol{x}\end{aligned}$$

is given, called **scalar multiplication**, *which satisfies*

   (a) *The multiplication is distributive:* $\lambda(\boldsymbol{x}+\boldsymbol{y}) = \lambda\boldsymbol{x} + \lambda\boldsymbol{y}$;
   (b) *Distributivity:* $(\lambda+\mu)\boldsymbol{x} = \lambda\boldsymbol{x} + \mu\boldsymbol{x}$;
   (c) *Distributivity:* $(\lambda\mu)\boldsymbol{x} = \lambda(\mu\boldsymbol{x})$;
   (d) $1 \in \mathbb{R}$ *is a multiplicative identity* $1\boldsymbol{x} = \boldsymbol{x}$,

for all $\lambda, \mu \in \mathbb{R}$ *and* $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in V$. *The elements of* V *are called* **vectors** *and the elements of* $\mathbb{R}$ *are, in this context, called* **scalars**.

Examples:

1. The set of 3-dimensional geometrical vectors (as in 20150–Vector Calculus) is a real vector space.

2. The set
$$\mathbb{R}^n = \{(x_1, \cdots x_n) \,|x_1, \cdots x_n \in \mathbb{R}\}$$
is a real vector space.

3. The set V_Ω of all real-valued functions,
$$V_\Omega = \{f | f : (\Omega \subset \mathbb{R}) \rightarrow \mathbb{R}\}$$
is a vector space, with vector addition
$$(f+g)(x) = f(x) + g(x),$$
and scalar multiplication
$$(\lambda f)(x) = \lambda f(x),$$
for all $x \in \Omega$, all $f, g \in V_\Omega$, and $\lambda \in \mathbb{R}$. Note that the addition operation is called **pointwise** because it is defined with reference to each point $x \in \Omega$.

Definition 2.2 *Let* $W \subset V$ *and let* V *be a real vector space. Then* W *is called a* **vector subspace of** V *if it is non-empty, and if,*

1. *Closure under addition:* $\boldsymbol{x}, \boldsymbol{y} \in W \implies \boldsymbol{x} + \boldsymbol{y} \in W$*;*

2. *Closure under scalar multiplication:* $\lambda \in \mathbb{R}\, \boldsymbol{x} \in W \implies \lambda \boldsymbol{x} \in W$*.*

Thus, W *is itself a real vector space.*

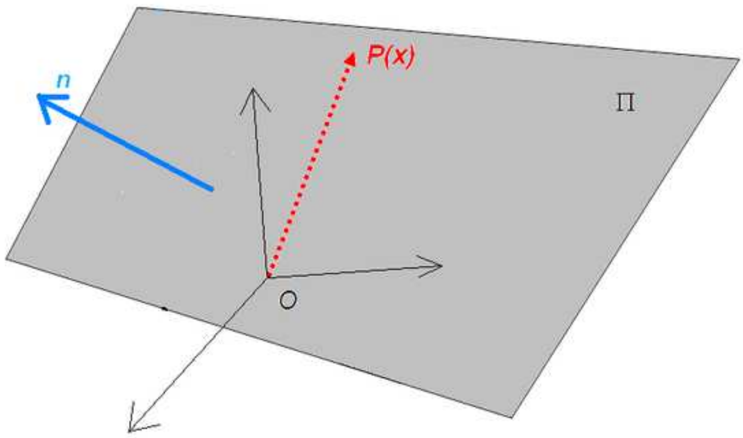


Figure 2.1: In ordinary three-dimensional space, a plane passing through the origin is a vector subspace.

Examples:

1. The set

$$\Pi(0,\boldsymbol{n}) = \{\boldsymbol{x} \in \mathbb{R}^3 | \boldsymbol{n} \cdot \boldsymbol{x} = 0\}$$

   is a vector subspace of $\mathbb{R}^3$ (Fig. 2.1). Proof: Let $\boldsymbol{x}$ and $\boldsymbol{y}$ be in the plane $\Pi(0,\boldsymbol{n})$:

$$\boldsymbol{n} \cdot \boldsymbol{x} = 0, \tag{2.1}$$

$$\boldsymbol{n} \cdot \boldsymbol{y} = 0. \tag{2.2}$$

   Adding gives $\boldsymbol{n} \cdot (\boldsymbol{x} + \boldsymbol{y}) = 0$, hence $\boldsymbol{x} + \boldsymbol{y} \in \Pi(0,\boldsymbol{n})$. Similarly for scalar multiplication.

2. Let Ω be an open subset of $\mathbb{R}$. As before, V_Ω is the set of all real-valued functions from Ω to $\mathbb{R}$. Then, $C^0(\Omega)$, the set of all **continuous** real-valued functions is a vector subspace of V_Ω.

3. C^∞, the space of all infinitely differentiable functions on Ω is a vector subspace of V_Ω.

4. The set of all solutions of the equation

$$\frac{d^2u}{dx^2} + u = 0$$

are a vector subspace of $C^\infty(\mathbb{R})$, and

$$(\text{Solution space}) \subset C^\infty(\mathbb{R}) \subset C^0(\mathbb{R}) \subset V_{\mathbb{R}}.$$

Definition 2.3 *Let $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_r$ be vectors in the real vector space V, and let $\lambda_1, \cdots \lambda_r$ be scalars. Then the vector*

$$\lambda_1 \boldsymbol{x}_1 + \cdots + \lambda_r \boldsymbol{x}_r$$

is called a **linear combination** *of $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_r$. We write*

$$\mathcal{S}(\boldsymbol{x}_1, \cdots \boldsymbol{x}_r) = \{\lambda_1 \boldsymbol{x}_1 + \cdots \lambda_r \boldsymbol{x}_r | \lambda_1, \cdots, \lambda_r \in \mathbb{R}\}$$

to denote the set of all linear combinations of $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_r$. $\mathcal{S}(\boldsymbol{x}_1, \cdots \boldsymbol{x}_r)$ is a vector subspace of V, and is called the subspace **spanned** *by $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_r$.*

If $\mathcal{S}(\boldsymbol{x}_1, \cdots \boldsymbol{x}_r) = V$, we say that $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_r$ **span** *the whole space V. Then, for each $\boldsymbol{x} \in M$, there exist scalars $\lambda_1, \cdots, \lambda_r$ such that*

$$\boldsymbol{x} = \lambda_1 \boldsymbol{x}_1 + \cdots \lambda_r \boldsymbol{x}_r.$$

Examples:

1. The vectors

$$\begin{aligned} \boldsymbol{e}_1 &= (1, 0, 0), \\ \boldsymbol{e}_2 &= (0, 1, 0), \\ \boldsymbol{e}_3 &= (0, 0, 1) \end{aligned}$$

span $\mathbb{R}^3$ because for any $\boldsymbol{x}$ in $\mathbb{R}^3$,

$$\begin{aligned} \boldsymbol{x} &= (x_1, x_2, x_3), \\ &= x_1 \boldsymbol{e}_1 + x_2 \boldsymbol{e}_2 + x_3 \boldsymbol{e}_3. \end{aligned}$$

2. The functions

$$\sin(x), \qquad \cos(x)$$

span the space of solutions of the equation

$$\frac{d^2 u}{dx^2} + u = 0,$$

because for any $u(x)$ in the solution space,

$$u(x) = \lambda_1 \sin(x) + \lambda_2 \cos(x).$$

Definition 2.4 *Let $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_r$ be vectors in a real vector space V. Then,*

1. *$\boldsymbol{x}_1, \cdots, \boldsymbol{x}_r$ are* **linearly dependent** *if there exist scalars $\lambda_1, \cdots, \lambda_r$* **not all zero** *such that*

$$\lambda_1 \boldsymbol{x}_1 + \cdots \lambda_r \boldsymbol{x}_r = 0.$$

2. *They are* **linearly independent** *if*

$$\lambda_1 \boldsymbol{x}_1 + \cdots \lambda_r \boldsymbol{x}_r = 0$$

implies that $\lambda_1 = \cdots = \lambda_r = 0$.

Example: $\sin(x)$ and $\cos(x)$ are linearly independent functions in $V_{\mathbb{R}}$. For, let us solve

$$\lambda \sin(x) + \mu \cos(x) = 0, \qquad \text{for all } x \in \mathbb{R}.$$

Since this expression must be true for all $x \in \mathbb{R}$, set $x = 0$. Since $\sin(0) = 0$ we have

$$\mu \cos(0) = \mu = 0.$$

Thus, we have

$$\lambda \sin(x) = 0, \qquad \text{for all } x \in \mathbb{R}.$$

Setting $x = \pi/2$ gives $\lambda = 0$ also, so $\sin(x)$ and $\cos(x)$ are linearly independent.

Note: If $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_r$ are linearly **dependent**, with

$$\lambda_1 \boldsymbol{x}_1 + \cdots \lambda_r \boldsymbol{x}_r = 0,$$

and $\lambda_1 \neq 0$ (say), then

$$\boldsymbol{x}_1 = -\frac{1}{\lambda_1} \left(\lambda_2 \boldsymbol{x}_2 + \cdots \lambda_r \boldsymbol{x}_r \right).$$

Thus, $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_r$ **are linearly dependent iff one of them is a linear combination of the others.**

Note: Let $\boldsymbol{x} \neq 0$ be a vector in V. Then the pair $\{0, \boldsymbol{x}\}$ are linearly dependent, because

$$1.0 + 0.\boldsymbol{x} = 0.$$

Thus, a list of linearly independent vectors can never contain the zero vector.

Definition 2.5 *A sequence of vectors $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n$ in a real vector space V is called a* **basis** *for V if,*

1. *$\boldsymbol{x}_1, \cdots, \boldsymbol{x}_r$ are linearly independent;*
2. *$\boldsymbol{x}_1, \cdots, \boldsymbol{x}_r$ span V.*

Thus, given a basis $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n$, we can write a vector $\boldsymbol{x} \in V$ as

$$\boldsymbol{x} = \alpha_1 \boldsymbol{x}_1 + \cdots \alpha_n \boldsymbol{x}_n, \qquad \alpha_1, \cdots \alpha_n \in \mathbb{R}.$$

The α_i's are called the **coordinates** of the vector.

Definition 2.6 *Let V be a real vector space spanned by a finite number of vectors. The minimal number of vectors required to span the vector space is called the* **dimsension** *of the space. The number of elements in a basis is equal to the dimension of the space.*

In this chapter, we are concerned with vector spaces that are spanned by a finite number of vectors. These are called **finite-dimensional vector spaces**.

Examples:

1. The vectors

$$\begin{aligned} \boldsymbol{e}_1 &= (1, 0, 0), \\ \boldsymbol{e}_2 &= (0, 1, 0), \\ \boldsymbol{e}_3 &= (0, 0, 1) \end{aligned}$$

   are a basis for $\mathbb{R}^3$. They certainly span $\mathbb{R}^3$:

$$\begin{aligned} \boldsymbol{x} &= (x_1, x_2, x_3), \\ &= x_1 \boldsymbol{e}_1 + x_2 \boldsymbol{e}_2 + x_3 \boldsymbol{e}_3. \end{aligned}$$

   They are also linearly independent:

$$\begin{aligned} a\boldsymbol{e}_1 + b\boldsymbol{e}_2 + c\boldsymbol{e}_3 &= 0, \\ (a, b, c) &= 0, \\ a = b = c &= 0. \end{aligned}$$

2. The functions $\sin(x)$ and $\cos(x)$ form a basis for the solution space of the differential equation

$$\frac{d^2u}{dx^2} + u = 0.$$

3. The $m \times n$ matrices

$$\begin{matrix} \uparrow \\ m \\ \downarrow \end{matrix} \overbrace{\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}}^{\leftarrow n \rightarrow}, \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix},$$

form a basis for $\mathbb{R}^{m \times n}$ as a real vector space.

Chapter 3

Scalar products

Overview

Just as the notion of vector space extends to spaces of functions, so too does the notion of scalar product. In this chapter we describe a scalar product for general, real vector spaces and extend the definition to spaces of (square-integrable) functions. In this way, we speak of 'the length of a function', or of two functions being orthogonal.

3.1 The definition

Definition 3.1 *Let V be a finite-dimensional real vector space. A scalar product on V is a map*

$$V \times V \to \mathbb{R},$$
$$(\boldsymbol{x}, \boldsymbol{y}) \to (\boldsymbol{x}|\boldsymbol{y}),$$

that is bilinear:

1. $(\lambda \boldsymbol{x} + \mu \boldsymbol{y}|\boldsymbol{z}) = \lambda(\boldsymbol{x}|\boldsymbol{z}) + \mu(\boldsymbol{y}|\boldsymbol{z})$,

2. $(\boldsymbol{x}|\lambda \boldsymbol{y} + \mu \boldsymbol{z}) = \lambda(\boldsymbol{x}|\boldsymbol{y}) + \mu(\boldsymbol{x}|\boldsymbol{z})$,

for all $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in V$ and $\lambda, \mu \in \mathbb{R}$.

3.2 The dot product on $\mathbb{R}^n$

Consider the usual basis on $\mathbb{R}^n$:

$$\begin{aligned}\boldsymbol{e}_1 &= (1, 0, \cdots, 0), \\ \boldsymbol{e}_2 &= (0, 1, \cdots, 0), \\ \vdots &= \vdots, \\ \boldsymbol{e}_n &= (0, 0, \cdots, 1).\end{aligned}$$

Define the **dot product** of two basis vectors:

$$\boldsymbol{e}_i \cdot \boldsymbol{e}_j = \delta_{ij}$$

where δ_{ij} is the Kronecker delta. Extend this definition by linearity two arbitrary vectors in $\mathbb{R}^n$:

$$\begin{aligned}\boldsymbol{a} &= a_1\boldsymbol{e}_1 + \cdots + a_n\boldsymbol{e}_n, \\ \boldsymbol{b} &= b_1\boldsymbol{e}_1 + \cdots + b_n\boldsymbol{e}_n, \\ \boldsymbol{a}\cdot\boldsymbol{b} &= (a_1\boldsymbol{e}_1 + \cdots + a_n\boldsymbol{e}_n)\cdot(b_1\boldsymbol{e}_1 + \cdots + b_n\boldsymbol{e}_n), \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} a_i\delta_{ij}b_j, \\ &= a_1b_1 + \cdots a_nb_n.\end{aligned}$$

Thus, for a vector $\boldsymbol{a} \in \mathbb{R}^n$, we define its length, $|\boldsymbol{a}|$:

$$\boldsymbol{a} = a_1\boldsymbol{e}_1 + \cdots + a_n\boldsymbol{e}_n,$$
$$|\boldsymbol{a}| := \sqrt{\boldsymbol{a}\cdot\boldsymbol{a}} = \sqrt{a_1^2 + \cdots + a_n^2}.$$

Theorem 3.1 *The scalar product satisfies the Cauchy–Schwartz inequality:*

$$|\boldsymbol{a}\cdot\boldsymbol{b}| \leq |\boldsymbol{a}||\boldsymbol{b}|.$$

Proof: Consider

$$\phi(\lambda) := (\lambda\boldsymbol{a} + \boldsymbol{b})^2 = (\lambda\boldsymbol{a} + \boldsymbol{b})\cdot(\lambda\boldsymbol{a} + \boldsymbol{b}).$$

We have $\phi(\lambda) \geq 0$ and

$$\phi(\lambda) = \lambda^2|\boldsymbol{a}|^2 + 2\lambda\boldsymbol{a}\cdot\boldsymbol{b} + |\boldsymbol{b}|^2.$$

This is a quadratic function in λ, with roots

$$\lambda_{\pm} = \frac{\boldsymbol{a}\cdot\boldsymbol{b} \pm \sqrt{(\boldsymbol{a}\cdot\boldsymbol{b})^2 - |\boldsymbol{a}|^2|\boldsymbol{b}|^2}}{|\boldsymbol{a}|^2}.$$

But $\phi(\lambda) \geq 0$, the quadratic function has at most one real root, so

$$(\boldsymbol{a}\cdot\boldsymbol{b})^2 - |\boldsymbol{a}|^2|\boldsymbol{b}|^2 \leq 0,$$

or

$$|\boldsymbol{a}\cdot\boldsymbol{b}| \leq |\boldsymbol{a}||\boldsymbol{b}|,$$

as required.

Note: The Cauchy–Schwartz inequality is true for any scalar product with the positive-definite property $(\boldsymbol{x}|\boldsymbol{x}) \geq 0$.

Definition 3.2 *Since* $|\boldsymbol{a}\cdot\boldsymbol{b}| \leq |\boldsymbol{a}||\boldsymbol{b}|$, *we define the* **angle between vectors** $\boldsymbol{a}$ *and* $\boldsymbol{b}$*:*

$$\cos\theta = \frac{\boldsymbol{a}\cdot\boldsymbol{b}}{|\boldsymbol{a}||\boldsymbol{b}|}, \qquad 0 \leq \theta \leq \pi.$$

Definition 3.3 *Two vectors are* **orthogonal** *if angle between them is* $\pi/2$ *(if their dot product is zero):*

$$\boldsymbol{a}\cdot\boldsymbol{b} = 0.$$

Changing basis: Consider now a basis $\boldsymbol{f}_1, \cdots \boldsymbol{f}_n$ for $\mathbb{R}^n$. Thus,

$$\boldsymbol{a} = \beta_1\boldsymbol{f}_1 + \cdots + \beta_n\boldsymbol{f}_n.$$

Suppose that we know the coordinates of $\boldsymbol{a}$ and $\boldsymbol{f}_1, \cdots \boldsymbol{f}_n$ w.r.t. the usual basis. How do we determine the scalars $\beta_1, \cdots, \beta_n$?

Solution: Let

$$\boldsymbol{a} = \alpha_1\boldsymbol{e}_1 + \cdots + \alpha_n\boldsymbol{e}_n.$$

Take the dot product of this expression with each new basis element $\boldsymbol{f}_i$:

$$\begin{aligned}
\boldsymbol{f}_1 \cdot \boldsymbol{a} &= \beta_1 \boldsymbol{f}_1 \cdot \boldsymbol{f}_1 + \cdots + \beta_n \boldsymbol{f}_1 \cdot \boldsymbol{f}_n, \\
\vdots &= \vdots \\
\boldsymbol{f}_i \cdot \boldsymbol{a} &= \beta_1 \boldsymbol{f}_i \cdot \boldsymbol{f}_1 + \cdots + \beta_n \boldsymbol{f}_i \cdot \boldsymbol{f}_n, \\
\vdots &= \vdots \\
\boldsymbol{f}_n \cdot \boldsymbol{a} &= \beta_1 \boldsymbol{f}_n \cdot \boldsymbol{f}_1 + \cdots + \beta_n \boldsymbol{f}_n \cdot \boldsymbol{f}_n.
\end{aligned}$$

In matrix form,

$$\begin{pmatrix} \boldsymbol{f}_1 \cdot \boldsymbol{a} \\ \vdots \\ \boldsymbol{f}_i \cdot \boldsymbol{a} \\ \vdots \\ \boldsymbol{f}_n \cdot \boldsymbol{a} \end{pmatrix} = \underbrace{\begin{pmatrix} \boldsymbol{f}_1 \cdot \boldsymbol{f}_1 & \cdots & \boldsymbol{f}_1 \cdot \boldsymbol{f}_n \\ \vdots & & \vdots \\ \boldsymbol{f}_i \cdot \boldsymbol{f}_1 & \cdots & \boldsymbol{f}_i \cdot \boldsymbol{f}_n \\ \vdots & & \vdots \\ \boldsymbol{f}_n \cdot \boldsymbol{f}_1 & \cdots & \boldsymbol{f}_n \cdot \boldsymbol{f}_n \end{pmatrix}}_{=G} \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_i \\ \vdots \\ \beta_n \end{pmatrix},$$

hence

$$\begin{pmatrix} \beta_1 \\ \vdots \\ \beta_i \\ \vdots \\ \beta_n \end{pmatrix} = G^{-1} \begin{pmatrix} \boldsymbol{f}_1 \cdot \boldsymbol{a} \\ \vdots \\ \boldsymbol{f}_i \cdot \boldsymbol{a} \\ \vdots \\ \boldsymbol{f}_n \cdot \boldsymbol{a} \end{pmatrix}, \qquad (*)$$

and everything on the RHS can be calculated in terms of the usual basis.

Examples:

1. Let

$$\begin{aligned}
\boldsymbol{f}_1 &= (1, 0), \\
\boldsymbol{f}_2 &= (1, 1).
\end{aligned}$$

   These vectors are linearly independent and form a basis for $\mathbb{R}^2$. Let

$$\begin{aligned}
\boldsymbol{a} &= \alpha_1 \boldsymbol{e}_1 + \alpha_2 \boldsymbol{e}_2, \\
&= \beta_1 \boldsymbol{f}_1 + \beta_2 \boldsymbol{f}_2.
\end{aligned}$$

   If α_1 and α_1 are known, what are β_1 and β_2?

Answer: We form the matrix

$$G = \begin{pmatrix} \boldsymbol{f}_1 \cdot \boldsymbol{f}_1 & \boldsymbol{f}_1 \cdot \boldsymbol{f}_2 \\ \boldsymbol{f}_2 \cdot \boldsymbol{f}_1 & \boldsymbol{f}_2 \cdot \boldsymbol{f}_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}.$$

Form also the column vector

$$\begin{pmatrix} \boldsymbol{f}_1 \cdot \boldsymbol{a} \\ \boldsymbol{f}_2 \cdot \boldsymbol{a} \end{pmatrix} = \begin{pmatrix} (1,0) \cdot (\alpha_1 \boldsymbol{e}_1 + \alpha_2 \boldsymbol{e}_2) \\ (1,1) \cdot (\alpha_1 \boldsymbol{e}_1 + \alpha_2 \boldsymbol{e}_2) \end{pmatrix} = \begin{pmatrix} (1,0) \cdot (\alpha_1, \alpha_2) \\ (1,1) \cdot (\alpha_1, \alpha_2) \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \alpha_1 + \alpha_2 \end{pmatrix}$$

From the equation (*), the β_i's are

$$\begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = G^{-1} \begin{pmatrix} \boldsymbol{f}_1 \cdot \boldsymbol{a} \\ \boldsymbol{f}_1 \cdot \boldsymbol{a} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}^{-1} \begin{pmatrix} \alpha_1, \\ \alpha_1 + \alpha_2 \end{pmatrix}$$

The final answer is

$$\begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = \begin{pmatrix} \alpha_1 - \alpha_2 \\ \alpha_2 \end{pmatrix}.$$

This example shows that basis elements do not have to be orthogonal – the usual basis $\{\boldsymbol{e}_1, \boldsymbol{e}_2\}$ is quite special.

2. Let

$$\boldsymbol{f}_1 = (\cos\theta, \sin\theta),$$
$$\boldsymbol{f}_2 = (-\sin\theta, \cos\theta).$$

These vectors are linearly independent and form a basis for $\mathbb{R}^2$. Let

$$\begin{aligned} \boldsymbol{a} &= \alpha_1 \boldsymbol{e}_1 + \alpha_2 \boldsymbol{e}_2, \\ &= \beta_1 \boldsymbol{f}_1 + \beta_2 \boldsymbol{f}_2. \end{aligned}$$

If α_1 and α_1 are known, what are β_1 and β_2?

Answer: As before, form the matrix G

$$G = \begin{pmatrix} \boldsymbol{f}_1 \cdot \boldsymbol{f}_1 & \boldsymbol{f}_1 \cdot \boldsymbol{f}_2 \\ \boldsymbol{f}_2 \cdot \boldsymbol{f}_1 & \boldsymbol{f}_2 \cdot \boldsymbol{f}_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Form also the column vector

$$\begin{pmatrix} \boldsymbol{f}_1 \cdot \boldsymbol{a} \\ \boldsymbol{f}_1 \cdot \boldsymbol{a} \end{pmatrix} = \begin{pmatrix} (\cos\theta, \sin\theta) \cdot (\alpha_1, \alpha_2), \\ (-\sin\theta, \cos\theta) \cdot (\alpha_1, \alpha_2) \end{pmatrix} = \begin{pmatrix} \alpha_1 \cos\theta + \alpha_2 \sin\theta, \\ -\alpha_1 \sin\theta + \alpha_2 \cos\theta \end{pmatrix}$$

Because G is the identity matrix, the β_i's are

$$\begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = \underbrace{\mathbb{I}}_{=G^{-1}} \begin{pmatrix} \boldsymbol{f}_1 \cdot \boldsymbol{a}, \\ \boldsymbol{f}_1 \cdot \boldsymbol{a} \end{pmatrix}$$

and the final answer is

$$\begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = \begin{pmatrix} \alpha_1 \cos\theta + \alpha_2 \sin\theta, \\ -\alpha_1 \sin\theta + \alpha_2 \cos\theta \end{pmatrix}$$

Definition 3.4 *A basis $\boldsymbol{f}_1, \cdots, \boldsymbol{f}_n$ for $\mathbb{R}^n$ is called* **orthonormal** *if*

$$\boldsymbol{f}_i \cdot \boldsymbol{f}_j = \delta_{ij}.$$

The components of an arbitrary vector $\boldsymbol{a}$ w.r.t. this basis are given by

$$\boldsymbol{a} = \beta_1 \boldsymbol{f}_1 + \cdots + \beta_n \boldsymbol{f}_n,$$

and

$$\beta_i = \boldsymbol{a} \cdot \boldsymbol{f}_i.$$

3.3 Spaces of functions

Recall, the set V_Ω of all real-valued functions,

$$V_\Omega = \{f | f : (\Omega \subset \mathbb{R}) \to \mathbb{R}\}$$

is a vector space, with pointwise operations of addition and scalar multiplication.

Definition 3.5 *The set*

$$L^2(\Omega) = \left\{ f \in V_\Omega \middle| \int_\Omega |f(x)|^2 \mathrm{d}x < \infty \right\}$$

is a vector subspace of V_Ω called the space of **square-integrable functions**.

Theorem 3.2 *The map*

$$(\cdot|\cdot) : L^2(\Omega) \times L^2(\Omega) \to \mathbb{R},$$
$$(f, g) \to \int_\Omega f(x)g(x)\mathrm{d}x$$

is a scalar product on the vector space $L^2(\Omega)$.

The proof is easy: all you do is show bi-linearity.

Definition 3.6 *Let $f \in L^2(\Omega)$. Then the* **length of the function** *f is denoted by $\|f\|_2$, and is defined by*

$$\|f\|_2^2 := (f|f) = \int_\Omega |f(x)|^2\mathrm{d}x.$$

Definition 3.7 *Let $f, g \in L^2(\Omega)$. These functions are* **orthogonal** *if*

$$(f|g) = \int_\Omega f(x)g(x)\mathrm{d}x = 0.$$

As an example of 'the length of a function', let $\Omega = [-\pi, \pi]$. The length of the function $\sin(x)$ is given by

$$\|\sin(x)\|_2^2 = \int_{-\pi}^{\pi} \sin^2(x)\mathrm{d}x = \pi.$$

The functions $\sin(x)$ and $\cos(x)$ are orthogonal because

$$\int_{-\pi}^{\pi} \sin(x)\cos(x) = 0.$$

We now focus on a further important example: We consider a vector space of functions on $\Omega = [-\pi, \pi]$ defined as follows:

$$F_n = \left\{ f(x) \in V_\Omega \middle| f(x) = a_0 + \sum_{i=1}^{n} [a_i \cos(ix) + b_i \sin(ix)] \right\},$$

where the a_i's and b_i's are ordinary real numbers. In other words,

$$F_n \subset V_\Omega, \qquad F_n = \mathcal{S}\Big(1, \cos(x), \cdots, \cos(nx), \sin(x), \cdots, \sin(nx)\Big)$$

Thus, a typical element in F_n is

$$f(x) = a_0 + \sum_{i=1}^{n} \left[a_i \cos(ix) + b_i \sin(ix)\right].$$

This is a square-integrable function, because

$$(f|f) = \pi \left[2a_0^2 + \sum_{i=1}^{n}(a_i^2 + b_i^2)\right].$$

We prove this statement now:

$$(f|f) = \int_{-\pi}^{\pi} \mathrm{d}x \left[a_0 + \sum_{i=1}^{n} \left[a_i \cos(ix) + b_i \sin(ix)\right]\right] \left[a_0 + \sum_{i=1}^{n} \left[a_i \cos(ix) + b_i \sin(ix)\right]\right],$$

$$\begin{aligned}(f|f) = 2\pi a_0^2 + 2a_0 \int_{-\pi}^{\pi} \mathrm{d}x \sum_{i=1}^{n} \left[a_i \cos(ix) + b_i \sin(ix)\right] \\ + \int_{-\pi}^{\pi} \mathrm{d}x \left[\sum_{i=1}^{n} \left[a_i \cos(ix) + b_i \sin(ix)\right]\right] \left[\sum_{i=1}^{n} \left[a_i \cos(ix) + b_i \sin(ix)\right]\right],\end{aligned}$$

$$(f|f) = 2\pi a_0^2 + 2a_0 \int_{-\pi}^{\pi} \mathrm{d}x \sum_{i=1}^{n} \left[a_i \cos(ix) + b_i \sin(ix)\right]$$

$$+\int_{-\pi}^{\pi} \mathrm{d}x \sum_{i=1}^{n} \sum_{j=1}^{n} \left[a_i a_j \cos(ix)\cos(jx) + a_i b_j \cos(ix)\sin(jx) + b_i a_j \cos(ix)\cos(jx) + b_i b_j \cos(ix)\cos(jx)\right],$$

Take the summation signs outside the integrals:

$$(f|f) = 2\pi a_0^2 + 2a_0 \sum_{i=1}^{n} \int_{-\pi}^{\pi} \mathrm{d}x \left[a_i \cos(ix) + b_i \sin(ix)\right]$$

$$+\sum_{i=1}^{n} \sum_{j=1}^{n} \int_{-\pi}^{\pi} \mathrm{d}x \left[a_i a_j \cos(ix)\cos(jx) + a_i b_j \cos(ix)\sin(jx) + b_i a_j \sin(ix)\cos(jx) + b_i b_j \cos(ix)\cos(jx)\right].$$

But

$$\int_{-\pi}^{\pi} \sin(px) = \int_{-\pi}^{\pi} \cos(px) = 0, \qquad p = \{1, 2, \cdots\}$$

hence

$$(f|f) = 2\pi a_0^2 + \sum_{i=1}^{n}\sum_{j=1}^{n}\int_{-\pi}^{\pi} \mathrm{d}x \left[a_i a_j \cos(ix)\cos(jx) + a_i b_j \cos(ix)\sin(jx) + b_i a_j \sin(ix)\cos(jx) + b_i b_j \cos(ix)\cos(jx)\right].$$

Let's tackle the other integrals:

$$\int_{-\pi}^{\pi} \sin(ix)\cos(jx) = \tfrac{1}{2}\int_{-\pi}^{\pi} \left[\sin((i+j)x) + \sin((i-j)x)\right] \mathrm{d}x.$$

If $i = j$, then this is

$$\tfrac{1}{2}\int_{-\pi}^{\pi} \sin(2ix)\mathrm{d}x = -\frac{1}{4i}\left[\cos(\pi i) - \cos(-\pi i)\right] = 0;$$

otherwise it is

$$-\tfrac{1}{2}\left[\frac{\cos((i+j)x)}{i+j} + \frac{\cos((i-j)x)}{i-j}\right]_{-\pi}^{\pi} = 0.$$

Thus, the cross terms vanish in the sum for $(f|f)$:

$$(f|f) = 2\pi a_0^2 + \sum_{i=1}^{n}\sum_{j=1}^{n}\int_{-\pi}^{\pi} \mathrm{d}x \left[a_i a_j \cos(ix)\cos(jx) + b_i b_j \cos(ix)\cos(jx)\right].$$

Let's tackle the first term:

$$\int_{-\pi}^{\pi} \cos(ix)\cos(jx) = \tfrac{1}{2}\int_{-\pi}^{\pi} \left[\cos((i-j)x) + \cos((i+j)x)\right] \mathrm{d}x.$$

If $i = j$ this is

$$\tfrac{1}{2}\int_{-\pi}^{\pi} \left[1 + \cos(2ix)\right] = \pi + \frac{1}{4i}\left[\sin(\pi i) - \sin(-\pi i)\right] = \pi + 0,$$

otherwise it is

$$\tfrac{1}{2}\left[\frac{\sin((i-j)x)}{i-j} + \frac{\sin((i-j)x)}{i-j}\right]_{-\pi}^{\pi} = 0.$$

In other words,

$$\int_{-\pi}^{\pi} \cos(ix)\cos(jx) = \pi\delta_{ij}.$$

The sine integral is identical. Thus,

$$(f|f) = 2\pi a_0^2 + \sum_{i=1}^{n}\sum_{j=1}^{n} \pi\delta_{ij}\left[a_i a_j + b_i b_j\right] = 2\pi a_0^2 + \pi\sum_{i=1}^{n}[a_i^2 + b_i^2].$$

Note: We have shown,

- $\int_{-\pi}^{\pi} \mathrm{d}x \sin(ix) = \int_{-\pi}^{\pi} \mathrm{d}x \cos(ix) = 0$,
- $\int_{-\pi}^{\pi} \mathrm{d}x \sin(ix) \cos(jx) = 0$,
- $\int_{-\pi}^{\pi} \mathrm{d}x \sin(ix) \sin(jx) = \int_{-\pi}^{\pi} \mathrm{d}x \cos(ix) \cos(jx) = \pi\delta_{ij}$.

Thus,

$$\begin{aligned} f_0(x) &= \tfrac{1}{\sqrt{2\pi}}, \\ f_1(x) &= \tfrac{1}{\sqrt{\pi}} \cos(x), \\ &\vdots \\ f_n(x) &= \tfrac{1}{\sqrt{\pi}} \cos(nx), \\ g_1(x) &= \tfrac{1}{\sqrt{\pi}} \sin(x), \\ &\vdots \\ g_n(x) &= \tfrac{1}{\sqrt{\pi}} \sin(nx) \end{aligned}$$

are linearly independent, span $F_n(\Omega)$ and form an **orthonormal basis** for the space. **By definition / result 3.4**, an arbitrary element $f(x)$ in $F_n(\Omega)$ has the representation

$$f(x) = \sum_{i=0}^{n} (f|f_i) f_i(x) + \sum_{i=1}^{n} (f|g_i) g_i(x).$$

$F_n(\Omega)$ is therefore a $2n + 1$-dimensional real vector space.

Chapter 4

Fourier series

Overview

In the last chapter, we considered a certain subset of square-integrable functions on the interval $[-\pi, \pi]$:

$$F_n(\Omega) = \mathcal{S}\Big(1, \cos(x), \cdots, \cos(nx), \sin(x), \cdots, \sin(nx)\Big).$$

In this chapter, we work out what happens as $n \to \infty$.

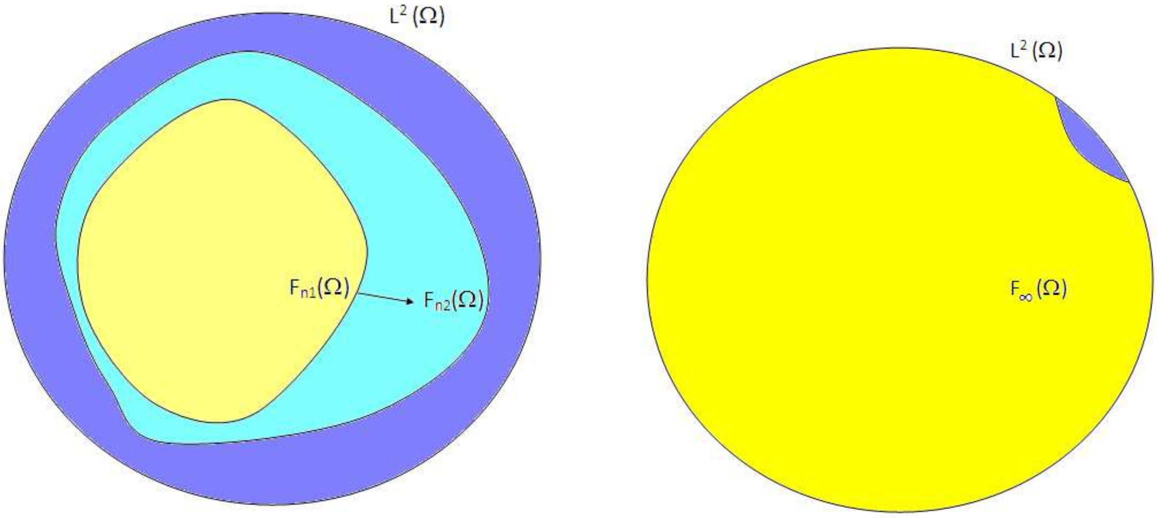
4.1 The limit as $n \to \infty$ of the set $F_n(\Omega)$

Recall the real vector-space

$$F_n(\Omega) = \mathcal{S}\Big(1, \cos(x), \cdots, \cos(nx), \sin(x), \cdots, \sin(nx)\Big).$$

for functions defined on the interval $\Omega = [-\pi, \pi]$. Recall, we formed the linearly independent set

$$\begin{aligned}
f_0(x) &= \tfrac{1}{\sqrt{2\pi}}, \\
f_1(x) &= \tfrac{1}{\sqrt{\pi}}\cos(x), \\
&\vdots \\
f_n(x) &= \tfrac{1}{\sqrt{\pi}}\cos(nx), \\
g_1(x) &= \tfrac{1}{\sqrt{\pi}}\sin(x), \\
&\vdots \\
g_n(x) &= \tfrac{1}{\sqrt{\pi}}\sin(nx)
\end{aligned}$$



(a) (b)

Figure 4.1: Convergence of Fourier series.

and showed that this was an orthonormal set,

$$(f_i|f_j) = \int_{\pi}^{\pi} f_i(x) f_j(x) \mathrm{d}x = \delta_{ij}, \qquad \text{\&c.}$$

Then, any function $f(x) \in F_n(\Omega)$ can be written as a linear combination of these orthonormal basis vectors

$$f(x) = \sum_{i=0}^{n} (f|f_i) f_i(x) + \sum_{i=1}^{n} (f|g_i) g_i(x).$$

We want to work out what happens as $n \to \infty$. The answer looks something like Fig. 4.1. As $2n+1$, the dimension of the space increases, more and more of $L^2(\Omega)$ is filled in. However, in the limit as $n \to \infty$, there is a small part of $L^2(\Omega)$ that is not filled in. This is the set of square-integrable functions that cannot be represented as an infinite Fourier sum. In this chapter we will describe this pathological set and then forget about it: $F_\infty(\Omega)$ is a big enough vector space to contain all of the PDE solutions we are interested in this course.

4.2 The Fourier series

Letting $n \to \infty$ in the definition of $F_n(\Omega)$, we obtain the set of all functions

$$F_\infty(\Omega) = \left\{ f(x) \in V_\Omega \middle| \begin{array}{c} f(x) = a_0 + \sum_{i=1}^{\infty} \left[a_i \cos(ix) + b_i \sin(ix) \right]; \\ \text{the Fourier series converges to its generating function } f \end{array} \right\}.$$

The coefficients a_0, a_i, and b_i can be obtained by taking the scalar product of $f(x)$ with the basis functions

$$\begin{aligned} f_0(x) &= \tfrac{1}{\sqrt{2\pi}}, \\ f_i(x) &= \tfrac{1}{\sqrt{\pi}}\cos(ix), \\ g_i(x) &= \tfrac{1}{\sqrt{\pi}}\sin(ix), \end{aligned}$$

where now $i \in \{1, 2, \cdots\}$ ranges over all positive integers:

$$\begin{aligned} a_i &= (f|f_i) = \int_{-\pi}^{\pi} f(x) f_i(x) \mathrm{d}x, \\ b_i &= (f|g_i) = \int_{-\pi}^{\pi} f(x) g_i(x) \mathrm{d}x. \end{aligned}$$

A series of the form $a_0 + \sum_{i=1}^{\infty} [a_i \cos(ix) + b_i \sin(ix)]$ is called a **Fourier series** and the coefficients a_i and b_i are called **Fourier coefficients**.
Provided $f(x)$ is square integrable (i.e. in $f \in L^2(\Omega)$), its Fourier coefficients can be calculated. It does not follow, however, that the corresponding Fourier series converges to $f(x)$. That is, the following diagram is not commutative (Fig. 4.2):

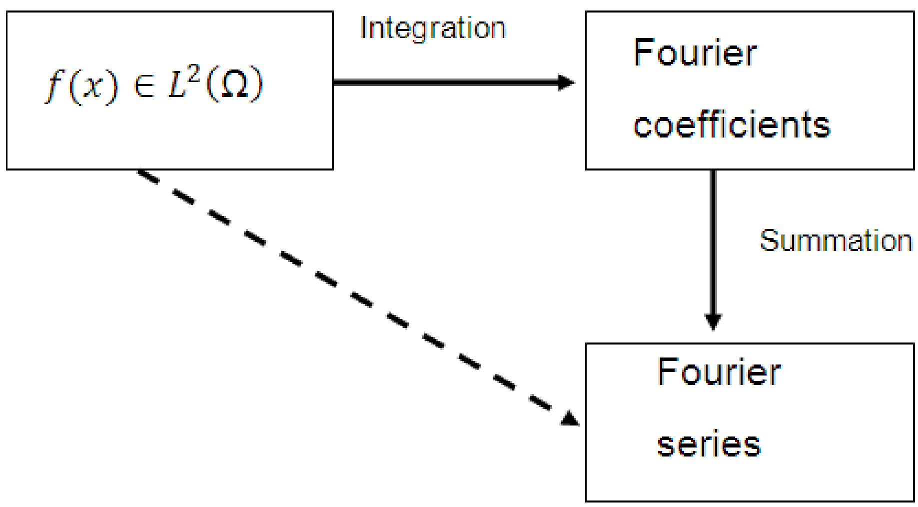


Fig. 4.2: The creation of a Fourier series from a generating function and the creation of functions from Fourier series are not **always** inverses: this is not a commutative diagram.

To ensure that this diagram is commutative, that is, to ensure that the Fourier series generated by a function's Fourier coefficients converges to the function itself, we need some stronger conditions than square-integrability.

Definition 4.1 *A function $f(x) \in L^2(\Omega)$, $\Omega = [a, b]$ is called piecewise smooth if there is a partition* *of* $[a, b]$,

$$a = x_0 < x_1 < x_2 < \cdots x_n = b,$$

such that f *has a continuous derivative (i.e.* C^1*) on each* **closed** *subinterval* $[x_m, x_{m+1}]$.

Example:

1. A function that is C^1 on $[a, b]$ is piecewise smooth on $[a, b]$.

2. The function

$$f(x) = \begin{cases} 2x, & 0 \leq x \leq \frac{1}{2}, \\ \frac{1}{2}, & \frac{1}{2} < x \leq 1 \end{cases}$$

   is piecewise smooth on $[0, 1]$ but is not continuous on $[0, 1]$.

3. The function

$$f(x) = |x|$$

   is both continuous and piecewise smooth on $[-1, 1]$, despite $f'(x)$ not being defined at $x = 0$. This is because we partition $[-1, 1]$ into two subintervals $[-1, 0]$ and $[0, 1]$. When worrying about $f'(x)$ near $x = 0$, note that on $[0, 1]$ we only care about the right limit,

$$f'(0^+) = \lim_{\epsilon \to 0, \epsilon > 0} \frac{f(0 + \epsilon) - f(0)}{\epsilon} = 1,$$

   while for $[-1, 0]$ we care only about the right limit,

$$f'(0^-) = \lim_{\epsilon \to 0, \epsilon > 0} \frac{f(0 - \epsilon) - f(0)}{\epsilon} = -1.$$

4. The function $f(x) = |x|^{1/2}$ is continuous on $[-1, 1]$ but is not piecewise smooth on $[-1, 1]$, since $f'(0^+)$ and $f'(0^-)$ do not exist.

Theorem 4.1 *Let* $f \in L^2((-\pi, \pi))$ *be piecewise smooth on the closed interval* $[-\pi, \pi]$ *and continuous on the open interval* $(-\pi, \pi)$*. Then, the Fourier series associated with* f *converges for all* $x \in [-\pi, \pi]$ *and converges to the generating function* $f(x)$ *for all* $x \in (-\pi, \pi)$.

No proof is given, but the proof (with copious hints) is examine in the homework (Assignment 1).

Now suppose instead that $f \in L^2((-\pi, \pi))$ is piecewise smooth on the closed interval $[-\pi, \pi]$ and piecewise continuous on $(-\pi, \pi)$, with a single discontinuity at $x = a$. Then the Fourier series converges to $f(x)$on $(-\pi, \pi)$, except at $x = a$, where it converges to

$$\frac{f(a-0)+f(a+0)}{2}.$$

4.3 A numerical example

Consider the function

$$f(x) = x^2$$

on the interval $[-\pi, \pi]$. This is a continuous function with continuous derivative. Thus, the function converges to its Fourier series on the interior of the interval, $(-\pi, \pi)$:

$$x^2 = a_0 f_0(x) + \sum_{i=1}^{\infty} \left[a_i f_i(x) + b_i g_i(x)\right],$$

where

$$\begin{aligned} a_0 &= \int_{-\pi}^{\pi} x^2 f_0(x), \qquad f_0(x) = \tfrac{1}{\sqrt{2\pi}}, \\ a_i &= \int_{-\pi}^{\pi} x^2 f_i(x), \qquad f_i(x) = \tfrac{1}{\sqrt{\pi}} \cos(ix), \\ b_i &= \int_{-\pi}^{\pi} x^2 g_i(x), \qquad g_i(x) = \tfrac{1}{\sqrt{\pi}} \sin(ix), \end{aligned}$$

Now

$$\begin{aligned} a_0 &= \tfrac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} x^2 \mathrm{d}x, \\ &= \tfrac{1}{\sqrt{2\pi}} \tfrac{2}{3}\pi^3, \\ a_i &= \tfrac{1}{\sqrt{\pi}} \int_{-\pi}^{\pi} x^2 \cos(ix) \mathrm{d}x, \\ &= \tfrac{1}{\sqrt{\pi}} \left[\frac{2x\cos(ix)}{i^2} + \frac{(-2+i^2x^2)\sin(ix)}{i^3}\right]_{-\pi}^{\pi}, \\ &= \tfrac{1}{\sqrt{\pi}} \frac{4\pi(-1)^i}{i^2}, \\ b_i &= \tfrac{1}{\sqrt{\pi}} \int_{-\pi}^{\pi} x^2 \sin(ix) \mathrm{d}x = 0. \end{aligned}$$

Hence,

$$\begin{aligned} x^2 &= \left(\tfrac{1}{\sqrt{2\pi}}\tfrac{2}{3}\pi^3\right)\tfrac{1}{\sqrt{2\pi}} + \sum_{i=1}^{\infty}\left(\tfrac{1}{\sqrt{\pi}}\frac{4\pi(-1)^i}{i^2}\right)\left(\tfrac{1}{\sqrt{\pi}}\cos(ix)\right), \\ &= \tfrac{\pi^2}{3} + 4\sum_{i=1}^{\infty}\tfrac{(-1)^i}{i^2}\cos(ix). \end{aligned}$$

Now, Theorem 4.1 guarantees that the Fourier series converges to its generating function on $(-\pi, \pi)$. The only reason why such a result cannot be extended **in general** to the closed interval $[-\pi, \pi]$ is because of possible discontinuities at the boundary points (here the comment below the theorem would apply). However, we can view $f(x) = x^2$ as a continuous function on all of $\mathbb{R}$, so long as we enforce the periodicity constraint (See Fig. 23.1). Thus, the Fourier series will converge to its generating function everywhere – including on the boundary points $x = \pm\pi$. We therefore set $x = \pi$

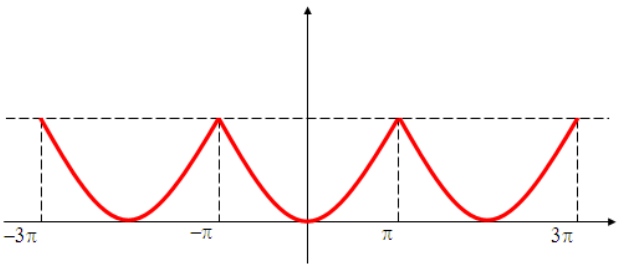


Figure 4.2: The function $f(x) = x^2$ viewed as a 2π-periodic function on all of $\mathbb{R}$.

in the Fourier series:

$$\begin{aligned} x^2 &= \tfrac{\pi^2}{3} + 4\sum_{i=1}^{\infty}\tfrac{(-1)^i}{i^2}\cos(ix), \\ \pi^2 &= \tfrac{\pi^2}{3} + 4\sum_{i=1}^{\infty}\tfrac{(-1)^i}{i^2}(-1)^i, \\ \tfrac{2}{3}\pi^2 &= 4\sum_{i=1}^{\infty}\tfrac{1}{i^2}, \\ \tfrac{1}{6}\pi^2 &= \sum_{i=1}^{\infty}\tfrac{1}{i^2}, \end{aligned}$$

a result first proved by Euler.

Note finally, a situation where convergence of the Fourier series to its generating function cannot be extended to the boundary points: $f(x) = x$ – See Fig. 4.3.

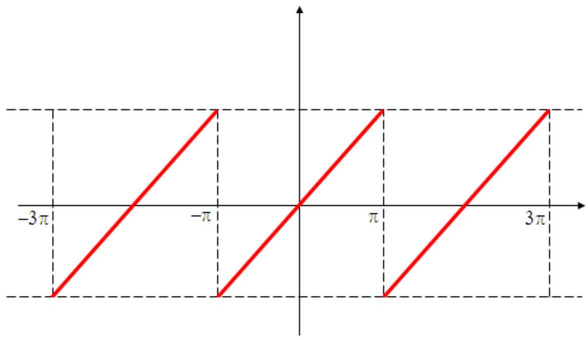


Figure 4.3: The function $f(x) = x$ viewed as a 2π-periodic function on all of $\mathbb{R}$: it possesses jump discontinuities at the boundary points.

Chapter 5

Differentiation and integration of Fourier series

Overview

It will be desirable to differentiate and integrate Fourier series in this course. The following definitions and results dictate when we can differentiate Fourier series term by term.

5.1 Uniform convergence

Definition 5.1 Uniform convergence: *The series*

$$\sum_{i=1}^{\infty} f_i(x)$$

of functions $f_i(x)$ defined on some interval $[a, b]$ **converges uniformly** *to a limit $f(x)$ if, for every $\epsilon > 0$, there exists a positive integer $N_0(\epsilon)$ depending on ϵ but* **independent of** *x such that*

$$\left| \left(\sum_{i=1}^{N} f_i(x) \right) - f(x) \right| < \epsilon \text{ whenever } N > N_0.$$

Theorem 5.1 Term-by-term differentiation: *If, on an interval $[a, b]$,*

1. *$f(x) = \sum_{i=1}^{\infty} f_i(x)$ converges uniformly,*
2. *$\sum_{i=1}^{\infty} f_i'(x)$ converges uniformly,*
3. *$f_i'(x)$ are continuous,*

then the series in (1) can be differentiated term-by-term:

$$f'(x) = \sum_{i=1}^{\infty} f_i'(x).$$

There is an easy way of checking if a series of functions converges uniformly:it is called the Weierstrass M-test:

Theorem 5.2 Weirstrass M-test: *Suppose $\{f_i(x)\}_{i=1}^{\infty}$ is a sequence of functions defined on an interval $[a, b]$, and suppose*

$$|f_i(x)| \le M_i,$$

where M_i is a real nonnegative number independent of x, and $i = 1, 2, \cdots$. Then the series of functions $\sum_{i=1}^{\infty} f_i(x)$ converges uniformly on $[a, b]$ if the series of numbers $\sum_{i=1}^{\infty} M_i$ converges.

In some cases, we will have a candidate for the M_i-series but want to test its convergence properties. Let's recall one useful test:

Theorem 5.3 The ratio test: *Let $\sum_{i=1}^{\infty} a_i$ be an infinite series. The series* **converges absolutely** *if the ratio*

$$r = \lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right|$$

is less than unity. The test is inconclusive if $r = 1$ and the series diverges if $r > 1$. Recall that absolute convergence means that the series $\sum_{i=1}^{\infty} |a_i|$ converges.

Proof: This result is easily proved by comparing the tail of the infinite series with a geometric series with geometric parameter r (Beals, p. 48).

Examples:

1. Consider the series

   $$\sum_{i=1}^{\infty} \cos(ix).$$

   There are certain points in the interval $[-\pi, \pi]$ where this series does not converge. For example, when $x = 0$, we have

   $$\sum_{i=1}^{\infty} \cos(ix) = \sum_{i=1}^{\infty} \cos(0) = 1 + 1 + 1 + \cdots,$$

   which is not a convergent series.

2. Consider, instead

   $$\sum_{i=1}^{\infty} \frac{\cos(ix)}{i^2}. \qquad (*)$$

The i^{th} term is

$$f_i(x) = \frac{\cos(ix)}{i^2}, \qquad |f_i(x)| \le \frac{1}{i^2},$$

and the series $\sum_{i=1}^{\infty} i^{-2}$ converges. Thus, by the M-test, the series (*) is uniformly convergent.

Proof of the Weirstrass M-test

Suppose $\{f_i(x)\}_{i=1}^{\infty}$ is a sequence of functions defined on an interval $[a, b]$, and suppose

$$|f_i(x)| \le M_i,$$

where M_i is a real non-negative number independent of x, and $i = 1, 2, \cdots$. Then the series of functions $\sum_{i=1}^{\infty} f_i(x)$ converges uniformly on $[a, b]$ if the series of numbers $\sum_{i=1}^{\infty} M_i$ converges.

Proof: We need to recall the following facts:

- The definition of a Cauchy sequence: A sequence $\{a_i\}_{i=1}^{\infty}$ of real numbers is Cauchy if, given $\epsilon > 0$, there exists a positive integer $N(\epsilon)$ such that

  $$|a_n - a_m| < \epsilon \text{ whenever } n,\ m > N(\epsilon).$$

  Recall, moreover, that a sequence of real numbers is convergent iff it is Cauchy (Beals, p. 36).

- A series $\sum_{i=1}^{\infty} a_i$ converges if the sequence formed by the N^{th} partial sum converges:

  $$\{S_i\}_{i=1}^{\infty}, \qquad S_N = \sum_{i=1}^{N} a_i, \qquad \text{a convergent sequence.}$$

Now consider $\sum_{i=1}^{\infty} f_i(x)$ We need to show first that the sequence $S_n(x) = \sum_{i=1}^{n} f_i(x)$ converges uniformly. It suffices to show that it is uniformly Cauchy. Assume, wlog, that $n > m$. Then,

$$\begin{aligned} |S_n(x) - S_m(x)| &= \left| \sum_{i=m+1}^{n} f_i(x) \right|, \\ &\le \sum_{i=m+1}^{n} |f_i(x)|, \\ &\le \sum_{i=m+1}^{n} M_i, \\ &= \left| \left(\sum_{i=1}^{n} M_i \right) - \left(\sum_{i=1}^{m} M_i \right) \right|. \end{aligned}$$

The series $\sum_{i=1}^{\infty} M_i$ converges, so the sequence of partial sums is Cauchy. Thus, given $\epsilon > 0$ there exists an $N_0(\epsilon)$ such that

$$\begin{aligned} |S_n(x) - S_m(x)| &\leq \left| \left(\sum_{i=1}^{n} M_i \right) - \left(\sum_{i=1}^{m} M_i \right) \right|, \\ &< \epsilon, \qquad \text{whenever } n, m > N_0(\epsilon). \end{aligned}$$

Since $N_0(\epsilon)$ is fixed entirely with reference to the M_i-series, it is independent of x, and the sequence of partial sums is uniformly Cauchy, and it is therefore uniformly convergent.

5.2 Integration of Fourier series

Theorem 5.4 Dominated convergence theorem: *Let $\{S_i(x)\}_{i=1}^{\infty}$ be a convergent sequence of square-integrable functions on a domain Ω, with limit $S(x)$. Then the limiting function $S(x)$ is square-integrable, and*

$$\lim_{n\to\infty} \int_\Omega \mathrm{d}x\, S_n(x) = \int_\Omega \mathrm{d}x\, S(x),$$

provided each term is dominated by a square-integrable function $g(x)$,

$$|S_i(x)| \leq g(x),$$

for all $i \in \{1, 2, \cdots\}$.

This proof is very technical and is not given here (Beals, p. 155).

Corollary: A series of functions $\sum_{i=1}^{\infty} f_i(x)$ can be integrated term-by-term if,

1. Each partial sum is square-integrable;

2. The partial sums are dominated by a a square-integrable function.

Example: Consider the function

$$f(x) = \sum_{i=1}^{\infty} \frac{\cos ix}{i^3}.$$

The partial sums are clearly square-integrable, because integration and summation commute when the summation is finite. The partial sums are also dominated by an integrable function:

$$
\begin{aligned}
|S_n(x)| &= \left|\sum_{i=1}^{n} \frac{\cos ix}{i^3}\right|, \\
&\leq \sum_{i=1}^{n} \left|\frac{\cos ix}{i^3}\right|, \\
&\leq \sum_{i=1}^{n} \frac{1}{i^3}, \\
&\leq \sum_{i=1}^{\infty} \frac{1}{i^3}, \\
&< \infty,
\end{aligned}
$$

and $\sum_{i=1}^{\infty} i^{-3}$ is, trivially, an integrable function. Thus, we can integrate term-by-term:

$$
\begin{aligned}
\int f(x)\,\mathrm{d}x &= \int \mathrm{d}x \left(\sum_{i=1}^{\infty} \frac{\cos ix}{i^3}\right), \\
&= \sum_{i=1}^{\infty} \frac{1}{i^3} \int \mathrm{d}x\ \cos(ix), \\
&= \sum_{i=1}^{\infty} \frac{1}{i^4} \sin(ix).
\end{aligned}
$$

Integrating the series term-by-term produces a series that converges faster than the original one!!

Note, moreover, that

$$
\begin{aligned}
|f_i(x)| &= \left|\frac{\cos(ix)}{i^3}\right|, \\
&\leq \frac{1}{i^3} := M_i, \\
|f_i'(x)| &= \left|\frac{\sin(ix)}{i^2}\right|, \\
&\leq \frac{1}{i^2} := N_i,
\end{aligned}
$$

and $\sum_{i=1}^{\infty}(M_i, N_i)$ are absolutely convergent. Thus, we can differentiate term-by-term:

$$\begin{aligned}\frac{df}{dx} &= \frac{d}{dx}\left(\sum_{i=1}^{\infty}\frac{\cos ix}{i^3}\right),\\ &= \sum_{i=1}^{\infty}\frac{d}{dx}\frac{\cos ix}{i^3},\\ &= -\sum_{i=1}^{\infty}\frac{\sin ix}{i^2},\end{aligned}$$

Differentiating the series term-by-term produces a series that converges slower than the original one. In very simple terms, then, 'differentiation bad, integration good'.

5.3 A numerical example

Consider the equation

$$\begin{aligned}u(x) &= \sum_{n=1}^{\infty} B_n \underbrace{\sin(n\pi x)e^{-n^2\pi^2 t}}_{=f_n(x)},\\ B_n &\leq B \text{ for all } n,\end{aligned}$$

where $t > 0$ is a positive parameter and $0 < x < 1$ is the domain of the function. Show that the series is uniformly convergent. We have,

$$|B_n \sin(n\pi x)e^{-n^2\pi^2 t}| \leq Be^{-n^2\pi^2 t}.$$

Consider the series $\sum_{n=1}^{\infty} M_n$, where $M_n = Be^{-n^2\pi^2 t}$. This is a convergent, x-independent series, because

$$\begin{aligned}\frac{M_{n+1}}{M_n} &= e^{-(n+1)^2\pi^2 t}e^{n^2\pi^2 t},\\ &= e^{-(2n+1)\pi^2 t},\\ &\leq e^{-\pi^2 t}, \qquad n \geq 1.\end{aligned}$$

Thus, the ratio of successive terms is always less than unity for $t > 0$, so the limit

$$\lim_{n\to\infty}(M_{n+1}/M_n) < 1,$$

and the series converges. By the Weirstrass M-test, the series in x is uniformly convergent. It can be similarly shown that the series

$$\sum_{n=1}^{\infty} B_n \left[\frac{d}{dx}\sin(n\pi x)\right] e^{-n^2\pi^2 t},$$
$$\sum_{n=1}^{\infty} B_n \left[\frac{d^2}{dx^2}\sin(n\pi x)\right] e^{-n^2\pi^2 t}.$$

are uniformly convergent and thus,

$$\begin{aligned}\frac{du}{dx} &= \sum_{n=1}^{\infty} B_n \left[\frac{d}{dx}\sin(n\pi x)\right] e^{-n^2\pi^2 t},\\ \frac{d^2u}{dx^2} &= \sum_{n=1}^{\infty} B_n \left[\frac{d^2}{dx^2}\sin(n\pi x)\right] e^{-n^2\pi^2 t}.\end{aligned}$$

5.4 Conclusions

In this chapter, we have done a lot of analysis, and formulated sufficient conditions to answer the following questions:

1. When does a Fourier series converge to the function that generates the Fourier coefficients?
2. When can we differentiate a series of functions term-by-term?
3. When can we integrate a series of functions term-by-term?
4. When does a series of functions converge uniformly?

However, in the rest of the course, we shall simply forget about the contents of this chapter, and do calculations. If we are unlucky enough that the calculations we do fail, we will revisit this chapter and see if we are violating any of the conditions set out here. This is the applied mathematics way.

Chapter 6

The 1-D diffusion equation: Solution

Overview

We formulate the theory of diffusion. Using Fourier-like series, we construct the solution of this equation.

6.1 Physical background

In this section solutions to the diffusion equation are presented. We saw the form of this equation in Ch. 1. Before we write it down again, let us remind ourselves of the physical ingredients in the equation. It describes the concentration $u(x,t)$ of particles undergoing Brownian motion, or the diffusion of heat in a metal rod (say). It is an evolutionary equation, so it contains a partial derivative w.r.t. time, $\partial u/\partial t$. Because the total number of particles

$$\int_\Omega u(x,t)\mathrm{d}x$$

is conserved (here $\Omega=[a,b]$ is the spatial domain of interest), it must be a **flux-conservative equation**:

$$\frac{\partial u}{\partial t}+\frac{\partial J}{\partial x}=0,$$

where J is the flux of particles (or the heat flux). Such an equation manifestly conserves the total particle number, for suitable boundary conditions:

$$\begin{aligned}\frac{d}{dt}\int_\Omega u(x,t)\mathrm{d}x &= \int_\Omega \mathrm{d}x\frac{\partial u}{\partial t},\\ &= -\int_\Omega \mathrm{d}x\frac{\partial J}{\partial x},\\ &= -\left[J(b,t)-J(a,t)\right].\end{aligned}$$

Thus, if $J(b,\cdot) - J(a,\cdot) = 0$, the total particle number is conserved. Now, we focus on the flux. If the flow of particles is proportional to minus the concentration gradient, particles will flow from regions of high concentration to regions of low concentration:

$$J \propto -\frac{\partial u}{\partial x}.$$

This is called **Fick's Law of diffusion** (call the constant of proportionality D). Thus,

$$J = -D\frac{\partial u}{\partial x}, \qquad \frac{\partial u}{\partial t} = -\frac{\partial J}{\partial x},$$

or

$$\frac{\partial u}{\partial t} = D\frac{\partial^2 u}{\partial x^2}. \tag{6.1}$$

This is the celebrated **diffusion equation**. Finally, let's look at the physical units of the proportionality factor D. It is,

$$\frac{[\text{Concentration}]}{[\text{Time}]} = [D]\,\frac{[\text{Concentration}]}{[\text{Length}]^2},$$

hence

$$D = \frac{[\text{Length}]^2}{[\text{Time}]}.$$

Boundary conditions

We will solve the diffusion equation on the interval $\Omega = [0, L]$. We pose the following boundary conditions:

$$\begin{aligned} \text{Initial condition:} \qquad & u(x, t = 0) = f(x), \qquad 0 < x < L, \\ \text{Boundary condition:} \qquad & u(x = 0, t) = u(x = L, t) = 0. \end{aligned}$$

The latter is certainly not the no-flux boundary condition: it is simpler, and makes the following exposition easier. Physically, this is a model of temperature in a metal rod, whose temperature is fixed at the end points.

Solving the diffusion equation with no-flux BCs will be left as an exercise.

6.2 Separation of variables

The diffusion equation is linear in the concentration $u(x,t)$, and the only coefficient in the equation is D, which is constant. When faced with a **linear, constant-coefficient** partial-differential equation, the following **separation of variables** procedure works: We make the following trial solution:

$$u(x,t) = X(x)T(t).$$

Substituting this ansatz into the PDE gives a set of ODEs – a much simpler problem:

$$X(x)\frac{dT}{dt} = DT(t)\frac{d^2X}{dx^2}.$$

Now divide out by $X(x)T(t)$. We obtain

$$\frac{1}{T}\frac{dT}{dt} = \frac{D}{X}\frac{d^2X}{dx^2} \quad (= -\lambda D).$$

But now the LHS is a function of t alone and the RHS is a function of x alone. The only way for this relation to be satisfied is if

$$\text{LHS} = \text{RHS} = \text{Const.} := -\lambda D.$$

Let us also substitute the trial solution into the BCs and the ICs:

$$\begin{aligned} \text{Initial condition:} \qquad & u(x,t=0) = X(x)T(0) = f(x), \qquad 0 < x < L, \\ \text{Boundary condition:} \qquad & T(t)X(0) = T(t)X(L) = 0 \end{aligned}$$

Solving for $X(x)$

Focussing on the $X(x)$-equations, we have:

$$\begin{aligned} \frac{1}{X}\frac{d^2X}{dx^2} &= -\lambda, \qquad 0 < x < L, \\ X(0) &= X(L) = 0. \end{aligned}$$

Equation in the **bulk** $0 < x < L$:

$$\frac{d^2X}{dx^2} + \lambda X = 0, \qquad (*)$$

Different possibilities for λ:

1. $\lambda = 0$. Then, the solution is $X(x) = Ax + B$, with $dX/dx = A$. However, the BCs specify $X(0) = 0$, hence $B = 0$. They also specify $X(L) = 0$, hence $A = 0$. Thus, only the trivial solution remains, in which we have no interest.

2. $\lambda < 0$. Then, the solution is $X(x) = Ae^{\mu x} + Be^{-\mu x}$, where $\mu = \sqrt{-\lambda}$. The BCs give

$$A + B = Ae^{\mu L} + Be^{-\mu L} = 0.$$

   Grouping the first two of these equations together gives

$$A = -B\frac{1 - e^{-\mu L}}{1 - e^{\mu L}}.$$

   But $A + B = 0$, hence

$$\begin{aligned} B\left[1 - \frac{1 - e^{-\mu L}}{1 - e^{\mu L}}\right] &= 0, \\ B\left[\frac{1 - e^{\mu L} - (1 - e^{-\mu L})}{1 - e^{\mu L}}\right] &= 0, \\ B\left[-e^{\mu L} + e^{-\mu L}\right] &= 0, \\ B\sinh(\mu L) &= 0, \end{aligned}$$

   which has only the trivial solution.

3. Thus, we are forced into the third option: $\lambda > 0$.

Solving Eq. (*) with $\lambda > 0$ gives

$$X(x) = A\cos(\sqrt{\lambda}x) + B\sin(\sqrt{\lambda}x),$$

with boundary condition

$$A \cdot 1 + B \cdot 0 = A\cos(\sqrt{\lambda}L) + B\sin(\sqrt{\lambda}L) = 0.$$

Hence, $A = 0$. Grouping the second and third equations in this string together therefore gives

$$B\sin(\sqrt{\lambda}L) = 0.$$

Of course, $B = 0$ is a solution, but this is the trivial one. Therefore, we must try to solve

$$\sin(\sqrt{\lambda}L) = 0.$$

This is possible, provided

$$\sqrt{\lambda}L = n\pi, \qquad n \in \{1, 2, \cdots\}.$$

Thus,

$$\lambda = \lambda_n = \frac{n^2\pi^2}{L^2},$$

and

$$X(x) = B_n \sin\left(\frac{n\pi x}{L}\right),$$

where B_n labels the constant of integration.

Solving for $T(t)$

Now substitute $\lambda_n = n^2\pi^2/L^2$ back into the $T(t)$-equation:

$$\frac{1}{T}\frac{dT}{dt} = -\lambda D = -\lambda_n D.$$

Solving give

$$T(t) = T(0)e^{-\lambda_n D t},$$

or

$$T(t) = T(0)e^{-n^2\pi^2 Dt/L^2}.$$

Putting it all together

Recall the ansatz:

$$u(x,t) = X(x)T(t).$$

Thus, we have a solution

$$X(x)T(t) = T(0)B_n \sin\left(\frac{n\pi x}{L}\right) e^{-n^2\pi^2 Dt/L^2}.$$

Calling $T(0)B_n := C_n$, this is

$$X_n(x)T_n(t) = C_n \sin\left(\frac{n\pi x}{L}\right) e^{-n^2\pi^2 Dt/L^2}.$$

The label n is just a label on the solution. However, each $n = 1, 2, \cdots$ produces a different solution, linearly independent of all the others. We can add all of these solutions together to obtain a **general**

solution of the PDE:

$$\begin{aligned} u(x,t) &= \sum_{n=1}^{\infty} X_n(x)T_n(t), \\ &= \sum_{n=1}^{\infty} C_n \sin\left(\frac{n\pi x}{L}\right) e^{-n^2\pi^2 Dt/L^2}. \end{aligned}$$

We are almost there. However, we still need to take care of the initial condition,

$$\begin{aligned} u(x,t=0) &= \sum_{n=1}^{\infty} C_n \sin\left(\frac{n\pi x}{L}\right), \\ &= f(x). \end{aligned}$$

But the functions

$$\left\{ \sin\left(\frac{n\pi x}{L}\right) \right\}_{n=1}^{\infty}$$

are orthogonal on $[0, L]$:

$$\begin{aligned} I_{n,m} &= \int_0^L \sin\left(\frac{n\pi x}{L}\right) \sin\left(\frac{m\pi x}{L}\right) \mathrm{d}x, \\ &= \frac{L}{\pi} \int_0^\pi \sin(ny)\sin(my)\mathrm{d}y, \qquad y = (\pi/L)x. \end{aligned}$$

If $n \neq m$ this is

$$\frac{L}{\pi}\left[\frac{\sin((m-n)x)}{2(m-n)} - \frac{\sin((m+n)x)}{2(m+n)}\right]_0^\pi = 0;$$

if $n = m$ it is

$$\frac{L}{\pi}\frac{\pi}{2} = \frac{L}{2},$$

hence $I_{nm} = (L/2)\delta_{nm}$. Thus, consider the IC again:

$$\begin{aligned} u(x,t=0) &= \sum_{n=1}^{\infty} C_n \sin\left(\frac{n\pi x}{L}\right), \\ &= f(x). \end{aligned}$$

Multiply both sides by $\sin(m\pi x/L)$ and integrate:

$$\begin{aligned}
\int_0^\pi f(x)\sin\left(\frac{m\pi x}{L}\right)\mathrm{d}x &= \int_0^\pi \sum_{n=1}^\infty C_n \sin\left(\frac{m\pi x}{L}\right)\sin\left(\frac{n\pi x}{L}\right), \\
&= \sum_{n=1}^\infty C_n \int_0^\pi \sin\left(\frac{m\pi x}{L}\right)\sin\left(\frac{n\pi x}{L}\right), \\
&= \sum_{n=1}^\infty C_n \frac{L}{2}\delta_{m,n}, \\
&= \frac{C_n L}{2}.
\end{aligned}$$

Hence,

$$C_n = \frac{2}{L}\int_0^\pi f(x)\sin\left(\frac{m\pi x}{L}\right)\mathrm{d}x,$$

and, substituting back into the general solution, we have

$$\begin{aligned}
u(x,t) &= \sum_{n=1}^\infty C_n \sin\left(\frac{n\pi x}{L}\right) e^{-n^2\pi^2 Dt/L^2}, \\
&= \sum_{n=1}^\infty \left[\frac{2}{L}\int_0^L f(s)\sin\left(\frac{m\pi s}{L}\right)\mathrm{d}s\right]\sin\left(\frac{n\pi x}{L}\right) e^{-n^2\pi^2 Dt/L^2}, \qquad (6.2)
\end{aligned}$$

which is a solution to the diffusion equation that satisfies the boundary and initial conditions.

Note finally, that as a result of Sec. 5.3, this series solution converges uniformly to the actual solution and can be differentiated term-by-term. Similarly, the series of derivatives in space and time (u_t and u_{xx}) converge uniformly to the derivatives of the solution.

Example: Cooling of a rod from a constant initial temperature

Suppose that we use the diffusion equation to model the temperature distribution in a metal rod. Suppose furthermore that the initial temperature distribution $f(x)$ in the rod is constant, i.e. $f(x) = u_0$. We wish to find the solution at later times. From Eq. (6.2), we have to work out

$$C_n = \frac{2}{L}\int_0^L f(x)\sin\left(\frac{m\pi x}{L}\right)\mathrm{d}x, \qquad f(x) = u_0.$$

That is,

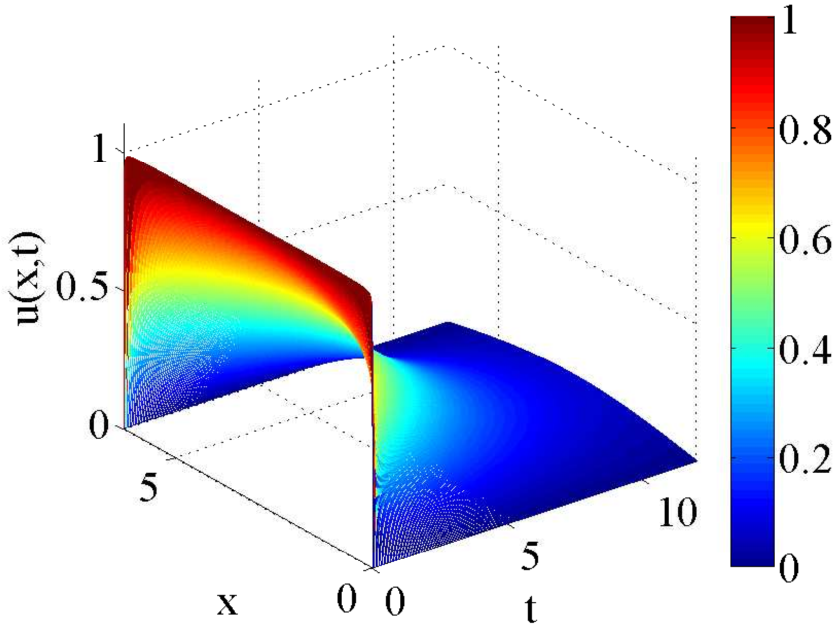


Figure 6.1: Solution for the diffusion equation with $u(x,t=0)=1$ and $u(0,t>0)=u(2\pi,t>0)=0$. I have set $D=u_0=1$.

$$\begin{aligned}
\frac{2u_0}{L}\int_0^L \sin\left(\frac{m\pi x}{L}\right)\mathrm{d}x &= \frac{2u_0}{\pi}\int_0^\pi \sin(my)\mathrm{d}y, \\
&= -\frac{2u_0}{\pi}\frac{\cos(my))}{m}\bigg|_0^\pi, \\
&= -\frac{2u_0}{\pi}\frac{(\cos(m\pi)-\cos(0))}{m}, \\
&= \frac{2u_0}{\pi}\frac{1-(-1)^m}{m}, \\
&= \begin{cases} 0, & m \text{ even}, \\ \frac{4u_0}{m\pi}, & m \text{ odd}. \end{cases}
\end{aligned}$$

Thus,

$$u(x,t) = \frac{4u_0}{\pi}\sum_{m=0}^{\infty}\frac{\sin\left(\frac{(2m+1)\pi x}{L}\right)}{2m+1}e^{-(2m+1)^2\pi^2 Dt/L^2}$$

I have plotted this solution in Fig. 6.1. The solution is rather odd:

- At time $t=0$ the solution is simply $u(x,t=0)=1$ everywhere in the domain.
- The solution then **instantaneously** adjusts so that $u(x,t>0)\approx 1$ in the bulk of the domain, and $u(x,t>0)=0$ on the boundary.

- Subsequently, the solution decays so that $u(x, t \to \infty) =$ boundary value $= 0$.

The code used to generate Fig. 6.1 is called diffusion__constant__temp.m

Chapter 7

The 1-D diffusion equation: Properties I

Overview

In Ch. 6 we solved the diffusion equation

$$\frac{\partial u}{\partial t} = D\frac{\partial^2 u}{\partial x^2}, \qquad \text{on } [0, L]$$

with a constant initial condition $u(x, t = 0) = u_0$ and null boundary conditions $u(0, t > 0) = u(L, t > 0) = 0$. We then plotted the solution and obtained some qualitative information about the solution. However, what if we did not have access to modern computer packages? How then can we extract qualitative information about the solution? The answer lies in finite-sum approximations.

Properties discussed: Finite truncations, visualisation of solutions, equilibrium state, temporal evolution towards equilibrium.

7.1 Finite-sum approximations

Consider the diffusion equation

$$\begin{aligned}
&\frac{\partial u}{\partial t} = D\frac{\partial^2 u}{\partial x^2}, \qquad \text{on } [0, L],\\
&u(x, t = 0) = f(x),\\
&u(x = 0, t) = u(x = L, t) = 0.
\end{aligned}$$

We have demonstrated that its solution is

$$
\begin{aligned}
u(x,t) &= \sum_{n=1}^{\infty} C_n \sin\left(\frac{n\pi x}{L}\right) e^{-n^2\pi^2 Dt/L^2}, \\
&= \sum_{n=1}^{\infty} \left[\frac{2}{L}\int_0^L f(x)\sin\left(\frac{m\pi s}{L}\right)\mathrm{d}s\right] \sin\left(\frac{n\pi x}{L}\right) e^{-n^2\pi^2 Dt/L^2},
\end{aligned} \tag{7.1}
$$

Although this solution is exact, it is not computable. In practical applications, we need to construct an approximate solution by taking finite-sum approximations of the solution (7.1). It is desirable to know what is the error incurred by doing this.

Suppose we work with a finite approximation

$$
u_N(x,t) = \sum_{n=1}^{N} C_n \sin\left(\frac{n\pi x}{L}\right) e^{-n^2\pi^2 Dt/L^2}.
$$

Then the error associated with this approximation is

$$
\begin{aligned}
\epsilon &= \left|u(x,t) - u_N(x,t)\right|, \\
&= \left|\sum_{n=N+1}^{\infty} C_n \sin\left(\frac{n\pi x}{L}\right) e^{-n^2\pi^2 Dt/L^2}\right|, \\
&\le \sum_{n=N+1}^{\infty} |C_n| e^{-n^2\pi^2 Dt/L^2}.
\end{aligned}
$$

Now

$$
\begin{aligned}
|C_n| &= \left|\frac{2}{L}\int_0^{\pi} \sin\left(\frac{n\pi x}{L}\right) f(x)\mathrm{d}x\right|, \\
&\le \frac{2}{L}\int_0^{\pi} \left|\sin\left(\frac{n\pi x}{L}\right) f(x)\right| \mathrm{d}x, \\
&\le \frac{2}{L}\int_0^{\pi} |f(x)|\,\mathrm{d}x, \\
&= C < \infty.
\end{aligned}
$$

Thus, the error is bounded by

$$
\begin{aligned}
\epsilon &\le C \sum_{n=N+1}^{\infty} e^{-n^2\pi^2 Dt/L^2}, \\
&= C e^{-N^2\pi^2 Dt/L^2} \sum_{n=N+1}^{\infty} e^{-(n^2-N^2)\pi^2 Dt/L^2}.
\end{aligned}
$$

Now

$$\begin{aligned} n^2 - N^2 &= (n-N)(n+N) \geq 2N(n-N) \geq 0, \\ -(n^2 - N^2) &\leq -2N(n-N), \\ e^{-(n^2-N^2)} &\leq e^{-2N(n-N)}, \end{aligned}$$

hence

$$\begin{aligned} \epsilon &\leq Ce^{-N^2\pi^2Dt/L^2} \sum_{n=N+1}^{\infty} e^{-2N(n-N)\pi^2Dt/L^2}, \\ &\leq C \sum_{n=N+1}^{\infty} e^{-[N^2+2N(n-N)]\pi^2Dt/L^2}, \\ &\leq C \sum_{n=N+1}^{\infty} e^{-[2Nn-N^2]\pi^2Dt/L^2}, \end{aligned}$$

$$\begin{aligned} \epsilon &\leq C \sum_{n=N+1}^{\infty} e^{+[N^2-2Nn]\pi^2Dt/L^2}, \\ &\leq Ce^{N^2\pi^2Dt/L^2} \sum_{n=N+1}^{\infty} e^{-2Nn\pi^2Dt/L^2}, \\ &\leq Ce^{N^2\pi^2Dt/L^2} \sum_{n=N+1}^{\infty} \left(e^{-2N\pi^2Dt/L^2}\right)^n, \\ &= Ce^{N^2\pi^2Dt/L^2} \frac{\left(e^{-2N\pi^2Dt/L^2}\right)^{N+1}}{1-e^{-2N\pi^2Dt/L^2}}, \\ &= Ce^{N^2\pi^2Dt/L^2} \frac{e^{-2N(N+1)\pi^2Dt/L^2}}{1-e^{-2N\pi^2Dt/L^2}}, \\ &= C\frac{e^{[-2N(N+1)+N^2]\pi^2Dt/L^2}}{1-e^{-2N\pi^2Dt/L^2}}, \\ &= Ce^{-2N\pi^2Dt/L^2}\left(\frac{e^{-N^2\pi^2Dt/L^2}}{1-e^{-2N\pi^2Dt/L^2}}\right). \end{aligned}$$

Let us work with the non-dimensional time $\tau = tD/L^2$. Thus,

$$\epsilon \leq Ce^{-2N\pi^2\tau}\left(\frac{e^{-N^2\pi^2\tau}}{1-e^{-2N\pi^2\tau}}\right).$$

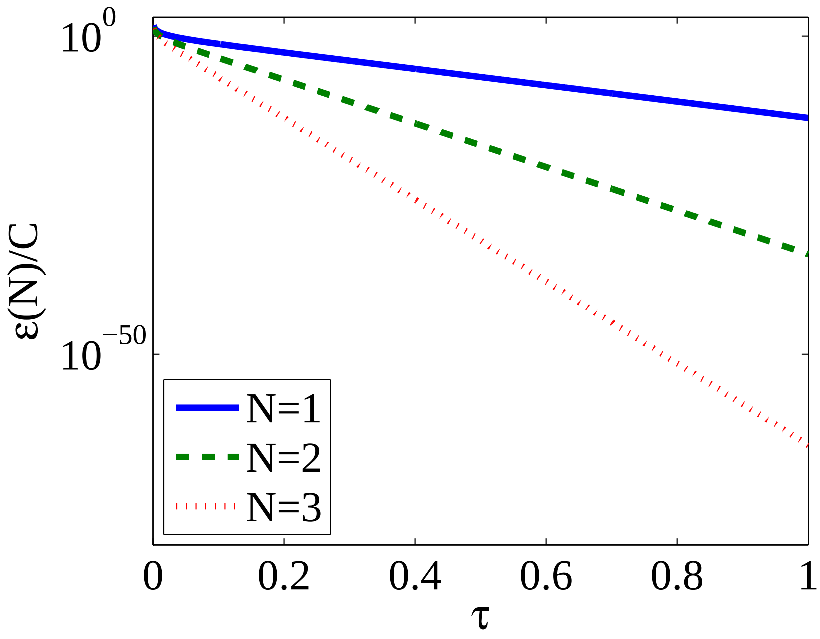


Figure 7.1: Error in the finite truncation of the diffusion equation, as a function of the truncation degree N, and as a function of elapsed time τ.

The error depends on time, and decreases exponentially fast as time progresses. The results are shown in Fig. 7.1 for various N-values: the larger N-values give more rapid convergence. For $N = 1$, we have

$$\epsilon(N=1) \leq Ce^{-2\pi^2\tau}\left(\frac{e^{-\pi^2\tau}}{1-e^{-2\pi^2\tau}}\right).$$

For $N = 1$ and $\tau = 1/\pi^2$ this is

$$\frac{\epsilon(N=1, \tau=1/\pi^2)}{C} = e^{-2}\left(\frac{e^{-1}}{1-e^{-2}}\right) = 0.0576,$$

a very small number. Thus,

> The error incurred by truncating the series after only the first term never exceeds
>
> $$\frac{2\times 0.0576}{L}\int_0^L |f(x)|\,\mathrm{d}x,$$
>
> when $\tau > 1/\pi^2$.

We use this result now to describe the solution for $u(x, t=0) = f(x) = u_0$ qualitatively.

7.2 Visualisation of the solution

Recall the solution to

$$\begin{aligned} &\frac{\partial u}{\partial t} = D\frac{\partial^2 u}{\partial x^2}, \qquad \text{on } [0,L], \\ &u(x,t=0) = u_0, \\ &u(x=0,t) = u(x=L,t) = 0; \end{aligned}$$

it was

$$u(x,t) = \frac{4u_0}{\pi}\sum_{m=0}^{\infty}\frac{\sin\left(\frac{(2m+1)\pi x}{L}\right)}{2m+1}e^{-(2m+1)^2\pi^2 Dt/L^2}$$

From the theory of error bounds just derived, the error in replacing the full series by

$$u(x,t) \approx \frac{4u_0}{\pi}e^{-\pi^2 Dt/L^2}\sin\left(\frac{\pi x}{L}\right) \tag{7.2}$$

never exceeds

$$\frac{2\times 0.0576}{L}\int_0^L |f(x)|\,\mathrm{d}x = 0.1152u_0$$

when $Dt/L^2 > 1/\pi^2$. Thus, let us work with this truncated solution and describe the solution qualitatively.

Spatial dependence

Fix $t = t_0 > L^2/(D\pi^2)$. Then the approximate solution (7.2) is

$$u(x,t_0) \approx \frac{4u_0}{\pi}e^{-\pi^2 Dt_0/L^2}\sin\left(\frac{\pi x}{L}\right).$$

This function has a maximum at $x = L/2$, since

$$\frac{d}{dx}\sin\left(\frac{\pi x}{L}\right)\bigg|_{x=L/2} = \frac{\pi}{L}\cos\left(\frac{\pi}{2}\right) = 0,$$

and $[d^2/dx^2\sin(\pi x/L)]_{x=L/2} < 0$. Indeed, the point $x = L/2$ is a point of symmetry, in the sense that

$$\sin\left[\frac{\pi}{L}\left(\frac{L}{2}-s\right)\right] = \sin\left[\frac{\pi}{L}\left(\frac{L}{2}+s\right)\right],$$

for $s \in [0, L/2]$. Moreover, these two properties are true of the full series solution

$$u(x,t) = \frac{4u_0}{\pi}\sum_{m=0}^{\infty}\frac{\sin\left(\frac{(2m+1)\pi x}{L}\right)}{2m+1}e^{-(2m+1)^2\pi^2 Dt/L^2};$$

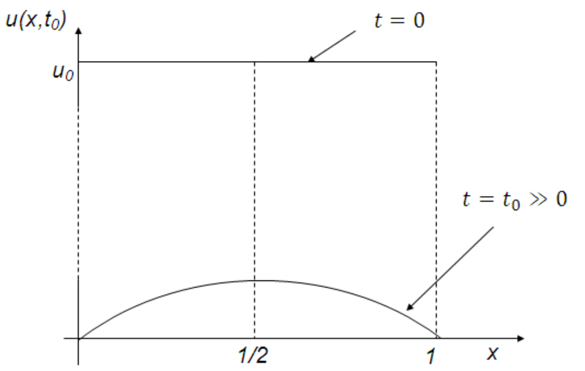


Figure 7.2: Snapshot of the solution at late time $t \gg 0$, compared to the initial non-smooth data.

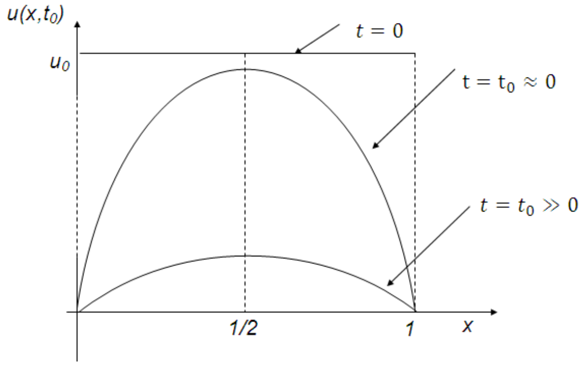


Figure 7.3: A further set of snapshots of the solution, showing intermediate and late configurations of the solution, and showing the evolution in time away from the non-smooth initial condition.

this can be verified by direct computation.

Thus, we have the following picture (Fig. 7.2).

- At $t = t_0 = 0$ the spatial profile is simply $u = u_0$.
- For $t = L^2/(D\pi^2)$ the spatial profile is a sine function, with the null BCS, and with a line of symmetry at $x = L/2$.
- We guess that at intermediate times, the spatial profile is,
  - A smooth curve sandwiched between the early- and late-time profiles;
  - One that respects the BCs $u(x = 0, \cdot) = u(x = L, \cdot) = 0$;
  - One with a maximum at a line of symmetry at $x = L/2$

Thus, we fill in the gap in the figure (Fig. 7.3).

Temporal dependence

Fix x_0. Then the approximate solution (7.2) is

$$u(x_0, t) \approx \frac{4u_0}{\pi} e^{-\pi^2 Dt/L^2} \sin\left(\frac{\pi x}{L}\right), \qquad t > L^2/(D\pi^2).$$

- The initial concentration is u_0, so

$$(4/\pi)e^{-\pi^2 Dt/L^2} \sin(\pi x_0/L) := (4/\pi)e^{-\pi^2 \tau} \sin(\pi x_0/L).$$

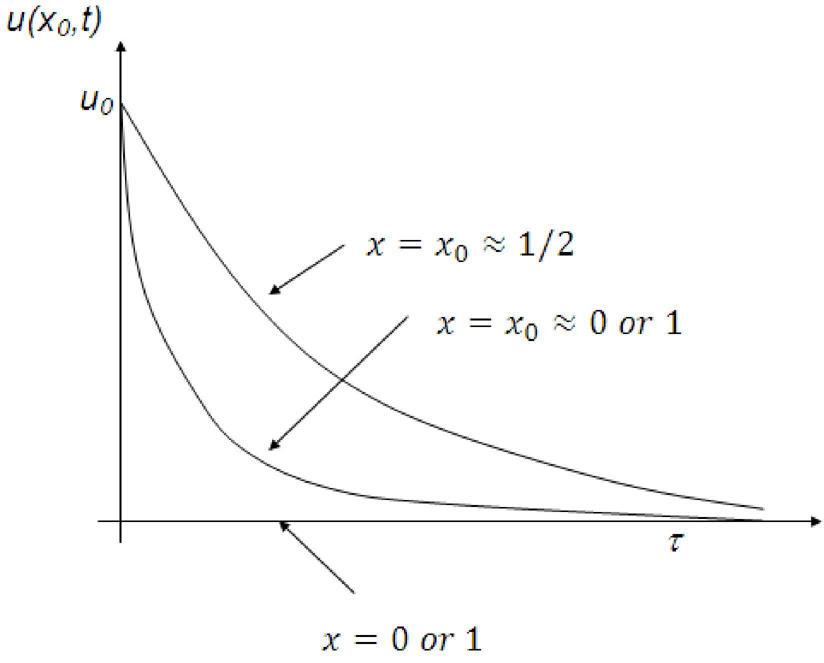


Figure 7.4: Visualization of the solution at a fixed location in space.

is a like a 'decay factor', damping the initial concentration level to zero as time goes by.

- However, the damping is faster at certain x_0-values.
  - The decay to zero faster near $x_0 = 0$ or 1 because the decay factor is smaller there.
  - At $x_0 = 0$ or $x_0 = 1$ the decay factor is zero because of the boundary conditions.
- We plot this decay factor for different values of x_0 in Fig. 7.4.

7.3 Rate of decay of $u(x,t)$

We have demonstrated conclusively that the solution to

$$\begin{aligned}&\frac{\partial u}{\partial t} = D\frac{\partial^2 u}{\partial x^2}, \qquad \text{on } [0,L], \qquad (*)\\ &u(x,t=0) = u_0,\\ &u(x=0,t) = u(x=L,t) = 0;\end{aligned}$$

decays to zero. Here, we answer the question, 'how fast?'

First, consider an equilibrium solution of the diffusion equation, $u_E(x)$. Then

$$0 = \frac{\partial u_E(x)}{\partial t} = D\frac{\partial^2 u_E}{\partial x^2}.$$

solving $d^2u_E(x)/dx^2 = 0$ gives $u_E(x) = Ax + B$. However, the zero boundary conditions force $A = B = 0$, and $u_E(x) = 0$. If we regard the diffusion equation (*) as modelling the temperature of a metal rod whose temperature is fixed at $0°$ at both ends, then physical intuition suggests that any initial temperature distribution must decay to zero – in agreement with our analytical result.

Now, recall again our solution to Eq. (*):

$$u(x,t) = \frac{4u_0}{\pi}\sum_{m=0}^{\infty}\frac{\sin\left(\frac{(2m+1)\pi x}{L}\right)}{2m+1}e^{-(2m+1)^2\pi^2 Dt/L^2}.$$

Let us try to bound the m^{th} term:

$$\left|\frac{\sin\left(\frac{(2m+1)\pi x}{L}\right)}{2m+1}e^{-(2m+1)^2\pi^2 Dt/L^2}\right| \le e^{-(2m+1)^2\pi^2 Dt/L^2} := e^{-(2m+1)^2\pi^2\tau}, \qquad \tau > 0.$$

Now

$$\begin{aligned}
(2m+1)^2\pi^2\tau &\ge (2m+1)\pi^2\tau, \\
-(2m+1)^2\pi^2\tau &\le -(2m+1)\pi^2\tau, \\
e^{-(2m+1)^2\pi^2\tau} &\le e^{-(2m+1)\pi^2\tau}, \\
&= \left(e^{-\pi^2\tau}\right)^{2m+1}, \\
&:= r^{2m+1},
\end{aligned}$$

hence

$$\begin{aligned}
|u(x,t)| &\le \frac{4u_0}{\pi}\sum_{m=0}^{\infty} r^{2m+1}, \\
&= \frac{4u_0 r}{\pi}\sum_{m=0}^{\infty}(r^2)^m, \\
&= \frac{4u_0}{\pi}\frac{r}{1-r^2}, \\
&= \frac{4u_0}{\pi}\frac{e^{-\pi^2\tau}}{1-e^{-2\pi^2\tau}}, \qquad \tau > 0.
\end{aligned}$$

Thus, the concentration (distribution) decays at least exponentially fast to zero.

Chapter 8

The 1-D diffusion equation: Properties II

Overview

Properties discussed: Well-posedness, uniqueness, various boundary conditions

8.1 Well-posed problems

The following reasonable conditions ought to be satisfied by a mathematical model of some physical system:

1. **Existence of solutions:** The mathematical model must possess at least one solution. Interpretation: the physical system exists over at least some finite time interval.
2. **Uniqueness of solutions:** For given boundary and initial conditions, the mathematical model has at most one solution. Interpretation: identical initial states of the system lead to the same outcome.
3. **Continuous dependence on parameters:** The solution of the mathematical model depends continuously on the initial conditions and parameters. Interpretation: Small changes in initial conditions or parameters lead to small changes in the outcome.

A model set of equations that satisfies these criteria is called **well posed**.

- The diffusion equation satisfies these criteria. Certainly at least one solution exists – we have constructed it using Fourier series. And, moreover, **by Result 5.3, this series can be differentiated term-by-term**.

- The continuous dependence on initial parameters is tricky and we do not consider it here. However, you may verify it numerically by looking at two numerical solutions whose initial conditions are 'similar' – the similarity between the two solutions will persist over time.

- Finally, there is the question of uniqueness, which we address below.

8.2 The diffusion equation: uniqueness of solutions

Let us define the sets

$$\Omega_t = \{(x,t)|x \in [0,L], t \in (0,\infty)\},$$
$$\overline{\Omega_t} = \{(x,t)|x \in [0,L], t \in [0,\infty)\},$$

and the function space

$$\mathcal{C}^2\left(\overline{\Omega_t}\right) = \left\{ u(x,t) \middle| u_{xx} \text{ is continuous in } \Omega_t \text{ and } u \text{ is continuous in } \overline{\Omega_t} \right\}.$$

We have the following theorem:

Theorem 8.1 *The diffusion equation studied so far, namely*

$$\frac{\partial u}{\partial t} = D\frac{\partial^2 u}{\partial x^2}, \qquad \textit{on } [0,L], \qquad (*)$$
$$u(x,t=0) = f(x),$$
$$u(x=0,t) = u(x=L,t) = 0;$$

has at most one solution in the space $\mathcal{C}^2\left(\overline{\Omega_t}\right)$.

Proof: Consider two solutions $u_1, u_2 \in \mathcal{C}^2\left(\overline{\Omega_t}\right)$ to the diffusion equation, with identical boundary and initial conditions. Form the difference $v = u_1 - u_2$. Then,

$$\begin{aligned} v_t &= (u_1-u_2)_t, \\ &= u_{1t} - u_{2t}, \\ &= Du_{1xx} - Du_{2xx}, \\ &= (u_1-u_2)_{xx}, \\ &= Dv_{xx}, \end{aligned}$$

and the difference of the solutions also satisfies the diffusion equation. Now consider the BCs and

ICs:

$$
\begin{aligned}
\text{BC}: &\quad v(0,t>0) = u_1(0,t>0) - u_2(0,t>0) = 0,\\
\text{BC}: &\quad v(L,t>0) = u_1(L,t>0) - u_2(L,t>0) = 0,\\
\text{IC}: &\quad v(x,t=0) = u_1(x,t=0) - u_2(x,t=0) = 0, \qquad 0<x<L.
\end{aligned}
$$

Let us form the $L^2(\Omega)$ norm of v and then differentiate it wrt time:

$$\|v\|_2^2 = \int_\Omega v^2(x,t)\,\mathrm{d}x.$$

$$
\begin{aligned}
\tfrac{1}{2}\frac{d}{dt}\|v\|_2^2 &= \tfrac{1}{2}\frac{d}{dt}\int_\Omega v(x,t)^2\,\mathrm{d}x,\\
&= \tfrac{1}{2}\int_\Omega \frac{\partial}{\partial t}v^2\,\mathrm{d}x,\\
&= \int_\Omega v v_t\,\mathrm{d}x,\\
&= D\int_\Omega v v_{xx}\,\mathrm{d}x,\\
&= D\int_\Omega \left[\frac{\partial}{\partial x}(v v_x) - v_x^2\right]\mathrm{d}x,\\
&= D\left[v(L)v_x(L) - v(0)v_x(0)\right] - D\int_\Omega v_x^2\,\mathrm{d}x,\\
&= -D\int_\Omega v_x^2,\mathrm{d}x.
\end{aligned}
$$

This is an ordinary differential equation in time and can be formally integrated:

$$\|v\|_2^2(t) = \|v\|_2^2(0) - 2D\int_0^t \|v_x\|_2^2(s)\,\mathrm{d}s,$$

hence

$$\|v\|_2^2(t) \le \|v\|_2^2(0) = 0.$$

Hence,

$$\|v\|_2^2(t) \le 0 \implies \|v\|_2^2(t) = 0.$$

The only way for this equation to be satisfied is if

$$v(x,t) = 0,$$

and the solutions agree, $u_1 = u_2$.

8.3 Variations on the boundary conditions

So far, we have solved the diffusion equation

$$\begin{aligned}
&\frac{\partial u}{\partial t} = D\frac{\partial^2 u}{\partial x^2}, \qquad \text{on } [0, L],\\
&u(x, t = 0) = f(x),\\
&u(x = 0, t) = u(x = L, t) = 0;
\end{aligned}$$

We have seen how these boundary conditions did not correspond very well to a model for particle diffusion, since there is a net flow of particles through the boundary. However, physical intuition does suggest it is a good model for the cooling of a metal rod, whose end points are held at a fixed temperature. Let us now discuss the different possibilities for boundary conditions in a more systematic way.

Dirichlet conditions

The function $u(x, t > 0)$ is specified on the boundaries:

$$\begin{aligned}
u(0, t > 0) &= g_1(t),\\
u(L, t > 0) &= g_2(t).
\end{aligned}$$

If the functions $g_1 = g_2 = 0$, then we have **homogeneous Dirichlet conditions**:

$$\begin{aligned}
u(0, t > 0) &= 0,\\
u(L, t > 0) &= 0
\end{aligned}$$

This is the case we have considered so far.

Neumann conditions

The **derivative** $u_x(x, t > 0)$ is specified on the boundaries:

$$\begin{aligned}
u_x(0, t > 0) &= g_1(t),\\
u_x(L, t > 0) &= g_2(t).
\end{aligned}$$

If the functions $g_1 = g_2 = 0$, then we have **homogeneous Neumann conditions**, corresponding to **no flux through the boundaries**. This case has been discussed briefly already:

$$\begin{aligned} u_x(0, t > 0) &= 0, \\ u_x(L, t > 0) &= 0. \end{aligned}$$

Mixed conditions

As the name suggests, this set is a mixture of Dirichlet and Neumann conditions:

$$\begin{aligned} \alpha_1 u_x(0, t > 0) + \alpha_2 u(0, t > 0) &= g_1(t), \\ \alpha_3 u_x(L, t > 0) + \alpha_4 u(L, t > 0) &= g_2(t). \end{aligned}$$

Periodic boundary conditions

The function $u(x, t > 0)$ has the same value on either boundary point:

$$u(0, t) = u(L, t), \qquad t > 0.$$

In practice, these are not very realistic boundary conditions but they are used in numerical experiments because they are easy to implement. However, they can be used to mimic an infinite domain, if the periodic length L is made long enough.

Examples

1. Solve the diffusion equation with homogeneous Neumann boundary conditions and constant initial condition

   $$u(x, t = 0) = u_0, \qquad 0 < x < L.$$

   A homework exercise!

2. Consider the diffusion equation with inhomogeneous Dirichlet conditions

$$\frac{\partial u}{\partial t} = D\frac{\partial^2 u}{\partial x^2}, \qquad \text{on } [0, L],$$

$$u(x, t = 0) = 0, \qquad 0 < x < L,$$
$$u(x = 0, t > 0) = 0,$$
$$u(x = L, t > 0) = u_1,$$

where u_1 is a constant.

- Separation of variables fails here. To see why, let us valiantly attempt the solution $u(x,t) = X(x)T(t)$. Then, at the boundary $x = L$, we would have $X(L)T(t) = u_1$, which implies that any coefficients in the $X(x)$-solution depend on time – impossible.

- Instead, we break up the solution into a bit that solves the diffusion equation with the given BCs (**particular integral**), and a bit that solves that solves the diffusion equation with zero BCs (**homogeneous solution**).

- Because the BCs are independent of time, we expect the PI to be independent of time also – call the PI $u_E(x)$:

$$\begin{aligned} u_E''(x) &= 0, \qquad 0 < x < L, \\ u_E(0) &= 0, \\ u_E(L) &= u_1. \end{aligned}$$

  This solution to the ODE is $u_E(x) = Ax + B$, and the BCs give $B = 0$ and $AL = u_1$, hence

$$u_E(x) = u_1(x/L).$$

- Now write

$$u(x,t) := u_0(x,t) + u_E(x) \implies u_0(x,t) = u(x,t) - u_E(x).$$

  where $u(x,t)$ is the solution to the full equation and u_0 is to be determined. By linearity, u_0 solves the diffusion equation:

$$u_{0t} = Du_{0xx}, \qquad 0 < x < L,$$

and

$$\begin{aligned}
&u_0(x=0,t>0)=0,\\
&u_0(x=L,t>0)=u(L,t)-u_E(L,t)=u_1-u_1=0,\\
&u_0(x,t=0)=u(0,t)-u_e(0,t)=0-u_1(x/L):=f(x),\qquad 0<x<L.
\end{aligned}$$

Thus, $u_0(x,t)$ satisfies the diffusion equation with homogeneous BCs – it is the **homogeneous solution**.

- But we know how to solve such an equation: the solution is

$$u_0(x,t)=\sum_{n=1}^{\infty}C_n\sin\left(\frac{n\pi x}{L}\right)e^{-n^2\pi^2Dt/L^2}.$$

Moreover,

$$\begin{aligned}
u_0(x,t=0) &= \sum_{n=1}^{\infty}C_n\sin\left(\frac{n\pi x}{L}\right),\\
&= u(x,t=0)-u_E(x),\\
&= 0-u_1(x/L).
\end{aligned}$$

Multiply both sides by $\sin(m\pi x/L)$ and integrate over $[0,L]$:

$$\begin{aligned}
-(u_1/L)\int_0^L x\sin\left(\frac{n\pi x}{L}\right)\mathrm{d}x &= \sum_{n=1}^{\infty}C_n\int_0^L\sin\left(\frac{n\pi x}{L}\right)\sin\left(\frac{m\pi x}{L}\right)\mathrm{d}x,\\
&= \sum_{n=1}^{\infty}C_n(L/2)\delta_{nm},\\
&= C_mL/2.
\end{aligned}$$

In other words,

$$\begin{aligned}
C_m &= -\frac{2u_1}{L^2}\int_0^L x\sin\left(\frac{m\pi x}{L}\right)\mathrm{d}x,\\
&= -\frac{2u_1}{L^2}\frac{L^2}{\pi^2}\int_0^\pi y\sin(my)\,\mathrm{d}y,\\
&= -\frac{2u_1}{\pi^2}\frac{1}{m^2}\left[\sin(my)-my\cos(my)\right]_{y=0}^{y=\pi},\\
&= -\frac{2u_1}{\pi^2}\frac{1}{m}\left[-\pi\cos(m\pi)\right],\\
&= \frac{2u_1}{m\pi}(-1)^m,
\end{aligned}$$

hence

$$C_m = \frac{2u_1}{m\pi}(-1)^m,$$

and

$$u_0(x,t) = \sum_{n=1}^{\infty} \frac{2u_1}{n\pi}(-1)^n \sin\left(\frac{n\pi x}{L}\right) e^{-n^2\pi^2 Dt/L^2}.$$

- Now put it all together:

$$u(x,t) = u_0(x,t) + u_E(x),$$

hence

$$u(x,t) = \frac{u_1}{L}x + \frac{2u_1}{n\pi}\sum_{n=1}^{\infty} \frac{(-1)^n}{n} \sin\left(\frac{n\pi x}{L}\right) e^{-n^2\pi^2 Dt/L^2}$$

- The homogeneous part of the solution is called the **transient part** and it decays to zero, leaving only the particular integral, or equilibrium part.

Chapter 9

The diffusion equation: Sources

9.1 Motivation for discussing diffusion with sources

We introduced the diffusion equation as a model for the 'smoothing-out' of concentration gradients over time, where the concentration measures the density of particles:

$$\frac{\partial u}{\partial t} = D\frac{\partial^2 u}{\partial x^2}, \qquad x \in (0, L),$$

with suitable boundary and initial conditions. What happens if particles are, at the same time, being injected into the system? Then, we study the equation

$$\frac{\partial u}{\partial t} = D\frac{\partial^2 u}{\partial x^2} + q, \qquad x \in (0, L),$$

where $q(x,t)$ is a function called the **source**. Dimensionally, we must have

$$\frac{[\text{Concentration}]}{[\text{Time}]} = [q]\,,$$

thus, the source has as its interpretation the **rate** at which matter (or, in other applications, heat) is being injected into the system. In this chapter we solve such equations.

9.2 Time-independent sources

In this section we assume $q = q(x)$ only. For definiteness, we assume the following BCs and ICs:

$$\begin{array}{ll} \text{BC} & u(0, t > 0) = b_1 = \text{Const.}, \\ \text{BC} & u(L, t > 0) = b_2 = \text{Const.}, \\ \text{IC} & u(x, t = 0) = f(x), \qquad 0 < x < L. \end{array}$$

As in the case of inhomogeneous BCs (Ch. 8), we introduce a **particular integral** to soak up the contribution from the source and the boundary conditions:

$$u(x,t) = u_0(x,t) + u_E(x),$$

where

$$\begin{aligned} u_E''(x) + q(x) &= 0, \qquad 0 < x < L \\ u_E(0) &= b_1, \\ u_E(L) &= b_2, \end{aligned}$$

This is a simple ODE, which we assume can be solved. Next, we solve for the homogeneous part $u_0(x,t)$, where

$$u(x,t) = u_0(x,t) + u_E(x) = \text{full solution},$$

and

$$u_0(x,t) = u(x,t) - u_E(x).$$

Study $u_0(x,t)$:

$$\begin{aligned} u_{0t} &= u_t(x,t) - 0, \\ &= Du_{xx} + q(x), \\ &= D\left(u_0 + u_E\right)_{xx} + q(x), \\ &= Du_{0xx} + Du_{Exx} + q(x), \\ &= Du_{0xx}, \end{aligned}$$

with BCs

$$
\begin{aligned}
&\text{BC} \qquad u_0(0, t>0) = 0, \\
&\text{BC} \qquad u_0(L, t>0) = 0, \\
&\text{IC} \qquad u_0(x, t=0) = f(x) - u_E(x), \qquad 0 < x < L.
\end{aligned}
$$

This is the homogeneous diffusion equation with homogeneous Dirichlet BCs. But we know the solution then:

$$
u_0(x,t) = \sum_{n=1}^{\infty} C_n \sin\left(\frac{\pi n x}{L}\right) \mathrm{e}^{-n^2\pi^2 Dt/L^2}.
$$

Moreover,

$$
\begin{aligned}
u_0(x, t=0) &= \sum_{n=1}^{\infty} C_n \sin\left(\frac{n\pi x}{L}\right), \\
&= f(x) - u_E(x).
\end{aligned}
$$

As before, multiply both sides by $\sin(m\pi x/L)$ and integrate over $[0, L]$:

$$
\begin{aligned}
\int_0^L [f(x) - u_E(x)] \sin\left(\frac{m\pi x}{L}\right) \mathrm{d}x &= \int_0^L \sum_{n=1}^{\infty} C_n \sin\left(\frac{n\pi x}{L}\right) \sin\left(\frac{m\pi x}{L}\right) \mathrm{d}x, \\
&= \sum_{n=1}^{\infty} C_n \, (L/2) \, \delta_{m,n},
\end{aligned}
$$

hence

$$
C_m = \frac{2}{L} \int_0^L [f(x) - u_E(x)] \sin\left(\frac{m\pi x}{L}\right) \mathrm{d}x.
$$

But

$$
u(x,t) = u_0(x,t) + u_E(x),
$$

hence

$$
\begin{aligned}
u(x,t) &= u_E(x) + \sum_{n=1}^{\infty} C_n \sin\left(\frac{\pi n x}{L}\right) \mathrm{e}^{-n^2\pi^2 Dt/L^2}, \\
C_n &= \frac{2}{L} \int_0^L [f(x) - u_E(x)] \sin\left(\frac{n\pi x}{L}\right) \mathrm{d}x.
\end{aligned}
$$

9.3 Time-dependent sources

We consider the following rather special case:

$$u_t = Du_{xx}, \qquad 0 < x < L,$$

with the following **inhomogeneous boundary conditions** (The inhomogeneity is, effectively, a source):

$$\begin{array}{ll} \text{BC} & u(0, t>0) = 0, \\ \text{BC} & u(L, t>0) = A\cos(\omega t), \\ \text{IC} & u(x, t=0) = f(x), \qquad 0 < x < L. \end{array}$$

- Physically, the BC at $x = L$ corresponds to a periodic modulation of the temperature at the end of the rod, or a time-periodic injection of particles into the domain.
- We cannot expect a time-independent state as $t \to \infty$ because in effect, there is a driving force.
- However, we do expect that whatever time-independence is left as $t \to \infty$ is periodic, with frequency ω.
- Thus, we make the following trial solution:

$$\begin{aligned} u(x,t) &= u_0(x,t) + u_{QS}(x,t), \\ &= u_0(x,t) + A(x)\cos(\omega t + \phi(x)), \end{aligned}$$

  where $\lim_{t\to\infty} u_0(x,t) = 0$ is a transient and $u_{QS}(x,t)$ is a **quasi-steady state**.

We construct the QS state such that it satisfies the BCs, and we construct the transient such that it satisfies the null BCs. To solve for the QS state, we work in complex numbers. Thus, the trial solution proposed for the quasi-steady part is

$$u_{QS} = \Re\left[U(x)\mathrm{e}^{\mathrm{i}\omega t}\right].$$

(To make a connection with the previous, non-complex form of the ansatz, take $U(x) = A(x)\mathrm{e}^{\mathrm{i}\phi(x)}$). We now substitute u_{QS} into the diffusion equation:

$$\frac{\partial u_{QS}}{\partial t} = D\frac{\partial^2 u_{QS}}{\partial x^2}, \qquad 0 < x < L, \qquad u_{QS}(x,t) \in \mathbb{C},$$

with BCs

$$\text{BC} \qquad u_{QS}(0, t > 0) = 0,$$
$$\text{BC} \qquad u_{QS}(L, t > 0) = A\mathrm{e}^{\mathrm{i}\omega t}.$$

But $u_{QS} = \mathrm{e}^{\mathrm{i}\omega t}U(x)$, hence, the diffusion equation becomes

$$\mathrm{i}\omega U(x) = DU''(x), \qquad 0 < x < L, \qquad u(x,t) \in \mathbb{C},$$

with BCs

$$\text{BC} \qquad U(0) = 0,$$
$$\text{BC} \qquad U(L) = A.$$

The ODE $\mathrm{i}\omega U(x) = DU''(x)$ has the general solution

$$U(x) = \alpha \mathrm{e}^{\left[\sqrt{\mathrm{i}\omega/D}\right]_1 x} + \beta \mathrm{e}^{\left[\sqrt{\mathrm{i}\omega/D}\right]_2 x},$$

where

$$\begin{aligned}
\sqrt{\mathrm{i}\omega/D} &= \left(\frac{\omega}{D}\mathrm{e}^{\mathrm{i}\pi/2+\mathrm{i}2n\pi}\right)^{1/2}, \\
&= \begin{cases} (\omega/D)^{1/2}\mathrm{e}^{\mathrm{i}\pi/4}, \\ (\omega/D)^{1/2}\mathrm{e}^{\mathrm{i}5\pi/4} \end{cases}, \\
&= \begin{cases} (\omega/2D)^{1/2}(1+\mathrm{i}), \\ (\omega/2D)^{1/2}(-1-\mathrm{i}) \end{cases}.
\end{aligned}$$

Hence,

$$\begin{aligned}
U(x) &= \alpha \mathrm{e}^{(\omega/2D)^{1/2}(1+\mathrm{i})x} + \beta \mathrm{e}^{(\omega/2D)^{1/2}(-1-\mathrm{i})x}, \\
&= \alpha \mathrm{e}^{(\omega L^2/2D)^{1/2}(1+\mathrm{i})(x/L)} + \beta \mathrm{e}^{(\omega L^2/2D)^{1/2}(-1-\mathrm{i})(x/L)}, \\
&= \alpha \mathrm{e}^{\mu(1+\mathrm{i})(x/L)} + \beta \mathrm{e}^{\mu(-1-\mathrm{i})(x/L)}, \qquad \mu = (\omega L^2/2D)^{1/2}.
\end{aligned}$$

Implement the BCs:

$$\begin{aligned}
U(0) &= \alpha + \beta, \\
&= 0.
\end{aligned}$$

Hence, $\alpha = -\beta$ and

$$\begin{aligned}
U(x) &= \alpha\left[\mathrm{e}^{\mu(1+\mathrm{i})(x/L)} - \mathrm{e}^{\mu(-1-\mathrm{i})(x/L)}\right], \\
&= \alpha\left[\mathrm{e}^{\mathrm{i}\mu(x/L)}\mathrm{e}^{\mu(x/L)} - \mathrm{e}^{-\mathrm{i}\mu(x/L)}\mathrm{e}^{-\mu(x/L)}\right], \\
&= \alpha\left[\mathrm{e}^{\mathrm{i}\mu(x/L)}\mathrm{e}^{\mu(x/L)} + \underbrace{\mathrm{e}^{\mathrm{i}\mu(x/L)}\mathrm{e}^{-\mu(x/L)} - \mathrm{e}^{\mathrm{i}\mu(x/L)}\mathrm{e}^{-\mu(x/L)}}_{\text{add and subtract}} - \mathrm{e}^{-\mathrm{i}\mu(x/L)}\mathrm{e}^{-\mu(x/L)}\right], \\
&= \alpha\left[\mathrm{e}^{\mathrm{i}\mu(x/L)}\left(\mathrm{e}^{\mu(x/L)} + \mathrm{e}^{-\mu(x/L)}\right) - \mathrm{e}^{-\mu(x/L)}\left(\mathrm{e}^{\mathrm{i}\mu(x/L)} + \mathrm{e}^{-\mathrm{i}\mu(x/L)}\right)\right], \\
&= 2\alpha\left[\mathrm{e}^{\mathrm{i}\mu x/L}\cosh(\mu x/L) - \mathrm{e}^{-\mu x/L}\cos(\mu x/L)\right].
\end{aligned}$$

At $x = L$,

$$\begin{aligned}
U(L) &= 2\alpha\left[\mathrm{e}^{\mathrm{i}\mu}\cosh(\mu) - \mathrm{e}^{-\mu}\cos(\mu)\right] \\
&= A,
\end{aligned}$$

hence

$$\alpha = \frac{A}{2\left[\mathrm{e}^{\mathrm{i}\mu}\cosh(\mu) - \mathrm{e}^{-\mu}\cos(\mu)\right]}.$$

and

$$U(x) = A\left(\frac{\mathrm{e}^{\mathrm{i}\mu x/L}\cosh(\mu x/L) - \mathrm{e}^{-\mu x/L}\cos(\mu x/L)}{\mathrm{e}^{\mathrm{i}\mu}\cosh(\mu) - \mathrm{e}^{-\mu}\cos(\mu)}\right).$$

Since $u_{QS}(x,t) = \Re\left(U(x)\mathrm{e}^{\mathrm{i}\omega t}\right)$, we have, finally,

$$u_{QS}(x,t) = \Re\left[A\mathrm{e}^{\mathrm{i}\omega t}\frac{\mathrm{e}^{\mathrm{i}\mu x/L}\cosh(\mu x/L) - \mathrm{e}^{\mathrm{i}\mu x/L}\cos(\mu x/L)}{\mathrm{e}^{\mathrm{i}\mu}\cosh(\mu) - \mathrm{e}^{-\mu}\cos(\mu)}\right].$$

The general solution is the particular integral ($u_{QS}(x,t)$), plus the general solution to the diffusion equation with zero boundary conditions:

$$
\begin{aligned}
u(x,t) &= u_{QS}(x,t)+\sum_{n=1}^{\infty}C_n\mathrm{e}^{-\pi^2n^2Dt/L^2}\sin\left(\frac{n\pi x}{L}\right),\\
&= \Re\left[A\mathrm{e}^{\mathrm{i}\omega t}\left(\frac{\mathrm{e}^{\mathrm{i}\mu x/L}\cosh(\mu x/L)-\mathrm{e}^{-\mu x/L}\cos(\mu x/L)}{\mathrm{e}^{\mathrm{i}\mu}\cosh(\mu)-\mathrm{e}^{-\mu}\cos(\mu)}\right)\right]\\
&\qquad +\sum_{n=1}^{\infty}C_n\mathrm{e}^{-\pi^2n^2Dt/L^2}\sin\left(\frac{n\pi x}{L}\right).
\end{aligned}
$$

The C_n-values are determined in the usual way:

$$
\begin{aligned}
u(x,t=0) &= f(x),\\
&= \Re\left[A\left(\frac{\mathrm{e}^{\mathrm{i}\mu x/L}\cosh(\mu x/L)-\mathrm{e}^{-\mu x/L}\cos(\mu x/L)}{\mathrm{e}^{\mathrm{i}\mu}\cosh(\mu)-\mathrm{e}^{-\mu}\cos(\mu)}\right)\right]+\sum_{n=1}^{\infty}C_n\sin\left(\frac{n\pi x}{L}\right),
\end{aligned}
$$

and

$$
C_n=\frac{2}{L}\int_0^L\mathrm{d}x\left\{f(x)-\Re\left[A\left(\frac{\mathrm{e}^{\mathrm{i}\mu x/L}\cosh(\mu x/L)-\mathrm{e}^{-\mu x/L}\cos(\mu x/L)}{\mathrm{e}^{\mathrm{i}\mu}\cosh(\mu)-\mathrm{e}^{-\mu}\cos(\mu)}\right)\right]\right\}.
$$

This is the final answer.

9.4 The perfect wine cellar

At what depth should we build a wine cellar? Take

$$D=2\times10^{-3}\text{cm}^2/\text{s}.$$

Answer: We model this problem as a diffusion equation. Let $x=0$ be some reference depth far below the surface of the earth, at which the temperature remains constant. Let the surface be at $x=L$ (See Fig. 9.1). We have the following boundary conditions for the temperature:

$$
\begin{aligned}
&\text{BC} \qquad T(0,t>0)=T_0=\text{Const.},\\
&\text{BC} \qquad T(L,t>0)=T_1+A\cos\omega t.
\end{aligned}
$$

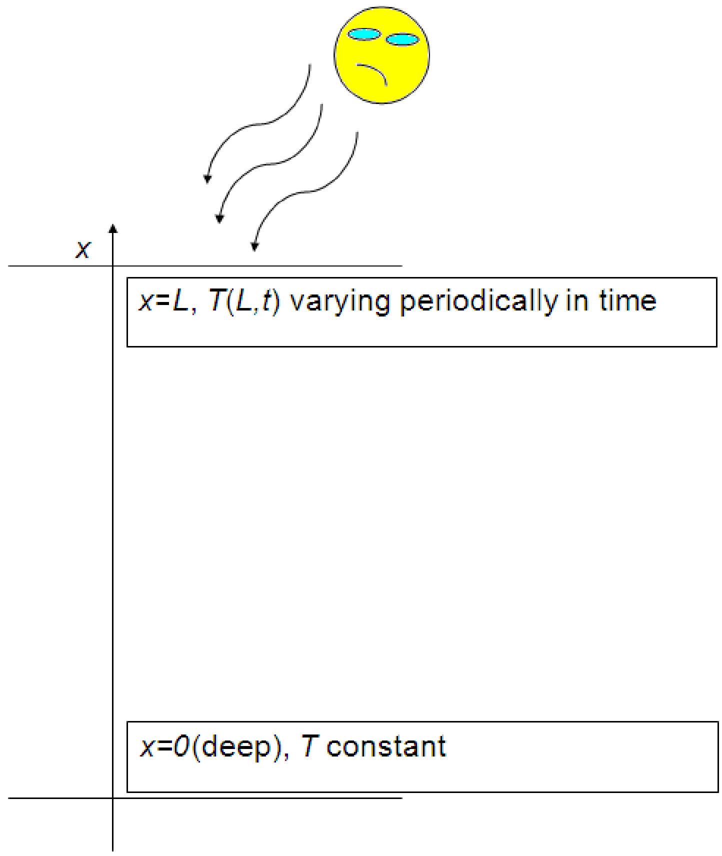


Figure 9.1: Coordinates for the cellar problem.

Here, $T(L,t) = T_1 + A\cos\omega t$ represents forcing by the sun on the earth's surface, with

$$\frac{2\pi}{\omega} = 1 \text{ year}.$$

We solve the diffusion equation. Because the diffusion equation is linear, we can subtract off a constant from the solution and from the BCs, and not change the answer. This corresponds to re-basing the temperature scale. Thus, we solve instead

$$\begin{aligned} &\text{BC} \qquad T(0, t>0) = 0 = \text{Const.}, \\ &\text{BC} \qquad T(L, t>0) = T_1 - T_0 + A\cos\omega t := T_s + A\cos\omega t. \end{aligned}$$

We are only interested in the particular integral $T_{QS}(x,t)$, as the transient effects will certainly have died off after less than one year! We break up the PI into two separate PIs, $T_{QS}(x,t) = v(x)+u(x,t)$, where

$$\frac{\partial v(x)}{\partial t} = D\frac{\partial^2 v}{\partial x^2},$$

$$\begin{aligned} &\text{BC} \qquad v(0, t>0) = 0 = \text{Const.}, \\ &\text{BC} \qquad v(L, t>0) = T_s. \end{aligned}$$

$$\begin{aligned} u(x,t) &= \Re\left(U(x)e^{i\omega t}\right), \\ \frac{\partial u}{\partial t} &= D\frac{\partial^2 u}{\partial x^2}, \end{aligned}$$

$$\begin{aligned} &\text{BC} \qquad U(0, t>0) = 0 = \text{Const.}, \\ &\text{BC} \qquad U(L, t>0) = A. \end{aligned}$$

The solution is available immediately:

$$\begin{aligned} v(x) &= T_s(x/L), \\ u(x,t) &= \Re\left[Ae^{i\omega t}\left(\frac{e^{i\mu x/L}\cosh(\mu x/L) - e^{-\mu x/L}\cos(\mu x/L)}{e^{i\mu}\cosh(\mu) - e^{-\mu}\cos(\mu)}\right)\right]. \end{aligned}$$

Now let's play a few final tricks. Let $z = L - x$, so $z = 0$ corresponds to the earth's surface, and

$z = L$ corresponds to the depths. We have

$$\begin{aligned} \mu^2 &= \frac{\omega L^2}{2D}, \\ &= \tfrac{1}{2}\frac{2\pi}{365 \times 24 \times \times 60 \times 60}\frac{1}{0.002} \approx \left(2.5 \times 10^{-5}\text{cm}^{-2}\right) L^2. \end{aligned}$$

If L is large, then μ^2 will be large too. We shall leave L arbitrary, but note: if $L \approx 100\text{m} = 10000\text{cm}$, then

$$\mu^2 \approx 2500,$$

and $\mu \approx 50$. In any case, we assume that $\mathrm{e}^{-\mu}$ is tiny, and $\cosh(\mu) \approx \mathrm{e}^{\mu}/2$ is large. Thus,

$$\begin{aligned} T(x,t) &\approx T_s(x/L) + \Re\left[A\mathrm{e}^{\mathrm{i}\omega t}\left(\frac{\mathrm{e}^{\mathrm{i}\mu x/L}\cosh(\mu x/L) - \mathrm{e}^{-\mu x/L}\cos(\mu x/L)}{\mathrm{e}^{\mathrm{i}\mu}\cosh(\mu) - \mathrm{e}^{-\mu}\cos(\mu)}\right)\right], \\ &= T_s(x/L) + 2\Re\left[A\mathrm{e}^{\mathrm{i}\omega t}\mathrm{e}^{-\mu}\frac{\cosh(\mu x/L)\mathrm{e}^{\mathrm{i}\mu x/L}}{\mathrm{e}^{\mathrm{i}\mu}}\right], \\ &= T_s\left(\frac{L-z}{L}\right) + 2\Re\left[A\mathrm{e}^{\mathrm{i}\omega t}\mathrm{e}^{-\mu}\frac{\cosh\left(\frac{\mu(L-z)}{L}\right)\exp\left(\mathrm{i}\frac{\mu(L-z)}{L}\right)}{\mathrm{e}^{\mathrm{i}\mu}}\right], \\ &= T_s[1-(z/L)] + 2\Re\left[A\mathrm{e}^{\mathrm{i}\omega t}\mathrm{e}^{-\mu}\cosh\left(\mu - (\mu z/L)\right)\mathrm{e}^{-\mathrm{i}\mu z/L}\right], \\ &= T_s[1-(z/L)] + \Re\left[A\mathrm{e}^{\mathrm{i}\omega t}\mathrm{e}^{-\mu}\left(\mathrm{e}^{\mu}\mathrm{e}^{-\mu z/L} + \mathrm{e}^{-\mu}\mathrm{e}^{\mu z/L}\right)\mathrm{e}^{-\mathrm{i}\mu z/L}\right]. \end{aligned}$$

Realistically, the we will have $z \ll L$ for the wine cellar, so $\mathrm{e}^{-\mu}\mathrm{e}^{\mu z/L} \approx \mathrm{e}^{-\mu}\mathrm{e}^{0} \ll 1$, so

$$u(x,t) \approx T_s + \Re\left[A\mathrm{e}^{\mathrm{i}\omega t}\mathrm{e}^{-\mu z/L}\mathrm{e}^{-\mathrm{i}\mu z/L}\right],$$

or

$$u(z,t) \approx T_s + A\mathrm{e}^{-\mu z/L}\cos\left(-\frac{\mu z}{L} + \omega t\right).$$

But the answer is independent of the arbitrary (but large) depth L, since

$$\mu/L = \sqrt{\omega/2D}.$$

Thus, the solution is

$$u(z,t) \approx T_s + A\mathrm{e}^{-\sqrt{\omega/2D}z}\cos\left(-\sqrt{\omega/2D}z + \omega t\right). \qquad (*)$$

Recall what the question asked: We need to find the optimal depth of a wine cellar. The cellar needs to be relatively cool in the summer and relatively warm in the winter. In other words, it needs to be exactly out of phase with the forcing term $A \cos \omega t$. This can be accomplished if we set

$$\pi = \sqrt{\omega/2} z,$$

or

$$z = \pi \sqrt{2D/\omega}.$$

Plugging in the values:

$$z = \pi \sqrt{2 \left(\frac{0.002\text{cm}^2/\text{sec}}{2\pi} \right) 365 \times 24 \times 60 \times 60\text{sec}} = 445\text{cm} \approx 4.45\text{m}.$$

Of course, a deeper cellar would be more desirable, because of the damping factor in Eq. (*). However, difficulties would then arise in excavating such a cellar. At the depth $z = 4.45$m,

- The **phase** of the temperature modulation is exactly opposite to that at the surface, meaning the cellar is relatively warm in the winter and relatively cool in the summer;
- The **amplitude** of the modulation (relative to the surface) is

$$A\mathrm{e}^{-\sqrt{\omega/2D} z} = A\mathrm{e}^{-\pi} \approx 0.04A,$$

  and these variations are only 4% of those at the surface.

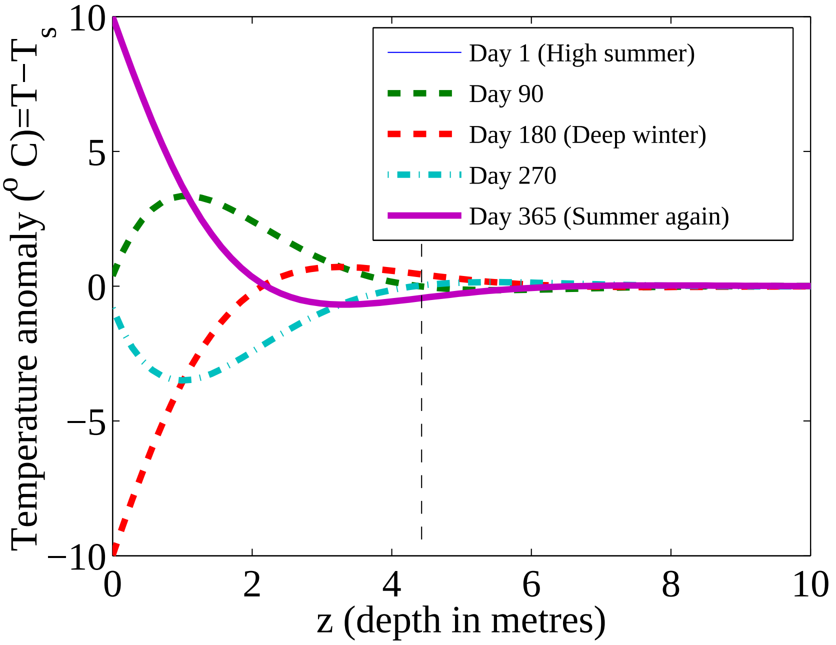


Figure 9.2: Temperature variations with depth.

Chapter 10

Linear PDEs: General formulation

Overview

So far our knowledge of PDEs is based solely on the diffusion equation. This is a good start! In this chapter we place the diffusion equation in the context of general, linear partial differential equations. As you have no doubt noticed, there is much overlap between the theory of vector spaces and the theory of linear PDEs. The overlap continues...

10.1 Linear operators

Definition 10.1 *Let V_1 and V_2 be real vector spaces. A linear operator T is a map*

$$\begin{aligned} T : V_1 &\rightarrow V_2, \\ \boldsymbol{x} &\rightarrow T\boldsymbol{x}, \end{aligned}$$

such that

$$\begin{aligned} T(\boldsymbol{x}+\boldsymbol{y}) &= T\boldsymbol{x}+T\boldsymbol{y}, \\ T(\lambda\boldsymbol{x}) &= \lambda T\boldsymbol{x}, \end{aligned}$$

for all $\boldsymbol{x}, \boldsymbol{y} \in V_1$ and $\lambda \in \mathbb{R}$.

Definition 10.2 *Let V_1 and V_2 be real vector spaces and let T be a linear operator, $T : V_1 \rightarrow V_2$. The* **kernel** *of the linear operator is the set*

$$\ker(T) = \{\boldsymbol{x} \in V_1 | T\boldsymbol{x} = 0\}.$$

The kernel of T is a vector subspace of V_1.

Examples:

- An $n \times n$ matrix is a linear operator on $\mathbb{R}^n$, and maps $\mathbb{R}^n$ to itself.

- The matrix
$$A = \begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix},$$
where a and b are nonzero real numbers is a linear operator that maps vectors $\boldsymbol{x} = (x, y)$ in $\mathbb{R}^2$ to vectors in $\mathbb{R}^2$. The kernel of A is the set of all vectors (x, y) such that
$$\begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 0.$$
In other words,
$$y = -(a/b)x,$$
which is a one-dimensional subspace of $\mathbb{R}^2$.

- Let $C^r(\omega)$ be the space of all r-times continuously differentiable functions on the open interval $\Omega \subset \mathbb{R}$. Then the usual derivative operation is a linear operator:
$$\begin{aligned} d/dx : C^r(\Omega) &\rightarrow C^{r-1}(\Omega), \\ f(x) &\rightarrow (df/dx), \end{aligned}$$
since
$$\begin{aligned} (d/dx)\,[f(x) + g(x)] &= (df/dx) + (dg/dx), \\ (d/dx)\,[\lambda f(x)] &= \lambda(df/dx), \end{aligned}$$
for all $f(x), g(x) \in C^r(\Omega)$ and $\lambda \in \mathbb{R}$.

The kernel of d/dx is the vector subspace of all constant functions.

- Let
$$\Omega_t = \{(x,t) | x \in [0,L], t \in (0,\infty)\},$$
$$\overline{\Omega_t} = \{(x,t) | x \in [0,L], t \in [0,\infty)\},$$

and let

$$\begin{aligned}
\mathcal{C}^{2,1}(\Omega_t) &= \left\{ u(x,t) \middle| u_{xx} \text{ is continuous in } \Omega_t \text{ and } u_t \text{ is continuous in } \Omega_t \right\}, \\
\mathcal{C}^{2,1}\left(\overline{\Omega_t}\right) &= \left\{ u(x,t) \middle| u_{xx} \text{ is continuous in } \overline{\Omega_t} \text{ and } u_t \text{ is continuous in } \overline{\Omega_t} \right\}.
\end{aligned}$$

Then the **diffusion operator**

$$\mathcal{L} = \frac{\partial}{\partial t} - D\frac{\partial^2}{\partial x^2}$$

is a linear operator that acts on the space $\mathcal{C}^{2,1}\left(\overline{\Omega_t}\right)$.

10.2 The principle of superposition

As before, consider the set

$$\Omega_t = \{(x,t)|x \in [0,L], t \in (0,\infty)\},$$

and consider the diffusion operator

$$\mathcal{L} = \frac{\partial}{\partial t} - D\frac{\partial^2}{\partial x^2}.$$

on the space of functions $C^{2,1}(\Omega_t)$. The kernel of the operator is the set of solutions of $\mathcal{L}u(x,t) = 0$:

$$u(x,t) \in \ker(\mathcal{L}) \text{ iff } \mathcal{L}u(x,t) = 0,$$

This is a vector subspace of $C^{2,1}(\Omega_t)$. But a vector subspace is closed under addition and scalar multiplication:

$$u_1, u_2 \in \ker(\mathcal{L}) \implies \lambda_1 u_1 + \lambda_2 u_2 \in (\ker \mathcal{L}).$$

This result is not unique to the diffusion operator. Indeed, we have the following **principle of superposition**:

Let $\mathcal{M}$ be a linear differential operator on some space of suitably differentiable functions $C(\Omega)$, where Ω is the domain of definition of the operator. Then,

$$u_1, u_2 \in \ker(\mathcal{M}) \implies \lambda_1 u_1 + \lambda_2 u_2 \in (\ker \mathcal{M});$$

in other words, **if u_1 and u_2 satisfy the linear PDE $\mathcal{M}u = 0$, then any linear combination of these two solutions also satisfies the PDE.**

Example: Let $u_1(x,t)$ solve

$$\mathcal{M}u = 0,$$

where $\mathcal{M}$ is some linear operator in space and time, subject to the following **linear** boundary and initial conditions:

$$\begin{aligned}
\text{BC}: &\qquad u(0,t>0) = 0,\\
\text{BC}: &\qquad u(L,t>0) = 0,\\
\text{IC}: &\qquad u(x,t=0) = f(x), \qquad 0<x<L,
\end{aligned}$$

and let $u_2(x,t)$ solve

$$\mathcal{M}u = 0,$$

subject to the following **linear** boundary and initial conditions:

$$\begin{aligned}
\text{BC}: &\qquad u(0,t>0) = g(t),\\
\text{BC}: &\qquad u(L,t>0) = 0,\\
\text{IC}: &\qquad u(x,t=0) = 0, \qquad 0<x<L,
\end{aligned}$$

Then the linear combination

$$\lambda_1 u_1 + \lambda_2 u_2$$

solves

$$\mathcal{M}u = 0,$$

with boundary and initial conditions:

$$\begin{aligned}
\text{BC}: &\qquad u(0,t>0) = \lambda_2 g(t),\\
\text{BC}: &\qquad u(L,t>0) = 0,\\
\text{IC}: &\qquad u(x,t=0) = \lambda_1 f(x), \qquad 0<x<L.
\end{aligned}$$

10.3 Beyond the diffusion equation

We have mentioned that the superposition principle holds not only for the diffusion equation, but for any linear operator. Let us think about other operators we might encounter in applied mathematics:

- The wave equation for a disturbance $\phi(x,t)$:

$$\frac{1}{c^2}\frac{\partial^2\phi}{\partial t^2} = \frac{\partial^2\phi}{\partial x^2},$$

- Laplace's equation in a two-dimensional domain, for a scalar field $\phi(x, y)$

$$\frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2} = 0,$$

- The linear advection equation for a concentration (scalar) field $\phi(x, t)$:

$$\frac{\partial \phi}{\partial t} + c(x, t)\frac{\partial \phi}{\partial x} = 0.$$

Definition 10.3 *Let $\phi(x, t)$ be a function of space and time.*

- *A second-order PDE in the function ϕ is a relationship of the form*

$$F\left[x, t, \phi, \partial_x \phi, \partial_t \phi, \partial_{xt}\phi, \partial_{xx}\phi, \partial_{tt}\phi\right] = 0.$$

The PDE is **linear** *if the function F is linear in its last six variables. It is called* **quasi-linear** *if F is linear in its last three variables.*

- *A first-order PDE in the function ϕ is a relationship of the form*

$$F\left[x, t, \phi, \partial_x \phi, \partial_t \phi\right] = 0.$$

The PDE is **linear** *if the function F is linear in its three variables. It is called* **quasi-linear** *if F is linear in its last two variables.*

1. The heat equation is a linear second-order PDE, with

$$F\left[\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6, \alpha_7, \alpha_8\right] = \alpha_5 - D\alpha_7.$$

2. The wave equation is a linear second-order PDE, with

$$F\left[\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6, \alpha_7, \alpha_8\right] = \frac{1}{c^2}\alpha_8 - \alpha_7$$

3. The linear advection equation is a linear first-order PDE, with

$$F\left[\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5\right] = \alpha_5 + c\left(\alpha_1, \alpha_2\right)\alpha_4.$$

4. Burgers equation,
$$\frac{\partial u}{\partial t}+u\frac{\partial u}{\partial x}=D\frac{\partial^2 u}{\partial x^2}$$
is a quasilinear PDE:
$$F\left[\alpha_1,\alpha_2,\alpha_3,\alpha_4,\alpha_5,\alpha_6,\alpha_7,\alpha_8\right]=\alpha_5+\alpha_3\alpha_4-D\alpha_7$$

Other, higher-order variants exist. In general, a PDE is called linear if the function ϕ and its derivatives appear in a linear way in the PDE relation. It is called quasi-linear if the highest-order derivatives appear in a linear way.

10.4 Classification of second-order PDEs

In the following treatment, let $\phi(t,x_1,x_2,\cdots,x_n)$ be a smooth function of t and the coordinates $\boldsymbol{x}=(x_1,\cdots,x_n)\in\Omega$, where Ω is some open subset of $\mathbb{R}^n$. **Furthermore, let $A_{ij}(\boldsymbol{x})$, or possibly $A_{ij}(\boldsymbol{x},t)$, be a real, symmetric matrix**.

1. No t-dependence: Consider the PDE
$$0=\sum_{i,j=1}^{n}A_{ij}(\boldsymbol{x})\frac{\partial^2\phi}{\partial x_i\partial x_j}+\sum_{i=1}^{n}b_i(\boldsymbol{x})\frac{\partial\phi}{\partial x_i}+c(\boldsymbol{x})\phi+d(\boldsymbol{x}).$$
The PDE is called **elliptic** if the eigenvalues of A_{ij} are real and all have the same sign, for all $\boldsymbol{x}\in\Omega$

2. Consider the PDE
$$\frac{\partial\phi}{\partial t}=\sum_{i,j=1}^{n}A_{ij}(\boldsymbol{x},t)\frac{\partial^2\phi}{\partial x_i\partial x_j}+\sum_{i=1}^{n}b_i(\boldsymbol{x},t)\frac{\partial\phi}{\partial x_i}+c(\boldsymbol{x},t)\phi+d(\boldsymbol{x},t).$$
The PDE is called **parabolic** if the eigenvalues of A_{ij} are all real and positive, for all t and for all $\boldsymbol{x}\in\Omega$.

3. Consider the PDE
$$\frac{\partial^2\phi}{\partial t^2}=\sum_{i,j=1}^{n}A_{ij}(\boldsymbol{x},t)\frac{\partial^2\phi}{\partial x_i x_j}+\sum_{i=1}^{n}b_i(\boldsymbol{x},t)\frac{\partial\phi}{\partial x_i}+c(\boldsymbol{x},t)\phi+d(\boldsymbol{x},t).$$
The PDE is called **hyperbolic** if the eigenvalues of A_{ij} are all real and positive, for all t and for all $\boldsymbol{x}\in\Omega$.

To see why these classifications are important, consider the following **anisotropic diffusion equation**:

$$\frac{\partial u}{\partial t} = \sum_{i,j=1}^{2} A_{ij} \frac{\partial^2 u}{\partial x_i \partial x_j} \tag{10.1}$$

where $(x_1, x_2) = (x, y) \in \mathbb{R}^2$ and A_{ij} is a constant, **symmetric** matrix. We have the following theorem:

Theorem 10.1 *Eq.* (10.1) *has a unique solution.*

To prove this, consider two solutions u_1 and u_2. Then form the difference $\phi = u_1 - u_2$. As an initial condition, the function ϕ is zero everywhere. Similarly, the boundary conditions on ϕ are zero: $\phi = 0$ on $\partial\Omega$. By linearity, ϕ satisfies the anisotropic diffusion equation (10.1):

$$\frac{\partial \phi}{\partial t} = \sum_{i,j=1}^{2} A_{ij} \frac{\partial^2 \phi}{\partial x_i \partial x_j}.$$

Now multiply by ϕ and integrate over the domain Ω:

$$\begin{aligned}
\phi \frac{\partial \phi}{\partial t} &= \sum_{i,j=1}^{2} A_{ij} \phi \frac{\partial^2 \phi}{\partial x_i \partial x_j}, \\
\int_\Omega \phi \frac{\partial \phi}{\partial t} \mathrm{d}x\, \mathrm{d}y &= \int_\Omega \sum_{i,j=1}^{2} A_{ij} \phi \frac{\partial^2 \phi}{\partial x_i \partial x_j} \mathrm{d}x\, \mathrm{d}y, \\
\tfrac{1}{2} \frac{d}{dt} \int_\Omega \phi^2 \, \mathrm{d}x\, \mathrm{d}y &= \int_\Omega \sum_{i,j=1}^{2} A_{ij} \phi \frac{\partial^2 \phi}{\partial x_i \partial x_j} \mathrm{d}x\, \mathrm{d}y, \\
\tfrac{1}{2} \frac{d}{dt} \|\phi\|_2^2 &= \int_\Omega \sum_{i,j=1}^{2} A_{ij} \left[\frac{\partial}{\partial x_i} \phi \frac{\partial \phi}{\partial x_j} - \frac{\partial \phi}{\partial x_i} \frac{\partial \phi}{\partial x_j} \right] \mathrm{d}x\, \mathrm{d}y,
\end{aligned}$$

Use Gauss's theorem in the plane (Fig. 10.1):

$$\begin{aligned}
\int_\Omega \nabla \cdot \boldsymbol{v} \, \mathrm{d}x \mathrm{d}y &= \int_{\partial\Omega} \boldsymbol{u} \cdot \hat{\boldsymbol{n}} \, \mathrm{d}\ell, \\
\int_\Omega \nabla \psi \, \mathrm{d}x \mathrm{d}y &= \int_{\partial\Omega} \psi \hat{\boldsymbol{n}} \, \mathrm{d}\ell, \\
\int_\Omega \partial_i v_j \, \mathrm{d}x \mathrm{d}y &= \int_{\partial\Omega} v_j n_i \, \mathrm{d}\ell.
\end{aligned}$$

where $\hat{\boldsymbol{n}} = (n_1, n_2)$ is the unit vector **normal** to the boundary curve $\partial\Omega$ which encloses the area Ω, and where $\mathrm{d}\ell$ is the line element along the curve. Hence,

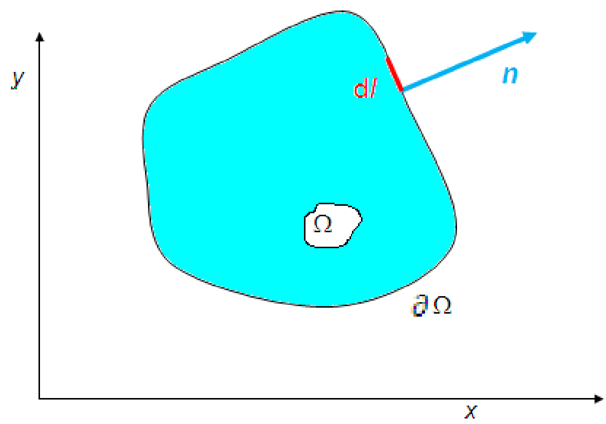


Figure 10.1: Gauss's theorem in the plane

$$\tfrac{1}{2}\frac{d}{dt}\|\phi\|_2^2 = \sum_{i,j=1}^{2} A_{ij}\left(\int_{\partial\Omega}\phi\frac{\partial\phi}{\partial x_j}n_i\,\mathrm{d}\ell - \sum_{i,j=1}^{2}\int_{\Omega}\frac{\partial\phi}{\partial x_i}\frac{\partial\phi}{\partial x_j}\mathrm{d}x\,\mathrm{d}y\right),$$

But $\phi = 0$ on $\partial\Omega$, hence

$$\begin{aligned}\tfrac{1}{2}\frac{d}{dt}\|\phi\|_2^2 &= -\sum_{i,j=1}^{2}\int_{\Omega}\frac{\partial\phi}{\partial x_i}\frac{\partial\phi}{\partial x_j}\mathrm{d}x\,\mathrm{d}y,\\ &:= -\int_{\Omega}(\nabla\phi|A|\nabla\phi)\mathrm{d}x\,\mathrm{d}y,\end{aligned}$$

where

$$(\boldsymbol{x}|A|\boldsymbol{x}) = (x,y)A(x,y)^T = \boldsymbol{x}\cdot(A\boldsymbol{x})$$

for all $\boldsymbol{x} := (x,y)\in\mathbb{R}^2$ is a **quadratic form**. However, we are told that the equation is parabolic, and hence A is a symmetric matrix with positive eigenvalues. That means that the quadratic form is **positive-definite**:

$$\begin{aligned}\boldsymbol{x} &= a_1\boldsymbol{f}_1 + a_2\boldsymbol{f}_2, \qquad A\boldsymbol{f}_{(i)} = \lambda_{(i)}\boldsymbol{f}_{(i)}, \qquad \boldsymbol{f}_{(i)}\cdot\boldsymbol{f}_{(j)} = \delta_{ij},\\ A\boldsymbol{x} &= a_1\lambda_{(1)}\boldsymbol{f}_{(1)} + a_2\lambda_{(2)}\boldsymbol{f}_{(2)},\\ \boldsymbol{x}\cdot(A\boldsymbol{x}) &= (a_1\boldsymbol{f}_1 + a_2\boldsymbol{f}_2)\cdot\left(a_1\lambda_{(1)}\boldsymbol{f}_{(1)} + a_2\lambda_{(2)}\boldsymbol{f}_{(2)}\right),\\ &= \lambda_1 a_1^2 + \lambda_2 a_2^2 \geq 0.\end{aligned}$$

Thus,

$$\begin{aligned}\tfrac{1}{2}\frac{d}{dt}\|\phi\|_2^2 &= -\int_{\Omega}(\nabla\phi|A|\nabla\phi)\mathrm{d}x\,\mathrm{d}y,\\ &\leq 0,\end{aligned}$$

and the L^2-norm of ϕ is decreasing. In particular,

$$0 \le \|\phi\|_2(T) \le \|\phi\|_2(0) = 0, \qquad T \ge 0.$$

hence

$$\|\phi\|_2(T) = 0.$$

For a smooth function, the only way to satisfy this equation is for $\phi = 0$, hence

$$u_1 = u_2.$$

Thus, we conclude that **the parabolic property is essential for obtaining the uniqueness of diffusion-type equations**. Similar conclusions can be drawn about the hyperbolic case.

Chapter 11

The 1-D wave equation

Overview

The wave equation describes linear oscillations in a generic field $u(x,t)$:

$$\frac{1}{c^2}\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2},$$

where c is the propagation speed of the oscillations. Topics: derivation; solution through separation of variables; energy conservation.

11.1 Derivation

Consider N identical particles arrayed in a line, and connected together by identical springs. The equilibrium position of the i^{th} particle is $x_i = i\Delta x$, with $i = 1, 2, \cdots N$, and the departure from equilibrium is small and equal to y_i. The potential energy of such a system is

$$\mathcal{U}(y_1, \cdots, y_N) = \tfrac{1}{2}k \sum_{i=2}^{N-1} (y_{i+1} - y_i)^2 + \text{Boundary terms.}$$

The boundary terms can be taken care of by forcing the displacements $y_1 = y_N = 0$. Thus,

$$\mathcal{U}(y_1, \cdots, y_N) = \tfrac{1}{2}k \sum_{i=1}^{N} (y_{i+1} - y_i)^2 .$$

Hence, at interior points, Newton's law gives

$$\begin{aligned} m\frac{d^2y_i}{dt^2} &= -\frac{\partial \mathcal{U}}{\partial y_i} = -k\left[-(y_{i+1}-y_i)+(y_i-y_{i-1})\right], \\ &= k(y_{i+1}-y_i)-k(y_i-y_{i-1}). \end{aligned}$$

The mass of each oscillator is $m=\rho\Delta x$, where ρ is the constant (linear density) of the system. Thus,

$$\frac{d^2y_i}{dt^2}=\frac{k\Delta x}{\rho}\frac{(y_{i+1}-y_i)-(y_i-y_{i-1})}{(\Delta x)^2}.$$

We identify $T=k\Delta x$ as the tension in the system of springs. Thus,

$$\frac{d^2y_i}{dt^2}=\frac{T}{\rho}\frac{(y_{i+1}-y_i)-(y_i-y_{i-1})}{(\Delta x)^2}.$$

Taking $\Delta x\to 0$ ($N\to\infty$) gives

$$\frac{\partial^2 y(x,t)}{\partial t^2}=\frac{T}{\rho}\frac{\partial^2 y(x,t)}{\partial x^2},$$

where we identify

$$c^2=\frac{T}{\rho}.$$

Now

$$[c]^2=\frac{\text{Force}}{\text{Mass/Length}}=\frac{\text{Mass Length/Time}^2}{\text{Mass/Length}}=\frac{\text{Length}^2}{\text{Time}^2},$$

and c is clearly a velocity: it is the velocity at which a wave of small oscillations propagates along the spring system. A similar treatment of other systems yields the same linear wave equation. For example, for small oscillations in a gas, the linear wave equation is satisfied, with

$$c_{\text{gas}}^2=\frac{\gamma p_0}{\rho},$$

where

- γ is the (nondimensional) ratio of specific heats;
- p_0 is the equilibrium pressure;
- ρ is the mass per unit volume.

In any case, the generic equation we study in this section is

$$\frac{1}{c^2}\frac{\partial^2 u(x,t)}{\partial t^2}=\frac{\partial^2 u(x,t)}{\partial x^2}.$$

11.2 Boundary and initial conditions

The most common kind of boundary conditions is the requirement that the oscillations at the end points of the domain $\Omega = [0, L]$ should be zero:

$$u(x = 0, t > 0) = u(x = L, t > 0) = 0.$$

In the language of Ch. 8, these are the **homogeneous Dirichlet conditions**. We need **two** boundary conditions because the equation is second-order in space. For the diffusion equation, we needed only one initial condition, because the equation was first-order in time. However, the wave equation is second-order in time, so we need **two initial conditions**, usually taken to be

$$\begin{aligned} u(x, t = 0) &= f(x), \qquad 0 < x < L, \\ u_t(x, t = 0) &= g(x), \qquad 0 < x < L. \end{aligned}$$

11.3 Separation of variables

Consider a taut string, such as a violin string, that is plucked according to the initial conditions

$$\begin{aligned} u(x, t = 0) &= f(x), \qquad 0 < x < L, \\ u_t(x, t = 0) &= g(x), \qquad 0 < x < L. \end{aligned}$$

The string is fixed at the end points, $u(0) = u(L) = 0$. Solve for the vibrations in the string.

We solve by separation of variables:

$$u(x, t) = X(x)T(t).$$

Substitution into the wave equation gives

$$\frac{1}{c^2} T''(t) X(x) = X''(x) T(t).$$

Dividing by $X(x)T(t)$ gives

$$\frac{1}{c^2} \frac{T''(t)}{T(t)} = \frac{X''(x)}{X(x)}.$$

Since the LHS is a function of time alone and the RHS is a function of space alone, the only way for this equation to be satisfied is if both sides are in fact equal to a constant:

$$\frac{1}{c^2} \frac{T''(t)}{T(t)} = \frac{X''(x)}{X(x)} = -\lambda.$$

Let us also substitute the trial solution into the BCs and the ICs:

$$
\begin{aligned}
\text{Initial condition:} &\quad u(x,t=0) = X(x)T(0) = f(x), \qquad 0 < x < L, \\
\text{Initial condition:} &\quad u_t(x,t=0) = X(x)T'(0) = g(x), \qquad 0 < x < L, \\
\text{Boundary condition:} &\quad T(t)X(0) = T(t)X(L) = 0
\end{aligned}
$$

Solving for $X(x)$

Focussing on the $X(x)$-equations, we have:

$$
\begin{aligned}
\frac{1}{X}\frac{d^2X}{dx^2} &= -\lambda, \qquad 0 < x < L, \\
X(0) &= X(L) = 0.
\end{aligned}
$$

Equation in the **bulk** $0 < x < L$:

$$\frac{d^2X}{dx^2} + \lambda X = 0, \qquad (*)$$

Different possibilities for λ:

1. $\lambda = 0$. Then, the solution is $X(x) = Ax + B$, with $dX/dx = A$. However, the BCs specify $X(0) = 0$, hence $B = 0$. They also specify $X(L) = 0$, hence $A = 0$. Thus, only the trivial solution remains, in which we have no interest.

2. $\lambda < 0$. Then, the solution is $X(x) = Ae^{\mu x} + Be^{-\mu x}$, where $\mu = \sqrt{-\lambda}$. The BCs give

   $$A + B = Ae^{\mu L} + Be^{-\mu L} = 0.$$

   Grouping the first two of these equations together gives

   $$A = -B\frac{1 - e^{-\mu L}}{1 - e^{\mu L}}.$$

   But $A + B = 0$, hence

   $$
   \begin{aligned}
   B\left[1 - \frac{1 - e^{-\mu L}}{1 - e^{\mu L}}\right] &= 0, \\
   B\left[\frac{1 - e^{\mu L} - (1 - e^{-\mu L})}{1 - e^{\mu L}}\right] &= 0, \\
   B\left[-e^{\mu L} + e^{-\mu L}\right] &= 0, \\
   B\sinh(\mu L) &= 0,
   \end{aligned}
   $$

which has only the trivial solution.

3. Thus, we are forced into the third option: $\lambda > 0$.

Solving Eq. (*) with $\lambda > 0$ gives

$$X(x) = A\cos(\sqrt{\lambda}x) + B\sin(\sqrt{\lambda}x),$$

with boundary condition

$$A \cdot 1 + B \cdot 0 = A\cos(\sqrt{\lambda}L) + B\sin(\sqrt{\lambda}L) = 0.$$

Hence, $A = 0$. Grouping the second and third equations in this string together therefore gives

$$B\sin(\sqrt{\lambda}L) = 0.$$

Of course, $B = 0$ is a solution, but this is the trivial one. Therefore, we must try to solve

$$\sin(\sqrt{\lambda}L) = 0.$$

This is possible, provided

$$\sqrt{\lambda}L = n\pi, \qquad n \in \{1, 2, \cdots\}.$$

Thus,

$$\lambda = \lambda_n = \frac{n^2\pi^2}{L^2},$$

and

$$X(x) = B\sin\left(\frac{n\pi x}{L}\right),$$

where B is a constant of integration.

Solving for $T(t)$

Now substitute $\lambda_n = n^2\pi^2/L^2$ back into the $T(t)$-equation:

$$\frac{1}{T}\frac{dT^2}{dt^2} = -\lambda c^2 = -\lambda_n c^2.$$

Solving give

$$T(t) = C\cos(c\sqrt{\lambda_n}t) + D\sin(c\sqrt{\lambda_n}t).$$

Putting it all together

Recall the ansatz:

$$u(x,t) = X(x)T(t).$$

Thus, we have a solution

$$X(x)T(t) = B\sin\left(\frac{n\pi x}{L}\right)\left[C\cos(c\sqrt{\lambda_n}t) + D\sin(c\sqrt{\lambda_n}t)\right].$$

Re-labelling the constants, this is

$$X_n(x)T_n(t) = \sin\left(\frac{n\pi x}{L}\right)\left[A_n\cos(c\sqrt{\lambda_n}t) + B_n\sin(c\sqrt{\lambda_n}t)\right].$$

The label n is just a label on the solution. However, each $n = 1, 2, \cdots$ produces a different solution, linearly independent of all the others. We can add all of these solutions together to obtain a **general solution** of the PDE:

$$\begin{aligned} u(x,t) &= \sum_{n=1}^{\infty} X_n(x)T_n(t), \\ &= \sum_{n=1}^{\infty} \sin\left(\frac{n\pi x}{L}\right)\left[A_n\cos(c\sqrt{\lambda_n}t) + B_n\sin(c\sqrt{\lambda_n}t)\right]. \end{aligned}$$

We are almost there. However, we still need to take care of the initial conditions. First IC:

$$\begin{aligned} u(x,t=0) &= \sum_{n=1}^{\infty} \sin\left(\frac{n\pi x}{L}\right)\left[A_n\cos(c\sqrt{\lambda_n}t) + B_n\sin(c\sqrt{\lambda_n}t)\right]_{t=0}, \\ &= \sum_{n=1}^{\infty} A_n \sin\left(\frac{n\pi x}{L}\right), \\ &= f(x). \end{aligned}$$

But the functions

$$\left\{\sin\left(\frac{n\pi x}{L}\right)\right\}_{n=1}^{\infty}$$

are orthogonal on $[0, L]$:

$$\int_0^L \sin\left(\frac{n\pi x}{L}\right)\sin\left(\frac{m\pi x}{L}\right)\mathrm{d}x = \frac{L}{2}\delta_{mn}.$$

Thus, multiply both sides by $\sin(m\pi x/L)$ and integrate:

$$\begin{aligned}
\int_0^\pi f(x)\sin\left(\frac{m\pi x}{L}\right)\mathrm{d}x &= \int_0^\pi \sum_{n=1}^\infty A_n \sin\left(\frac{m\pi x}{L}\right)\sin\left(\frac{n\pi x}{L}\right), \\
&= \sum_{n=1}^\infty A_n \int_0^\pi \sin\left(\frac{m\pi x}{L}\right)\sin\left(\frac{n\pi x}{L}\right), \\
&= \sum_{n=1}^\infty A_n \frac{L}{2}\delta_{m,n}, \\
&= \frac{A_n L}{2}.
\end{aligned}$$

Hence,

$$A_n = \frac{2}{L}\int_0^\pi f(x)\sin\left(\frac{m\pi x}{L}\right)\mathrm{d}x.$$

Second IC:

$$\begin{aligned}
u_t(x,t=0) &= \sum_{n=1}^\infty \sin\left(\frac{n\pi x}{L}\right)\frac{d}{dt}\left\{\left[A_n\cos(c\sqrt{\lambda_n}t)+B_n\sin(c\sqrt{\lambda_n}t)\right]\right\}_{t=0}, \\
&= \sum_{n=1}^\infty B_n c\sqrt{\lambda_n}\sin\left(\frac{n\pi x}{L}\right), \\
&= \sum_{n=1}^\infty \frac{B_n n\pi c}{L}\sin\left(\frac{n\pi x}{L}\right), \\
&= g(x).
\end{aligned}$$

Taking the scalar product with $\sin(m\pi x/L)$, we get

$$\int_0^L g(x)\sin\left(\frac{m\pi x}{L}\right)\mathrm{d}x = \frac{L}{2}\frac{B_n n\pi c}{L},$$

hence

$$B_n = \frac{L}{n\pi c}\frac{2}{L}\int_0^L g(x)\sin\left(\frac{m\pi x}{L}\right)\mathrm{d}x.$$

and, substituting back into the general solution, we have

$$\begin{aligned}
u(x,t) &= \sum_{n=1}^\infty \sin\left(\frac{n\pi x}{L}\right)\left[A_n\cos(c\sqrt{\lambda_n}t)+B_n\sin(c\sqrt{\lambda_n}t)\right], \qquad (11.1)\\
A_n &= \frac{2}{L}\int_0^L \sin\left(\frac{n\pi x}{L}\right)f(x)\,\mathrm{d}x, \\
B_n &= \frac{L}{n\pi c}\frac{2}{L}\int_0^L \sin\left(\frac{n\pi x}{L}\right)g(x)\,\mathrm{d}x.
\end{aligned}$$

which is a solution to the wave equation that satisfies the boundary and initial conditions.

Note: Proving that this series converges to the solution is difficult, because we do not have decaying exponentials like $e^{-n^2\pi^2 Dt/L^2}$ as in the diffusion equation, thus making it difficult to apply the Weirstrass M-test

11.4 Physical interpretation of solution

We have found the following solution to the wave equation:

$$u(x,t) = \sum_{n=1}^{\infty} \sin\left(\frac{n\pi x}{L}\right)\left[A_n \cos\left(n\frac{c\pi}{L}t\right) + B_n \sin\left(n\frac{c\pi}{L}t\right)\right],$$

which vanishes at the boundaries $u(0) = u(L) = 0$.

- The component
$$\sin\left(\frac{n\pi x}{L}\right)\left[A_n \cos\left(n\frac{c\pi}{L}t\right) + B_n \sin\left(n\frac{c\pi}{L}t\right)\right]$$
is called the n^{th} **normal mode of vibration**.
- The solution is a sum over all normal modes.
- Each normal mode is a periodic function of time, with period
$$n\frac{c\pi}{L}\tau_n = 2\pi \implies \tau_n = \frac{2}{n}\frac{L}{c},$$
- The frequency of a normal mode is given by
$$\begin{aligned}\omega_n &= \frac{2\pi}{\tau_n},\\ &= 2\pi\frac{nc}{2L},\\ &= 2\pi\frac{n}{2L}\sqrt{\frac{T}{\rho}},\end{aligned}$$
upon restoration of the original interpretation of the wave speed. This is probably the nicest result of high-school physics: **Modes of vibration of a string are periodic, and each frequency is an integer multiple of a basic or fundamental frequency, given by**
$$\omega_1 = 2\pi\frac{1}{2L}\sqrt{\frac{T}{\rho}},$$

- In a given complex disturbance (i.e. multiple frequencies), each mode is characterised by its frequency ω_n and by the quantities A_n and B_n, which tell us the intensity of the contribution made by the n^{th} normal mode. However, we can re-write the disturbance:

$$\begin{aligned} u(x,t) &= \sum_{n=1}^{\infty} \sin\left(\frac{n\pi x}{L}\right)\left[A_n \cos\left(n\frac{c\pi}{L}t\right) + B_n \sin\left(n\frac{c\pi}{L}t\right)\right], \\ &= \sum_{n=1}^{\infty} \sin\left(\frac{n\pi x}{L}\right) C_n \sin\left(n\frac{c\pi}{L}t + \gamma_n\right). \end{aligned}$$

The quantity

$$C_n^2 = A_n^2 + B_n^2$$

is thus the **amplitude** of the n^{th} normal mode and

$$\gamma_n = \arctan(B_n/A_n)$$

is its **phase**.

11.5 Energy

For the diffusion equation $u_t = Du_{xx}$, either

$$E_1 = \int_\Omega u(x,t)\mathrm{d}x,$$

or

$$E_2 = \tfrac{1}{2}\int_\Omega u^2(x,t)\mathrm{d}x,$$

has the interpretation of energy, depending on the physical context. Both of these are decreasing functions of time, since the general solution is

$$|u(x,t)| = \left|\sum_{n=1}^{\infty} B_n \sin\left(\frac{n\pi x}{L}\right) e^{-n^2\pi^2 Dt/L^2}\right| \le |u(x,0)|.$$

In this section we formulate an energy for the wave equation and show that it is conserved.

To do this, we recall the discrete starting point for the wave equation. We took N identical particles arrayed on a line, connected together by identical springs. The equilibrium position of the i^{th} particle is $x_i = i\Delta x$, and the departure from equilibrium is small and equal to y_i. The potential energy of such a system is

$$\begin{aligned}\mathcal{U}(y_1,\cdots,y_N) &= \tfrac{1}{2}k\sum_{i=2}^{N-1}(y_{i+1}-y_i)^2 + \text{Boundary terms},\\ &= \tfrac{1}{2}k\sum_{i=1}^{N}(y_{i+1}-y_i)^2, \qquad y_0=y_N=0.\end{aligned}$$

At interior points, Newton's law gives

$$m\frac{d^2y_i}{dt^2} = -\frac{\partial\mathcal{U}}{\partial y_i} = k\left[(y_{i+1}-y_i)-(y_i-y_{i-1})\right].$$

This is an equation of the type

$$m\frac{d^2\boldsymbol{y}}{dt^2} = -\nabla_y\mathcal{U}(\boldsymbol{y}).$$

If we take the dot product of this equation with $d\boldsymbol{y}/dt$ we obtain

$$\begin{aligned}m\frac{d\boldsymbol{y}}{dt}\cdot\frac{d^2\boldsymbol{y}}{dt^2} &= -\frac{d\boldsymbol{y}}{dt}\cdot\nabla_y\mathcal{U}(\boldsymbol{y}),\\ m\frac{d}{dt}\left(\frac{d\boldsymbol{y}}{dt}\right)^2 &= -\frac{d}{dt}\mathcal{U}(\boldsymbol{y}),\end{aligned}$$

or

$$\tfrac{1}{2}m\left(\frac{d\boldsymbol{y}}{dt}\right)^2 + \mathcal{U}(\boldsymbol{y}) = E = \text{Const}.$$

In other words,

$$\tfrac{1}{2}m\sum_{i=2}^{N-1}\left(\frac{dy_i}{dt}\right)^2 + \tfrac{1}{2}k\sum_{i=2}^{N-1}(y_{i+1}-y_i)^2 = E = \text{Const}.$$

As before, let $m=\rho\Delta x$ and let $k\Delta x = T = \text{Const.}$. Hence, $k = T/\Delta x$ and

$$\tfrac{1}{2}\rho\sum_{i=2}^{N-1}\Delta x\left(\frac{dy_i}{dt}\right)^2 + \tfrac{1}{2}T\sum_{i=2}^{N-1}\Delta x\frac{(y_{i+1}-y_i)^2}{(\Delta x)^2} = E.$$

Now take $\Delta x\to 0$. The sums become Riemann integrals and the finite differences become derivatives.

$$\tfrac{1}{2}\int_\Omega \mathrm{d}x\,\rho\left(\frac{\partial u}{\partial t}\right)^2 + \tfrac{1}{2}T\int_\Omega \mathrm{d}x\,\left(\frac{\partial u}{\partial x}\right)^2 = E.$$

Thus, our candidate for conserved pseudo-energy is

$$\mathcal{E} := \int_\Omega\left[\frac{1}{c^2}\left(\frac{\partial u}{\partial t}\right)^2 + \left(\frac{\partial u}{\partial x}\right)^2\right]\mathrm{d}x.$$

(I call it a pseudo-energy because strictly speaking, it does not have dimensions of energy.) Now finally, let's double check that the wave equation $c^{-2}\partial_{tt}u = \partial_{xx}$ with the zero BCs conserves the pseudo-energy:

$$
\begin{aligned}
\frac{d\mathcal{E}}{dt} &= \int_\Omega \left[\frac{1}{c^2}\frac{\partial u}{\partial t}\frac{\partial^2 u}{\partial t^2} + \frac{\partial u}{\partial x}\frac{\partial^2 u}{\partial t \partial x}\right], \\
&= \int_\Omega \left[\frac{1}{c^2}\frac{\partial u}{\partial t}\frac{\partial^2 u}{\partial t^2} + \frac{\partial u}{\partial x}\frac{\partial}{\partial x}\frac{\partial u}{\partial t}\right] \mathrm{d}x, \\
&= \int_\Omega \left[\frac{1}{c^2}\frac{\partial u}{\partial t}\frac{\partial^2 u}{\partial t^2} + \frac{\partial}{\partial x}\left(\frac{\partial u}{\partial t}\frac{\partial u}{\partial x}\right) - \frac{\partial u}{\partial t}\frac{\partial^2 u}{\partial x^2}\right] \mathrm{d}x, \\
&= \int_\Omega \left[\frac{1}{c^2}\frac{\partial u}{\partial t}\frac{\partial^2 u}{\partial t^2} - \frac{\partial u}{\partial t}\frac{\partial^2 u}{\partial x^2}\right] \mathrm{d}x + \left(\frac{\partial u}{\partial t}\frac{\partial u}{\partial x}\right)_0^L, \\
&= \int_\Omega \left[\frac{1}{c^2}\frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial x^2}\right]\frac{\partial u}{\partial t}\mathrm{d}x + \underbrace{\frac{\partial u(L,t)}{\partial t}}_{=0}\left(\frac{\partial u}{\partial x}\right)_{x=L} - \underbrace{\frac{\partial u(0,t)}{\partial t}}_{=0}\left(\frac{\partial u}{\partial x}\right)_{x=0}, \\
&= 0 - 0.
\end{aligned}
$$

Chapter 12

Interlude: the chain rule

Overview

We review the chain rule of multivariate calculus.

12.1 The chain rule: Theory

Let

$$\phi : \mathbb{R}^2 \to \mathbb{R}$$

be *a function of two slots*, $\phi = \phi(\cdot,\cdot)$. Furthermore, let

$$\begin{aligned} u : \mathbb{R}^2 &\to \mathbb{R} \\ (x,y) &\to u(x,y), \\ v : \mathbb{R}^2 &\to \mathbb{R} \\ (x,y) &\to v(x,y) \end{aligned}$$

be two-slot functions of the variables x and y. Let us place $u(x,y)$ and $v(x,y)$ into the two slots of $\phi(\cdot,\cdot)$ and investigate how ϕ changes as x and y change. Form the difference

$$\begin{aligned} \phi\left[u(x+\delta x,y), v(x+\delta x,y)\right] - \phi\left[u(x,y),v(x,y)\right] &= \\ \phi\left[u(x,y)+u_x(x,y)\delta x, y, v(x,y)+v_x(x,y)\delta x\right] &- \phi\left[u(x,y),v(x,y)\right], \\ = \left[\frac{\partial \phi}{\partial s_1}\right]_{u(x,y),v(x,y)} u_x(x,y)\delta x &+ \left[\frac{\partial \phi}{\partial s_2}\right]_{u(x,y),v(x,y)} v_x(x,y)\delta x, \end{aligned}$$

where $\partial\phi/\partial s_1$ means the partial derivative of ϕ w.r.t. the first slot, and $\partial\phi/\partial s_2$ means the partial derivative of ϕ w.r.t. the second slot. Now, dividing the difference by δx and taking $\delta x \to 0$, we obtain

$$\left[\frac{\partial\phi}{\partial x}\right]_{u(x,y),v(x,y)} = \left[\frac{\partial\phi}{\partial s_1}\right]_{u(x,y),v(x,y)} u_x(x,y) + \left[\frac{\partial\phi}{\partial s_2}\right]_{u(x,y),v(x,y)} v_x(x,y).$$

Although not technically correct, this is sometimes written as

$$\frac{\partial\phi}{\partial x} = \frac{\partial\phi}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial\phi}{\partial v}\frac{\partial v}{\partial x}.$$

Similarly,

$$\frac{\partial\phi}{\partial y} = \frac{\partial\phi}{\partial u}\frac{\partial u}{\partial y} + \frac{\partial\phi}{\partial v}\frac{\partial v}{\partial y}.$$

This abuse of notation is fine, so long as you remember that you are computing the partial derivatives of ϕ w.r.t. the first and second slots, and these are independent of what goes into these slots (i.e. independent of u and v).

12.2 The chain rule: Examples

1. Let

$$\phi(s_1,s_2) = \tfrac{1}{2}\left(s_1^2+s_2^2\right).$$

We wish to compute

$$\frac{\partial}{\partial x}\phi\left[u(x,y),v(x,y)\right], \qquad \frac{\partial}{\partial y}\phi\left[u(x,y),v(x,y)\right],$$

where

$$u(x,y) = 1+\sin(x), \qquad v(x,y) = xe^y.$$

We write

$$\phi(u,v) = \tfrac{1}{2}\left(u^2+v^2\right)$$

and we use the formula

$$\begin{aligned} \frac{\partial\phi}{\partial x} &= \frac{\partial\phi}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial\phi}{\partial v}\frac{\partial v}{\partial x}, \\ &= u\frac{\partial}{\partial x}\left(1+\sin(x)\right) + v\frac{\partial}{\partial x}\left(xe^y\right), \\ &= (1+\sin(x))\cos(x) + xe^y e^y, \\ &= (1+\sin(x))\cos(x) + xe^{2y}. \end{aligned}$$

Similarly,

$$\begin{aligned}
\frac{\partial \phi}{\partial y} &= \frac{\partial \phi}{\partial u}\frac{\partial u}{\partial y}\frac{\partial \phi}{\partial v}\frac{\partial v}{\partial y}, \\
&= u\frac{\partial}{\partial y}\left(1+\sin(x)\right) + v\frac{\partial}{\partial y}\left(xe^{y}\right), \\
&= (1+\sin(x))0 + xe^{y}(xe^{y}), \\
&= x^2 e^{2y}.
\end{aligned}$$

2. Let $\phi(s_1, s_2) = s_1 s_2$. We wish to compute

$$\frac{\partial}{\partial x}\phi\left[u(x,y), v(x,y)\right], \qquad \frac{\partial^2}{\partial x^2}\phi\left[u(x,y), v(x,y)\right],$$

where

$$u(x,y) = y + \sin(x), \qquad v(x,y) = x\cos(y).$$

We write $\phi(u,v) = uv$, and we use the formula

$$\begin{aligned}
\frac{\partial \phi}{\partial x} &= \frac{\partial \phi}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial \phi}{\partial v}\frac{\partial v}{\partial x}, \\
&= v\frac{\partial}{\partial x}\left(y+\sin(x)\right) + u\frac{\partial}{\partial x}\left(x\cos(y)\right), \\
&= x\cos(y)\cos(x) + [y+\sin(x)]\cos(y), \\
&= x\cos(x)\cos(y) + \sin(x)\cos(y) + y\cos(y).
\end{aligned}$$

Also, by straightforward calculation,

$$\frac{\partial^2 \phi}{\partial x^2} = \cos(y)\left[-x\sin(x)+\cos(x)\right] + \cos(x)\cos(y) = -x\sin(x)\cos(y) + 2\cos(x)\cos(y).$$

Chapter 13

The 1-D wave equation: d'Alembert's solution

Overview

The d'Alembert solution to the wave equation $c^{-2}u_{tt} = u_{xx}$ is a function of the form

$$P(x + ct) + Q(x - ct),$$

where $P(\cdot)$ and $Q(\cdot)$ are functions of a single slot. In this chapter we explain how such a solution arises.

13.1 Motivation

We have found the following solution to the wave equation:

$$u(x,t) = \sum_{n=1}^{\infty} \sin\left(\frac{n\pi x}{L}\right)\left[A_n \cos\left(n\frac{c\pi}{L}t\right) + B_n \sin\left(n\frac{c\pi}{L}t\right)\right],$$

which vanishes at the boundaries $u(0) = u(L) = 0$. We call the component

$$u_n(x,t) := \sin\left(\frac{n\pi x}{L}\right)\left[A_n \cos\left(n\frac{c\pi}{L}t\right) + B_n \sin\left(n\frac{c\pi}{L}t\right)\right]$$

is called the n^{th} **normal mode of vibration**. For ease of exposition, this can be re-written as

$$u_n(x,t) = A_n \sin(k_n x)\cos(\omega_n t) + B_n \sin(k_n x)\sin(\omega_n t),$$

with $\omega = nc\pi/L$ and $k_n = n\pi/L$. Let's apply the following sum to product identities:

$$\tfrac{1}{2}\left[\sin(a+b)+\sin(a-b)\right] = \sin a\cos b,$$
$$\tfrac{1}{2}\left[\cos(a+b)+\cos(a-b)\right] = \sin a\sin b.$$

Thus, the n^{th} normal mode becomes

$$u_n(x,t) = \tfrac{1}{2}A_n\left[\sin(k_nx+\omega_nt)+\sin(k_nx-\omega_nt)\right] + \tfrac{1}{2}B_n\left[\cos(k_nx+\omega_nt)+\cos(k_nx-\omega_nt)\right].$$

But

$$c = \omega_n/k_n$$

independent of n, hence $\omega_n = ck_n$, and

$$u_n(x,t) = \tfrac{1}{2}A_n\left[\sin(k_n(x+ct))+\sin(k_n(x-ct)\right] + \tfrac{1}{2}B_n\left[\cos(k_n(x+ct))+\cos(k_n(x-ct))\right].$$

This is a linear combination of functions that depend on $x-ct$ and $x+ct$. The $x-ct$ contribution corresponds to a wave propagating to the **right**; the other contribution is a wave propagating to the **left** (note the signs!!).

13.2 Change of variables; physical interpretation

From now on, we assume that we are working with waves on the whole of the real line, and that boundary conditions are unimportant, provided the energy is finite. Our result for the single-mode case, namely that the solution is a linear combination of single-variable functions in the variables $x-ct$ and $x+ct$, suggests that we introduce new variables

$$\begin{aligned}\eta &= x-ct,\\ \xi &= x+ct,\end{aligned}$$

For the solution $u(x,t)$ to the wave equation,

$$\begin{aligned}\frac{\partial u}{\partial x} &= \frac{\partial u}{\partial \eta}\frac{\partial \eta}{\partial x}+\frac{\partial u}{\partial \xi}\frac{\partial \xi}{\partial x},\\ &= \frac{\partial u}{\partial \eta}+\frac{\partial u}{\partial \xi},\end{aligned}$$

and

$$\begin{aligned}
\frac{\partial^2 u}{\partial x^2} &= \left[\frac{\partial}{\partial \eta}\left(\frac{\partial u}{\partial \eta}\right)\frac{\partial \eta}{\partial x} + \frac{\partial}{\partial \xi}\left(\frac{\partial u}{\partial \eta}\right)\frac{\partial \xi}{\partial x}\right] + \left[\frac{\partial}{\partial \eta}\left(\frac{\partial u}{\partial \xi}\right)\frac{\partial \eta}{\partial x} + \frac{\partial}{\partial \xi}\left(\frac{\partial u}{\partial \xi}\right)\frac{\partial \xi}{\partial x}\right], \\
&= \left[\frac{\partial}{\partial \eta}\left(\frac{\partial u}{\partial \eta}\right) + \frac{\partial}{\partial \xi}\left(\frac{\partial u}{\partial \eta}\right)\right] + \left[\frac{\partial}{\partial \eta}\left(\frac{\partial u}{\partial \xi}\right) + \frac{\partial}{\partial \xi}\left(\frac{\partial u}{\partial \xi}\right)\right], \\
&= \frac{\partial^2 u}{\partial \eta^2} + \frac{\partial^2 u}{\partial \xi \partial \eta} + \frac{\partial^2 u}{\partial \eta \partial \xi} + \frac{\partial^2 u}{\partial \xi^2}. \qquad (*)
\end{aligned}$$

Similarly,

$$\begin{aligned}
\frac{\partial u}{\partial t} &= \frac{\partial u}{\partial \eta}\frac{\partial \eta}{\partial t} + \frac{\partial u}{\partial \xi}\frac{\partial \xi}{\partial t}, \\
&= -c\frac{\partial u}{\partial \eta} + c\frac{\partial u}{\partial \xi},
\end{aligned}$$

and

$$\begin{aligned}
\frac{\partial^2 u}{\partial t^2} &= -c\left[\frac{\partial}{\partial \eta}\left(\frac{\partial u}{\partial \eta}\right)\frac{\partial \eta}{\partial t} + \frac{\partial}{\partial \xi}\left(\frac{\partial u}{\partial \eta}\right)\frac{\partial \xi}{\partial t}\right] + c\left[\frac{\partial}{\partial \eta}\left(\frac{\partial u}{\partial \xi}\right)\frac{\partial \eta}{\partial t} + \frac{\partial}{\partial \xi}\left(\frac{\partial u}{\partial \xi}\right)\frac{\partial \xi}{\partial t}\right], \\
&= -c\left[-c\frac{\partial}{\partial \eta}\left(\frac{\partial u}{\partial \eta}\right) + c\frac{\partial}{\partial \xi}\left(\frac{\partial u}{\partial \eta}\right)\right] + c\left[-c\frac{\partial}{\partial \eta}\left(\frac{\partial u}{\partial \xi}\right) + c\frac{\partial}{\partial \xi}\left(\frac{\partial u}{\partial \xi}\right)\right], \\
&= c^2\frac{\partial^2 u}{\partial \eta^2} - c^2\frac{\partial^2 u}{\partial \xi \partial \eta} - c^2\frac{\partial^2 u}{\partial \eta \partial \xi} + c^2\frac{\partial^2 u}{\partial \xi^2}. \qquad (**)
\end{aligned}$$

Putting it all together, i.e. taking Eqs. (*) and (**),

$$\begin{aligned}
\frac{\partial u}{\partial x^2} &= \frac{\partial^2 u}{\partial \eta^2} + 2\frac{\partial^2 u}{\partial \xi \partial \eta} + \frac{\partial^2 u}{\partial \xi^2}, \\
\frac{1}{c^2}\frac{\partial u}{\partial t^2} &= \frac{\partial^2 u}{\partial \eta^2} - 2\frac{\partial^2 u}{\partial \xi \partial \eta} + \frac{\partial^2 u}{\partial \xi^2}.
\end{aligned}$$

Using the wave equation $c^{-2}u_{tt} = u_{xx}$, this is

$$\frac{\partial^2 u}{\partial \xi \partial \eta} = 0.$$

In other words,

$$\frac{\partial}{\partial \xi}\left(\frac{\partial u}{\partial \eta}\right) = 0.$$

This means that $\partial u/\partial\eta$ is a function of η alone. Doing the integration gives

$$\frac{\partial u}{\partial \eta} = p(\eta),$$

Integrating again gives

$$\begin{aligned} u(\eta,\xi) &= \int^{\eta} p(s)\mathrm{d}s + \text{function of } \xi \text{ alone}, \\ &= P(\eta) + Q(\xi). \end{aligned}$$

Restoring the $x - t$ notation,

$$u(x,t) = P(x-ct) + Q(x+ct).$$

Physical interpretation

- The component $Q(x-ct)$ corresponds to a **right-travelling wave**. A disturbance that is initially localised at $x = 0$ will have propagated a distance $x = ct$ to the right after a time t.
- The component $Q(x+ct)$ corresponds to a **left-travelling wave**. A disturbance that is initially localised at $x = 0$ will have propagated a distance $x = -ct$ to the right after a time t.
- The lines $x \pm ct$ = Const in $x - t$ space are called **characteristics**, and information is propagated along these lines.

We will make these ideas more precise in the following sections.

13.3 Putting it all together: d'Alembert's formula

Recall the wave equation to solve:

$$c^{-2}u_{tt} = u_{xx},$$

$$\begin{aligned} u(x,t=0) &= f(x), \\ u_t(x,t=0) &= g(x). \end{aligned}$$

We know the solution is of the form

$$u(x,t) = P(x-ct) + Q(x+ct).$$

We now fix $P(\cdot)$ and $Q(\cdot)$ i.t.o. the initial conditions. We have,

$$\begin{aligned} u(x,t=0) &= P(x) + Q(x) = f(x), \\ u_t(x,t=0) &= -cP'(x) + cQ'(x) = g(x). \end{aligned}$$

Differentiating the first equation and multiplying by c gives

$$\begin{aligned} cP'(x) + cQ'(x) &= cf'(x), \qquad (*) \\ -cP'(x) + cQ'(x) &= g(x). \qquad (**) \end{aligned}$$

Adding (*)+(**) gives

$$2\left[Q(x) - Q(0)\right] = f(x) - f(0) + \frac{1}{c}\int_0^x g(s)\,\mathrm{d}s.$$

or

$$Q(x) = \tfrac{1}{2}f(x) + \frac{1}{2c}\int_0^x g(s)\mathrm{d}s + \left[Q(0) - \tfrac{1}{2}f(0)\right].$$

But this relation holds for all values of the real number x in the domain. Thus, we may replace x with $x+ct$ in this formula:

$$Q(x+ct) = \tfrac{1}{2}f(x+ct) + \frac{1}{2c}\int_0^{x+ct} g(s)\mathrm{d}s + \left[Q(0) - \tfrac{1}{2}f(0)\right].$$

Similarly, subtracting (*)-(**) gives

$$2\left[P(x) - P(0)\right] = f(x) - f(0) - \frac{1}{c}\int_0^x g(s)\,\mathrm{d}s.$$

or

$$P(x) = \tfrac{1}{2}f(x) - \frac{1}{2c}\int_0^x g(s)\mathrm{d}s + \left[P(0) - \tfrac{1}{2}f(0)\right].$$

Again, we replace x with $x-ct$ to give

$$P(x-ct) = \tfrac{1}{2}f(x-ct) - \frac{1}{2c}\int_0^{x-ct} g(s)\mathrm{d}s + \left[P(0) - \tfrac{1}{2}f(0)\right].$$

Add these two results:

$$
\begin{aligned}
u(x,t) &= P(x-ct)+Q(x+ct),\\
&= \tfrac{1}{2}f(x-ct)-\frac{1}{2c}\int_0^{x-ct} g(s)\mathrm{d}s+\left[P(0)-\tfrac{1}{2}f(0)\right]\\
&\quad +\tfrac{1}{2}f(x+ct)+\frac{1}{2c}\int_0^{x+ct} g(s)\mathrm{d}s+\left[Q(0)-\tfrac{1}{2}f(0)\right],\\
&= \tfrac{f(x-ct)+f(x+ct)}{2}+\frac{1}{2c}\int_{x-ct}^{x+ct} g(s)\mathrm{d}s+\left[P(0)+Q(0)-f(0)\right],
\end{aligned}
$$

and the term in the square brackets is identically zero, leaving

$$u(x,t)=\frac{f(x-ct)+f(x+ct)}{2}+\frac{1}{2c}\int_{x-ct}^{x+ct} g(s)\mathrm{d}s.$$

This is **d'Alembert's formula** for the solution of the one-dimensional wave equation on the line.

Physical interpretation

Suppose we are at a location (x, ct) in the spacetime plane. To find the solution of the wave equation, we need to know f and $\int g(s)\mathrm{d}s$ in the range

$$x-ct \le x_0 \le x+ct.$$

- In this context, 'information' means initial conditions.
- To know $u(x,t)$ we need only a finite amount of information.
- The amount of information required can be obtained by tracing two lines from the point (x, ct) to the points $(x-ct, 0)$ and $(x+ct, 0)$ in x-ct space.
- These lines have slopes of unity and enclose a triangular region.
- In ordinary x-t space, these lines have slope $\pm c$ – they are characteristics.
- The range $x-ct \le x_0 \le x+ct$ is called the **domain of dependence** (Fig. 13.1).

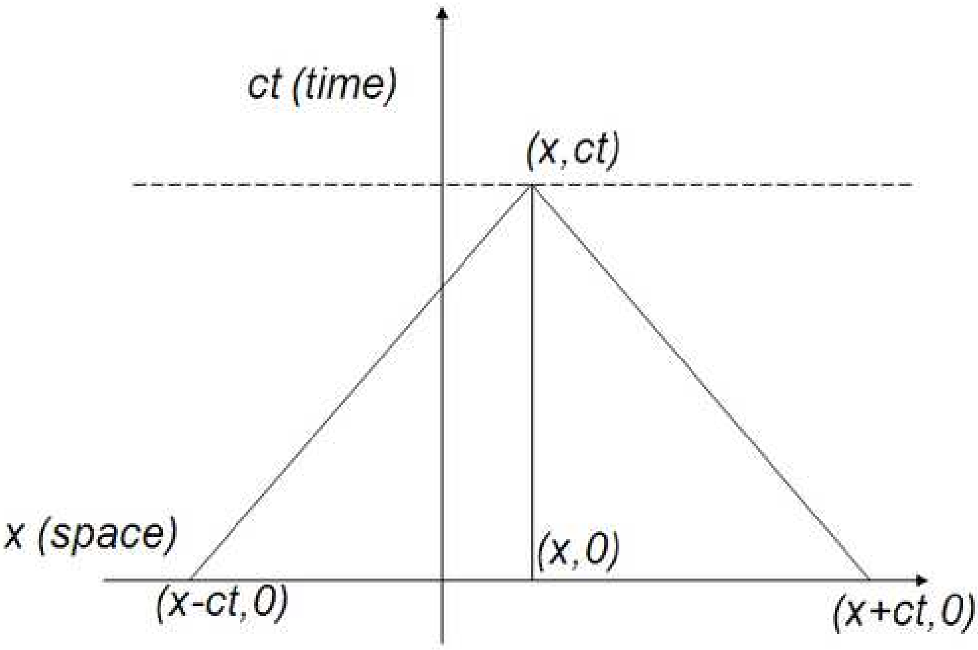


Figure 13.1: The domain of dependence for the wave equation

Chapter 14

The 1-D wave equation: Causality

Overview

In this section we show, by examples, that information is propagated at speed c in the wave equation. Here, 'information' means initial data; we will see shortly what is meant by propagation.

14.1 Example I

Consider the wave equation

$$u_{tt} = u_{xx}$$

with initial data

$$\begin{aligned} u(x, t=0) &= f(x) = \begin{cases} F(x), & |x| \leq 1, \\ 0, & |x| > 1 \end{cases}, \\ u_t(x, t=0) &= g(x) = 0. \end{aligned}$$

and with wave speed $c = 1$. We are to solve for the wave. By d'Alembert's formula, the solution is

$$u(x,t) = \tfrac{1}{2}\left[f(x+t) + f(x-t)\right].$$

We need to identify where $|x+t|$ and $|x-t|$ are less than one; outside of these regions the solution is zero.

- Case 1: $|x+t| \leq 1$ AND $|x-t| \leq 1$. Along the lines where the **equalities** hold, $|x+t| = 1$

and $|x-t|=1$. These lines represent the boundaries of the region of interest:

$$-1 \le x-t \le 1, \qquad -1 \le x+t \le 1.$$

Note also that $dx/dt = \pm 1 = \pm c$ along these lines, i.e. they are **characteristic lines** that give a trajectory moving at the wave speed.

We pick out the pertinent boundary lines:

$$t \le x+1, \qquad t \le 1-x.$$

These are lines with slopes ± 1 and a $y-axis$ intercept at 1. The region R_1 is below these lines, and above the x-axis ($t=0$).

- Case 2: $|x-t| \le 1$ only. In other words, $-1 \le x-t \le 1$. The boundaries of this region are

$$t \le x+1,$$

$$t \ge x-1.$$

  These are characteristics.

- Case 3: $|x+t| \le 1$ only. In other words, $-1 \le x+t \le 1$. The boundaries of this region are

$$t \le 1-x$$

$$t \ge -1-x.$$

Next, we plot these different regions in spacetime (Fig. 14.1).

- In region R_1, $|x+t| \le 1$ AND $|x-t| \le 1$;
- In region R_2, $|x-t| \le 1$ only;
- In region R_3, $|x+t| \le 1$ only;
- Outside of these regions, $|x+t|$ AND $|x-t|$ both exceed 1 (>1).

Thus,

$$u(x,t) = \begin{cases} \frac{1}{2}\left[F(x+t)+F(x-t)\right], & (x,t) \in R_1, \\ \frac{1}{2}F(x-t), & (x,t) \in R_2, \\ \frac{1}{2}F(x+t), & (x,t) \in R_3, \\ 0, & \text{otherwise.} \end{cases}$$

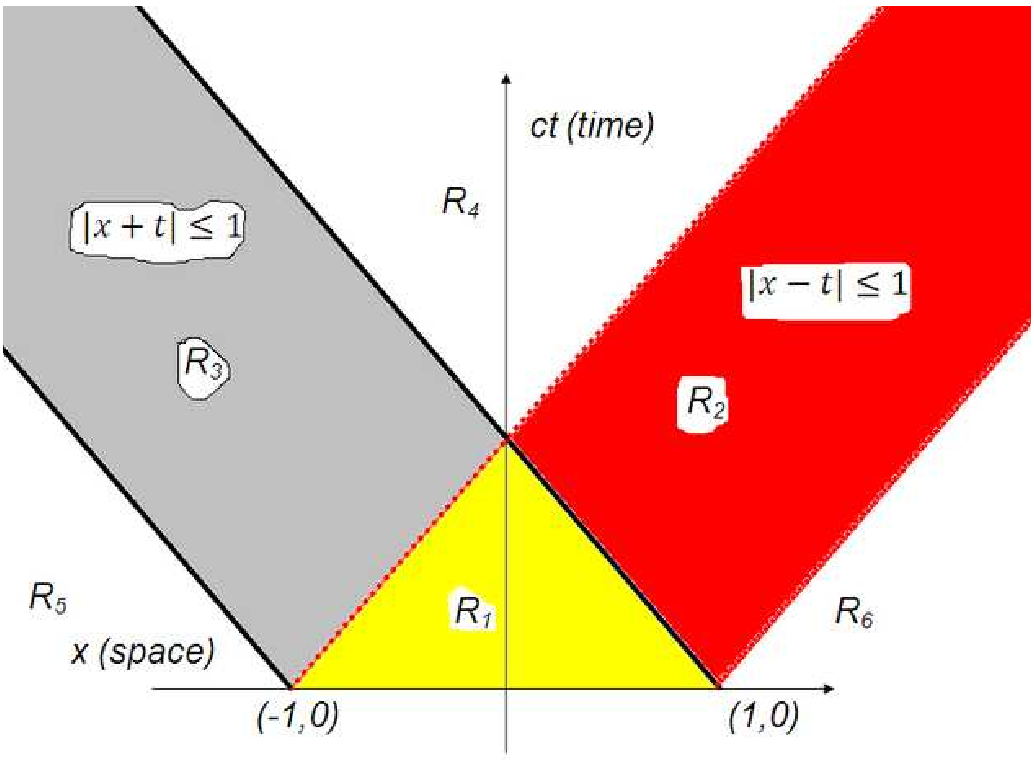


Figure 14.1: The different regions where $|x \pm t| \leq 1$.

Physical interpretation

- The initial, compactly-supported disturbance remains compactly supported for all time. The support never exceeds $x = 1 + ct$ and $x = -1 - ct$, which are characteristics $dx/dt = \pm c$.
- In other words, the equations $x = x_{0R} + ct = 1 + t$ and $x = x_{0L} - ct = -1 - t$ are an envelope within which information is carried forwards in time.
- Outside of this envelope, no information is carried forwards.
- This is the notion of **causality:** The initial solution affects the solution at a later time, within the boundaries set by the characteristics $x = x_{0R} + ct$ and $x = x_{0L} - ct$.
- In other words, causality demands that a compactly-supported initial condition always remain compactly supported, and that support should depend on the initial conditions and the characteristics.
- As can be seen from Fig. 14.1, $F(x - ct)$ represents a right-travelling disturbance, because the domain $|x - ct| \leq 1$ extends into the right half of the spacetime plane.

14.2 Example II

Consider the wave equation

$$u_{tt} = u_{xx}$$

with initial data

$$\begin{aligned} u(x,t=0) &= f(x) = 0, \\ u_t(x,t=0) &= g(x) = \begin{cases} G(x) := \cos^2(\pi x/2), & |x| \le 1, \\ 0, & |x| > 1 \end{cases}, \end{aligned}$$

and with wave speed $c = 1$. We are to solve for the wave. By d'Alembert's formula, the solution is

$$u(x,t) = \tfrac{1}{2}\left[f(x+t) + f(x-t)\right] + \frac{1}{2}\int_{x-t}^{x+t} g(s)\,\mathrm{d}s.$$

We need to identify where $|x+t|$ and $|x-t|$ are less than one; outside of these regions the solution is zero. But we have already done this:

1. In region R_1, $|x+t| \le 1$ AND $|x-ct| \le 1$;

2. In region R_2, $|x-t| \le 1$ only;

3. In region R_3, $|x+t| \le 1$ only;

4. In region R_4, $x-t \le -1$ and $x+t \ge 1$;

5. In region R_5, $x+t \le -1$;

6. In region R_6, $x-t \ge 1$.

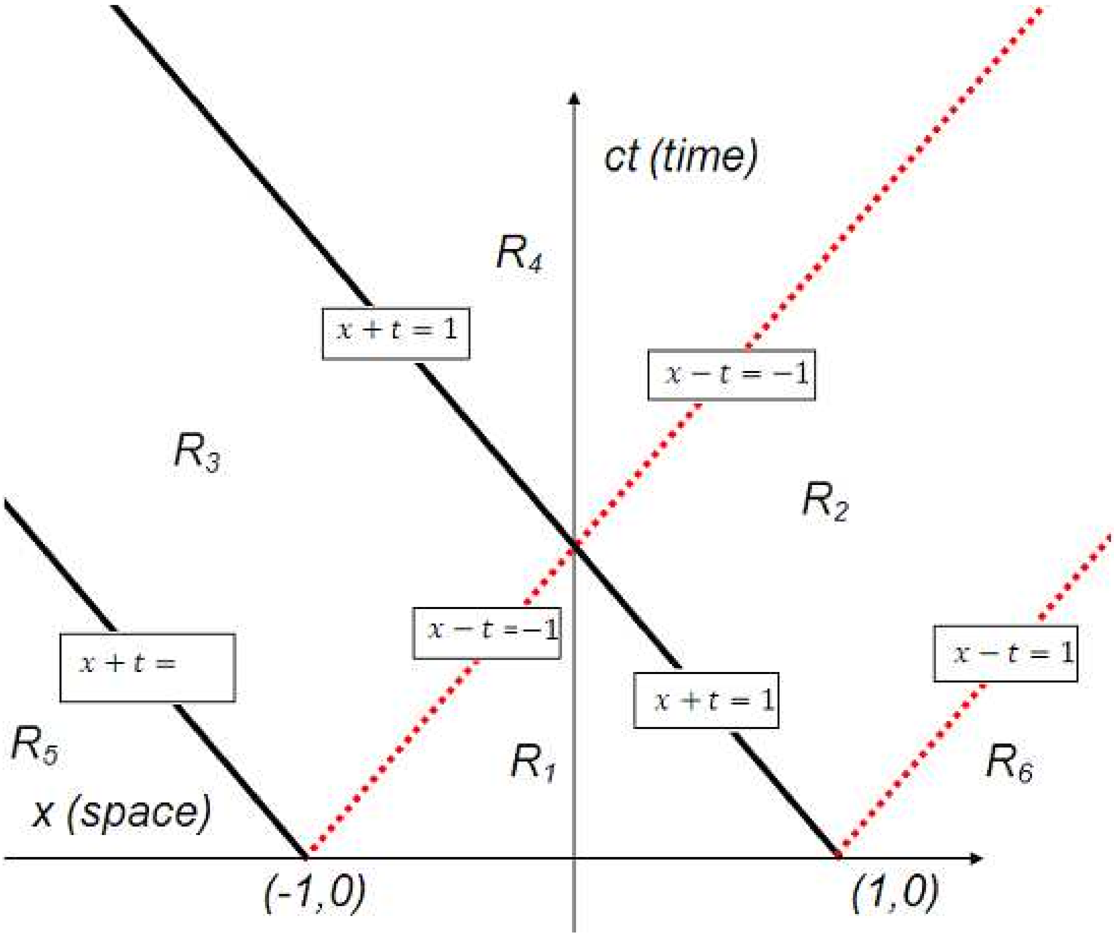
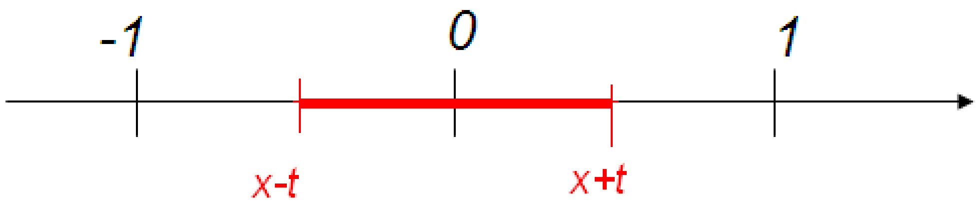


Figure 14.2: The different regions where $|x \pm t| \leq 1$.



Region 1: $|x + t| \le 1$ and $|x - t| \le 1$. This implies that

$$-1 \le x - t \le x + t \le 1.$$

Do the G-integral. Note:

$$\int_a^b \cos^2(\pi x/2) \mathrm{d}x = \int_a^b \tfrac{1}{2}\left[1 + \cos(\pi x)\right] \mathrm{d}x = \tfrac{1}{2}(b-a) + \tfrac{1}{2}\frac{\sin(\pi b) - \sin(\pi a)}{\pi}.$$

Inside region 1,

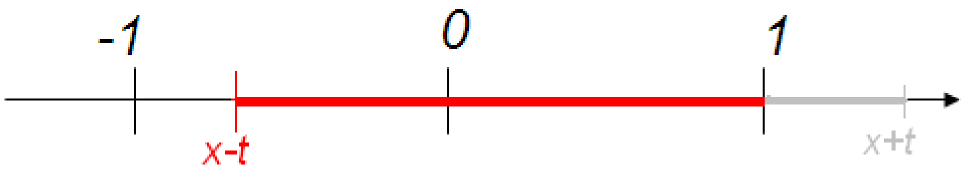
$$-1 \le x - t \le x + t \le 1,$$

so $G(s) = g(s)$ everywhere in the integral:

$$\begin{aligned}
\int_{x-t}^{x+t} G(s)\,\mathrm{d}s &= \int_{x-t}^{x+t} \cos^2(\pi s/2)\,\mathrm{d}s, \\
&= t + \frac{\sin\left(\pi(x+t)\right) - \sin\left(\pi(x-t)\right)}{2\pi}, \\
&= t + \frac{1}{\pi}\cos(\pi x)\sin(\pi t).
\end{aligned}$$

Finally, in region 1,

$$u_1(x,t) = 0 + \tfrac{1}{2}t + \frac{1}{2\pi}\cos(\pi x)\sin(\pi t).$$



Region 2: $|x - t| \leq 1$. Inspection of Fig. 14.2 shows that the region boundaries are

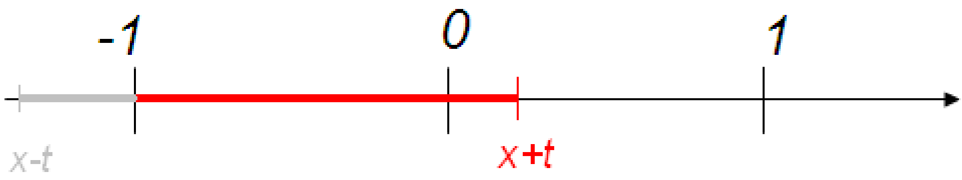
$$-1 \leq x - t \leq 1 \text{ AND } x + t > 1,$$

so that $x - t \geq -1$ and $x + t \geq 1$. Now $G(s) = g(s)$ for $s \in [x - t, 1]$ and is zero elsewhere. Hence,

$$\begin{aligned}
\int_{x-t}^{x+t} G(s)\,\mathrm{d}s &= \int_{x-t}^{1} \cos^2(\pi s/2)\,\mathrm{d}s, \\
&= \tfrac{1}{2}\left[1 - (x - t)\right] + \frac{\sin(\pi) - \sin(\pi(x - t))}{2\pi}, \\
&= \tfrac{1}{2}\left[1 - x + t\right] - \frac{\sin(\pi(x - t))}{2\pi}.
\end{aligned}$$

Finally, in region 2,

$$u_2(x, t) = \tfrac{1}{4}\left[1 - x + t\right] - \frac{\sin(\pi(x - t))}{4\pi}.$$



Region 3: $|x+t| \leq 1$. Inspection of Fig. 14.2 shows that the region boundaries are

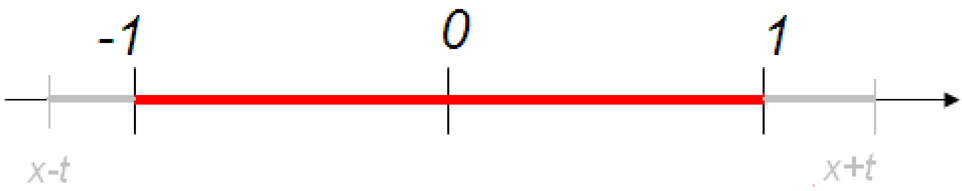
$$-1 \leq x+t \leq 1 \text{ AND } x-t < -1,$$

so that $x-t \leq -1$ and $x+t \leq 1$. Now $G(s) = g(s)$ for $s \in [-1, x+t]$ and is zero elsewhere. Hence,

$$\begin{aligned}
\int_{x-t}^{x+t} G(s)\,\mathrm{d}s &= \int_{-1}^{x+t} \cos^2(\pi s/2)\,\mathrm{d}s, \\
&= \tfrac{1}{2}\left[x+t-(-1)\right] + \frac{\sin\left(\pi(x+t)\right) - \sin\left(\pi(-1)\right)}{2\pi}, \\
&= \tfrac{1}{2}\left[x+t+1\right] + \frac{\sin\left(\pi(x+t)\right)}{2\pi}.
\end{aligned}$$

Finally, in region 3,

$$u_2(x,t) = \tfrac{1}{4}\left[x+t+1\right] + \frac{\sin\left(\pi(x+t)\right)}{4\pi}.$$



Region 4: Fig. 14.2,

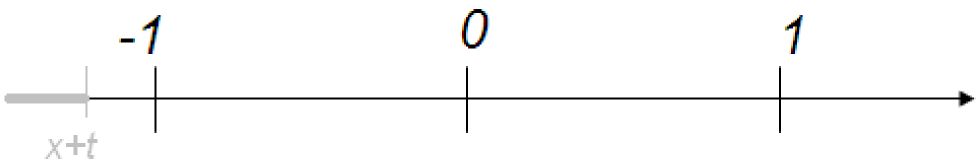
$$x - t \leq -1, \qquad x + t \geq 1$$

so $G(s) = g(s)$ for $s \in [-1, 1]$ and is zero elsewhere.

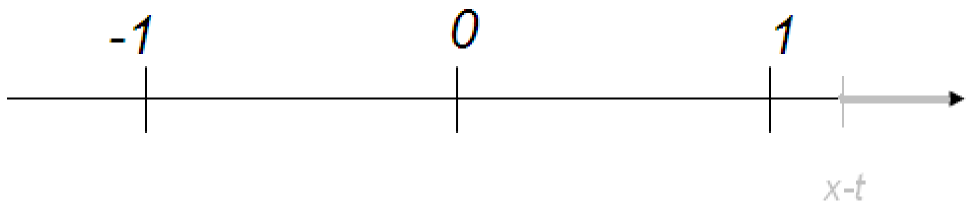
$$\begin{aligned}
\int_{x-t}^{x+t} G(s)\,\mathrm{d}s &= \int_{-1}^{1} \cos^2(\pi s/2)\,\mathrm{d}s, \\
&= \tfrac{1}{2}\left[1 - (-1)\right] + \frac{\sin(\pi) - \sin(\pi(-1))}{2\pi}, \\
&= 1
\end{aligned}$$

Finally, in region 4,

$$u_4(x, t) = \tfrac{1}{2}.$$



Region 5: $x+t<-1$, hence $G(s)=0$.



Region 6: $x - t > 1$, hence $G(s) = 0$.

Putting it all together,

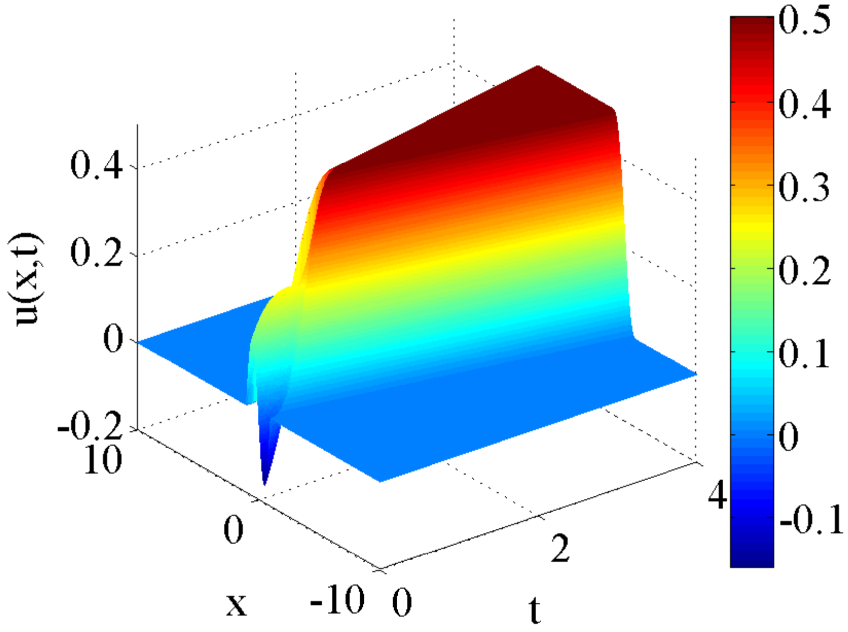
$$u(x,t) = \begin{cases} u_1(x,t), & (x,t) \in R_1, \\ u_2(x,t), & (x,t) \in R_2, \\ u_3(x,t), & (x,t) \in R_3, \\ u_4(x,t), & (x,t) \in R_4, \\ u_5(x,t), & (x,t) \in R_5, \\ u_6(x,t), & (x,t) \in R_6, \end{cases}$$

or,

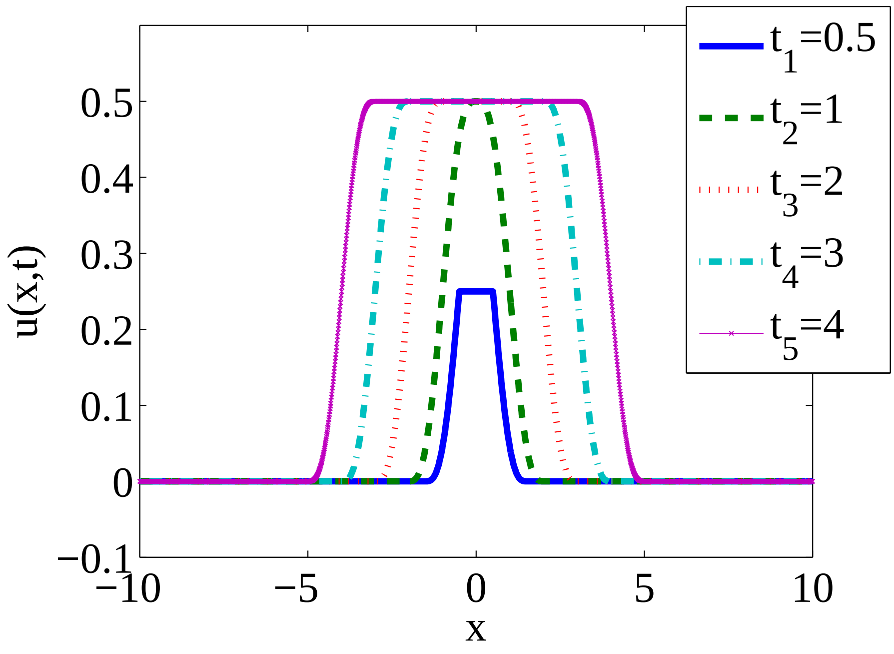
$$u(x,t) = \begin{cases} \frac{1}{2}t + \frac{1}{2\pi}\cos(\pi x)\sin(\pi t), & (x,t) \in R_1, \\ \frac{1}{4}\left[1 - x + t\right] - \frac{\sin(\pi(x-t))}{4\pi}, & (x,t) \in R_2, \\ \frac{1}{4}\left[x + t + 1\right] + \frac{\sin(\pi(x+t))}{4\pi}, & (x,t) \in R_3, \\ \frac{1}{2}, & (x,t) \in R_4, \\ 0, & (x,t) \in R_5, \\ 0, & (x,t) \in R_6, \end{cases}$$

Notes:

- Region 4 gives a contribution here. If there is no initial velocity ($u_t(x,t=0)=0$), there is no contribution from this region.
- I have sketched the d'Alembert solution in Fig. 14.3, using the code wavesolve__exact.m.
- However, there is also available my webpage, a finite-difference code integrate__sde.m. Both these codes agree at late times, but for early times they disagree. Can you explain why? **This is a very subtle but important point!**



(a)



(b)

Figure 14.3: The d'Alembert solution

Chapter 15

Fourier transforms: The definition

15.1 Motivation and basic idea

Recall our earlier result for Fourier series (Ch. 4, Theorem 4.1):

> Let $f \in L^2((-\pi,\pi))$ be piecewise smooth on the closed interval $[-\pi,\pi]$ and continuous on the open interval $(-\pi,\pi)$. Then, the Fourier series associated with f converges for all $x \in [-\pi,\pi]$ and converges to the sum $f(x)$ for all $x \in (-\pi,\pi)$.

Thus, the theorem says that

$$f(x) = \tfrac{1}{2}a_0 + \sum_{n=1}^{\infty} \left[a_n \cos(nx) + b_n \sin(nx)\right],$$

where

$$\begin{aligned} a_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} \cos(nx) f(x), \qquad n = 0, 1, \cdots \\ b_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} \cos(nx) f(x), \qquad n = 1, 2, \cdots . \end{aligned}$$

However, there is nothing special about the interval $[-\pi,\pi]$. Thus, on an interval $[-L/2, L/2]$, we can write

$$f(x) = \tfrac{1}{2}a_0 + \sum_{n=1}^{\infty} \left[a_n \cos\left(\frac{2\pi n}{L}x\right) + b_n \sin\left(\frac{2\pi n}{L}x\right)\right],$$

where

$$\begin{aligned} a_n &= \frac{2}{L}\int_{-L/2}^{L/2} \cos\left(\frac{2\pi n}{L}x\right) f(x)\mathrm{d}x, \qquad n = 0, 1, \cdots \\ b_n &= \frac{2}{L}\int_{-L/2}^{L/2} \sin\left(\frac{2\pi n}{L}x\right) f(x)\mathrm{d}x, \qquad n = 1, 2, \cdots . \end{aligned}$$

In this chapter, we are going to show what happens when $L \to \infty$.

15.2 Complex notation

First, we re-write the Fourier expansion in complex notation:

$$\begin{aligned} f(x) &= c_0 + \sum_{\substack{n=-\infty \\ n\neq 0}}^{\infty} c_n e^{\mathrm{i}(2\pi/L)nx}, \\ c_n &= \frac{1}{L}\int_{-L/2}^{L/2} f(x) e^{-\mathrm{i}(2\pi/L)nx}\mathrm{d}x, \end{aligned}$$

since

$$\int_{-L/2}^{L/2} e^{\mathrm{i}(2\pi/L)nx} e^{-\mathrm{i}(2\pi/L)mx}\mathrm{d}x = L\delta_{nm}.$$

We are going to equate the c_n's of the complex expansion with the a_n's and b_n's of the real expansion. First, we prove the following theorem:

Theorem 15.1 *Let $f(x)$ be a function with a complex Fourier expansion, with coefficients c_n, and $n \in \mathbb{Z}$. If $f(x)$ is a real-valued function, then*

$$c_n^* = c_{-n}.$$

Proof: Note first that

$$\begin{aligned} f(x) &= c_0 + \sum_{\substack{n=-\infty \\ n\neq 0}}^{\infty} c_n e^{\mathrm{i}(2\pi/L)nx}, \\ &= c_0 + \sum_{n=1}^{\infty} \left[c_n e^{\mathrm{i}(2\pi/L)nx} + c_{-n} e^{-\mathrm{i}(2\pi/L)nx}\right], \end{aligned}$$

But, for a real-valued function

$$f(x) = [f(x)]^*,$$

or

$$
\begin{aligned}
f(x) &= c_0^* + \left\{ \sum_{n=1}^{\infty} \left[c_n e^{\mathrm{i}(2\pi/L)nx} + c_{-n} e^{-\mathrm{i}(2\pi/L)nx} \right] \right\}^*, \\
&= c_0^* + \sum_{n=1}^{\infty} \left[c_n^* e^{-\mathrm{i}(2\pi/L)nx} + c_{-n}^* e^{\mathrm{i}(2\pi/L)nx} \right], \\
&= c_0^* + \sum_{n=1}^{\infty} \left[c_{-n}^* e^{\mathrm{i}(2\pi/L)nx} + c_n^* e^{-\mathrm{i}(2\pi/L)nx} \right], \\
&= c_0 + \sum_{n=1}^{\infty} \left[c_n e^{\mathrm{i}(2\pi/L)nx} + c_{-n} e^{-\mathrm{i}(2\pi/L)nx} \right].
\end{aligned}
$$

By orthogonality, we can equate each term:

$$
f(x) = [f(x)]^* \implies c_0 = c_0^*, \text{ and } c_n^* = c_{-n},\ n \geq 1.
$$

as required.

Now, take

$$
\begin{aligned}
f(x) &= c_0 + \sum_{n=1}^{\infty} \left[c_n e^{\mathrm{i}(2\pi/L)nx} + c_{-n} e^{-\mathrm{i}(2\pi/L)nx} \right], \\
&= c_0 + \sum_{n=1}^{\infty} \left[c_n e^{\mathrm{i}(2\pi/L)nx} + \text{c.c.} \right], \\
&= c_0 + \sum_{n=1}^{\infty} 2\Re\left(c_n e^{\mathrm{i}(2\pi/L)nx} \right).
\end{aligned}
$$

Use

$$
\Re(ab) = \Re(a)\Re(b) - \Im(a)\Im(b)
$$

to get

$$
f(x) = c_0 + \sum_{n=0}^{\infty} \left[2\Re(c_n) \cos\left(\frac{2\pi}{L} nx \right) - 2\Im(c_n) \sin\left(\frac{2\pi}{L} nx \right) \right].
$$

Thus, the real Fourier expansion

$$
f(x) = \tfrac{1}{2} a_0 + \sum_{n=1}^{\infty} \left[a_n \cos\left(\frac{2\pi}{L} nx \right) + b_n \sin\left(\frac{2\pi}{L} nx \right) \right]
$$

and the complex one are equivalent, with

$$\begin{aligned} a_0 &= 2c_0, \\ a_n &= 2\Re(c_n), \qquad n = 1, 2, \cdots, \\ b_n &= -2\Im(c_n), \qquad n = 1, 2, \cdots. \end{aligned}$$

15.3 The limit

We have

$$\begin{aligned} f(x) &= \sum_{n=-\infty}^{\infty} c_n e^{\mathrm{i}(2\pi/L)nx}, \\ &= \sum_{n=-\infty}^{\infty} \frac{1}{L}\left[\int_{-L/2}^{L/2} e^{-\mathrm{i}(2\pi/L)ns} f(s)\mathrm{d}s\right] e^{\mathrm{i}(2\pi/L)nx}. \end{aligned}$$

Call

$$k_n = \frac{2\pi}{L}n,$$

and

$$\Delta k = k_{n+1} - k_n = \frac{2\pi}{L}.$$

Hence,

$$f(x) = \sum_{n=-\infty}^{\infty} \frac{\Delta k}{2\pi}\left[\int_{-L/2}^{L/2} e^{-\mathrm{i}k_n s} f(s)\mathrm{d}s\right] e^{\mathrm{i}k_n x}.$$

But this is a straightforward Riemann sum! Letting $L \to \infty$ corresponds to $\Delta k \to 0$, and the sum converts into an integral:

$$f(x) = \int_{-\infty}^{\infty} \frac{\mathrm{d}k}{2\pi}\left[\int_{-\infty}^{\infty} f(s)e^{-\mathrm{i}ks}\mathrm{d}s\right] e^{\mathrm{i}kx}.$$

Thus, we have the function $f(x)$ and its Fourier transform:

$$\begin{aligned} \widehat{f}_k :&= \int_{-\infty}^{\infty} e^{-\mathrm{i}ks} f(s)\mathrm{d}s, \\ f(x) &= \int_{-\infty}^{\infty} \frac{\mathrm{d}k}{2\pi} \widehat{f}_k e^{\mathrm{i}kx}. \end{aligned}$$

This procedure is perfectly rigorous, provided the original conditions on the Fourier expansion are satisfied in the limit as $L \to \infty$, namely

- $f(x)$ should be square-integrable as the domain is extended infinitely:

$$\int_{-\infty}^{\infty} |f(x)|^2 \mathrm{d}x < \infty,$$

- $f(x)$ should be piecewise differentiable (C^1) on the whole of $\mathbb{R}$.

15.4 Example I

Compute the FT of $f(x) = e^{-x^2/2\sigma^2}$. We compute

$$\begin{aligned}
\widehat{f}_k &= \int_{-\infty}^{\infty} e^{-\mathrm{i}kx} f(x)\mathrm{d}x, \\
&= \int_{-\infty}^{\infty} e^{-\mathrm{i}kx} e^{-x^2/2\sigma^2}\mathrm{d}x, \\
&= \int_{-\infty}^{\infty} e^{-(x^2/2\sigma^2)-\mathrm{i}kx}\mathrm{d}x, \\
&= \int_{-\infty}^{\infty} e^{-(1/2\sigma^2)\left[x^2+2\sigma^2\mathrm{i}kx\right]}\mathrm{d}x,
\end{aligned}$$

Now complete the square:

$$\begin{aligned}
\widehat{f}_k &= \int_{-\infty}^{\infty} e^{-(1/2\sigma^2)\left[(x+\sigma^2\mathrm{i}k)^2+\sigma^4k^2\right]}\mathrm{d}x, \\
&= e^{-k^2\sigma^2/2}\int_{-\infty}^{\infty} e^{-(1/2\sigma^2)(x+\sigma^2\mathrm{i}k)^2}\mathrm{d}x, \\
&= e^{-k^2\sigma^2/2}\int_{-\infty}^{\infty} e^{-y^2/2\sigma^2}\mathrm{d}y, \\
&= \sqrt{2\sigma^2}e^{-k^2\sigma^2/2}\int_{-\infty}^{\infty} e^{-z^2}\mathrm{d}z, \\
&= \sqrt{2\pi\sigma^2}e^{-k^2\sigma^2/2}.
\end{aligned}$$

Thus, the Fourier transform of a Gaussian is itself a Gaussian.

Chapter 16

Interlude: functions of a single complex variable

16.1 A very brief summary of complex analysis

In this chapter, let $z = x + \mathrm{i}y \in \mathbb{C}$ and let

$$\begin{aligned} f : \mathbb{C} &\rightarrow \mathbb{C}, \\ z &\rightarrow f(z) \end{aligned}$$

be a complex-valued function of a single complex variable. An example of such a function is a power series,

$$f(z) = \sum_{n=0}^{\infty} a_n z^n,$$

assumed to be convergent on a disc D centred at the origin, with some non-zero radius. Such a series is differentiable as a complex function (analytic):

$$f'(z) = \sum_{n=1}^{\infty} n a_n z^{n-1};$$

a remarkable fact of complex analysis is that the power series for $f(z)$ is infinitely differentiable on D. Moreover, the integral of any convergent power series around a closed contour C in the complex plane is zero:

$$
\begin{aligned}
\int_C \mathrm{d}z f(z) &= \sum_{n=0}^{\infty} a_n \int_C \mathrm{d}z z^n, \\
&= \sum_{n=0}^{\infty} a_n \int_C \mathrm{d}z z^n, \\
&= \sum_{n=0}^{\infty} a_n \frac{z^{n+1}}{n+1}\bigg|_a^a, \\
&= 0.
\end{aligned}
$$

Any analytic function in the complex plane has a power series valid on some disc of finite radius. Thus, this result can be extended to all analytic functions:

$$\int_C \mathrm{d}z f(z) = 0, \qquad \text{for all analytic functions, and any closed contour } C.$$

Now let's look at what happens when $f(z)$ has a simple pole at $z=0$. Thus, $f(z)$ splits into an analytic part and a singular part:

$$f(z) = \frac{A}{z} + f_1(z),$$

where $f_1(z)$ is analytic. Let us integrate this function around the circle of radius R centred at zero. This is a closed curve, hence $\int_C \mathrm{d}z f_1(z) = 0$, and

$$\int_C \mathrm{d}z f(z) = A \int \frac{\mathrm{d}z}{z}.$$

But $z = Re^{\mathrm{i}\theta}$, hence

$$\mathrm{d}z = \mathrm{d}Re^{\mathrm{i}\theta} + Re^{\mathrm{i}\theta}\mathrm{i}\mathrm{d}\theta = 0 + Re^{\mathrm{i}\theta}\mathrm{i}\mathrm{d}\theta.$$

Thus,

$$
\begin{aligned}
\int_C \mathrm{d}z f(z) &= A \int_C \frac{\mathrm{d}z}{z}, \\
&= \int_0^{2\pi} A \frac{R\mathrm{i}\mathrm{d}\theta}{Re^{\mathrm{i}\theta}}, \\
&= \int_0^{2\pi} A\mathrm{i}\mathrm{d}\theta, \\
&= 2\pi\mathrm{i}A.
\end{aligned}
$$

More generally, let $f(z)$ be a complex-valued function of the form

$$f(z) = \frac{A}{z-a} + f_1(z),$$

where $a \in \mathbb{C}$ is a constant and $f_1(z)$ is a regular function. Then, A is defined to be the **residue** of $f(z)$ at $z = a$:

$$\text{Res}(f, a) = A = \lim_{z \to a} \left[f(z)\,(z-a) \right],$$

and the following result holds:

$$\int_C f(z) \mathrm{d}z = 2\pi A,$$

for **any closed contour** C **that encloses the point** a.

This is called **the residue theorem**.

16.2 Contour integrals

Evaluate

$$I = \int_{-\infty}^{\infty} \frac{\mathrm{d}x}{1+x^2}.$$

We complexify the problem and consider instead the integral

$$I(R) := \int_C \frac{\mathrm{d}z}{1+z^2},$$

where C is a semicircle of radius R whose diameter lies on the x-axis between $-R$ and R, and which extends into the upper half-plane. In this domain, the function $f(z) = (1+z^2)^{-1}$ has a pole at $z = \mathrm{i}$, since

$$f(z) = \frac{1}{(z+\mathrm{i})\,(z-\mathrm{i})}.$$

We write

$$f(z) = \frac{A}{z-\mathrm{i}} + f_1(z),$$

where $f_1(z)$ is regular. We also compute

$$f(z)(z-\mathrm{i}) = A + f_1(z)(z-\mathrm{i}),$$

hence

$$\lim_{z \to \mathrm{i}} \left[f(z)(z-\mathrm{i}) \right] = A + \lim_{z \to \mathrm{i}} \left[f_1(z)(z-\mathrm{i}) \right] = A + 0.$$

But

$$\begin{aligned}\lim_{z\to \mathrm{i}}\left[f(z)(z-\mathrm{i})\right] &= \lim_{z\to \mathrm{i}}\left[\frac{1}{(z+\mathrm{i})\,(z-\mathrm{i})}\times(z-\mathrm{i})\right],\\ &= \lim_{z\to \mathrm{i}}\frac{1}{z+\mathrm{i}},\\ &= \tfrac{1}{2\mathrm{i}} = A;\end{aligned}$$

$$\mathsf{Res}(f,\mathrm{i}) = A = \frac{1}{2\mathrm{i}}.$$

From the last section, it follows that

$$I(R) = \int_C f(z)\mathrm{d}z = 2\pi\mathrm{i}A = 2\pi\mathrm{i}\left(\frac{1}{2\mathrm{i}}\right) = \pi.$$

Now C can be broken up into two parts (Fig. 16.1): First, a semi-circle starting at $(x,y)=(R,0)$ $(\theta=0)$ and ending at $(x,y)=(-R,0)$ $(\theta=\pi)$. Along this semi-circle,

$$\frac{\mathrm{d}z}{1+z^2} = \frac{\mathrm{i}Re^{\mathrm{i}\theta}}{1+R^2e^{2\mathrm{i}\theta}} \to 0 \text{ as } R\to\infty.$$

Second, a line segment going from $(x,y)=(-R,0)$ to $(x,y)=(R,0)$. Along this line segment,

$$\frac{\mathrm{d}z}{1+z^2} = \frac{\mathrm{d}x}{1+x^2}.$$

Hence,

$$\pi = I(R) = \left[\int_{\text{Semi-circle}} + \int_{\text{Line}}\right]\frac{\mathrm{d}z}{1+z^2} \to \int_{-\infty}^{\infty}\frac{\mathrm{d}x}{1+x^2}, \text{ as } R\to\infty,$$

and finally,

$$\pi = \int_{-\infty}^{\infty}\frac{\mathrm{d}x}{1+x^2}.$$

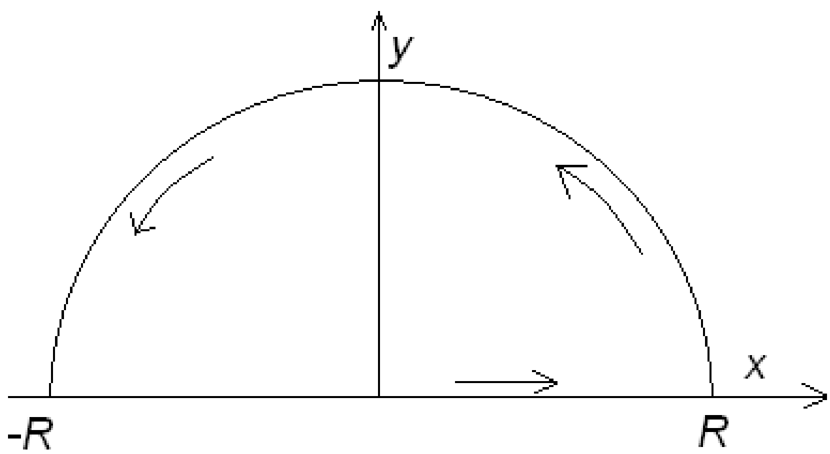


Figure 16.1: Contour for $\int_{-\infty}^{\infty} \mathrm{d}x(1+x^2)^{-1}$

Chapter 17

Fourier transforms continued

Overview

This chapter continues with examples of computing Fourier transforms. The integrals are tricky, and rely on the complex-variable theory introduced in Ch. 16.

17.1 An example

Compute the FT of

$$f(x) = \frac{1}{1+x^2}.$$

We compute

$$\begin{aligned} \widehat{f}_k &= \int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}kx} f(x)\mathrm{d}x, \\ &= \int_{-\infty}^{\infty} \frac{\mathrm{e}^{-\mathrm{i}kx}}{1+x^2}\mathrm{d}x. \qquad (*) \end{aligned}$$

As in Ch. 16, we complexify the problem and solve the following apparently more difficult problem:

$$I := \int_C \frac{\mathrm{e}^{-\mathrm{i}kz}}{1+z^2}\mathrm{d}z := \int_C f(z)\mathrm{d}z.$$

where $z = x + \mathrm{i}y$ and C is a closed path in the complex plane to be determined. The choice of C depends on the sign of k.

Case 1: We have $k > 0$. Then,

$$-\mathrm{i}kz = -\mathrm{i}k(x + \mathrm{i}y) = -\mathrm{i}kx + ky.$$

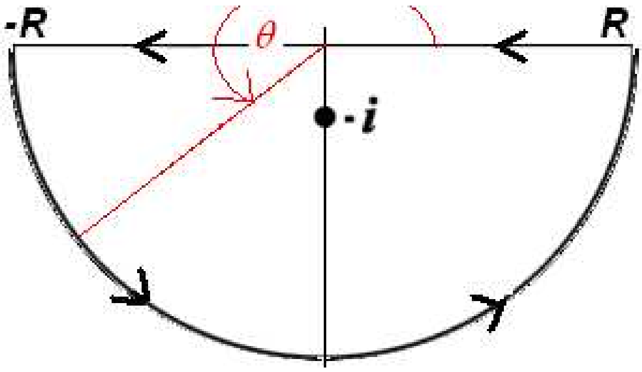


Figure 17.1: The contour for $k > 0$ (Case 1)

To get a convergent exponential, $\mathrm{e}^{-\mathrm{i}kz}$, the contour must be in the **lower half-plane**. We therefore take the path to be a line segment $C_1 = [-R, R]$ along the real line, together with a semicircle C_2 of radius R in the lower half-plane. See Fig. 17.1. The integral therefore splits into two parts:

$$C_1: \qquad \mathrm{d}z = \mathrm{d}x, \qquad y = 0,$$
$$C_2: \qquad \mathrm{d}z = \mathrm{i}R\mathrm{e}^{\mathrm{i}\theta}\mathrm{d}\theta, \qquad \pi < \theta < 2\pi.$$

We have,

$$I = \int_C f(z)\mathrm{d}z = \int_{C_1} f(z)\mathrm{d}z + \int_{C_2} f(z)\mathrm{d}z := I_1 + I_2.$$

Consider I_1. Letting $R \to \infty$ gives

$$I_1 \to \int_{+\infty}^{-\infty} f(x)\mathrm{d}x = -\int_{-\infty}^{+\infty} f(x)\mathrm{d}x,$$

which is the required integral. Letting $R \to \infty$ in I_2 gives

$$I_2 = \int_{\pi}^{2\pi} \frac{\mathrm{e}^{-\mathrm{i}kR\cos\theta}\mathrm{e}^{kR\sin\theta}}{1 + R^2\mathrm{e}^{2i\theta}}\mathrm{i}R\mathrm{e}^{\mathrm{i}\theta}\mathrm{d}\theta,$$

However, the exponential $\mathrm{e}^{kR\sin\theta}$ is a decaying function of R since $k > 0$ and $\sin\theta < 0$. Thus,

$$I_2 \to 0, \text{ as } R \to \infty.$$

Hence, as $R \to \infty$,

$$I = I_1 = -\int_{-\infty}^{\infty} f(x)\mathrm{d}x.$$

Now let us examine if the integral has any singularities in the space enclosed by the contour C.

Indeed, in this domain, the function $f(z) = \mathrm{e}^{-\mathrm{i}kz}/(1+z^2)$ has a pole at $z = -\mathrm{i}$, since

$$f(z) = \frac{\mathrm{e}^{-\mathrm{i}kz}}{(z+\mathrm{i})(z-\mathrm{i})}.$$

We write

$$f(z) = \frac{A}{z+\mathrm{i}} + f_1(z),$$

where $f_1(z)$ is regular. We also compute

$$f(z)(z+\mathrm{i}) = A + f_1(z)(z+\mathrm{i}),$$

hence

$$\lim_{z\to-\mathrm{i}} [f(z)(z+\mathrm{i})] = A + \lim_{z\to-\mathrm{i}} [f_1(z)(z+\mathrm{i})] = A + 0.$$

Now

$$\begin{aligned}
\lim_{z\to-\mathrm{i}} [f(z)(z+\mathrm{i})] &= \lim_{z\to-\mathrm{i}} \left[\frac{\mathrm{e}^{-\mathrm{i}kz}}{(z+\mathrm{i})(z-\mathrm{i})} \times (z+\mathrm{i})\right], \\
&= \lim_{z\to-\mathrm{i}} \frac{\mathrm{e}^{-\mathrm{i}kz}}{z-\mathrm{i}}, \\
&= -\frac{\mathrm{e}^{-k}}{2\mathrm{i}}, \\
&= A.
\end{aligned}$$

From Ch. 16, we have

$$I = \int_C f(z)\mathrm{d}z = 2\pi\mathrm{i}A,$$

But

$$I = I_1 = -\int_{-\infty}^{\infty} f(x)\mathrm{d}x.$$

hence

$$\int_{-\infty}^{\infty} f(x)\mathrm{d}x = +\pi\mathrm{e}^{-k},$$

or

$$\int_{-\infty}^{\infty} \frac{\mathrm{e}^{-\mathrm{i}kx}}{1+x^2}\mathrm{d}x = \pi\mathrm{e}^{-k}, \qquad k > 0.$$

Case 2: $k < 0$. Then,

$$-\mathrm{i}kz = -\mathrm{i}k(x+\mathrm{i}y) = -\mathrm{i}kz + ky.$$

To get a convergent exponential, $\mathrm{e}^{-\mathrm{i}kz}$, the contour must be in the **upper half-plane**. We therefore take the path to be a line segment $C_1 = [-R, R]$ along the real line, together with a semicircle C_2

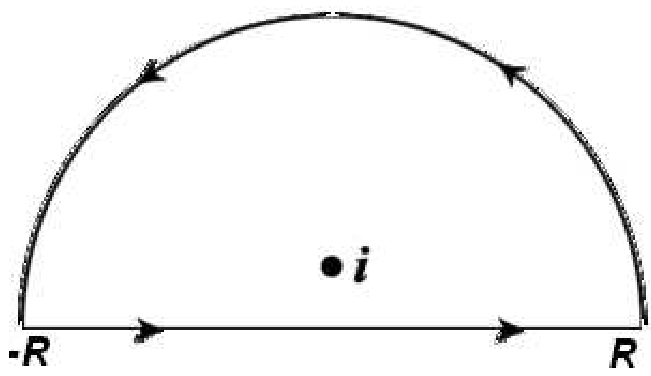


Figure 17.2: The contour for $k < 0$ (Case 2)

of radius R in the upper half-plane. See Fig. 17.2. The integral therefore splits into two parts:

$$\begin{aligned} C_1 : &\qquad \mathrm{d}z = \mathrm{d}x, \qquad y = 0, \\ C_2 : &\qquad \mathrm{d}z = \mathrm{i}R\mathrm{e}^{\mathrm{i}\theta}\mathrm{d}\theta, \qquad 0 < \theta < \pi. \end{aligned}$$

We have,

$$I = \int_C f(z)\mathrm{d}z = \int_{C_1} f(z)\mathrm{d}z + \int_{C_2} f(z)\mathrm{d}z := I_1 + I_2.$$

Consider I_1. As before (but watch for the change in sign)

$$\begin{aligned} I_1 &\rightarrow +\int_{-\infty}^{\infty} f(x)\mathrm{d}x, \\ I_2 &\rightarrow 0 \end{aligned}$$

as as $R \to \infty$. Hence, as $R \to \infty$,

$$I = I_1 = \int_{-\infty}^{\infty} f(x)\mathrm{d}x.$$

Now let us check if the integral has any singularities in the space enclosed by the contour C. In this domain, the function $f(z) = \mathrm{e}^{-\mathrm{i}kz}/(1+z^2)$ has a pole at $z = +\mathrm{i}$, since

$$f(z) = \frac{\mathrm{e}^{-\mathrm{i}kz}}{(z+\mathrm{i})(z-\mathrm{i})}.$$

We write

$$f(z) = \frac{A}{z-\mathrm{i}} + f_1(z),$$

where $f_1(z)$ is regular. We also compute

$$f(z)(z-\mathrm{i}) = A + f_1(z)(z-\mathrm{i}),$$

hence

$$\lim_{z\to\mathrm{i}}\left[f(z)(z-\mathrm{i})\right] = A + \lim_{z\to\mathrm{i}}\left[f_1(z)(z-\mathrm{i})\right] = A + 0.$$

Now

$$\begin{aligned}
\lim_{z\to\mathrm{i}}\left[f(z)(z-\mathrm{i})\right] &= \lim_{z\to\mathrm{i}}\left[\frac{\mathrm{e}^{-\mathrm{i}kz}}{(z+\mathrm{i})\,(z-\mathrm{i})}\times(z-\mathrm{i})\right],\\
&= \lim_{z\to\mathrm{i}}\frac{\mathrm{e}^{-\mathrm{i}kz}}{z+\mathrm{i}},\\
&= +\frac{\mathrm{e}^{+k}}{2\mathrm{i}},\\
&= A.
\end{aligned}$$

From Ch. 16, we have

$$I = \int_C f(z)\mathrm{d}z = 2\pi\mathrm{i}A,$$

But

$$I = I_1 = +\int_{-\infty}^{\infty} f(x)\mathrm{d}x.$$

hence

$$\int_{-\infty}^{\infty} f(x)\mathrm{d}x = +\pi\mathrm{e}^{+k},$$

or

$$\int_{-\infty}^{\infty}\frac{\mathrm{e}^{-\mathrm{i}kx}}{1+x^2}\mathrm{d}x = \pi\mathrm{e}^{k}, \qquad k<0.$$

Putting the two cases together gives

$$\int_{-\infty}^{\infty}\frac{\mathrm{e}^{-\mathrm{i}kx}}{1+x^2}\mathrm{d}x = \pi\mathrm{e}^{-|k|}.$$

Chapter 18

The delta function; convolution theorem

18.1 The definition

Definition 18.1 *The* **delta function** *is defined by the following three properties:*

1. $$\delta(x) = \begin{cases} 0, & x \neq 0, \\ \infty, & x = 0; \end{cases}$$

2. $$\int_{-\infty}^{\infty} \delta(x)\, \mathrm{d}x = 1;$$

3. $$\int_{-\infty}^{\infty} f(x)\delta(x-a)\, \mathrm{d}x = f(a),$$

   for any function $f(x)$ that decays rapidly as $|x| \to \infty$.

As you can see from these properties, the delta function is not, strictly speaking, a function, but rather it is a **distribution**. However, in this chapter we show how to construct the delta function as a limit of a sequence of real-valued functions.

Example 1

Consider the sequence of functions

$$\delta_n(x) = \frac{n}{\sqrt{\pi}} \mathrm{e}^{-n^2x^2}, \qquad n = 1, 2, \cdots .$$

Property 2:

$$\begin{aligned}
\int_{-\infty}^{\infty}\delta_n(x)\,\mathrm{d}x &= \frac{n}{\sqrt{\pi}}\int_{-\infty}^{\infty}\mathrm{e}^{-n^2x^2}\mathrm{d}x,\\
&= \frac{n}{\sqrt{\pi}}\frac{1}{n}\int_{-\infty}^{\infty}\mathrm{e}^{-z^2}\mathrm{d}z,\\
&= \frac{n}{\sqrt{\pi}}\frac{1}{n}\sqrt{\pi}=1.
\end{aligned}$$

Property 3: Let $f(\cdot)$ have a Taylor series. Then,

$$\begin{aligned}
\int_{-\infty}^{\infty} f(x)\delta_n(x-a)\,\mathrm{d}x &= \frac{n}{\sqrt{\pi}}\int_{-\infty}^{\infty}\mathrm{e}^{-n^2(x-a)^2}f(x)\,\mathrm{d}x,\\
&= \frac{n}{\sqrt{\pi}}\int_{-\infty}^{\infty}\mathrm{e}^{-n^2y^2}f(y+a)\,\mathrm{d}y, \qquad y=x-a,\\
&= \frac{n}{\sqrt{\pi}}\int_{-\infty}^{\infty}\mathrm{e}^{-n^2y^2}\left(\sum_{p=0}^{\infty}\frac{f^{(p)}(a)}{p!}y^p\right)\mathrm{d}y,\\
&= \frac{n}{\sqrt{\pi}}\left(\sum_{p=0}^{\infty}\frac{f^{(p)}(a)}{p!}\int_{-\infty}^{\infty}\mathrm{e}^{-n^2y^2}y^p\right)\mathrm{d}y,\\
&= f(a)\frac{n}{\sqrt{\pi}}\int_{-\infty}^{\infty}\mathrm{e}^{-n^2y^2}\,\mathrm{d}y+\sum_{p=1}^{\infty}\frac{n}{\sqrt{\pi}}\left(\frac{f^{(p)}(a)}{p!}\int_{-\infty}^{\infty}\mathrm{e}^{-n^2y^2}y^p\right)\mathrm{d}y,\\
&= f(a)+\sum_{2,4,6,\cdots}\frac{n}{\sqrt{\pi}}\left(\frac{f^{(p)}(a)}{p!}2\int_{0}^{\infty}\mathrm{e}^{-n^2y^2}y^p\right)\mathrm{d}y,
\end{aligned}$$

and it remains to show that the second term vanishes as $n\to\infty$. Call the second term I_n and introduce $z=ny$. Thus,

$$\begin{aligned}
I_n &= \sum_{2,4,6,\cdots}\frac{1}{\sqrt{\pi}}\frac{f^{(p)}(a)}{p!}\frac{1}{n^p}2\int_0^{\infty}\mathrm{e}^{-z^2}z^p\mathrm{d}z,\\
&= \sum_{2,4,6,\cdots}\frac{1}{\sqrt{\pi}}\frac{f^{(p)}(a)}{p!}\frac{1}{n^p}\Gamma\left((p+1)/2\right).
\end{aligned}$$

Taking $n\to\infty$ gives $I_n\to 0$, and property 3 is proved.

Example 2

Consider the family of functions

$$\delta_L(x)=\int_{-L}^{L}\mathrm{e}^{\mathrm{i}kx}\frac{\mathrm{d}k}{2\pi}, \qquad L>0.$$

First, let us re-write $\delta_L(x)$, by doing the integral:

$$\begin{aligned}
\delta_L(x) &= \frac{1}{2\pi}\frac{1}{\mathrm{i}x}\left(\mathrm{e}^{\mathrm{i}xL}-\mathrm{e}^{-\mathrm{i}xL}\right), \\
&= \frac{1}{2\pi}\times 2L\times\left(\frac{\mathrm{e}^{\mathrm{i}xL}-\mathrm{e}^{-\mathrm{i}xL}}{2iLx}\right), \\
&= \frac{L}{\pi}\frac{\sin(Lx)}{Lx}.
\end{aligned}$$

Property 2:

$$\int_{-\infty}^{\infty}\delta_L(x)\,\mathrm{d}x = \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\sin(s)}{s}\mathrm{d}s, \qquad s = Lx,$$

Using contour integration or other methods (homework), this integral can be shown to be unity:

$$\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\sin(s)}{s}\mathrm{d}y = 1.$$

Property 3: As before, take

$$\begin{aligned}
\int_{-\infty}^{\infty} f(x)\delta_L(x-a)\,\mathrm{d}x &= \frac{L}{\pi}\int_{-\infty}^{\infty}\frac{\sin(L(x-a))}{L(x-a)}f(x)\,\mathrm{d}x, \\
&= \frac{L}{\pi}\int_{-\infty}^{\infty}\frac{\sin(Ly)}{Ly}f(y+a)\,\mathrm{d}y, \\
&= \frac{L}{\pi}\int_{-\infty}^{\infty}\frac{\sin(Ly)}{Ly}\left(\sum_{p=0}^{\infty}\frac{f^{(p)}(a)}{p!}y^p\right)\,\mathrm{d}y, \\
&= \frac{L}{\pi}\left(\sum_{p=0}^{\infty}\frac{f^{(p)}(a)}{p!}\int_{-\infty}^{\infty}\frac{\sin(Ly)}{Ly}y^p\right)\,\mathrm{d}y, \\
&= f(0)\frac{L}{\pi}\int_{-\infty}^{\infty}\frac{\sin(Ly)}{Ly}\,\mathrm{d}y + \sum_{p=1}^{\infty}\frac{L}{\pi}\left(\frac{f^{(p)}(a)}{p!}\int_{-\infty}^{\infty}\frac{\sin(Ly)}{Ly}y^p\right)\,\mathrm{d}y, \\
&= f(0) + \sum_{2,4,6,\cdots}\frac{L}{\pi}\left(\frac{f^{(p)}(a)}{p!}2\int_{0}^{\infty}\frac{\sin(Ly)}{Ly}y^p\right)\,\mathrm{d}y,
\end{aligned}$$

and it remains to show that the second term vanishes as $n\rightarrow\infty$. Call the second term $I(L)$ and introduce $s = Ly$. Thus,

$$I(L) = \sum_{2,4,6,\cdots}\frac{1}{\pi}\frac{f^{(p)}(a)}{p!}\frac{1}{L^p}\int_{-\infty}^{\infty}\frac{\sin(s)}{s}s^p\mathrm{d}s.$$

Properly interpreted, the integral

$$I_q := \int_0^{\infty}\sin(s)s^q\mathrm{d}s, \qquad q = 1,3,5,\cdots$$

is finite. E.g.

$$I_q := (-1)^q \lim_{\lambda\to 0}\left[\frac{d^q}{d\lambda^q}\int_0^\infty \mathrm{e}^{-\lambda s}\sin(s)\mathrm{d}s\right].$$

Hence,

$$I(L) \to 0 \text{ as } L \to \infty,$$

and property 3 is proved.

Corollary

We have shown,

$$\delta(x) = \int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}kx}\frac{\mathrm{d}k}{2\pi}. \qquad (*)$$

But recall the Fourier decomposition of a function:

$$\delta(x) = \int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}kx}\widehat{\delta}_k\frac{\mathrm{d}k}{2\pi}. \qquad (**)$$

Equating coefficients in (*) and (**) gives

$$\widehat{\delta}_k = 1.$$

Thus, **the Fourier transform of a delta function is unity.** In other words, the delta function is a sum of all normal modes, where each mode is given the same amplitude.

18.2 The convolution theorem

Let $f(x)$ and $g(x)$ be two real-valued functions. Define the convolution,

$$(f*g)(x) = \int_{-\infty}^{\infty} f(x-y)g(y)\mathrm{d}y.$$

It follows that

$$(f*g)(x) = \int_{-\infty}^{\infty}\frac{\mathrm{d}k}{2\pi}\mathrm{e}^{\mathrm{i}kx}\widehat{f}_k\widehat{g}_k.$$

Proof: By direct computation,

$$\begin{aligned}
(f*g)(x) &= \int_{-\infty}^{\infty} f(x-y)g(y)\mathrm{d}y, \\
&= \int_{-\infty}^{\infty} \mathrm{d}y \left(\int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}k(x-y)} \widehat{f}_k \frac{\mathrm{d}k}{2\pi} \right) \left(\int_{-\infty}^{\infty} \mathrm{e}^{ipy} \widehat{g}_p \frac{\mathrm{d}p}{2\pi} \right), \\
&= \int_{-\infty}^{\infty} \mathrm{d}y \left(\int_{-\infty}^{\infty} \frac{\mathrm{d}k}{2\pi} \int_{-\infty}^{\infty} \frac{\mathrm{d}p}{2\pi} \widehat{f}_k \widehat{g}_p \mathrm{e}^{\mathrm{i}k(x-y)} \mathrm{e}^{ipy} \right), \\
&= \int_{-\infty}^{\infty} \frac{\mathrm{d}k}{2\pi} \int_{-\infty}^{\infty} \frac{\mathrm{d}p}{2\pi} \widehat{f}_k \widehat{g}_p \left(\int_{-\infty}^{\infty} \mathrm{d}y\, \mathrm{e}^{\mathrm{i}k(x-y)} \mathrm{e}^{ipy} \right), \\
&= \int_{-\infty}^{\infty} \frac{\mathrm{d}k}{2\pi} \int_{-\infty}^{\infty} \mathrm{d}p\, \widehat{f}_k \widehat{g}_p \mathrm{e}^{\mathrm{i}kx} \left(\int_{-\infty}^{\infty} \frac{\mathrm{d}y}{2\pi}\, \mathrm{e}^{iy(p-k)} \right), \\
&= \int_{-\infty}^{\infty} \frac{\mathrm{d}k}{2\pi} \int_{-\infty}^{\infty} \mathrm{d}p\, \widehat{f}_k \widehat{g}_p \mathrm{e}^{\mathrm{i}kx} \delta(p-k), \\
&= \int_{-\infty}^{\infty} \frac{\mathrm{d}k}{2\pi} \widehat{f}_k \widehat{g}_k \mathrm{e}^{\mathrm{i}kx}.
\end{aligned}$$

In other words, **convolution in real space is merely multiplication in Fourier space**.

Chapter 19

Fourier transforms: Applications

19.1 Laplace's equation on a semi-infinite domain

Solve **Laplace's equation**

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0,$$

on the semi-infinite domain

$$\Omega = \{(x, y) | -\infty < x < \infty, 0 < y < \infty\},$$

with BCs

$$|u(x, y)| \to 0 \text{ as } y \to \infty,$$
$$|u(x, y)| \to 0 \text{ as } |x| \to \infty.$$

Because of the shape of the boundary in the x-direction (an infinite line), we attempt a Fourier transform in the x-direction only:

$$\widehat{u}_k(y) = \int_{-\infty}^{\infty} u(x, y) \mathrm{e}^{-\mathrm{i}kx} \mathrm{d}x.$$

Let's also FT the equation itself:

$$\begin{aligned}
0 &= \int_{-\infty}^{\infty} [u_{yy} + u_{xx}]\, \mathrm{e}^{-\mathrm{i}kx} \mathrm{d}x, \\
&= \int_{-\infty}^{\infty} u_{yy} \mathrm{e}^{-\mathrm{i}kx} \mathrm{d}x + \int_{-\infty}^{\infty} u_{xx} \mathrm{e}^{-\mathrm{i}kx} \mathrm{d}x, \\
&= \frac{\partial^2}{\partial y^2} \int_{-\infty}^{\infty} u \mathrm{e}^{-\mathrm{i}kx} \mathrm{d}x + \int_{-\infty}^{\infty} \left[\frac{\partial}{\partial x} \left(u_x \mathrm{e}^{-\mathrm{i}kx} \right) + \mathrm{i}k u_x \mathrm{e}^{-\mathrm{i}kx} \right] \mathrm{d}x, \\
&= \frac{\partial^2}{\partial y^2} \widehat{u}_k(y) + \left(u_x \mathrm{e}^{-\mathrm{i}kx} \right)_{-\infty}^{\infty} + \mathrm{i}k \int_{-\infty}^{\infty} u_x \mathrm{e}^{-\mathrm{i}kx} \mathrm{d}x
\end{aligned}$$

If the function decays sufficiently fast at the boundaries (and we assume it does), then this is

$$\begin{aligned}
0 &= \frac{\partial^2}{\partial y^2} \widehat{u}_k(y) + \mathrm{i}k \int_{-\infty}^{\infty} u_x \mathrm{e}^{-\mathrm{i}kx} \mathrm{d}x, \\
&= \frac{\partial^2}{\partial y^2} \widehat{u}_k(y) + \mathrm{i}k \int_{-\infty}^{\infty} \left[\frac{\partial}{\partial x} \left(u \mathrm{e}^{-\mathrm{i}kx} \right) + \mathrm{i}k u \mathrm{e}^{-\mathrm{i}kx} \right] \mathrm{d}x, \\
&= \frac{\partial^2}{\partial y^2} \widehat{u}_k(y) + \left(u \mathrm{e}^{-\mathrm{i}kx} \right)_{-\infty}^{\infty} - k^2 \int_{-\infty}^{\infty} u \mathrm{e}^{-\mathrm{i}kx} \mathrm{d}x, \\
&= \frac{\partial^2}{\partial y^2} \widehat{u}_k(y) + 0 - k^2 \widehat{u}_k(y).
\end{aligned}$$

Thus, we are reduced to solving

$$\frac{\partial^2 \widehat{u}_k(y)}{\partial y^2} - k^2 \widehat{u}_k(y) = 0.$$

We know the solution straight away:

$$\widehat{u}_k(y) = A \mathrm{e}^{-ky} + B \mathrm{e}^{ky}.$$

In order for the solution to be bounded as $y \to \infty$, we must take

$$\widehat{u}_k(y) = A \mathrm{e}^{-ky},$$

if $k > 0$ and

$$\widehat{u}_k(y) = B \mathrm{e}^{ky},$$

if $k < 0$. In other words, we have

$$\widehat{u}_k(y) = A_k \mathrm{e}^{-|k|y},$$

where the subscript k on the constant of integration labels the mode $\mathrm{e}^{\mathrm{i}kx}$ in the Fourier transform. Now we sum over all such solutions to obtain the general solution. This amounts to taking the inverse Fourier transform:

$$u(x, y) = \int_{-\infty}^{\infty} \frac{\mathrm{d}k}{2\pi} \mathrm{e}^{\mathrm{i}kx} A_k \mathrm{e}^{-|k|y}.$$

Now we must fix the precise values of A_k. We have,

$$
\begin{aligned}
u(x,y=0) &= \int_{-\infty}^{\infty}\frac{\mathrm{d}k}{2\pi}\mathrm{e}^{\mathrm{i}kx}A_k,\\
&= f(x),\\
&= \int_{-\infty}^{\infty}\frac{\mathrm{d}k}{2\pi}\mathrm{e}^{\mathrm{i}kx}\widehat{f}_k.
\end{aligned}
$$

Matching up Fourier coefficients gives

$$\widehat{f}_k = A_k.$$

Thus, the final answer is

$$u(x,y) = \int_{-\infty}^{\infty}\frac{\mathrm{d}k}{2\pi}\mathrm{e}^{\mathrm{i}kx}\mathrm{e}^{-y|k|}\widehat{f}_k. \tag{19.1}$$

This is a perfectly acceptable final answer. But, recall Ch. 17, where we found

$$\left(\widehat{\frac{1}{\pi}\frac{y}{y^2+x^2}}\right)_k = \mathrm{e}^{-y|k|}, \qquad y>0.$$

The multiplication of the two Fourier transforms in Eq. (19.1) therefore corresponds to a convolution:

$$
\begin{aligned}
u(x,y) &= \left[\left(\frac{1}{\pi}\frac{y}{y^2+x^2}\right)*f\right]_x,\\
&= \frac{y}{\pi}\int_{-\infty}^{\infty}\frac{f(s)}{y^2+(x-s)^2}\mathrm{d}s.
\end{aligned}
$$

19.2 The diffusion equation on the line

Solve the diffusion equation

$$u_t = Du_{xx}$$

on the line $-\infty < x < \infty$ with initial condition

$$u(x,t=0) = f(x).$$

Because of the shape of the boundary in the x-direction (an infinite line), we attempt a Fourier transform in the x-direction only:

$$\widehat{u}_k(t) = \int_{-\infty}^{\infty} u(x,t)\mathrm{e}^{-\mathrm{i}kx}\mathrm{d}x.$$

Let's also FT the equation itself:

$$\begin{aligned}
\int_{-\infty}^{\infty} u_t \mathrm{e}^{-\mathrm{i}kx}\mathrm{d}x &= D\int_{-\infty}^{\infty} u_{xx}\mathrm{e}^{-\mathrm{i}kx}\mathrm{d}x,\\
\frac{\partial}{\partial t}\int_{-\infty}^{\infty} u\mathrm{e}^{-\mathrm{i}kx}\mathrm{d}x &= D\int_{-\infty}^{\infty} u_{xx}\mathrm{e}^{-\mathrm{i}kx}\mathrm{d}x,\\
\frac{\partial}{\partial t}\widehat{u}_k(t) &= D\int_{-\infty}^{\infty}\left[\frac{\partial}{\partial x}\left(u_x\mathrm{e}^{-\mathrm{i}kx}\right)+\mathrm{i}ku_x\mathrm{e}^{-\mathrm{i}kx}\right]\mathrm{d}x,\\
&= D\left(u_x\mathrm{e}^{-\mathrm{i}kx}\right)_{-\infty}^{\infty}+\mathrm{i}kD\int_{-\infty}^{\infty}u_x\mathrm{e}^{-\mathrm{i}kx}\mathrm{d}x
\end{aligned}$$

If the function decays sufficiently fast at the boundaries (and we assume it does), then this is

$$\begin{aligned}
\frac{\partial}{\partial t}\widehat{u}_k(t) &= \mathrm{i}kD\int_{-\infty}^{\infty}u_x\mathrm{e}^{-\mathrm{i}kx}\mathrm{d}x,\\
&= \mathrm{i}kD\int_{-\infty}^{\infty}\left[\frac{\partial}{\partial x}\left(u\mathrm{e}^{-\mathrm{i}kx}\right)+\mathrm{i}ku\mathrm{e}^{-\mathrm{i}kx}\right]\mathrm{d}x,\\
&= D\left(u\mathrm{e}^{-\mathrm{i}kx}\right)_{-\infty}^{\infty}-Dk^2\int_{-\infty}^{\infty}u\mathrm{e}^{-\mathrm{i}kx}\mathrm{d}x,\\
&= -Dk^2\widehat{u}_k(y).
\end{aligned}$$

Thus, we are reduced to solving

$$\frac{\partial\widehat{u}_k(t)}{\partial t}-Dk^2\widehat{u}_k(y)=0,$$

with solution

$$\widehat{u}_k(t)=A_k\mathrm{e}^{-Dk^2t}.$$

Now we sum over all such solutions to obtain the general solution. This amounts to taking the inverse Fourier transform:

$$u(x,t)=\int_{-\infty}^{\infty}\frac{\mathrm{d}k}{2\pi}\mathrm{e}^{\mathrm{i}kx}A_k\mathrm{e}^{-Dk^2t}.$$

Now we must fix the precise values of A_k. We have,

$$\begin{aligned}
u(x,t=0) &= \int_{-\infty}^{\infty}\frac{\mathrm{d}k}{2\pi}\mathrm{e}^{\mathrm{i}kx}A_k,\\
&= f(x),\\
&= \int_{-\infty}^{\infty}\frac{\mathrm{d}k}{2\pi}\mathrm{e}^{\mathrm{i}kx}\widehat{f}_k.
\end{aligned}$$

Matching up Fourier coefficients gives

$$\widehat{f}_k=A_k.$$

Thus, the final answer is

$$u(x,y) = \int_{-\infty}^{\infty} \frac{\mathrm{d}k}{2\pi} \mathrm{e}^{\mathrm{i}kx} \mathrm{e}^{-Dk^2 t} \widehat{f}_k. \qquad (*)$$

This is the final answer, although it can be manipulated into an integral over real space in the following manner:

- However, recall the result
$$(\widehat{\mathrm{e}^{-x^2/2\sigma^2}})_k = \sqrt{2\pi\sigma^2}\mathrm{e}^{-k^2\sigma^2/2}.$$
Identify $2\sigma^2 = 4Dt$, $\sigma^2 = 2Dt$, or $2\pi\sigma^2 = 4\pi Dt$. Thus,
$$(\widehat{\mathrm{e}^{-x^2/4Dt}})_k = \sqrt{4\pi Dt}\mathrm{e}^{-k^2 Dt},$$
or
$$\left(\widehat{\frac{1}{\sqrt{4\pi Dt}}\mathrm{e}^{-x^2/4Dt}}\right)_k = \mathrm{e}^{-k^2 Dt},$$
- Calling
$$K(x,t) := \frac{1}{\sqrt{4\pi Dt}}\mathrm{e}^{-x^2/4Dt},$$
we have
$$\widehat{K}_k(t) = \mathrm{e}^{-k^2 Dt}.$$
- Recall the convolution theorem:
$$\begin{aligned}(F * G)(x) &:= \int_{-\infty}^{\infty} F(y-x)G(y)\mathrm{d}y,\\ &= \int_{-\infty}^{\infty} \frac{\mathrm{d}k}{2\pi}\mathrm{e}^{\mathrm{i}kx}\widehat{F}_k\widehat{G}_k.\end{aligned}$$
- Applying this theorem to the result (*) gives

$$\begin{aligned}u(x,t) &= \int_{-\infty}^{\infty} K(y-x)f(y)\mathrm{d}y,\\ &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{4\pi Dt}}\mathrm{e}^{-(x-y)^2/4Dt}f(y)\mathrm{d}y.\end{aligned}$$

The function
$$K(s,t) := \frac{1}{\sqrt{4\pi Dt}}\mathrm{e}^{-s^2/4Dt}$$

is called the **heat kernel**.

Hot spot on an infinite rod Consider the diffusion equation on the line with initial condition

$$u(x, t = 0) = f(x) = \begin{cases} u_0/\epsilon, & |x| \le \epsilon/2, \\ 0, & |x| > \epsilon/2. \end{cases}$$

Let's apply the heat kernel and work out the resulting integrals:

$$\begin{aligned} u(x,t) &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{4\pi D t}} \mathrm{e}^{-(x-y)^2/4Dt} f(y)\mathrm{d}y, \\ &= \frac{u_0}{\epsilon} \int_{-\epsilon/2}^{\epsilon/2} \frac{1}{\sqrt{4\pi D t}} \mathrm{e}^{-(x-y)^2/4Dt} \mathrm{d}y. \end{aligned}$$

Let

$$z = \frac{y - x}{\sqrt{4Dt}}.$$

Then,

$$u(x,t) = \frac{u_0}{\epsilon} \frac{1}{\sqrt{\pi}} \int_{(\epsilon/2 - x)/\sqrt{4Dt}}^{(-\epsilon/2 - x)/\sqrt{4Dt}} \mathrm{e}^{-z^2} \mathrm{d}z.$$

Let

$$\mathrm{erf}\,(x) := \frac{2}{\sqrt{\pi}} \int_0^x \mathrm{e}^{-z^2} \mathrm{d}z.$$

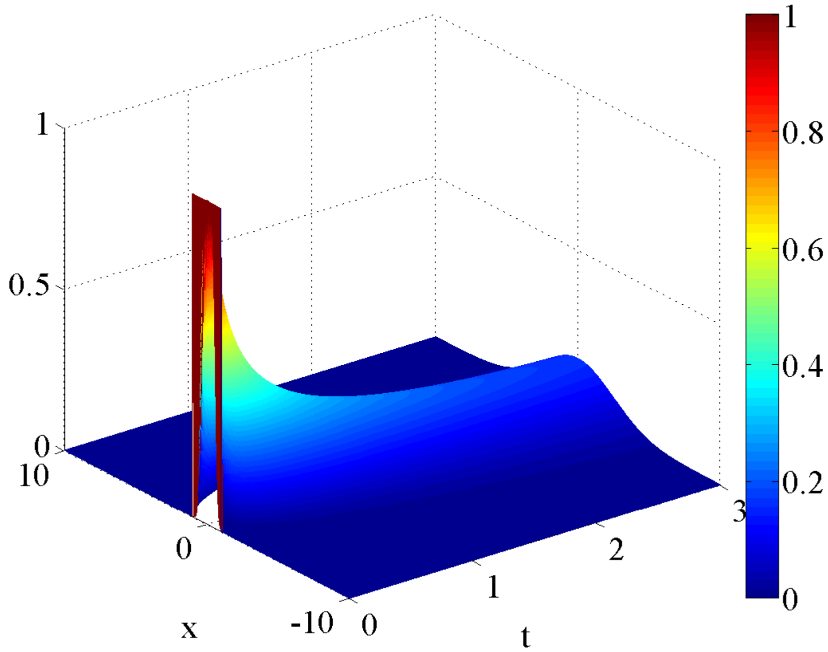
Thus,

$$\begin{aligned} u(x,t) &= \frac{u_0}{\epsilon} \frac{1}{\sqrt{\pi}} \int_{(\epsilon/2 - x)/\sqrt{4Dt}}^{(-\epsilon/2 - x)/\sqrt{4Dt}} \mathrm{e}^{-z^2} \mathrm{d}z, \\ &= \frac{u_0}{\epsilon} \frac{1}{\sqrt{\pi}} \left[\int_0^{(\epsilon/2 - x)/\sqrt{4Dt}} \mathrm{e}^{-z^2} \mathrm{d}z - \int_0^{(-\epsilon/2 - x)/\sqrt{4Dt}} \mathrm{e}^{-z^2} \mathrm{d}z \right], \\ &= \frac{u_0}{2\epsilon} \left[\mathrm{erf}\left(\frac{\epsilon/2 - x}{\sqrt{4Dt}} \right) - \mathrm{erf}\left(\frac{-\epsilon/2 - x}{\sqrt{4Dt}} \right) \right]. \end{aligned}$$

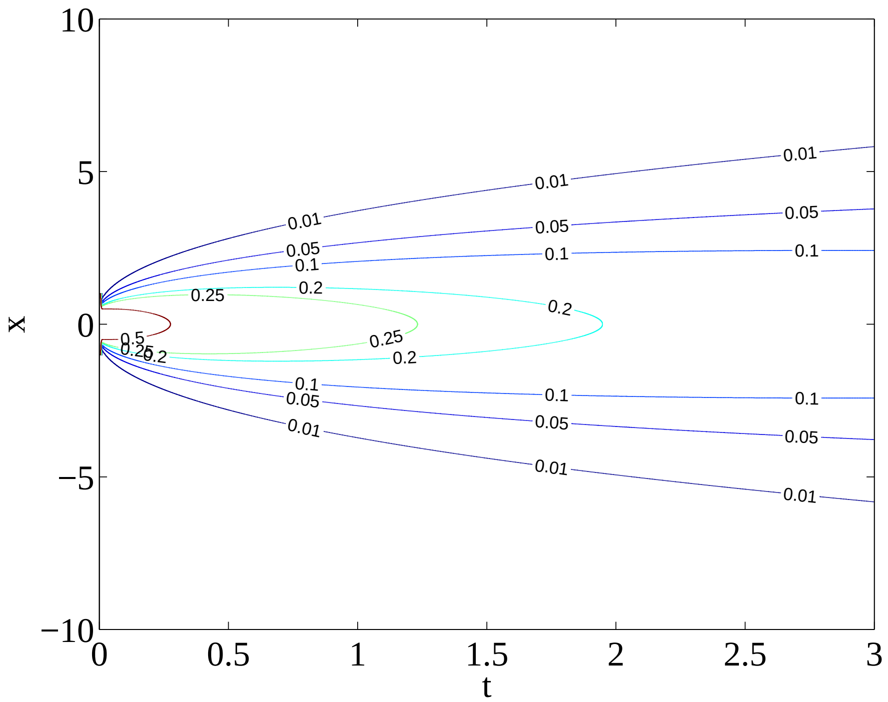
But erf is odd in its argument, so this is

$$u(x,t) = \frac{u_0}{2\epsilon} \left[\mathrm{erf}\left(\frac{x + \epsilon/2}{\sqrt{4Dt}} \right) - \mathrm{erf}\left(\frac{x - \epsilon/2}{\sqrt{4Dt}} \right) \right].$$

Some pictures of this solution are shown in Fig. 19.1.



(a) Spacetime plot of the solution



(b) Contour plot

Figure 19.1: The hot spot problem; all parameters set to unity

19.3 The diffusion equation on the half-line

Consider the diffusion equation

$$\begin{aligned} u_t &= Du_{xx}, \qquad 0 \le x < \infty, \\ u(x, t=0) &= f(x), \end{aligned}$$

where $u(x,t)$ and $f(x)$ decay rapidly as $x \to \infty$. We apply the Dirichlet BC

$$u(x=0, t>0) = 0.$$

To solve this problem, recall the solution on the whole line:

$$\begin{aligned} u_{WL}(x,t) &= \int_{-\infty}^{\infty} K(y-x) f(y)\, \mathrm{d}y, \\ &= \frac{1}{\sqrt{4\pi Dt}} \int_{-\infty}^{\infty} \mathrm{e}^{-\frac{(x-y)^2}{4Dt}} f(y)\, \mathrm{d}y, \end{aligned}$$

and

$$u_{WL}(0,t) \frac{1}{\sqrt{4\pi Dt}} \int_{-\infty}^{\infty} \mathrm{e}^{-\frac{y^2}{4Dt}} f(y)\, \mathrm{d}y.$$

If could somehow make $f(y)$ odd in y, then this integral would vanish, and the restriction of the whole-line solution to the half line would solve or problem. Thus, define

$$\widetilde{f}(y) = \begin{cases} f(y), & y > 0, \\ 0, & y = 0, \\ -f(-y), & y < 0. \end{cases}$$

Thus,

$$u(x,t) \frac{1}{\sqrt{4\pi Dt}} \int_{-\infty}^{\infty} \mathrm{e}^{-\frac{(x-y)^2}{4Dt}} \widetilde{f}(y)\, \mathrm{d}y$$

solves the heat equation on the half line $x \ge 0$ and solves the BC at zero.

A similar solution can be constructed for the homogeneous Neumann condition (exercise).

Chapter 20

Green's functions on infinite domains

20.1 Motivation: a systematic way of dealing with sources

There have been several instances so far where we have had to solve

$$\mathcal{L}u(x) = 0 + \text{inhomogeneous boundary conditions,}$$

or

$$\mathcal{L}u(x) = q(x) + \text{homogeneous boundary conditions,}$$

where $q(x)$ is a source term and $\mathcal{L}$ is some linear operator. Examples:

- The inhomogeneous diffusion equation in equilibrium:

$$0 = u_t = Du_{xx} + q(x);$$

- Laplace's equation on the half plane $x \in \mathbb{R}$ and $y > 0$:

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0, \qquad -\infty < x < \infty, \qquad y > 0,$$

  with BC

$$u(x, y = 0) = f(x).$$

In this section we outline a generic method to find the **particular integral** to the problem

$$\mathcal{L}u(x_1, \cdots, x_n) = q(x_1, \cdots, x_n).$$

This is called the **Green's function technique**.

20.2 Fourier transforms in n dimensions

Given a function $f(x_1, \cdots x_n)$ defined on $\mathbb{R}^n$, we define the FT of f in an obvious way:

$$\widehat{f}_{k_1,\cdots,k_n} = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \mathrm{d}x_1 \cdots \mathrm{d}x_n \mathrm{e}^{-\mathrm{i}k_1 x_1} \cdots \mathrm{e}^{-\mathrm{i}k_n x_n} f(x_1, \cdots, x_n).$$

That is, we do n copies of the Fourier transform, one copy on each real-space variable x_i. Sometimes we use vector notation for this operation:

$$\widehat{f}_{\boldsymbol{k}} = \int_{-\infty}^{\infty} \mathrm{d}^n x \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}} f(\boldsymbol{x}).$$

Because the transform-inverse-transform pair is valid for each variable x_i, an inverse FT exists for the function of several variables:

$$f(\boldsymbol{x}) = \int_{-\infty}^{\infty} \frac{\mathrm{d}^n k}{(2\pi)^n} \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}} \widehat{f}_{\boldsymbol{k}}$$

(NOTE THE SIGN!!).

The delta function in n dimensions

We define the delta function in $\mathbb{R}^n$ as follows:

$$\delta(\boldsymbol{x}) = \delta(x_1, \cdots, x_n) := \delta(x_1) \cdots \delta(x_n).$$

Thus, the n-dimensional delta function inherits the properties of the one-dimensional variant:

$$\int_{-\infty}^{\infty} \mathrm{d}^n x \, \delta(\boldsymbol{x}) = 1,$$

$$\int_{-\infty}^{\infty} \mathrm{d}^n x \, \delta(\boldsymbol{x} - \boldsymbol{a}) f(\boldsymbol{x}) = f(\boldsymbol{a}),$$

for any smooth function f. Thus,

$$\widehat{\delta}_{\boldsymbol{k}} = \int_{-\infty}^{\infty} \mathrm{d}^n x \, \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}} \delta(x) = \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot 0} = 1,$$

hence,

$$\begin{aligned}
\widehat{\delta}_{\boldsymbol{k}} &= 1, \\
\delta(\boldsymbol{x}) &= \int_{-\infty}^{\infty} \frac{\mathrm{d}^n k}{(2\pi)^n} (\widehat{\delta}_{\boldsymbol{k}} = 1) \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}, \\
&= \int_{-\infty}^{\infty} \frac{\mathrm{d}^n k}{(2\pi)^n} \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}.
\end{aligned}$$

20.3 The definition

Let $\mathcal{L}$ be a linear operator mapping smooth functions on $\mathbb{R}^n$ to $\mathbb{R}$. The **Green's function** associated with $\mathcal{L}$ is the function G which satisfies

$$\mathcal{L}G(\boldsymbol{x}) = \delta(\boldsymbol{x}),$$

such that

$$\lim_{|\boldsymbol{x}|\to\infty} G(\boldsymbol{x}) = 0.$$

The idea behind the definition

Why have we bothered with this definition? Well, it turns out that knowledge of the Green's function is sufficient to solve the problem

$$\mathcal{L}u(\boldsymbol{x}) = q(\boldsymbol{x}), \qquad \boldsymbol{x} \in \mathbb{R}^n.$$

For, consider

$$u(\boldsymbol{x}) = \int \mathrm{d}^n y\, G(\boldsymbol{x}-\boldsymbol{y}) q(\boldsymbol{y}).$$

Operate on both sides with the linear operator $\mathcal{L}$:

$$\begin{aligned}
\mathcal{L}_x u(\boldsymbol{x}) &= \mathcal{L}_x \int \mathrm{d}^n y\, G(\boldsymbol{x}-\boldsymbol{y}) q(\boldsymbol{y}), \\
&= \int \mathrm{d}^n y\, \mathcal{L}_x \left[G(\boldsymbol{x}-\boldsymbol{y}) q(\boldsymbol{y}) \right], \\
&= \int \mathrm{d}^n y\, \left[\mathcal{L}_{x-y} G(\boldsymbol{x}-\boldsymbol{y}) \right] q(\boldsymbol{y}), \\
&= \int \mathrm{d}^n y\, \delta(\boldsymbol{x}-\boldsymbol{y}) q(\boldsymbol{y}), \\
&= q(\boldsymbol{x}).
\end{aligned}$$

Thus, the full particular integral can be obtained from the Green's function by a simple convolution.

Example

The function

$$G(\boldsymbol{x}) = -\frac{1}{4\pi}\frac{1}{|\boldsymbol{x}|}$$

is the Green's function for the Laplacian in three dimensions.

For, let $\boldsymbol{x} \neq 0$. Then,

$$G(\boldsymbol{x}) = -\frac{1}{4\pi}\frac{1}{r}, \qquad r = |\boldsymbol{x}| = \sqrt{x^2+y^2+z^2}.$$

The function G is radially symmetric; the radially symmetric Laplace operator is

$$\nabla^2(\cdot) = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\cdot\right).$$

Hence,

$$\begin{aligned}
\nabla^2 G &= -\frac{1}{4\pi}\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\frac{1}{r}\right), \\
&= \frac{1}{4\pi}\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{1}{r^2}\right), \\
&= \frac{1}{4\pi}\frac{1}{r^2}\frac{\partial}{\partial r}(1), \\
&= 0.
\end{aligned}$$

Now let's integrate the $\nabla^2 G$ over a region including the origin. For definiteness, let's take the region to be a ball of radius R centred at 0; call it B.

$$\begin{aligned}
\int_B \nabla^2 G \, \mathrm{d}^3\boldsymbol{x} &= \int_B \nabla\cdot(\nabla G)\, \mathrm{d}^3\boldsymbol{x}, \\
&= \int_S \hat{\boldsymbol{n}}\cdot\nabla G \, \mathrm{d}S,
\end{aligned}$$

where

- S is the boundary of B – a sphere of radius R, centred at 0;
- $\boldsymbol{n}$ is the outward-pointing unit normal to the sphere;
- $\mathrm{d}S$ is an element of area.

This is an application of Gauss's theorem.

But we know these quantities:

- $\boldsymbol{n} = \hat{\boldsymbol{r}}$ is unit vector in the radial direction;
- $\mathrm{d}S = R^2 \sin\theta \, \mathrm{d}\theta \, \mathrm{d}\varphi$ is the element of area in the usual spherical polar coordinates.

Hence,

$$\hat{\boldsymbol{n}} \cdot \nabla G = \hat{\boldsymbol{r}} \cdot \nabla G = \frac{\partial G}{\partial r}.$$

So,

$$\begin{aligned}
\int_B \nabla^2 G \, \mathrm{d}^3\boldsymbol{x} &= \int_S \hat{\boldsymbol{n}} \cdot \nabla G \, \mathrm{d}S, \\
&= \int_S \mathrm{d}S \, \frac{\partial G}{\partial r}\bigg|_{r=R}, \\
&= -\frac{1}{4\pi} \int_S \mathrm{d}S \, \frac{\partial r^{-1}}{\partial r}\bigg|_{r=R}, \\
&= +\frac{1}{4\pi} \int_S R^2 \sin\theta \, \mathrm{d}\theta \, \mathrm{d}\varphi \frac{1}{R^2}, \\
&= \frac{1}{4\pi} \int_S \sin\theta \, \mathrm{d}\theta \, \mathrm{d}\varphi, \\
&= 1.
\end{aligned}$$

In summary,

- $\nabla^2 G = 0$, except at zero, where it is infinite;
- $\int_B \nabla^2 G = 1$, for any region $B \subset \mathbb{R}^n$ containing 0;
- $G \propto 1/r$ decays to zero as $r \to \infty$.

It follows that $\nabla^2 G$ is a delta function, or that

$$G = -\frac{1}{4\pi} \frac{1}{r}$$

is a Green's function for the Laplace operator **in three dimensions**.

Of course, this example is rather contrived – how did we have such excellent foresight in picking candidates for the delta function? In the next section, we develop a systematic way to compute Green's functions.

20.4 Green's functions by Fourier transform

Consider the problem

$$\nabla^2 G(\boldsymbol{x}) = \delta(\boldsymbol{x}), \qquad \boldsymbol{x} \in \mathbb{R}^3,$$

to be solved in **three dimensions**. Let's multiply both sides of the equation by $\mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}$ and integrate over space:

$$
\begin{aligned}
\int \mathrm{d}^3x\delta(\boldsymbol{x})\mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}} &= \int \mathrm{d}^3x\,\nabla^2 G(\boldsymbol{x})\mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}},\\
1 &= \int \mathrm{d}^3x\left[\nabla\cdot\left(\mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}\nabla G\right)+\mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}\mathrm{i}\boldsymbol{k}\cdot\nabla G\right],\\
&= \text{Vanishing boundary terms} + \int \mathrm{d}^3x\mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}\mathrm{i}\boldsymbol{k}\cdot\nabla G,\\
&= \int \mathrm{d}^3x\mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}\mathrm{i}\boldsymbol{k}\cdot\nabla G,\\
&= \mathrm{i}\boldsymbol{k}\cdot\int \mathrm{d}^3x\mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}\nabla G,\\
&= \mathrm{i}\boldsymbol{k}\cdot\int \mathrm{d}^3x\left[\nabla\left(\mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}G\right)+\mathrm{i}\boldsymbol{k}G\mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}\right],\\
&= \text{V.B.T.s} + (\mathrm{i}\boldsymbol{k})\cdot(\mathrm{i}\boldsymbol{k})\int \mathrm{d}^3x\mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}G(\boldsymbol{x}),\\
&= -k^2\widehat{G}_{\boldsymbol{k}}.
\end{aligned}
$$

Thus, we have solved for the Green's function, at least in Fourier space:

$$\widehat{G}_{\boldsymbol{k}} = -\frac{1}{k^2}.$$

But now we can re-construct the Green's function by the inverse transform:

$$
\begin{aligned}
G(\boldsymbol{x}) &= \int \frac{\mathrm{d}^3k}{(2\pi)^3}\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}\widehat{G}_{\boldsymbol{k}},\\
&= -\int \frac{\mathrm{d}^3k}{(2\pi)^3}\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}\frac{1}{k^2}.
\end{aligned}
$$

Thus, a hard problem, namely solving

$$\nabla^2 G(\boldsymbol{x}) = \delta(\boldsymbol{x})$$

has been reduced to (relatively) simple integration, namely

$$G(\boldsymbol{x}) = -\int \frac{\mathrm{d}^3k}{(2\pi)^3}\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}\frac{1}{k^2}.$$

20.5 The integral

We have seen the equivalence between solving

$$\nabla^2 G(\boldsymbol{x}) = \delta(\boldsymbol{x})$$

and simple integration, namely

$$G(\boldsymbol{x}) = -\int \frac{\mathrm{d}^3 k}{(2\pi)^3} \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}} \frac{1}{k^2}.$$

We now do this integral.

Three-dimensional space In three dimensions,

$$\begin{aligned} G(\boldsymbol{x}) &= -\int \frac{\mathrm{d}^3 k}{(2\pi)^3} \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}} \frac{1}{k^2}, \\ &= -\frac{1}{(2\pi)^3} \int_0^\infty k^2 \mathrm{d}k \int_0^{2\pi} \mathrm{d}\varphi \int_0^\pi \sin\theta \, \mathrm{d}\theta \, \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}} \frac{1}{k^2}, \end{aligned}$$

where we have introduced the standard spherical polar coordinates for $\boldsymbol{k}$-space. Now we choose a coordinate system such that the vector $\boldsymbol{x}$ aligns with the z-direction of $\boldsymbol{k}$-space. Thus,

$$\boldsymbol{k} \cdot \boldsymbol{x} = k|\boldsymbol{x}| \cos\theta := kr \cos\theta, \qquad r \geq 0,$$

and

$$\begin{aligned} G(\boldsymbol{x}) &= -\frac{1}{(2\pi)^3} \int_0^\infty k^2 \mathrm{d}k \int_0^{2\pi} \mathrm{d}\varphi \int_0^\pi \sin\theta \, \mathrm{d}\theta \, \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}} \frac{1}{k^2}, \\ &= -\frac{1}{(2\pi)^3} \int_0^\infty \mathrm{d}k \int_0^{2\pi} \mathrm{d}\varphi \int_0^\pi \sin\theta \, \mathrm{d}\theta \, \mathrm{e}^{\mathrm{i}kr\cos\theta}, \\ &= -\frac{1}{(2\pi)^2} \int_0^\infty \mathrm{d}k \int_0^\pi \sin\theta \, \mathrm{d}\theta \, \mathrm{e}^{\mathrm{i}kr\cos\theta}, \\ &= -\frac{1}{(2\pi)^2} \int_0^\infty \mathrm{d}k \int_0^\pi \left(\frac{-1}{\mathrm{i}kr}\right) \left(\frac{d}{d\theta} \mathrm{e}^{\mathrm{i}kr\cos\theta}\right) \mathrm{d}\theta, \\ &= -\frac{1}{(2\pi)^2} \int_0^\infty \mathrm{d}k \left(\frac{-1}{\mathrm{i}kr}\right) \left(\mathrm{e}^{\mathrm{i}kr\cos\theta}\right)_0^\pi, \\ &= -\frac{1}{(2\pi)^2} \int_0^\infty \mathrm{d}k \left(\frac{1}{\mathrm{i}kr}\right) \left(\mathrm{e}^{\mathrm{i}kr} - \mathrm{e}^{-\mathrm{i}kr}\right), \\ &= -\frac{2}{(2\pi)^2} \int_0^\infty \mathrm{d}k \frac{\sin(kr)}{kr}, \end{aligned}$$

$$
\begin{aligned}
G(\boldsymbol{x}) &= -\frac{2}{(2\pi)^2}\frac{1}{2}\int_{-\infty}^{\infty}\mathrm{d}(kr)\frac{1}{r}\frac{\sin(kr)}{kr},\\
&= -\frac{1}{(2\pi)^2}\frac{1}{r}\int_{-\infty}^{\infty}\mathrm{d}s\frac{\sin s}{s},\\
&= -\frac{1}{(2\pi)^2}\frac{1}{r}\pi,\\
&= -\frac{1}{4\pi^2}\frac{1}{r},
\end{aligned}
$$

as advertised previously.

Two-dimensional space: Remarkably, the problem is harder to solve in two dimensions. We have,

$$
\begin{aligned}
G(\boldsymbol{x}) &= -\int\frac{\mathrm{d}^2k}{(2\pi)^2}\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}\frac{1}{k^2},\\
&= -\frac{1}{(2\pi)^2}\int_0^{\infty}k\mathrm{d}k\int_0^{2\pi}\mathrm{d}\varphi\,\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}\frac{1}{k^2},
\end{aligned}
$$

where we have introduced the standard polar coordinates in the two-dimensional $\boldsymbol{k}$-space. Now we choose a coordinate system such that the vector $\boldsymbol{x}$ aligns with the y-direction (note!!) of $\boldsymbol{k}$-space. Thus,

$$
\boldsymbol{k}\cdot\boldsymbol{x} = k|\boldsymbol{x}|\sin\varphi := kr\sin\varphi,\qquad r=\sqrt{x^2+y^2}\geq 0,
$$

and

$$
\begin{aligned}
G(\boldsymbol{x}) &= -\frac{1}{(2\pi)^2}\int_0^{\infty}k\mathrm{d}k\int_0^{2\pi}\mathrm{d}\varphi\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}\frac{1}{k^2},\\
&= -\frac{1}{(2\pi)^2}\int_0^{\infty}\mathrm{d}k\frac{1}{k}\int_0^{2\pi}\mathrm{d}\varphi\,\mathrm{e}^{\mathrm{i}kr\sin\varphi}.
\end{aligned}
$$

But we have the identity

$$
\int_0^{2\pi}\mathrm{d}\varphi\,\mathrm{e}^{\mathrm{i}kr\sin\varphi} = 2\pi J_0(kr),
$$

where $J_0(\cdot)$ is the zeroth Bessel function. Thus,

$$
\begin{aligned}
G(\boldsymbol{x}) &= -\frac{1}{(2\pi)^2}\int_0^{\infty}\mathrm{d}k\frac{1}{k}\int_0^{2\pi}\mathrm{d}\varphi\,\mathrm{e}^{\mathrm{i}kr\sin\varphi},\\
&= -\frac{1}{2\pi}\int_0^{\infty}\mathrm{d}k\frac{1}{k}J_0(kr).
\end{aligned}
$$

We must be careful doing this integral because of the apparent singularity at $k = 0$. Let us take

$$\begin{aligned}
\frac{dG}{dr} &= -\frac{1}{2\pi}\frac{d}{dr}\int_0^\infty \mathrm{d}k\frac{1}{k}J_0(kr), \\
&= -\frac{1}{2\pi}\int_0^\infty \mathrm{d}k\,\frac{1}{k}\frac{d(kr)}{dr}\frac{d}{d(kr)}J_0(kr), \\
&= -\frac{1}{2\pi}\int_0^\infty \mathrm{d}k\,\frac{1}{k}kJ_0'(kr), \\
&= -\frac{1}{2\pi}\int_0^\infty \mathrm{d}k\,J_0'(kr), \\
&= -\frac{1}{2\pi}\frac{1}{r}\int_0^\infty \mathrm{d}(kr)J_0'(kr).
\end{aligned}$$

Now the integrand is totally regular, and we can do a change of variable:

$$\begin{aligned}
\frac{dG}{dr} &= -\frac{1}{2\pi}\frac{1}{r}\int_0^\infty \mathrm{d}(kr)J_0'(kr), \\
&= -\frac{1}{2\pi}\frac{1}{r}\int_0^\infty \mathrm{d}sJ_0'(s), \qquad s = kr, \\
&= -\frac{1}{2\pi}\frac{1}{r}\left[J_0(\infty) - J_0(0)\right], \\
&= -\frac{1}{2\pi}\frac{1}{r}\left[0 - 1\right],
\end{aligned}$$

Hence,

$$\frac{dG}{dr} = \frac{1}{2\pi r}.$$

Integrating once gives

$$G(r) = \frac{1}{2\pi}\log r,$$

where the constant of integration is set to zero because the Green's function must vanish as $r \to \infty$. Some important conclusions:

- The Green's function depends not only on the form of the linear operator, but also on the dimension of the space;
- The Green's function for the Laplacian does not exist in one dimension (i.e. for $(-\infty < x < \infty)$).
- Working out the Green's function in this way is not confined to the Laplacian; any linear operator can be studied in this way.
- For example, as homework, try to work out the Green's function for

$$\left(+\alpha^2 - \nabla^2\right) G(\boldsymbol{x}) = \delta(\boldsymbol{x})$$

in three dimensions. The answer should be

$$G(\boldsymbol{x}) = \frac{\mathrm{e}^{-\alpha|r|}}{4\pi r}, \qquad r = \sqrt{x^2 + y^2 + z^2}, \qquad \alpha \in \mathbb{R}.$$

This pair is called the **Helmholtz operator** and the **Helmholtz kernel**, respectively.

20.6 Table of Green's Functions

Dimension/Operator	Laplace, ∇^2	Helmholtz, $-\nabla^2 - \alpha^2$	Modified Helmholtz, $-\nabla^2 + \alpha^2$
1	No solution on $(-\infty, \infty)$	$\frac{\mathrm{i}}{2\alpha}\mathrm{e}^{\mathrm{i}\alpha|\boldsymbol{x}|}$	$\frac{1}{2\alpha}\mathrm{e}^{-\alpha|\boldsymbol{x}|}$
2	$\frac{1}{2\pi}\log|\boldsymbol{x}|$	$\frac{\mathrm{i}}{4}H_0^{(1)}(\alpha|\boldsymbol{x}|)$	$\frac{1}{2\pi}K_0(\alpha|\boldsymbol{x}|)$
3	$-\frac{1}{4\pi|\boldsymbol{x}|}$	$+\frac{\mathrm{e}^{\mathrm{i}\alpha|\boldsymbol{x}|}}{4\pi|\boldsymbol{x}|}$	$+\frac{\mathrm{e}^{-\alpha|\boldsymbol{x}|}}{4\pi|\boldsymbol{x}|}$

Chapter 21

Green's functions on domains with boundaries

Overview

In the last chapter, we solved the Green's function problem

$$\nabla^2 G_0(\boldsymbol{x}) = \delta(\boldsymbol{x}), \qquad \boldsymbol{x} \in \mathbb{R}^2 \text{ or } \mathbb{R}^3,$$

where the domain of the PDE is the full space. The solution is

$$G_0(\boldsymbol{x}) = -\int \frac{\mathrm{d}^n k}{(2\pi)^n} \frac{\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}}{k^2}.$$

In two and three dimensions $n = 2, 3$, the solution satisfies

$$\lim_{|\boldsymbol{x}|\to\infty} G_0(\boldsymbol{x}) = 0.$$

We also saw how to generalise the Green's function problem to arbitrary linear operators. In this chapter, we study a method to obtain the Green's function when the domain is compact ('finite').

21.1 Finite domains

For definiteness, in this chapter we work in two dimensions. Let Ω be a finite domain in $\mathbb{R}^2$, with smooth boundary $\partial\Omega$. We are interested in solving

$$\begin{aligned} \nabla_x^2 G(\boldsymbol{x};\boldsymbol{y}) &= \delta(\boldsymbol{x}-\boldsymbol{y}), \qquad \boldsymbol{x} \text{ in } \Omega, \\ G(\boldsymbol{x};\boldsymbol{y}) &= 0, \qquad \boldsymbol{x} \text{ in } \partial\Omega. \end{aligned}$$

- Now the function $G_0(\boldsymbol{x}-\boldsymbol{y})$ will satisfy the first of these criteria.

- To construct a G that satisfies both criteria, simply add a smooth function to G_0:

$$G(\boldsymbol{x};\boldsymbol{y}) = G_0(\boldsymbol{x}-\boldsymbol{y}) + h(\boldsymbol{x};\boldsymbol{y}).$$

- There are some conditions on h:

$$\begin{aligned} \nabla_x^2 h(\boldsymbol{x};\boldsymbol{y}) &= 0, \qquad \boldsymbol{x} \text{ in } \Omega, \\ G_0(\boldsymbol{x}-\boldsymbol{y}) + h(\boldsymbol{x};\boldsymbol{y}) &= 0, \qquad \boldsymbol{x} \text{ in } \partial\Omega. \end{aligned}$$

- Having constructed a Green's function on the full space, solving for the Green's function in the compact domain Ω is relatively straightforward - just add a function that satisfies $\nabla_x^2 h = 0$, together with some BCs.

Let's check that our solution makes sense. Consider

$$\begin{aligned} \nabla_x^2 u(\boldsymbol{x}) &= q(\boldsymbol{x}), \qquad \text{in } \Omega, \\ u(\boldsymbol{x}) &= 0, \qquad \text{on } \partial\Omega. \end{aligned}$$

and propose the solution

$$u(\boldsymbol{x}) = \int_\Omega \mathrm{d}^2 y\, G(\boldsymbol{x};\boldsymbol{y}) q(\boldsymbol{y}), \qquad G = G_0 + h.$$

Operate on both sides with ∇_x^2:

$$\begin{aligned}
\nabla_x^2 u(\boldsymbol{x}) &= \nabla_x^2 \int_\Omega \mathrm{d}^2 y\, G(\boldsymbol{x};\boldsymbol{y}) q(\boldsymbol{y}), \\
&= \int_\Omega \mathrm{d}^2 y\, \nabla_x^2 \left[G(\boldsymbol{x};\boldsymbol{y}) q(\boldsymbol{y}) \right], \\
&= \int_\Omega \mathrm{d}^2 y\, \left[\nabla_x^2 G(\boldsymbol{x};\boldsymbol{y}) \right] q(\boldsymbol{y}), \\
&= \int_\Omega \mathrm{d}^2 y\, \left[\nabla_x^2 G_0(\boldsymbol{x}-\boldsymbol{y}) + \nabla_x^2 h(\boldsymbol{x};\boldsymbol{y}) \right] q(\boldsymbol{y}), \\
&= \int_\Omega \mathrm{d}^2 y\, \left[\delta(\boldsymbol{x}-\boldsymbol{y}) + 0 \right] q(\boldsymbol{y}), \\
&= q(\boldsymbol{x}).
\end{aligned}$$

On the boundary,

$$\begin{aligned}
u(\boldsymbol{x} \in \partial\Omega) &= \int_\Omega \mathrm{d}^2 y\, G(\boldsymbol{x} \in \partial\Omega;\boldsymbol{y}) q(\boldsymbol{y}), \\
&= 0.
\end{aligned}$$

21.2 Example: The Laplacian on the half plane

Solve

$$\begin{aligned}
\nabla_x^2 G(\boldsymbol{x};\boldsymbol{y}) &= \delta(\boldsymbol{x}-\boldsymbol{y}), \qquad \boldsymbol{x} \text{ in } \Omega, \\
G(\boldsymbol{x};\boldsymbol{y}) &= 0, \qquad \boldsymbol{x} \text{ in } \partial\Omega,
\end{aligned}$$

where

$$\Omega = \{(x_1, x_2) | -\infty < x_1 < \infty,\, x_2 > 0\}.$$

We already know the fundamental Green's function:

$$G_0(\boldsymbol{x};\boldsymbol{y}) = G_0(\boldsymbol{x}-\boldsymbol{y}) = \frac{1}{2\pi} \log |\boldsymbol{x}-\boldsymbol{y}|.$$

Now we must solve

$$\begin{aligned}
\nabla_x^2 h(\boldsymbol{x};\boldsymbol{y}) &= 0, \qquad \boldsymbol{x} \in \Omega, \\
h(x_1, x_2 = 0;\boldsymbol{y}) &= -\frac{1}{2\pi} \log |(x_1, 0) - \boldsymbol{y}|.
\end{aligned}$$

We introduce an 'image point' $\tilde{\boldsymbol{x}} = (x_1, -x_2)$ and we try the solution

$$h(\boldsymbol{x};\boldsymbol{y}) = -\frac{1}{2\pi}\log|\tilde{\boldsymbol{x}}-\boldsymbol{y}|.$$

We have,

$$\begin{aligned}
\frac{\partial h}{\partial x_1} &= -\frac{1}{2\pi}\frac{1}{|\tilde{\boldsymbol{x}}-\boldsymbol{y}|}\frac{\partial}{\partial x_1}\left[(x_1-y_1)^2+(-x_2-y_2)^2\right]^{1/2},\\
&= -\frac{1}{4\pi}\frac{1}{|\tilde{\boldsymbol{x}}-\boldsymbol{y}|}\left[(x_1-y_1)^2+(-x_2-y_2)^2\right]^{-1/2}\left[2(x_1-y_1)\right],\\
&= -\frac{1}{2\pi}\frac{x_1-y_1}{(x_1-y_1)^2+(-x_2-y_2)^2};
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 h}{\partial x_1^2} &= -\frac{1}{2\pi}\frac{1}{(x_1-y_1)^2+(-x_2-y_2)^2}+\frac{1}{2\pi}\frac{x_1-y_1}{[(x_1-y_1)^2+(-x_2-y_2)^2]^2}\left[2(x_1-y_1)\right],\\
&= -\frac{1}{2\pi}\frac{1}{(x_1-y_1)^2+(-x_2-y_2)^2}+\frac{1}{2\pi}\frac{2(x_1-y_1)^2}{[(x_1-y_1)^2+(-x_2-y_2)^2]^2}.
\end{aligned}$$

Also,

$$\begin{aligned}
\frac{\partial h}{\partial x_2} &= -\frac{1}{2\pi}\frac{1}{|\tilde{\boldsymbol{x}}-\boldsymbol{y}|}\frac{\partial}{\partial x_2}\left[(x_1-y_1)^2+(-x_2-y_2)^2\right]^{1/2},\\
&= -\frac{1}{2\pi}\frac{1}{|\tilde{\boldsymbol{x}}-\boldsymbol{y}|}\frac{\partial}{\partial x_2}\left[(x_1-y_1)^2+(x_2+y_2)^2\right]^{1/2},\\
&= -\frac{1}{2\pi}\frac{1}{|\tilde{\boldsymbol{x}}-\boldsymbol{y}|}\left(\tfrac{1}{2}\frac{1}{[(x_1-y_1)^2+(x_2+y_2)^2]^{1/2}}\frac{\partial}{\partial x_2}(x_2+y_2)^2\right),\\
&= -\frac{1}{2\pi}\frac{x_2+y_2}{(x_1-y_1)^2+(x_2+y_2)^2};
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 h}{\partial x_2^2} &= -\frac{1}{2\pi}\frac{1}{(x_1-y_1)^2+(-x_2-y_2)^2}+\frac{1}{2\pi}\frac{x_2+y_2}{[(x_1-y_1)^2+(-x_2-y_2)^2]^2}\left[2(x_2+y_2)\right],\\
&= -\frac{1}{2\pi}\frac{1}{(x_1-y_1)^2+(x_2+y_2)^2}+\frac{1}{2\pi}\frac{2(x_2+y_2)^2}{[(x_1-y_1)^2+(x_2+y_2)^2]^2}.
\end{aligned}$$

Putting these together, we have

$$\begin{aligned}
\frac{\partial^2 h}{\partial x_1^2}+\frac{\partial^2 h}{\partial x_2^2} = &-\frac{1}{2\pi}\frac{1}{(x_1-y_1)^2+(-x_2-y_2)^2}+\frac{1}{2\pi}\frac{2(x_1-y_1)^2}{[(x_1-y_1)^2+(-x_2-y_2)^2]^2}\\
&-\frac{1}{2\pi}\frac{1}{(x_1-y_1)^2+(x_2+y_2)^2}+\frac{1}{2\pi}\frac{2(x_2+y_2)^2}{[(x_1-y_1)^2+(x_2+y_2)^2]^2},
\end{aligned}$$

or

$$\frac{\partial^2 h}{\partial x_1^2}+\frac{\partial^2 h}{\partial x_2^2}=-\frac{1}{2\pi}\frac{2}{(x_1-y_1)^2+(-x_2-y_2)^2}+\frac{1}{2\pi}\frac{2(x_1-y_1)^2+2(x_2+y_2)^2}{[(x_1-y_1)^2+(x_2+y_2)^2]^2}=0.$$

Note, moreover, that this calculation can be carried out everywhere in the domain Ω because the singularity is located at

$$(x_1-y_1)^2+(x_2+y_2)^2=0,$$

or

$$x_1=y_1,\qquad x_2=-y_2.$$

Provided $\boldsymbol{y}\in\Omega$ also, the singularity is outside of the domain Ω, and h is therefore entirely regular. Moreover, on the boundary,

$$h(x_1,0;\boldsymbol{y})=-\frac{1}{2\pi}\log\left[(x_1-y_1)^2+(y_2)^2\right]^{1/2},$$

while

$$G_0(x_1,0;\boldsymbol{y})=\frac{1}{2\pi}\log\left[(x_1-y_1)^2+(y_2)^2\right]^{1/2}=-h(x_1,0;\boldsymbol{y}).$$

Thus, we have found the Green's function for the upper half-plane, and the solution to the inhomogeneous problem is thus

$$\begin{aligned}u(x_1,x_2) &= \frac{1}{2\pi}\int_\Omega\left\{\log\left[(x_1-y_1)^2+(x_2-y_2)^2\right]^{1/2}-\log\left[(x_1-y_1)^2+(x_2+y_2)^2\right]^{1/2}\right\}q(\boldsymbol{y})\,\mathrm{d}^2y,\\ &= \frac{1}{4\pi}\int_\Omega\log\frac{(x_1-y_1)^2+(x_2-y_2)^2}{(x_1-y_1)^2+(x_2+y_2)^2}q(\boldsymbol{y})\,\mathrm{d}^2y,\end{aligned}$$

21.3 Inhomogeneous boundary conditions: Laplacian on the half plane

Suppose instead we are to solve

$$\begin{aligned}\mathcal{L} &:= \nabla^2,\\ \mathcal{L}u(\boldsymbol{x}) &= q(\boldsymbol{x}),\qquad \boldsymbol{x}\text{ in }\Omega,\\ u(\boldsymbol{x}) &= f(\boldsymbol{x}),\qquad \boldsymbol{x}\text{ in }\partial\Omega,\end{aligned}$$

on some domain Ω. We already know how to find the Green's function:

1. Construct the fundamental solution on the whole space by Fourier transforms;

2. Add a regular solution that solves $\mathcal{L}=0$ to soak up the boundary conditions.

3. Call the answer $G(\boldsymbol{x};\boldsymbol{y})$. Then,

$$\begin{aligned}\mathcal{L}_x G(\boldsymbol{x};\boldsymbol{y}) &= \delta(\boldsymbol{x}-\boldsymbol{y}), \qquad \boldsymbol{x} \text{ in } \Omega,\\ G(\boldsymbol{x};\boldsymbol{y}) &= 0, \qquad \boldsymbol{x} \text{ in } \partial\Omega,\end{aligned}$$

To tackle inhomogeneous problems, we propose the following convolution solution:

$$u(\boldsymbol{x}) = \int_\Omega G(\boldsymbol{x};\boldsymbol{y})q(\boldsymbol{y})\,\mathrm{d}^n y + \int_{\partial\Omega} f(\boldsymbol{y})\hat{\boldsymbol{n}}(\boldsymbol{y})\cdot\nabla_y G(\boldsymbol{x};\boldsymbol{y})\,\mathrm{d}S_y,$$

where Ω is some n-dimensional volume, $\hat{\boldsymbol{n}}(\boldsymbol{y})$ is the outward-pointing normal on the boundary $\partial\Omega$, and $\mathrm{d}S_y$ is an element of area. Let's check this ansatz. We work with $\boldsymbol{x}\in\Omega$:

$$\begin{aligned}\mathcal{L}_x u(\boldsymbol{x}) &= \mathcal{L}_x \int_\Omega G(\boldsymbol{x};\boldsymbol{y})q(\boldsymbol{y})\,\mathrm{d}^n y + \int_{\partial\Omega} \mathcal{L}_x f(\boldsymbol{y})\hat{\boldsymbol{n}}(\boldsymbol{y})\cdot\nabla_y G(\boldsymbol{x};\boldsymbol{y})\,\mathrm{d}S_y,\\ &= \int_\Omega \left[\mathcal{L}_x G(\boldsymbol{x};\boldsymbol{y})\right] q(\boldsymbol{y})\,\mathrm{d}^n y + \int_{\partial\Omega} f(\boldsymbol{y})\hat{\boldsymbol{n}}(\boldsymbol{y})\cdot\nabla\left[\mathcal{L}_x G(\boldsymbol{x};\boldsymbol{y})\right]\,\mathrm{d}S_y,\\ &= \int_\Omega \delta(\boldsymbol{x}-\boldsymbol{y})q(\boldsymbol{y})\,\mathrm{d}^n y + \int_{\partial\Omega} f(\boldsymbol{y})\hat{\boldsymbol{n}}(\boldsymbol{y})\cdot\nabla\delta(\boldsymbol{x}-\boldsymbol{y})\,\mathrm{d}S_y,\end{aligned}$$

We assume $\boldsymbol{x}\in\Omega$, hence, if $\boldsymbol{y}\in\partial\Omega$ it is impossble for $\boldsymbol{x}-\boldsymbol{y}=0$, since a boundary point and an interior point cannot coincide. Thus, $\delta(\boldsymbol{x}-\boldsymbol{y})=0$ in the second integral, and

$$\begin{aligned}\mathcal{L}_x u(\boldsymbol{x}) &= \int_\Omega \delta(\boldsymbol{x}-\boldsymbol{y})q(\boldsymbol{y})\,\mathrm{d}^n y + 0,\\ &= q(\boldsymbol{x}).\end{aligned}$$

Moreover, consider again

$$u(\boldsymbol{x}) = \int_\Omega G(\boldsymbol{x};\boldsymbol{y})q(\boldsymbol{y})\,\mathrm{d}^n y + \int_{\partial\Omega} f(\boldsymbol{y})\hat{\boldsymbol{n}}(\boldsymbol{y})\cdot\nabla_y G(\boldsymbol{x};\boldsymbol{y})\,\mathrm{d}S_y,$$

with

$$\boldsymbol{x}\in\partial\Omega.$$

By construction,

$$G(\boldsymbol{x}\in\partial\Omega;\boldsymbol{y}) = 0,$$

hence

$$u(\boldsymbol{x}\in\partial\Omega) = \int_{\partial\Omega} f(\boldsymbol{y})\hat{\boldsymbol{n}}(\boldsymbol{y})\cdot\nabla_y G(\boldsymbol{x};\boldsymbol{y})\,\mathrm{d}S_y.$$

But we have,

$$\begin{aligned}
\int_{\partial\Omega} f(\boldsymbol{y})\hat{\boldsymbol{n}}(\boldsymbol{y})\cdot\nabla_y G(\boldsymbol{x};\boldsymbol{y})\,\mathrm{d}S_y &= \int_{\partial\Omega} f(\boldsymbol{y})\hat{\boldsymbol{n}}(\boldsymbol{y})\cdot\nabla_y G(\boldsymbol{x};\boldsymbol{y})\,\mathrm{d}S_y - \underbrace{\int_{\partial\Omega} G(\boldsymbol{x};\boldsymbol{y})\hat{\boldsymbol{n}}(\boldsymbol{y})\cdot\nabla_y f(\boldsymbol{y})\,\mathrm{d}S_y}_{=0,\ \boldsymbol{x}\in\partial\Omega}, \\
&= \int_{\partial\Omega} \hat{\boldsymbol{n}}(\boldsymbol{y})\cdot\left[f(\boldsymbol{y})\nabla_y G(\boldsymbol{x};\boldsymbol{y}) - G(\boldsymbol{x};\boldsymbol{y})\nabla_y f(\boldsymbol{y})\right]\mathrm{d}S_y, \\
&= \int_{\Omega} \nabla_y\cdot\left[f(\boldsymbol{y})\nabla_y G(\boldsymbol{x};\boldsymbol{y}) - G(\boldsymbol{x};\boldsymbol{y})\nabla_y f(\boldsymbol{y})\right]\mathrm{d}^n y, \\
&\qquad\qquad \text{(by Gauss's theorem)} \\
&= \int_{\Omega} \left[f(\boldsymbol{y})\nabla_y^2 G(\boldsymbol{x};\boldsymbol{y}) - G(\boldsymbol{x};\boldsymbol{y})\nabla_y^2 f(\boldsymbol{y})\right]\mathrm{d}^n y, \\
&= \int_{\Omega} \left[f(\boldsymbol{y})\nabla_y^2 G(\boldsymbol{x};\boldsymbol{y}) - 0\times\nabla_y^2 f(\boldsymbol{y})\right]\mathrm{d}^n y, \\
&= \int_{\Omega} \left\{f(\boldsymbol{y})\left[\nabla_y^2 G_0(\boldsymbol{x};\boldsymbol{y}) + \nabla_y^2 h(\boldsymbol{x};\boldsymbol{y})\right]\right\}\mathrm{d}^n y, \\
&= \int_{\Omega} \left\{f(\boldsymbol{y})\left[\nabla_y^2 G_0(\boldsymbol{x};\boldsymbol{y}) + 0\right]\right\}\mathrm{d}^n y, \\
&= \int_{\Omega} \left[f(\boldsymbol{y})\delta(\boldsymbol{x}-\boldsymbol{y})\right]\mathrm{d}^n y, \\
&= f(\boldsymbol{y}).
\end{aligned}$$

Now we solve

$$\begin{aligned}
\nabla_x^2 u(\boldsymbol{x}) &= 0, \qquad \boldsymbol{x} \text{ in } \Omega, \\
u(\boldsymbol{x}) &= f(\boldsymbol{x}), \qquad \boldsymbol{x} \text{ in } \partial\Omega,
\end{aligned}$$

where

$$\Omega = \{(x_1, x_2) | -\infty < x_1 < \infty,\ x_2 > 0\}.$$

We need to solve for the Green's function

$$\begin{aligned}
\nabla_x^2 G(\boldsymbol{x};\boldsymbol{y}) &= \delta(\boldsymbol{x}-\boldsymbol{y}), \qquad \boldsymbol{x} \text{ in } \Omega, \\
G(\boldsymbol{x};\boldsymbol{y}) &= 0, \qquad \boldsymbol{x} \text{ in } \partial\Omega.
\end{aligned}$$

But we already know this: the fundamental solution is

$$G_0(\boldsymbol{x};\boldsymbol{y}) = G_0(\boldsymbol{x}-\boldsymbol{y}) = \frac{1}{2\pi}\log|\boldsymbol{x}-\boldsymbol{y}|,$$

and the bit that soaks up the BCs is

$$h(\boldsymbol{x};\boldsymbol{y}) = -\frac{1}{2\pi}\log\left[(x_1-y_1)^2 + (x_2+y_2)^2\right]^{1/2},$$

The required solution is thus

$$u(\boldsymbol{x}) = -\frac{1}{2\pi}\int_{y_1=-\infty}^{y_1=\infty} \mathrm{d}y_1 (-1)\frac{\partial}{\partial y_2}\log\left[(x_1-y_1)^2+(x_2+y_2)^2\right]^{1/2}\bigg|_{y_2=0} f(y_1).$$

Doing the differentiation, we have

$$\begin{aligned} u(x_1,x_2) &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{d}y_1 \frac{2x_2 f(y_1)}{(x_1-y_1)^2+x_2^2}, \\ &= \frac{x_2}{\pi}\int_{-\infty}^{\infty}\mathrm{d}y_1 \frac{f(y_1)}{(x_1-y_1)^2+x_2^2}. \end{aligned}$$

However, if we treat x_2 as a mere parameter, this is a convolution:

$$u(x_1,\cdot) = (F * f)(x_1),$$

where

$$F(s,\cdot) = \frac{1}{\pi}\frac{x_2}{s^2+x_2^2}.$$

In other words,

$$u(x_1;\cdot) = \int_{-\infty}^{\infty} F(s-y_1)f(y_1)\,\mathrm{d}y_1.$$

But by the convolution theorem, this is

$$u(x_1;\cdot) = \int_{-\infty}^{\infty}\frac{\mathrm{d}k}{2\pi}\widehat{F}_k\widehat{f}_k \mathrm{e}^{ikx_1},$$

where

$$\begin{aligned} \widehat{F}_k &= \int_{-\infty}^{\infty}\mathrm{d}s \mathrm{e}^{-\mathrm{i}ks}F(s), \\ &= \int_{-\infty}^{\infty}\mathrm{d}s\,\mathrm{e}^{-\mathrm{i}ks}\frac{x_2}{\pi}\frac{1}{s^2+x_2^2}. \end{aligned}$$

Let's do the contour integral. For example, take $k < 0$. Then, we must close the semicircular contour in the upper half-plane, and integrate in an anticlockwise sense around the contour. The only singularity is a simple pole at $s = \mathrm{i}x_2$. Thus, the only residue is at this point, and

$$\begin{aligned}
\mathsf{Res}(\mathrm{i}x_2) &= \int_{\mathsf{Circle}} \left[\mathrm{e}^{-\mathrm{i}ks}\frac{1}{\pi}\frac{x_2}{s^2+x_2^2}\mathrm{d}s\right]_{s=\mathrm{i}x_2+\epsilon\mathrm{e}^{\mathrm{i}\theta}}, \\
&= \frac{x_2}{\pi}\int_0^{2\pi}\frac{\mathrm{e}^{-\mathrm{i}k(\mathrm{i}x_2+\epsilon\mathrm{e}^{\mathrm{i}\theta})}}{2\mathrm{i}x_2\epsilon\mathrm{e}^{\mathrm{i}\theta}+\epsilon^2\mathrm{e}^{2\mathrm{i}\theta}}\mathrm{i}\epsilon\mathrm{e}^{\mathrm{i}\theta}\,\mathrm{d}\theta, \\
&= \frac{x_2}{\pi}\int_0^{2\pi}\left[\mathrm{e}^{kx_2}+O(\epsilon)\right]\mathrm{d}\theta, \\
&= \frac{x_2}{\pi}\frac{\mathrm{e}^{kx_2}}{2x_2}2\pi, \qquad \epsilon\to 0, \\
&= \mathrm{e}^{-|k|x_2}.
\end{aligned}$$

Similarly for the upper half-plane. Thus,

$$\widehat{F}_k = \mathrm{e}^{-|k|x_2}$$

and

$$\begin{aligned}
u(x_1;\cdot) &= \int_{-\infty}^{\infty}\frac{\mathrm{d}k}{2\pi}\widehat{F}_k\widehat{f}_k\mathrm{e}^{ikx_1}, \\
u(x_1;x_2) &= \int_{-\infty}^{\infty}\frac{\mathrm{d}k}{2\pi}\mathrm{e}^{-|k|x_2}\widehat{f}_k\mathrm{e}^{\mathrm{i}kx_1},
\end{aligned}$$

but this is precisely the result obtained by FT methods, in Ch. 19.

Chapter 22

The 1-D linear advection equation

Overview

The advection equation describes the transport of a scalar $u(x,t)$ by a prescribed velocity field $c(x,t)$:

$$\frac{\partial u}{\partial t} + c(x,t)\frac{\partial u}{\partial x} = 0.$$

Topics: discussion of $c =$ Const.; solution in the general case; derivation. **In this chapter, the equation is to be solved on the real line, unless otherwise stated.**

22.1 Constant velocity

We consider the equation

$$\frac{\partial u}{\partial t} + c\frac{\partial u}{\partial x} = 0, \qquad c \in \mathbb{R} > 0.$$

There are three ways of solving this.

1. **A clever ansatz:** This is the most contrived way. We take

$$u(x,t) = f(x - ct)$$

as an ansatz, and verify that it satisfies the linear advection equation:

$$\begin{aligned}
\frac{\partial u}{\partial t} &= \frac{\partial \eta}{\partial t}\frac{df(\eta)}{d\eta}, \qquad \eta = x - ct, \\
&= -cf'(\eta), \\
\frac{\partial u}{\partial x} &= \frac{\partial \eta}{\partial x}\frac{df(\eta)}{d\eta}, \\
&= +f'(\eta).
\end{aligned}$$

Adding them,

$$u_t + cu_x = -cf'(\eta) + cf'(\eta) = 0.$$

Note that

$$u(x, t = 0) = f(x)$$

is the initial condition.

2. **By consideration of the second-order wave equation:** Recall the **second-order** wave equation

$$\frac{\partial^2 u}{\partial t^2} - c^2\frac{\partial^2 u}{\partial x^2} = 0.$$

We solved this by changing coordinates:

$$\frac{\partial}{\partial \eta}\frac{\partial u}{\partial \xi} = 0,$$

where

$$\eta = x - ct, \qquad \xi = x + ct.$$

which gives a solution

$$\begin{aligned}
u(x,t) &= [\text{Left-moving wave}] + [\text{Right-moving wave}] + [\text{part depending on initial velocity}], \\
&= \tfrac{1}{2}f(x+ct) + \tfrac{1}{2}f(x-ct) + [\text{part depending on initial velocity}],
\end{aligned}$$

where $u(x, t = 0) = f(x)$ is the initial condition.

Note, however, the form of the wave equation again:

$$\left(\frac{\partial}{\partial t} - c\frac{\partial}{\partial x}\right)\left(\frac{\partial}{\partial t} + c\frac{\partial}{\partial x}\right)u(x,t) = 0.$$

Thus, solutions to the linear advection equation are solutions of

$$\left(\frac{\partial}{\partial t} + c\frac{\partial}{\partial x}\right)u = 0,$$

and are therefore right-moving solutions of the second-order wave equation. In other words,

$$\left(\frac{\partial}{\partial t} + c\frac{\partial}{\partial x}\right) u = 0 \implies u(x,t) = f(x - ct).$$

3. **By physical intuition:**

   - If you are familiar with fluid mechanics, you will know what advection is.
   - It is the idea that material property, such as the concentration of a pollutant, is carried with the flow.
   - Thus, a patch of pollutant, of density u, and initially located at x_0, will, at a later time, retain its density, but move to a new location $x = x_0 + ct$.
   - In other words,
     $$u(x_0, 0) = u(x = x_0 + ct, t).$$
   - A solution to this equation is
     $$u(x,t) = f(x - ct).$$

     Check:
     $$\begin{aligned} u(x_0, 0) &= f(x_0), \\ u(x_0 + ct, t) &= f(x_0 + ct - ct) = f(x_0) = u(x_0, 0). \end{aligned}$$

     Thus, the statement that the density u is conserved along trajectories $dx/dt = c$ is equivalent to the advection equation,
     $$\text{density conserved along trajectories} \iff \frac{\partial u}{\partial t} + c\frac{\partial u}{\partial x} = 0.$$

To understand the next section, it will be helpful to be conversant with all three interpretations.

22.2 Velocities depending on space and time

Consider the problem
$$\frac{\partial u}{\partial t} + c(x,t)\frac{\partial u}{\partial x} = 0.$$

- We can give $c(x,t)$ the interpretation of a fluid velocity: $dx/dt = c(x,t)$
- Also, interpret $u(x,t)$ as the density of a pollutant.

- Thus, a patch of pollutant, of density u, and initially located at x_0, will, at a later time, retain its density, but move to a new location given by the solution of

$$\frac{dx}{dt} = c(x,t), \qquad x(0) = x_0.$$

In most cases [1], this equation will have a solution. We write this in generic form:

$$x(t) = x(t; x_0),$$

where the x_0 on the RHS indicates the parametric dependence on the initial condition. Suppose we can invert this identity, to give

$$x_0 = \eta(x,t),$$

Then, the solution to the advection equation is

$$u(x,t) = f(\eta(x,t)),$$

where $f(x)$ is the initial data.

- Let's check if this makes sense.

$$\begin{aligned}\frac{\partial u}{\partial t} &= f'(\eta)\frac{\partial \eta}{\partial t} = \eta_t f'(\eta),\\ \frac{\partial u}{\partial x} &= f'(\eta)\frac{\partial \eta}{\partial t} = \eta_x f'(\eta).\end{aligned}$$

Add:

$$\frac{\partial u}{\partial t} + c(x,t)\frac{\partial u}{\partial x} = f'(\eta)\left[\eta_t + c(x,t)\eta_x\right] = 0.$$

But $x_0 = \eta(x,t) =$ Const. Differentiate w.r.t. time:

$$\begin{aligned}0 &= \frac{dx_0}{dt} = \eta_t + \eta_x\frac{dx}{dt},\\ &= \eta_t + c(x,t)\eta_x,\end{aligned}$$

and the trial solution works.

[1] That is, if $c(x,t)$ has the Lipschitz condition

22.3 Examples

Linear velocity field

Solve the equation

$$\frac{\partial u}{\partial t} - ax\frac{\partial u}{\partial x} = 0.$$

We first of all solve the **characteristic equation** $dx/dt = c(x,t)$, or

$$\frac{dx}{dt} = -ax$$

Separating the variables gives

$$\int_{x_0}^{x} \frac{dx}{x} = -at.$$

Doing the integration gives

$$x = x_0 e^{-at}.$$

hence

$$x_0 = xe^{at}$$

or

$$x_0 = xe^{at} = \eta(x,t).$$

The solution is

$$\begin{aligned} u(x_0, t=0) &= u(x,t), \\ &= f(x_0). \end{aligned}$$

But $x_0 = xe^{at} = \eta(x,t)$, hence

$$u(x,t) = f\left(xe^{at}\right).$$

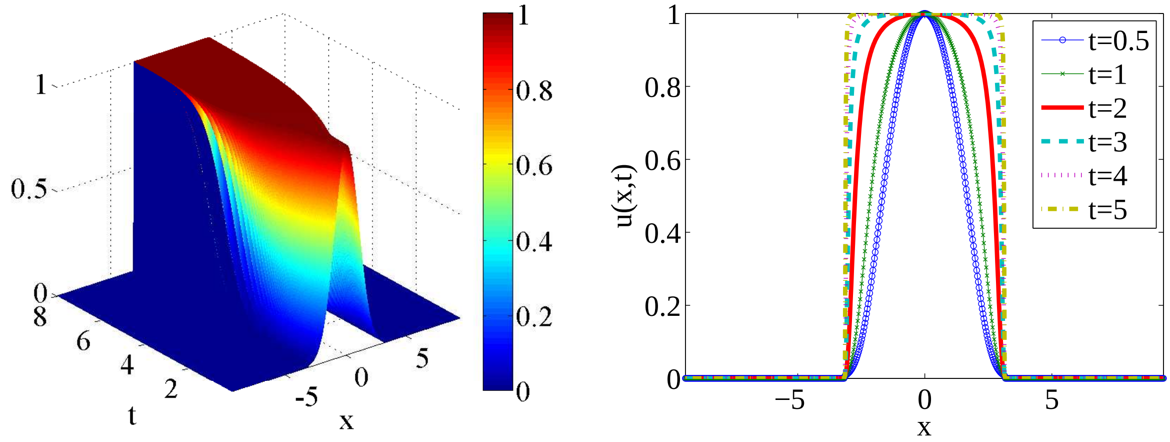
Sinusoidal velocity field

Solve the equation

$$\frac{\partial u}{\partial t} + a\sin(x)\frac{\partial u}{\partial x} = 0.$$

We first of all solve the **characteristic equation**

$$\frac{dx}{dt} = a\sin(x).$$



(a) (b)

Figure 22.1: Advection with the flow $c(x,t) = a\sin(x)$. I have set $a = 1$.

Separating the variables gives

$$\int_{x_0}^{x} \frac{dx}{\sin(x)} = at.$$

Doing the integration gives

$$\log\tan(x/2)\Big|_{x_0}^{x} = at,$$

hence

$$\log\tan(x/2) = \log\tan(x_0/2) + at,$$

and

$$\log\tan(x_0/2) = \log\tan(x/2) - at.$$

Thus, we have

$$\begin{aligned} \tan(x_0/2) &= \tan(x/2)e^{-at}, \\ x_0 &= \arctan\left(\tan(x/2)e^{-at}\right) := \eta(x,t). \end{aligned}$$

The solution is

$$u(x_0, t=0) = f(x_0) = u(x,t).$$

But $x_0 = \arctan\left(\tan(x/2)e^{-at}\right) := \eta(x,t)$, hence

$$u(x,t) = f\left[\arctan\left(\tan(x/2)e^{-at}\right)\right]. \qquad (*)$$

I have plotted the solution in Fig. 22.1, with $a = 1$ and the ICs

$$f(x) = \begin{cases} e^{-x^2/(2\times 0.5^2)}, & |x| < \pi, \\ 0, & |x| > \pi. \end{cases}$$

Something very troubling happens, which we can describe by studying the equation

$$dx/dt = a\sin(x).$$

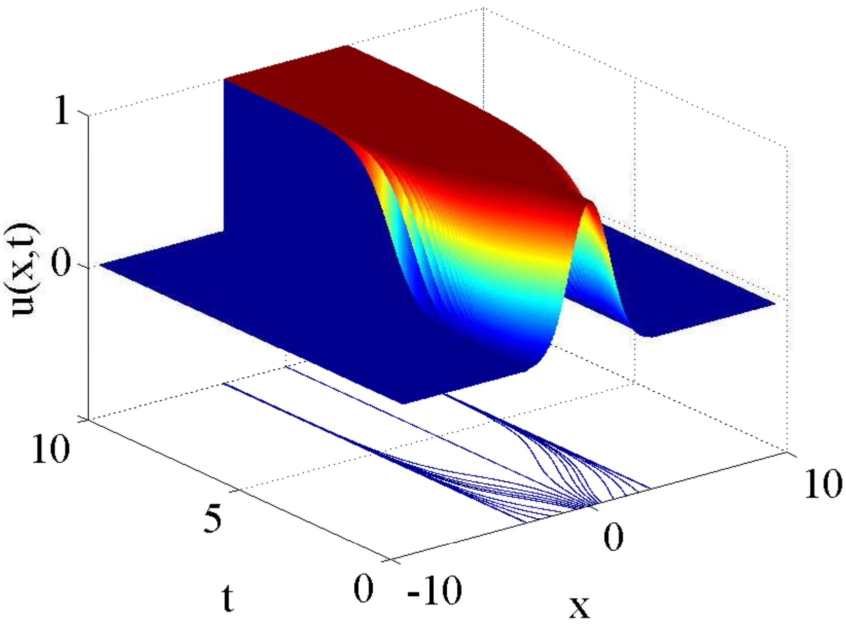
- This equation has fixed points (equilibria) at $x = 0$, $x = \pm\pi$.
- The point $x = 0$ is an unstable fixed point, because in the neighbourhood of this point, $x = 0 + \delta$ the equation is $d\delta/dt = \delta$.
- The points $x = \pm\pi$ are stable fixed points, because in the neighbourhood $x = \pi + \delta$ of these points, the equation is $d\delta/dt = -\delta$.
- Thus, an initial condition initially near $x = 0$ are repelled from there, and are attracted to $x = \pm\pi$.
- Our initial condition for the concentration has a maximum near $x = 0$. This corresponds to a collection of particles clustered here.
- The particles are repelled from zero and end up at $x = \pm\pi$. This corresponds to matter'piling up' at the end points. No matter can pass the end points because $\dot{x} = 0$ there.
- The piling up of matter produces at $x = \pm\pi$ leads to a steepening in the solution.
- Finally, and with reference to the solution (*), the asymptotic ($t \to \infty$) solution is

  $$u(x,t) \sim f(\arctan(0)) = \text{Const.} = 1,$$

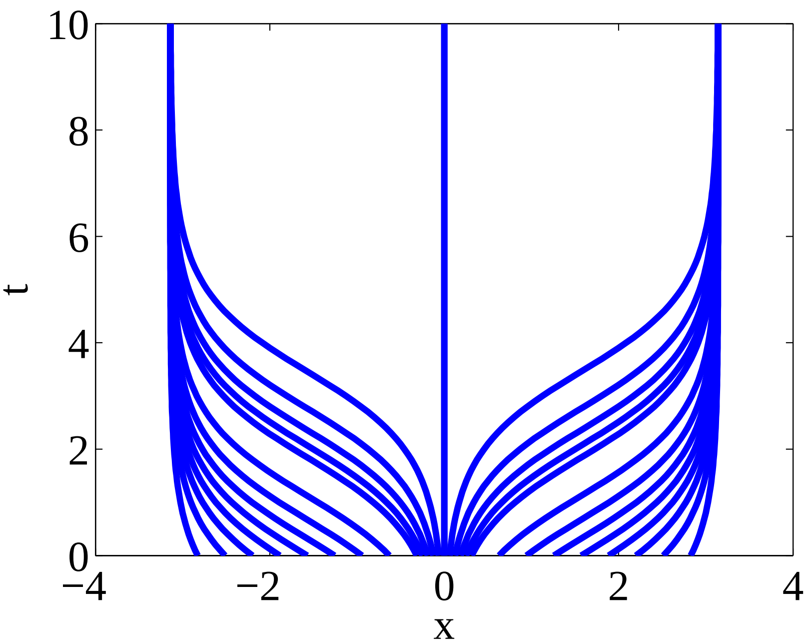
  unless we are outside $|x| < \pi$, where it is

  $$u(x,t) = 0.$$

  Thus, the problem develops a discontinuity in infinite time. This is what I called 'troubling'. Later on, we shall investigate a more worrying phenomenon, namely a discontinuity in finite time – even more troubling.



(a)



(b)

Figure 22.2: Characteristics for the problem $dx/dt = a \sin x$, with $a = 1$.

One last thing

By construction, $u =$ const on $dx/dt = c(x) = a\sin(x)$. For, let

$$v(t) := u(x(t), x), \qquad dx/dt = c(x).$$

Differentiate w.r.t. time:

$$\begin{aligned}\frac{dv}{dt} &= \frac{\partial u}{\partial x}\frac{dx}{dt} + \frac{\partial u}{\partial t}, \\ &= \frac{\partial u}{\partial x}c(x) + \frac{\partial u}{\partial t}, \\ &= 0.\end{aligned}$$

In other words, contours of $u(x,t)$ plotted in spacetime satisfy $dx/dt = c(x,t)$. This is shown in Fig. 22.2:

- Figure 22.2 (a) shows the contours $u(x,t) =$ const. in spacetime.
- Along $t = 0$ (the x-axis), the contours intersect the x-axis points I have chosen:

$$\begin{aligned}x_0 = \{&0, \pm 0.0628, \pm 0.1257, \pm 0.1885, \pm 0.2513, \pm 0.3142, \pm 0.6283, \pm 0.9425, \\ &\pm 1.2566, \pm 1.5708, \pm 1.8850, \pm 2.1991, \pm 2.5133, \pm 2.8274\}. \qquad (*)\end{aligned}$$

- Figure 22.2 (b) shows the solutions of $dx/dt = \sin(x)$ for the initial data (*).
- The curves in (a) and (b) are exactly the same: $u =$ const along $dx/dt = c(x)$.

The implicit solution

$$u = \text{const along } dx/dt = c(x)$$

is called the **solution by the method of characteristics**.

22.4 The continuity equation

In this section we give a physical motivation for the linear advection equation **with constant velocity**.

Let $u(x,t)$ be the concentration (mass/area) of a pollutant contained in an infinitely long channel

$$-\infty < x < \infty, \qquad 0 < y < L.$$

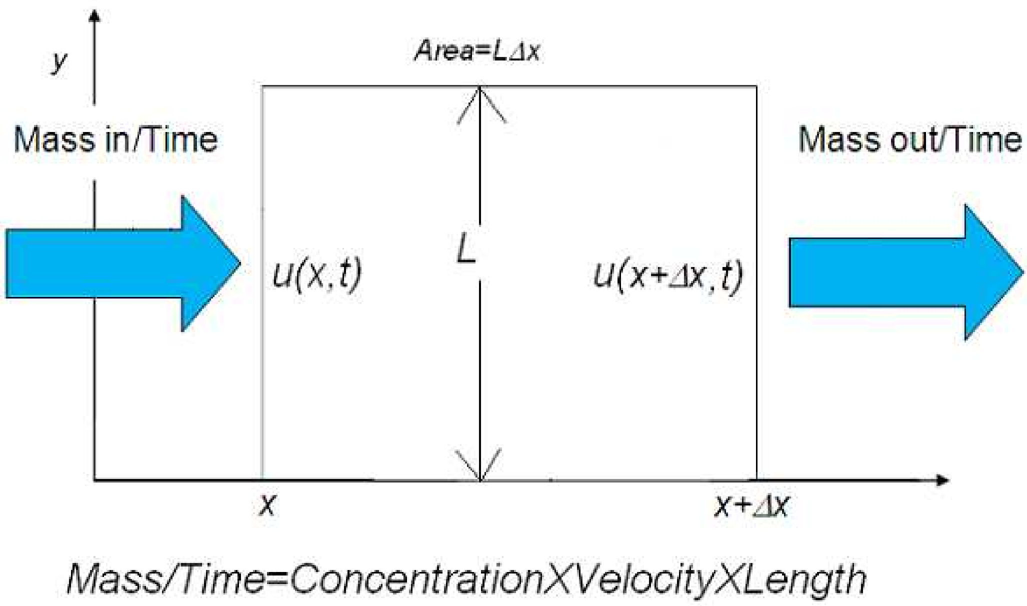


Figure 22.3: One-dimensional continuity equation

Let $c > 0$ be the constant velocity of flow, such that the pollutant is carried along the channel in the x-direction. Furthermore, consider a control area

$$L\Delta x,$$

located at x (See Fig. 22.3). The amount of matter entering through x in a a time interval Δt is

$$m_{\text{in}} = u(x)Lc\Delta t.$$

Similarly, the amount of matter leaving the control patch through $x + \Delta x$ in a a time interval Δt is

$$m_{\text{out}} = u(x + \Delta x)Lc\Delta t.$$

The net increase in the amount of matter is thus

$$\begin{aligned} \Delta m &= m_{\text{in}} - m_{\text{out}} = u(x)Lc\Delta t - u(x + \Delta x)Lc\Delta t, \\ &\approx -c\frac{\partial u}{\partial x}L\Delta x\Delta t. \end{aligned}$$

Dividing by Δt, this is

$$\frac{\Delta m}{\Delta t} = -c\frac{\partial u}{\partial x}L\Delta x.$$

However, the instantaneous amount of matter in the control patch is

$$m(t) = \int_{\text{patch}} u(x,t) L \mathrm{d}x \approx u(x,t) L \Delta x,$$

hence

$$\frac{\Delta m}{\Delta t} \approx \frac{\partial u(x,t)}{\partial t} L \Delta x.$$

Equating the two formulas gives

$$\frac{\partial u}{\partial t} = -c \frac{\partial u}{\partial x},$$

which is the linear advection equation with constant velocity. Notes:

- For appropriate boundary conditions, the total mass

  $$M = \int_{-\infty}^{\infty} u(x,t) \mathrm{d}x$$

  is conserved:

  $$\begin{aligned} \frac{dM}{dt} &= \frac{d}{dt} \int_{-\infty}^{\infty} u(x,t) \mathrm{d}x, \\ &= \int_{-\infty}^{\infty} \frac{\partial u(x,t)}{\partial t} \mathrm{d}x, \\ &= -c \int_{-\infty}^{\infty} \frac{\partial u(x,t)}{\partial x} \mathrm{d}x, \\ &= -c \left[u(\infty,t) - u(-\infty,t) \right], \\ &= 0, \end{aligned}$$

  if the concentration at both ends is zero.

- In one dimension, this physical interpretation of the one-dimensional advection equation breaks down in c depends on space (try to see why).

Chapter 23

Burgers' equation: Introduction

Overview

Burgers' equation is a nonlinear variant of the advection equation previously studied

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = D\frac{\partial^2 u}{\partial x^2}.$$

Topics: solution of the equation by the method of characteristics; breaking waves; Cole–Hopf transformation.

In this chapter, the equation is to be solved on the real line, unless otherwise stated.

23.1 Characteristic solution

In this section we work with $D = 0$, and we are to solve

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = 0,$$

with initial data $u(x, t = 0) = f(x)$.

- Recall the linear advection equation, $u_t + c(x,t)u_x = 0$, with solution $u =$ const on $dx/dt = c(x,t)$. Let's propose the same thing here:

$$v(t) := u(x(t), t), \qquad dx/dt = u.$$

Differentiate wrt time to show that $v(t) = \text{const.}$:

$$\begin{aligned}\frac{dv}{dt} &= \frac{\partial u}{\partial x}\frac{dx}{dt} + \frac{\partial u}{\partial t},\\ &= \frac{\partial u}{\partial x}u + \frac{\partial u}{\partial t},\\ &= 0.\end{aligned}$$

- Hence,

$$\begin{aligned}u(x_0, 0) &= u_0,\\ u(x,t) &= u_0, \qquad dx/dt = u(x,t), x(0) = x_0.\end{aligned}$$

Solving the characteristic equation gives

$$x = x(t; x_0).$$

Inverting gives

$$x_0 = \eta(x,t).$$

But

$$u(x,t) = u_0 = f(x_0) = f(\eta(x,t)).$$

- For Burgers' equation, the solution can also be written in **implicit form**:

$$u(x,t) = f(x - u(x,t)t).$$

Proof:

$$\begin{aligned}u_t + uu_x &= f'(x-ut)\left[\frac{\partial}{\partial t}(x-ut) + u\frac{\partial}{\partial x}(x-ut)\right],\\ &= f'(x-ut)\left[-u - u_t t + u(1-u_x t)\right],\\ &= f'(x-ut)\left[-u_t - uu_x\right]t,\\ &= 0.\end{aligned}$$

The latter is not much use in practical applications, although in gives rise to certain theoretical insights later on – so remember it!

For now, we shall consider instead the characteristic solution only: $u = \text{const}$ along $dx/dt = u$.

Examples

1. Solve the PDE
$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = 0,$$
for $t > 0$, subject to the initial condition $u = x$ on $t = 0$.

   The characteristics are given by
$$\frac{dx}{dt} = u = \text{const.} = u_0,$$
   hence
$$x = x_0 + u_0 t.$$
   But $u_0 = f(x_0) = x_0$, hence
$$x = x_0 + x_0 t.$$
   Inverting gives
$$x_0 = \frac{x}{1+t},$$
   and
$$u(x,t) = f(x_0) = f\left(\frac{x}{1+t}\right) = \frac{x}{1+t}.$$

2. Solve the PDE
$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = 1,$$
for $t > 0$, subject to the initial condition $u = x$ on $t = 0$.

   Even though this is not Burgers' equation, let's try characteristics:
$$v(t) = u(x(t),t), \qquad dx/dt = u.$$
   Then,
$$\begin{aligned} \frac{dv}{dt} &= \frac{\partial u}{\partial x}\frac{dx}{dt} + \frac{\partial u}{\partial t}, \\ &= \frac{\partial u}{\partial x}u + \frac{\partial u}{\partial t}, \\ &= 1. \end{aligned}$$
   Hence,
$$\begin{aligned} v(t) &= v_0 + t = u_0 + t, \\ &= u(x,t). \end{aligned}$$

along characteristics $dx/dt = u$. Thus, characteristics are given by

$$\frac{dx}{dt} = u = v = u_0 + t,$$

hence

$$x = x_0 + u_0 t + \tfrac{1}{2}t^2.$$

But $u_0 = f(x_0) = x_0$, hence

$$x = x_0 + x_0 t + \tfrac{1}{2}t^2. \qquad (*)$$

Inverting gives

$$x_0 = \frac{x - \frac{1}{2}t^2}{1+t},$$

and

$$\begin{aligned} u(x,t) &= u_0 + t, \\ &= x_0 + t, \\ &= \frac{x - \frac{1}{2}t^2}{1+t} + t, \\ &= \frac{x + t + \frac{1}{2}t^2}{1+t}. \end{aligned}$$

The characteristic curves are shown schematically in Fig. 23.1.

Note: for plotting a t-x curve, invert Eq. (*) to obtain

$$t = -x_0 + \sqrt{x_0^2 - 4\tfrac{1}{2}\left(x_0 - x\right)} = -x_0 + \sqrt{x_0^2 - 2x_0 + x}.$$

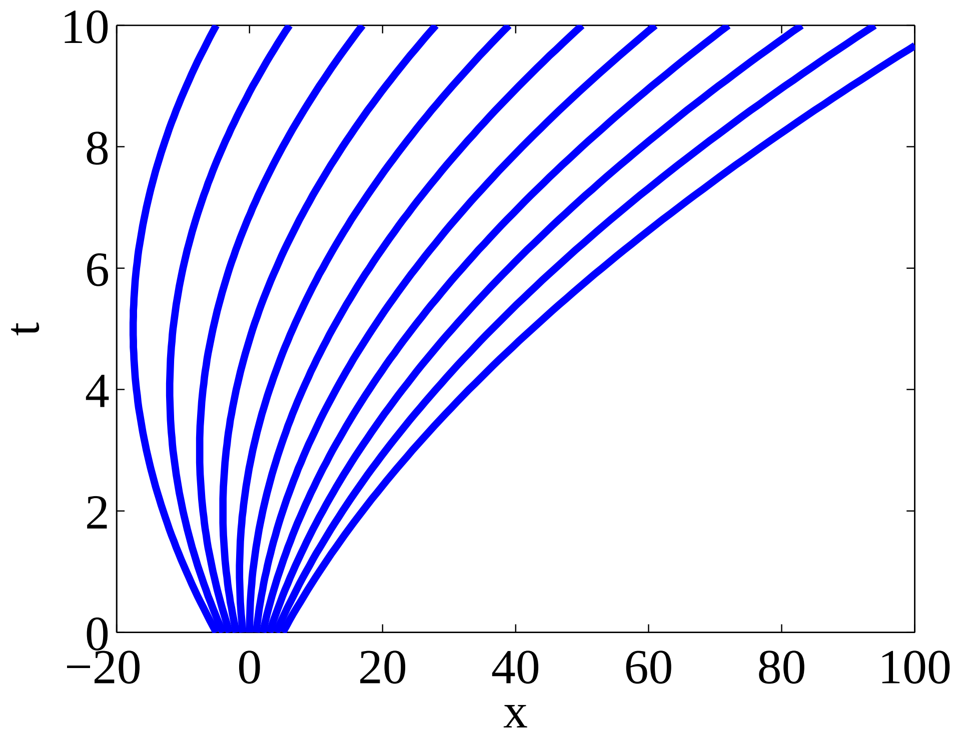


Figure 23.1: Characteristic curves for $u_t + uu_x = 1$, initial data $u(x, t = 0) = x$. The curves are given by Eq. (*).

Chapter 24

Burgers' equation: Breaking waves

24.1 Breaking waves

Consider the PDE

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = 0,$$

for $t > 0$, subject to the initial condition $u = f(x)$ on $t = 0$

The characteristics are given by

$$\frac{dx}{dt} = u = \text{const.} = u_0,$$

hence

$$x = x_0 + u_0 t.$$

But $u_0 = f(x_0)$, hence

$$x = x_0 + f(x_0)t.$$

We need to invert this and solve for $x_0 = \eta(x, t)$. Is this possible? Consider two curves:

$$\begin{aligned} y_1(x_0) &= x \text{ (a positive constant in } x_0\text{-}y \text{ space)}, \\ y_2(x_0) &= x_0 + f(x_0)t, \end{aligned}$$

in $x_0 - y$ space (we take $x > 0$ here; this is allowed, since the equation is to be solved on the whole real line). A sufficient condition for a solution to the problem $y_1 = y_2$ to exist is if

$$\frac{dy_2}{dx_0} > 0, \qquad x_0 \in \mathbb{R},$$

since then the curve $y_2(x_0)$ is increasing and is bound to cross the line $y_1 = x$ eventually (Fig. 24.1).

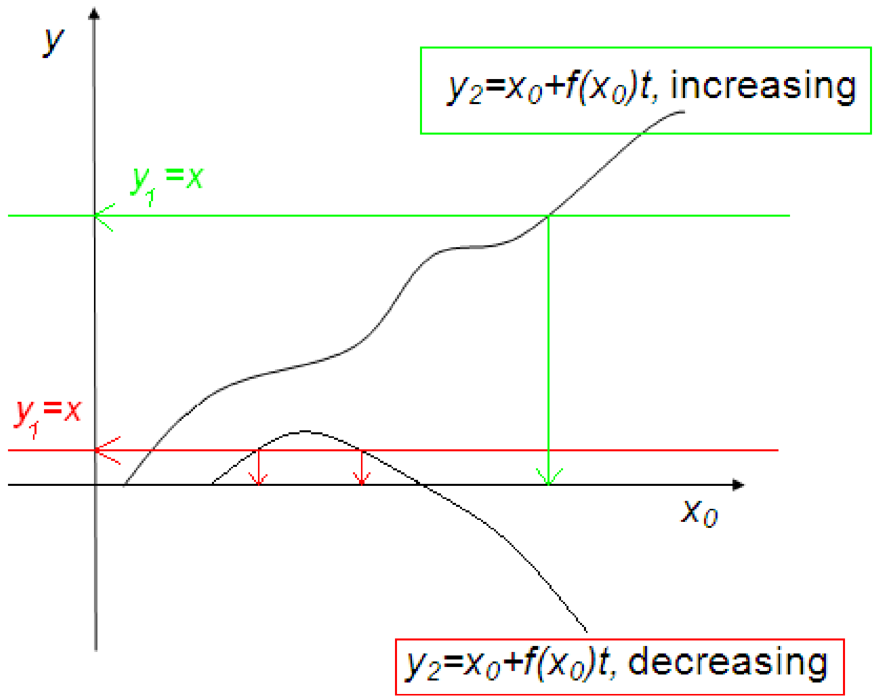


Figure 24.1: The failure of the inversion for f decreasing.

However, we see that

$$\min_{x_0} \frac{dy_2}{dx_0} = \min_{x_0} \left[1 + f'(x_0)t\right],$$

which might be negative. This suggests that the method of characteristics **will fail in general** if

$$\min_{x_0} \left[1 + f'(x_0)t\right] \leq 0.$$

The failure happens at the critical time

$$t_c = -\min_{x_0} \frac{1}{f'(x_0)}.$$

Thus, we are left with the following solution:

$$u = \begin{cases} f(x_0) \text{ along } x(t) = x_0 + f(x_0)t, & 0 < t < \infty \text{ if } t_c < 0, \\ f(x_0) \text{ along } x(t) = x_0 + f(x_0)t, & 0 < t < t_c \text{ if } t_c > 0. \end{cases}$$

What happens after $t = t_c$?

Suppose that $f'(s) <$ somewhere. Then, there is an interval I where $f(s)$ is a decreasing function. Let's take two points $x_1, x_2 \in I$. We will try to compute if and where the characteristic curves cross. We have,

$$\begin{aligned} x &= x_0 + f(x_0)t, \\ x &= x_1 + f(x_1)t. \end{aligned}$$

At the crossing point, the two x-coordinates agree, and the times are the same, so

$$x_0 + f(x_0)t = x_1 + f(x_1)t.$$

Solving for t gives

$$t = \frac{x_0 - x_1}{f(x_1) - f(x_0)},$$

or

$$t = -\left(\frac{x_1 - x_0}{f(x_1) - f(x_0)}\right) > 0$$

which is positive because $f(s)$ is decreasing on I. Thus, the failure of the inversion

$$x = x_0 + f(x_0)t \implies x_0 = \eta(x, t)$$

is equivalent to the crossing of characteristics.

Finally, recall the implicit solution of the problem:

$$u = f(x - ut).$$

Along characteristics, consider the gradient of u:

$$\begin{aligned} g &:= \left.\frac{\partial u}{\partial x}\right|_{x(t)}, \\ &= f'(x - u(x,t)t)\left(1 - \frac{\partial u}{\partial x}t\right), \\ &= f'(x - ut)\left(1 - gt\right), \end{aligned}$$

Hence,

$$g = \left.\frac{f'(x - ut)}{1 + f'(x - ut)t}\right|_{x(t)}.$$

Call $X := x - ut$. Thus,

$$g = \left.\frac{f'(X)}{1 + f'(X)t}\right|_{x(t)}.$$

Thus, if

$$f'(X)t = -1, \qquad (*)$$

for some X, then the slope goes to infinity and the solution loses its smoothness. But (*) is precisely the criterion for the inversion to fail. Thus, the slope g goes to infinity if the method of characteristics fails. Thus, we have the following equivalent criteria:

The method of characteristics fails iff the inversion procedure

$$x = x_0 + f(x_0)t \implies x_0\eta(x,t)$$

fails,

- iff $f(x_0)$ is decreasing somewhere on the domain,
- iff characteristics cross in finite time,
- iff the slope of the solution tends to infinity in finite time.

This phenomenon is **wave breaking**.

Example

Solve the PDE

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = 0,$$

for $t > 0$, subject to the initial condition $u = \sin(x)$ on $t = 0$. Restrict your attention to $x \in [0, 2\pi]$. Sketch the characteristic curves up to the breaking point.

The characteristics are given by

$$\frac{dx}{dt} = u = \text{const.} = u_0,$$

hence

$$x = x_0 + u_0 t.$$

But $u_0 = \sin(x_0)$, hence

$$x = x_0 + \sin(x_0)t.$$

We need to invert this and solve for $x_0 = \eta(x,t)$. Is this possible? The time at which the inversion fails is given by

$$1 + \cos(x_0)t = 0,$$

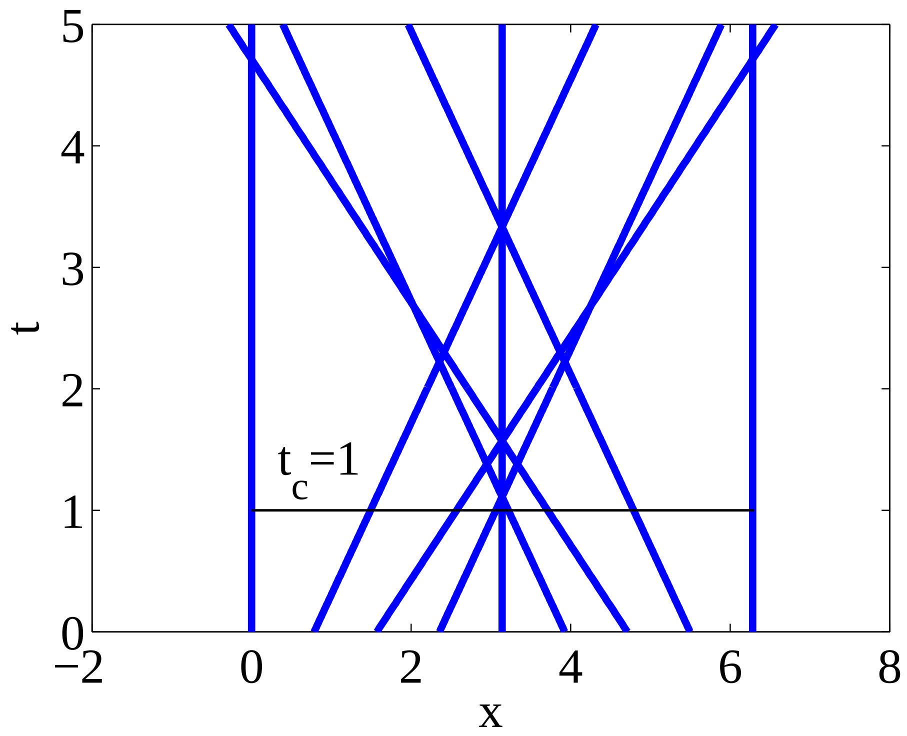
hence

$$t = -1/\cos(x_0).$$

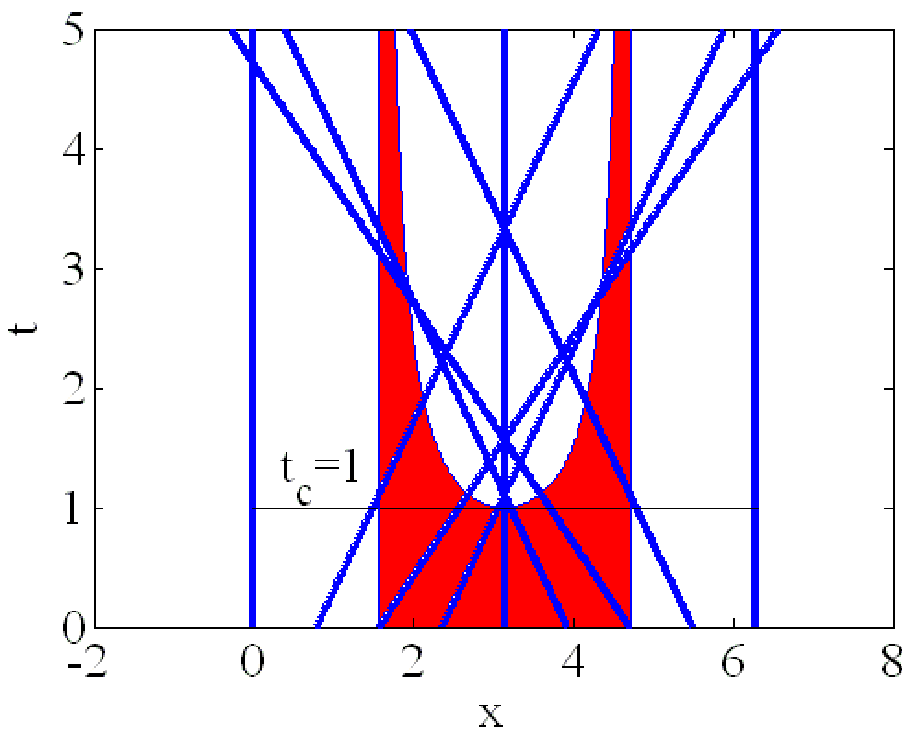
For $\pi/2 < x_0 < 3\pi/2$ the function $\cos(x_0)$ is negative, and $t > 0$. Thus,

$$t_c = \min_{x_0 \in [0,2\pi]} (-1/\cos(x_0)) = 1.$$

Sketches of the characteristics and the breaking time are shown in Fig. 24.2.



(a) Characteristics up to breaking time



(b) Breaking time as a function of initial location

Figure 24.2:

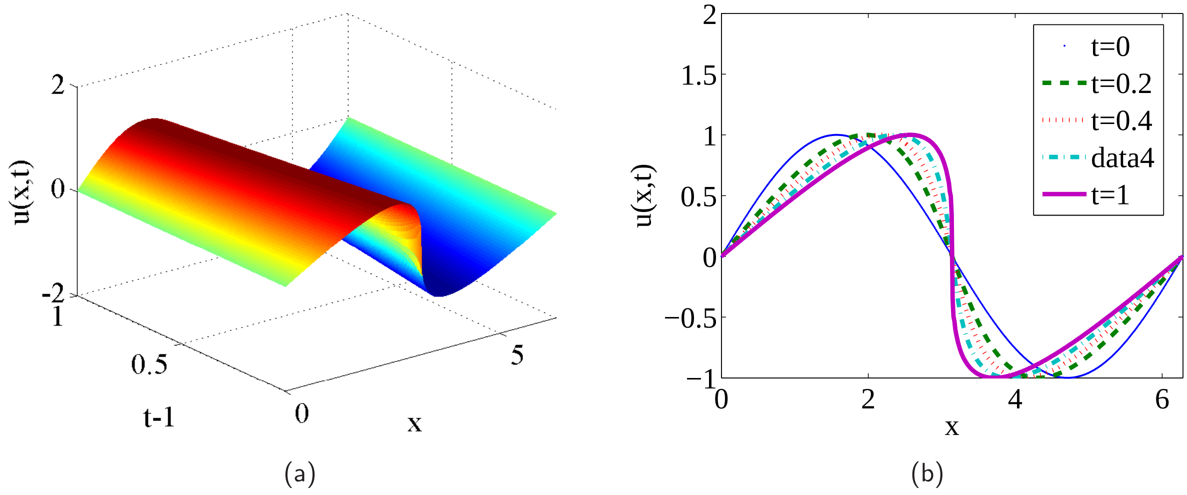


Figure 24.3: Numerical simulation of Burgers equation with $u(x, t = 0) = \sin(x)$.

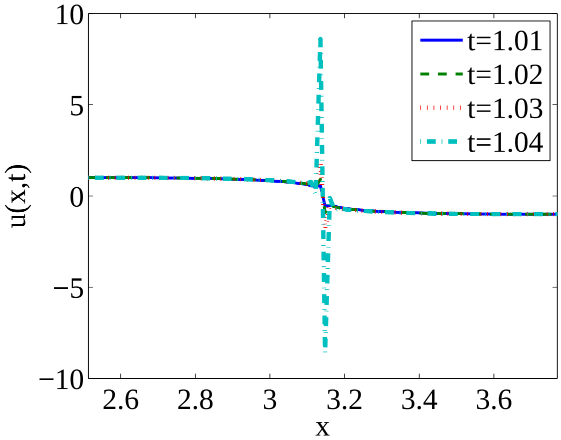


Figure 24.4: Failure of the numerical simulation of Burgers equation with $u(x, t = 0) = \sin(x)$, after $t = 1$.

Numerical simulation

I have done a numerical simulation of the last problem (Fig. 24.3). It is a standard finite-difference code ('burgers.m'), and it works. The wave steepens around $x = \pi$, until it develops a very large slope at $x = 1$. Now look at what happens if you try to integrate beyond $t = 1$ (Fig. 24.4). **The code fails!!** It is impossible to integrate beyond a breaking point with a standard finite-difference code. Don't try it!!

24.2 Riemann problems

For the purposes of this module, a **Riemann problem** is a Burgers problem with initial data defined in a piecewise manner. The idea can be extended to other PDEs of conservation-law type.

Example 1

Consider the PDE

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = 0,$$

for $t > 0$, subject to the initial condition $u(x, t = 0) = f(x)$,

$$f(x) = \begin{cases} 0, & x < 0, \\ x, & 0 \leq x \leq 1, \\ 1, & x > 1. \end{cases}$$

Solve the equation using the method of characteristics, up to the breaking point (if it exists).

Divide the answer into three regions.

Region 1: Here $x < 0$. Along characteristics,

$$dx/dt = \text{const} = u(x_0) = 0,$$

hence

$$x = x_0, \qquad \text{Region 1.}$$

Region 2: Here $0 \leq x \leq 1$. Along characteristics,

$$dx/dt = \text{const} = u(x_0) = x_0,$$

hence

$$\begin{aligned} x &= x_0(1+t), \\ t &= \frac{x}{x_0} - 1, \qquad \text{Region 2.} \end{aligned}$$

Region 3: Here $x > 1$. Along characteristics,

$$dx/dt = \text{const} = u(x_0) = 1,$$

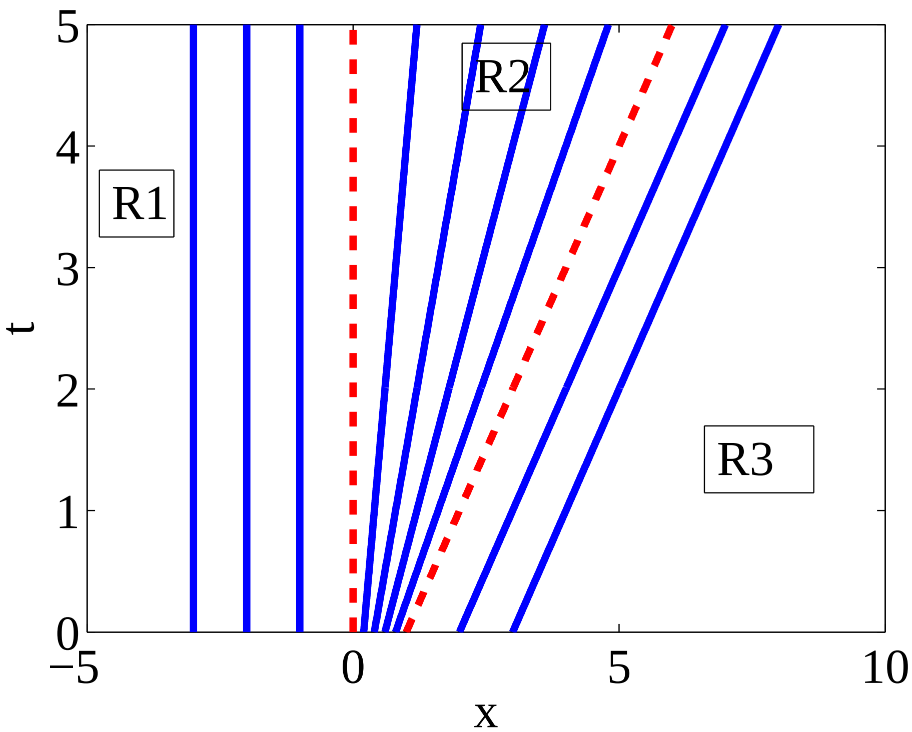


Figure 24.5: Characteristics of the Riemann problem with piecewise increasing initial data

hence

$$\begin{aligned} x &= x_0 + t, \\ t &= x - x_0, \qquad \text{Region 2.} \end{aligned}$$

Now sketch these curves, but pay close attention to the boundaries:

Region 1–Region 2 boundary: Here $x_0 = 0$. The region-1 curve gives $x = 0$, as does the region-2 curve. This is a continuous boundary.

Region 2–Region 3 boundary: Here $x_0 = 1$. The region-2 curve gives $x = 1 + t$, as does the region-3 curve. This is a continuous boundary.

Sketch the results in an $x - t$ plane for a variety of initial values x_0 (Fig. 24.5).

Example 2

Consider the PDE

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} = 0,$$

for $t > 0$, subject to the initial condition $u(x, t = 0) = f(x)$,

$$f(x) = \begin{cases} 1, & x < 0, \\ 1 - x, & 0 \le x \le 1, \\ 0, & x > 1. \end{cases}$$

Solve the equation using the method of characteristics, up to the breaking point (if it exists).

Divide the answer into three regions.

Region 1: Here $x < 0$. Along characteristics,

$$dx/dt = \text{const} = u(x_0) = 1,$$

hence

$$x = x_0 + t, \qquad \text{Region 1.}$$

Region 2: Here $0 \le x \le 1$. Along characteristics,

$$dx/dt = \text{const} = u(x_0) = 1 - x_0,$$

hence

$$\begin{aligned} x &= x_0 + (1 - x_0)t, \\ t &= \frac{x - x_0}{1 - x_0}, \qquad \text{Interior of region 2.} \end{aligned}$$

Region 3: Here $x > 1$. Along characteristics,

$$dx/dt = \text{const} = u(x_0) = 0,$$

hence

$$x = x_0, \qquad \text{Region 2.}$$

Now sketch these curves, but pay close attention to the boundaries:

Region 1–Region 2 boundary: Here $x_0 = 0$. The region-1 curve gives $x = t$, as does the region-2 curve. This is a continuous boundary.

Region 2–Region 3 boundary: Here $x_0 = 1$. The region-2 curve gives $x = 1$, as does the region-3 curve. This is a continuous boundary.

Sketch the results in an $x - t$ plane for a variety of initial values x_0 (Fig. 24.6). **The characteristics from region 2 all collide at $t = 1$, and a breaking wave occurs**. Thus, $t_c = 1$ for the piecewise

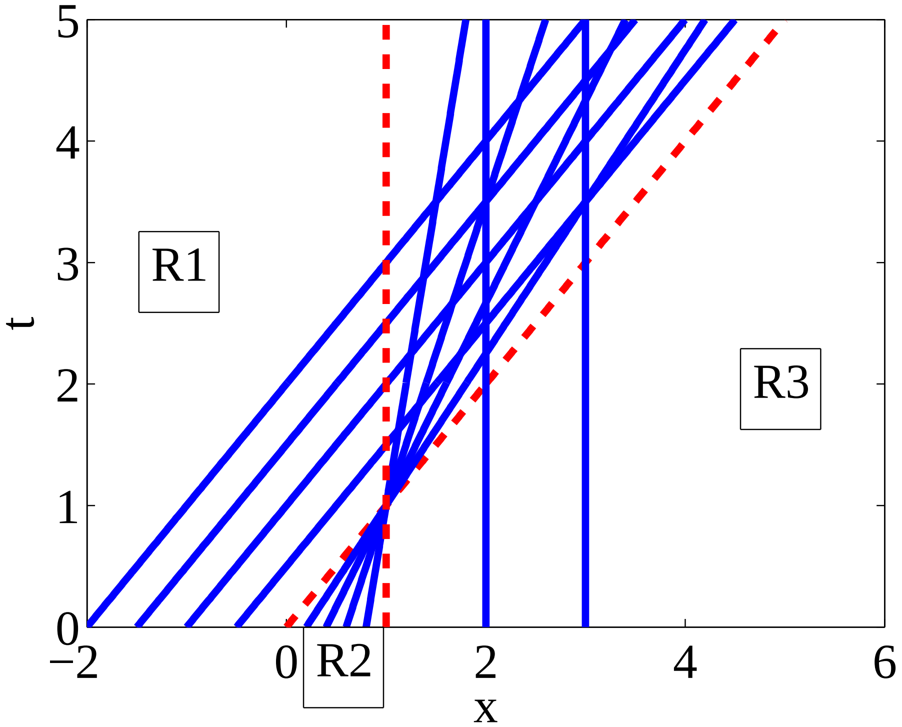


Figure 24.6: Characteristics of the Riemann problem with piecewise decreasing initial data

decreasing data.

Chapter 25

The spectrum of Burgers' equation

Overview

In this chapter we compute the Fourier coefficients of the solution of Burgers' equation in closed form for the initial data

$$u(x, t=0) = f(x) = 1 + \cos(x). \tag{25.1}$$

The Fourier coefficients of the gradient $G := u_x$ then follows immediately. We are able to show that for a certain value of $x - t$, the gradient blows up. This calculation was pointed out to me by Miguel Bustamante. **In this chapter, the boundary conditions are periodic and the spatial domain is $[0, 2\pi]$.**

25.1 The derivation

Start with

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = 0. \tag{25.2}$$

The solution is known to be

$$u(x,t) = f(\eta), \qquad \eta = x - u(x,t)t,$$

where $u(x, t=0) = f(x)$ is the IC. The spatial domain of the PDE is $[0, 2\pi]$, and the boundary and initial conditions are periodic. Thus, define the Fourier coefficient or u:

$$\hat{u}_p(t) = \frac{1}{2\pi}\int_0^{2\pi} u(x,t)\mathrm{e}^{-\mathrm{i}px}\,\mathrm{d}x.$$

Now

$$\begin{aligned}\frac{1}{2\pi}\int_0^{2\pi}\frac{\partial u}{\partial x}\mathrm{e}^{-\mathrm{i}px}\mathrm{d}x &= -\frac{1}{2\pi}\int_0^{2\pi}u\partial_x\left(\mathrm{e}^{-\mathrm{i}px}\right)\mathrm{d}x,\\ &= \frac{\mathrm{i}p}{2\pi}\int_0^{2\pi}u\mathrm{e}^{-\mathrm{i}px},\\ &= \mathrm{i}p\hat{u}_p(t),\end{aligned}$$

hence

$$\mathrm{i}p\hat{u}_p(t) = \frac{1}{2\pi}\int_0^{2\pi}\partial_x u(x,t)\mathrm{e}^{-\mathrm{i}px}\,\mathrm{d}x.$$

Consider

$$\begin{aligned}\frac{\partial u}{\partial x} &= f'(\eta)\frac{\partial\eta}{\partial x},\\ \frac{\partial u}{\partial x}\mathrm{d}x &= f'(\eta)\frac{\partial\eta}{\partial x}\mathrm{d}x,\\ \frac{\partial u}{\partial x}\mathrm{d}x &= f'(\eta)\mathrm{d}\eta,\end{aligned}$$

Also, $\eta = x - u(x,t)\eta$, $x = \eta + u(x,t)t$, $x = \eta + f(\eta)t$. Hence,

$$\begin{aligned}\mathrm{i}p\hat{u}_p(t) &= \frac{1}{2\pi}\int_0^{2\pi}\mathrm{e}^{-\mathrm{i}px}\frac{\partial u}{\partial x}\mathrm{d}x,\\ &= \frac{1}{2\pi}\int_0^{2\pi}\mathrm{e}^{-\mathrm{i}p(\eta+f(\eta)t)}f'(\eta)\mathrm{d}\eta.\end{aligned}$$

We are going to choose the IC

$$f(\eta) = 1+\cos(\eta),$$

for which $f'(\eta) = -\sin(\eta)$. Hence,

$$ip\hat{u}_p(t) = -\frac{1}{2\pi}\int_0^{2\pi}\sin\eta\mathrm{e}^{-\mathrm{i}p\eta}\mathrm{e}^{-\mathrm{i}pt}\mathrm{e}^{-\mathrm{i}pt\cos\eta}\mathrm{d}\eta.$$

Effect a change of variable:

$$\begin{aligned}y &= \tfrac{1}{2}\pi - \eta,\\ \cos\eta &= \sin y,\\ \sin\eta &= \cos y.\end{aligned}$$

Hence,

$$\mathrm{i}p\hat{u}_p(t) = -\frac{1}{2\pi}\int_{-3\pi/2}^{\pi/2}\cos y\,\mathrm{e}^{-\mathrm{i}p(\pi/2-y)}\mathrm{e}^{-\mathrm{i}pt}\mathrm{e}^{-\mathrm{i}pt\sin y}\mathrm{d}y.$$

Take the constants outside the integral:

$$\mathrm{i}p\hat{u}_p(t) = -\frac{\mathrm{e}^{-\mathrm{i}p\pi/2}\mathrm{e}^{-\mathrm{i}pt}}{2\pi}\int_{-3\pi/2}^{\pi/2}\cos y\,\mathrm{e}^{\mathrm{i}py}\mathrm{e}^{-\mathrm{i}pt\sin y}\mathrm{d}y.$$

The integrand is a 2π-periodic function and the limits of integration can therefore be shifted at will:

$$\mathrm{i}p\hat{u}_p(t) = -\frac{\mathrm{e}^{-\mathrm{i}p\pi/2}\mathrm{e}^{-\mathrm{i}pt}}{2\pi}\int_{-\pi}^{\pi}\cos y\,\mathrm{e}^{\mathrm{i}py}\mathrm{e}^{-\mathrm{i}pt\sin y}\mathrm{d}y.$$

Now consider the integral

$$\begin{aligned} I &:= \tfrac{1}{2\pi}\int_{-\pi}^{\pi}\cos y\,\mathrm{e}^{\mathrm{i}py}\,\mathrm{e}^{-\mathrm{i}pt\sin y}\,\mathrm{d}y, \\ &= \tfrac{1}{2\pi}\int_{-\pi}^{\pi}\left(\frac{\mathrm{e}^{\mathrm{i}y}+\mathrm{e}^{-\mathrm{i}y}}{2}\right)\mathrm{e}^{-\mathrm{i}pt\sin y}\,\mathrm{d}y, \\ &= \tfrac{1}{2}\tfrac{1}{2\pi}\int_{-\pi}^{\pi}\mathrm{e}^{\mathrm{i}y(p+1)}\mathrm{e}^{-\mathrm{i}pt\sin y}\,\mathrm{d}y + \tfrac{1}{2}\tfrac{1}{2\pi}\int_{-\pi}^{\pi}\mathrm{e}^{\mathrm{i}y(p-1)}\mathrm{e}^{-\mathrm{i}pt\sin y}\,\mathrm{d}y. \end{aligned}$$

But

$$J_n(\theta) = \tfrac{1}{2\pi}\int_{-\pi}^{\pi}\mathrm{e}^{\mathrm{i}yn}\,\mathrm{e}^{-\mathrm{i}\theta\sin y}\,\mathrm{d}y, \qquad \theta\in\mathbb{R}.$$

Therefore, if we take $\theta = pt$ and $n = p\pm 1$, we get

$$I = \tfrac{1}{2}J_{p+1}(pt) + \tfrac{1}{2}J_{p-1}(pt),$$

and

$$\mathrm{i}p\hat{u}_p(t) = -\frac{\mathrm{e}^{-\mathrm{i}p\pi/2}\mathrm{e}^{-\mathrm{i}pt}}{2}\left(J_{p+1}(pt) + J_{p+1}(pt)\right).$$

The **power spectrum** $|\hat{u}_p|^2$ of Burgers' equation (for the prescribed initial conditions) is thus

$$|\hat{u}_p|^2 = \tfrac{1}{4p^2}\left|J_{p+1}(pt) + J_{p-1}(pt)\right|^2, \qquad p\neq 0.$$

25.2 The Fourier coefficients of the gradient

Now Fourier coefficients of the gradient are

$$\hat{G}_p = \mathrm{i}p\hat{u}_p = -\frac{\mathrm{e}^{-\mathrm{i}p\pi/2}\mathrm{e}^{-\mathrm{i}pt}}{2}\left[J_{p+1}(pt) + J_{p+1}(pt)\right],$$

hence

$$
\begin{aligned}
G := u_x &= \sum_{\substack{p=-\infty \\ p\neq 0}}^{\infty} \mathrm{e}^{\mathrm{i}px} \hat{G}_p, \\
&= -\sum_{\substack{p=-\infty \\ p\neq 0}}^{\infty} \frac{\mathrm{e}^{-\mathrm{i}p\pi/2}\mathrm{e}^{-\mathrm{i}pt}\mathrm{e}^{\mathrm{i}px}}{2} \left[J_{p+1}(pt) + J_{p+1}(pt)\right], \\
&= -\tfrac{1}{2} \sum_{\substack{p=-\infty \\ p\neq 0}}^{\infty} \mathrm{e}^{\mathrm{i}p(-\pi/2-t+x)} \left[J_{p+1}(pt) + J_{p-1}(pt)\right].
\end{aligned}
$$

Consider what happens when the phases

$$
\mathrm{e}^{\mathrm{i}p(-\pi/2-t+x)}
$$

are all the same, in other words, when $x = t + \pi/2$:

$$
G(x = t + \pi/2, t) = -\tfrac{1}{2} \sum_{\substack{p=-\infty \\ p\neq 0}}^{\infty} \left[J_{p+1}(pt) + J_{p+1}(pt)\right]
$$

However, we have the asymptotic formulas

$$
J_\nu(z) \sim \frac{1}{\sqrt{2\pi\nu}} \left(\frac{ez}{2\nu}\right)^{\nu}, \qquad z \to \infty,
$$

hence

$$
J_{p\pm 1}(pt) \sim \left(\frac{et}{2}\right)^{p\pm 1} \frac{p^{-1/2}}{\sqrt{2\pi}} \qquad p \to \infty.
$$

Now $\sum p^{-1/2}$ diverges by the integral comparison test and hence, the tail of the series (*) is bounded from below by a divergent series and therefore itself diverges.

As in previous chapters, the minimum blow-up time is

$$
t_{\mathrm{c}} = \min_{x_0} \left[-1/f'(x_0)\right] = 1,
$$

and the location of the blow-up is therefore $x = \pi/2 + 1$. In a way, **this entire calculation is a fancy way of proving that the gradient blows up** – even though the shape of the initial data anyway guaranteed such a result. However, the behaviour of the phases at blowup is a very interesting phenomenon: it is as if 'constructive interference' happens, such that all Fourier waves in the gradient align and add explosively – to produce blowup. We therefore have the following theorem:

Theorem 25.1 *The gradient in this simple example of Burgers' equation (Eq.* (25.2)*, initial con-*

ditions (25.1)*) blows up when the phases of the Fourier coefficients all align. This happens when* $x = t + \pi/2$.

Chapter 26

Burgers' equation: Regularisation

26.1 Finite diffusion; Cole–Hopf transformation

Without more information, it is impossible to integrate beyond the breaking time. There are ways of doing this – **weak solutions** and **jump conditions**. We do not pursue this approach here. Instead, we consider the **regularized Burgers' equation**,

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = D\frac{\partial^2 u}{\partial x^2}, \qquad D > 0.$$

It is straightforward to show that this equation is regular, at least in a mean sense. To show this, multiply both sides by u and integrate. Assume that u vanishes at the boundaries of the domain Ω:

$$\begin{aligned}
\int_\Omega u u_t \,\mathrm{d}x + \int_\Omega u^2 u_x \,\mathrm{d}x &= D\int_\Omega u u_{xx}\,\mathrm{d}x, \\
\tfrac{1}{2}\partial_t \int_\Omega u^2\,\mathrm{d}x + \tfrac{1}{3}\int_\omega \partial_x u^3\,\mathrm{d}x &= \int_\Omega u u_{xx}\,\mathrm{d}x, \\
\tfrac{1}{2}\frac{d}{dt}\|u\|_2^2 + \tfrac{1}{3}u^3\Big|_{\partial\Omega} &= D\int_\Omega \left[\partial_x (u u_x) - u_x^2\right]\,\mathrm{d}x, \\
\tfrac{1}{2}\frac{d}{dt}\|u\|_2^2 &= u u_x\Big|_{\partial\Omega} - D\int_\Omega \left[\partial_x (u u_x)\, u_x^2\right]\,\mathrm{d}x.
\end{aligned}$$

Finally,

$$\frac{d}{dt}\|u\|_2^2 = -2D\|u_x\|_2^2.$$

Integrating over a finite period of time gives

$$\|u\|_2^2(T) = \|u\|_2^2(0) - 2D\int_0^T \|u_x\|_2^2(t)\mathrm{d}t. \qquad (*)$$

Thus,

$$\|u\|_2^2(T) \leq \|u\|_2^2(0),$$

and the solution at later times $T > 0$ is bounded in the L^2-norm. Call

$$\|u\|_2^2(T) = a(T)\|u\|_2^2(0), \qquad a(T) \leq 1.$$

Refer back to Eq. (*). We have

$$\|u\|_2^2(T) = a(T)\|u\|_2^2(0) = \|u\|_2^2(0) - 2D\int_0^T \|u_x\|_2^2(t)\mathrm{d}t.$$

Re-arranging gives

$$2D\int_0^T \|u_x\|_2^2(t)\mathrm{d}t = (1 - a(T))\|u\|_2^2(0) \geq 0.$$

Thus, we have a sandwich result:

$$0 \leq 2D\int_0^T \|u_x\|_2^2(t)\mathrm{d}t \leq B(T) := (1+\epsilon)(1 - a(T))\|u\|_2^2(0),$$

and the gradient is bounded in the L^2 norm in space and the L^1 norm in time, for any finite interval of time $[0, T]$. Thus, finite diffusion $D > 0$ prevents blow-up of the gradient.

There is one further nice feature of the regularised Burgers' equation: it has **an exact solution**. With $20 - 20$ hindsight, we suggest the transformation

$$u(x,t) = -\frac{2D}{\phi}\phi_x.$$

Thus,

$$u_t = -\frac{2D}{\phi}\phi_{xt} + \frac{2D}{\phi^2}\phi_x\phi_t,$$

and

$$u_x = -\frac{2D}{\phi}\phi_{xx} + \frac{2D}{\phi^2}\phi_x\phi_x.$$

Put them together:

$$\begin{aligned} u_t + uu_x &= -\frac{2D}{\phi}\phi_{xt} + \frac{2D}{\phi^2}\phi_x\phi_t + \left(-\frac{2D}{\phi}\phi_x\right)\left(-\frac{2D}{\phi}\phi_{xx} + \frac{2D}{\phi^2}\phi_x\phi_x\right), \\ &= -\frac{2D}{\phi}\phi_{xt} + \frac{2D}{\phi^2}\phi_x\phi_t + \frac{4D^2}{\phi^2}\phi_x\phi_{xx} - \frac{4D^2}{\phi^3}\phi_x^3. \end{aligned}$$

Now let's work out u_{xx}:

$$u_{xx} = -\frac{2D}{\phi}\phi_{xxx} + \frac{2D}{\phi^2}\phi_{xx}\phi_x + \frac{4D}{\phi^2}\phi_x\phi_{xx} - \frac{4D}{\phi^3}\phi_x^3,$$

and

$$Du_{xx} = -\frac{2D^2}{\phi}\phi_{xxx} + \frac{2D^2}{\phi^2}\phi_{xx}\phi_x + \frac{4D^2}{\phi^2}\phi_x\phi_{xx} - \frac{4D^2}{\phi^3}\phi_x^3.$$

Compare both sides:

$$\begin{aligned} u_t + uu_x &= Du_{xx}, \\ -\frac{2D}{\phi}\phi_{xt} + \frac{2D}{\phi^2}\phi_x\phi_t + \underbrace{\frac{4D^2}{\phi^2}\phi_x\phi_{xx}}_{***} - \underbrace{\frac{4D^2}{\phi^3}\phi_x^3}_{ooo} &= -\frac{2D^2}{\phi}\phi_{xxx} + \frac{2D^2}{\phi^2}\phi_{xx}\phi_x + \underbrace{\frac{4D^2}{\phi^2}\phi_x\phi_{xx}}_{**} - \underbrace{\frac{4D^2}{\phi^3}\phi_x^3}_{ooo}. \end{aligned}$$

Cancel like terms:

$$-\frac{2D}{\phi}\phi_{xt} + \frac{2D}{\phi^2}\phi_x\phi_t = -\frac{2D^2}{\phi}\phi_{xxx} + \frac{2D^2}{\phi^2}\phi_{xx}\phi_x$$

More cancellation:

$$-\phi_{xt} + \frac{1}{\phi}\phi_x\phi_t = -D\phi_{xxx} + \frac{D}{\phi}\phi_{xx}\phi_x$$

Factorise the terms in ϕ_x/ϕ:

$$-\phi_{xt} + D\phi_{xxx} = \frac{\phi_x}{\phi}\left[-\phi_t + D\phi_{xx}\right],$$

multiply both sides by -1:

$$\phi_{xt} - D\phi_{xxx} = \frac{\phi_x}{\phi}\left[\phi_t - D\phi_{xx}\right];$$

'factorize' by ∂_x on the LHS:

$$\partial_x\left(\phi_t - D\phi_{xx}\right) = \frac{\phi_x}{\phi}\left[\phi_t - D\phi_{xx}\right],$$

one solution of which is the **linear diffusion equation:**

$$\phi_t = D\phi_{xx}.$$

But we know the solution to the diffusion equation on the line:

$$\phi(x,t) = \int_{-\infty}^{\infty} K(x-y,t)\phi_0(y)\mathrm{d}y,$$

where

$$K(x-y,t) = \frac{1}{\sqrt{4\pi Dt}}e^{-(x-y)^2/4Dt}$$

is the heat kernel and

$$\phi_0(x) = \phi(x, t=0).$$

But

$$u(x,t=0)=f(x)=-\frac{2D}{\phi_0(x)}\frac{\partial\phi_0}{\partial x}=-2D\frac{\partial}{\partial x}\log\phi_0(x).$$

Solving for $\phi_0(x)$ gives

$$\int_0^x\frac{-f(s)}{2D}\mathrm{d}s=\log(\phi_0(x))-\log(\phi_0(0)),$$

hence

$$\phi_0(x)=\phi_0(0)\exp\left[-\frac{1}{2D}\int_0^x f(s)\mathrm{d}s\right].$$

But recall the heat-kernel solution:

$$\phi(x,t)=\frac{1}{\sqrt{4\pi Dt}}\int_{-\infty}^{\infty}e^{-(x-y)^2/4Dt}\phi_0(y)\mathrm{d}y.$$

Putting the last two equations together gives

$$\begin{aligned}\phi(x,t) &= \frac{1}{\sqrt{4\pi Dt}}\int_{-\infty}^{\infty}e^{-(x-y)^2/4Dt}\phi_0(0)\exp\left[-\frac{1}{2D}\int_0^y f(s)\mathrm{d}s\right]\mathrm{d}y,\\ &= \frac{\phi_0(0)}{\sqrt{4\pi Dt}}\int_{-\infty}^{\infty}\exp\left[-\frac{(x-y)^2}{4Dt}-\frac{1}{2D}\int_0^y f(s)\mathrm{d}s\right]\mathrm{d}y,\end{aligned}$$

and

$$u(x,t)=-2D\frac{\partial}{\partial x}\log\left\{\frac{\phi_0(0)}{\sqrt{4\pi Dt}}\int_{-\infty}^{\infty}\exp\left[-\frac{(x-y)^2}{4Dt}-\frac{1}{2D}\int_0^y f(s)\mathrm{d}s\right]\mathrm{d}y\right\}.$$

The logarithmic derivative is invariant under rescaling, so the final answer is

$$\boxed{u(x,t)=-2D\frac{\partial}{\partial x}\log\left\{\frac{1}{\sqrt{4\pi Dt}}\int_{-\infty}^{\infty}\exp\left[-\frac{(x-y)^2}{4Dt}-\frac{1}{2D}\int_0^y f(s)\mathrm{d}s\right]\mathrm{d}y\right\}.}$$

26.2 Perspectives

The Burgers' equation is nice: it is nonlinear but causal. It admits analytical solutions. The regularised equation is even better: the solution never blows up, and an exact, analytical formula exists. It was once thought that such simple equations could shed light on turbulence. Turbulence manifests itself in the Navier–Stokes equations

$$\begin{aligned}
\frac{\partial \boldsymbol{u}}{\partial t} + \boldsymbol{u} \cdot \nabla \boldsymbol{u} &= -\nabla p + D\nabla^2 \boldsymbol{u}, \\
\nabla \cdot \boldsymbol{u} &= 0, \\
\boldsymbol{u}(\boldsymbol{x}, t) &\in \mathbb{R}^3, \\
\boldsymbol{x} &\in \Omega \subset \mathbb{R}^3.
\end{aligned}$$

Unfortunately, the Burgers model does not shed light on the following open question:

> Prove or give a counter-example of the following statement: In three space dimensions and time, given an initial velocity field, there exists a vector velocity $\boldsymbol{u}(\boldsymbol{x}, t)$ and a scalar pressure field $p(\boldsymbol{x}, t)$, which are both smooth and globally defined, that solve the Navier-Stokes equations.

Correct answers, sent on a postcard or otherwise, to the Clay Mathematics Institutes in Cambridge MA will earn you a one million dollar prize.

Appendix A

List of .m codes

- brownian.m – solves a stochastic differential equation for a particle doing Brownian motion on a finite interval.
- diffusion__one__d.m – solves the one-dimensional diffusion equation with periodic boundary conditions and a spectral method.
- cahnhilliard__solve.m – solves the two-dimensional quasi-linear Cahn–Hilliard equation with periodic boundary conditions and a spectral method.
- diffusion__constant__temp.m – computes the series solution for the diffusion equation with zero boundary conditions and a constant initial temperature.
- wave__solve__exact.m – solves the wave equation using the d'Alembert formula for certain piecewise initial data (Ch. 14, Sec. 14.2).
- integrate__sde.m – a one-dimensional finite difference solver for the wave equation. Needs C^2 initial data.
- burgers.m – a one-dimensional finite difference code for the Burgers equation. It cannot handle shocks.