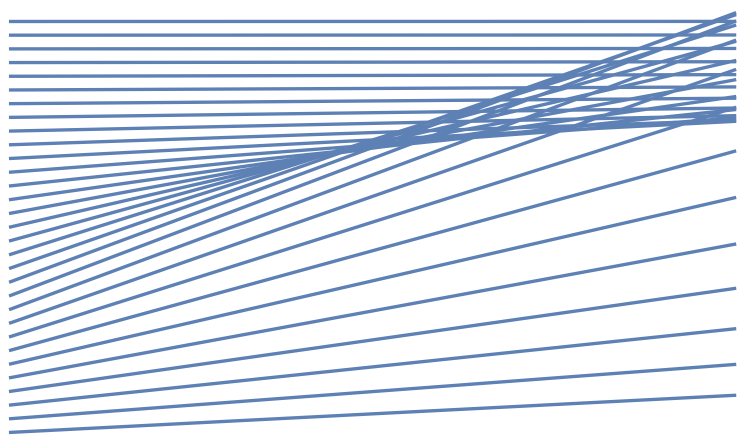
Partial Differential Equations

From classical to modern

Gerald Teschl



Gerald Teschl
Fakultät für Mathematik
Oskar-Mogenstern-Platz 1
Universität Wien
1090 Wien, Austria

E-mail: Gerald.Teschl@univie.ac.at
URL: `http://www.mat.univie.ac.at/~gerald/`

2010 Mathematics subject classification. 35-01

Abstract. This textbook provides a brief introduction to Partial Differential Equations. The first part is intended as a first introduction and does neither require any functional analytic tools nor Lebesgue or Sobolev spaces. The second part deals with advanced techniques from functional analysis.

Keywords and phrases. Method of characteristics, Cauchy–Kovalevskaya theorem, Separation of variables, Fourier transform, Dispersion, Laplace equation, Harmonic functions, Dirichlet problem, Heat equation, Wave equation Sobolev spaces, Elliptic regularity, Operator semigroups, Reaction diffusion equations, Calculus of variations, Nonlinear Schrödinger equation, Strichartz estimates

Typeset by $\mathcal{AMS}$-LaTeX, Ti*k*Z, and Makeindex.
Version: April 2, 2024

Contents

Preface

The present manuscript are the notes for my courses *Partial Differential Equations* and *Advanced Partial Differential Equations* given at the University of Vienna in winter 2020 and 2021, respectively. When preparing a course on such a classical topic, the first question is of course why another set of notes. However, while this question is common, so is the answer: Because I could not find existing notes meeting all my expectations. To help you find out if the present notes will meet yours, I will describe its content and explain some of my motivations below.

Aims and features

The first question one has to answer when designing such a course is the question about the prerequisites and the overall direction (in most cases these items will in fact already have been decided for you). In my case the prerequisites for the first part are only multi-variable calculus (including some elementary facts about ordinary differential equations). Neither functional analysis nor measure theory (Lebesgue spaces) are assumed to be known (and it is also not the purpose to introduce them during the first part). In particular, rushing through the classical theory and getting to functional analytic tools as fast as possible is *not* the aim of the present course. Nevertheless, this is still supposed to be a modern introduction and hence I tried to focus on those topics which either enhance the understanding of partial differential equations (PDE) or can still be considered useful nowadays. The main features from my perspective are the following:

- An important point I want to make is that this subject is part of Analysis. Complete and detailed proofs of all results are a core ingredient, even if some of them are challenging. Nevertheless, I have strived to avoid mathematical *elegance* which might impress the expert reader by its brevity but leaves the beginner puzzled.
- I wanted to start gentle but still cover a good amount of key results during the course. Hence I spent a lot of time streamlining the presentation and searching for the simpelst possible proofs. I found it surprising that some of them, even though already quite old, still did not make it into the standard textbook literature. The most striking example being the proof for the strong maximum principle for general parabolic equations.
- Math books are rarely read (or taught) cover to cover and hence the material is presented to minimize interdependence such that certain parts can be easily skipped.
- Examples within the text. From easy standard examples, over neat applications, to (sometimes tricky) counterexamples, explaining why results fail once some of their assumptions are dropped.
- Problems. Again from easy standard *drill* problems, over neat applications, to additional results extending the material presented in the main text. All problems are student-tested and the harder ones usually have hints to get you started.
- Don't be afraid of special functions: The only difference between trigonometric and Bessel functions is that your pocket calculator might not have a button for the latter. However, all modern computer algebra systems have them built in and a myriad of formulas is available. While these formulas can be readily looked up in standard references like the NIST Digital Library of Mathematical Functions[1], I still find it interesting how they can be derived. Hence for every formula there is both a reference and a problem (in case you want to get your hands dirty).

Contents

First Part

I have chosen to start with the method of characteristics not only because this stands historically at the beginning of PDE theory, but also because it serves as a gentle entry point. In particular, this already allows us to discuss some interesting examples and get a first flavor of what PDEs are about. We

[1] `https://dlmf.nist.gov`

also have a first look at second order equations starting with the case of two variables and discussing canonical forms as well as the common classification scheme.

Of course the method of characteristics raises the question about a general existence and uniqueness result for the Cauchy problem (an anlog of the Picard–Lindelöf theorem for ordinary differential equations (ODE)). This endeavor turned out too much to hope for and consequently had to be shelved. Still I find it interesting to understand to what extend this is possible and where it fails. For this reason the Cauchy–Kovalevskaya theorem is discussed and it is applied to the classical equations (the Holy Trinity) in order to explain its limitations. It requires some familiarity with real analytic functions; the necessary background is collected in an appendix. Finally, there is Lewy's example to show that one cannot go beyond. This chapter (in fact the first two chapters) is independent of the rest and could be skipped.

The third classical method is separation of variables. It is possible (and indeed quite frequent) to do a full course only using this method. I have decided to devote limited time to this topic dealing only with the heat and wave equation (in one spatial variable) in somewhat detail. This is done with the extension to higher dimensions in mind and hence special emphasis on the fundamental solution, the maximum principle, and energy methods is made. In addition to an interval, the case of a rectangle and a disc are also briefly discussed.

The fourth chapter introduces the Fourier transform as the premier tool for evolution problems on the real line. In addition to the heat and wave equation, we discuss dispersion including the stationary phase method. The emphasis is on a precise treatment avoiding vague arguments often found in this context. There is also a brief section on symmetry groups. There is nothing substantial here, but I at least wanted the reader to know that these methods exist.

The fifth chapter makes the step to arbitrary dimensions and in principle one could even start here. We begin with harmonic functions, discuss the Newton potential, Green's function, and the Dirichlet principle. Many authors stop at this point since the modern way of solving the Dirichlet problem is via the calculus of variations, which is far beyond the current toolbox. However, I found it disappointing to say so much about the structure of solutions without ever establishing existence (except for a ball and a half space). Hence I have decided to include Perron's method to solve the Dirichlet problem. Finally, there is a glimpse at general elliptic equations including the strong maximum principle.

The sixth chapter deals with the heat equation. To this end the Fourier transform is extended to higher dimensions. There is also a brief discussion on the fundamental solution for a domain which is used to motivate the mean value formula. I prefer this approach over (the slightly shorter approach of) postulating the formula and then verifying it via an opaque calculation. As a consequence we obtain the maximum principles, which are extended to unbounded domains including a discussion of Tikhonov's counterexample. Again there is a glimpse at general parabolic equations including the strong maximum principle.

The seventh chapter derives the Kirchhoff and the Poisson formula for the wave equation using the Fourier transform. The method of spherical averages is used to obtain formulas for arbitrary dimensions. Also energy methods are briefly discussed.

Second Part

The second part covers some selected topics using advanced tools from functional analysis and measure theory. It is to a large part independent of the first part but of course assumes some basic familiarity with partial differential equations. Moreover, the reader is assumed to be familiar with the Lebesgue integral as well as Lebesgue spaces. The relevant results are collected in Section B.2. Similarly, we require some basic familiarity with Functional Analysis. Again the required results are collected in the appendix.

The eighth chapter is intended as a motivation for what is to come. Here constant coefficient equations on $\mathbb{R}^n$ are discussed by virtue of the Fourier transform. Weak derivatives and the semigroup point of view for evolution equations are introduced in this context.

Since the course is not limited to problems on $\mathbb{R}^n$, general Sobolev spaces are introduced in chapter nine. While you will not see a single partial differential equation in this chapter, it will lay they technical foundation for the following chapters. Apart from the basic results, trace operators as well as the usual embedding theorems are established.

The tenth chapter discusses elliptic problems via the Lax–Milgram approach. We start out with the Poisson problem and introduce the key ideas using this simplest case. We also discuss domain issues for the associated operator, a topic which is sometimes swept under the carpet. Everything is then extended to elliptic operators. In particular, the maximum principle and elliptic regularity are discussed in detail.

The eleventh chapter turns to evolution problems. I have decided to take the abstract approach via operator semigroups. Since this is an extremely versatile tool, I feel that it is worth while covering in such a course. To facilitate the extra abstract hurdle I start with discussing abstract Cauchy problems before turning to the question of generator theorems. While contraction semigroups (on Hilbert spaces) would be sufficient for the intended applications, I have decided to develop the general theory, since the simplifications would only be minimal. Parabolic and hyperbolic equations are then handled as straightforward applications.

The twelfth chapter leaves the realm of linear equations. We look at the simplest case of semilinear equations again within the framework of semigroup theory. Apart from the basic local existence result, blowup criteria and persistence of regularity are discussed. One prototypical example, reaction diffusion equations, are discussed in some detail (including a comparison principle).

The thirteenth chapter aims at nonlinear elliptic problems via the calculus of variations. The direct method is discussed including the case of constraints.

The final chapter is intended as an outlook to even more advanced techniques. We discuss one prominent example, the nonlinear Schrödinger equation. We show how the basic semigroup theory provides local existence in $H^r(\mathbb{R}^n)$ provided $r > \frac{n}{2}$ exploiting the algebra structure available in this case. We discuss global existence and blowup in detail for the one-dimensional case, since this explains all the main ideas and since the previous local result covers H^1 in this case. Establishing local existence in L^2 and H^1 (without the restriction to one dimension) requires the use of Strichartz estimates. Of course this raises the technical bar even further (in particular, since a clean mathematical treatment requires the Bochner integral), but the results will hopefully reward the reader for the efforts.

Problems relevant for the main text are marked with a "*".

The manuscript is updated whenever I find some errors and extended from time to time. Hence you might want to make sure that you have the most recent version, which is available from

`http://www.mat.univie.ac.at/~gerald/ftp/book-pde/`

Please do not redistribute this file or put a copy on your personal webpage but link to the page above.

Acknowledgments

I wish to thank my readers, Teresa Engelbrecht, Darius Erös, Viktoria Hablas, Bastian Hilder, Jakob Holböck, Janik Hrubant, Daniel Karner, Clemens Kinn, Jan, Patrick Pollek, Emma Schwentner, Jonas Quenzer, Sophie Rosenmeier, Johannes Rotheneder, Miloš Vujičić, Eric Wahlén, Anton Zand who have pointed out several typos and made useful suggestions for improvements. Moreover, I am most grateful to Iryna Karpenko who read large parts of the manuscript, provided long lists of typos, and also contributed some of the problems.

The graphics in these notes were produced using Ti*k*Z and *Mathematica*.[2]

Finally, no book is free of errors. So if you find one, or if you have comments or suggestions (no matter how small), please let me know.

Gerald Teschl

Vienna, Austria
January, 2022

[2] *Mathematica*® is a registered trademark of Wolfram Research, Inc.

Part 1

Classical Partial Differential Equations

Chapter 1

The method of characteristics

1.1. Introduction and motivation

As you might have already guessed from the name, a **partial differential equation (PDE)** is an equation involving an unknown function u and its derivatives. Unlike the case of ordinary differential equations, u will depend on more than one variable and hence we have to deal with partial derivatives. As with ordinary differential equations, the **order** of the differential equation is the order of the highest derivative appearing in the equation.

To get a first feeling for partial differential equations, let us look at a specific example. The simplest example is of course a first order homogenous linear equation (with constant coefficients) for an unknown (differentiable, C^1 say) scalar function u depending on two independent variables (x, y):

$$a\, u_x + b\, u_y + c\, u = 0. \tag{1.1}$$

Here we use the convenient notation

$$u_x := \frac{\partial u}{\partial x}, \qquad u_y := \frac{\partial u}{\partial y} \tag{1.2}$$

for partial derivatives. To ensure that this is really a differential equation, we need to assume that at least one of the coefficients a or b is nonzero. We will assume $b \neq 0$ and then divide by b such that we can assume $b = 1$ without loss of generality. Moreover, to begin with we will also assume $c = 0$. Interpreting the second variable y as time t, this is the famous **transport equation**

$$u_t + a\, u_x = 0. \tag{1.3}$$

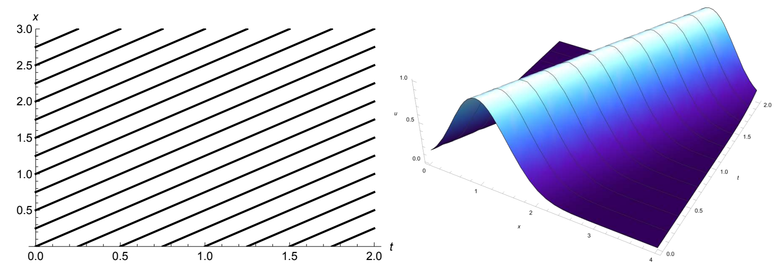


Figure 1.1. Transport equation: Characteristics (left) and solution (right) for a Gaussian initial datum

The equation tells us that a certain directional derivative (namely the derivative in the direction $(1, a)$ in the (t, x)-plane) vanishes. Hence a solution u does not change in this direction. So if $u(t_0, x_0) = u_0$ we also have $u(t_0 + s, x_0 + as) = u_0$ for all $s \in \mathbb{R}$. Consequently, if we prescribe the values of u at some time t_0,

$$u(t_0, x) = g(x), \tag{1.4}$$

then we obtain the value of u at a given point $(t, x) \in \mathbb{R}^2$ by following the characteristic line (along which u is constant) $(t + s, x + as)$ through this point backwards until we hit the line $t = t_0$, where we have specified our initial condition. This implies the condition $t + s = t_0$ from which we get

$$u(t, x) = u(t+s, x+as)\big|_{s=t_0-t} = u(t_0, x-a(t-t_0)) = g(x-a(t-t_0)). \tag{1.5}$$

Hence the effect of our equation is to take the **initial condition** g and translate (*transport*) it by a distance of $a(t - t_0)$. Hence the profile g travels to the right with constant speed a.

Note that we could specify the initial conditions along pretty much any curve as long as the characteristic line through any point (t, x) has precisely one point of intersection with this curve. In particular, we cannot prescribe the value of u on (part of) a characteristic line since this will lead to a contradiction in general. Moreover, even if it works, we could loose uniqueness — Problem 1.1.

Even without specifying an initial condition, we could read our findings as any solution being of the form $u(t, x) = g(x - at)$, where g is an arbitrary differentiable function. Hence in contradistinction to ordinary differential equations, the general solution of a partial differential equations will typically depend on unknown functions, rather than unknown constants.

Motivated by our success, let us return to the original problem with $c \neq 0$:

$$u_t + a\, u_x + c\, u = 0. \tag{1.6}$$

Now u is no longer constant along characteristic lines, but we still know how it changes. More precisely, setting $z(s) := u(t+s, x+as)$ our partial differential equations gives us an ordinary differential equation for the change of u along characteristic lines:

$$\begin{aligned}\dot{z}(s) &= u_t(t+s, x+as) + a\, u_x(t+s, x+as) = -c\, u(t+s, x+as)\\ &= -c\, z(s),\end{aligned} \tag{1.7}$$

where the dot indicates a derivative with respect to s. Solving this differential equation we obtain

$$z(s) = z(s_0)\mathrm{e}^{-c(s-s_0)}. \tag{1.8}$$

Moreover, choosing s_0 such that $t + s_0 = t_0$ we obtain

$$\begin{aligned}u(t,x) &= z(0) = z(s_0)\mathrm{e}^{cs_0} = u(t_0, x - a(t-t_0))\mathrm{e}^{-c(t-t_0)}\\ &= g(x - a(t-t_0))\mathrm{e}^{-c(t-t_0)}.\end{aligned} \tag{1.9}$$

This clearly raises the question how far this approach can be pushed. Well, first of all note that we can easily replace the term $c\,u$ by an arbitrary function $c(t,x,u)$. Evaluating u along the characteristic lines as before gives

$$\dot{z} = -c(t+s, x+as, z), \tag{1.10}$$

which has to be regarded as an ordinary differential equation for z as a function of s with t and x as fixed parameters. The only remaining problem is solving this ordinary differential equation. For example, this can be used to solve the inhomogeneous problem where $c(t,x,u) = c\,u + f(t,x)$, see Problem 1.2.

But what if the coefficients are no longer constant? We will consider this case in the next section.

Problem* 1.1. *Find two different solutions of*

$$u_t + u_x = 1$$

with initial conditions $u(s,s) = s$ along a characteristic line.

Problem 1.2. *Establish the solution formula*

$$u(t,x) = g(x - at) + \int_0^t f(s, x + a(s-t))ds \tag{1.11}$$

for the inhomogeneous transport equation

$$u_t + a\, u_x = f(t,x), \qquad u(0,x) = g(x). \tag{1.12}$$

Problem 1.3. *Derive* ***d'Alembert's formula***

$$u(t,x) = \frac{g(x+ct) + g(x-ct)}{2} + \frac{1}{2c}\int_{x-ct}^{x+ct} h(y)dy$$

for the solution of the wave equation

$$u_{tt} - c^2 u_{xx} = 0, \qquad u(0,x) = g(x), \quad u_t(0,x) = h(x).$$

(Hint: Note that the wave equation can be factorized as two transport equations

$$(\partial_t - c\partial_x)(\partial_t + c\partial_x)\, u = (\partial_t + c\partial_x)(\partial_t - c\partial_x)\, u = 0$$

and use Problem 1.2.)

1.2. Semilinear equations

Let U be some nonempty open subest of $\mathbb{R}^n$. We now look at the more general problem of a **semilinear equation** for an unknown differentiable scalar function u depending on n independent variables $x \in U$:

$$a_1(x)u_{x_1} + \cdots + a_n(x)u_{x_n} + c(x,u) = 0. \tag{1.13}$$

Here the name semilinear refers to the fact that the equation is linear with respect to the (highest) derivatives with the coefficients depending only on the independent variables, whereas the lower oder terms are allowed to depend on u (and possibly on lower order derivatives, if there were any).

Again we can regard this as a directional derivative

$$a(x) \cdot \nabla u = -c(x,u), \tag{1.14}$$

where we use the notation

$$\nabla u := (u_{x_1}, \dots, u_{x_n}) \tag{1.15}$$

for the **gradient** of u and a dot to indicate the scalar product in $\mathbb{R}^n$. If $c = 0$ we can interpret this as the level sets of the solution u being tangential to the vector field a. Hence it suggests itself to evaluate the solution u along characteristic curves, which are now given as the integral curves $x(s)$ of the vector field a, that is, the solutions of the system of ordinary differential equations

$$\dot{x} = a(x). \tag{1.16}$$

Then, evaluating u along this characteristic curves and setting $z(s) := u(x(s))$, we obtain from the chain rule

$$\dot{z} = a(x(s)) \cdot \nabla u(x(s)) = -c(x(s), z). \tag{1.17}$$

This clearly gives a procedure for solving (1.13) with given initial conditions on an $n-1$ dimensional surface $\Gamma \subset U$.

To describe it in detail we first need to describe Γ. Recall that there are basically two ways of describing a hyper-surface: Either implicitly as the zero set of a given function or explicitly given some parametrization. By the implicit function function theorem both ways are of course equivalent at least locally. Here we will assume that we have a parametrization $\Gamma = \{h(y) | y \in$

$V\}$, where V is some nonempty open subset in $\mathbb{R}^{n-1}$ and $h \in C^1(V, \Gamma)$ is a differentiable homeomorphism whose Jacobi[1] matrix $\frac{\partial h}{\partial y}$ is injective for every $y \in V$. In particular, the tangent vectors $\frac{\partial h}{\partial y_j}(y)$, $1 \le j \le n-1$, are linearly independent and span the tangent space of Γ at $x_0 := h(y)$. A vector $\nu(x_0)$ orthogonal to these tangent vectors is called the **normal vector** of Γ. If we normalize ν it is unique up to its direction. If Γ is defined implicitly, then the normal vector is given by the gradient. In addition to h let us assume that also $a \in C^1(U, \mathbb{R}^n)$ and $c \in C^1(U \times \mathbb{R}, \mathbb{R})$.

First solve (1.16) to obtain the characteristic curves $\xi(s, y)$ corresponding to initial conditions $\xi(0, y) = h(y) \in \Gamma$. Please recall that by standard ODE theory our assumption $a \in C^1(U, \mathbb{R}^n)$ implies that for every y there is a unique solution which exists at least locally for s in a neighborhood of 0 (the Picard–Lindelöf theorem). Moreover, this solution depends differentiable on the initial condition and the solution ξ will be C^1 in a neighborhood of $(0, y_0)$ for any $y_0 \in V$ (e.g. Theorem 2.10 from **[33]**). Now fix $x_0 = h(y_0) \in \Gamma$ and assume that $a(x_0)$ is not in the tangent space of Γ at x_0. That is, assume the **non-characteristic condition**

$$\nu(x_0) \cdot a(x_0) \neq 0, \tag{1.18}$$

where $\nu(x_0)$ is the normal vector of Γ at x_0. Then the Jacobi matrix of $\xi(s, y)$ at $(0, y_0)$ is given by

$$\frac{\partial \xi}{\partial(s, y)}(0, y_0) = \left(a(x_0), \frac{\partial h}{\partial y_1}(y_0), \cdots, \frac{\partial h}{\partial y_{n-1}}(y_0) \right) \tag{1.19}$$

and will be invertible thanks to (1.18). Hence by the inverse function theorem $\xi(s, y)$ will be a diffeomorphism when restricted to a small neighborhood $(-\varepsilon, \varepsilon) \times V(y_0) \to U(x_0)$ by the inverse function theorem. Note that geometrically this condition is clear since the characteristic curves must not stay within Γ for our procedure to work. Moreover, if they intersect Γ non-tangentially, they form a mesh which covers a neighborhood of x_0 without leaving any holes.

Hence for any $x \in U(x_0)$ we can find some corresponding $(s, y) \in (-\varepsilon, \varepsilon) \times V(x_0)$ such that the characteristic curve $\xi(., y)$ starts at $h(y) \in \Gamma$ and hits $x = \xi(s, y)$. Finally, compute z by solving (1.17) with $x(.) = \xi(., y)$ and initial condition $z(0) = u(\xi(0, y)) = u(h(y)) = g(h(y))$, from which $u(x) = z(s)$ follows.

In summary (compare also Figure 1.2):

- Solve $\dot{x} = a(x)$ with initial conditions $x(0) = h(y) \in \Gamma$ to obtain the characteristic curves $\xi(s, y)$.
- Solve $\dot{z} = -c(\xi(s, y), z)$ with initial condition $z(0) = g(h(y))$.

[1] Carl Gustav Jacob Jacobi (1804–1851), German mathematician

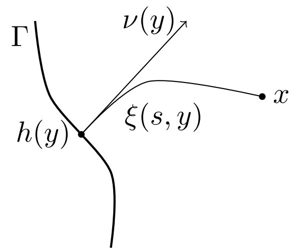


Figure 1.2. Cartoon for the method of characteristics: To find the value of the solution $u(x)$ we first find the characteristic curve ξ which leads from Γ to x and then we compute the values of u along ξ.

- Solve the nonlinear system $\xi(s, y) = x$ for (s, y) and insert the solution into $u(x) = z(s, y)$.

As pointed out above, the last step will be possible in a neighborhood of x_0 by the inverse function theorem provided (1.18) holds.

Theorem 1.1 (Method of Characteristics – semilinear case)**.** *Let U be some nonempty open subest of $\mathbb{R}^n$. Suppose $a \in C^1(U, \mathbb{R}^n)$ and $c \in C^1(U \times \mathbb{R}, \mathbb{R})$. Let $\Gamma \subset U$ be an $n-1$ dimensional C^1 surface such that the non-characteristic condition* (1.18) *holds at some point $x_0 \in \Gamma$. Then, the semi-linear partial differential equation* (1.13) *has a unique solution u in a sufficiently small neighborhood U_0 of x_0 for any given initial data $g \in C^1(\Gamma \cap U_0)$. This solution can be obtained by solving the characteristic differential equations as described above.*

Proof. Of course our procedure gives the necessary form of a solution and hence it remains to verify that the function u constructed in this way is indeed a solution. First of all note that u is C^1 and by the chain rule we have $\frac{\partial u}{\partial x} = \dot{z}\frac{\partial s}{\partial x} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial x}$. Now observe that, by the inverse function theorem, the Jacobi matrix of the inverse of the map $\xi(s, y)$ is the inverse of the Jacobi matrix of $\xi(s, y)$. Furthermore, the first column of the Jacobi matrix of $\xi(s, y)$ is $a(\xi(s, y)) = a(x)$ and hence its inverse maps $a(x)$ to the vector $(1, 0, \dots, 0)$ implying $\frac{\partial u}{\partial x}(x) \cdot a(x) = \dot{z}(s, y) = -c(x, u(x))$. □

Example 1.1. Let $\Gamma = \{x \in \mathbb{R}^2 | x_2 = x_1^2\}$ and solve

$$u_{x_1} + x_1 u_{x_2} + x_1 u = 0, \qquad u(y, y^2) = 1.$$

The normal vector is given by the gradient of $x_2 - x_1^2$,

$$\nu(y, y^2) = \frac{1}{\sqrt{1 + 4y^2}} \begin{pmatrix} -2y \\ 1 \end{pmatrix},$$

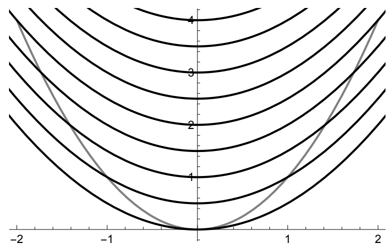


Figure 1.3. Initial surface Γ (black) and characteristics for Example 1.1

and the non-characteristic condition reads ($a(x) = (1, x_1)$)

$$\nu(y, y^2) \cdot a(y, y^2) = \frac{1}{\sqrt{1+4y^2}} \begin{pmatrix} -2y \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ y \end{pmatrix} = \frac{-y}{\sqrt{1+4y^2}} \neq 0,$$

which is satisfied for $y \neq 0$.

The characteristic equations read

$$\dot{x}_1 = 1, \quad \dot{x}_2 = x_1, \qquad \dot{z} = -x_1 z$$

and the initial conditions are $x_1(0) = y$, $x_2(0) = y^2$, $z(0) = 1$. The solution is

$$\xi_1(s, y) = y + s, \quad \xi_2(s, y) = y^2 + ys + \frac{s^2}{2}, \qquad z(s, y) = \mathrm{e}^{-ys - s^2/2}.$$

Solving $x = \xi(s, y)$ gives $(s, y) = (x_1 \mp \sqrt{2x_2 - x_1^2}, \pm\sqrt{2x_2 - x_1^2})$ for $\pm y > 0$. Consequently

$$u(x) = z(s) = \mathrm{e}^{x_2 - x_1^2}.$$

The characteristics are shown in Figure 1.3. Observe that the characteristics are indeed tangential to Γ at $(0, 0)$. Moreover, while the characteristic curve starting at $(0, 0)$ leaves Γ, the characteristics still do not cover a neighborhood of this point. In particular, note that the characteristics starting on Γ do not cover the region $x_2 < \frac{x_1^2}{2}$ and hence our method cannot give us the values of u in this region. Indeed, our system $x = \xi(s, y)$ only has a (real-valued) solution provided $x_2 \geq \frac{x_1^2}{2}$. The fact that the final formula for u makes sense on all of $\mathbb{R}^2$ just means that the solution we have found in the region $x_2 \geq \frac{x_1^2}{2}$ has an analytic continuation to all of $\mathbb{R}^2$. Finally, observe that the characteristic curves intersect Γ twice and such a situation will typically lead to a contradiction if you try to prescribe the values at both intersection points. Indeed, in our case this will work if we prescribe symmetric initial conditions $g(y) = g(-y)$, while it will fail otherwise. For example, if we look for a solution satisfying $u(y, y^2) = y$ and start our characteristics on $\Gamma_\pm = \{(x_1, x_1^2) | \pm x_1 > 0\}$ we get $u_\pm(x) = \pm\sqrt{2x_2 - x_1^2}\mathrm{e}^{x_2 - x_1^2}$. Hence $u_\pm(y, y^2) = \pm|y|$ and $u_\pm(y, y^2) = y$ on $\Gamma_\pm$ while $u_\pm(y, y^2) = -y$ on $\Gamma_\mp$. ⋄

Example 1.2. The **Liouville equation**[2]

$$u_t + v\, u_x + \frac{1}{m} F(x) u_v = 0, \qquad u(0, x, v) = g(x, v),$$

describes the density $u(t, x, v)$ of an ensemble of identical particles of mass m in phase space $(x, v) \in \mathbb{R}^2$ (position and velocity) at time $t \in \mathbb{R}$.

The characteristic equations are

$$\dot{t} = 1, \qquad \dot{x} = v, \qquad \dot{v} = \frac{1}{m} F(x), \qquad \dot{z} = 0.$$

Hence the density is constant along the trajectories of the individual particles given by Newton's[3] second law of motion $m\ddot{x} = F(x)$. ◇

Of course this method reduces our original partial differential equation to a system of ordinary differential equations and, provided we can solve this system, we get an implicit form of the solution. If we can also solve the resulting nonlinear system of equations, we even get an explicit solution. So from this point of view we can record this as a full success. Moreover, the fact that we only get local solvability is not surprising either, since we do not have global solutions for ordinary differential equations in general either. However, note that even if the system of ordinary differential equations has global solutions, we might still have further trouble since the characteristic curves might not reach all points.

Problem 1.4. *Let* $\Gamma = \{(t, x) | t = 0\}$ *and solve*

$$u_t + t\, u_x = 0, \qquad u(0, x) = g(x).$$

Problem 1.5. *Let* $\Gamma = \{(t, x) | t = 0\}$ *and solve*

$$u_t + x\, u_x = 0, \qquad u(0, x) = g(x).$$

Problem 1.6. *Let* $\Gamma = \{x \in \mathbb{R}^2 | x_2 = \mathrm{e}^{x_1}\}$ *and solve*

$$x_1 u_{x_1} + x_2 u_{x_2} + (x_1 - 2) u = 0, \qquad u(y, \mathrm{e}^y) = y.$$

Problem 1.7. *Let* $\Gamma = \{x \in \mathbb{R}^3 | x_3 = 0\}$ *and solve*

$$x_1 u_{x_1} + x_2 u_{x_2} + x_1 x_2 u_{x_3} = 0, \qquad u(y_1, y_2, 0) = y_1^2 + y_2^2.$$

Note that the solution is constant along characteristics. Nevertheless the formula for the solution gives $u(1, 1, 1) = -2$ *while the initial conditions are nonnegative. How can this be explained?*

Problem 1.8. *Solve the Liouville equation for the harmonic oscillator* $F(x) = -m\omega^2 x$, $\omega > 0$, *with the initial density* $g(x, v)$.

---

[2] Joseph Liouville (1809–1882), French mathematician and engineer

[3] Isaac Newton (1643–1727), English mathematician, physicist, astronomer, and theologian

Problem 1.9. *Find the general solution of* ***Euler's homogeneity relation***

$$x_1 u_{x_1} + \cdots + x_n u_{x_n} = c\, u.$$

Problem 1.10. *Let* $\Gamma = \{(x, y) \in \mathbb{R}^2 | y = 0\}$ *and solve*

$$u_x - u_y = 1, \qquad u v_x - v_y = v, \qquad u(x, 0) = x, \quad v(x, 0) = g(x).$$

1.3. Quasilinear equations

An important class of first order equations are the **scalar conservation laws**

$$u_t + F(u)_x = 0, \qquad u(0, x) = g(x), \tag{1.20}$$

where $F \in C^2(\mathbb{R})$ is some given function. Explicitly the equation reads

$$u_t + F'(u) u_x = 0. \tag{1.21}$$

The idea behind this equation is that $u(t, x)$ corresponds to the density of some substance at time t and a location x. If one assumes that the quantity is conserved (i.e. it is neither destroyed nor created) within any given interval (a, b) as time evolves, the total amount of the substance within this interval $\int_a^b u(t, x) dx$ can only change by the net flux through the two endpoints. This net flux is given by $v(t, a)u(t, a) - v(t, b)u(t, b)$, where v is the velocity with which the substance is flowing. If the velocity can be expressed as a function of u, the expression $F(u) = v(u)u$ is called the **flux function**. In this case we obtain

$$\frac{d}{dt} \int_a^b u(t, x) dx = F(u(t, a)) - F(u(t, b)). \tag{1.22}$$

Interchanging integration and differentiation on the left-hand side and invoking the fundamental theorem of calculus on the right-hand side gives

$$\int_a^b \big(u_t(t, x) + F(u(t, x))_x\big) dx = 0. \tag{1.23}$$

Since this should hold for arbitrary intervals, we obtain our conservation law. Note that one could also assume that the velocity field is a given function $v(t, x)$ or, even more general, that we have a relation $v(t, x, u)$. This would lead to a flux function $F(t, x, u)$ and a corresponding equation

$$u_t + F(t, x, u)_x = u_t + F_x(t, x, u) + F_u(t, x, u) u_x = 0. \tag{1.24}$$

Example 1.3. Assuming that the velocity is constant, $v(u) = a$, one obtains the transport equation. ◇

At first sight this type of equation does not fit within the framework of the previous section since a was not allowed to depend on u. However, it is straightforward to check that the method of characteristics extends to **quasilinear equations** of the type

$$a(x,u)\cdot\nabla u = -c(x,u). \tag{1.25}$$

Here the name quasilinear refers to the fact that the equation is linear with respect to the (highest) derivatives, where in contradistinction to a semiliner equation, both the coefficients and the inhomogeneous term are now allowed to depend on u (and possibly on lower order derivatives).

The characteristic equations now read

$$\dot{x} = a(x,z), \qquad \dot{z} = -c(x,z) \tag{1.26}$$

and the only difference is that this is now a coupled system which has to be solved simultaneously. Moreover, the non-characteristic condition now reads

$$\nu(x_0)\cdot a(x_0, g(x_0)) \neq 0 \tag{1.27}$$

and also involves the initial condition g and not only the surface Γ. Other than that the argument carries over verbatim and we obtain:

Theorem 1.2 (Method of Characteristics – quasilinear case)**.** *Let U be some nonempty open subest of $\mathbb{R}^n$. Suppose $a\in C^1(U\times\mathbb{R},\mathbb{R}^n)$ and $c\in C^1(U\times\mathbb{R},\mathbb{R})$. Let $\Gamma\subset U$ be an $n-1$ dimensional C^1 surface and $g\in C^1(\Gamma)$ such that the non-characteristic condition* (1.27) *holds at some point $x_0\in\Gamma$. Then, the quasi-linear partial differential equation* (1.25) *has a unique solution u in a sufficiently small neighborhood U_0 of x_0 satisfying the given initial data g. This solution can be obtained by solving the characteristic differential equations as described above.*

This method, developed in the 18th century, goes back to Charpit de Villecourt[4] and Lagrange[5] and is also known as Lagrange–Charpit method.

In the case of our conservation law the non-characteristic condition (1.27) reads

$$\begin{pmatrix}1\\0\end{pmatrix}\cdot\begin{pmatrix}1\\F'(g(x_0))\end{pmatrix} = 1\neq 0 \tag{1.28}$$

and hence is always satisfied. The characteristic equations read

$$\dot{t}=1, \qquad \dot{x}=F'(z), \qquad \dot{z}=0. \tag{1.29}$$

The solution is

$$t(s)=s, \qquad x(s)=x_0+F'(g(x_0))s, \qquad z(s)=g(x_0) \tag{1.30}$$

and the solution of our conservation law is implicitly given by

$$u(t,x)=z(s)=g(x_0)=g(x-F'(u(t,x))t). \tag{1.31}$$

[4]Paul Charpit de Villecourt (ca. 1750–1784), French mathematician
[5]Joseph Louis Lagrange (1736–1839), French mathematician

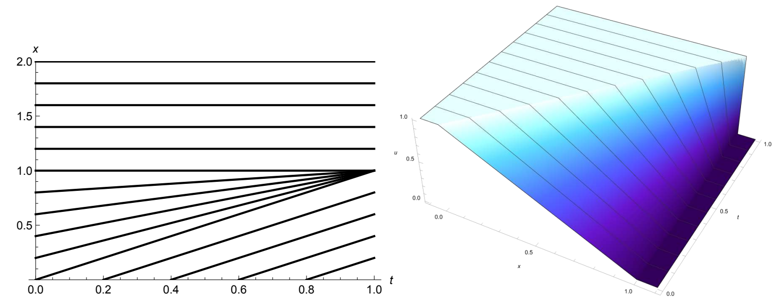


Figure 1.4. Shock wave for Burgers' equation: Characteristics (left) and solution (right) for steplike initial data

So the situation is quite similar to the transport equation with the only difference that the speed depends on the hight of the solution. Moreover, note that while in the semi-linear case, the uniqueness part of the Picard–Lindelöf theorem[6] implies that different characteristics $x(s)$ cannot intersect, this no longer applies here.

Example 1.4. Assuming that the velocity is proportional to the density, we obtain the **Burgers' equation**[7] corresponding to $F(u) = \frac{u^2}{2}$:

$$u_t + uu_x = 0.$$

Let us try to solve it with the initial condition

$$g(x) = \begin{cases} 1, & x \le 0, \\ 1 - x, & 0 < x < 1, \\ 0, & 1 \le x. \end{cases}$$

Of course this initial condition is only piecewise C^1 but let us ignore this fact for now (one could work with a smoothed out version, but this will only make the formulas more messy). So g equals one on the left, zero on the right, and has a linear transition region in the middle. In particular, since the higher values travel faster, they will eventually collide with the lower ones. In fact, the characteristics starting on the interval $0 \le x \le 1$ are given by

$$x(s) = x_0 + (1 - x_0)s$$

and they all hit at the point $x(1) = 1$. In fact, the linear transition will become steeper and steeper until it gets vertical at $t = 1$. This is illustrated in Figure 1.4. Hence at time $t = 1$ the solution has formed a discontinuity,

[6]Émile Picard (1856–1941), French mathematician
[6]Ernst Lindelöf (1870–1946), Finnish mathematician
[7]Jan Burgers (1895–1981), Dutch physicist

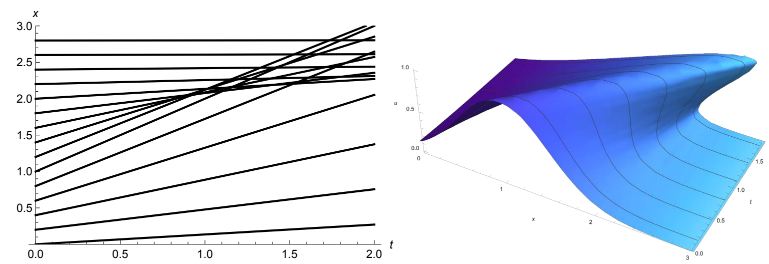


Figure 1.5. Wave breaking for Burgers' equation: Characteristics (left) and solution (right) for a Gaussian initial data (drawn beyond its classical time of existence)

being 1 for $x < 1$ and 0 for $x > 1$. This is known as the formation of a **shock**. Explicitly we have

$$u(t,x) = \begin{cases} 1, & x \le t, \\ \frac{1-x}{1-t}, & t < x < 1, \\ 0, & 1 \le x. \end{cases}$$

There is no way to extend this solution for $t \ge 1$ in a classical way. Moreover, note that this is not related to the fact that g is only piecewise C^1 since the same phenomena will occur no matter how smooth you chose g on the transition region in between (the characteristics emanating from the boundary points will always intersect at $t = 1$). Note also that this kind of behavior is typical for nonlinear waves and can be observed, e.g., on a beach, when wave fronts will steepen until they break (Figure 1.5). It is also known as **blowup** of solutions, even though it is not the solution itself but its gradient which will become unbounded. Note that if two characteristics corresponding to values u_0 and u_1 intersect, the solution attains all values between u_0 and u_1 within the cone formed by these crossing characteristics (by the intermediate value theorem). Hence the gradient must get vertical as one approaches the intersection. $\diamond$

Example 1.5 (Traffic flow)**.** The scalar conservation law provides a simple model for automobile traffic: In this case let $u(t,x)$ describe the density of cars on a road (considering one lane in one direction with no possibility to overtake). Since the number of cars on the road can only change if the number of cars entering at one side differs from the number of cars exiting at the other side (assuming cars are neither created nor destroyed in between — no accidents, no exits), such a situation can be described using a scalar conservation law with the flux function given by $F = u\,v$, where $v(t,x)$ is the velocity distribution of the cars. In order to turn this into a PDE, one needs to make an assumption on how the velocity depends on the density. One simple assumption is that the velocity will be equal to the maximal velocity v_{max} (speed limit) when the density is low and tend to 0 when the density approaches the maximal density u_{max} (bumper to bumper traffic, when the density equals one over the average car length):

$$v(u) := v_{max}\big(1 - \frac{u}{u_{max}}\big).$$

In summary,

$$u_t + v_{max}\big(1 - \frac{2u}{u_{max}}\big)u_x = 0. \qquad \diamond$$

Note that one can easily extend this to higher dimensions. Again for any given (sufficiently smooth) domain U the change of the total amount within U equals the net flux through the boundary

$$\frac{d}{dt}\int_U u(t,x)d^n x = -\int_{\partial U} F(t,x)\cdot \nu dS(x), \tag{1.32}$$

where ν is the outward pointing unit normal vector of the surface ∂U and dS is the surface measure. Using the divergence theorem this turns into

$$\int_u (u_t(t,x) + \operatorname{div}_x F(t,x))\, d^n x = 0 \tag{1.33}$$

and since this must hold for any (sufficiently smooth) domain U, we arrive at the general form of a conservation law in $\mathbb{R}^n$

$$u_t + \operatorname{div}_x(F) = 0. \tag{1.34}$$

If the flux function is of the from $F(u) = v\,u$ with constant velocity $v(t,x) = a$, we again get the transport equation.

Problem 1.11. *Solve*

$$u_t + u\,u_x = 2t, \qquad u(0,x) = x.$$

Problem 1.12. *Solve*

$$u_t + \sqrt{u}u_x = 0, \qquad u(0,x) = x, \qquad x > 0.$$

Problem 1.13. *Consider a conservation law with $F \in C^2(\mathbb{R})$. Show that for bounded initial conditions $g \in C^1(\mathbb{R})$, the gradient of the solution remains bounded on bounded positive time intervals (as long as the solution exists) if F is convex and g is increasing. (Hint: Use implicit differentiation to find a formula for u_x.)*

Problem 1.14. *Consider a conservation law with $F \in C^2(\mathbb{R})$. Show that for bounded initial conditions $g \in C^1(\mathbb{R})$ with compact support* $\operatorname{supp}(g) \subseteq [-R,R]$ *we have that the solution u has compact support* $\operatorname{supp}(u(.,t)) \subseteq [-R-ct, R+ct]$ *for some constant c as long as it exists. Moreover, show that*

$$\int_{\mathbb{R}} u(t,x)dx = \int_{\mathbb{R}} g(x)dx.$$

Problem 1.15. *Solve the conservation law for* $(t, x) \in \mathbb{R}^{1+2}$ *with velocity field* $v(t, x) := x$.

Problem 1.16. *Derive a formula for the solution of the conservation law in* $\mathbb{R}^n$ *with a given velocity field* $v(x)$ *in terms of the solution of the characteristic equation. (Hint:* ***Jacobi's formula*** $\frac{d}{ds}\det(A(s)) = \operatorname{tr}(\dot{A}(s)A^{-1}(s))\det(A(s))$.*)*

1.4. Weak solutions

Recall the Burgers equation from Example 1.4, where we have seen (cf. Figure 1.5) that certain initial conditions can lead to the formation of a discontinuity in the solution. In the case of water waves such a situation can be observed at a beach when waves eventually break. Our present model is not suitable for describing such a behavior and thus there is no point in trying to extend the solution beyond this point. On the other hand, there are cases where u models the density of some substance and one observes that after the formation of such a discontinuity, this discontinuity continues to travel along a certain trajectory. This is commonly referred to as a shock wave and it raises the question whether it is possible to extend our notion of solution to cover such situations.

The idea is borrowed from distribution theory and to formulate it, we first assume that we have a classical solution $u \in C([0, \infty) \times \mathbb{R}, \mathbb{R}) \cap C^1((0, \infty) \times \mathbb{R}, \mathbb{R})$ of our scalar conservation law (1.20). Now we multiply this equation with a smooth **test function**

$$\varphi \in C_c^\infty([0, \infty) \times \mathbb{R}, \mathbb{R}) \tag{1.35}$$

with compact support and integrate over $[0, \infty) \times \mathbb{R}$ to obtain

$$\int_0^\infty \int_{\mathbb{R}} \big(u_t(t, x) + F(u(t, x))_x\big)\varphi(t, x)dx\, dt = 0. \tag{1.36}$$

So far nothing interesting has happened. However, note that if this equation holds for all test functions $\varphi \in C_c^\infty([0, \infty) \times \mathbb{R}, \mathbb{R})$, then u must solve our equation. Now here comes the trick, we use integration by parts and insert our initial condition to obtain

$$\int_0^\infty \int_{\mathbb{R}} \big(u(t, x)\varphi_t(t, x) + F(u(t, x))\varphi_x(t, x)\big)dx\, dt + \int_{\mathbb{R}} g(x)\varphi(0, x)dx = 0. \tag{1.37}$$

Again, as long as u is sufficiently smooth, we can undo the integration by parts to conclude that a sufficiently smooth function u satisfies (1.20) if and only if it satisfies (1.37) for all test functions $\varphi \in C_c^\infty([0, \infty) \times \mathbb{R}, \mathbb{R})$. However, since no smoothness is required for (1.37) to make sense, we can use it to generalize our notion of solution:

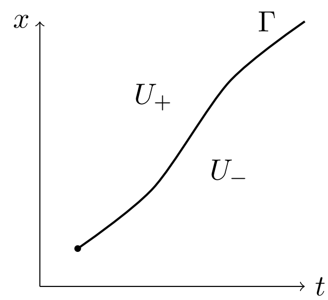


Figure 1.6. Notation for a weak solution with a discontinuity along a curve Γ.

A locally integrable function $u : [0,\infty) \times \mathbb{R} \to \mathbb{R}$ satisfying (1.37) for all test functions $\varphi \in C_c^\infty([0,\infty) \times \mathbb{R}, \mathbb{R})$ is called a **weak solution** of (1.20). The following example shows that this is indeed a useful definition:

Example 1.6. Consider the transport equation $F(u) := au$ for some $a \in \mathbb{R}$. Then for any locally integrable function $g : \mathbb{R} \to \mathbb{R}$ the function $u(t,x) := g(x - at)$ is a weak solution. Indeed using Fubini we have

$$\begin{aligned}
&\int_0^\infty \int_{\mathbb{R}} g(x-at)\big(\varphi_t(t,x) + a\varphi_x(t,x)\big) dx\, dt = \\
&\quad = \int_0^\infty \int_{\mathbb{R}} g(y)\big(\varphi_t(t,y+at) + a\varphi_x(t,y+at)\big) dy\, dt \\
&\quad = \int_{\mathbb{R}} g(y) \int_0^\infty \frac{\partial}{\partial t}\varphi(t,y+at) dt\, dy = -\int_{\mathbb{R}} g(y)\varphi(0,y) dy. \qquad \diamond
\end{aligned}$$

Next, let us try to apply this definition to our problem alluded to at the beginning. So we want to look at the case where a solution is smooth away from a smooth curve $\Gamma := \{(t,x) | x = \gamma(t), t \ge t_0\}$. More precisely, if we split $U := [t_0, \infty) \times \mathbb{R}$ according to (cf. Figure 1.6)

$$U_\pm := \{(x,t) | \pm (x - \gamma(t)) > 0, t \ge t_0\}, \tag{1.38}$$

then we will assume that u is smooth in the interior of $U_\pm$ with a continuous extension to its closure $\overline{U_\pm}$. The limits of u towards Γ from $U_\pm$ will be denoted by

$$u_\pm(t, \gamma(t)) = \lim_{\varepsilon \downarrow 0} u(t, \gamma(t) \pm \varepsilon). \tag{1.39}$$

Of course, if we avoid Γ and choose a test function φ with support away from Γ, we conclude that u satisfies (1.20) in $[0,\infty) \times \mathbb{R} \setminus \Gamma$.

If the support of φ includes Γ we have to work harder and split the domain of integration according to $V_- \cup V_+$, where $V_\pm$ contains $U_\pm$, respectively (to this end one could assume $t_0 = 0$ by extending C without loss of generality). Then using the integration by parts (Corollary A.14) one obtains

$$
\begin{aligned}
0 &= \int_0^\infty \int_{\mathbb{R}} \big(u\varphi_t + F(u)\varphi_x\big) dx\, dt + \int_{\mathbb{R}} g\varphi_0 dx \\
&= -\iint_{V_-} \big(u_t + F(u)_x\big)\varphi\, dx\, dt + \int_{\partial V_-} \big(u_-\nu_1 + F(u_-)\nu_2\big)\varphi\, dS \\
&\quad - \iint_{V_+} \big(u_t + F(u)_x\big)\varphi\, dx\, dt + \int_{\partial V_+} \big(u_+\nu_1 + F(u_+)\nu_2\big)\varphi\, dS \\
&\quad + \int_{\mathbb{R}} g\varphi_0 dx.
\end{aligned}
$$

Here ν is the unit normal and $\varphi_0(x) := \varphi(0,x)$. Now by our previous consideration the integrals over $V_\pm$ both vanish and the integrals over $\partial V_\pm$ along $t = 0$ cancel with the last integral. This leaves us with the integrals over $\partial V_\pm$ along Γ which, since these are oriented in opposite directions, leaves us with

$$
0 = \int_\Gamma \Big((u_+ - u_-)\nu_1 - (F(u_+) - F(u_-)\nu_2\Big)\varphi\, dS. \tag{1.40}
$$

As usual, since this must hold for arbitrary test functions, we arrive at the **Rankine–Hugoniot condition**[8]

$$
F(u_+) - F(u_-) = \dot\gamma\big(u_+ - u_-\big), \qquad (t,x) \in \Gamma. \tag{1.41}
$$

Such a discontinuity is known as a **shock** and the Rankine–Hugoniot condition says the jump in the flux equals the speed of the shock times the hight of the shock.

Example 1.7. Now we are able to continue the solution from Example 1.4 to $t \ge 1$. To this end choose $\gamma(t) := \frac{1+t}{2}$ and set

$$
u(t,x) := \begin{cases} 1, & x < \gamma(s), \\ 0, & \gamma(s) < x. \end{cases}
$$

We leave it as an exercise to verify the Rankine–Hugoniot condition. The solution is depicted in Figure 1.7. ⋄

Example 1.8. Let us again consider the Burgers equation, but now with the initial condition

$$
g(x) := \begin{cases} 0, & x < 0, \\ 1, & x > 0. \end{cases}
$$

[8]William Rankine (1820–1872), Scottish mechanical engineer

[8]Pierre Henri Hugoniot (1851–1887), French mathematician

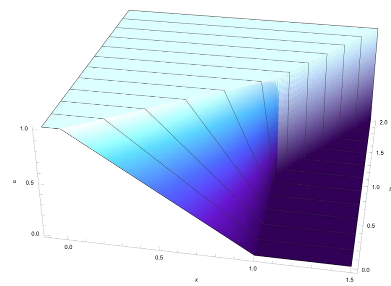


Figure 1.7. Shock wave for Burgers' equation

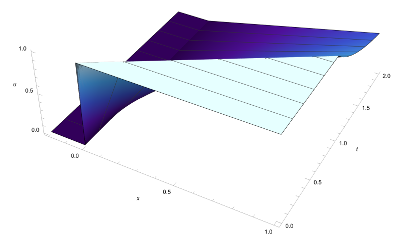


Figure 1.8. Rarefaction wave for Burgers' equation

In this case the characteristics do not cover the region $0 < x < t$ but we can choose $\gamma(t) := \frac{t}{2}$ and set

$$u(t, x) := \begin{cases} 0, & x < \gamma(s), \\ 1, & x > \gamma(s), \end{cases}$$

such that the Rankine–Hugoniot condition holds. However, note that here it is also possible to patch the solution $\frac{x}{t}$ into the missing region such that the result is continuous along the boundary of this region

$$v(t, x) := \begin{cases} 0, & x < 0, \\ \frac{x}{t}, & 0 < x < t, \\ 1, & x > t. \end{cases}$$

Naturally the Rankine–Hugoniot condition holds if the solution is continuous along the curve. So we get two weak solutions, but the latter is preferable in applications due to its continuity. It is known as a **rarefaction wave**. The solution is depicted in Figure 1.8.

Hence an additional condition is needed to ensure uniqueness of weak solutions for the Burgers equation. We will however not pursue this here. ⋄

Problem 1.17. *Show that the solution from Example 1.4 satisfies the Rankine–Hugoniot condition.*

Problem 1.18. *Find further weak solutions corresponding to the initial conditions from Example 1.8. (Hint: Split the shock into two jumps from (e.g.) 0 to $\frac{1}{2}$ to 1.)*

Problem 1.19. *Find a continuous weak solution of the traffic model from Example 1.5 corresponding to the initial conditions*

$$g(x) := \begin{cases} u_{max}, & x < 0, \\ 0, & x > 0, \end{cases}$$

modeling the situation when cars start at a traffic light turning green. (Hint: Look at a transformation $\alpha\, v(t, \beta t - x)$.)

1.5. Fully nonlinear equations

Finally, quasilinear equations are not the end of the story either. In fact, by adding the derivatives to the system of characteristic equations one can extend this method to arbitrary first order equations

$$F(x, u, \nabla u) = 0. \tag{1.42}$$

As before, the idea is to evaluate the solution u along some characteristic curves $x(s)$ (which are to be determined). Hence we define

$$z(s) := u(x(s)), \qquad p(s) := (\nabla u)(x(s)). \tag{1.43}$$

Differentiating z we obtain

$$\dot{z} = \sum_{k=1}^{n} u_{x_k} \dot{x}_k = \sum_{k=1}^{n} p_k \dot{x}_k. \tag{1.44}$$

Differentiating p we obtain

$$\dot{p}_j = \sum_{k=1}^{n} u_{x_j x_k} \dot{x}_k \tag{1.45}$$

and we need to get rid of the second derivatives if we want to get a closed system of differential equations for (x, z, p). To this end we assume $F \in C^2(\mathbb{R}^{2n+1}, \mathbb{R})$ and differentiate (1.42) with respect to x_j:

$$\sum_{k=1}^{n} F_{p_k}(\nabla u, u, x) u_{x_k x_j} + F_z(\nabla u, u, x) u_{x_j} + F_{x_j}(\nabla u, u, x) = 0. \tag{1.46}$$

Now if we assume (which is compatible with what we had in the quasi-linear case)

$$\dot{x}_j = F_{p_j}(p, z, x), \tag{1.47}$$

then we can use this equation to eliminate the second derivatives in the equation for $\dot{p}$ leading to

$$\dot{p}_j = -F_{x_j}(p, z, x) - F_z(p, z, x) p_j. \tag{1.48}$$

In summary, this leads us to the characteristic equations

$$\begin{aligned}\dot{x}_j &= F_{p_j}(p,z,x),\\ \dot{z} &= \sum_{k=1}^{n} p_k F_{p_k}(p,z,x),\\ \dot{p}_j &= -F_{x_j}(p,z,x) - F_z(p,z,x)p_j. \end{aligned} \tag{1.49}$$

Example 1.9. Consider the **Hamilton–Jacobi equation**[9]

$$u_t + H(\nabla_x u, x) = 0.$$

In this case the characteristic equations for x and $p := \nabla_x u$ are precisely the Hamilton equations from classical mechanics

$$\dot{x} = \nabla_p H(p,x), \qquad \dot{p} = -\nabla_x H(p,x).$$

Of course the equation for t is just $\dot{t} = 1$ (i.e., $t = s$). Setting $r(s) := u_t(s, x(s))$ the associated equation is $\dot{r} = 0$ implying that $r(t) = r(0) = -H(p(0), x(0))$. This corresponds to the well-known fact that the Hamiltonian H is conserved:

$$H(p(s), x(s)) = H(p(0), x(0)).$$

The equation for z is $\dot{z} = p \cdot \nabla_p H + r$.

Let me remark that the interesting fact about the Hamilton–Jacobi equation is not so much that you can solve it via the method of characteristics if you can solve the Hamilton equations, but the fact, known as **Jacobi's theorem**, that you can also go the other way: If you manage to find a solution, via some clever ansatz say, depending on n parameters, then you can obtain the solution of Hamilton's equations. ⋄

Example 1.10. The **eikonal equation**

$$|\nabla u| = n(x)$$

introduced by Hamilton provides the foundation of geometrical optics. Here $n(x)$ is the refraction index of the medium (the reciprocal of the wave speed). Taking squares we have $F(p,z,x) = \frac{1}{2}(|p|^2 - n(x)^2)$ and the characteristic equations are

$$\dot{x} = p, \qquad \dot{z} = |p|^2 = n(x)^2, \qquad \dot{p} = n(x)\nabla n(x).$$

In particular, we can eliminate p by considering

$$\ddot{x} = \frac{1}{2}\nabla n(x)^2.$$

Note that if the refraction index is constant, the characteristics will be straight lines. In fact, this way of solving the eikonal equation goes back

[9] William Rowan Hamilton (1805 –1865), Irish mathematician and astronomer

to Hamilton who also coined the name characteristics for the solutions of the above equation. ◇

Now what about initial conditions? Of course we will start at some point $x_0 \in \Gamma$, that is,

$$x(0) = x_0, \qquad z(0) = g(x_0). \tag{1.50}$$

But what about the derivatives? Again the derivatives within the tangent plane of Γ at x_0 can be computed from g, but the normal derivative is missing. This missing derivative must be computed from the differential equation. Once one has a triple of admissible initial conditions, the non-characteristic condition

$$\nabla_p F(p_0, z_0, x_0) \cdot \nu(x_0) \neq 0 \tag{1.51}$$

will ensure that we can obtain admissible initial conditions in a neighborhood of (p_0, z_0, x_0) via the implicit function theorem. Moreover, as in the previous section, this condition will also guarantee, that we can solve the resulting nonlinear equation to obtain u. Finally, one can show that the function u constructed in this way will indeed be a solution. We refer to [**10**, Section 3.2] for details.

Example 1.11. Let us solve the Hamilton–Jacobi equation for the harmonic oscillator with

$$H(p,x) = \frac{p^2 + x^2}{2}, \qquad g(x_0) = p_0 x_0.$$

We set $p = u_x$ and $r = u_t$. Then $F(r,p,z,t,x) = r + H(p,x)$ and the characteristic equations read

$$\dot{r} = 0, \quad \dot{p} = -x, \quad \dot{z} = p^2 + r, \quad \dot{t} = 1, \quad \dot{x} = p$$

and the initial conditions are

$$r(0) = -\frac{g'(y)^2 + y^2}{2}, \quad p(0) = g'(y), \quad z(0) = g(y), \quad t(0) = 0, \quad x(0) = y.$$

Solving for x, p we obtain

$$x(s) = y\cos(s) + g'(y)\sin(s), \qquad p(s) = g'(y)\cos(s) - y\sin(s)$$

and of course we have $t = s$ as well as $r(s) = r(0)$. Plugging this into $\dot{z}$ gives

$$z(s) = g(y) + \frac{g'(y)^2 - y^2}{4}\sin(2s) - yg'(y)\sin(s)^2.$$

Solving $x = y\cos(s) + p_0\sin(s)$ for y gives

$$u(t,x) = z(s) = \frac{p_0 x}{\cos(t)} - \frac{p_0^2 + x^2}{2}\tan(t).$$ ◇

Note that the method of characteristics breaks down if we allow u to be vector-valued, that is, as soon as we look at first order systems. However, this would be crucial for higher order equations since arbitrary partial differential equations can always be reduced to first order systems by adding the higher derivatives to the dependent variables.

Problem 1.20. *Let* $\Gamma = \{x \in \mathbb{R}^2 | x_1 = 0\}$ *and solve*

$$u_{x_1} u_{x_2} = 4u, \qquad u(0, x_2) = x_2^2.$$

Problem 1.21. *Solve the Hamilton–Jacobi equation with*

$$H(p, x) = \frac{p^2 + x^2}{2}, \qquad g(x_0) = \frac{x_0^2}{2}.$$

Problem 1.22. *The Hamilton–Jacobi equation for the Kepler problem in polar coordinates reads*

$$u_t + \frac{1}{2}\big(u_r^2 + \frac{1}{r^2} u_\theta^2\big) = \frac{k^2}{r}.$$

Find solutions which are linear in t *and* θ*.*

1.6. Classification and canonical forms

Since we have been quite successful with first order equations, it seems natural to make the next step and increase the order by one. For simplicity we will only look at the case of a second order quasi-linear equation

$$A_{11}(x, y) u_{xx} + 2A_{12}(x, y) u_{xy} + A_{22}(x, y) u_{yy} + b(x, y, u_x, u_y) = 0 \qquad (1.52)$$

in two variables. Given an initial surface Γ, we can try to prescribe initial conditions. Since our equation is second order, we expect that we need to prescribe both the values of u and its first derivatives on Γ. Of course derivatives in a tangential direction can be computed directly from the initial values and hence only the derivative in the normal direction needs to be prescribed. That is, if ν is the unit normal of Γ, then the initial conditions read

$$u\big|_\Gamma = g, \qquad \frac{\partial u}{\partial \nu}\big|_\Gamma = h, \qquad (1.53)$$

where

$$\frac{\partial u}{\partial \nu} := \nu \cdot \nabla u \qquad (1.54)$$

denotes the normal derivative. In the first order case the non-characteristic condition ensured that we can compute the missing normal derivative from the differential equation. Hence we expect that there should be a similar condition which ensures that we can compute all second order derivatives from the differential equation. To investigate this it will be convenient to straighten out the boundary. To this end recall (cf. Appendix A.2) that (after possibly permutating x and y) we can assume that $\Gamma = \{(x,y)|y = \gamma(x)\}$. Choosing new coordinates $\tilde{x} = x$, $\tilde{y} = y - \gamma(x)$, such that in this new coordinates we have $\tilde{\Gamma} = \{(\tilde{x},\tilde{y})|\tilde{y} = 0\}$, our equation will still be of the same form with coefficients

$$\tilde{A}_{11} = A_{11}, \quad \tilde{A}_{12} = A_{12} - \gamma' A_{11}, \quad \tilde{A}_{22} = A_{22} - 2\gamma' A_{12} + (\gamma')^2 A_{11}. \tag{1.55}$$

Hence we can assume that $\Gamma = \{(x,y)|y = 0\}$ without loss of generality in which case our initial conditions read

$$u(x,0) = g(x), \qquad u_y(x,0) = h(x). \tag{1.56}$$

Of course $u_x(x,0) = g'(x)$ and for the second derivatives we get $u_{xx}(x,0) = g''(x)$, $u_{xy}(x,0) = h'(x)$ as well as

$$\begin{aligned} A_{11}(x,0)g''(x) + 2A_{12}(x,0)h'(x) + A_{22}(x,0)u_{yy}(x,0) \\ + b(x,0,g'(x),h(x)) = 0. \end{aligned} \tag{1.57}$$

Hence to be able to solve for the missing derivative $u_{yy}(x,0)$ we clearly need $A_{22}(x,0) \neq 0$. In fact, we could continue in this way and compute arbitrary high derivatives of u by differentiating the differential equation (assuming that the coefficients as well as the initial conditions have sufficiently many derivatives). Surprisingly, the condition $A_{22}(x,0) \neq 0$ will also guarantee solvability for the higher derivatives (try to compute the third order derivatives and you will see why). In fact, this shows that we could try to obtain the solution by computing its Taylor series at a given point on Γ. This is the strategy of the Cauchy–Kovalevskaya theorem to be discussed in the next section.

These considerations suggest that the condition $A_{22}(x,0) \neq 0$ takes the role of the non-characteristic condition for (1.52). Translating back to a general surface Γ, we hence define the **non-characteristic condition** as

$$\nu \cdot A\nu = A_{11}\nu_1^2 + 2A_{12}\nu_1\nu_2 + A_{22}\nu_2^2 \neq 0. \tag{1.58}$$

As in the first order case this condition depends on the point on Γ. Nevertheless, if it is satisfied at one point, it will be satisfied in a neighborhood of this point by continuity. Conversely, one calls Γ characteristic if this condition is violated on all of Γ. While in the case of a first order equation there is always precisely one characteristic direction (and a corresponding characteristic line emanating from a given point) unless the vector field a vanishes at this point, now there are three cases depending on the signature of the quadratic form associated with A:

If A is positive (or negative) definite (that is, both eigenvalues are positive (or negative)), then there are no characteristic directions and we call the equation **elliptic** in this case. If A has one positive and one negative eigenvalue, there are two characteristic directions and the equation is called

hyperbolic in this case. In the degenerate case, where one eigenvalue of A vanishes, the equation is called **parabolic**. Of course there is also the case when both eigenvalues vanish (and hence A is zero). But if this happens on an open set, then (1.52) is in fact a first order equation and hence this case has no practical relevance. Indeed, this classification is only of interest if (1.52) is of a given type in an open set.

So while these considerations suggest that we should expect trouble if we specify initial conditions on a characteristic line, they do not really help us with our task of finding a solution. But this is unfortuantely the gist of the matter, there is no easy way of solving second order equations analogous to the method of characteristics for first order equations. At this point you might object that we can simply transform our second order equation to a first order system by adding the first order derivatives to the dependent variables, just like one does in the case of ordinary differential equations. And indeed, we can of course do this, but as the method of characteristics does *not* extend to systems, there is no point unless we know how to handle first order systems.

Another approach is to use a change of variables to reduce (1.52) to a simpler form. For simplicity we first look at the case of a linear equation with constant coefficients,

$$A_{11}u_{xx} + 2A_{12}u_{xy} + A_{22}u_{yy} + b_1u_x + b_2u_y + cu = 0, \tag{1.59}$$

or equivalently

$$\big(\nabla \cdot A\nabla + b \cdot \nabla + c\big)u = 0. \tag{1.60}$$

First of all, a rotation will diagonalize A and we can assume that the first eigenvalue equals one (if both were zero, the equation would be of first order). Moreover, setting $u = \mathrm{e}^{-b_1x/2}v$ we can even eliminate the u_x term. Scaling y we can assume that the second eigenvalue is ± 1 and eliminate u_y as before, or, if the second eigenvalue vanishes, assume $b_2 = -1$. This eventually transforms our equation into one of the following standard forms (assuming our equation contains derivatives with respect to both variables; otherwise it would be an ordinary differential equation):

- Elliptic case (two eigenvalues of equal sign, $\det(A) > 0$):

$$u_{xx} + u_{yy} + cu = 0.$$

- Hyperbolic case (two eigenvalues of opposite sign, $\det(A) < 0$):

$$u_{xx} - u_{yy} + cu = 0.$$

- Parabolic case (one vanishing eigenvalue, $\det(A) = 0$):

$$u_{xx} - u_y = 0.$$

In the elliptic case this equation is the **Helmholtz equation**[10] and reduces to the **Laplace equation**[11] for $c = 0$. In the hyperbolic case this is the **Klein–Gordon equation**[12] and reduces to the **wave equation** for $c = 0$. In the parabolic case this is the **heat equation**.

For $c = 0$ these equations are the three most ubiquitous partial differential equations (also known has the *Holy Trinity*) and understanding them will be one of our main tasks during this course. They were introduced in the 18th and early 19th century. The one-dimensional wave equation was introduced as a model for vibrating strings and analyzed by d'Alembert[13] in 1752. It was extended to higher dimensions as a model of acoustic waves by Euler[14] in 1759 and Bernoulli[15] in 1762. The Laplace equation was first studied by Laplace in connection with gravitational fields around 1780 and the heat equation was introduced by Fourier[16] during his investigation of heat conduction between 1810 and 1822.

It turns out that there is not too much we can do at this point, except for the hyperbolic case, where a different canonical form is more beneficial: If one changes coordinates such that the characteristic lines coincide with the coordinate directions, which implies that both A_{11} and A_{22} will vanish.

Example 1.12. For the wave equation $c^{-2}u_{tt} = u_{xx}$ the characteristic directions are

$$\nu_1^2 - c^2\nu_2^2 = (\nu_1 - c\nu_2)(\nu_1 + c\nu_2) = 0.$$

Hence we choose new coordinates $\xi = x + ct$, $\eta = x - ct$ such that $v(\xi, \eta) := u(t, x)$ satisfies

$$v_{\xi\eta} = 0.$$

From this form it is easy to see that the general solution is $v(\xi, \eta) = F(\xi) + G(\eta)$ which implies that the general solution of the wave equation is $u(t, x) = F(x + ct) + G(x - ct)$. If $u(t, x)$ should satisfy the initial conditions

$$u(0, x) = g(x) \in C^2(\mathbb{R}), \qquad u_t(0, x) = h(x) \in C^1(\mathbb{R}),$$

we must have $F(x)+G(x) = g(x)$ and $c(F'(x)-G'(x)) = h(x)$. So if $H(x) = \int h(x)dx$ we have $F(x) = \frac{1}{2}(g(x) + c^{-1}H(x))$, $G(x) = \frac{1}{2}(g(x) - c^{-1}H(x))$. Note that while H is only defined up to a constant, this constant will cancel

[10] Hermann von Helmholtz (1821–1894), German physicist

[11] Pierre-Simon Laplace (1749–1827), French mathematician

[12] Oskar Klein (1894–1977), Swedish theoretical physicist

[12] Walter Gordon (1893–1939), German theoretical physicist

[13] Jean le Rond d'Alembert (1717–1783), French mathematician

[14] Leonhard Euler (1707–1783), Swiss mathematician, physicist, astronomer, geographer, logician and engineer

[15] Daniel Bernoulli (1700–1782), Swiss mathematician and physicist

[16] Joseph Fourier (1768–1830), French mathematician and physicist

in $F+G$. Hence we obtain **d'Alembert's formula**

$$u(t,x) = \frac{g(x+ct)+g(x-ct)}{2} + \frac{1}{2c}\int_{x-ct}^{x+ct} h(y)dy$$

also found in Problem 1.3. ◇

While this last canonical form also works for the Klein–Gordon equation, it does not lead to a solution formula in this case (which exists, but is much harder to find — see Problem 4.22). Using new coordinates based on the characteristic directions also works for equations with non-constant coefficients. However, unless this eventually happens to lead to the wave equation, it will be of little practical help.

Example 1.13. For the **Tricomi equation**[17]

$$u_{xx} + x\,u_{yy} = 0$$

the characteristic directions are

$$\nu_1^2 + x\nu_2^2 = 0.$$

Hence it is hyperbolic for $x<0$, parabolic for $x=0$ (though this set is not open), and elliptic for $x>0$. To get the characteristic lines in the hyperbolic case, we look for a curve $(x(s),y(s))$ such that the normal vector $\nu(s) = (\dot{y}(s), -\dot{x}(s))$ satisfies the characteristic condition

$$\dot{y}(s)^2 + x(s)\dot{x}(s)^2 = 0.$$

Eliminating s and writing the curve as $y(x)$ we get

$$y'(x) = \pm\sqrt{-x}$$

and hence the two characteristic curves are

$$\xi = y + \frac{2}{3}(-x)^{3/2}, \qquad \eta = y - \frac{2}{3}(-x)^{3/2}.$$

Using (ξ,η) as new variables and setting $v(\xi,\eta) := u(t,x)$ we get

$$v_{\xi\eta} - \frac{v_\xi - v_\eta}{6(\xi-\eta)} = 0.$$

In the elliptic case $x>0$ the two characteristic curves are complex. However, we can still get a canonical form if we take real and imaginary part as new coordinates:

$$\xi = y, \qquad \eta = \frac{2}{3}x^{3/2}.$$

Setting again $v(\xi,\eta) := u(t,x)$ we get

$$v_{\xi\xi} + v_{\eta\eta} + \frac{v_\eta}{3\eta} = 0.$$ ◇

[17] Francesco Tricomi (1897–1978), Italian mathematician

Note that this classification is invariant under a change of coordinates (Problem 1.26). A more in-depth discussion of normal forms for second order equations can be found in [**24**, Chapter 3].

The next obvious step would be to generalize these considerations to more than two variables. First of all, the non-characteristic condition can be derived by a straightforward generalization of the present arguments. Of course the classification gets more cumbersome since there are much more possibilities for the signature of the quadratic form of A. Here we will only discuss the special cases which are most relevant (and where the definitions are universally accepted): The equation is called elliptic, if it has no characteristic directions, that is, all eigenvalues of A have the same sign. It is called hyperbolic if all except for one eigenvalue have the same sign and the remaining one is of opposite sign. It is parabolic if one eigenvalue vanishes while the others are of the same sign.

However, from a practical perspective, this way of classifying partial differential equation can also be viewed as generalizing the three prototypical equations, representing more complicated models which behave similarly and can be tackled using similar techniques. Specifically, let us look at the most general elliptic linear differential operator

$$Lu(x) = -\sum_{j,k=1}^{n} A_{jk}(x)\partial_{x_j x_k} u(x) + \sum_{j=1}^{n} b_j(x)\partial_{x_j} u(x) + c(x)u(x), \tag{1.61}$$

where A is a positive definite matrix. Then

$$Lu = 0 \tag{1.62}$$

is **elliptic**,

$$u_t = -Lu \tag{1.63}$$

is **parabolic**, and

$$u_{tt} = -Lu \tag{1.64}$$

is **hyperbolic**. In the latter two cases we have denoted the last independent variable by t since it corresponds to time in most applications (and in this case the coefficients A, b, and c are allowed to depend on t as well). It is folk wisdom that these equations behave similar to their prototypical representative, where $L = -\Delta$, the Laplace, heat, and wave equation.

Problem 1.23. *Show that the Cauchy problem*

$$u_{xx} + y^2 u_{yy} = 2u, \qquad u(x,0) = 0,\ u_y(x,0) = 0.$$

fails the non-characteristic condition. Find a nontrivial solution.

Problem 1.24. *Find the solution of the problem*

$$u_{xx} - 4u_{xy} + 3u_{yy} + 4u_x - 12u = 0, \quad u(x,0) = g(x),\ u_y(x,0) = h(x).$$

Problem 1.25. *Transform*

$$xu_{xx} - 4x^3 u_{yy} - u_x = 0$$

to a canonical form by using the characteristic directions as new coordinates. Can you find the general solution?

Problem* 1.26. *Consider a diffeomorphism $\xi(x,y)$, $\eta(x,y)$. Show that the coefficient matrix A of* (1.52) *in the new coordinates is given by*

$$\tilde{A} = \frac{\partial(\xi,\eta)}{\partial(x,y)} A \frac{\partial(\xi,\eta)}{\partial(x,y)}^T, \qquad \frac{\partial(\xi,\eta)}{\partial(x,y)} = \begin{pmatrix} \xi_x & \xi_y \\ \eta_x & \eta_y \end{pmatrix}.$$

Conclude that the classification is invariant under a change of coordinates.

Chapter 2

The Cauchy–Kovalevskaya theorem

2.1. First order systems

The main limitation of the method of characteristics is the fact that it does not extend to systems. The Cauchy–Kovalevskaya theorem[1] removes this limitation and hence can be considered as an analog of the Picard–Lindelöf theorem from the theory of ordinary differential equations. The price one has to pay is that it only works in the realm of real analytic functions. Lewy's example will show that it breaks down once we leave this realm. However, there are further reasons why the Cauchy–Kovalevskaya theorem does not play the same role for partial differential equations like the Picard–Lindelöf theorem does for ordinary ones. Namely, it does not distinguish between positive and negative times and hence cannot handle the (physically relevant) Cauchy problem for the heat equation. Moreover, as it only settles the Cauchy problem, it does not help with boundary value problems which play a far more important role for partial differential equations than for ordinary ones. In particular, the results in this chapter will not be used in the rest of this book.

We recall that a function of several variables is called **real analytic** if it has a convergent Taylor[2] expansion in a neighborhood of every point

[1]Augustin-Louis Cauchy (1789–1857), French mathematician
[1]Sofya Kovalevskaya (1850–1891), Russian mathematician
[2]Brook Taylor (1685–1731), English mathematician

of its domain (cf. Appendix A.1). If the coefficients are real analytic, this opens another approach, which also works for systems. Then the partial differential equation can be used (by recursive differentiation of the equation) to determine the derivatives (and hence the Taylor series) of the solution. Of course this will only work in a small neighborhood and the corresponding result will be of a local nature.

To be more specific, let us look at the Cauchy problem for a quasilinear first order system

$$\sum_{k=1}^{n} A_k(x,u)u_{x_k} + b(x,u) = 0, \quad u(y) = g(y),\ y \in \Gamma. \tag{2.1}$$

Here $u(x) = (u^1(x), \dots, u^m(x))$ is now vector-valued and $A_k(x,u)$ are $m\times m$ matrices while $b(x,u)$ is a vector. The coefficients A_k and b (i.e. all entries) as well as Γ are assumed real analytic. We will first make a few simplifications which can be done without loss of generality (at least locally and after a possible permutation of the coordinates).

The assumption that Γ is real analytic implies that, once we restrict x to some sufficiently small open set $U \subseteq \mathbb{R}^n$, it is given in the form $\Gamma = \{x \in U | x_n = \gamma(x_1,\dots,x_{n-1})\}$, where γ is real analytic. Moreover, we can straighten out Γ by making a change of variables $y := (x_1,\dots,x_{n-1}, x_n - \gamma(x_1,\dots,x_{n-1}))$ such that in this new variables Γ will be the hyperplane $y_n = 0$. The resulting equation will still be quasilinear with coefficients given by

$$\tilde{A}_k = \begin{cases} A_k & 1 \le k < n, \\ A_n - \sum_{l=1}^{m-1} A_l \partial_l \gamma, & k = n, \end{cases} \qquad \tilde{b} = b. \tag{2.2}$$

Hence we can assume $\Gamma = \{x \in U | x_n = 0\}$ without loss of generality and we will write $x = (\bar{x}, x_n) \in \mathbb{R}^n$, where $\bar{x} = (x_1,\dots,x_{n-1})$.

For the second step we recall, that we want to solve our system by computing the Taylor series of $u(x)$ at some given point $y_0 \in \Gamma$. Without loss of generality we can assume $y_0 = 0$. Now $u(0) = g(0)$ follows from the initial condition and so will all first order derivatives $u_{x_j}(0) = g_{x_j}(0)$ for all $1 \le j \le n-1$. To obtain the remaining derivatives $u_{x_n}(0)$ we need to use our differential equation. In particular, if we evaluate the system at $x = 0$ and insert what we already know, we must be able to solve for the missing derivatives. This leads to the **non-characteristic condition**

$$\det\big(A_n(0, g(0))\big) \neq 0. \tag{2.3}$$

Observe that in the case $m = 1$ this of course agrees with our previous definition for the scalar case (1.25), where $\det(A_n) = A_n = a_n \neq 0$. Moreover, if the non-characteristic condition holds at a point $(0, g(0))$, continuity implies that we can reduce U such that it holds on all of U. In particular, we can multiply our system by the inverse of A_n reducing it to the form

$$u_{x_n} = \sum_{k=1}^{n-1} A_k(x,u) u_{x_k} + b(x,u), \quad u(\bar{x},0) = g(\bar{x}). \tag{2.4}$$

Here we have already solved the system for the derivatives with respect to x_n.

Theorem 2.1 (Cauchy–Kovalevskaya). *Let the coefficients A_k, b as well as the surface Γ and the initial condition g be real analytic in a neighborhood of a point $x_0 \in \Gamma$. Suppose the* ***non-characteristic condition***

$$\det\left(\sum_{k=1}^{n} \nu_k(x_0) A_k(x_0, g(x_0))\right) \neq 0, \tag{2.5}$$

where $\nu(x_0)$ is the normal vector of Γ at x_0, holds. Then the initial value problem (2.1) *has a unique real analytic solution in a neighborhood of x_0.*

Proof. We have already seen that we can straighten the boundary and assume our system to be of the form (2.4) without loss of generality. Furthermore, by adding another dependent variable $u^{m+1}_{x_n} = 1$, $u^{m+1}(\bar{x},0) = 0$ to the system we can assume that the coefficients A_k and b do not depend on x_n. Finally, replacing u by $u - g$ we can also assume $g = 0$ without loss of generality.

Now we first observe that all derivatives $(\partial^\alpha u)(0)$ can be computed recursively using the differential equation. To see this we use induction on the number r of derivatives with respect to x_n. If $r = \alpha_n = 0$, then $(\partial^\alpha u)(0) = 0$ since g vanishes. If $\alpha_n = r+1$ we use $\partial^\alpha u = \partial^\beta(\partial_{x_n} u)$ with $\beta_n = r$ and insert the differential equation for $\partial_{x_n} u$. This shows that we can express $(\partial^\alpha u)(0)$ in terms of derivatives which contain at most r derivatives with respect to x_n and establishes the induction step.

So we can write down a Taylor series for u and we need to show that this series converges. To this end we will use majorants $\tilde{A}_k$ and $\tilde{b}$ for A_k and b, respectively. Then, if U solves

$$U_{x_n} = \sum_{k=1}^{n-1} \tilde{A}_k(\bar{x}, U) U_{x_k} + \tilde{b}(\bar{x}, U), \quad U(\bar{x},0) = 0$$

the component U^i will be a majorant for u^i. Indeed, to see this observe that the derivatives of U can be computed from its differential equation following the same procedure as for u. Since in every step the expression for a derivative of U majorizes the corresponding expression for u, the claim follows.

Moreover, since the series for A_k^{ij} and b^i converge on a common rectangle $\mathcal{R}((r,\dots,r))$ we can choose a common majorant

$$\tilde{A}_k^{ij}(\bar{x},u) = \tilde{b}^i(\bar{x},u) := \frac{Mr}{r-(x_1+\cdots+x_{n-1}+u_1+\cdots+u_m)}.$$

Then the corresponding differential equation is

$$U^i_{x_n} = \frac{Mr}{r-(x_1+\cdots+x_{n-1}+U^1+\cdots+U^m)}\left(\sum_{j=1}^{m}\sum_{k=1}^{n-1}U^j_{x_k}+1\right)$$

and we will get a solution $U^i(x) = V(x_1+\cdots+x_n, x_n)$ provided $V(y,t)$ solves

$$V_t = \frac{Mr}{r-y-mV}(1+m(n-1)V_y), \qquad V(y,0)=0.$$

The solution of this scalar equation can be found by the method of characteristics to be

$$V(y,t) = \frac{1}{mn}\big(r-y-\sqrt{(r-y)^2-2mnMrt}\big).$$

Hence we have found a real analytic majorant and thus the Taylor series for u converges in a neighborhood of 0.

It remains to show that u is a solution of our original problem. First of all note that by construction all $\bar{x}$ derivatives are 0 and hence u satisfies the initial condition. Moreover, if we insert u into the differential equation, we get a real analytic function of x in a neighborhood of the origin. Moreover, by construction all derivatives vanish at the origin and thus this real analytic function vanishes identically, that is, u solves the differential equation. □

A surface Γ is called a **characteristic surface** if the non-characteristic condition fails on every point of Γ.

Of course, if we prescribe initial conditions at each point of Γ, we can patch the local solutions by virtue of the unique continuation principle to obtain a solution in a neighborhood of Γ. The main problem with this result is that most applications will require existence of a solution on a given domain and not just in a neighborhood of its boundary.

Let me also remark that uniqueness applies only within the class of real analytic solutions. At least for linear equations there is a stronger uniqueness theorem due to Holmgren[3] which establishes uniqueness within the class of C^1 solutions [**15**, Section 3.5], [**27**, Section 2.3].

Example 2.1. Consider the **Cauchy–Riemann equations**[4]

$$v_x = -w_y, \quad w_x = v_y.$$

[3]Erik Albert Holmgren (1872–1943), Swedish mathematician

[4]Bernhard Riemann (1826–1866), German mathematician

In our notation the system reads

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} v_x \\ w_x \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} v_y \\ w_y \end{pmatrix} = 0$$

and the non-characteristic condition

$$\det \begin{pmatrix} \nu_1 & \nu_2 \\ -\nu_2 & \nu_1 \end{pmatrix} = |\nu|^2 \neq 0$$

is always satisfied.

If we choose for example $\Gamma = \{y = 0\}$ and $v(x,0) = g(x)$, $w(x,0) = h(x)$, then using $v_{yy} = -v_{xx}$ and $w_{yy} = -w_{xx}$ one concludes

$$\partial_x^m \partial_y^n v(0,0) = \begin{cases} (-1)^k g^{(m+n)}(0), & n = 2k, \\ (-1)^k h^{(m+n)}(0), & n = 2k+1, \end{cases}$$

and

$$\partial_x^m \partial_y^n w(0,0) = \begin{cases} (-1)^k h^{(m+n)}(0), & n = 2k, \\ -(-1)^k g^{(m+n)}(0), & n = 2k+1. \end{cases}$$

Consequently

$$\begin{aligned} v(x,y) &= \sum_{m=0}^{\infty} \sum_{k=0}^{\infty} \frac{(-1)^k}{m!(2k)!} \left(g^{(m+2k)}(0) + h^{(m+2k+1)}(0) \frac{y}{2k+1} \right) x^m y^{2k} \\ &= \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} \left(g^{(2k)}(x) + h^{(2k+1)}(x) \frac{y}{2k+1} \right) y^{2k}. \end{aligned}$$

Extending the initial conditions to the complex plane this can be written more compactly as

$$\begin{aligned} v(x,y) &= \frac{g(x+\mathrm{i}y) + g(x-\mathrm{i}y)}{2} + \frac{h(x+\mathrm{i}y) - h(x-\mathrm{i}y)}{2\mathrm{i}} \\ &= \mathrm{Re}\big(g(x+\mathrm{i}y) - \mathrm{i}h(x+\mathrm{i}y)\big). \end{aligned}$$

Similarly we obtain

$$\begin{aligned} w(x,y) &= \sum_{m=0}^{\infty} \sum_{k=0}^{\infty} \frac{(-1)^k}{m!(2k)!} \left(h^{(m+2k)}(0) - g^{(m+2k+1)}(0) \frac{y}{2k+1} \right) x^m y^{2k} \\ &= \mathrm{Re}\big(h(x+\mathrm{i}y) + \mathrm{i}g(x+\mathrm{i}y)\big). \end{aligned}$$

Note that the Cauchy–Riemann equations can be thought of as a transport equation $u_x + \mathrm{i}u_y = 0$ with a complex speed if we set $u := v + \mathrm{i}w$. In this context one also writes $\frac{\partial}{\partial \bar{z}} u = 0$, where $z = x + \mathrm{i}y$ and $\frac{\partial}{\partial \bar{z}} = \frac{\partial}{\partial x} + \mathrm{i}\frac{\partial}{\partial y}$. ◇

Example 2.2. As another example let us try to solve the system

$$u_x = v, \qquad u_t = v_x.$$

Observe that inserting the first equation into the second shows that the first component u will solve the **heat equation**

$$u_t = u_{xx}.$$

In our notation the system reads

$$\begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} u_t \\ v_t \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} u_x \\ v_x \end{pmatrix} - \begin{pmatrix} v \\ 0 \end{pmatrix} = 0$$

and the non-characteristic condition is

$$\det \begin{pmatrix} \nu_2 & 0 \\ -\nu_1 & \nu_2 \end{pmatrix} = \nu_2^2 \neq 0.$$

Hence $\Gamma = \{x = 0\}$ is non-characteristic and we can choose corresponding initial conditions $u(t,0) = g(t)$, $v(t,0) = h(t)$. Now observe that

$$\partial_x^n u = \partial_x^{n-1} v = \partial_x^{n-2} u_t = \partial_t \partial_x^{n-2} u.$$

Consequently

$$\partial_x^n u(t,0) = \begin{cases} g^{(m)}(t), & n = 2m, \\ h^{(m)}(t), & n = 2m+1, \end{cases}$$

implying

$$\begin{aligned} u(t,x) &= \sum_{m=0}^{\infty} \frac{g^{(m)}(t)}{(2m)!} x^{2m} + \sum_{m=0}^{\infty} \frac{h^{(m)}(t)}{(2m+1)!} x^{2m+1} \\ &= \sum_{m,n=0}^{\infty} \frac{g^{(m+n)}(0)}{(2m)!n!} x^{2m} t^n + \sum_{m,n=0}^{\infty} \frac{h^{(m+n)}(0)}{(2m+1)!n!} x^{2m+1} t^n. \end{aligned}$$

While this provides some nontrivial solutions of the heat equation, it is of limited interest since the typical setting in applications is to specify an initial condition $u(0,x) = f(x)$ at $t = 0$. It is however interesting to note that one does not need real analyticity of g, h for the above series for u (first form) to converge (Problem 2.3).

Now if we choose the standard setting $\Gamma = \{t = 0\}$, then this surface is everywhere characteristic and the Cauchy–Kovalevskaya theorem does not apply. This is not surprising since the coefficient matrix $A_2 = \left(\begin{smallmatrix} 0 & 0 \\ -1 & 0 \end{smallmatrix}\right)$ is not invertible and hence we cannot solve the system for the t derivatives. So while at first it seems that we only have $u_t = v_x$, we can get the missing derivative by differentiating the first equation with respect to t, producing $v_t = u_{xt} = v_{xx}$. This shows that given $u(0,x) = f(x)$ (note that $v(0,x) = f'(x)$) we can compute all partial derivatives at (e.g.) $x_0 = 0$ even though $\Gamma = \{t = 0\}$ is characteristic. However, it is important to emphasize that this alone does not imply that there is a real analytic solution. Indeed, proceeding as pointed out before, one obtains the formal solution

$$u(t,x) = \sum_{m,n=0}^{\infty} \frac{f^{(m+2n)}(0)}{m!n!} x^m t^n.$$

If this series converges, it will solve the heat equation. For example, choosing $f(x) = x^k$ gives the **heat polynomials**

$$P_k(t,x) := \sum_{n=0}^{\lfloor k/2 \rfloor} \frac{k!}{(k-2n)!n!} x^{k-2n} t^n.$$

To see that this series does not always converge take for example (due to Kovalevskaya)

$$f(x) = \frac{1}{1+x^2}.$$

Then one has $f^{(2n)}(0) = (-1)^n (2n)!$ and $f^{(2n+1)}(0) = 0$ and hence

$$u(t,0) = \sum_{n=0}^{\infty} \frac{(2n)!}{n!} (-t)^n,$$

whose radius of convergence is zero. In particular, there is no real analytic solution in a neighborhood of the origin. We will see later that solutions of the heat equation typically only exist for positive times. ◇

Example 2.3. Consider the system

$$v_x = w_t, \quad w_x = v_t.$$

Note that differentiating the first equation with respect to x and the second equation with respect to t shows that v satisfies the **wave equation** $v_{tt} = v_{xx}$. Similarly one obtains $w_{tt} = w_{xx}$.

In our notation the system reads

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} v_x \\ w_x \end{pmatrix} - \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} v_t \\ w_t \end{pmatrix} = 0$$

and the non-characteristic condition reads

$$\det \begin{pmatrix} \nu_1 & -\nu_2 \\ -\nu_2 & \nu_1 \end{pmatrix} = (\nu_1 + \nu_2)(\nu_1 - \nu_2) \neq 0.$$

In particular, there are two characteristic surfaces $\Gamma = \{x = t\}$ and $\Gamma = \{x = -t\}$.

If we choose for example $\Gamma = \{t = 0\}$ and $v(0, x) = g(x)$, $w(0, x) = h(x)$ one obtains as before

$$\begin{aligned} v(t, x) &= \sum_{m=0}^{\infty}\sum_{k=0}^{\infty} \frac{1}{m!(2k)!}\left(g^{(m+2k)}(0) + h^{(m+2k+1)}(0)\frac{t}{2k+1}\right) x^m t^{2k} \\ &= \sum_{k=0}^{\infty} \frac{g^{(2k)}(x)}{(2k)!} t^{2k} + \sum_{k=0}^{\infty} \frac{h^{(2k+1)}(x)}{(2k+1)!} t^{2k+1} \\ &= \frac{g(x+t) + g(x-t)}{2} + \frac{h(x+t) - h(x-t)}{2}. \end{aligned}$$

Note that $v(t, x)$ satisfies the initial conditions $v(0, x) = g(x)$ and $v_t(0, x) = w_x(0, x) = h'(x)$. Similarly

$$w(t, x) = \frac{h(x+t) + h(x-t)}{2} + \frac{g(x+t) - g(x-t)}{2}. \qquad \diamond$$

Note that the above three cases in some sense cover everything which can happen for a two dimensional system: In the first case there are no characteristic directions, in the second case there is one characteristic direction, while in the last case there are two characteristic directions. Accordingly, a system is called **elliptic** in the first case, **parabolic** in the second and **hyperbolic** in the last.

Moreover, note that prescribing initial conditions on a characteristic surface will lead to contradictions in general.

Example 2.4. Consider the system

$$u_x = v, \qquad v_y = 0.$$

In our notation the system reads

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} u_x \\ v_x \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} u_y \\ v_y \end{pmatrix} - \begin{pmatrix} v \\ 0 \end{pmatrix} = 0$$

and the non-characteristic condition is

$$\det\begin{pmatrix} \nu_1 & 0 \\ 0 & \nu_2 \end{pmatrix} = \nu_1\nu_2 \neq 0.$$

In particular the coordinate directions are characteristic. Trying to prescribe the initial conditions $u(x, 0) = g(x)$ and $v(x, 0) = h(x)$ we get $u_x(x, 0) = g'(x)$ which contradicts the system unless $g'(x) = h(x)$. If we satisfy this constraint and prescribe $u(x, 0) = g(x)$ and $v(x, 0) = g'(x)$, we get many solutions $u(x, y) = g(x) + yf(y)$, $v(x, y) = g'(x)$, where f is arbitrary. $\diamond$

Our considerations so far might lead to the false impression that, at least locally, one can always find a solution. Of course for a nonlinear equation you need to assume that you can solve for the highest derivatives (otherwise you can construct trivial examples like $\exp(u_x) = 0$ which have no solution). But even in this case you need to be able to find a non-characteristic direction.

Example 2.5. Consider the equation

$$yu_x - xu_y + u = 0.$$

The characteristic curves are circles around the origin and nontrivial solutions must grow exponentially along the characteristics. Consequently you cannot get a nontrivial solution defined in a neighborhood of the origin. ◇

However, even if we exclude these obvious obstructions, there are cases where we have no solution. Of course this must be a system (since for scalar equations the method of characteristics provides a solution) and cannot be analytic (in which case the Cauchy–Kovalevskaya theorem applies). In fact, for a long time the dream of proving a general existence result at least for linear equations with smooth coefficients lured around until it was finally shattered by Lewy with the following example. This came as a surprise for many people including Lewy himself.

Example 2.6. Lewy's example[5] in complex notation is

$$\frac{\partial U}{\partial \bar{z}} - 2\mathrm{i}z\frac{\partial U}{\partial t} = f(t), \tag{2.6}$$

where $\frac{\partial}{\partial \bar{z}} = \frac{\partial}{\partial x} + \mathrm{i}\frac{\partial}{\partial y}$ and f is real-valued. Writing $z = x + \mathrm{i}y$, $U = u + \mathrm{i}v$ the system spelled out explicitly reads

$$u_x = v_y - 2y\, u_t - 2x\, v_t - f, \qquad v_x = -u_y + 2x\, u_t - 2y\, v_t \tag{2.7}$$

and hence the Cauchy–Kovalevskaya theorem implies existence of solutions provided f is real analytic. Lewy showed, that if there is a C^1 solution in a neighborhood of the origin, then f must be real analytic.

Later even simpler examples of the form $U_x + \mathrm{i}x\, U_y = F(x, y)$ appeared, which have no solution for certain F (see [**8**]). This is as simple as it gets since any such counterexample must have non-constant coefficients. Indeed in the constant coefficient case the solution of the inhomogeneous equation can be expressed with the help of the fundamental solution (to be discussed in case of the Poisson equation in Section 5.3), and the fact that every linear constant coefficient equation has a fundamental solution is the celebrated **Malgrange–Ehrenpreis theorem**.[6] ◇

Lemma 2.2 (Lewy)**.** *The partial differential equation* (2.6) *has a* C^1 *solution in a neighborhood of the origin if and only if* f *is real analytic.*

[5]Hans Lewy (1904–1988), German-American mathematician
[6]Bernhard Malgrange (1928), French mathematician
[6]Leon Ehrenpreis (1930–2010), American mathematician

Proof. Introduce $V(r,\varphi,t) := \sqrt{r}\mathrm{e}^{\mathrm{i}\varphi}U(\sqrt{r}\cos(\varphi),\sqrt{r}\sin(\varphi),t)$ and note that

$$2V_r + \frac{\mathrm{i}}{r}V_\varphi - 2\mathrm{i}V_t = \frac{\partial U}{\partial \bar{z}} - 2\mathrm{i}z\frac{\partial U}{\partial t} = f.$$

Next, consider $\bar{V}(r,t) = \frac{1}{\pi}\int_0^{2\pi} V(r,\varphi,t)d\varphi$ and note (using periodicity of V with respect to φ)

$$\bar{V}_t + \mathrm{i}\bar{V}_r = \frac{1}{\pi}\int_0^{2\pi}\big(V_t - \frac{1}{2r}V_\varphi + \mathrm{i}V_r\big)d\varphi = \mathrm{i}f.$$

Using $\zeta = t + \mathrm{i}r$, this last equations says $\frac{\partial}{\partial\bar{\zeta}}W(\zeta) = 0$, where $W(\zeta) := \bar{V}(r,t) - \mathrm{i}F(t)$ and F is a primitive of f (i.e. $F' = f$). Thus W satisfies the Cauchy–Riemann equations and hence is analytic in the upper half plane $r > 0$ and extends continuously to the boundary $r = 0$ with $W(t,0) = -\mathrm{i}F(t)$. Thus W can be extended to the lower half plane using $W(\bar{\zeta}) := -\overline{W(\zeta)}$. By the Schwarz reflection principle this extension is analytic in a neighborhood of the origin which implies that $W(t,0) = -\mathrm{i}F(t)$ is real analytic and so is $f(t) = F'(t)$. □

Problem 2.1. *Solve the Cauchy–Riemann equations with initial conditions* $v(x,0) = x$ *and* $w(x,0) = x$.

Problem 2.2. *Solve the heat equation with initial condition* $u(0,x) = \mathrm{e}^{\alpha x}$.

Problem* 2.3. *A function* $f \in C^\infty(\mathbb{R})$ *is in the* ***Gevrey class***[7] *of order* θ *if for every* $r > 0$*, there are some constants* M, a *such that*

$$|f^{(m)}(t)| \le M a^m (m!)^\theta, \qquad |t| < r.$$

Note that $\theta = 1$ *gives the class of real analytic functions, while for* $\theta > 1$ *the function* f *will no longer be real analytic in general.*

Show that if $\theta < 2$*, then*

$$u(t,x) = \sum_{m=0}^\infty \frac{f^{(m)}(t)}{(2m)!}x^{2m}$$

converges for all $x \in \mathbb{R}$ *and defines a solution of the heat equation. Note that this class of functions contains functions with compact support (see Problem 2.4).*

Problem* 2.4. *Show that*

$$\varphi(t) := \begin{cases} \mathrm{e}^{-1/t^2}, & t > 0,\\ 0, & t \le 0,\end{cases}$$

[7] Maurice Gevrey (1884–1957), French mathematician

is in the Gevrey class of order $\theta = \frac{3}{2}$. *(Hint: Use the Cauchy integral formula*

$$\varphi^{(m)}(t) = \frac{m!}{2\pi \mathrm{i}} \oint_\gamma \frac{\mathrm{e}^{-z^{-2}}}{(z-t)^{m+1}} dz$$

with $\gamma = \{t + \frac{t}{2}\mathrm{e}^{\mathrm{i}\vartheta} | 0 \le \vartheta \le 2\pi\}$*.)*

Problem 2.5. *Show that if* f_1*,* f_2 *are in the Gevrey class of order* θ_1*,* θ_2*, respectively, then* $f_1 f_2$ *is in the Gevrey class of order* $\max(\theta_1, \theta_2)$*.*

2.2. Second order equations

One of the big advantages of the Cauchy–Kovalevskaya theorem over the method of characteristics is that it enables us to solve higher order equations by turning them into first order systems. We illustrate this by an example first.

Example 2.7. Consider the **Laplace equation**

$$u_{xx} + u_{yy} = 0.$$

To transform it into a system suitable for the Cauchy–Kovalevskaya theorem we first introduce two new dependent variables

$$v := u_x, \quad w := u_y,$$

such that the system for (u, v, w) reads

$$u_x - v = 0, \quad u_y - w = 0, \quad v_x + w_y = 0.$$

However, for this system all surfaces are characteristic and so it looks like we are out of luck. On the other hand, using $v_y = u_{xy} = w_x$, we could add this equation to our system and drop $u_x = v$ which gives

$$u_y = w, \quad v_y = w_x, \quad w_y = -v_x.$$

Obviously the Cauchy–Kovalevskaya theorem applies to this new system if we choose the initial conditions

$$u(x,0) = u_0(x), \qquad v(x,0) = v_0(x), \qquad w(x,0) = w_0(x)$$

on $\Gamma = \{y = 0\}$. Moreover, observe that the equations for v and w are just the Cauchy–Riemann equations considered in Example 2.1. In particular, these equations do not involve u and we could apply the Cauchy–Kovalevskaya theorem to this smaller system and then determine u by a simple integration.

But what about the equation $u_x = v$ we have dropped? Are we allowed to do this? That is, is the new system equivalent to the original one? To shed some light on this, note that the new system implies $u_{xy} = w_x = v_y$ and hence we have $v = u_x$ if and only if $v(x,0) = u_x(x,0)$, that is, $v_0(x) = u_0'(x)$. But this is precisely what we get if we choose

$$u(x,0) = u_0(x), \qquad u_y(x,0) = w_0(x)$$

as the initial conditions for our original problem, which lead to

$$u(x,0) = u_0(x), \qquad v(x,0) = u_0'(x), \qquad w(x,0) = w_0(x)$$

for the corresponding system. In summary, choosing these initial conditions Example 2.1 implies $v(x,y) = \mathrm{Re}(u_0'(x+\mathrm{i}y) - \mathrm{i}w_0(x+\mathrm{i}y))$ and hence

$$u(x,y) = \mathrm{Re}\big(u_0(x+\mathrm{i}y) - \mathrm{i}W_0(x+\mathrm{i}y)\big),$$

where $W_0(x) = \int w_0(x)dx$ is a primitive of w_0.

In summary, there is a unique real analytic solution of the Laplace equation satisfying given real analytic initial conditions $u(x,0) = u_0(x)$, $u_y(x,0) = w_0(x)$ which is defined in a neighborhood of the plane $y = 0$ (or some part of this plane, if the initial conditions are defined only on some part). ◇

Now suppose we have a second order quasiliner equation

$$\sum_{j,k} A_{jk}(x,u,\nabla u)u_{x_j x_k} + b(x,u,\nabla u) = 0 \tag{2.8}$$

with initial conditions

$$u(y) = g(y), \qquad \nu(y)\cdot\nabla u(y) = h(y), \qquad y \in \Gamma. \tag{2.9}$$

That is, the value of u and the value of the normal derivative $\frac{\partial u}{\partial \nu}$ are given on Γ. Of course the matrix A can (and will) be assumed symmetric without loss of generality.

Following the same strategy as in our example we obtain:

Theorem 2.3. *Let the coefficients A_{jk}, b as well as the surface Γ and the initial conditions g, h be real analytic in a neighborhood of a point $x_0 \in \Gamma$. Suppose the **non-characteristic condition***

$$\nu(x_0)\cdot A(x_0, g(x_0), \nabla u(x_0))\nu(x_0) \neq 0, \tag{2.10}$$

where $\nu(x_0)$ is the normal vector of Γ at x_0 and $\nabla u(x_0)$ has to be expressed in terms of the tangential derivatives of g and of h at x_0, holds. Then the initial value problem (2.8), (2.9) has a unique real analytic solution in a neighborhood of x_0.

Proof. As in the case of systems, it will be convenient to first straighten out the boundary using $y := (x_1, \dots, x_{n-1}, x_n - \gamma(x_1, \dots, x_{n-1}))$. Setting

$v(y) := u(x)$ we have (of course $\gamma_{y_n} = 0$, but keeping this term avoids case distinctions)

$$\begin{aligned} u_{x_j} &= v_{y_j} - v_{y_n}\gamma_{y_j}, \\ u_{x_jx_k} &= v_{y_jy_k} - (v_{y_jy_n}\gamma_{y_k} + v_{y_ky_n}\gamma_{y_j}) + v_{y_ny_n}\gamma_{y_j}\gamma_{y_k} - v_{y_n}\gamma_{y_jy_k} \end{aligned}$$

and the new system is of the same form

$$\sum_{j,k} \tilde{A}_{jk}(y, v, \nabla v)v_{y_jy_k} + \tilde{b}(y, v, \nabla v) = 0$$

with initial conditions

$$v(\bar{y}, 0) = \tilde{g}(\bar{y}), \qquad v_{y_n}(\bar{y}, 0) = \tilde{h}(\bar{y}).$$

We will not need expressions for the new coefficients except for

$$\tilde{A}_{nn} = A_{nn} - 2\sum_{l=1}^{n-1} A_{ln}\gamma_{y_l} + \sum_{l,m=1}^{n-1} A_{lm}\gamma_{y_l}\gamma_{y_m},$$

which is required to relate the non-characteristic condition between both coordinates. Indeed, since $\nu = (\gamma_{x_1}, \dots, \gamma_{x_{n-1}}, -1)$ and $\tilde{\nu} = (0, \dots, 0, -1)$, the non-characteristic condition in the new coordinates is just $\tilde{A}_{nn} \neq 0$.

In summary, we can assume $\Gamma = \{x_n = 0\}$ without loss of generality. Now before turning (2.8) into a system, we note that we have

$$u(x) = g(x), \qquad u_{x_j}(x) = \begin{cases} g_{x_j}(x), & 1 \le j < n, \\ h(x), & j = n, \end{cases},$$

and

$$u_{x_jx_k}(x) = \begin{cases} g_{x_jx_k}(x), & 1 \le j, k < n, \\ h_{x_k}(x), & j = n, 1 \le k < n, \end{cases}$$

for $x \in \Gamma$. The missing derivative $u_{x_nx_n}$ follows from the differential equation (2.8) provided we have the non-characteristic condition

$$A_{nn} \neq 0.$$

Now we set

$$v^0 := u, \quad v^1 := u_{x_1}, \quad \dots, \quad v^n := u_{x_n}.$$

Then we have

$$v^j_{x_n} = \begin{cases} u_{x_n} = v^n, & j = 0, \\ u_{x_jx_n} = u_{x_nx_j} = v^n_{x_j}, & 1 \le j < n, \\ u_{x_nx_n} = \frac{-1}{A_{nn}(x,v)}\left(\sum_{j,k\neq n,n} A_{jk}(x, v)v^j_{x_k} + b(x, v)\right), & j = n, \end{cases}$$

and the initial condition read

$$v^j(x) = \begin{cases} g(x), & j = 0, \\ g_{x_j}(x), & 1 \le j < n, \\ h(x), & j = n, \end{cases}$$

for $x \in \Gamma$. Finally, the Cauchy–Kovalevskaya theorem ensures that there is a solution of this system and the first component $u := v^0$ will solve (2.8) together with the corresponding initial conditions (2.9). Indeed the system implies $u_{x_n} = v^n$ as well as $v^j_{x_n} = v^n_{x_j} = u_{x_n x_j}$, that is $(v^j - u_{x_j})_{x_n} = 0$. From the initial condition we deduce $v^j = u_{x_j}$ on Γ and hence $v^j = u_{x_j}$ for $1 \le j < n$ as well. Finally, the last equation of the system shows that u solves (2.8). □

A surface Γ on which the non-characteristic condition fails throughout is called a **characteristic surface**. Note that on a characteristic surface the values of the differential equations is determined by the initial data alone (since the coefficient in front of the missing second order normal derivative is zero by the very definition of a characteristic surface). Hence it is not surprising that prescribing initial values on a characteristic surface will lead to contradictions in general.

Example 2.8. Let us solve

$$u_{xy} = 0$$

with initial conditions $u(x,0) = g(x)$, $u_y(x,0) = h(x)$ on $\Gamma = \{y = 0\}$. Note that in this case $\nu = (0,1)$, $A = \frac{1}{2}\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ and the non-characteristic condition

$$\frac{1}{2} \begin{pmatrix} 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 0$$

is violated. Indeed differentiating the second initial condition with respect to x gives $u_{xy}(x,0) = h'(x)$ which will violate the differential equation unless $h' = 0$. Hence if $h' \neq 0$ there is no solution and if $h' = 0$ there are many solutions: $u(x,y) = g(x) + h(0)y + f(y)y^2$. ◇

Example 2.9. Let us use the Cauchy–Kovalevskaya theorem to solve the **wave equation**

$$u_{tt} = u_{xx}$$

with initial conditions $u(0,x) = g(x)$, $u_t(0,x) = h(x)$ on $\Gamma = \{t = 0\}$. We have $\nu = (1,0)$, $A = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$ and the non-characteristic condition

$$\begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 1$$

holds. Note that there is no real need to check it, since we would also notice that it is violated when we fail to set up the system.

Using $v := u_x$, $w := u_t$ the system reads

$$u_t = w, \qquad v_t = u_{xt} = w_x, \qquad w_t = u_{tt} = u_{xx} = v_x$$

with initial conditions

$$u(0,x) = g(x), \qquad v(0,x) = g'(x), \qquad w(0,x) = h(x).$$

The system for v, w was solved in Example 2.3, where we found

$$w(t,x) = \frac{h(x+t) + h(x-t)}{2} + \frac{g'(x+t) - g'(x-t)}{2}.$$

Integration with respect to t and using the initial condition for u we again obtain **d'Alembert's formula**

$$u(t,x) = \frac{g(x+t) + g(x-t)}{2} + \frac{1}{2}\int_{x-t}^{x+t} h(y)dy.$$

already found in Example 1.12. $\diamond$

Example 2.10. Note that while in two dimensions the characteristics of a constant coefficient linear equation are just straight lines, this changes drastically if we increase the dimension. For example, let us look at the characteristic surfaces of the wave equation

$$u_{tt} = \Delta u = \sum_{j=1}^{n} u_{x_j x_j}.$$

If we assume that a characteristic surface is implicitly given as the level set of a smooth function, $S(t,x) = 0$, then the normal vector is given by the gradient, $\nu = (S_t, \nabla_x S)$ and the characteristic condition reads

$$S_t^2 - |\nabla_x S|^2 = 0.$$

Assuming that we can solve S for t, that is, that the surface is given by $S(t,x) = t - v(x) = 0$, we obtain the **eikonal equation**

$$|\nabla v| = 1$$

for a medium with constant refraction index 1 discussed in Example 1.10.

Two simple examples of characteristic surfaces are the planes $S(t,x) = t \pm (r + a \cdot x)$, where $|a| = 1$, or the cones $S(t,x) = t \pm (r - |x - x_0|)$. $\diamond$

Hence the Cauchy–Kovalevskaya theorem allowed us to solve the three most ubiquitous partial differential equations: The Laplace equation, the heat equation, and the wave equation. However, while for the wave equation d'Alembert's formula is precisely what one wants, this is not the case for the other two equations. It was already pointed out in Example 2.2 that initial conditions for the heat equation are naturally posed for $t = 0$, but this surface is characteristic and hence cannot be handled within the current framework of the Cauchy–Kovalevskaya theorem. Similarly, while Example 2.7 might seem like a nice result, it is again of little practical use. The problem is not only the fact that the solution is merely defined locally, but also that this is again not the *right* kind of setting for the Laplace equation. The usual problem is to find solutions in a (bounded) domain which attain prescribed values at the boundary, while here we prescribe both the values and the values of the normal derivative. That this is not the *right* kind of setting is also confirmed by the fact that this type of problem is ill-posed as pointed out by Hadamard[8]:

Example 2.11. The Laplace equation with initial conditions $u(x,0) = \varepsilon\sin(\frac{x}{\varepsilon})$, $u_y(x,0) = 0$ has the solution

$$u(x,y) = \mathrm{Re}(u(x+\mathrm{i}y,0)) = \varepsilon\cosh(\frac{y}{\varepsilon})\sin(\frac{x}{\varepsilon}).$$

Note that even though the initial conditions is continuous with respect to ε, the solution is not. In this sense the problem is **ill-posed**. ◇

Example 2.12. The Laplace equation with initial conditions $u(x,0) = \frac{-\varepsilon}{x^2+\varepsilon^2}$, $u_y(x,0) = \frac{x^2-\varepsilon^2}{(x^2+\varepsilon^2)^2}$ has the solution

$$u(x,y) = \mathrm{Re}(\frac{\mathrm{i}}{x+\mathrm{i}(y-\varepsilon)}) = \frac{y-a}{x^2+(y-\varepsilon)^2}.$$

Note that even though the initial conditions is perfectly nice, the solution blows up at $y=\varepsilon$. Hence a linear constant coefficient equation might not have global solutions. This is in contradistinction to ordinary differential equations, where linear equations always have global solutions (cf. Theorem 2.17 in [**33**]). ◇

So the results in this section are not the end of the story and we will derive different methods in the following chapters.

We end this section with the remark, that it is possible to treat higher order equations as well as fully nonlinear systems in a similar manner, see [**11**].

Problem 2.6. *Solve the Laplace equation with initial conditions* $u(0,y) = 0$, $u_x(0,y) = y$.

Problem 2.7. *Transform the equation* $u_{tt}+u_{xx}-u_{xt}+u_t = 0$, $u(0,x) = g(x)$, $u_t(0,x) = h(x)$ *into a system suitable for the Cauchy–Kovalevskaya theorem. What are the initial conditions such that solutions of the resulting system also solve the original equation.*

Problem 2.8. *Consider the* ***Stokes system***[9] *which describes the stationary velocity* (u,v) *field and pressure* p *of a two dimensional incompressible*

[8]Jacques Hadamard (1865–1963), French mathematician

[9]Sir George Stokes (1819–1903), Anglo-Irish physicist and mathematician

Newtonian fluid:

$$u_{xx} + u_{yy} = p_x, \qquad v_{xx} + v_{yy} = p_y, \qquad u_x + v_y = 0.$$

Transform it into a system suitable for the Cauchy–Kovalevskaya theorem and discuss the corresponding initial conditions.

Chapter 3

Separation of variables

3.1. The heat equation for a thin rod

During his investigation of heat conduction Fourier studied the simple model of a thin rod. Let $u(t,x)$ denote the temperature distribution at time $t \in \mathbb{R}$ at the point $x \in [0,1]$. Assuming that there are no heat sources within the rod, energy conservation implies that the problem can be described by a scalar conservation law (1.21). Assuming that the flux is proportional to the temperature gradient (**Fourier's law**), $F(u) = -k\, u_x$ (the minus reflects the fact that heat is transferred to regions with a lower temperature) leads to the (one-dimensional) **heat equation** $u_t = k\, u_{xx}$. By scaling the time variable we can assume $k = 1$ without loss of generality. Indeed, if u solves the heat equation with $k = 1$, then $v(t,x) := u(kt,x)$ solves the heat equation with general $k \in \mathbb{R}$. Hence we will consider

$$u_t = u_{xx}. \tag{3.1}$$

The very same equation arises when u models the concentration of some substance, where again one assumes that the flux is proportional to the gradient of the concentration (**Fick's law**[1] of diffusion). In this context the above equation is also known as **diffusion equation**.

It is usually assumed, that the temperature at $x = 0$ and $x = 1$ is fixed, say $u(t,0) = a_0$ and $u(t,1) = a_1$. By considering $u(t,x) \to u(t,x) - a_0 - (a_1 - a_0)x$ it is clearly no restriction to assume $a_0 = a_1 = 0$. That is, we assume the **Dirichlet boundary conditions**[2]

$$u(t,0) = u(t,1) = 0. \tag{3.2}$$

[1] Adolf Eugen Fick (1829–1901), German physiologist

[2] Peter Gustav Lejeune Dirichlet (1805 –1859), German mathematician

Moreover, the initial temperature distribution $u(0,x) = g(x)$ is assumed to be known as well.

Since finding the solution seems at first sight unfeasible, we could try to find at least some solutions of (3.1). For example, we could make an ansatz for $u(t,x)$ as a product of two functions, each of which depends on only one variable, that is,

$$u(t,x) := w(t)y(x). \tag{3.3}$$

Plugging this ansatz into the heat equation we arrive at

$$\dot{w}(t)y(x) = y''(x)w(t), \tag{3.4}$$

where the dot refers to differentiation with respect to t and the prime to differentiation with respect to x. Bringing all t, x dependent terms to the left, right side, respectively, we obtain

$$\frac{\dot{w}(t)}{w(t)} = \frac{y''(x)}{y(x)}. \tag{3.5}$$

Accordingly, this ansatz is known as **separation of variables**. This method was originally introduced by d'Alembert (1747) and Euler (1748) for the wave equation (to be discussed in Section 3.3).

Now if this equation should hold for all t and x, the quotients must be equal to a constant $-\lambda$ (we choose $-\lambda$ instead of λ for convenience later on). That is, we are led to the equations

$$-\dot{w}(t) = \lambda w(t) \tag{3.6}$$

and

$$-y''(x) = \lambda y(x), \qquad y(0) = y(1) = 0, \tag{3.7}$$

which can easily be solved. The first one gives

$$w(t) = c_1 \mathrm{e}^{-\lambda t} \tag{3.8}$$

and the second one

$$y(x) = c_2 \cos(\sqrt{\lambda}x) + c_3 \sin(\sqrt{\lambda}x). \tag{3.9}$$

However, $y(x)$ must also satisfy the boundary conditions $y(0) = y(1) = 0$. The first one $y(0) = 0$ is satisfied if $c_2 = 0$ and the second one yields (c_3 can be absorbed by $w(t)$)

$$\sin(\sqrt{\lambda}) = 0, \tag{3.10}$$

which holds if $\lambda = (\pi n)^2$, $n \in \mathbb{N}$ (in the case $\lambda < 0$ we get $\sinh(\sqrt{-\lambda}) = 0$, which cannot be satisfied and explains our choice of sign above). In summary, we obtain the family of solutions

$$u_n(t,x) := c_n \mathrm{e}^{-(\pi n)^2 t} \sin(n\pi x), \qquad n \in \mathbb{N}. \tag{3.11}$$

So we have found a large number of solutions, but we still have not dealt with our initial condition $u(0,x) = g(x)$. This can be done using the **superposition principle** which holds since our equation is linear: Any finite linear combination of the above solutions will again be a solution. Moreover, under suitable conditions on the coefficients, we can even consider infinite linear combinations. In fact, choosing

$$u(t,x) := \sum_{n=1}^{\infty} c_n \mathrm{e}^{-(\pi n)^2 t} \sin(n\pi x), \tag{3.12}$$

where the coefficients c_n decay sufficiently fast (e.g. absolutely summable), we obtain further solutions of our equation. Of course for this last statement to hold we need to ensure that we can interchange summation and differentiation.

Lemma 3.1. *Suppose c_n is an absolutely summable sequence*

$$\sum_{n=1}^{\infty} |c_n| < \infty. \tag{3.13}$$

Then u defined by (3.12) *is in $C([0,\infty)\times[0,1]) \cap C^\infty((0,\infty)\times[0,1])$ and solves the heat equation* (3.1) *for $(t,x) \in (0,\infty)\times[0,1]$ as well as the boundary conditions $u(t,0) = u(t,1) = 0$.*

Proof. Using $|c_n \mathrm{e}^{-(\pi n)^2 t} \sin(n\pi x)| \le |c_n|$ shows that the series (3.12) converges uniformly for $(t,x) \in [0,\infty)\times[0,1]$ and hence $u(t,x)$ is continuous there by the Weierstrass[3] M-test.

Moreover,

$$|\frac{\partial}{\partial t} c_n \mathrm{e}^{-(\pi n)^2 t} \sin(n\pi x)| = |-(\pi n)^2 c_n \mathrm{e}^{-(\pi n)^2 t} \sin(n\pi x)| \le \pi^2 n^2 |c_n| \mathrm{e}^{-(\pi n)^2 t_0}$$

for $t \ge t_0 > 0$ shows that the derivative of the partial sum converges uniformly and hence (3.12) is continuously differentiable with respect to t for $(t,x) \in [t_0,\infty)\times[0,1]$ with

$$u_t(t,x) = -\sum_{n=1}^{\infty} (\pi n)^2 c_n \mathrm{e}^{-(\pi n)^2 t} \sin(n\pi x).$$

In particular, we can interchange differentiation and summation. The same argument applies to the partial derivative with respect to x as well as to any higher order derivatives. Since $t_0 > 0$ is arbitrary the claim follows. □

Note that our rod will attain the temperature specified at the boundary exponentially fast (**Newton's law of cooling**):

[3]Karl Weierstrass (1815–1897), German mathematician

Corollary 3.2. *The solution* (3.12) *satisfies*

$$|u(t,x)| \le \left(\sum_{n=1}^{\infty} |c_n| \right) \mathrm{e}^{-\pi^2 t}. \tag{3.14}$$

Finally, the remaining question is how to take the initial condition into account. Setting $t = 0$ in (3.12) we see

$$u(0,x) = \sum_{n=1}^{\infty} c_n \sin(n\pi x) \tag{3.15}$$

and finding the correct coefficients c_n to satisfy our initial condition $u(0,x) = g(x)$ boils down to expanding g into a Fourier sine series

$$g(x) = \sum_{n=1}^{\infty} g_n \sin(n\pi x), \tag{3.16}$$

where the Fourier coefficients are given by (Problem 3.1)

$$g_n = 2 \int_0^1 \sin(n\pi x) g(x) dx. \tag{3.17}$$

Then (3.12) will satisfy our initial condition if we choose $c_n = g_n$.

That such an expansion (3.16) is possible was first postulated by Fourier. He was however disputed by other leading mathematicians, in particular Cauchy, until a few years later Dirichlet showed that this is indeed possible at least for piecewise continuously differentiable functions. Determining when a given function g can be expanded into a convergent sine expansion with absolutely convergent coefficients is a formidable task and we will not address it here. The set of functions with this property is known as **Wiener algebra**[4] and convenient sufficient conditions are known. For example Bernstein's theorem[5], which states that g is in the Wiener algebra if g is Hölder[6] continuous of exponent $\gamma > \frac{1}{2}$ and vanishes at the boundary points, $g(0) = g(1) = 0$. We refer to Appendix A.3 for more information.

Note that combining (3.12) and (3.16) gives

$$\begin{aligned} u(t,x) &= 2 \sum_{n=1}^{\infty} \mathrm{e}^{-(\pi n)^2 t} \sin(n\pi x) \int_0^1 \sin(n\pi y) g(y) dy \\ &= \int_0^1 \left(2 \sum_{n=1}^{\infty} \mathrm{e}^{-(\pi n)^2 t} \sin(n\pi x) \sin(n\pi y) \right) g(y) dy \end{aligned}$$

[4]Norbert Wiener (1894–1964), American mathematician and philosopher

[5]Sergei Natanovich Bernstein (1880–1968), Russian mathematician

[6]Otto Hölder (1859–1937), German mathematician

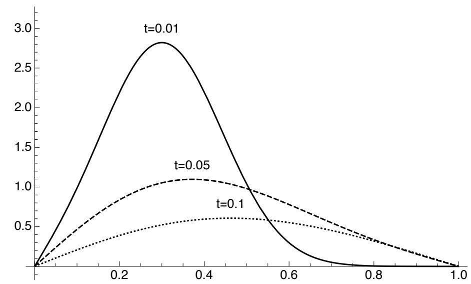


Figure 3.1. The heat kernel $K(t,x,y)$ for $t = 0.01$, 0.05, 0.1 and $y = 0.3$.

and hence the solution can be written as

$$u(t,x) = \int_0^1 K(t,x,y)g(y)dy, \qquad t > 0, \tag{3.18}$$

where the **heat kernel** is given by (see Figure 3.1)

$$\begin{aligned} K(t,x,y) &:= 2\sum_{n=1}^{\infty} \mathrm{e}^{-(\pi n)^2 t} \sin(n\pi x)\sin(n\pi y) \\ &= \frac{1}{2}\big(\vartheta(\frac{x-y}{2}, \mathrm{i}\pi t) - \vartheta(\frac{x+y}{2}, \mathrm{i}\pi t)\big). \end{aligned} \tag{3.19}$$

Here

$$\vartheta(z,\tau) := \sum_{n\in\mathbb{Z}} \mathrm{e}^{\mathrm{i}\pi n^2\tau + 2\pi \mathrm{i} n z} = 1 + 2\sum_{n\in\mathbb{N}} \mathrm{e}^{\mathrm{i}\pi n^2 \tau}\cos(2\pi n z), \quad \mathrm{Im}(\tau) > 0, \tag{3.20}$$

is the **Jacobi theta function**. The theta function is entire with respect to z and satisfies

$$\vartheta(z + m + n\tau, \tau) = \mathrm{e}^{-2\pi \mathrm{i} n z - \pi \mathrm{i} n^2 \tau}\vartheta(z,\tau), \quad \vartheta(-z) = \vartheta(z), \tag{3.21}$$

$m, n \in \mathbb{Z}$.

Note that by construction $(t,x) \mapsto K(t,x,y)$ satisfies the heat equation (which can of course also be verified directly by differentiating the series for the theta function) and we have the symmetries

$$K(t,x,y) = K(t,y,x), \qquad K(t,1-x,y) = K(t,x,1-y). \tag{3.22}$$

In particular, (3.18) will solve the heat equation under the sole assumption that g is integrable. However, it is not clear in what sense (3.18) will satisfy the initial conditions (note that $K(t,x,y)$ is not well-defined for $t = 0$) unless we assume that g is in the Wiener algebra such that we can resort to our original arguments. Also note that since ϑ is entire with respect to z, so is (3.18) for $t > 0$.

So we have found a solution to the initial value problem, but is this the only solution? Before we try to answer this question we need to specify what precisely we mean by a solution. Since the heat equation involves only one time derivative but two spatial derivatives, we set

$$C^{1;2}(I \times U) := \{u \in C(I \times U) | u_t, u_x, u_{xx} \in C(I \times U)\}, \tag{3.23}$$

where $I, U \subseteq \mathbb{R}$ are some intervals. Then we will call a function $u \in C^{1;2}((0,\infty) \times (0,1)) \cap C([0,\infty) \times [0,1])$ a solution if it satisfies the heat equation in $(0,\infty) \times (0,1)$, the boundary conditions and any given initial conditions.

Now we look at the **energy** functional associated with a solution $u \in C^{1;2}((0,\infty) \times [0,1]) \cap C([0,\infty) \times [0,1])$,

$$E(t) := \frac{1}{2} \int_0^1 u(t,x)^2 dx \geq 0. \tag{3.24}$$

Differentiating this expression we get

$$\begin{aligned} \frac{d}{dt} E(t) &= \int_0^1 u_t(t,x) u(t,x) dx = \int_0^1 u_{xx}(t,x) u(t,x) dx \\ &= -\int_0^1 u_x(t,x)^2 dx \leq 0, \end{aligned} \tag{3.25}$$

where we have used integration by parts together with the Dirichlet boundary conditions to obtain the last equality. In particular, the energy is nonincreasing

$$E(t) \leq E(0) = \frac{1}{2} \int_0^1 g(x)^2 dx. \tag{3.26}$$

Note that using the Poincaré inequality one can even get exponential decay (Problem 3.2). Hence we have energy **dissipation**.

Consequently, the heat equation with vanishing initial condition $g = 0$ has only the trivial solution and since the difference of two solutions corresponding to the same initial condition g will be a solution vanishing at $t = 0$, we obtain:

Lemma 3.3. *The heat equation with Dirichlet boundary condition and prescribed initial condition g has at most one solution $u \in C^{1;2}((0,\infty) \times [0,1]) \cap C([0,\infty) \times [0,1])$.*

In fact, taking differences shows that we have stability with respect to the initial condition in the sense that

$$\int_0^1 \big(u_1(t,x) - u_2(t,x)\big)^2 dx \leq \int_0^1 \big(g_1(x) - g_2(x)\big)^2 dx, \tag{3.27}$$

where u_1, u_2 are the solutions corresponding to the initial conditions g_1, g_2, respectively. In this sense the problem is **well-posed**.

Another interesting property is that a maximum will either be assumed initially at $t = 0$ or otherwise at one of the boundary points $x = 0$ or $x = 1$.

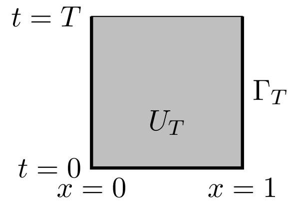


Figure 3.2. Parabolic boundary (thick line)

In fact, this property will not require any fixed boundary values. The key observation is that at a maximum (t_0, x_0) the gradient must vanish while the Hesse matrix must be negative definite. In particular, in our situation this implies $u_t = 0$ and $u_{xx} \le 0$. But this is incompatible with the heat equation at least if we have a strict inequality $u_{xx} < 0$. To handle the limiting case we make the observation that for this argument to work, we do not need the precise heat equation $u_t = u_{xx}$, but an inequality $v_t < v_{xx}$ will suffice. So we add an extra term to get such an inequality and then investigate the limit, when this extra term disappears.

Set

$$U_T := (0,T] \times (0,1), \qquad \Gamma_T := \overline{U_T} \setminus U_T, \tag{3.28}$$

where Γ_T is known as the **parabolic boundary** (see Figure 3.2). Moreover, we call v a **subsolution** of the heat equation if v satisfies $v_t \le v_{xx}$. Similarly, we call v a **supersolution** of the heat equation if v satisfies $v_t \ge v_{xx}$. Clearly v will be a subsolution if $-v$ is a supersolution and vice versa.

Theorem 3.4 (**Maximum principle**). *Let $v \in C(\overline{U_T}) \cap C^{1;2}(U_T)$ be a subsolution of the heat equation. Then*

$$\max_{\overline{U_T}} v \le \max_{\Gamma_T} v. \tag{3.29}$$

Proof. As outlined we consider $v^\varepsilon(t,x) := v(t,x) - \frac{\varepsilon}{2}(1-x)x$ such that $v_t^\varepsilon = v_t$ and $v_{xx}^\varepsilon = v_{xx} + \varepsilon$. In particular, $v_t^\varepsilon \le v_{xx}^\varepsilon - \varepsilon$ on U_T. Then, as argued before, v^ε cannot attain an interior maximum in U_T. Hence the maximum must be either on Γ_T or on the line $\{T\} \times (0,1)$. At $t = T$ the gradient might not vanish, but we at least have $v_t^\varepsilon \ge 0$ (since otherwise, v^ε would attain larger values for $t < T$) and $v_{xx}^\varepsilon \le 0$. Consequently, v^ε attains its maximum on Γ_T, that is $v \le v^\varepsilon \le \max_{\Gamma_T} v^\varepsilon$. Taking $\varepsilon \to 0$ establishes the claim. $\square$

Of course if v is a supersolution we get a corresponding **minimum principle**

$$\min_{\overline{U_T}} v \ge \min_{\Gamma_T} v \tag{3.30}$$

by applying the maximum principle to $-v$. Clearly, for solutions of the heat equation both the maximum and the minimum principle hold. This is particularly relevant for applications modeling a diffusion process, where u corresponds to the concentration of a substance. This concentration should always remain positive, which is ensured by the minimum principle.

The maximum/minimum principle is a key tool whose power must not be underestimated as we are going to demonstrate with a few simple consequences. We start with an **a priori bound** for the initial/boundary value problem.

Corollary 3.5. *Let $u \in C(\overline{U_T}) \cap C^{1;2}(U_T)$ solve*

$$u_t = u_{xx}, \qquad \begin{cases} u(0,x) = g(x), & x \in (0,1), \\ u(t,0) = a_0(t),\ u(t,1) = a_1(t), & t \in [0,T]. \end{cases} \tag{3.31}$$

Then

$$|u| \le \max_{[0,1]} |g| + \max_{[0,T]} |a_0| + \max_{[0,T]} |a_1|. \tag{3.32}$$

Applying this to the difference of two solutions gives an alternate proof for uniqueness as well as stability with respect to small changes in the initial or boundary data.

Next we have the following **comparison principle** which explains the name subsolution:

Corollary 3.6. *If u is a solution of the heat equation and v a subsolution, then $v \le u$ on the parabolic boundary Γ_T implies $v \le u$ on all of $\overline{U_T}$.*

Proof. Apply the maximum principle to the subsolution $v - u$. □

There is also a **strong maximum principle** which states that if a solution of the heat equation attains a maximum at an interior point $(t_0, x_0) \in U_T$, then u is actually constant on $\overline{U_{t_0}}$, that is, it is constant up to the time t_0. But this is harder to prove (see Theorem 6.15).

As an interesting application of these circle of ideas we now obtain some basic properties of the heat kernel which are not so obvious from its definition.

Lemma 3.7. *The heat kernel has the following properties:*

(i) $K(t,x,y) > 0$ *for* $x, y \in (0,1)$.

(ii) $\int_0^1 K(t,x,y)dy \le 1$.

(iii) *For arbitrary open intervals I, J with $\overline{J} \subset I \subseteq (0,1)$ we have $\lim_{t\to 0} \int_I K(t,x,y)dy = 1$ for all $x \in I$ and uniformly with respect to $x \in J$.*

Proof. For a given interval J with $\overline{J} \subset (0,1)$ let g_J be some function in the Wiener algebra satisfying $0 \le g \le 1$ with $g(y) = 1$ for $y \in J$ (e.g. a smooth function with compact support in $(0,1)$). Denote by $u_J(t,x) \in C([0,\infty)\times[0,1]) \cap C^\infty((0,\infty)\times[0,1])$ the corresponding solution of the heat equation with initial condition g_J (which for $t>0$ is given by (3.18)).

If $K(t_0,x_0,y_0) < 0$ we would have $K(t_0,x_0,y) < 0$ for y in some neighborhood J of y_0. Then u_J would be negative at (t_0,x_0) contradicting the fact that $u_J \ge 0$ by the minimum principle. Hence $K(t,x,y) \ge 0$. For the strict inequality see Problem 3.3.

To see the second claim suppose there is some (t,x) such that this integral is strictly larger than one. Choosing a sufficiently large interval J we will have $1 < \int_J K(t,x,y)dy \le u_J(t,x)$ contradicting the maximum principle.

For the last claim we choose again g_J but make the additional requirement that g_J has support in I. Then

$$\liminf_{t\to 0} \int_I K(t,x,y)dy \ge \lim_{t\to 0} u_J(t,x) = 1. \qquad \square$$

Note that property (i) shows that if a nonvanishing initial condition $g \ge 0$ has compact support, $u(t,x)$ will be strictly positive for all $t > 0$. In particular, a small change of the initial condition in a small neighborhood will be immediately propagated to the entire interval. In this sense the heat equation exhibits **infinite propagation speed**.

This lemma tells us that $x \mapsto K(t,x,y)$ concentrates more and more around y as $t \downarrow 0$ in the sense that the integral over arbitrarily small neighborhoods around y tends to 1 while the integral over the rest must tend to 0 (the sum of both parts is bounded by 1 by property (ii)). Informally we could also describe this as $K(t,x,y)$ converging to a delta function centered at y as $t \downarrow 0$. More precisely, we have:

Theorem 3.8. *Let $g \in C([0,1])$ with $g(0) = g(1) = 0$. Then u defined via (3.18) for $t > 0$ and $u(0,x) := g(x)$ is in $C([0,\infty)\times[0,1]) \cap C^\infty((0,\infty)\times[0,1])$, solves the heat equation, and satisfies the Dirichlet boundary conditions.*

Proof. That u is smooth for $t > 0$ and satisfies the heat equation follows from the corresponding properties of K upon interchanging differentiation and integration (Lemma A.6). Also that the Dirichlet boundary conditions are satisfied is immediate. To see continuity for $t = 0$ fix $\varepsilon > 0$, $x_0 \in (0,1)$ and abbreviate $M := \max|g|$. Choose a sufficiently small open interval I containing x_0 such that $|g(x) - g(y)| \le \varepsilon$ for $x, y \in I$. Moreover, let J be another open interval containing x_0 with $\overline{J} \subset I$. Then

$$\begin{aligned}
&\int_0^1 K(t,x,y)|g(y)-g(x)|dy \\
&\quad = \int_I K(t,x,y)|g(y)-g(x)|dy + \int_{[0,1]\setminus I} K(t,x,y)|g(y)-g(x)|dy \\
&\quad \le \varepsilon + 2M \int_{[0,1]\setminus I} K(t,x,y)dy
\end{aligned}$$

for $x \in I$. Now by property *(iii)* this last integral tends to 0 uniformly with respect to $x \in J$. Since ε is arbitrary the same is true for the original integral:

$$|u(t,x)-g(x)| \le \int_0^1 K(t,x,y)|g(y)-g(x)|dy + M\left|\int_0^1 K(t,x,y)dy - 1\right|$$

and the claim follows from property (iii). The case when $x_0 \in \{0,1\}$ is similar and left as an exercise. □

Note that this does not imply that the Fourier sine series of a continuous function converges! It is only a certain *regularization* (in this context sometimes known as heat kernel regularization) which converges. However, it incidentally shows that the Fourier coefficients of a continuous function uniquely determine the function. It also shows that if the Fourier series of a continuous function g converges at some point x, then it converges to $g(x)$.

Now that we have a quite good understanding of the homogenous equation, we turn to the inhomogeneous problem

$$u_t = u_{xx} + f(t,x), \qquad \begin{cases} u(t,0) = u(t,1) = 0, \\ u(0,x) = g(x). \end{cases} \tag{3.33}$$

Our considerations thus far suggest to make an ansatz as a Fourier series

$$u(t,x) = \sum_{n=1}^{\infty} c_n(t)\sin(n\pi x), \qquad f(t,x) = \sum_{n=1}^{\infty} f_n(t)\sin(n\pi x) \tag{3.34}$$

which gives

$$\dot{c}_n(t) = -(\pi n)^2 c_n(t) + f_n(t). \tag{3.35}$$

The solution of this ordinary differential equation is given by

$$c_n(t) = g_n \mathrm{e}^{-(\pi n)^2 t} + \int_0^t \mathrm{e}^{-(\pi n)^2(t-s)} f_n(s)ds \tag{3.36}$$

and we see that this approach indeed provides a solution if we assume

$$|f_n(t)| \le M_n, \qquad \sum_{n=1}^{\infty} M_n < \infty. \tag{3.37}$$

To see this observe $|c_n(t)| \le |g_n| \mathrm{e}^{-(\pi n)^2 t} + \frac{M_n}{(\pi n)^2}$.

As always with inhomogeneous linear equations, the solution is a sum of a solution of the homogenous problem plus a particular solution of the inhomogeneous equation. In the case of ordinary differential equations a particular solution is given by the variation of constants formula as displayed in (3.36). Informally speaking it amounts to computing the homogenous time evolution starting at s, with the inhomogeneous term as initial condition, up to t, and then integrating the result with respect to s from 0 to t. When summing up the Fourier coefficients (3.36) this structure is preserved and we can express our result in terms of the heat kernel as

$$u(t,x) = \int_0^1 K(t,x,y) g(y) dy + \int_0^t \int_0^1 K(t-s,x,y) f(s,y) dy\, ds, \quad t > 0. \tag{3.38}$$

This extension of the variation of constants formula from ODEs to PDEs is know as **Duhamel principle**.[7]

Our solution is such that the first term satisfies the initial condition $u(0,x) = g(x)$ while the second vanishes at $t = 0$. However, note that the integrand has to be interpreted with care at the upper integration limit $s = t$. We have $\int_0^1 K(t-s,x,y) f(s,y) dy \to f(t,x)$ as $s \to t$ and hence all is well-defined, but things get tricky when one tries to differentiate this formula. We will show in Theorem 4.6 below that this works if f is uniformly Hölder continuous with respect to x. For now we are happy with (3.37).

Instead of Dirichlet boundary conditions we could have also assumed **Robin boundary conditions**[8]

$$u_x(t,0) - a_0 u(t,0) = u_x(t,1) + a_1 u(t,1) = 0 \tag{3.39}$$

which model the case that energy is lost at the boundary at a rate proportional to the temperature. The special case $a_0 = a_1 = 0$ corresponds to insulated ends and is known as **Neumann boundary conditions**[9]

$$u_x(t,0) = u_x(t,1) = 0. \tag{3.40}$$

Of course it is also possible to have different boundary conditions at different endpoints. Another possibility is **periodic boundary conditions**

$$u(t,-1) = u(t,1), \qquad u_x(t,-1) = u_x(t,1). \tag{3.41}$$

Note that in this case you get precisely the usual Fourier series and we have chosen the interval $[-1,1]$ to be consistent with (A.61) (with $L = 1$).

---

[7] Jean-Marie Duhamel (1797–1872), French mathematician and physicist

[8] Victor Gustave Robin (1855–1897), French mathematician

[9] Carl Neumann (1832 –1925), German mathematician

It is straightforward to extend the analysis of this section to these cases. You will always get a sequence of eigenvalues together with orthogonal eigenfunctions (see the next section; in particular Problem 3.12). Note however, that most cases lead to a trigonometric equation for the eigenvalues which cannot be solved explicitly. A nice discussion for Robin boundary conditions can be found in [**30**, Section 2.4].

Finally, note that all these boundary conditions are homogenous, in the sense that the resulting problem is linear such that the superposition principle holds. If one prescribes for example the temperature at the endpoints, one obtains **inhomogeneous Dirichlet boundary conditions**

$$u(t,0) = a_0(t), \qquad u(t,1) = a_1(t). \tag{3.42}$$

This problem can be reduced to an inhomogeneous equation with homogenous boundary conditions by a simple transformation — Problem 3.10. Similarly one can consider inhomogeneous Neumann or Robin boundary conditions. For example, an energy loss proportional to the difference to the temperature $\bar{u}$ of the surrounding medium at the boundary points is modeled by the inhomogeneous Robin boundary conditions

$$u_x(t,0) = a_0(u(t,0) - \bar{u}), \qquad u_x(t,1) = -a_1(u(t,1) - \bar{u}). \tag{3.43}$$

Of course this case can be reduced to the homogenous case by considering $v(t,x) = u(t,x) - \bar{u}$.

Problem* 3.1. *Show that for $n, m \in \mathbb{N}$ we have*

$$2\int_0^1 \sin(n\pi x)\sin(m\pi x)dx = \begin{cases} 1, & n = m, \\ 0, & n \neq m. \end{cases}$$

Conclude that the Fourier sine coefficients g_n of $g(x)$ are given by (3.17) *provided the sum in* (3.16) *converges uniformly.*

Problem* 3.2. *Show that for $u \in C^1[0,1]$ with $u(0) = u(1) = 0$ we have the **Poincaré inequality**[10]*

$$\int_0^1 u(x)^2 dx \le C \int_0^1 u'(x)^2 dx$$

for some $C > 0$. (Hint: Insert $u(x) = \int_0^x u'(y)dy$ one the left. This gives the inequality with $C = \frac{1}{2}$. Using Fourier series one can show that the optimal constant is $C = \frac{1}{\pi^2}$, which is the reciprocal of the square of the lowest eigenvalue. In fact, the minimum is attained for the corresponding eigenfunction $\sin(\pi x)$ — see Problem 5.28.)

[10] Henri Poincaré (1854–1912), French mathematician, theoretical physicist, engineer, and philosopher of science

Problem* 3.3. *Show that the heat kernel satisfies*

$$K(t+s,x,y) = \int_0^1 K(t,x,r)K(s,r,y)dr, \qquad t,s>0.$$

Conclude that $K(t,x,y)>0$ for $x,y \in (0,1)$ since K is analytic with respect to x and hence cannot vanish on an interval.

Problem 3.4. *Show uniqueness for the heat equation with Robin boundary conditions provided $a_0, a_1 \geq 0$.*

Problem 3.5. *Solve the heat equation with Neumann boundary conditions $u_x(t,0) = u_x(t,1) = 0$. Show that the solution converges to the average temperature at an exponential rate. Show that the solution is unique.*

Problem 3.6. *Solve the heat equation with inhomogeneous mixed Dirichlet/Neumann boundary conditions $u(t,0)=0$, $u_x(t,1)=1$.*

Problem 3.7. *Consider the heat equation with mixed Dirichlet/Neumann boundary conditions $u_x(t,0) = u(t,1) = 0$. Derive a maximum principle:* $\min g \leq u(t,x) \leq \max g$.

Problem 3.8. *Find transformations which reduce*

- $u_t = u_{xx} + cu$ *(**cable equation**)*
- $u_t = u_{xx} - au_x$ *(**convection-diffusion equation**)*

to the heat equation. (Hint: For the first multiply u by a suitable function. For the second equation switch to a moving frame $y = x - at$.)

Problem 3.9. *Compute the heat kernel for the case of periodic boundary conditions on the interval* $[-1,1]$.

Problem* 3.10. *Show that the heat equation with (vanishing initial conditions and) inhomogeneous Dirichlet boundary conditions*

$$u(t,0) = a_0(t),\ u(t,1) = a_1(t)$$

can be reduced to an inhomogeneous equation with (homogenous) Dirichlet boundary conditions if $a_0, a_1 \in C^1$. Use this connection to derive a formula for the solution:

$$u(t,x) = -\int_0^t a_0(s)K_y(t-s,x,0)ds + \int_0^t a_1(s)K_y(t-s,x,1)ds.$$

(Hint: Integration by parts; as many times as possible.)

Problem* 3.11. *Let $u \in C(\overline{U_T}) \cap C^{1;2}(U_T)$ solve*

$$u_t = u_{xx} + f, \qquad \begin{cases} u(0,x) = g(x), & x \in (0,1), \\ u(t,0) = a_0(t),\ u(t,1) = a_1(t), & t \in [0,T]. \end{cases}$$

Show

$$|u| \le \max_{[0,1]} |g| + \max_{[0,T]} |a_0| + \max_{[0,T]} |a_1| + T \max_{[0,T]\times[0,1]} |f|.$$

(Hint: Apply the maximum principle to $v := u - tF$, where F is a suitably chosen constant.)

3.2. Outlook: The reaction diffusion equation

The following model is known as the **reaction-diffusion equation**

$$\begin{aligned} &u_t(t,x) - u_{xx}(t,x) + q(x)u(t,x) = 0, \\ &u(0,x) = g(x), \\ &u(t,a) = u(t,b) = 0. \end{aligned} \tag{3.44}$$

Here $u(t,x)$ could be the density of some gas in a pipe and $q(x) \ge 0$ describes that a certain amount per time is removed (e.g., by a chemical reaction). The case $q = 0$ is of course the diffusion equation (aka heat equation) from the previous section.

Applying separation of variables leads to the investigation of the following problem

$$Ly(x) = \lambda y(x), \qquad L := -\frac{d^2}{dx^2} + q(x), \qquad x \in (a,b), \tag{3.45}$$

subject to the **boundary conditions**

$$\cos(\alpha)y(a) = \sin(\alpha)y'(a), \quad \cos(\beta)y(b) = \sin(\beta)y'(b), \tag{3.46}$$

$\alpha, \beta \in \mathbb{R}$. Note that it is crucial that the above boundary conditions are homogenous, such that the superposition principle holds. Inhomogeneous boundary conditions have to be reduced to homogenous ones by a suitable transformation (resulting in an inhomogeneous equation with homogenous boundary conditions).

Of course there is nothing much we can do at this point unless we make some specific choice for q. However, rather than treating some special cases for q, let us reflect about this problem in somewhat more generality.

First of all, you might notice that the problem reassembles finding the eigenvalues of a matrix L. However, in our case L is not a matrix but a linear differential operator. Nevertheless, if (3.45) has a nontrivial solution u satisfying the boundary conditions (3.46), then $\lambda \in \mathbb{C}$ will be called an **eigenvalue** and u will be the corresponding **eigenfunction**. We allow eigenvalues to be complex since it is well-known from linear algebra that the proper setting for eigenvalue problems is over $\mathbb{C}$ rather than over $\mathbb{R}$.

Such a problem is called a **Sturm–Liouville boundary value problem**[11] and the case $q = 0$ suggest that the following facts should be true for this more general problem:

(i) The Sturm–Liouville problem has a countable number of eigenvalues E_n with corresponding eigenfunctions u_n, that is, u_n satisfies the boundary conditions and $Lu_n = E_n u_n$.

(ii) The normalized eigenfunctions u_n form an orthonormal basis, that is, any *nice* function g can be expanded into a generalized Fourier series

$$g(x) = \sum_{n=1}^{\infty} g_n u_n(x).$$

Here the underlying scalar product is

$$\langle f, g\rangle := \int_a^b f^*(x) g(x) dx, \tag{3.47}$$

with '$*$' denoting complex conjugation. The eigenfunctions are orthogonal (i.e., $\langle u_m, u_n\rangle = 0$ for $m \neq n$) and (assuming they are also normalized, $\langle u_n, u_n\rangle = 1$) the Fourier coefficients are given by

$$g_n = \langle u_n, g\rangle. \tag{3.48}$$

Now this problem is very similar to the eigenvalue problem of a symmetric matrix and we are looking for a generalization of the well-known fact that every symmetric matrix has an orthonormal basis of eigenvectors. But since our linear operator L is acting on some space of functions which is not finite dimensional we are in the realm of functional analysis and we will not pursue this problem here. Indeed, while orthogonality of the eigenfunctions is easy to see (Problem 3.12), completeness is much harder. We refer to [**33**, Chapter 5] for further details. We will also establish this result later in Theorem 10.13.

Of course, once we have solved the eigenvalue problem associated with L, the solution of the reaction-diffusion equation is

$$u(t, x) = \sum_{n=1}^{\infty} g_n \mathrm{e}^{-E_n t} u_n(x) \tag{3.49}$$

provided g is such that the Fourier coefficients g_n decay sufficiently fast. The corresponding energy is

$$E(t) := \frac{1}{2}\int_a^b u(t, x)^2 dx = \sum_{n=1}^{\infty} |g_n|^2 \mathrm{e}^{-2E_n t}, \tag{3.50}$$

[11] Jacques Charles François Sturm (1803–1855), French mathematician

where the last equality is owed to the fact that the eigenfunctions form an orthonormal basis. The energy will decay provided all eigenvalues are negative and the rate is given by the lowest eigenvalue. Moreover, the fast decay of $\mathrm{e}^{-E_n t}$ will also imply that we can differentiate the sum (3.49) termwise and hence the solution will be as smooth as the eigenfunctions u_n, which will have two more derivatives than q.

One can also write down a corresponding heat kernel

$$K(t,x,y) = \sum_{n=1}^{\infty} \mathrm{e}^{-E_n t} u_n(x) u_n(y) \tag{3.51}$$

and solve the inhomogeneous equation via Duhamel's principle. Hence the solution of the Cauchy problem

$$u_t + Lu = 0, \qquad u(0,x) = g(x), \tag{3.52}$$

subject to the boundary conditions

$$\cos(\alpha)u(t,a) = \sin(\alpha)u_x(t,a), \quad \cos(\beta)u(t,b) = \sin(\beta)u_x(t,b) \tag{3.53}$$

is given by

$$u(t,x) = \int_a^b K(t,x,y) g(y) dy. \tag{3.54}$$

Note that the eigenvalues of the solution operator are precisely given by $\mathrm{e}^{-E_n t}$ and the solution can formally be written as

$$u(t) = \mathrm{e}^{-tL} g. \tag{3.55}$$

In the case of a system of ordinary differential equations L would be a matrix and e^{-tL} would be defined as the matrix exponential. In our infinite dimensional setting this is more tricky and this is the starting point of the theory of **strongly continuous semigroups**. As this requires tools from functional analysis we will not pursue it here (we will return to this in Chapter 11).

We conclude with the observation that the maximum principle continues to hold, when formulated suitably:

Theorem 3.9. *Let $u \in C(\overline{U_T}) \cap C^{1;2}(U_T)$ and suppose $q \geq 0$. If $u_t + Lu \leq 0$ then*

$$\max_{\overline{U_T}} u \leq \max_{\Gamma_T} u^+. \tag{3.56}$$

If $u_t + Lu \geq 0$ then

$$\min_{\overline{U_T}} u \geq -\max_{\Gamma_T} u^-. \tag{3.57}$$

Here $u^{\pm} = \max(\pm u, 0)$ is the positive, negative part of u, respectively.

Proof. We use a similar strategy as in the proof of Theorem 3.4. Let $v(t,x) := u(t,x) - t\varepsilon$ such that $v_t + Lv = -\varepsilon - \varepsilon q(x)t \le -\varepsilon$. Then v cannot attain an interior maximum at a point x_0 with $v(x_0) \ge 0$. Consequently $v \le \max_{\Gamma_T} v^+ \le \max_{\Gamma_T} u^+$. Hence the first claim follows. For the second claim replace u by $-u$. □

If $u_t + Lu = 0$ we can combine both estimates to obtain again the a priori bound (3.32). In fact, we even have the stronger inequality

$$-\max_{\Gamma_T} u^- \le u(t,x) \le \max_{\Gamma_T} u^+, \tag{3.58}$$

which shows for example $0 \le u \le 1$ if this inequality holds for the initial and boundary data.

Problem* 3.12. *Show that for twice differentiable functions f, g satisfying the boundary conditions* (3.46) *we have*

$$\langle f, Lg \rangle = \langle Lf, g \rangle.$$

Use this to show that all eigenvalues E_n are real and eigenvectors corresponding to different eigenvalues are orthogonal.

Problem 3.13. *Assume that the eigenfunctions u_n are bounded and $E_n \to \infty$ sufficiently fast. Integrate* (3.51) *to obtain the* ***trace formula***

$$\int_a^b K(t,x,x)dx = \sum_{n=1}^\infty \mathrm{e}^{-E_n t}.$$

Use this to show that the Jacobi theta function satisfies

$$\int_0^1 \vartheta(x,\tau)dx = 1.$$

Problem 3.14. *Show directly that the energy is decreasing provided $q \ge 0$.*

Problem 3.15. *Show that Theorem 3.9 holds for more general operators of the form*

$$L = -r(x)\frac{d^2}{dx^2} + p(x)\frac{d}{dx} + q(x)$$

provided $r(x) > 0$ and $q(x) \ge 0$.

3.3. The wave equation for a string

The vibrations of an elastic string can be described by its displacement $u(t,x)$ at the point x and time t. Looking at a tiny element of the string, Newton's law of motion implies that mass times acceleration equals the force acting on this particle. The string is in its equilibrium when it is straight. Once it is bent, the elastic forces will try to bring it back to its equilibrium position. Hence a natural assumption is that this force is proportional to the amount

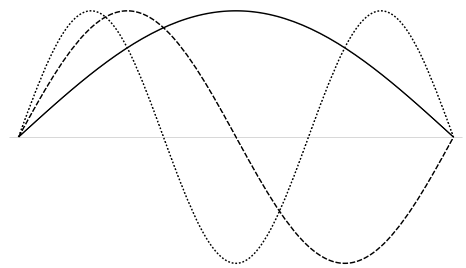


Figure 3.3. Fundamental tone and two overtones of a string

of bending and hence to the second spatial derivative of u. This leads to the one-dimensional **wave equation** $c^{-2}u_{tt}(t,x) = u_{xx}(t,x)$, where $c > 0$ is the propagation speed of waves in our string. By scaling the time variable we can assume $c = 1$ without loss of generality. Indeed, if u solves the wave equation with $c = 1$, then $v(t,x) := u(ct,x)$ solves the wave equation with general $c > 0$. Hence we will consider

$$u_{tt} = u_{xx}. \tag{3.59}$$

Moreover, we will assume that the string is fixed at both endpoints, that is, $x \in [0,1]$ and $u(t,0) = u(t,1) = 0$, and that the initial displacement $u(0,x) = g(x)$ and the initial velocity $u_t(0,x) = h(x)$ are given.

As before, the separation of constants ansatz

$$u(t,x) = w(t)y(x) \tag{3.60}$$

leads to

$$\frac{\ddot{w}(t)}{w(t)} = \frac{y''(x)}{y(x)} \equiv -\lambda, \tag{3.61}$$

which can be solved as in the previous section. In summary, we obtain the solutions

$$u(t,x) = (c_1 \cos(n\pi t) + c_2 \sin(n\pi t)) \sin(n\pi x), \qquad n \in \mathbb{N}. \tag{3.62}$$

In particular, the string can vibrate only with certain fixed frequencies! See Figure 3.3, where the fundamental tone ($n = 1$) and two overtones ($n = 2,3$) are shown. If you pluck the string in the middle, you will never get only the fundamental tone but you will always excite some overtones as well. The combinations of these overtones make up the characteristic sound of the instrument. If you want to change the fundamental tone (frequency), you need to change the length of the string. If you want to play different tones you either need different lengths for each tone (piano) or you change the length on the fly by fixing a point inside (e.g. with your fingers — guitar).

Taking linear combinations we get as in the case of the heat equation (Problem 3.17):

Lemma 3.10. *Suppose* $c_{1,n}$ *and* $c_{2,n}$ *are sequences satisfying*

$$\sum_{n=1}^{\infty} n^2 |c_{1,n}| < \infty, \qquad \sum_{n=1}^{\infty} n^2 |c_{2,n}| < \infty. \tag{3.63}$$

Then

$$u(t,x) = \sum_{n=1}^{\infty} (c_{1,n} \cos(n\pi t) + c_{2,n} \sin(n\pi t)) \sin(n\pi x) \tag{3.64}$$

is in $C^2(\mathbb{R} \times [0,1])$ *and satisfies the wave equation* (3.59) *as well as the boundary conditions* $u(t,0) = u(t,1) = 0$.

Next, under the assumptions (3.63), the proof of the previous lemma also shows

$$u(0,x) = \sum_{n=1}^{\infty} c_{1,n} \sin(n\pi x), \qquad u_t(0,x) = \sum_{n=1}^{\infty} n\pi c_{2,n} \sin(n\pi x). \tag{3.65}$$

Now observe that the sums on the right-hand side are again Fourier sine series. Hence expanding the initial conditions into Fourier sine series

$$g(x) = \sum_{n=1}^{\infty} g_n \sin(n\pi x), \qquad h(x) = \sum_{n=1}^{\infty} h_n \sin(n\pi x), \tag{3.66}$$

where

$$g_n = 2\int_0^1 \sin(n\pi x) g(x) dx, \quad h_n = 2\int_0^1 \sin(n\pi x) h(x) dx, \tag{3.67}$$

we see that the solution of our original problem is given by (3.64) with $c_{1,n} = g_n$ and $c_{2,n} = \frac{h_n}{n\pi}$, provided the Fourier coefficients satisfy

$$\sum_{n=1}^{\infty} n^2 |g_n| < \infty, \qquad \sum_{n=1}^{\infty} n |h_n| < \infty. \tag{3.68}$$

It can be shown that this last condition holds if $g \in C^3[0,1]$ with $g(0) = g''(0) = g(1) = g''(1) = 0$ and $h \in C^2[0,1]$ with $h(0) = h(1) = 0$ (see Appendix A.3).

However, note that this time it is not easily possible to write down a corresponding integral kernel since the corresponding sums

$$2\sum_{n=1}^{\infty} \cos(n\pi t) \sin(n\pi x) \sin(n\pi y), \quad 2\sum_{n=1}^{\infty} \frac{\sin(n\pi t)}{n\pi} \sin(n\pi x) \sin(n\pi y) \tag{3.69}$$

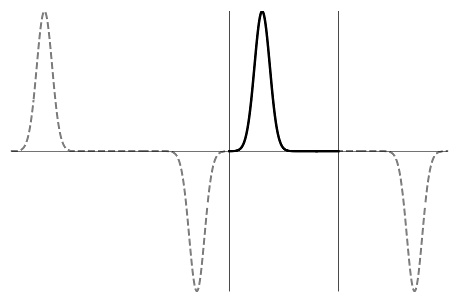


Figure 3.4. The odd periodic extension of a bump

do not converge properly. Nevertheless we can use some trigonometric identities to obtain

$$\begin{aligned} u(t,x) &= \sum_{n=1}^{\infty} \left(g_n \cos(n\pi t) + \frac{h_n}{n\pi} \sin(n\pi t) \right) \sin(n\pi x) \\ &= \frac{1}{2} \sum_{n=1}^{\infty} g_n \big(\sin(n\pi(x+t)) + \sin(n\pi(x-t)) \big) \\ &\qquad - \frac{1}{2} \sum_{n=1}^{\infty} \frac{h_n}{n\pi} \big(\cos(n\pi(x+t)) - \cos(n\pi(x-t)) \big) \\ &= \frac{g(x+t)+g(x-t)}{2} + \frac{1}{2} \int_{x-t}^{x+t} h(y) dy. \end{aligned} \tag{3.70}$$

In fact, for the last equality to hold we have extended both g and h from $[0,1]$ to all of $\mathbb{R}$ using (3.66). That is, we use an odd periodic extension.

Hence we have again obtained **d'Alembert's formula**

$$u(t,x) = \frac{g(x+t)+g(x-t)}{2} + \frac{1}{2} \int_{x-t}^{x+t} h(y) dy. \tag{3.71}$$

Since this formula provides a solution for $g \in C^2(\mathbb{R})$, $h \in C^1(\mathbb{R})$, there is no need to expand the initial data into Fourier sine series and we can use it as our starting point.

To obtain a solution on $x \in [0,1]$ satisfying the boundary conditions $u(t,0) = u(t,1) = 0$ our analysis suggests to use the following reflection technique: Extend the initial condition $g(x) \in C^2[0,1]$ to $[-1,1]$ using reflection $g(-x) = -g(x)$ and then to $\mathbb{R}$ using periodicity $g(x+2) = g(x)$ (see Figure 3.4). Hence by construction this odd extension satisfies $g(-x) = -g(x)$ and $g(x+2) = g(x)$ for all $x \in \mathbb{R}$. This is of course precisely what you get if one regards the Fourier sine expansion (3.66) as a function on $\mathbb{R}$.

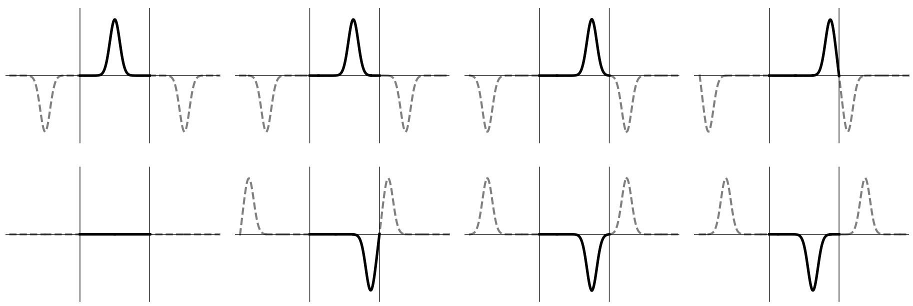


Figure 3.5. A traveling bump is being reflected at the ends

Since a continuous odd function must vanish at 0, the odd periodic extension of g will be twice differentiable at 0 if and only if $g(0) = g''(0) = 0$. Similarly, periodicity requires $-g(1) = g(-1) = g(1)$ implying $g(1) = 0$. Applying the same argument to the derivatives shows $g''(1) = 0$. In summary, the odd extension of g will be $C^2(\mathbb{R})$ provided $g(0) = g''(0) = g(1) = g''(1) = 0$. Similarly we can extend $h \in C^1[0,1]$ to an odd periodic function $h \in C^1(\mathbb{R})$ provided $h(0) = h(1) = 0$. Consequently, d'Alembert's formula gives a solution $u(t,x) \in C^2(\mathbb{R})$ with $u(t,-x) = -u(t,x)$ and $u(t,x+2) = u(t,x)$. By construction $u(t,0) = u(t,1) = 0$. Note that the solution is periodic in time with period $T = 2$:

$$u(t+T,x) = u(t,x), \qquad T = 2. \tag{3.72}$$

Hence we have the following picture: If we think of $u(t,x)$ as a little bump which travels to the right, say, then there is a corresponding reflected negative bump in $[1,2]$ traveling to the left. Eventually they will meet at the boundary $x = 1$ and the reflected bump will enter our interval, while the original one will leave. Hence our little bump now has its sign changed and travels to the left until it hits its symmetric image from the interval $[-2,-1]$ at the other side $x = 0$, where the same happens again. See Figure 3.5.

For a bump starting at x_0 its path is shown on the left in Figure 3.6. The two lines emanating from x_0 correspond to the two cases when the bump travels to the left, right, respectively. More precisely, if we set $h = 0$, then the value $g(x_0)$ will be transported along these lines. Of course we can also trace backwards and find the two points which contribute to the value $u(t,x)$; right in Figure 3.6. For the contribution from h the situation is similar, but now the values of h within the cone formed by the two lines will contribute.

In summary, we see that the effect of a small perturbation in a neighborhood of x_0 can only propagate with a finite speed $c = 1$ to the left and to the right. Hence in contradistinction to the heat equation, we have a **finite**

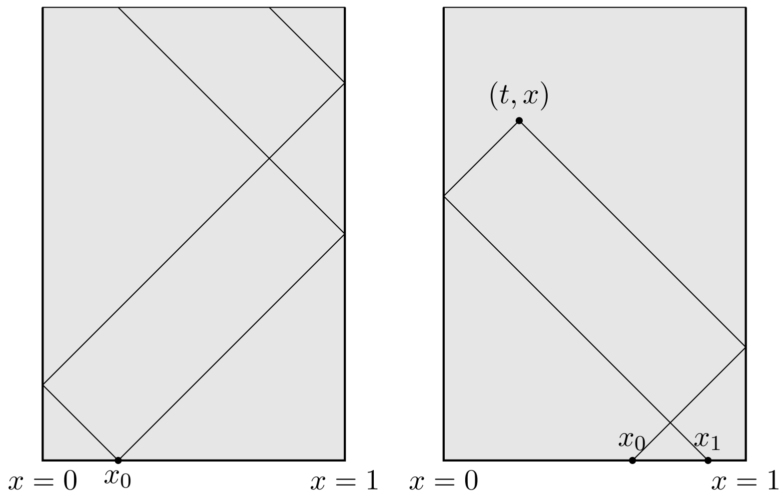


Figure 3.6. Domain of influence for the wave equation

propagation speed $c = 1$. Also note that the wave equation does not improve the smoothness of the initial condition. On the other hand, solutions are defined for both positive and negative times.

To see uniqueness we introduce again a suitable **energy** functional

$$E(t) = \frac{1}{2}\int_0^1 \big(u_t(t,x)^2 + u_x(t,x)^2\big)dx. \tag{3.73}$$

It turns out that the energy is conserved,

$$\begin{aligned}\frac{d}{dt}E(t) &= \int_0^1 \big(u_t(t,x)u_{tt}(t,x) + u_x(t,x)u_{tx}(t,x)\big)dx\\ &= \int_0^1 \big(u_t(t,x)u_{xx}(t,x) + u_x(t,x)u_{tx}(t,x)\big)dx\\ &= \int_0^1 \big(u_t(t,x)u_{xx}(t,x) - u_{xx}(t,x)u_t(t,x)\big)dx = 0,\end{aligned} \tag{3.74}$$

where we have used integration by parts together with the Dirichlet boundary conditions, which imply $u_t(t,0) = u_t(t,1) = 0$. Hence any solution with vanishing initial conditions $g = h = 0$ has zero energy. But this implies $u_x = u_t = 0$ and thus u must be constant. Finally, since it vanishes at $t = 0$, it must be zero.

At this point you might ask why we have not started with d'Alembert's formula in the first place? One reason is that while d'Alembert's formula tells you what you see when you look at the wave, the series solution tells you what you hear when you listen to it. Moreover, the series solution is more robust in the sense that it also applies to the somewhat more general problem

$$\begin{aligned}&u_{tt}(t,x) - u_{xx}(t,x) + q(x)u(t,x) = 0,\\&u(0,x) = g(x), \qquad u_t(0,x) = h(x),\\&u(t,0) = u(t,1) = 0.\end{aligned} \tag{3.75}$$

This problem be handled as in Section 3.2 which leads to the solution

$$u(t,x) = \sum_{n=1}^{\infty} \left(g_n \cos(\sqrt{E_n} t) + h_n \frac{\sin(\sqrt{E_n} t)}{\sqrt{E_n}} \right) u_n(x) \tag{3.76}$$

provided g and h are such that the coefficients g_n and h_n decay sufficiently fast. Also the energy argument extends to this situation – Problem 3.16.

Finally, the inhomogeneous problem

$$u_{tt} = u_{xx} + f(t,x), \qquad \begin{cases} u(t,0) = u(t,1) = 0, \\ u(0,x) = g(x), u_t(0,x) = h(x), \end{cases} \tag{3.77}$$

can also be handled by an ansatz as a Fourier series

$$u(t,x) = \sum_{n=1}^{\infty} a_n(t) \sin(n\pi x), \qquad f(t,x) = \sum_{n=1}^{\infty} f_n(t) \sin(n\pi x) \tag{3.78}$$

which gives

$$\ddot{a}_n(t) = -(\pi n)^2 a_n(t) + f_n(t). \tag{3.79}$$

The solution of this ordinary differential equation is given by

$$a_n(t) = g_n \cos(n\pi t) + h_n \frac{\sin(n\pi t)}{n\pi} + \int_0^t \frac{\sin(n\pi(t-s))}{n\pi} f_n(s) ds \tag{3.80}$$

and we see that this approach indeed provides a solution if we assume

$$|f_n(t)| \le M_n, \qquad \sum_{n=1}^{\infty} n\, M_n < \infty. \tag{3.81}$$

However, we can also use trigonometric identities as before to show that the solution is given by the **Duhamel principle** and is given by

$$\begin{aligned}u(t,x) = {}& \frac{g(x+t) + g(x-t)}{2} + \frac{1}{2} \int_{x-t}^{x+t} h(y) dy\\&+ \frac{1}{2} \int_0^t \int_{x-(t-s)}^{x+(t-s)} f(s,y) dy\, ds,\end{aligned} \tag{3.82}$$

where again on has to take the odd periodic extension of f with respect to x. This will be a solution provided $f \in C^{0;1}(\mathbb{R} \times [0,1])$ with $f(s,0) = f(t,1) = 0$ (Problem 4.23). Note that since we have a second order equation, when choosing initial conditions for the Duhamel principle, we have to choose f as h and g to be zero.

Problem* 3.16. *Show that* (3.75) *preserves the energy*

$$E(t) = \frac{1}{2}\int_0^1 \big(u_t(t,x)^2 + u_x(t,x)^2 + q(x)u(t,x)^2\big)dx.$$

Conclude that solutions are unique if $q(x) \ge 0$.

Problem* 3.17. *Prove Lemma 3.10.*

Problem 3.18. *Solve the* ***telegraph equation***

$$u_{tt}(t,x) + 2\eta u_t + \gamma u = c^2 u_{xx}, \qquad 0 \le \eta < \sqrt{(c\pi)^2 + \gamma},\ 0 \le \gamma,$$

with Dirichlet boundary conditions on $[0,1]$. *Show that the solutions converge to* 0. *Show that the energy is non-increasing and conclude that solutions are unique.*

Problem 3.19. *Explain how d'Alembert's formula can be used to obtain solutions which satisfy Neumann boundary conditions* $u_x(t,0) = u_x(t,1) = 0$. *Discuss what happens to a small bump traveling to the right.*

Problem 3.20. *Consider the* ***beam equation***

$$u_{tt} = -u_{xxxx}$$

with boundary conditions

$$u(t,0) = u_x(t,0) = 0, \qquad u_{xx}(t,1) = u_{xxx}(t,1) = 0$$

corresponding to a cantilevered beam on $[0,1]$. *Find the eigenfrequencies and the corresponding solutions of the beam. (Hint: Show that a corresponding eigenfunction* $-u'''' = Eu$ *satisfying the boundary conditions must satisfy* $\int_0^1 (u'')^2 dx = -E\int_0^1 u^2 dx$ *and hence* $E < 0$. *It is not possible to obtain analytic expressions for the eigenfrequencies.)*

3.4. The wave equation on a rectangle and on a disc

Consider the vibrations of a rectangular membrane which is fixed at the boundary. The motion is described by the two dimensional wave equation

$$u_{tt}(t,x_1,x_2) = u_{x_1x_1}(t,x_1,x_2) + u_{x_2x_2}(t,x_1,x_2) \tag{3.83}$$

together with the Dirichlet boundary conditions

$$u(t,0,x_2) = u(t,1,x_2) = u(t,x_1,0) = u(t,x_1,1) = 0, \quad x_j \in [0,1]. \tag{3.84}$$

Looking for solutions of the form $u(t,x_1,x_2) = w(t)u(x_1,x_2)$ yields

$$w(t) = c_1\cos(\sqrt{\lambda}t) + c_2\sin(\sqrt{\lambda}t) \tag{3.85}$$

and

$$u_{x_1x_1}(x_1,x_2) + u_{x_2x_2}(x_1,x_2) = -\lambda u(x_1,x_2). \tag{3.86}$$

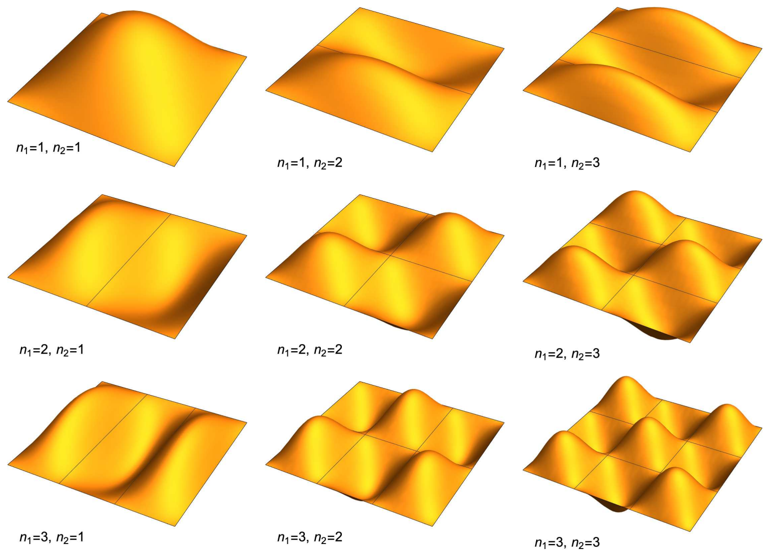


Figure 3.7. Fundamental tone and some overtones of a rectangular membrane

Our simple geometry suggests another ansatz $u(x_1, x_2) = u_1(x_1)u_2(x_2)$ which leads to identical equations

$$u_j'' = -\lambda_j u_j, \qquad u_j(0) = u_j(1) = 0, \tag{3.87}$$

where $\lambda_1 + \lambda_2 = \lambda$. Since we already know the solution, we get

$$u(t, x_1, x_2) = \sum_{n_1,n_2=1}^{\infty} \left(c_{1,n_1,n_2} \cos(\sqrt{n_1^2 + n_2^2}\pi t) + c_{2,n_1,n_2} \sin(\sqrt{n_1^2 + n_2^2}\pi t) \right) \cdot \cdot \sin(n_1 \pi x_1) \sin(n_2 \pi x_2). \tag{3.88}$$

Initial conditions can be handled by expanding them into two-dimensional Fourier series much like we did in the one-dimensional case. Hence we omit further details at this point.

From a physical point of view it is interesting, that our membrane can only vibrate with certain fixed frequencies $\frac{1}{2}\sqrt{n_1^2 + n_2^2}$, $(n_1, n_2) \in \mathbb{N}^2$. The same turns out to be true for a differently shaped membrane (where separation of variables fails and the frequencies can no longer be computed explicitly in general) and these frequencies are characteristic for the sound of the membrane. This lead the mathematician Mark Kac[12] to the famous question: "Can One Hear the Shape of a Drum?" Mathematically, the question is whether the eigenfrequencies determine the shape and it took two decades

[12] Marc Kac (1914–1984), Polish American mathematician

until this question was finally answered in the negative. It is however true for convex domains with analytic boundary.

If we consider the same problem on the unit disc, the first step will work as before, but the second will fail since the ansatz $u(x_1,x_2) = u_1(x_1)u_2(x_2)$ is not compatible with the geometry of our domain. However, it suggests itself to switch to polar coordinates

$$x_1 = r\cos(\varphi), \qquad x_2 = r\sin(\varphi) \tag{3.89}$$

such that our differential equation for

$$v(r,\varphi) := u(r\cos(\varphi), r\sin(\varphi)) \tag{3.90}$$

reads

$$v_{rr} + \frac{1}{r}v_r + \frac{1}{r^2}v_{\varphi\varphi} = -\lambda v. \tag{3.91}$$

Since in polar coordinates our domain is again rectangular, we can try separation of variables $v(r,\varphi) = \rho(r)\theta(\varphi)$. Our boundary condition translates into $\rho(1) = 0$ and θ must be periodic with period 2π. The equations read

$$\rho'' + \frac{1}{r}\rho' - \frac{\eta}{r^2}\rho = -\lambda\rho, \qquad \theta'' = -\eta\theta, \tag{3.92}$$

where η is another constant. The second equation can be easily solved

$$\theta(\varphi) = a_n\cos(n\varphi) + b_n\sin(n\varphi), \tag{3.93}$$

where $\eta = n^2$ and $n \in \mathbb{N}_0$ to satisfy our periodicity requirement. The first equation is a version of Bessel's equation[13] and a solution is (Problem 3.21)

$$\rho(r) = J_n(\sqrt{\lambda}r), \tag{3.94}$$

where

$$J_\nu(z) := \sum_{j=0}^\infty \frac{(-1)^j}{j!\Gamma(\nu+j+1)}\left(\frac{z}{2}\right)^{2j+\nu} \tag{3.95}$$

is the **Bessel function** of order ν, [**25**, (10.2.2)]. There is a second linearly independent solution (the Weber[14] function $Y_\nu(z)$, [**25**, (10.2.3)]) but it can be ruled out for our purpose since it has a singularity at $z = 0$ and hence will not lead to a continuous solution of our original problem. See for example [**33**] for how to solve Bessel's equation. Our boundary condition requires $\rho(1) = J_n(\sqrt{\lambda}) = 0$ and hence we need $\lambda = j_{n,k}^2$, where $j_{n,k}$ is the k'th positive zero of J_n. It can be shown that ([**25**, (10.17.3)], Problem 4.32)

$$J_\nu(z) = \sqrt{\frac{2}{\pi z}}\left(\cos\left(z - \frac{\pi}{2}\nu - \frac{\pi}{4}\right) + O(z^{-1})\right) \tag{3.96}$$

as $z \to \infty$ and hence there are infinitely many zeros for each n. Note that for $\lambda < 0$ the solution would be $\mathrm{i}^{-n}J_n(\mathrm{i}\sqrt{|\lambda|}r)$ (which is real-valued) but since

[13]Friedrich Bessel (1784–1846), German astronomer, mathematician, physicist, and geodesist

[14]Heinrich Friedrich Weber (1843–1912), German physicist

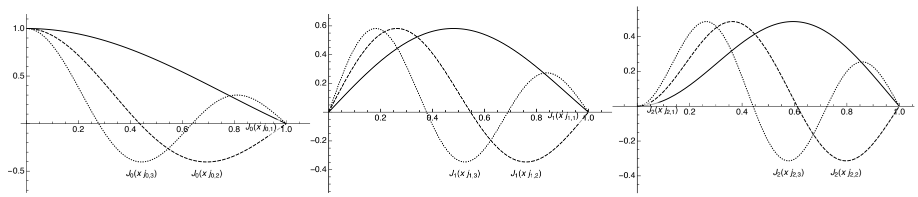


Figure 3.8. Bessel–Fourier modes $J_0(j_{0,k}x)$ (left), $J_1(j_{1,k}x)$ (middle), and $J_2(j_{2,k}x)$ (right) for $k = 1, 2, 3$

all coefficients of the power series are positive in this case, this solution has no zeros for $r > 0$. It can also be shown that there is a corresponding **Fourier–Bessel series**. The first few modes are shown in Figure 3.8. It is not hard to see that these functions are orthogonal (Problem 3.23) but completeness is a more involved task which requires advanced tools from functional analysis and is beyond our scope. In this respect observe that the Bessel equation on $[0, 1]$ can be written in the form $Ly = \lambda y$ with $L = -\frac{d^2}{dx^2} + q(x)$ upon using $y(x) = \sqrt{x}\rho(x)$. However in this case $q(x) = \frac{\nu^2 - \frac{1}{4}}{x^2}$ has a singularity at $x = 0$ and hence the standard results for regular Sturm–Liouville problems do not apply.

The fundamental tones and overtones of the disc are hence given by

$$u^{\mathrm{c}}_{n,k}(x_1, x_2) = J_n(j_{n,k}r)\cos(n\varphi), \quad u^{\mathrm{s}}_{n,k}(x_1, x_2) = J_n(j_{n,k}r)\sin(n\varphi) \quad (3.97)$$

and a few instances of $u^{\mathrm{c}}_{n,k}$ are shown in Figure 3.9 (clearly the corresponding picture for $u^{\mathrm{s}}_{n,k}$ is just rotated by $-\frac{\pi}{2n}$).

Problem* 3.21. *Show that the Bessel function* (3.95) *solves the* ***Bessel equation***

$$z^2 u'' + z u' + (z^2 - \nu^2) u = 0.$$

Show that ***Lommel's equation***[15]

$$w'' + \frac{1 - 2a}{z} w' + \left((bcz^{c-1})^2 + \frac{a^2 - \nu^2 c^2}{z^2} \right) w = 0$$

can be transformed to the Bessel equation via $w(z) = z^a u(bz^c)$.

Problem 3.22. *Prove the following properties of the Bessel functions* (3.95).

(i) $(z^{\pm\nu} J_\nu(z))' = \pm z^{\pm\nu} J_{\nu\mp 1}(z)$.

(ii) $J_{\nu-1}(z) + J_{\nu+1}(z) = \frac{2\nu}{z} J_\nu(z)$.

(iii) $J_{\nu-1}(z) - J_{\nu+1}(z) = 2J'_\nu(z)$.

Problem* 3.23. *Let* ρ_1, ρ_2 *be two solutions of*

$$\rho'' + \frac{1}{r}\rho' - \frac{\nu^2}{r^2}\rho = -\lambda\rho, \qquad \nu \geq 0,$$

[15] Eugen von Lommel (1837–1899), German physicist

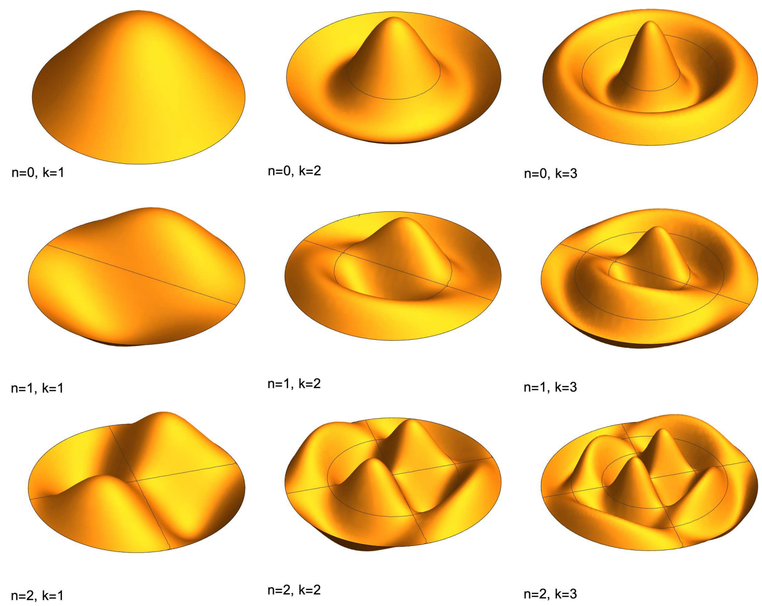


Figure 3.9. Fundamental tone and some overtones of a disc shaped membrane

corresponding to λ_1, λ_2, *respectively. Show that*

$$\frac{d}{dr} r\big(\rho_1(r)\rho_2'(r) - \rho_1'(r)\rho_2(r)\big) = (\lambda_1 - \lambda_2) r\, \rho_1(r)\rho_2(r).$$

Conclude

$$\int_0^1 J_\nu(j_{\nu,k}r) J_\nu(j_{\nu,l}r) r\, dr = \begin{cases} \frac{1}{2} J_\nu'(j_{\nu,k})^2, & l = k, \\ 0, & l \neq k. \end{cases}$$

Note that $J_\nu'(j_{\nu,k}) \neq 0$ *since if for a solution of a second order linear equation both the function and its derivative would vanish, it would be the zero solution.*

Problem* 3.24. *Establish the following integral representation for the Bessel function of integer order*

$$J_n(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathrm{e}^{\mathrm{i}(nt - \sin(t)x)} dt, \qquad n \in \mathbb{N}_0.$$

(Hint: Split the exponential and insert the Taylor series for the one containing x*. Then use the residue theorem to evaluate the resulting integral.)*

Problem 3.25 (Hanging chain)**.** *Consider the vibrations of a chain of length* 1 *suspended at* $x = 1$*. Denote the displacement by* $u(t,x)$*. Then the motion is described by the equation*

$$u_{tt}(t,x) = g \frac{\partial}{\partial x} x \frac{\partial}{\partial x} u(t,x), \quad x \in [0,1],$$

with boundary conditions $u(t, 1) = 0$*, where* $g > 0$ *is a constant. Apply separation of variables to find the eigenvalues and eigenfunctions.*

Problem 3.26. *How would one solve the wave equation on a semidisc* $\{x | |x| < 1, x_2 > 0\}$ *with Dirichlet conditions on the boundary?*

3.5. The Laplace equation on a disc

As our next example we look at the Laplace equation on the unit disc with given values on the boundary:

$$u_{xx} + u_{yy} = 0, \qquad u(\cos(\varphi), \sin(\varphi)) = g(\varphi). \tag{3.98}$$

As in the previous section, separation of variables in polar coordinates $u(x, y) = \rho(r)\theta(\varphi)$ gives

$$\theta(\varphi) = a_n \cos(n\varphi) + b_n \sin(n\varphi), \qquad n \in \mathbb{N}_0, \tag{3.99}$$

and the corresponding equation for $\rho(r)$ reads

$$\rho'' + \frac{1}{r}\rho' - \frac{n^2}{r^2}\rho = 0. \tag{3.100}$$

It is of Euler type and its solution is

$$\rho(r) = \begin{cases} c_1 + c_2 \log(r), & n = 0, \\ c_1 r^n + c_2 r^{-n}, & n > 0. \end{cases} \tag{3.101}$$

Since our solution u must be continuous at the origin we need $c_2 = 0$ and using $v(r, \varphi) := u(r\cos(\varphi), r\sin(\varphi))$ we get

$$v(r, \varphi) = \frac{a_0}{2} + \sum_{n=1}^{\infty} r^n \big(a_n \cos(n\varphi) + b_n \sin(n\varphi)\big). \tag{3.102}$$

Setting $r = 1$ we see that a_n, b_n are the Fourier coefficients of g:

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} g(\varphi) \cos(n\varphi) d\varphi, \qquad b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} g(\varphi) \sin(n\varphi) d\varphi. \tag{3.103}$$

Using complex notation we can rewrite this as

$$v(r, \varphi) = \frac{a_0}{2} + \mathrm{Re} \sum_{n=1}^{\infty} r^n (a_n - \mathrm{i} b_n) \mathrm{e}^{\mathrm{i} n\varphi} = \frac{a_0}{2} + \mathrm{Re} \sum_{n=1}^{\infty} (a_n - \mathrm{i} b_n) z^n, \tag{3.104}$$

where $z = r\mathrm{e}^{\mathrm{i}\varphi} = x + \mathrm{i}y$. Hence the solution is the real part of a holomorphic function inside the unit disc. Conversely, it is a well-known fact from complex analysis, that real (and imaginary) part of holomorphic functions satisfy the Laplace equation.

From this formula we can even get an integral representation for the solution. To this end we use

$$a_n - \mathrm{i}b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} g(\vartheta)\mathrm{e}^{-\mathrm{i}n\vartheta}d\vartheta \tag{3.105}$$

to obtain

$$\begin{aligned} v(r,\varphi) &= \frac{a_0}{2} + \frac{1}{\pi}\mathrm{Re}\sum_{n=1}^{\infty}\int_{-\pi}^{\pi} g(\vartheta)r^n\mathrm{e}^{\mathrm{i}n(\varphi-\vartheta)}d\vartheta \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} P_r(\varphi-\vartheta)g(\vartheta)d\vartheta, \qquad 0 \le r < 1, \end{aligned} \tag{3.106}$$

where

$$P_r(\varphi) := \mathrm{Re}\left(1 + 2\sum_{n=1}^{\infty} r^n\mathrm{e}^{\mathrm{i}n\varphi}\right) = \mathrm{Re}\left(\frac{1+r\mathrm{e}^{\mathrm{i}\varphi}}{1-r\mathrm{e}^{\mathrm{i}\varphi}}\right) = \frac{1-r^2}{1-2r\cos(\varphi)+r^2} \tag{3.107}$$

is the **Poisson kernel**[16] for the unit disc. To obtain the closed form of the series note that this is just a geometric series.

Theorem 3.11. *Suppose $g \in C[0,2\pi]$ is periodic. Then the **Poisson integral***

$$v(r,\varphi) = \frac{1}{2\pi}\int_{-\pi}^{\pi} P_r(\varphi-\vartheta)g(\vartheta)d\vartheta, \qquad 0 \le r < 1, \tag{3.108}$$

is harmonic inside the unit disc and satisfies

$$\lim_{r\uparrow 1} v(r,\varphi) = g(\varphi). \tag{3.109}$$

Proof. That v solves the Laplace equation follows by interchanging differentiation and summation in (3.102) which is permissible for $r \le r_0 < 1$ since the Fourier coefficients of a continuous function are bounded.

Next, it is straightforward to verify the following properties of the Poisson kernel:

- $\frac{1}{2\pi}\int_{-\pi}^{\pi} P_r(\varphi)d\varphi = 1$.
- $P_r(\varphi) = P_r(-\varphi)$.
- $\frac{1-r}{1+r} \le P_r(\varphi) \le \frac{1+r}{1-r}$ with the unique maximum at $\varphi = 0$, the unique minimum at $\varphi = \pm\pi$, and monotone in between.
- $\lim_{r\to 1} P_r(\varphi) = 0$ for $\varphi \ne 0$.

Without loss of generality it suffices to show convergence at $\varphi = 0$. Moreover, given ε, choose $0 < \delta < \pi$ such that $|g(\vartheta) - g(0)| \le \frac{\varepsilon}{2}$ for $|\vartheta| < \delta$ and r_0 such

[16] Siméon Denis Poisson (1781–1840), French mathematician, engineer, and physicist

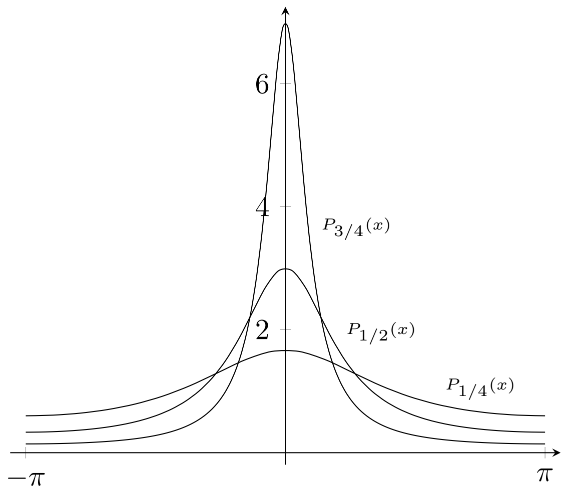


Figure 3.10. The Poisson kernels $P_{1/4}$, $P_{1/2}$, and $P_{3/4}$

that $P_r(\delta) \le \frac{\varepsilon}{4M}$ for $r \ge r_0$, where $M = \max_{|\vartheta|\le\pi} |g(\vartheta)|$. Then

$$v(r,0) - g(0) = \frac{1}{2\pi}\int_{-\pi}^{\pi} P_r(\vartheta)(g(\vartheta) - g(0))d\vartheta$$

and hence

$$\begin{aligned}|v(r,0) - g(0)| &\le \frac{\varepsilon}{4\pi}\int_{|\vartheta|\le\delta} P_r(\vartheta)d\vartheta + \frac{P_r(\delta)}{2\pi}\int_{|\vartheta|\ge\delta} |g(\vartheta) - g(0)|d\vartheta\\ &\le \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon\end{aligned}$$

for $r \ge r_0$. □

Note that the special case $r = 0$ says that the value at the center equals the average at the boundary.

Problem 3.27. *Find the solution of the Laplace equation on the unit disc which satisfies $u(x, y) = x^2$ for $x^2 + y^2 = 1$. Give the solution in Cartesian coordinates.*

Problem 3.28. *Find the solution u of the Laplace equation on a ring $r_0 < \sqrt{x^2 + y^2} < r_1$ (with $0 < r_0 < r_1$) satisfying the boundary conditions $u(x, y) = y$ for $x^2 + y^2 = r_0^2$ and $u(x, y) = x$ for $x^2 + y^2 = r_1^2$. Give the solution in Cartesian coordinates.*

Problem 3.29. *Find the solution of the Laplace equation on the square $U := (0, 1) \times (0, 1)$ satisfying the boundary conditions*

$$u(x, 0) = 0, \quad u(0, y) = 0, \quad u(1, y) = 0, \quad u(x, 1) = g(x).$$

Chapter 4

The Fourier transform and problems on the line

4.1. Motivation

Now we want to look at the case of the heat equation on the line. In contradistinction to a finite interval we have no boundary conditions and hence there are no restrictions on the value of λ. We keep however the restriction $\lambda > 0$, since for $\lambda < 0$ the corresponding solutions are exponentially growing. Setting $\lambda = k^2$, $k \in \mathbb{R}$ and using the complex versions of the trigonometric functions we hence have the solutions

$$u(t,x) = \mathrm{e}^{-\mathrm{i}kx-k^2t}. \tag{4.1}$$

Again we can take linear combinations of these solutions and, since k is continuous, we can even take a *continuous* sum:

$$u(t,x) = \int_{\mathbb{R}} c(k)\mathrm{e}^{-\mathrm{i}kx-k^2t} dk. \tag{4.2}$$

Assuming that c is integrable, one can verify that $u \in C([0,\infty)\times\mathbb{R}) \cap C^\infty((0,\infty)\times\mathbb{R})$ satisfies the heat equation with initial condition

$$u(0,x) = \int_{\mathbb{R}} c(k)\mathrm{e}^{-\mathrm{i}kx} dk. \tag{4.3}$$

Hence $u(0,x)$ is up to a normalization constant equal to the Fourier transform of c. In order to express c in terms of the initial condition, we need to compute the inverse Fourier transform.

4.2. The Fourier transform in one dimension

The **Fourier transform** of an integrable function f is defined as

$$\mathcal{F}(f)(k) \equiv \hat{f}(k) := \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \mathrm{e}^{-\mathrm{i}kx} f(x) dx. \tag{4.4}$$

Note that $\hat{f}$ is continuous by Lemma A.6. The fact which makes it a key tool in the study of partial differential equation is the property that it turns differentiation into multiplication and vice versa:

Lemma 4.1. *Suppose* $f \in C^1(\mathbb{R})$ *such that* $\lim_{|x|\to\infty} f(x) = 0$ *and* f, f' *integrable. Then*

$$(f')^\wedge(k) = \mathrm{i}k\hat{f}(k). \tag{4.5}$$

Similarly, if $f(x), xf(x)$ *are integrable, then* $\hat{f}(k)$ *is differentiable and*

$$(x\, f(x))^\wedge(k) = \mathrm{i}\hat{f}'(k). \tag{4.6}$$

Proof. First of all, by integration by parts, we see

$$\begin{aligned}(f')^\wedge(k) &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \mathrm{e}^{-\mathrm{i}kx} f'(x) dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \mathrm{i}k\, \mathrm{e}^{-\mathrm{i}kx} f(x) dx = \mathrm{i}k\, \hat{f}(k).\end{aligned}$$

Similarly, the second formula follows from

$$\begin{aligned}(x\, f(x))^\wedge(k) &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} x\mathrm{e}^{-\mathrm{i}kx} f(x) dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \left(\mathrm{i}\frac{\partial}{\partial k}\mathrm{e}^{-\mathrm{i}kx}\right) f(x) dx = \mathrm{i}\hat{f}'(k),\end{aligned}$$

where interchanging the derivative and integral is permissible by Lemma A.7. In particular, $\hat{f}(k)$ is differentiable. □

This result immediately extends to higher derivatives. Roughly speaking this shows that the decay of a function is related to the smoothness of its Fourier transform and the smoothness of a function is related to the decay of its Fourier transform.

Another key property is the fact that the Fourier transform of a Gaussian[1] is again a Gaussian:

Lemma 4.2. *The Fourier transform of* $\mathrm{e}^{-tx^2/2}$ *for* $t > 0$ *is given by*

$$\mathcal{F}(\mathrm{e}^{-tx^2/2})(k) = \frac{1}{t^{1/2}}\mathrm{e}^{-k^2/(2t)}. \tag{4.7}$$

[1] Carl Friedrich Gauss (1777–1855), German mathematician and physicist

Proof. Let $\phi_t(x) := \exp(-tx^2/2)$. Then $\phi_t'(x) + tx\phi_t(x) = 0$ and hence $\mathrm{i}(k\hat{\phi}_t(k) + t\hat{\phi}_t'(k)) = 0$. Thus $\hat{\phi}_t(k) = c\phi_{1/t}(k)$ and (Problem A.3)

$$c = \hat{\phi}_t(0) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \mathrm{e}^{-tx^2/2} dx = \frac{1}{\sqrt{t}}. \qquad \square$$

Now we can compute the inverse of the Fourier transform:

Theorem 4.3. *Suppose f is continuous such that $f, \hat{f}$ are integrable. Then*

$$(\hat{f})^\vee = f, \tag{4.8}$$

where

$$\check{f}(k) := \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}kx} f(x) dx = \hat{f}(-k). \tag{4.9}$$

Proof. Abbreviate $\phi_\varepsilon(x) := (2\pi)^{-1/2} \exp(-\varepsilon x^2/2)$. Then

$$\int_{\mathbb{R}} \phi_\varepsilon(k) \mathrm{e}^{\mathrm{i}kx} \hat{f}(k) dk = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \int_{\mathbb{R}} \phi_\varepsilon(k) \mathrm{e}^{\mathrm{i}kx} f(y) \mathrm{e}^{-\mathrm{i}ky} dy\, dk$$

and, invoking Fubini[2], Lemma 4.2, and $(\mathrm{e}^{\mathrm{i}xa} f(x))^\wedge(k) = \hat{f}(k-a)$ we further see that this is equal to

$$= \int_{\mathbb{R}} (\phi_\varepsilon(k) \mathrm{e}^{\mathrm{i}kx})^\wedge(y) f(y) dy = \int_{\mathbb{R}} \frac{1}{\sqrt{\varepsilon}} \phi_{1/\varepsilon}(y-x) f(y) dy.$$

Letting $\varepsilon \to 0$ the integral we have started with converges to $(\hat{f})^\vee(x)$ while the last one converges to $f(x)$ by Lemma A.12. $\square$

Problem* 4.1. *Compute the Fourier transform of the following functions $f : \mathbb{R} \to \mathbb{C}$:*

(i) $f(x) = \chi_{(-1,1)}(x)$. (ii) $f(x) = \frac{\mathrm{e}^{-a|x|}}{a}$, $\mathrm{Re}(a) > 0$.

Problem 4.2. *Show that*

$$\psi_n(x) := H_n(x) \mathrm{e}^{-\frac{x^2}{2}},$$

*where $H_n(x)$ is the **Hermite polynomial**[3] [**25**, (12.7.2)] of degree n given by*

$$H_n(x) := \mathrm{e}^{\frac{x^2}{2}} \left(x - \frac{d}{dx} \right)^n \mathrm{e}^{-\frac{x^2}{2}},$$

are eigenfunctions of the Fourier transform: $\hat{\psi}_n(k) = (-\mathrm{i})^n \psi_n(k)$.

[2] Guido Fubini (1879–1943), Italian mathematician

[3] Charles Hermite (1822–1901), French mathematician

Problem 4.3. *Prove the* ***Poisson summation formula***

$$\sum_{n\in\mathbb{Z}} f(n)\mathrm{e}^{-\mathrm{i}nx} = \sqrt{2\pi}\sum_{m\in\mathbb{Z}} \hat{f}(x+2\pi m),$$

where f satisfies $|f(x)|+|\hat{f}(x)| \le \frac{C}{(1+|x|)^\alpha}$ for some $\alpha>1$. (Hint: Compute the Fourier coefficients of the right-hand side. To this end observe that the integrals over $[-\pi,\pi]$ give a tiling of $\mathbb{R}$ when m runs through all values in $\mathbb{Z}$.)

Problem 4.4. *Prove the* ***Whittaker–Shannon interpolation formula:***[4]

$$f(x) = \sum_{n\in\mathbb{Z}} f(n)\,\mathrm{sinc}(\pi(x-n))$$

provided $\mathrm{supp}(\hat{f}) \subseteq [-\pi,\pi]$. *Here* $\mathrm{sinc}(x) := \frac{\sin(x)}{x}$. *(Hint: Use the Poisson summation formula to express $\hat{f}$ and take the inverse Fourier transform.)*

4.3. The heat equation on the line

Returning to our original problem this shows

$$c(k) = \frac{1}{2\pi}\int_{\mathbb{R}} g(y)\mathrm{e}^{\mathrm{i}ky}dy \tag{4.10}$$

and hence

$$u(t,x) = \frac{1}{2\pi}\int_{\mathbb{R}}\int_{\mathbb{R}} \mathrm{e}^{-\mathrm{i}kx-k^2t} g(y)\mathrm{e}^{\mathrm{i}ky}dy\,dk. \tag{4.11}$$

Using Fubini we obtain

$$u(t,x) = \int_{\mathbb{R}} \Phi(t,x-y)g(y)dy, \tag{4.12}$$

where

$$\Phi(t,x) := \frac{1}{2\pi}\int_{\mathbb{R}} \mathrm{e}^{-\mathrm{i}kx-k^2t}dk = \frac{1}{\sqrt{4\pi t}}\mathrm{e}^{-x^2/(4t)}, \quad t>0, \tag{4.13}$$

is the **fundamental solution of the heat equation**.

Theorem 4.4. *Suppose g is bounded. Then*

$$u(t,x) := \begin{cases} \int_{\mathbb{R}} \Phi(t,x-y)g(y)dy, & t>0,\\ g(x), & t=0, \end{cases} \tag{4.14}$$

defines a solution of the heat equation which satisfies $u\in C^\infty((0,\infty)\times\mathbb{R})\cap C([0,\infty)\times\mathbb{R})$. Moreover,

$$\inf g \le u(t,x) \le \sup g. \tag{4.15}$$

[4]Edmund Taylor Whittaker (1873–1956), British mathematician, physicist, and historian of science

[4]Claude Shannon (1916–2001), American mathematician, electrical engineer, and cryptographer known as "the father of information theory"

Proof. That $u \in C^\infty$ follows since $\Phi \in C^\infty$ upon interchanging differentiation and integration using Lemma A.7. This also shows that u solves the heat equation. It remains to show that u is continuous on the boundary $t = 0$. But this follows from Lemma A.12 since (by Problem A.3)

$$\int_{\mathbb{R}} \Phi(t,x)dx = 1. \tag{4.16}$$

The last claim is immediate from (4.16) and $\Phi(t,x) > 0$. □

Note that since $\Phi > 0$, the inequality in (4.15) is strict for $t > 0$ unless g is constant, which again implies infinite propagation speed. Moreover, even though (4.15) hints at a maximum principle, this is more subtle since the proof of Theorem 3.4 does not extend to unbounded domains. We will address this problem, as well as the question of uniqueness, in Theorem 6.18.

Two more properties are immediate from (4.12):

Corollary 4.5. *Suppose g is integrable. Then the solutions has the following properties:*

(i) *(Mass conservation)*

$$\int_{\mathbb{R}} u(t,x)dx = \int_{\mathbb{R}} g(x)dx. \tag{4.17}$$

(ii)

$$|u(t,x)| \le \frac{1}{\sqrt{4\pi t}} \int_{\mathbb{R}} |g(x)|dx. \tag{4.18}$$

Proof. (ii) is immediate and (i) follows from Fubini. □

The solution of the inhomogeneous equation

$$u_t = u_{xx} + f \tag{4.19}$$

follows from the **Duhamel principle**:

Theorem 4.6. *Suppose $f \in C([0,\infty)\times\mathbb{R})$ is bounded and uniformly Hölder continuous with respect to the second argument on compact sets with respect to the first argument:*

$$|f(t,x) - f(t,y)| \le C_T|x-y|^\gamma, \qquad 0 \le t \le T. \tag{4.20}$$

Then

$$u(t,x) := \int_0^t \int_{\mathbb{R}} \Phi(t-s,x-y)f(s,y)dy\,ds, \qquad t \ge 0, \tag{4.21}$$

is in $C^{1;2}((0,\infty)\times\mathbb{R}) \cap C([0,\infty)\times\mathbb{R})$ and solves the inhomogeneous heat equation with initial condition $u(0,x) = 0$.

Proof. It is not hard to see that $u \in C([0,\infty)\times\mathbb{R})$ since the inner integral converges to $f(t,x)$ as $s \to t$. It is less obvious what happens, when we consider the derivatives. In order to avoid the problems at $s = t$ we consider a cutoff

$$u^\varepsilon(t,x) := \int_0^{t-\varepsilon}\int_{\mathbb{R}} \Phi(t-s,x-y) f(s,y) dy\, ds.$$

Then

$$u_t^\varepsilon(t,x) := \int_{\mathbb{R}} \Phi(\varepsilon, x-y) f(t-\varepsilon,y) dy + \int_0^{t-\varepsilon}\int_{\mathbb{R}} \Phi_t(t-s,x-y) f(s,y) dy\, ds.$$

The first term is unproblematic and converges to $f(t,x)$ as $\varepsilon \to 0$, uniformly on compact sets. To handle the second term we use $\int_{\mathbb{R}} \Phi_t(s,x-y)dy = 0$ such that we can write

$$\begin{aligned}\int_0^{t-\varepsilon}&\int_{\mathbb{R}} \Phi_t(t-s,x-y) f(s,y) dy\, ds\\ &= \int_0^{t-\varepsilon}\int_{\mathbb{R}} \Phi_t(t-s,x-y) \big(f(s,y)-f(s,x)\big) dy\, ds.\end{aligned}$$

Now we use our Hölder estimate to investigate this last integrand

$$\begin{aligned}\Big|\int_0^{t-\varepsilon}&\int_{\mathbb{R}} \Phi_t(t-s,x-y) \big(f(s,y)-f(s,x)\big) dy\, ds\Big|\\ &\le C_T \int_0^T \int_{\mathbb{R}} |\Phi_t(s,y)||y|^\gamma dy\, ds < \infty\end{aligned}$$

since $\Phi_t(t,x) = (\frac{x^2}{4t^2} - \frac{1}{2t})\Phi(t,x)$ and

$$\int_{\mathbb{R}} |y|^\gamma \Big(\frac{y^2}{4s^2} + \frac{1}{2s}\Big)\Phi(s,y) dy = 2(\gamma+2)\frac{\Gamma((\gamma+1)/2)}{\sqrt{\pi}}(4s)^{\gamma/2-1}.$$

Hence, u_t^ε converges uniformly on compact sets as $\varepsilon \to 0$ which shows that u is differentiable with respect to t with derivative given by

$$u_t(t,x) = f(t,x) + \int_0^t \int_{\mathbb{R}} \Phi_t(t-s,x-y)\big(f(s,y)-f(s,x)\big) dy\, ds.$$

Similarly one shows

$$\begin{aligned}u_x(t,x) &= \int_0^t \int_{\mathbb{R}} \Phi_x(t-s,x-y)\big(f(s,y)-f(s,x)\big) dy\, ds,\\ u_{xx}(t,x) &= \int_0^t \int_{\mathbb{R}} \Phi_{xx}(t-s,x-y)\big(f(s,y)-f(s,x)\big) dy\, ds.\end{aligned}$$

This finishes the claim since $\Phi_t = \Phi_{xx}$. □

If an initial condition $u(0,x) = g(x)$ is given, then we can of course combine the above solution with the corresponding solution of the homogenous equation:

$$u(t,x) = \int_{\mathbb{R}} \Phi(t,x-y)g(y)dy + \int_0^t \int_{\mathbb{R}} \Phi(t-s,x-y)f(s,y)dy\,ds, \quad t>0. \tag{4.22}$$

Moreover, like in the case of the wave equation, this result also provides the corresponding solution for a finite interval with Dirichlet (or Neumann) boundary conditions by considering the odd (even) periodic extension of f to $\mathbb{R}$.

Problem 4.5. *Find the solution of the heat equation on* $(0,\infty)$ *with a Dirichlet boundary condition at* 0. *What about a Neumann boundary condition? (Hint: Reflection.)*

Problem 4.6. *The* ***Black–Scholes equation***[5]

$$u_t = -\frac{\sigma^2}{2}x^2 u_{xx} - rx\,u_x + r\,u$$

models the price evolution of a European call or put option for some financial asset as a function of time t. *Here* $\sigma > 0$ *is the asset's volatility,* x *the asset's price and* $r > 0$ *is the interest rate. A financial institution needs to find the price* $u(0,x)$ *it should charge for the option given the value of the option at a terminal time* T. *For a call option, where the asset is to be bought at the exercise price* $p > 0$, *the final condition is* $u(T,x) = \max(x-p,0)$, *while for a put option, where the asset is to be sold at the exercise price* $p > 0$, *the final condition is* $u(T,x) = \max(p-x,0)$.

Show that this problem can be reduced to the heat equation as follows:

- *Introduce a new time* $s := T - t$ *to turn it into an initial value problem.*
- *Observe that the right-hand side (for fixed* t*) is an Euler differential equation, which can be solved by introducing* $y := \log(x)$ *as a new dependent variable.*
- *Finally use Problem 3.8.*

Find the solutions corresponding to a call, put option, respectively.

Problem 4.7. *Show that the solution of the heat equation on* $[0,\infty)\times[0,\infty)$ *with*

$$u(t,0) = a(t), \qquad u(0,x) = 0$$

---

[5] Fischer Black (1938–1995), American economist

[5] Myron Scholes (1941), Canadian-American financial economist

(in particular $a(0) = 0$) is given by

$$u(t,x) = \frac{x}{\sqrt{4\pi}} \int_0^t \frac{1}{(t-s)^{3/2}} \mathrm{e}^{-x^2/(4(t-s))} a(s) ds.$$

(Hint: Assume $a \in C^1$ with $a(0) = 0$ and reduce it to an inhomogeneous problem.)

Problem 4.8. *Compute the energy of the fundamental solution Φ of the heat equation. Show that it does not decay exponentially, and conclude that there is no Poincaré inequality on $\mathbb{R}$.*

Problem 4.9. *Let $g(x) \in C[0,1]$ with $g(0) = g(1)$ be given. Extend the initial condition g to $[-1,1]$ using reflection $g(-x) = -g(x)$ and then to $\mathbb{R}$ using periodicity $g(x+2) = g(x)$. Show that (4.12) defines a solution $u \in C^\infty((0,\infty) \times \mathbb{R})$ which satisfies $u(t,-x) = -u(t,x)$ and $u(t,x+2) = u(t,x)$ for all $t > 0$. Conclude that $u(t,0) = u(t,1) = 0$ for all $t > 0$. Moreover, show that the heat kernel $K(t,x,y)$ for $[0,1]$ can be represented as*

$$K(t,x,y) = \sum_{n \in \mathbb{Z}} \big(\Phi(t, x-y+2n) - \Phi(t, x+y+2n)\big).$$

Note that this last identity is an instance of the Poisson summation formula from Problem 4.3.

Problem 4.10. *Show that if g has compact support, then $u = \Phi_t * g$ is an entire function for $t > 0$. Use this to prove the* ***Weierstrass approximation theorem****: Every $g \in C[0,1]$ can be uniformly approximated by polynomials. (Hint: Extend g to a continuous function on $\mathbb{R}$ with compact support.)*

4.4. The wave equation on the line

Finally we turn to the **wave equation**

$$u_{tt} = u_{xx}, \qquad u(0,x) = g(x), \quad u_t(0,x) = h(x). \tag{4.23}$$

After applying the Fourier transform with respect to x the equation reads

$$\hat{u}_{tt} = -k^2 \hat{u}, \qquad \hat{u}(0,k) = \hat{g}(k), \quad \hat{u}_t(0,k) = \hat{h}(k), \tag{4.24}$$

and the solution is given by

$$\hat{u}(t,k) = \cos(tk)\hat{g}(k) + \frac{\sin(tk)}{k}\hat{h}(k). \tag{4.25}$$

When trying to compute the corresponding fundamental solution one again runs into the problem that neither $\cos(tk)$ nor $\frac{\sin(tk)}{k}$ are integrable. However, the latter is not too far away from being integrable (it is at least square integrable) and one can check that the Fourier transform of $\sqrt{\frac{\pi}{2}}\chi_{[-1,1]}(x/t)$ is $\frac{\sin(tk)}{k}$. Fortunately, the argument leading to the fundamental solution of the heat equation can be extended to such a situation:

Lemma 4.7. *Suppose f, g are integrable. Then*

$$\mathcal{F}^{-1}(\hat{f}g)(x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x-y)\check{g}(y)dy. \tag{4.26}$$

Proof. This is a straightforward application of Fubini:

$$\begin{aligned}\mathcal{F}^{-1}(\hat{f}g)(x) &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \hat{f}(k)g(k)\mathrm{e}^{\mathrm{i}kx}dk = \frac{1}{2\pi} \int_{\mathbb{R}} \int_{\mathbb{R}} f(y)g(k)\mathrm{e}^{\mathrm{i}k(x-y)}dy\, dk \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} f(y) \int_{\mathbb{R}} g(k)\mathrm{e}^{\mathrm{i}k(x-y)}dk\, dy = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x-y)\check{g}(y)dy. \square\end{aligned}$$

Using this lemma we obtain

$$u(t,x) = \int_{\mathbb{R}} \frac{1}{2}\chi_{[-t,t]}(x-y)h(y)dy = \frac{1}{2}\int_{x-t}^{x+t} h(y)dy$$

in the case $g = 0$. But the corresponding expression for $h = 0$ is just the time derivative of this expression, which is known as **Stokes' rule** (cf. Problem 4.20), and thus we obtain again d'Alembert's formula

$$\begin{aligned}u(t,x) &= \frac{1}{2}\frac{\partial}{\partial t}\int_{x-t}^{x+t} g(y)dy + \frac{1}{2}\int_{x-t}^{x+t} h(y)dy \\ &= \frac{g(x+t)+g(x-t)}{2} + \frac{1}{2}\int_{x-t}^{x+t} h(y)dy.\end{aligned} \tag{4.27}$$

Of course there is no point in trying to find conditions such that $u(t,x)$ defined as the inverse Fourier transform of (4.25) solves the wave equation, since it is immediate that d'Alembert's formula gives a solution provided $h \in C^1(\mathbb{R})$ and $g \in C^2(\mathbb{R})$. Concerning uniqueness one can use the factorization from Problem 1.3 and reduce it to uniqueness for the transport equation (Theorem 1.1).

The solution of the inhomogeneous equation

$$u_{tt} - u_{xx} = f, \qquad u(0,x) = 0,\ u_t(0,x) = 0, \tag{4.28}$$

follows from the **Duhamel principle** and is given by

$$u(t,x) = \frac{1}{2}\int_0^t \int_{x-(t-s)}^{x+(t-s)} f(s,y)dy\, ds \tag{4.29}$$

provided $f \in C^{0;1}(\mathbb{R}^2)$ (Problem 4.23).

Without going into details, let me remark that the fact that the wave equation has finite speed of propagation, while the heat equation has not, can also be understood using the Fourier picture. In fact, if f has compact support, then the Fourier integral used to define $\hat{f}$ will make sense not only for real k, but for all $k \in \mathbb{C}$. Moreover, $\hat{f}$ will be an entire function satisfying the growth estimate

$$|\hat{f}(k)| \le \frac{\mathrm{e}^{R|k|}}{\sqrt{2\pi}} \int_{-R}^{R} |f(x)| dx \tag{4.30}$$

provided $\operatorname{supp}(f) \subseteq [-R, R]$. Such functions are said to be of **exponential type** R and it can be shown that the converse also holds. This is know as the **Paley–Wiener theorem**[6] [**29**, Theorem 3.3]. In particular, multiplication with $\cos(tk)$ or $\frac{\sin(tk)}{k}$ preserves such an estimate (increasing the constant R in the estimate, and hence the support), while multiplication with e^{-tk^2} destroys such an estimate. Moreover, this shows that the fact that the wave equation has finite propagation speed is not tied to its particularly simple form, but is shared by other equations as well. For example by the Klein–Gordon equation (Problem 4.18), where it is much harder to derive the fundamental solution (Problem 4.22).

Problem 4.11. *Find the solution of the wave equation satisfying* $u(0, x) = x^2$ *and* $u_t(0, x) = 1$.

Problem 4.12. *Establish d'Alembert's formula for* $h = 0$ *directly from* (4.25) *by finding a formula for* $\mathrm{e}^{\mathrm{i}ak}\hat{f}(k)$.

Problem 4.13. *Show that solutions* u *of the wave equation satisfy the* ***parallelogram property****: If* $a = (t, x)$, $b = a + s(1, 1)$, $c = b + d - a$, $d = a + r(1, -1)$ *are vertices of a parallelogram, then*

$$u(a) + u(c) = u(b) + u(d).$$

Problem 4.14. *Establish* ***Huygens' principle****:*[7] *Suppose* g, h *are supported in* $[a, b] \subset \mathbb{R}$. *Then the solution* (4.27) *of the wave equation has support in* $\{(t, x) | x \in [a - t, b + t]\}$. *If* $\int_a^b h(x) dx = 0$ *the support is in* $\{(t, x) | x \in [a - t, b - t] \cup [a + t, b + t]\}$.

Problem 4.15. *Suppose* g, h *have compact support. Then the solution* (4.27) *of the wave equation satisfies*

$$\int_{\mathbb{R}} u(t, x) dx = \int_{\mathbb{R}} g(x) dx + t \int_{\mathbb{R}} h(x) dx.$$

Problem* 4.16. *Find the solution of the wave equation on* $(0, \infty)$ *with a Dirichlet boundary condition at* 0. *What about a Neumann boundary condition? (Hint: Reflection.)*

---

[6] Raymond Paley (1907–1933)), English mathematician

[7] Christiaan Huygens (1629–1695), Dutch physicist, mathematician, astronomer, and inventor

Problem 4.17. *Derive a formula for the Fourier transform $\hat{u}(t,k)$ of a solution of the* ***beam equation***

$$u_{tt} = -u_{xxxx}, \qquad u(0,x) = g(x), \quad u_t(0,x) = h(x).$$

Do you expect finite or infinite propagation speed?

Problem 4.18. *Derive a formula for the Fourier transform $\hat{u}(t,k)$ of a solution of the* ***Klein–Gordon equation***

$$u_{tt} = u_{xx} - m^2 u, \qquad u(0,x) = g(x), \quad u_t(0,x) = h(x),$$

where m is a constant. Do you expect finite or infinite propagation speed?

Problem 4.19. *Find a transformation which reduces the* ***telegraph equation***

$$u_{tt}(t,x) + 2\eta u_t + \gamma u = c^2 u_{xx}$$

to the Klein–Gordon equation.

Problem* 4.20 (Stokes' rule)**.** *Let $u \in C^3(\mathbb{R})$ be a solution of the Klein–Gordon equation with initial condition $u(0,x) = 0$, $u_t(0,x) = h(x)$. Then $v := u_t$ solves the Klein–Gordon equation with initial condition $v(0,x) = h(x)$, $v_t(0,x) = 0$.*

Problem 4.21. *Show that if both g and h are even (or odd), then so is the solution $u(t,x)$ of the wave equation.*

Problem 4.22. *Show*

$$\frac{1}{2}\int_{-1}^{1} J_0(m\sqrt{1-x^2})\mathrm{e}^{-\mathrm{i}xk}dx = \frac{\sin(\sqrt{k^2+m^2})}{\sqrt{k^2+m^2}},$$

where $J_0(z)$ is the Bessel function of order 0. Conclude that the fundamental solution of the Klein–Gordon equation is given by

$$K(t,x) = \frac{1}{2}J_0(m\sqrt{t^2-x^2})\chi_{[-t,t]}(x)$$

and the solution of the initial value problem is

$$\begin{aligned} u(t,x) =& \frac{1}{2}\frac{\partial}{\partial t}\int_{x-t}^{x+t} J_0(m\sqrt{t^2-(x-y)^2})g(y)dy \\ &+ \frac{1}{2}\int_{x-t}^{x+t} J_0(m\sqrt{t^2-(x-y)^2})h(y)dy. \end{aligned}$$

(Hint: To compute the integral insert the power series for J_0 and express it as a sum of the integrals $I_j(k) := \frac{1}{2}\int_{-1}^{1}(1-x^2)^j\mathrm{e}^{-\mathrm{i}xk}dx$. Then use integration by parts to express I_j' in terms of I_{j+1}.)

Problem* 4.23. *Verify that (4.29) is in C^2 and solves the inhomogeneous wave equation provided $f \in C^{0;1}(\mathbb{R}^2)$.*

Problem 4.24. *Consider the wave equation with initial conditions g, h supported in a compact interval $[-R, R]$. Show that for $t > R$ one has equipartition of energy:*

$$\int_{\mathbb{R}} u_t(t,x)^2 dx = \int_{\mathbb{R}} u_x(t,x)^2 dx.$$

4.5. Dispersion

Of course any linear partial differential equation with constant coefficients can in principle be solved using the Fourier transform. In addition to explicit solution formulas this can also provide profound insight into the behavior of these solutions. To explain this further, suppose we have a linear equation with constant coefficients. The Fourier transform writes the solution as a superposition of **sinusodial waves** (also **plane waves**) of the form

$$u(t,x) = \mathrm{e}^{\mathrm{i}(kx-\omega t)}. \tag{4.31}$$

In principal, the solution of the ordinary differential equation could also involve powers of t, but we will ignore this here. In this context k is known as the **wave number** and ω as the **angular frequency**. The wave length is then $\frac{2\pi}{|k|}$ and the period is $\frac{2\pi}{|\omega|}$. The wave will travel with speed

$$c = \frac{\omega}{k} \tag{4.32}$$

either to the right ($c > 0$) or to the left ($c < 0$). Inserting this ansatz into for example the transport equation $u_t + c\, u_x = 0$ gives

$$\omega(k) = ck. \tag{4.33}$$

This connection between the angular frequency and the wave number is characteristic for the equation and is known as **dispersion relation**. Hence we have sinusodial waves for all wave numbers k and all travel with the same speed. In particular, taking superpositions we have

$$u(t,x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \hat{g}(k) \mathrm{e}^{\mathrm{i}(kx-\omega(k)t)} dk = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \hat{g}(k) \mathrm{e}^{\mathrm{i}k(x-ct)} dk = g(x-ct), \tag{4.34}$$

which just shows that any superposition also travels with the same speed. A result we have already derived before.

Example 4.1. In the case of the wave equation the dispersion relation reads

$$\omega(k)^2 = c^2 k^2 \tag{4.35}$$

and again our sinusodial waves either travel to the left or right, but all with the same speed c. So taking superpositions as in the case of the transport equation, we get the sum of two waves, one traveling to the left and one traveling to the right. Again in agreement with our previous findings. ⋄

Example 4.2. In the case of the heat equation the dispersion relation reads

$$\omega(k) = -\mathrm{i}k^2. \tag{4.36}$$

Hence we do not get sinusoidal waves but decaying solutions $u(t,x) = \mathrm{e}^{\mathrm{i}kx-k^2t}$ (or $\cos(kx)\mathrm{e}^{-k^2t}$ and $\sin(kx)\mathrm{e}^{-k^2t}$ if you prefer real solutions). This is known as **dissipation**. ⋄

Next let us look at the **Korteweg–de Vries (KdV) equation**[8]

$$u_t + u_{xxx} + 6uu_x = 0, \tag{4.37}$$

which appears as a model for waves in shallow water. To obtain a linear equation we will consider waves with small amplitude and neglect the non-linear term:

$$u_t + u_{xxx} = 0. \tag{4.38}$$

In this case the dispersion relation reads

$$\omega(k) = -k^3. \tag{4.39}$$

Now the big difference is that sinusodial waves with different wave numbers k will travel with different speeds

$$c(k) = \frac{\omega(k)}{k} = -k^2 \tag{4.40}$$

and this is known as **dispersion**. As with the transport equation, all waves travel in the same direction and hence this model is unidirectional. The effect of dispersion is that general wave packets will no longer travel with constant speed but will dissolve with time. To understand this mathematically, we take again superpositions

$$u(t,x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \hat{g}(k)\mathrm{e}^{\mathrm{i}(kx-\omega(k)t)}dk = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \hat{g}(k)\mathrm{e}^{\mathrm{i}(kx+k^3t)}dk. \tag{4.41}$$

Rather than trying to *compute* this integral, we will try to understand the asymptotic behavior of the solution as $t \to \infty$. Since we expect wave profiles to travel, it does not make much sense to look at this limit for fixed x. Instead we will fix a speed $c = \frac{x}{t}$ and look at the limit along the ray $x = ct$. Hence we are interested in the asymptotics of the following **oscillatory integral**

$$I(t) := \int_{\mathbb{R}} A(k)\mathrm{e}^{\mathrm{i}\phi(k)t}dk, \tag{4.42}$$

where in our case the amplitude and phase are given by

$$A(k) = \frac{\hat{g}(k)}{\sqrt{2\pi}}, \quad \text{and} \quad \phi(k) = kc + k^3. \tag{4.43}$$

[8]Diederik Korteweg (1848–1941), Dutch mathematician
[8]Gustav de Vries (1866–1934), Dutch mathematician

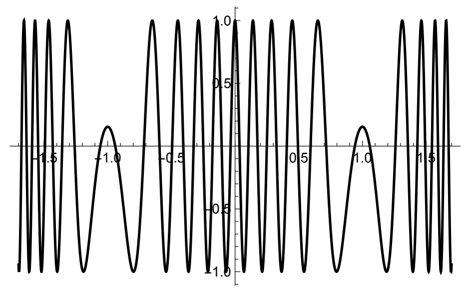


Figure 4.1. Real part of $\mathrm{e}^{\mathrm{i}\phi(k)t}$ for the linearized KdV equation at $t = 10$ with stationary phase points at $k = \pm 1$

Expanding ϕ into a Taylor series shows that the local frequency of $\exp(\mathrm{i}\phi(k)t)$ at k_0 is given by $\phi'(k_0)t$. These oscillations will increase with t and lead to cancellations in the integral. In particular, this factor will oscillate slower in neighborhoods of points where $\phi'(k_0) = 0$. These points are known as **stationary phase points** (cf. Figure 4.1). In our example the stationary phase points are given by

$$\pm k_0, \qquad k_0 := \sqrt{-\frac{c}{3}}, \quad c \le 0, \tag{4.44}$$

and there are no stationary phase points (at least on the real line, where integration takes place) for $c > 0$. For $c < 0$ we have two points with $\phi''(k_0) = 6k_0 \neq 0$ while for $c = 0$ we have $\phi''(0) = 0$ but at least $\phi'''(0) = 6$ is nonvanishing.

Lemma 4.8 (van der Corput[9]). *Suppose $\phi \in C^{n+1}([a,b])$ is real-valued and satisfies $|\phi^{(n)}(k)| \ge \varepsilon > 0$ plus $\phi'(k)$ monotone if $n = 1$. Then*

$$\left| \int_a^b \mathrm{e}^{\mathrm{i}\phi(k)t} dk \right| \le \frac{C_n}{(\varepsilon t)^{1/n}}, \qquad t > 0. \tag{4.45}$$

Here the constant depends only on n and not on the interval (a,b).

Proof. By a simple scaling we can assume $\varepsilon = 1$ without loss of generality. We use induction and start with the case $n = 1$. Using integration by parts shows

$$\begin{aligned}\int_a^b \mathrm{e}^{\mathrm{i}\phi(k)t} dk &= \int_a^b \frac{1}{\mathrm{i}\phi'(k)t} \left(\frac{d}{dk} \mathrm{e}^{\mathrm{i}\phi(k)t} \right) dk \\ &= \frac{\mathrm{e}^{\mathrm{i}\phi(b)t}}{\mathrm{i}\phi'(b)t} - \frac{\mathrm{e}^{\mathrm{i}\phi(a)t}}{\mathrm{i}\phi'(a)t} - \int_a^b \left(\frac{d}{dk} \frac{1}{\mathrm{i}\phi'(k)t} \right) \mathrm{e}^{\mathrm{i}\phi(k)t} dk.\end{aligned}$$

[9]Johannes van der Corput (1890–1975), Dutch mathematician

Since $|\phi'(k)| \geq 1$ the two boundary terms can be estiamted by $\frac{1}{t}$ and for the integral we obtain (using that ϕ'' is of one sign since ϕ' is assumed monotone)

$$\begin{aligned}\left|\int_a^b \left(\frac{d}{dk}\frac{1}{\mathrm{i}\phi'(k)t}\right)\mathrm{e}^{\mathrm{i}\phi(k)t}dk\right| &\leq \frac{1}{t}\left|\int_a^b \left(\frac{d}{dk}\frac{1}{\phi'(k)}\right)dk\right| \\ &= \frac{1}{t}\left|\frac{1}{\phi'(b)} - \frac{1}{\phi'(a)}\right| \leq \frac{2}{t}.\end{aligned}$$

Hence the claim holds with $C_1 = 4$.

Now consider the induction step and assume $|\phi^{(n+1)}(k)| \geq 1$. Then $\phi^{(n)}(k)$ is monotone and can have at most one zero. Let k_0 be this zero or, if there is no zero, choose k_0 such that $|\phi^{(n)}(k_0)|$ gets minimal (i.e., one of the boundary points). Then we have $|\phi^{(n)}(k)| \geq \delta$ for $|k-k_0| \geq \delta$ since $|\phi^{(n+1)}(k)| \geq 1$.

Hence the induction hypothesis implies

$$\left|\int_{|k-k_0|\geq\delta} \mathrm{e}^{\mathrm{i}\phi(k)t}dk\right| \leq \frac{2C_n}{(\delta t)^{1/n}},$$

which combined with the trivial estimate $\left|\int_{|k-k_0|\leq\delta} \mathrm{e}^{\mathrm{i}\phi(k)t}dk\right| \leq 2\delta$ gives

$$\left|\int_a^b \mathrm{e}^{\mathrm{i}\phi(k)t}dk\right| \leq 2\delta + \frac{2C_n}{(\delta t)^{1/n}}.$$

Choosing $\delta = t^{-1/(n+1)}$ establishes the claim with $C_{n+1} = 2(C_n + 1)$. □

Now this immediately gives us a corresponding estimate for our oscillatory integral $I(t)$.

Theorem 4.9. *Suppose the phase ϕ satisfies the assumptions of the van der Corput lemma and suppose $A \in C^1([a,b])$ is bounded and has an integrable derivative. Then*

$$\left|\int_a^b A(k)\mathrm{e}^{\mathrm{i}\phi(k)t}dk\right| \leq \frac{C_n\big(|A(a)| + |A(b)| + \int_a^b |A'(k)|dk\big)}{(\varepsilon t)^{1/n}}. \tag{4.46}$$

In the case $A \in C^1(\mathbb{R})$ with both A' integrable and $\lim_{|k|\to\infty} A(k) = 0$ we have

$$\left|\int_{\mathbb{R}} A(k)\mathrm{e}^{\mathrm{i}\phi(k)t}dk\right| \leq \frac{C_n \int_{\mathbb{R}} |A'(k)|dk}{(\varepsilon t)^{1/n}}. \tag{4.47}$$

Proof. Abbreviate $\Phi(k) := \int_{k_0}^k \mathrm{e}^{\mathrm{i}\phi(r)t}dr$, where $k_0 \in (a,b)$ and note that $|\Phi(k)| \leq C_n(\varepsilon t)^{-1/n}$ by the van der Corput lemma. Now

$$\left|\int_a^b A(k)\Phi'(k)dk\right| = \left|A(b)\Phi(b) - A(a)\Phi(a) - \int_a^b A'(k)\Phi(k)dk\right|$$

from which the first claim follows. In the case of $\mathbb{R}$ take limits $a \to -\infty$ and $b \to \infty$. □

For our linearized KdV equation this implies

$$|u(t,x)| \le \frac{C_3}{(6t)^{1/3}} \int_{\mathbb{R}} |\hat{g}'(k)| dk, \tag{4.48}$$

which applies for all $x \in \mathbb{R}$. If we look at the sector $\frac{x}{t} \ge c_0 > 0$ we have $\phi'(k) = c + 3k^2 \ge c_0$ and hence we have the stronger estimate (since ϕ' is not monotone we need to split the interval)

$$|u(t,x)| \le \frac{2C_1}{c_0 t} \int_{\mathbb{R}} |\hat{g}'(k)| dk, \qquad \frac{x}{t} \ge c_0. \tag{4.49}$$

Similarly, we obtain a stronger estimate for the sector $\frac{x}{t} \le -c_0 < 0$. To do this we split the integral into neighborhoods of the stationary phase points, where we apply our result with $n = 2$, and the rest, where we apply our result with $n = 1$. This shows

$$|u(t,x)| \le \frac{C}{\sqrt{t}} \int_{\mathbb{R}} |\hat{g}'(k)| dk, \qquad -\frac{x}{t} \ge c_0. \tag{4.50}$$

In fact, since only the contributions of the small neighborhoods decay like $O(t^{-1/2})$, while the rest contributes $O(t^{-1})$, this suggests to investigate the contribution of such a stationary phase point. This is known as the **method of stationary phase**.

Theorem 4.10. *Suppose $\phi \in C^4((a,b))$ and $A \in C^2((a,b))$ with A bounded and A' integrable. Suppose (a,b) contains precisely one stationary phase point k_0 with $\phi''(k_0) \ne 0$ and $|\phi'(k)| \ge \varepsilon > 0$ outside a neighborhood of k_0. Then*

$$\int_a^b A(k) \mathrm{e}^{\mathrm{i}\phi(k)t} dk = A(k_0) \mathrm{e}^{\mathrm{i}\,\mathrm{sign}(\phi''(k_0))\frac{\pi}{4} + \mathrm{i}\phi(k_0)t} \sqrt{\frac{2\pi}{|\phi''(k_0)|t}} + O(t^{-1}). \tag{4.51}$$

Proof. As argued above we can take (a,b) to be an arbitrary small neighborhood of k_0. Moreover, we can use a change of coordinates $\kappa \in C^3$ such that $\phi(k)$ has the simple standard form (Problem 4.28)

$$\phi(k) - \phi(k_0) = \frac{\sigma}{2} \kappa(k - k_0)^2, \qquad \sigma := \mathrm{sign}(\phi''(k_0)),$$

and such that $(b - k_0, a - k_0)$ is mapped to $(-\delta, \delta)$. In this new coordinate our integral reads

$$I_0(t) = \mathrm{e}^{\mathrm{i}\phi(k_0)t} \int_{-\delta}^{\delta} B(\kappa) \mathrm{e}^{\mathrm{i}\frac{\sigma t}{2}\kappa^2} d\kappa, \qquad B(\kappa) := \frac{A(k_0 + k)}{\kappa'(k)}.$$

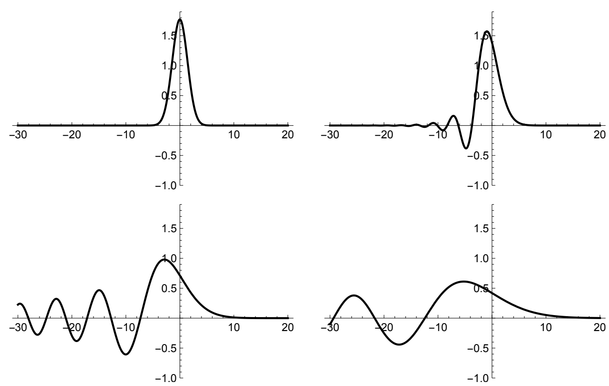


Figure 4.2. A Gaussian initial condition evolving according to the linearized KdV equation, forming a dispersive tail.

Here $B \in C^2$ with $B(0) = \frac{A(k_0)}{\sqrt{|\phi''(k_0)|}}$. Now write $B(\kappa) = B(0) + \kappa C(\kappa)$ with $C \in C^1$. Then, using integration by parts, we see

$$\int_{-\delta}^{\delta} C(\kappa)\kappa \mathrm{e}^{\mathrm{i}\frac{\sigma t}{2}\kappa^2} d\kappa = \frac{1}{\mathrm{i}\sigma t}\left(C(\kappa)\mathrm{e}^{\mathrm{i}\frac{\sigma t}{2}\kappa^2}\Big|_{-\delta}^{\delta} - \int_{-\delta}^{\delta} C'(\kappa)\mathrm{e}^{\mathrm{i}\frac{\sigma t}{2}\kappa^2} d\kappa \right) = O(t^{-1}).$$

Thus

$$I_0(t) = \frac{2A(k_0)}{\sqrt{|\phi''(k_0)|t}} \mathrm{e}^{\mathrm{i}\phi(k_0)t} \int_0^{\delta\sqrt{t}} \mathrm{e}^{\mathrm{i}\frac{\sigma}{2}\kappa^2} d\kappa + O(t^{-1}).$$

Now the integral on the right is the famous Fresnel integral and its limit as $t \to \infty$ is $\mathrm{e}^{\mathrm{i}\sigma\pi/4}\sqrt{\frac{\pi}{2}}$ (Problem 4.29). Moreover, invoking once more the van der Corput lemma, the difference between this integral and its limit is $O(t^{-1})$, which finishes the proof. □

Of course this result is just the tip of the iceberg and we refer to **[23]** for more on the stationary phase method.

In our case this gives (recall $k_0 = \sqrt{-\frac{x}{3t}}$)

$$u(t,x) = \mathrm{Re}\left(\hat{g}(k_0)\mathrm{e}^{\mathrm{i}(k_0 x + t k_0^3 + \frac{\pi}{4})}\right)\sqrt{\frac{2}{3k_0 t}} + O(t^{-1}), \qquad -\frac{x}{t} \geq c_0, \quad (4.52)$$

and is illustrated in Figure 4.2. The initial condition is a Gaussian which travels to the left thereby dissolving into a dispersive tail.

Finally, let us compute the fundamental solution of the linearized KdV equation. The problem is that we cannot take the inverse Fourier transform of $\mathrm{e}^{\mathrm{i}tk^3}$. Hence we investigate the truncated integral

$$K_r(t,x) := \frac{1}{2\pi}\int_{-r}^{r} \mathrm{e}^{\mathrm{i}tk^3 + \mathrm{i}xk} dk. \quad (4.53)$$

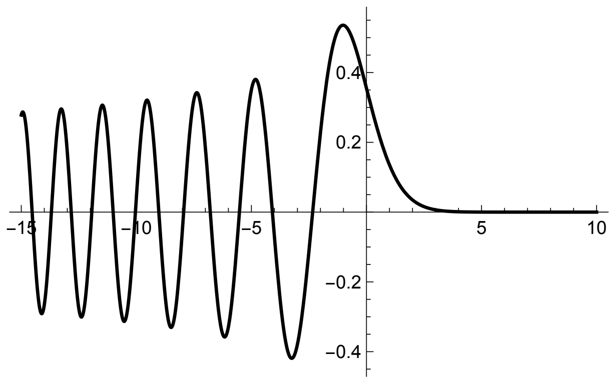


Figure 4.3. The Airy function $\mathrm{Ai}(x)$.

Using $\phi(k) = k^3 + ck$ with $c = \frac{x}{t}$ fixed, the van der Corput lemma gives us an a priori bound

$$|K_r(t,x)| \le \frac{C_3}{2\pi(6|t|)^{1/3}}. \tag{4.54}$$

Moreover, integration by parts shows

$$\int_a^r \mathrm{e}^{\mathrm{i}t\phi(k)} dk = \frac{\mathrm{e}^{\mathrm{i}t\phi(k)}}{\mathrm{i}t\phi'(k)}\Big|_{k=a}^r + \frac{1}{\mathrm{i}t}\int_a^r \frac{\phi''(k)}{\phi'(k)^2}\mathrm{e}^{\mathrm{i}t\phi(k)} dk. \tag{4.55}$$

Hence the limit

$$K(t,x) := \lim_{r\to\infty} K_r(t,x) \tag{4.56}$$

exist and if g, $\hat{g}$ are integrable we get

$$\begin{aligned} u(t,x) &= \lim_{r\to\infty} \frac{1}{\sqrt{2\pi}} \int_{-r}^{r} \mathrm{e}^{\mathrm{i}tk^3+\mathrm{i}xk}\hat{g}(k)dk \\ &= \lim_{r\to\infty} \int_{\mathbb{R}} K_r(t,x-y)g(y)dy = \int_{\mathbb{R}} K(t,x-y)g(y)dy \end{aligned}$$

thanks to our a priori bound. In summary we obtain that the fundamental solution is given by

$$K(t,x) = (3t)^{-1/3}\,\mathrm{Ai}((3t)^{-1/3}x), \tag{4.57}$$

where

$$\mathrm{Ai}(x) := \frac{1}{2\pi}\int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}k^3/3+\mathrm{i}xk} dk = \frac{1}{\pi}\int_0^\infty \cos(k^3/3 + xk)dk \tag{4.58}$$

is the **Airy function**[10] (see Figure 4.3). Of course the integral is to be understood as an improper Riemann integral (i.e., $\int_{\mathbb{R}} = \lim_{r\to\infty}\int_{-r}^{r}$) as explained above.

[10]George Biddell Airy (1801–1892), English mathematician and astronomer

We close this section with a remark about the KdV equation. We have seen that the linearized version causes solutions to decay. On the other hand, if we drop the u_{xxx} term, we obtain Burgers' equation which leads to breaking of waves. It turns out that in the KdV equation both effects can compensate each other and lead to stable wave profiles, known as **solitons**, which evolve without changing their shape. In fact, you can easy check that

$$u(t,x) = \frac{c}{2\cosh\left(\frac{\sqrt{c}}{2}(x-ct-x_0)\right)^2} \tag{4.59}$$

solves the KdV equation. Note that the amplitude is proportional to the speed c, while the width of the wave is proportional to $c^{-1/2}$. Taller waves travel faster and are more narrow.

One can show that any *nice* decaying initial condition will split into a dispersive tail traveling to the left plus a finite number of solitons (Figure 4.4). In this sense the solitons are the stable parts of arbitrary initial conditions. In fact, solitons were first observed by the naval engineer John Scott Russell[11] in the Union Canal in Scotland. However, at his time not much attention was paid to this discovery. Following experiments by Russell there were theoretical investigations by Rayleigh[12] and Boussinesq[13]. The latter derived the KdV equation which was later rediscovered by Korteweg and his student de Vries. However, the KdV equation still did not receive much attention after this until Zabusky[14] and Kruskal[15] discovered numerically that solutions decompose into solitons and used this to explain the famous **Fermi–Pasta–Ulam–Tsingou experiment**.[16] At that moment the theory of solitons started to explode and turned into one of the most active areas in mathematical physics. While the KdV equation was soon formally solved by Gardner[17], Greene[18], Kruskal and Miura[19] by what is today known as the **inverse scattering transform**, a rigorous mathematical treatment establishing the aforementioned soliton asymptotics is a quite formidable task (known as **soliton resolution conjecture**). Needless to say, that this is well beyond our mathematical tool box.

---

[11] John Scott Russell (1808–1882), Scottish civil engineer, naval architect and shipbuilder
[12] John William Strutt, 3rd Baron Rayleigh (1842–1919), English physicist
[13] Joseph Boussinesq (1842–1929), French mathematician and physicist
[14] Norman Zabusky (1929–2018), American physicist
[15] Martin David Kruskal (1925–2006), American mathematician and physicist
[16] Enrico Fermi (1901–1954), Italian (later naturalized American) physicist
[16] John Pasta (1918–1981), American computational physicist and computer scientist
[16] Stanislaw Ulam (1909–1984), Polish-American mathematician and nuclear physicist
[16] Mary Tsingou (*1928), American physicist and mathematician
[17] Clifford S. Gardner (1924–2013), American mathematician
[18] John M. Greene (1928–2007), American theoretical physicist and applied mathematician
[19] Robert M. Miura (1938–2018), American mathematician

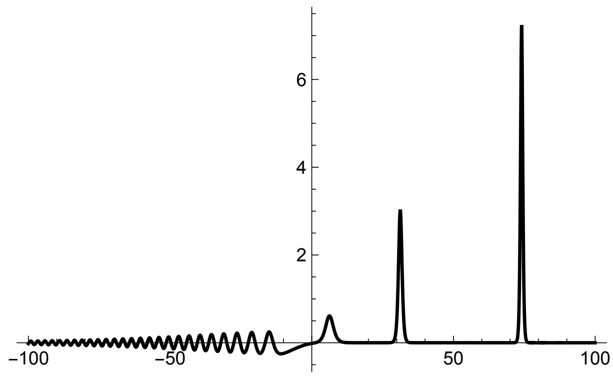


Figure 4.4. A solution of the KdV equation has split into three solitons traveling to the right and a dispersive tail traveling to the left.

Problem 4.25. *Find the dispersion relation of the* ***beam equation***

$$u_{tt} + u_{xxxx} = 0.$$

Problem 4.26. *Consider the partial differential equation*

$$u_t = \sum_{j=0}^{m} c_j \partial_x^j u, \qquad u(0,x) = g(x),$$

whose solution in Fourier space is formally given by $\hat{u}(t,k) = \mathrm{e}^{-\mathrm{i}\omega(k)t}\hat{g}(k)$*. Compute* ω *and answer the following questions in terms of* ω *(assuming the solution is well defined, say* $\hat{g}$ *rapidly decaying and* $\mathrm{Im}(\omega) \le 0$*):*

- *Is* $\int_{\mathbb{R}} u(t,x)dx$ *a constant of motion?*
- *Is* $\int_{\mathbb{R}} |u(t,x)|^2 dx$ *a constant of motion?*
- *Is the propagation speed finite or infinite?*

Problem 4.27. *Find the asymptotics of the oscillatory integral*

$$I(t) := \int_{\mathbb{R}} \frac{1}{1+k^2} \mathrm{e}^{\mathrm{i}(k^3/3 - k^2/2)t} dk.$$

Problem* 4.28. *Let* $\phi \in C^3$ *with* $\phi'(k_0) = 0$ *and* $\phi''(k_0) \neq 0$*. Show that there is a local change of coordinates* $\kappa \in C^2$ *such that* $\phi(k) - \phi(k_0) = \frac{\sigma}{2}\kappa(k-k_0)^2$*,* $\sigma := \mathrm{sign}(\phi''(k_0))$*, holds in a neighborhood of* k_0*. In particular,* $\kappa(0) = 0$*,* $\kappa'(0) = \sqrt{|\phi''(k_0)|}$*, and* $\kappa''(0) = \frac{\sigma}{3\sqrt{|\phi''(k_0)|}}\phi'''(k_0)$*. Moreover, if* $\phi \in C^4$*, then* $\kappa \in C^3$*.*

Problem* 4.29. *Compute the* ***Fresnel integral***[20]

$$\lim_{R\to\infty} \int_0^R \mathrm{e}^{\mathrm{i}\frac{x^2}{2}} dx = \mathrm{e}^{\mathrm{i}\pi/4}\sqrt{\frac{\pi}{2}}.$$

[20]Augustin-Jean Fresnel (1788 –1827), French civil engineer and physicist

(Hint: Consider $\sqrt{\frac{2}{\pi}}\int_0^r\int_0^\infty \mathrm{e}^{y(\mathrm{i}-t^2)}dt\,dy$ *and use Fubini.)*

Problem 4.30. *Consider the oscillatory integral* (4.42) *under the assumptions of the van der Corput lemma with* $n \geq 2$. *Suppose* A *is in the Wiener algebra, that is,*

$$A(k) = \int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}kx}\ f(x)dx$$

with f *integrable. Then*

$$|I(t)| \leq \frac{C_n}{(\varepsilon t)^{1/n}} \int_{\mathbb{R}} |f(x)|dx, \quad t > 0.$$

where C_n *is the constant from the van der Corput lemma. (Hint: Fubini.)*

Problem 4.31. *Consider the* ***Schrödinger equation***[21]

$$-\mathrm{i}u_t = u_{xx}, \qquad u(0,x) = g(x).$$

Compute the dispersion relation and establish the dispersive estimate

$$|u(t,x)| \leq \frac{C_2}{\sqrt{2|t|}} \int_{\mathbb{R}} |g(x)|dx.$$

(Hint: Problem 4.30.)

Problem* 4.32. *Use the integral representation for the Bessel function* $J_n(x)$ *from Problem 3.24 to establish the asymptotics* (3.96) *(for integer order as* $x \to \infty$ *with error* $O(x^{-1/2})$ *instead of* $O(x^{-1})$*).*

Problem 4.33. *Show that*

$$\mathrm{Ai}(x) := \frac{1}{2\pi} \int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}(k+\mathrm{i}\varepsilon)^3/3 + \mathrm{i}x(k+\mathrm{i}\varepsilon)} dk, \qquad \varepsilon > 0,$$

where now the integrand is integrable. Use this representation to show that Ai *solves* ***Airy's equation***

$$u'' - x\,u = 0.$$

Note that this is a special case ($a = \frac{1}{2}$, $b = \frac{2}{3}\mathrm{i}$, $c = \frac{3}{2}$, $\nu = \frac{1}{3}$) *of Lommel's equation from Problem 3.21. Moreover, with* $\omega = \mathrm{e}^{\mathrm{i}2\pi/3}$ *show that* $\mathrm{Ai}(\omega x)$, $\mathrm{Ai}(\omega^2 x)$ *are again solutions and*

$$\mathrm{Ai}(x) + \omega\,\mathrm{Ai}(\omega x) + \omega^2\,\mathrm{Ai}(\omega^2 x) = 0.$$

Problem 4.34. *Show that the Airy function satisfies*

$$\mathrm{Ai}(x) = \frac{\mathrm{e}^{-2x^{3/2}/3}}{2\pi} \int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}k^3/3 - \sqrt{x}k^2} dk, \qquad |\arg(x)| < \pi.$$

[21] Erwin Schrödinger (1887–1961), Austrian physicist

Use this to show

$$|\operatorname{Ai}(x)| \le \frac{e^{-2x^{3/2}/3}}{2\sqrt{\pi}x^{1/4}}, \qquad \operatorname{Ai}(x) = \frac{e^{-2x^{3/2}/3}}{2\sqrt{\pi}x^{1/4}}\Big(1 + O(x^{-6/4})\Big)$$

for $x > 0$ *and*

$$|\operatorname{Ai}(x)| \le \frac{1}{\sqrt{\pi}(-x)^{1/4}}, \quad \operatorname{Ai}(x) = \frac{\sin(\frac{2}{3}(-x)^{3/2} + \frac{\pi}{4})}{\sqrt{\pi}(-x)^{1/4}} + O((-x)^{-7/4})$$

for $x < 0$. *(Hint: Suppose* $x > 0$ *and choose* $\varepsilon = \sqrt{x}$ *in the integral representation from the previous problem. To get it for* $|\arg(x)| < \pi$ *invoke the identity theorem. Now for* $x > 0$ *use a Taylor expansion for* $e^{ik^3/3}$ *and for* $x < 0$ *use* $\operatorname{Ai}(-x) = -\omega \operatorname{Ai}(-\omega x) - \omega^2 \operatorname{Ai}(-\omega^2 x)$ *established in the previous problem.)*

Problem 4.35. *Find all traveling wave solutions, that is, solutions of the form* $u(t, x) = v(x - ct)$, *of the KdV equation (4.37). (Hint: Insert this ansatz and integrate once to obtain* $v'' + 3v^2 - cv + a = 0$. *Use a translation to eliminate the linear term and note that this can be viewed as Newton's equation for a particle in a potential.)*

4.6. Symmetry groups

Many partial differential equations are invariant under certain group actions. In such a situation one can apply the group action in order to get new solutions out of old ones. Let us for example look at the heat equation $u_t = u_{xx}$. Two obvious **symmetries** are space and time translations

$$(t, x) \mapsto (t + \varepsilon, x) \qquad (t, x) \mapsto (t, x + \varepsilon), \tag{4.60}$$

which show that if u is a solution, so will be

$$u(t + \varepsilon, x) \qquad \text{and} \qquad u(t, x + \varepsilon). \tag{4.61}$$

Also easy to spot is the scaling symmetry

$$(t, x) \mapsto (e^{2\varepsilon} t, e^{\varepsilon} x), \tag{4.62}$$

which shows that if u is a solution, so will be

$$u(e^{2\varepsilon} t, e^{\varepsilon} x). \tag{4.63}$$

However, apart from linearity, there are two more group actions which leave solutions invariant:

$$e^{-\varepsilon x + \varepsilon^2 t} u(t, x - 2\varepsilon t), \qquad \frac{1}{\sqrt{1 + 4\varepsilon t}} e^{-\frac{\varepsilon x^2}{1 + 4\varepsilon t}} u\Big(\frac{t}{1 + 4\varepsilon t}, \frac{x}{1 + 4\varepsilon t}\Big). \tag{4.64}$$

While it is straightforward to verify these formulas (please do it!), it remains a mystery how they were found. It turns out there is a fully fledged theory behind this developed by Sophus Lie[22]. In fact, he invented **Lie groups** precisely for this purpose! This theory is beyond our scope and I refer to [**26**] for further details (the heat equation is discussed in Example 2.41 of [**26**]). Also note that many computer algebra systems have packages for computing the (continuous) symmetry groups of a given differential equation.

Here we want to point out another fruitful approach, namely to look for group invariant solutions. For example, consider a linear combination of the scaling symmetry from above and the scaling of the solution u:

$$(t, x, u) \mapsto (\mathrm{e}^{2\varepsilon} t, \mathrm{e}^{\varepsilon} x, \mathrm{e}^{2\alpha\varepsilon} u), \qquad \alpha \in \mathbb{R}. \tag{4.65}$$

Such solutions are known as **similarity solutions**. This suggests to make a change of coordinates $y = \frac{x}{\sqrt{t}}$, $v = t^{-\alpha} u$ such that the new coordinates are invariant. This gives the ordinary differential equation

$$v_{yy} + \frac{y}{2} v_y - \alpha v = 0. \tag{4.66}$$

In the case $\alpha = 0$ this leads to the solution

$$u(t, x) = c_1 + c_2 \operatorname{erf}(\frac{x}{2\sqrt{t}}), \tag{4.67}$$

where $\operatorname{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z \mathrm{e}^{-t^2} dt$ is the Gauss **error function**. Differentiating with respect to x and choosing $c_2 = \frac{1}{2}$ gives the fundamental solution Φ. In the general case one uses the standard transformation $w(y) = \mathrm{e}^{1/2 \int (y/2) dy} v(y) = \mathrm{e}^{y^2/8} v(y)$ to eliminate the first order derivative,

$$w_{yy} = \big(\alpha + \frac{1}{4} + \frac{y^2}{16}\big) w, \tag{4.68}$$

which is **Weber's equation**. The solutions are the **parabolic cylinder functions** [**25**, (12.2.1)]. If $\alpha = \frac{n+1}{2}$ with $n \in \mathbb{N}_0$ they can be expressed in terms of Hermite polynomials H_n (cf. Problem 4.2) giving

$$u(t, x) = t^{-(n+1)/2} H_n(\frac{x}{2\sqrt{t}}) \mathrm{e}^{-x^2/(4t)}. \tag{4.69}$$

Again the case $n = 0$ gives the fundamental solution since $H_0(x) = 1$. See [**26**, Chapter 3] for more on group invariant solutions (the heat equation is discussed in Example 3.3).

Problem 4.36. *Show that if $u \in C^3$ is a solution of the heat equation, so is*

$$v(t, x) := x u_x(t, x) + 2t u_t(t, x).$$

(Hint: Use (4.63)*.)*

Problem 4.37. *Look for similarity solutions of the linearized Korteweg–de Vries equation* (4.38)*. Derive the corresponding differential equation (you do not need to solve it).*

---

[22] Sophus Lie (1842–1899), Norwegian mathematician

Chapter 5

The Laplace equation

Before we begin, we recall some basic facts and fix our notation:

A nonempty and open subset $U \subseteq \mathbb{R}^n$ is called a **domain**. We will write $B_r(x) := \{y \in \mathbb{R}^n | \, |x - y| < r\}$ for the open **ball** of radius r centered at $x \in \mathbb{R}^n$ (we do not display the dimension n explicitly as it will be always clear from the context). Of course we use the usual Euclidean distance $|x| := \sqrt{\sum_{j=1}^n x_j^2}$ on $\mathbb{R}^n$. The corresponding closed ball will be denoted by $\bar{B}_r(x) := \{y \in \mathbb{R}^n | \, |x - y| \le r\}$. The **boundary** of U will be denoted by $\partial U := \overline{U} \setminus U$, where $\overline{U}$ is the **closure** of U. The boundary of a ball is of course $\partial B_r(x) := \{y \in \mathbb{R}^n | \, |x - y| = r\}$. Recall that the **distance**

$$\operatorname{dist}(x, U) := \inf_{y \in U} |y - x| \tag{5.1}$$

is a continuous functions and hence attains its minimum on every compact set K by the extreme value theorem of Weierstrass. Thus if $K \subset U$ is compact, then all points from K are a positive distance away from ∂U since the boundary consists precisely of those point for which the distance to U is zero. In particular we have $\bar{B}_r(x) \subset U$ if and only if $\operatorname{dist}(x, \partial U) > r$.

A set U will be called connected if there is no nontrivial subset which is both open and closed (in the relative topology). In other words, if U is connected, then any nonempty subset which is both open and closed is equal to U.

We will also employ the usual multi-index notation for partial derivatives (see Appendix A.1) and some basic facts from multi-dimensional integration theory, in particular surface measure and the Gauss–Green[1] theorem (see Appendix A.2). In this context we will use the term *integrable* in a naive

[1]George Green (1793–1841), British mathematical physicist

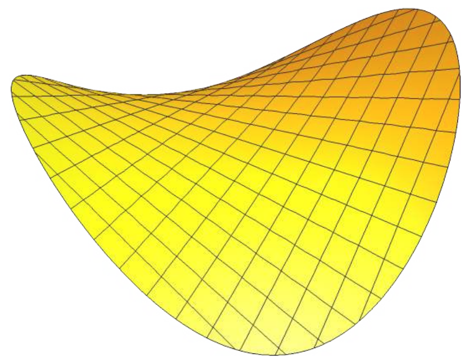


Figure 5.1. The harmonic function xy

way. So you can read it as either (absolutely) Riemann integrable or, if you are familiar with the Lebesgue[2] integral, as measurable with a finite integral of the absolute value. In particular, we always understand it as *absolutely* integrable, that is, in case of an improper Riemann integral we assume it to converge absolutely.

5.1. Harmonic functions

We have already encountered the Laplace equation

$$\Delta u = \sum_{j=1}^{n} u_{x_j x_j} = 0, \tag{5.2}$$

discovered by Laplace in his study of celestial mechanics.

In one dimension the situation is quite simple since the only solutions are affine functions. Hence we will assume $n \geq 2$ throughout this chapter. In higher dimensions it is easy to come up with further examples. For example

$$u(x, y) = x^2 - y^2 \tag{5.3}$$

is a solution in two dimensions. Since the equation is linear, linear combinations of solutions will be again solutions and so will be any translation, scaling, or rotation of a solution since the Laplace equation is invariant under these operations. For example, if we rotate the above solution by 45 degrees we get the new solution

$$v(x, y) = u\big(\frac{x-y}{\sqrt{2}}, \frac{x+y}{\sqrt{2}}\big) = -2xy. \tag{5.4}$$

Moreover, if a solution is sufficiently smooth, its derivatives will again be solutions since partial derivatives commute by Schwarz' lemma.[3]

[2] Henri Lebesgue (1875-1941), French mathematician

[3] Hermann Schwarz (1843–1921), German mathematician

In fact, in two dimensions it is easy to produce many more examples since the Cauchy–Riemann differential equations from complex analysis imply that both real and imaginary part of a holomorphic function solve the Laplace equation. So it is not surprising, at least in two dimensions, that solutions of the Laplace equation share many of the nice properties of holomorphic functions. The stunning fact is, that this remains true in higher dimensions.

In one dimension the value of a solution in the middle of an interval will be precisely the mean value of its boundary values and we call such functions harmonic. Similarly, in two dimensions the Poisson integral (3.108) applied to a ball $B_r(x)$ and evaluated in the center (note $P_0(r,\vartheta)=1$) gives

$$u(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} u(x + r(\cos(\vartheta), \sin(\vartheta)))d\vartheta. \tag{5.5}$$

It turns out that this property is characteristic for solutions of the Laplace equation in arbitrary dimensions and we will take it as the starting point for our investigations.

Let $U \subseteq \mathbb{R}^n$ be a domain. A function $u \in C(U)$ is called **harmonic** if it satisfies the Gauss **mean value property**

$$u(x) = \frac{1}{nV_n}\int_{S^{n-1}} u(x + r\omega)d\sigma^{n-1}(\omega) \tag{5.6}$$

for every ball $\bar{B}_r(x) \subset U$. Here V_n denotes the volume of the unit ball $B_1(0)$ in $\mathbb{R}^n$ and consequently nV_n is the surface of the unit sphere $S^{n-1} = \partial B_1(0)$. The fact that we require $\bar{B}_r(x) \subset U$ implies that the boundary of this ball $\partial B_r(x)$ is within U and hence u is well-defined on $\partial B_r(x)$. In this context let us emphasize, that u is not required to have an continuous extension to $\overline{U}$. It could get arbitrarily wild when we approach the boundary of U.

Note that if u is harmonic, it also automatically satisfies a corresponding mean value property where the integral is taken over the ball rather than over its boundary. The volume of a ball of radius r is of course $|B_r(x)| = V_n r^n$ and we will frequently omit the center and just write $|B_r|$. Similarly, the surface area of a sphere of radius r is $|\partial B_r(x)| = nV_n r^{n-1}$ and we will again write $|\partial B_r|$ for brevity.

Lemma 5.1. *Let u be harmonic in U. Then*

$$u(x) = \frac{1}{|B_r|}\int_{B_r(x)} u(y)d^n y \tag{5.7}$$

for every ball $\bar{B}_r(x) \subset U$.

Proof. A simple calculation using

$$\begin{aligned}\frac{1}{|B_r|}\int_{B_r(x)} u(y)d^n y &= \frac{1}{V_n r^n}\int_0^r \left(\int_{S^{n-1}} u(x+s\omega)d\sigma^{n-1}(\omega)\right) s^{n-1}ds \\ &= \frac{1}{V_n r^n}\int_0^r \left(nV_n u(x)\right) s^{n-1}ds = u(x). \qquad \square\end{aligned}$$

We first show that harmonic functions are in fact smooth.

Lemma 5.2. *Let u be harmonic in U. Then $u \in C^\infty(U)$.*

Proof. Take a nonnegative radial function $\eta_r \in C_c^\infty(B_r(0))$ normalized such that we have $\int_{B_r(0)} \eta_r(x)d^n x = 1$ and consider the convolution $u_r := \eta_r * u$ (setting $u(x) = 0$ for $x \notin U$). Then $u_r \in C^\infty(\mathbb{R}^n)$ since $\partial_\alpha(\eta_r * u) = (\partial_\alpha \eta_r) * u$. Moreover, if $\bar{B}_r(x) \subseteq U$ we have

$$\begin{aligned}u_r(x) &= \int_{B_r(0)} \eta_r(y)u(x-y)d^n y = \int_0^r \eta_r(s)s^{n-1}\int_{S^{n-1}} u(x-s\omega)d\sigma^{n-1}(\omega)ds \\ &= nV_n\int_0^r \eta_r(s)s^{n-1}u(x)ds = u(x).\end{aligned}$$

In particular, this holds for every x with $\operatorname{dist}(x, \partial U) > r$ and hence u is smooth on this set. Since $r > 0$ is arbitrary, the claim follows. $\square$

We leave it to the reader familiar with Lebesgue integration to check that the same proof works if we require u to be merely locally integrable instead of continuous in the definition of a harmonic function.

Since continuity as well as the mean value property are preserved under uniform limits we get:

Theorem 5.3. *Let u_k be a sequence of harmonic functions in U which converges uniformly to some function u on compact subsets of U. Then u is harmonic on U.*

Now we can show that the harmonic functions are precisely the solutions of the Laplace equation.

Theorem 5.4 (Gauss–Koebe[4])**.** *A function u is harmonic in U if and only if it solves the Laplace equation in U.*

Proof. Let $u \in C^2(U)$, fix $x \in U$ and set

$$\phi(r) := \frac{1}{nV_n}\int_{S^{n-1}} u(x+r\omega)d\sigma^{n-1}(\omega)$$

[4]Paul Koebe (1882–1945), German mathematician

for $r < r_0 := \mathrm{dist}(x, \partial U)$. Then $\phi \in C([0,r_0)) \cap C^1(0,r_0)$ with derivative given by

$$\phi'(r) = \frac{1}{nV_n} \int_{S^{n-1}} \nabla u(x + r\omega) \cdot \omega \, d\sigma^{n-1}(\omega)$$

and

$$\phi(0) = \lim_{r \downarrow 0} \phi(r) = \frac{1}{nV_n} \int_{S^{n-1}} u(x) d\sigma^{n-1}(\omega) = u(x).$$

Moreover, observe that the integrand is a normal derivative and hence using Green's first identity we obtain

$$\phi'(r) = \frac{1}{nV_n r^{n-1}} \int_{\partial B_r(x)} \frac{\partial u}{\partial \nu}(y) dS(y) = \frac{1}{nV_n r^{n-1}} \int_{B_r(x)} \Delta u(y) d^n y.$$

Consequently ϕ is constant if u is a solution of the Laplace equation and hence $\phi(r) = \phi(0) = u(x)$ implies that u is harmonic. Conversely, if u is harmonic we have $\phi(r) = u(x)$ implying $0 = \phi'(r) = \frac{1}{nV_n r^{n-1}} \int_{B_r(x)} \Delta u(y) d^n y$ for any ball $\bar{B}_r(x) \subset U$. Hence $\Delta u(y) = 0$ for all $y \in U$. □

As an immediate consequence one obtains the strong maximum principle

Theorem 5.5 (Strong maximum principle)**.** *Suppose $U \subseteq \mathbb{R}^n$ is connected and u is harmonic in U. If u attains its maximum in U, then u is constant.*

Proof. Suppose $u(x_0) = M := \sup_{x \in U} u(x)$ for some $x_0 \in U$. Let $\bar{B}_r(x_0) \subset U$ and observe that (5.7) implies that in fact $u(x) = M$ on $B_r(x_0)$ (if it were strictly smaller at some point, it would be smaller on a neighborhood by continuity and thus the whole integral would be smaller). Hence the set $\{x \in U | u(x) = M\}$ is open. By continuity of u this set is also (relatively) closed in U. Hence it must be all of U since U is assumed connected. □

Note that connectedness is crucial since if U consists of (at least) two components, we can choose (for example) two different constants on the two components to see that Theorem 5.5 fails in such a situation.

Since $-u$ satisfies the same assumptions we also have a corresponding minimum principle, that is, Theorem 5.5 holds if maximum is replaced by minimum. The following version is also frequently used.

Corollary 5.6 (Maximum principle)**.** *Suppose U is bounded and $u \in C(\overline{U})$ is harmonic. Then u attains its maximum on the boundary:*

$$\max_{x \in \overline{U}} u(x) = \max_{x \in \partial U} u(x). \tag{5.8}$$

Proof. Since $\overline{U}$ is compact, the Weierstrass theorem implies that u attains its maximum on $\overline{U}$. Now if U is connected and the maximum is attained at an interior point, then u is constant by Theorem 5.5 and hence the maximum is also attained on the boundary. Hence the claim holds if U is connected. In the general case we can write $U = \dot\bigcup_j U_j$ as a disjoint union of its connected components U_j. Note that there can be at most countably many. By the first part we have $\sup_{x\in U_j} u(x) = \max_{x\in\partial U_j} u(x)$ for every connected component and since $\partial U_j \subseteq \partial U$ the claim follows. $\square$

Again we have a corresponding minimum principle. Moreover, the assumption that U is bounded is crucial. Indeed, if $U = \mathbb{R}^n$, then $\partial U = \emptyset$ and there is no way the boundary values can control the maximum of u. But even if we have some boundary, this is not sufficient. Consider for example $u(x) = x_1 x_2$ which vanishes on the boundary of $U = \{x \in \mathbb{R}^2 | x_1 > 0\}$. If we add a condition to control the behavior at ∞, we can also cover this case:

Corollary 5.7. *Let u be harmonic on U with*

$$\limsup_{k\to\infty} u(x_k) \le M \tag{5.9}$$

for every sequence $x_k \in U$ converging either to a point in ∂U or to ∞. Then $u \le M$ on U.

Here a sequence will be said to converge to ∞ if $|x_k| \to \infty$.

Proof. Pick a sequence y_k such that $u(y_k) \to M_0 := \sup_{x\in U} u(x)$. If y_k is unbounded, it has a subsequence converging to ∞ and hence $M_0 \le M$ by assumption. Otherwise y_k is bounded and has subsequence converging to some point x. If $x \in \partial U$ we obtain again $M_0 \le M$. Otherwise $x \in U$ with $u(x) = M_0$. Hence $u = M_0$ on the connected component U_0 containing x and we can find a sequence $x_k \in U_0$ converging either to $\partial U_0 \subseteq \partial U$ (if this set is nonempty) or to ∞ and thus again $M_0 \le M$. $\square$

Of course we get a corresponding minimum principle by reversing the inequalities and replacing lim sup by lim inf.

Applying the mean value property to $\partial_j u$ and invoking the Gauss–Green theorem we get

$$\partial_j u(x) = \frac{1}{V_n r^n} \int_{B_r(x)} (\partial_j u) d^n y = \frac{1}{V_n r^n} \int_{\partial B_r(x)} u \nu_j dS \tag{5.10}$$

and hence we obtain (using the Jensen[5] inequality)

$$|\nabla u(x)| \le \frac{n}{r} \max_{\partial B_r(x)} |u|. \tag{5.11}$$

In fact, as long as we stay a fixed distance away from the boundary, we can get a uniform estimate by taking the sup over the entire domain

$$\sup_{U_r} |\nabla u| \le \frac{n}{r} \sup_U |u|, \qquad U_r := \{x \in U | \operatorname{dist}(x, \partial U) \ge r\}. \tag{5.12}$$

[5]Johan Jensen (1859–1925)), Danish mathematician and engineer

Iterating this result on layers with equal distance gives

Lemma 5.8. *Suppose u is harmonic in U. Then*

$$\sup_{U_r} |\partial^\alpha u| \le \left(\frac{n|\alpha|}{r}\right)^{|\alpha|} \sup_U |u|. \tag{5.13}$$

In particular, applying this on a ball $\bar{B}_r(x) \subset U$ we get

$$|\partial^\alpha u(x)| \le \left(\frac{n|\alpha|}{r}\right)^{|\alpha|} \max_{\partial B_r(x)} |u|. \tag{5.14}$$

Proof. Indeed, let $m := |\alpha|$ and let $\partial^\alpha = \partial_{j_1} \cdots \partial_{j_m}$. Then

$$\begin{aligned} |\partial^\alpha u(x)| &\le \frac{nm}{r} \sup_{U_{r(1-1/m)}} |\partial_{j_2} \cdots \partial_{j_m} u| \le \left(\frac{nm}{r}\right)^2 \sup_{U_{r(1-2/m)}} |\partial_{j_3} \cdots \partial_{j_m} u| \\ &\le \cdots \le \left(\frac{nm}{r}\right)^m \sup_U |u|. \end{aligned}$$

□

As a consequence we get

Theorem 5.9 (Liouville)**.** *Suppose u is harmonic on $\mathbb{R}^n$ and satisfies*

$$\lim_{|x|\to\infty} \frac{u(x)}{|x|^{m+1}} = 0 \tag{5.15}$$

for some $m \in \mathbb{N}_0$. Then u is a polynomial of degree at most m.

Proof. Fix $x \in \mathbb{R}^n$ and note that our assumption implies

$$\lim_{r\to\infty} \frac{1}{r^{m+1}} \max_{\partial B_r(x)} |u| = 0.$$

Letting $r \to \infty$ in (5.14) we get $\partial^\alpha u(x) = 0$ for $|\alpha| > m$. Hence u is a polynomial of degree at most m by Taylor's theorem. □

Of course the classical version, that a bounded harmonic function on $\mathbb{R}^n$ is constant, follows immediately from the $m = 0$ case.

As another consequence we get that harmonic functions are **real analytic**, that is, they can be expanded into absolutely convergent power series in the neighborhood of any point from U.

Theorem 5.10. *A harmonic function $u : U \to \mathbb{R}$ is real analytic in U.*

Proof. We will show that u is given by its Taylor series

$$u(x) = \sum_{\alpha\in\mathbb{N}_0^n} \frac{\partial^\alpha u(x_0)}{\alpha!} (x - x_0)^\alpha.$$

To see this we will compare with the series (using the multinomial theorem)

$$\frac{1}{1-(x_1+\cdots+x_n)} = \sum_{k=0}^{\infty}(x_1+\cdots+x_n)^k = \sum_{k=0}^{\infty}\sum_{|\alpha|=k}\frac{k!}{\alpha!}x^\alpha = \sum_{\alpha}\frac{|\alpha|!}{\alpha!}x^\alpha$$

which is obviously convergent for $|x_1|+\cdots+|x_n|<1$. Choosing a ball $\bar{B}_{2r}(x_0)\subset U$ the estimate (5.13) gives

$$\max_{x\in\bar{B}_r(x_0)}|\partial^\alpha u(x)| \le M\left(\frac{n|\alpha|}{r}\right)^{|\alpha|} \le M\left(\frac{n\mathrm{e}}{r}\right)^{|\alpha|}|\alpha|!, \qquad M:=\max_{\partial B_{2r}(x_0)}|u|,$$

since $\frac{k^k}{k!}\le \mathrm{e}^k$ (cf. Problem 5.4). This shows that our Taylor series converges absolutely whenever $|x_1-x_{0,1}|+\cdots+|x_n-x_{0,n}|<\frac{r}{n\mathrm{e}}$ as well as that the remainder converges to zero and hence u is given by its Taylor series. $\square$

As an important consequence recall the unique continuation principle of real analytic functions (Theorem A.4).

Corollary 5.11 (Unique continuation principle)**.** *Let U be connected and u harmonic on U. If all derivatives of u vanish at some point in U (e.g., u vanishes in a neighborhood of this point), then u vanishes on U.*

Finally we look at the case of positive harmonic functions. The striking fact is that the ratio between the maximum and the minimum on every compact subset is bounded by a constant depending only on the subset!

Theorem 5.12 (Harnack inequality[6])**.** *Suppose U is connected. Then for every compact set $K\subset U$ there is a constant C such that*

$$\frac{1}{C}\le\frac{u(y)}{u(x)}\le C, \qquad x,y\in K, \tag{5.16}$$

for all positive harmonic functions u in U.

Proof. Since we can exchange the roles of x and y it suffices to establish the upper bound. The key observation is that for given x,y with $|x-y|<r<\frac{1}{2}\operatorname{dist}(x,\partial U)$ (such that $B_r(y)\subset B_{2r}(x)\subset U$) we have

$$u(y)=\frac{1}{|B_r|}\int_{B_r(y)}u\,d^nz \le \frac{2^n}{|B_{2r}|}\int_{B_{2r}(x)}u\,d^nz = 2^n u(x)$$

for all positive harmonic functions on U. Moreover, by the triangle inequality $\operatorname{dist}(y,\partial U)\le|x-y|+\operatorname{dist}(x,\partial U)$ and hence $|x-y|<r<\frac{1}{3}\operatorname{dist}(y,\partial U)$ implies $|x-y|<r<\frac{1}{2}\operatorname{dist}(x,\partial U)$ and the above inequality still holds.

Now set

$$S(x,y):=\sup\left\{\frac{u(y)}{u(x)}\,\middle|\, u \text{ is positive and harmonic on } U\right\}$$

[6] Carl Gustav Axel Harnack (1851–1888), Baltic German mathematician

and let $F_x := \{y \in U | S(x, y) < \infty\}$. Since $S(x, x) = 1$ this set is clearly nonempty. It is also open, since if $y \in F_x$, then our above estimate shows $S(x, z) \le 2^n S(x, y)$ for all z with $|z - y| < r < \frac{1}{2}\,\mathrm{dist}(y, \partial U)$. Hence F_x is open. Moreover, F_x is also closed. Indeed, let $y \in \overline{F_x}$ and choose $z \in F_x$ with $|z - y| < r < \frac{1}{3}\,\mathrm{dist}(y, \partial U)$. Then $S(x, y) \le 2^n S(x, z)$ implying $y \in F_x$. Consequently $F_x = U$ since U is connected.

Finally, for every $(x, y) \in K \times K$ there is a neighborhood $B_r(x) \times B_s(y)$ (with $r < \frac{1}{3}\,\mathrm{dist}(x, \partial U)$ and $s < \frac{1}{2}\,\mathrm{dist}(y, \partial U)$) such that $S(\tilde{x}, \tilde{y}) \le 4^n S(x, y)$ for all $(\tilde{x}, \tilde{y})$ in this neighborhood. By compactness finitely many of these neighborhoods cover $K \times K$ and hence S is bounded on $K \times K$ as desired. □

Note that if we multiply (5.16) by $u(x)$, then the claim holds for nonnegative harmonic functions. However, if a nonnegative harmonic function attains 0 on U, it vanishes identically by the strong maximum principle.

As an application we show

Theorem 5.13 (Harnack principle). *Suppose U is connected and let u_k be an increasing sequence of harmonic functions in U. Then either u_k converges uniformly on compact subsets of U to a harmonic function u in U or $u_k(x) \to \infty$ for all $x \in U$.*

Proof. If $u(x) := \lim_{k\to\infty} u_k(x)$ is finite for some $x \in U$, then applying Harnack's inequality to the nonnegative function $u_k - u_j$, $k \ge j$, implies for any compact set $K \subset U$ that

$$u_k(y) - u_j(y) \le C(u_k(x) - u_j(x)), \qquad y \in K, \quad j \le k,$$

where C is the Harnack constant for $K \cup \{x\}$. Hence u_k converges uniformly on compact subsets and its limit is harmonic by Theorem 5.3.

Otherwise, if $\lim_{k\to\infty} u_k(x) = \infty$ we can choose $K = \{x, y\}$ and invoke again Harnack's inequality to obtain $u_k(y) - u_1(y) \ge C(u_k(x) - u_1(x))$, which finishes the proof. □

This is only the tip of the iceberg. If you want to find out more about harmonic functions I warmly recommend [**3**].

Problem 5.1. *Show that for $f \in C(\mathbb{R}^n)$ and $x \in \mathbb{R}^n$ we have*

$$\lim_{r\downarrow 0} \frac{1}{V_n r^n} \int_{B_r(x)} f(y)\, d^n y = \lim_{r\downarrow 0} \frac{1}{n V_n r^{n-1}} \int_{\partial B_r(x)} f(y)\, dS(y) = f(x).$$

Problem 5.2. *Find all integrable harmonic functions u on $\mathbb{R}^n$.*

Problem 5.3. *Show that zeros of harmonic functions are never isolated.*

Problem* 5.4. *Establish the bounds*

$$\mathrm{e}\left(\frac{n}{\mathrm{e}}\right)^n \le n! \le \frac{\mathrm{e}^2}{4}\left(\frac{n+1}{\mathrm{e}}\right)^{n+1}$$

for the factorial of integers. (Hint: Take logarithms and estimate the sums by integrals.)

Problem 5.5. *Show that a nonnegative harmonic function on $\mathbb{R}^n$ is constant. Moreover, a harmonic function on $\mathbb{R}^n$ is constant if it is bounded from above or from below. (Hint: Fix two points x, y and note that $B_r(x) \subset B_{r+d}(y)$ for $d := |x-y|$.)*

Problem 5.6. *Find all harmonic functions u on $\mathbb{R}^2$ such that $u(x,y) \le x^2 - y^2$.*

Problem 5.7. *Find all harmonic functions u on $\mathbb{R}^2$ such that $u_x(x,y) < u_y(x,y)$.*

Problem 5.8 (differential Harnack inequality). *Suppose u is a positive harmonic function on $B_r(x)$. Show*

$$|\nabla u(x)| \le \frac{n}{r} u(x).$$

This gives another proof that a positive harmonic function on $\mathbb{R}^n$ is constant.

Problem 5.9. *An eigenfunction of $-\Delta$ is a function $u \in C^2(U)$ satisfying*

$$-\Delta u = \lambda u, \qquad \lambda \in \mathbb{R}.$$

Show that u is real analytic. (Hint: Consider the function $v(x_1, \dots, x_{n+1}) = u(x_1, \dots, x_n)\mathrm{e}^{\sqrt{\lambda} x_n}$.)

5.2. Subharmonic functions

We begin with the simple observation, that the proof of the maximum principle does not require the full mean value property, but an inequality will suffice. This suggests the following extension:

A continuous function v is called **subharmonic** in U if it satisfies the **submean property**

$$v(x) \le \frac{1}{nV_n} \int_{S^{n-1}} v(x + r\omega) d\sigma^{n-1}(\omega) \tag{5.17}$$

whenever $r \le r_0(x) < \operatorname{dist}(x, \partial U)$. The fact that we require the submean property only for sufficiently small balls will simplify some arguments later on. Moreover, the submean property for all $r < \operatorname{dist}(x, \partial U)$ will then follow automatically, as we will see below.

Example 5.1. In one dimension the harmonic functions are precisely the linear functions and the subharmonic functions are the convex functions. In this respect observe that the submean property in one dimension amounts to midpoint-convexity, $v(\frac{x+y}{2}) \le \frac{v(x)+v(y)}{2}$, and Lemma 5.16 below will establish convexity. ◇

One calls v **superharmonic** if the inequality in the definition is reversed, that is, if $-v$ is subharmonic. We will formulate the following simple facts only for the case of subharmonic functions and urge the reader to find the corresponding formulation for superharmonic functions.

Inspecting the proofs for harmonic functions one infers:

Lemma 5.14. *We have:*

- *A function $v \in C^2(U)$ is subharmonic if and only if $\Delta v \ge 0$.*
- *A subharmonic function satisfies the strong maximum principle, that is, both Theorem 5.5 and Corollary 5.6 hold for subharmonic functions.*

Of course there is a price we have to pay since a subharmonic function will *not* satisfy the minimum principle. On the other hand subharmonic functions are much more flexible:

Lemma 5.15. *The following constructions give again subharmonic functions:*

- *If v is subharmonic, so is $\varphi(v)$ for every convex and non-decreasing function $\varphi \in C(\mathbb{R})$.*
- *If v_1, v_2 are subharmonic, so is $\max(v_1, v_2)$ and $\alpha_1 v_1 + \alpha_2 v_2$ if $\alpha_1, \alpha_2 \ge 0$.*

Proof. The first item follows from the fact that convex functions are continuous and Jensen's inequality:

$$\varphi(v(x)) \le \varphi\Big(\frac{1}{|\partial B_r|}\int_{\partial B_r(x)} v\, dS\Big) \le \frac{1}{|\partial B_r|}\int_{\partial B_r(x)} \varphi(v)\, dS.$$

The second is immediate from the definition. □

Moreover, the following simple fact sheds some further light on the naming:

Lemma 5.16 (Subharmonic functions are subsolutions)**.** *If $u \in C(\overline{U})$ is harmonic and $v \in C(\overline{U})$ subharmonic, then $v \le u$ on ∂U implies $v \le u$ on all of U.*

Proof. Apply the maximum principle to the subharmonic function $v-u$. □

Conversely, if for a ball $\bar{B}_r(x) \subset U$ we have $v \le u$ on $B_r(x)$, where u is the harmonic function which coincides with v on $\partial B_r(x)$ (given by the Poisson integral, to be established in Theorem 5.25 below), then

$$v(x) \le u(x) = \frac{1}{nV_n} \int_{S^{n-1}} v(x + r\omega) d\sigma^{n-1}(\omega). \tag{5.18}$$

Hence the property from Lemma 5.16 uniquely characterizes subharmonic functions. Incidentally this also shows that the submean property holds for all balls $\bar{B}_r(x) \subseteq U$.

Finally, note that subharmonic functions are not only far more flexible, but it is also easy to find concrete examples and invoke Lemma 5.16 to obtain useful bounds.

Example 5.2. Let $u \in C^2(B_1(0)) \cap C(\bar{B}_1(0))$. Then

$$u(x) \le \max_{\partial B_1(0)} u + \frac{\sup_{B_1(0)}(-\Delta u)}{2n}(1 - |x|^2).$$

To see this denote the right-hand side of the inequality by v and note that $\Delta v = -\sup_{B_1(0)}(-\Delta u)$. Hence $u - v$ is subharmonic and hence $u - v \le 0$ as this holds on the boundary.

Note that applying the same argument to $-u$ shows

$$\max_{B_1(0)} |u| \le \max_{\partial B_1(0)} |u| + \frac{1}{2n} \sup_{B_1(0)} |\Delta u|. \qquad \diamond$$

Problem 5.10. *Show that if u is harmonic, then $\varphi(u)$ is subharmonic for every convex function $\varphi \in C(\mathbb{R})$.*

Problem 5.11. *Show that v is subharmonic if and only if*

$$v(x) \le \frac{1}{|B_r|} \int_{B_r(x)} v(y) d^n y$$

for every ball with sufficiently small radius $r \le r_0(x) < \operatorname{dist}(x, \partial U)$. Moreover, in this case this holds for all $r < \operatorname{dist}(x, \partial U)$.

5.3. The Newton potential and the Poisson equation on $\mathbb{R}^n$

We now look at the inhomogeneous problem first studied by Poisson (a student of Laplace and Lagrange) and hence known as **Poisson equation**

$$-\Delta u = f. \tag{5.19}$$

We will consider the problem in $\mathbb{R}^n$ since this case can be solved explicitly. But at this point it is of course unclear how one should obtain such a formula. If you seek the advice of a physicist, you will get the hint to look for the solution of the special problem

$$-\Delta \Phi = \delta, \tag{5.20}$$

where δ is the Dirac delta function and then take the convolution to get the solution of the original problem:

$$u(x) := (\Phi * f)(x) := \int_{\mathbb{R}^n} \Phi(x-y) f(y) d^n y. \tag{5.21}$$

Indeed, a simple calculation verifies $-\Delta u = (-\Delta \Phi) * f = \delta * f = f$. While this might not make much sense from a rigorous point of view (at least at this stage, without the theory of distributions at our disposal), we can still try to follow this advice and see where it leads us.

So let us try to find the **fundamental solution** Φ first. Our friend from physics tells us that the δ function is zero on $\mathbb{R}^n \setminus \{0\}$ and since it is also radially symmetric, we look for a radial harmonic function which is allowed to be singular at the origin. So let us try to find all such functions and hope that the fundamental solution is among them. To this end we set $u(x) = \varphi(r)$, where $r = |x|$ and compute

$$\Delta u = \varphi'' + \frac{n-1}{r} \varphi' = 0, \qquad r > 0. \tag{5.22}$$

This ordinary differential equation is of Euler type and its solution is easily seen to be

$$\varphi(r) = \begin{cases} a \log(r) + b, & n = 2, \\ \frac{a}{r^{n-2}} + b, & n \geq 3. \end{cases} \tag{5.23}$$

So fortunately this does not leave us much choice for Φ and hence we define

$$\Phi(x) := \begin{cases} -\frac{1}{2\pi} \log(|x|), & n = 2, \\ \frac{1}{n(n-2)V_n |x|^{n-2}}, & n \geq 3, \end{cases} \tag{5.24}$$

to be the fundamental solution of the Laplace equation. The choice of the normalization constants will be justified in our theorem below. Note that we have $\Phi \in C^\infty(\mathbb{R}^n \setminus \{0\})$ with

$$\begin{aligned} \partial_j \Phi(x) &= -\frac{1}{nV_n} \frac{x_j}{|x|^n}, \\ \partial_j \partial_k \Phi(x) &= -\frac{1}{nV_n} \left(\frac{\delta_{j,k}}{|x|^n} - n \frac{x_j x_k}{|x|^{n+2}} \right). \end{aligned} \tag{5.25}$$

Moreover, we have the estimates

$$|\partial_j \Phi(x)| \leq \frac{1}{nV_n |x|^{n-1}}, \qquad |\partial_j \partial_k \Phi(x)| \leq \frac{1}{V_n |x|^n}. \tag{5.26}$$

In particular, Φ and $\partial_j \Phi(x)$ are locally integrable while $\partial_j \partial_k \Phi(x)$ is not.

The function u defined in (5.21) is called the **Newton potential** associated with f. Recall that in classical mechanics a conservative force field can be written as the negative gradient of a potential field. According to Newton's theory of gravitation, the fundamental solution in $\mathbb{R}^3$ is (up to physical constants) the gravitational potential of a point mass, while u is the potential of a mass distribution given by the density f. Similarly, in electrostatics, the fundamental solution is (up to physical constants) the electrostatic potential of a point charge, while u is interpreted as the electrostatic potential of the charge density f.

Newton was the first to observe that the fundamental solution is harmonic away from the origin. Accordingly one obtains that the Newton potential is harmonic away from the support of f. To show this we record:

Lemma 5.17. *Suppose $f(x,y): U \times V \subseteq \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ is harmonic with respect to $x \in U$ for every $y \in V$ and integrable with respect to $y \in V$ for every $x \in U$. Then if $x \mapsto \int_V |f(x,y)| d^m y$ is locally integrable,*

$$\int_V f(x,y) d^m y \tag{5.27}$$

is harmonic in U.

Proof. Use Fubini to verify that the mean value property holds. □

Corollary 5.18. *Suppose $n = 2$ and $f(y)\log(\mathrm{e} + |y|)$ is integrable or $n \ge 3$ and $f(y)(1+|y|)^{-n+2}$ is integrable. Then the Newton potential of f is harmonic away from the support of f.*

Proof. Abbreviate $g(y) := \log(\mathrm{e} + y)$ for $n = 2$ and $g(y) := (1+y)^{-n+2}$ for $n \ge 3$. Then

$$\int_{B_r(x_0)} |\Phi(x-y)| d^n x \le C_r g(|x_0 - y|) \le C_{r,x_0} g(|y|).$$

To see the first estimate we can chose $x_0 = 0$ without loss of generality. Now if $|y| \le r+1$ we can estimate the integral by $\int_{B_{2r+1}(0)} |\Phi(x)| d^n x$. And if $|y| > r+1$ we can estimate the integrand (up to a constant) by $g(|y| + r - \mathrm{e})$ if $n = 2$ and $g(|y| - r - 1)$ if $n \ge 3$. Since $\frac{g(|y|+r-\mathrm{e})}{g(|y|)}$ for $n = 2$ and $\frac{g(|y|-r-1)}{g(|y|)}$ for $n \ge 3$ is bounded for $|y| > r+1$, the first estimate follows. The last argument also shows the second one.

Thus

$$\int_{B_r(x_0)} \int_{\mathbb{R}^n} |\Phi(x-y)||f(y)| d^n y d^n x \le C_{r,x_0} \int_{\mathbb{R}^n} g(|y|)|f(y)| d^n y$$

is finite and we can apply the previous lemma. □

That the Newton potential solves the Poisson equation was established by Gauss if f is C^1. After further investigations by Riemann, Dirichlet, and Clausius[7], Hölder established the following result:

[7] Rudolf Clausius (1822–1888), German physicist and mathematician

Theorem 5.19 (Hölder)**.** *Let U be a bounded domain and suppose $f \in C(U)$ is bounded and locally Hölder continuous. Set $f = 0$ for $\mathbb{R}^n \setminus U$. Then the Newton potential (5.21) of f is $C^1(\mathbb{R}^n) \cap C^2(U)$ and satisfies $-\Delta u = f$ on U. If $V \supseteq U$ is some domain where we can apply the Gauss–Green theorem (e.g. a large ball) the derivatives are given by*

$$\begin{aligned}(\partial_j u)(x) &= \int_U (\partial_j \Phi)(x-y) f(y) d^n y,\\ (\partial_k \partial_j u)(x) &= \int_V (\partial_k \partial_j \Phi)(x-y)\big(f(y) - f(x)\big) d^n y\\ &\qquad - f(x) \int_{\partial V} (\partial_j \Phi)(x-y) \nu_k(y) dS(y),\end{aligned} \tag{5.28}$$

where the first formula is valid for $x \in \mathbb{R}^n$ and the second for $x \in U$.

Proof. Denote the right-hand sides of the formulas for the claimed first and second derivatives of u by v_j and $w_{k,j}$, respectively. Let $|f| \le M$. We introduce a monotone cutoff function $\phi \in C^1(\mathbb{R})$ which is zero for $x \le 1$ and one for $x \ge 2$ with $0 \le \phi'(x) \le 2$. Set $\phi_\varepsilon(x) = \phi(|x|/\varepsilon)$ and note $|\partial_j \phi_\varepsilon(x)| \le \frac{2}{\varepsilon}$. Then

$$u_\varepsilon(x) := \int_U (\phi_\varepsilon \Phi)(x-y) f(y) d^n y, \quad u_{\varepsilon,j}(x) := \int_U (\partial_j \phi_\varepsilon \Phi)(x-y) f(y) d^n y$$

satisfy $u_\varepsilon \in C^2(\mathbb{R}^n)$ with $\partial_j u_\varepsilon = u_{\varepsilon,j}$. Moreover, taking $\varepsilon \to 0$ we have

$$\begin{aligned}|u_\varepsilon(x) - u(x)| &\le \int_{B_{2\varepsilon}(x)} ((1-\phi_\varepsilon)\Phi)(x-y)|f(y)| d^n y \le M \int_{B_{2\varepsilon}} \Phi(y) d^n y\\ &= \frac{M}{n-2} \int_0^{2\varepsilon} \frac{r^{n-1}}{r^{n-2}} dr = \frac{2M\varepsilon^2}{n-2},\end{aligned}$$

where we have assumed $n \ge 3$ for notational simplicity. The easy adaptions for the case $n = 2$ are left as an exercise. Hence $u_\varepsilon \to u$ uniformly. Similarly one has (again $n \ge 3$)

$$\begin{aligned}|u_{\varepsilon,j}(x) - v_j(x)| &\le \int_{B_{2\varepsilon}(x)} (\partial_j (1-\phi_\varepsilon)\Phi)(x-y)|f(y)| d^n y\\ &\le M \int_{B_{2\varepsilon}} (\frac{2}{\varepsilon}\Phi(y) + |\partial_j \Phi(y)|) d^n y \le \frac{2Mn\varepsilon}{n-2}\end{aligned}$$

and hence $u_{\varepsilon,j} \to v_j$ implying $u \in C^1(\mathbb{R}^n)$ with $\partial_j u = v_j$ as claimed.

For the second derivatives we proceed similarly and consider

$$v_{\varepsilon,j} := \int_V (\phi_\varepsilon \partial_j \Phi)(x-y) f(y) d^n y, \quad v_{\varepsilon,k,j} := \int_V (\partial_k \phi_\varepsilon \partial_j \Phi)(x-y) f(y) d^n y.$$

As before we have $\partial_k v_{\varepsilon,j} = v_{\varepsilon,k,j}$ and one checks $v_{\varepsilon,j} \to v_j$. However, when looking at $v_{\varepsilon,j,k}$, we run into the problem that the second derivatives of Φ

are no longer integrable near 0. Hence we need to cheat in an extra term to fix this:

$$\begin{aligned}
v_{\varepsilon,k,j}(x) &= \int_V (\partial_k\phi_\varepsilon\partial_j\Phi)(x-y)(f(y)-f(x))d^n y \\
&\qquad + f(x)\int_V (\partial_k\phi_\varepsilon\partial_j\Phi)(x-y)d^n y \\
&= \int_V (\partial_k\phi_\varepsilon\partial_j\Phi)(x-y)(f(y)-f(x))d^n y \\
&\qquad - f(x)\int_{\partial V} (\partial_j\Phi)(x-y)\nu_k(y)dS(y),
\end{aligned}$$

where we have used the Gauss–Green theorem and assumed that $2\varepsilon < \operatorname{dist}(x,\partial V)$ in the last step. Now since f is Hölder continuous we can use

$$|f(x)-f(y)| \le C|x-y|^\alpha$$

to obtain

$$\begin{aligned}
|v_{\varepsilon,k,j}(x) - w_{k,j}(x)| &\le C\int_{B_{2\varepsilon}} \big(|\partial_k\partial_j\Phi(y)| + \frac{2}{\varepsilon}|\partial_j\Phi(y)|\big)|y|^\alpha \\
&\le C\left(4+\frac{n}{\alpha}\right)(2\varepsilon)^\alpha.
\end{aligned}$$

Letting $\varepsilon \to 0$ establishes $u \in C^2(U)$ together with the formula for the second derivatives. Finally, to compute Δu note that the formula for $w_{k,j}$ remains unchanged if we replace V by any ball containing V (recall that $f = 0$ outside U) and hence we can replace it by $B_R(x)$ to obtain

$$\begin{aligned}
-\Delta u(x) &= f(x)\sum_{j=1}^n \int_{\partial B_R(x)} (\partial_j\Phi)(x-y)\nu_j(y)dS(y) \\
&= \frac{f(x)}{nV_nR^{n-1}}\int_{\partial B_R(0)} \nu(y)\cdot\nu(y)dS(y) = f(x). \qquad \square
\end{aligned}$$

Note that in combination with Corollary 5.18 we see that this is really a local result. That is, if we consider the Newton potential of a function f as in Corollary 5.18, then it will be C^2 in a neighborhood of some fixed point provided f is Hölder continuous in a neighborhood of this point. Moreover, if $f \in C^k$ and the highest derivatives satisfy the assumptions of the previous theorem, then $u \in C^{k+2}$. In fact, as just pointed out this is a local result and hence one can assume that f has compact support and use $\partial_j\Phi * f = \Phi * (\partial_j f)$.

While the above theorem seems quite natural, it still leaves the question to what extend the Hölder condition can be improved. In fact, one can replace it by any modulus of continuity which ensures integrability of the first term in the formula for the second derivative. However, it was an open question for quite some time if continuity alone is also sufficient. Petrini[8] was the first to come up with a counterexample (his original example can be found in [**36**, Satz 4.3.1], we present a more explicit one).

Example 5.3. Consider

$$u(x,y) := (x^2 - y^2)\log(\log(r^{-1})), \qquad r := \sqrt{x^2+y^2},$$

which is clearly in $C^\infty(B_{1/2}(0) \setminus \{0\}) \cap C(B_{1/2}(0))$ if we set $u(0,0) := 0$. Moreover, a straightforward computation shows

$$u_x(x,y) = 2x\log(\log(r^{-1})) + \frac{x(x^2-y^2)}{r^2\log(r)}$$

and since, by symmetry, $u_y(x,y) = -u_x(y,x)$, we conclude $u \in C^1(B_{1/2}(0))$ if we set $u_x(0,0) = u_y(0,0) = 0$. Next, we have

$$u_{xx}(x,y) = 2\log(\log(r^{-1})) + \frac{5x^2-y^2}{r^2\log(r)} - x^2(x^2-y^2)\frac{1+2\log(r)}{r^4\log(r)^2}$$

and this function is unbounded near 0 (take the limit along the diagonal $x = y$). Hence $u \notin C^2(B_{1/2}(0))$. However, again by symmetry, $u_{yy}(x,y) = -u_{xx}(y,x)$ and hence

$$\Delta u(x,y) = (x^2-y^2)\frac{1-4\log(r)}{r^2\log(r)^2} =: f(x,y) \in C(B_{1/2}(0))).$$

Now let $v := \Phi * f$ be the Newton potential of f. Then $u - v$ is harmonic on $B_{1/2}(0) \setminus \{0\}$ and continuous at 0. Hence by the removable singularity theorem for harmonic functions (Problem 5.34 below), $u - v$ is harmonic on $B_{1/2}(0)$ and the Newton potential of f is not C^2. Similarly we can conclude that the corresponding Poisson equation does not have a solution in C^2. $\diamond$

Corollary 5.20. *Suppose $n \ge 3$ and $f(y)(1+|y|)^{-n+2}$ is integrable, f vanishes at ∞ (i.e. $\lim_{|y|\to\infty} f(y) = 0$), and is locally Hölder continuous. Then the Newton potential solves the Poisson problem. Moreover, it is the only solution vanishing at ∞.*

Proof. Fix $r > 0$ and split $f = f_1 + f_2$ where $f_1 = \chi_{B_r(0)} f$. Then the Newton potential u_1 of f_1 is $C^2(B_r(0))$ and satisfies $-\Delta u_1 = f$ on $B_r(0)$ by Lemma 5.19 while the Newton potential u_2 of f_2 is harmonic on $B_r(0)$ by Corollary 5.18. Since r is arbitrary we see that the Newton potential $u = u_1 + u_2$ of f is $C^2(\mathbb{R}^n)$ and solves $\Delta u = f$.

To see $u(x) \to 0$ as $|x| \to \infty$ observe that $|u_1(x)| \le \frac{C}{(|x|-r)^{n-2})}$ for $|x| > r$ and $|u_2(x)| \le C\sup_{|y|>r}|f(y)|$. Hence $\limsup_{|x|\to\infty}|u(x)| \le C\sup_{|y|>r}|f(y)|$ and since r is arbitrary, we conclude $u(x) \to 0$ as $|x| \to \infty$.

[8]Henrik Petrini (1863–1957), Swedish mathematician

Finally, uniqueness follows from Liouville since the difference of two solutions is bounded and vanishes at ∞. $\square$

In two dimensions we still get a solution, but the Newton potential will in general not vanish at ∞ but grow logarithmically (Problem 5.14).

Problem 5.12. *What is the fundamental solution for $n = 1$? Under what conditions does Hölder's theorem hold?*

Problem 5.13. *Find a fundamental solution for the Helmholtz equation*

$$-\Delta u + u = f$$

in $n = 3$ dimensions. (Hint: You need a radial solution of the form $\Phi(x) = \frac{\varphi(|x|)}{4\pi|x|}$, with $\varphi(0) = 1$.)

Problem* 5.14. *Suppose $n = 2$ and f bounded with $f(y)\log(\mathrm{e} + |y|)$ integrable. Then the Newton potential solves the Poisson problem. Moreover, it is the only solution growing at most like $u(x) = o(|x|)$ as $x \to \infty$ up to a constant.*

Problem 5.15. *Suppose f is integrable with compact support. Then the Newton potential satisfies*

$$u(x) = C\Phi(x) + O(|x|^{-n+1})$$

as $|x| \to \infty$, where $C := \int_{\mathbb{R}^n} f(y)d^n y$. In particular, it is the only solution with these asymptotics. If in addition f is rotationally symmetric, then $u(x) = C\Phi(x)$ for x outside the support of f. (Hint: The inverse triangle inequality $||x| - |y|| \le |x - y|$ might be useful.)

Problem 5.16. *Show that the Newton potential of $\chi_{B_r(0)}$ is given by*

$$u(x) = \begin{cases} \frac{1}{2}\left(\frac{r^2}{n-2} - \frac{|x|^2}{n}\right), & |x| \le r, \\ \frac{r^2}{n(n-2)}\left(\frac{r}{|x|}\right)^{n-2}, & |x| \ge r, \end{cases}$$

for $n \ge 3$ and

$$u(x) = \begin{cases} -\frac{r^2}{2}\log(r) + \frac{r^2 - |x|^2}{4}, & |x| \le r, \\ -\frac{r^2}{2}\log(|x|), & |x| \ge r, \end{cases}$$

for $n = 2$. (Hint: Use the fact that u must be rotationally symmetric.)

Problem 5.17. *Show that the Newton potential of a continuous, rotationally symmetric function $f(x) = F(r)$, $r = |x|$, with $\int_1^\infty F(s)s\,ds < \infty$ for $n \ge 3$ and $\int_1^\infty \log(s)F(s)s\,ds < \infty$ for $n = 2$ is given by*

$$u(x) = \frac{1}{n-2}\int_0^\infty \min(1, s/r)^{n-2} F(s)s\,ds$$

for $n \geq 3$ *and*

$$u(x) = -\int_0^\infty \log(\max(s,r))F(s)s\,ds$$

for $n = 2$.

Problem 5.18. *A differentiable function* $K : \mathbb{R}^n \to \mathbb{C}$ *is called a (strong)* ***Calderón–Zygmund kernel***[9] *provided*

(i) $|K(x)| \leq \frac{C}{|x|^n}$, *for all* $x \in \mathbb{R}^n$,

(ii) $|\nabla K(x)| \leq \frac{C}{|x|^{n+1}}$,

(iii) $\int_{r<|x|<R} K(x) d^n x = 0$ *for all* $0 < r < R$.

Show that $(\partial_k \partial_j \Phi)(x)$ *is a Calderón–Zygmund kernel. Convolution with this kernel is known as* ***double Riesz transform***.[10]

Problem 5.19. *Suppose* f *is Hölder continuous and satisfies the assumptions of Corollary 5.18. Show that the Newton potential* u *is in* $C^2(\mathbb{R}^n)$ *and satisfies*

$$(\partial_j u)(x) = \int_{\mathbb{R}^n} (\partial_j \Phi)(x-y) f(y) d^n y,$$

$$(\partial_k \partial_j u)(x) = \lim_{\varepsilon \downarrow 0} \int_{\varepsilon < |x-y|} (\partial_k \partial_j \Phi)(x-y) f(y) d^n y - \frac{1}{n} f(x) \delta_{jk}.$$

(Hint: Item (iii) from the previous problem.)

5.4. The Poisson equation on a domain and Green's function

Next we want to look at the Poisson equation on a bounded domain U;

$$-\Delta u = f, \qquad u|_{\partial U} = g, \tag{5.29}$$

where $f \in C(U)$ and $g \in C(\partial U)$.

Example 5.4. In the special case, where $U := B_r(0)$ is a ball and $f := f_0$ and $g := g_0$ are constant, the solution is given by

$$u(x) = g_0 + \frac{f_0}{2n}(r^2 - |x|^2). \qquad \diamond$$

Note that the maximum principle implies uniqueness and using the above solution in combination with our comparison principle from Lemma 5.16 even gives us an **a priori bound**:

[9] Alberto Calderón (1920–1998), Argentinian mathematician
[9] Antoni Zygmund (1900–1992), Polish mathematician
[10] Marcel Riesz (1886–1969), Hungarian mathematician

Theorem 5.21. *The problem* (5.29) *has at most one solution* $u \in C^2(U) \cap C(\overline{U})$. *A solution satisfies*

$$\max_{\overline{U}} |u| \le \max_{\partial U} |g| + \frac{r^2}{2n} \sup_U |f|, \tag{5.30}$$

if U *is contained in a ball of radius* r.

Proof. The difference of two solutions is harmonic and vanishes on the boundary. Hence it is zero by the maximum principle. This establishes uniqueness.

To see the bound, note that after a translation we can assume $U \subseteq B_r(0)$. Consider the function $v(x) := G + \frac{F}{2n}(r^2 - |x|^2)$ with $G := \max_{\partial U} |g|$ and $F := \sup_U |f|$ and note that we have $\Delta v = -F$ as well as $v \ge G$ on ∂U since $r^2 - |x|^2 \ge 0$ for $x \in U \subseteq B_r(0)$. Hence $u - v$ is subharmonic on U and $u - v \le 0$ on ∂U implies $u \le v$ on $\overline{U}$ by Lemma 5.16. Applying the same argument to $-u$ establishes the claim. □

Again taking differences of solutions corresponding to different data shows that the solution depends continuously on f and g. In this sense the problem is **well-posed**.

It remains to investigate if there is a solution at all. Clearly the Newton potential associated with f will solve the differential equation, but it will not have the required boundary values in general. However, by subtracting the Newton potential this reduces our problem to the **Dirichlet problem** of finding a harmonic function with prescribed boundary values. In this respect recall that we have already found the solution for a two dimensional ball in Section 3.5. To find such a formula for arbitrary domains, we will play with **Green's second identity**

$$\int_U (u\Delta v - v\Delta u) d^n y = \int_{\partial U} \left(u \frac{\partial v}{\partial \nu} - v \frac{\partial u}{\partial \nu} \right) dS. \tag{5.31}$$

Of course, for this identity to be valid, we need to assume that the Gauss–Green theorem holds for U (e.g. U has a C^1 boundary) as well as $u, v \in C^2(\overline{U})$. Choosing $v = \Phi(x - .)$ equation (5.20) formally implies

$$u(x) = -\int_U \Phi(x-.)(\Delta u) d^n y - \int_{\partial U} \left(u \frac{\partial \Phi(x - .)}{\partial \nu} - \Phi(x - .) \frac{\partial u}{\partial \nu} \right) dS. \tag{5.32}$$

Here the first boundary integral over ∂U involving $\frac{\partial \Phi}{\partial \nu}$ is called the **double layer potential** (it is interpreted as the electrostatic potential of a dipole density on the surface ∂U) while the second one involving Φ is called the **single layer potential** (it is interpreted as the electrostatic potential of a charge density on the surface ∂U). The single layer potential is continuous when x crosses the boundary while the double layer potential will have a jump in general. In two dimensions the double layer potential is the Cauchy integral operator.

This already looks quite promising except for the fact that it not only involves the boundary values of u, but also of the normal derivative $\frac{\partial u}{\partial \nu}$. To get rid of this term, we invoke the fact that our argument still holds true if we add a harmonic function to $\Phi(x - .)$. Choosing this harmonic correction term in such a way that the resulting function vanishes on the boundary, we obtain the desired effect. Explicitly, let ϕ^x be the solution of the Dirichlet problem

$$\Delta \phi^x = 0, \qquad \phi^x|_{\partial U} = \Phi(x - .) \tag{5.33}$$

and define the **Green function** of U as

$$G(x, y) := \Phi(x - y) - \phi^x(y), \qquad x \neq y \in U \times \overline{U}. \tag{5.34}$$

We will say that the Green function exists if the above Dirichlet problem has a solution $\phi^x \in C^2(\overline{U})$ for all $x \in U$. In this case

$$K(x, y) := -\frac{\partial G(x, y)}{\partial \nu}, \qquad (x, y) \in U \times \partial U, \tag{5.35}$$

is called the **Poisson kernel** for U. Then we have

$$u(x) = -\int_U G(x, y) \Delta u(y) d^n y + \int_{\partial U} K(x, y) u(y) dS(y). \tag{5.36}$$

The second integral is known as **Poisson integral**. Of course our derivation of this formula is still a bit wacky as it relies on (5.20). However, this can be easily fixed.

Lemma 5.22. *Suppose U is a bounded domain for which the Gauss–Green theorem holds. If the Green function for U exists, then a function $u \in C^2(\overline{U})$ can be represented as* (5.36).

Proof. To make our above argument rigorous we apply Green's second identity with $v = G(x, .)$ on the domain $U \setminus B_\varepsilon(x)$ where $\varepsilon > 0$ is sufficiently small such that $\bar{B}_\varepsilon(x) \subset U$. This gives (note that $\partial(U \setminus B_\varepsilon(x)) = \partial U - \partial B_\varepsilon(x)$ when taking the orientation into account)

$$\begin{aligned} -\int_{U \setminus B_\varepsilon(x)} G(x, y) \Delta u(y) d^n y = &-\int_{\partial U} K(x, y) u(y) dS(y) \\ &+ \int_{\partial B_\varepsilon(x)} \left(K(x, y) u(y) - G(x, y) \frac{\partial u(y)}{\partial \nu} \right) dS(y). \end{aligned}$$

Taking $\varepsilon \to 0$ and using that Φ and hence $G(x,.)$ is integrable, we obtain

$$\begin{aligned} - \int_U G(x,y)\Delta u(y) d^n y + \int_{\partial U} K(x,y) u(y) dS(y) = \\ = \lim_{\varepsilon\to 0} \int_{\partial B_\varepsilon(x)} \left(K(x,y)u(y) - G(x,y)\frac{\partial u(y)}{\partial \nu}\right) dS(y) \\ = - \lim_{\varepsilon\to 0} \int_{\partial B_\varepsilon(0)} \left(\frac{\partial \Phi(y)}{\partial \nu} u(x+y) + \Phi(y)\frac{\partial u(x+y)}{\partial \nu}\right) dS(y). \end{aligned}$$

Now using $\frac{\partial\Phi(y)}{\partial\nu} = -\frac{1}{nV_n\varepsilon^{n-1}}$ we see that the first part of the integral gives

$$\lim_{\varepsilon\to 0} \frac{1}{nV_n\varepsilon^{n-1}} \int_{\partial B_\varepsilon(0)} u(x+y) dS(y) = u(x)$$

while $\lim_{\varepsilon\to 0}\Phi(\varepsilon)\varepsilon^{n-1} = 0$ shows that the second part of the integral vanishes in the limit. □

So in summary we have reduced problem (5.29) to establishing existence for the associated Dirichlet problem. In the next section we will look at some simple domains where the Green function can be computed explicitly. In this context note, that even if one has found the Green function, this does not automatically imply solvability of the Dirichlet problem. Indeed, (5.36) only gives us a necessary form of the solution whose existence has to be assumed in the outset. Hence one needs to verify that the Poisson integral is a harmonic function attaining the required boundary values.

A few important properties of the Green function and the Poisson kernel are collected below:

Lemma 5.23. *Suppose $U \subset \mathbb{R}^n$ is a bounded and connected domain. The Green function is symmetric $G(x,y) = G(y,x)$, $x,y \in \overline{U}\times\overline{U}$ and positive for $x \neq y \in U$. The Poisson kernel $K(x,y)$ is a nonnegative harmonic function for $x \in U$ and satisfies*

$$\int_{\partial U} K(x,y) dS(y) = 1. \tag{5.37}$$

Proof. To see symmetry we fix x, y and introduce the two functions $u(z) := G(x,z)$ and $v(z) := G(y,z)$. Applying Green's second identity on the domian $U \setminus (B_\varepsilon(x) \cup B_\varepsilon(y))$ gives

$$\int_{\partial B_\varepsilon(x)} \left(v\frac{\partial u}{\partial\nu} - u\frac{\partial v}{\partial \nu}\right) dS = \int_{\partial B_\varepsilon(y)} \left(u\frac{\partial v}{\partial\nu} - v\frac{\partial u}{\partial \nu}\right) dS.$$

Letting $\varepsilon \to 0$ on the left-hand side gives (using the same argument as in the proof of Lemma 5.22)

$$\lim_{\varepsilon\to 0}\int_{\partial B_\varepsilon(x)} \left(v\frac{\partial u}{\partial\nu} - u\frac{\partial v}{\partial \nu}\right) dS = \lim_{\varepsilon\to 0}\int_{\partial B_\varepsilon(0)} \frac{\partial\Phi(y)}{\partial\nu} v(x+y) dS = v(x)$$

and similarly the limit on the right equals $u(y)$. Consequently $G(y,x) = v(x) = u(y) = G(x,y)$.

To see positivity fix $x \in U$ and consider $u(y) := G(x,y)$ which is harmonic in $U \setminus \{x\}$. Since $\lim_{y\to x} u(y) = +\infty$ we can choose $\varepsilon > 0$ such that u is positive on $\bar{B}_\varepsilon(x) \setminus \{x\}$. Applying the strong minimum principle on $U \setminus \bar{B}_\varepsilon(x)$ we see that u is positive on $U \setminus \bar{B}_\varepsilon(x)$ as well.

Since G is positive within U and vanishes on ∂U, its outward pointing normal derivative must be nonpositive and hence $K(x,y) \ge 0$. To see the last claim choose $u = 1$ in the representation formula. □

The connectedness assumption is not essential and was only made to get positivity of G. The general case boils down to finding a Green function for every connected component. We then get the Green function for U by setting $G(x,y) = 0$ whenever x and y lie in different components. Moreover, note that $K(x,y)$ will in fact be positive at every point of the boundary which satisfies an interior sphere condition (i.e. if U contains a sphere which touches the boundary at the point under consideration). This will follow from the Hopf–Oleinik lemma (see Example 5.8).

Let me remark that similar considerations can be made for the associated **Neumann problem**

$$-\Delta u = f, \qquad \frac{\partial u}{\partial \nu}\Big|_{\partial U} = g, \tag{5.38}$$

where the normal derivative is prescribed on the boundary. Our considerations suggest that we should add a harmonic correction term $\psi^x(y)$ to $\Phi(x-y)$ such that the normal derivative of the resulting function vanishes at the boundary. However, since (5.32) holds for $G_N(x,y) = \Phi(x-y) - \psi^x(y)$ whenever $\psi^x \in C^2(\overline{U})$ is harmonic on U, we can apply this to the constant function $u = 1$ to obtain

$$1 = -\int_{\partial U} \left(\frac{\partial G_N(x,.)}{\partial \nu} \right) dS. \tag{5.39}$$

This shows that there is no choice for ψ^x which will make the normal derivative vanish on the boundary! Hence the next best option seems to require the normal derivative to be constant:

$$\frac{\partial G_N(x,.)}{\partial \nu} = -\frac{1}{|\partial U|}. \tag{5.40}$$

Equivalently, ψ^x should solve

$$-\Delta \psi^x = 0, \qquad \frac{\partial \psi^x}{\partial \nu}\Big|_{\partial U} = \frac{\partial \Phi(x-.)}{\partial \nu} + \frac{1}{|\partial U|}. \tag{5.41}$$

This then implies

$$u(x) = -\int_U G_N(x,.)(\Delta u)d^n y + \int_{\partial U} G_N(x,.)\frac{\partial u}{\partial \nu} dS + \frac{1}{|\partial U|}\int_{\partial U} u\, dS. \quad (5.42)$$

The last (constant) term in this representation reflects the fact that the solution of the Neumann problem is not unique since you can always add a constant. Moreover, applying Green's second identity with $v = 1$ shows that a necessary condition for the Neumann problem to be solvable is

$$\int_U f\, d^n y = -\int_{\partial U} g\, dS. \quad (5.43)$$

We end this section with the remark that again a convenient setting for the discussion of these issues here is operator theory, as discussed in Section 3.2. In fact, if we look at square integrable functions with the scalar product

$$\langle f, g\rangle := \int_U f^*(x) g(x) d^n x, \quad (5.44)$$

then Green's second identity (5.31) tells us that the Laplace operator $L := -\Delta$ subject to Dirichlet (or Neumann) boundary conditions is symmetric $\langle Lf, g\rangle = \langle f, Lg\rangle$. In fact, we have already computed its eigenfunctions in case of a rectangle and a disc in Section 3.4. Clearly, for a general domain U there is no hope to find the eigenfunctions explicitly. However, note that solving the Poisson problem (5.29) with vanishing boundary values (i.e. $g = 0$) amounts to inverting L, that is, the solution is $u = L^{-1} f$. So our findings in this section suggest that the inverse of L is an integral operator whose kernel is the Green function. Moreover, an integral operator on a bounded domain with a continuous kernel is compact and all claims made about the eigenvalues and eigenfunctions of a Sturm–Liouville problem in Section 3.2 would follow from the spectral theorem for compact operators (known as Hilbert–Schmidt[11] theorem). In particular, one can write

$$u(x) = L^{-1} f(x) = \sum_{n=1}^{\infty} \frac{\langle u_n, f\rangle}{E_n} u_n(x). \quad (5.45)$$

Hence a possible strategy for the Poisson problem is to establish compactness of L^{-1}. This approach hinges on the use of Sobolev spaces and the Rellich[12] compactness theorem; see Chapter 10.1.

Note that in the case of Dirichlet boundary conditions 0 is not an eigenvalue and hence L is injective. In the case of Neumann boundary conditions $u = 1$ is an eigenvalue and we cannot invert L in this case. However, symmetry implies that the range of L is orthogonal to the kernel of L and hence this

[11] David Hilbert (1862–1943), German mathematician

[11] Erhard Schmidt (1876–1959), Baltic German mathematician

[12] Franz Rellich (1906–1955), Austrian-German mathematician

explains the solvability condition found for the Neumann problem. Moreover, once you switch to the orthogonal complement of the kernel, you can proceed as before.

Problem 5.20. *Find the Green function for an interval $(a, b) \subseteq \mathbb{R}$.*

Problem 5.21. *Show the generalized mean value formula*

$$u(0) = \frac{1}{|B_r|} \int_{\partial B_r(0)} u(y) dS(y) - \int_{B_r(0)} \big(\Phi(y) - \Phi(r)\big) \Delta u(y) d^n y$$

whenever $u \in C^2(\bar{B}_r(0))$.

Problem 5.22. *Show that the Green function satisfies*

$$G(x, y) \le \Phi(x - y)$$

for $n \ge 3$. What about $n = 2$?

Problem 5.23. *Show that the Neumann problem (5.41) for ψ^x satisfies the solvability condition (5.43).*

Problem 5.24. *Let $U_1 \subseteq U_2$ be two bounded domains and G_1, G_2 be the corresponding Green functions. Show $G_1(x, y) \le G_2(x, y)$ for $x \neq y \in U_1$.*

Problem 5.25. *Let U be a bounded C^1 domain and let a partition of its boundary $\partial U = V_1 \uplus V_2$ be given. Show that solutions $u \in C^2(\overline{U})$ of the mixed Dirichlet/Neumann problem*

$$-\Delta u = f, \qquad u|_{V_1} = g_1, \quad \frac{\partial u}{\partial \nu}\Big|_{V_2} = g_2,$$

differ by at most a constant. Moreover, this constant is zero if V_1 is nonempty. (Hint: Green's first identity with both functions equal.)

Problem 5.26. *Let U be a bounded C^1 domain and $a \in C(\partial U)$. Show that solutions $u \in C^2(\overline{U})$ of the Robin problem*

$$-\Delta u = f, \qquad \Big(\frac{\partial u}{\partial \nu} + a u\Big)\Big|_{\partial U} = g,$$

differ by at most a constant if $a \ge 0$. Moreover, this constant is zero unless $a = 0$. (Hint: Green's first identity with both functions equal.)

Problem 5.27. *Let U be a bounded C^1 domain and $a \in C(\partial U)$. Show that the Laplace operator with (homogenous) Robin boundary conditions*

$$\Big(\frac{\partial u}{\partial \nu} + a u\Big)\Big|_{\partial U} = 0$$

is symmetric. That is, if $u, v \in C^2(\overline{U})$ both satisfy the boundary conditions, then

$$\int_U (u \Delta v) d^n y = \int_U (v \Delta u) d^n y.$$

5.5. The Dirichlet principle

In this section we will have a brief look at another approach for the Dirichlet problem (5.29). The Dirichlet problem means that we are looking for a solution of the Poisson equation

$$-\Delta u = f \tag{5.46}$$

among the admissible functions

$$\mathcal{A}_g := \{v \in C^2(\overline{U}) |\, v|_{\partial U} = g\}. \tag{5.47}$$

If $u \in \mathcal{A}_g$ is a solution, we can multiply $-\Delta u - f = 0$ by an admissible function v and integrate over U to obtain

$$0 = \int_U (-\Delta u - f) v\, d^n x = \int_U (\nabla u \cdot \nabla v - f v) d^n x - \int_{\partial U} g \frac{\partial u}{\partial \nu} dS, \tag{5.48}$$

where we have used Green's first identity. Now by Cauchy–Schwarz we have $\nabla u \cdot \nabla v \le \frac{1}{2}|\nabla u|^2 + \frac{1}{2}|\nabla v|^2$ with equality for $v = u$. Hence subtracting these two cases (v arbitrary and $v = u$) we obtain the **Dirichlet principle**

$$\int_U \big(\frac{1}{2}|\nabla u|^2 - f u\big) d^n x \le \int_U \big(\frac{1}{2}|\nabla v|^2 - f v\big) d^n x. \tag{5.49}$$

This suggests to find the solution by minimizing the Dirichlet functional

$$I(v) := \int_U \big(\frac{1}{2}|\nabla v|^2 - f v\big) d^n x \tag{5.50}$$

among all admissible functions.

We remark that the Dirichlet functional also arises directly when considering (e.g.) the deformation of a thin membrane under the pressure of some external force f. In this case I is interpreted as the energy (which the equilibrium of the membrane has to minimize according to physical principles) with the first summand giving the deformation energy and the second summand the potential energy.

Of course this raises the question if there could be other minimizers which do not solve our original problem. To this end we use the fact that if u is a minimizer, then $t \mapsto I(u + t\, w)$ with fixed $w \in \mathcal{A}_0$ (the functions which vanish at the boundary such that $u + t\, w \in \mathcal{A}_g$ for all $t \in \mathbb{R}$) must have a minimum at $t = 0$ and thus its **variational derivative**

$$\frac{d}{dt}\Big|_{t=0} I(u + t\, w) \stackrel{!}{=} 0, \qquad w \in \mathcal{A}_0, \tag{5.51}$$

must vanish. But $I(t)$ is a quadratic function in t and hence one easily obtains

$$0 = \frac{d}{dt}\Big|_{t=0} I(u + t\, w) = \int_U (\nabla u \cdot \nabla w - f w) d^n x = \int_U (-\Delta u - f) w\, d^n x \tag{5.52}$$

for all $w \in \mathcal{A}_0$. Consequently $-\Delta u - f = 0$ as desired.

This reduces our problem to the problem of finding a minimizer of the Dirichlet functional I. This approach was already suggested by Lord Kelvin[13] and Dirichlet, but it was only much later until Riemann (who coined the name Dirichlet problem in honor of his teacher) was able to solve this problem. However, Weierstrass pointed out that Riemann had not proven the existence of a minimizer and gave a counterexample (Problem 5.29) of a similar functional which had no minimizers. Moreover, Prym[14] found a continuous boundary datum g such that there is no solution with finite Dirichlet integral. Thus, the legitimacy of the Dirichlet principle was unclear for several decades until eventually Arzelà[15] and Hilbert (independently) were able to give a rigorous proof. Hilbert's approach is now known as the direct method in the calculus of variations and it motived him to formulate his 20th problem concerning existence of solutions of partial differential equations when the values on the boundary of the region are prescribed (which is now considered solved).

It turns out that $C^2(\overline{U})$ is not a natural space for these kind of problems witnessed by the fact that Theorem 5.19 fails for $f \in C(\overline{U})$. Moreover, Theorem 5.19 suggests that Hölder spaces should provide a better setting and this was indeed shown by Kellogg[16] and Schauder[17]. However, nowadays one usually looks for minimizers in some Sobolev spaces. While this simplifies the problem of establishing existence of a minimizer, which is called a weak solution in this context, it leaves the problem of checking that this weak solution is C^2 as a separate task. The functional analytic tools required for this approach are beyond our present scope and hence we will not pursue this idea now. We will return to this problem in Chapter 10.

Problem 5.28. *Show that the minimizer u (provided one exists) of*

$$I(v) := \int_U |\nabla v|^2 d^n x, \quad v \in \mathcal{A} := \{v \in C^2(\overline{U}) | v|_{\partial U} = 0, \int_U |v|^2 d^n x = 1\}$$

is an eigenfunction of the Laplace operator and the corresponding minimum is the eigenvalue. In particular, this must then be the smallest eigenvalue.

Problem 5.29 (Weierstrass). *Consider the problem to minimize*

$$I(v) := \int_{-1}^{1} (x\, v'(x))^2 dx$$

[13]William Thomson, 1st Baron Kelvin (1824–1907), British mathematical physicist and engineer

[14]Friedrich Prym (1841–1915), German mathematician

[15]Cesare Arzelà (1847–1912), Italian mathematician

[16]Oliver Dimon Kellogg (1878–1932), American mathematician

[17]Juliusz Schauder (1899-1943), Polish mathematician

among the admissible functions

$$\mathcal{A}_{a,b} := \{v \in C^2[-1,1] \,|\, v(-1) = a,\ v(1) = b\}. \tag{5.53}$$

Show that there is no minimizer if $a \neq b$. (Hint: Look at $\arctan(x/\varepsilon)$*.)*

5.6. Solution for a half space and for a ball

In the case U has a simple geometry it is possible to compute Green's function via a technique known as mirror charges. The idea is that the fundamental solution can be interpreted as the potential of a single point charge located at the origin. Now suppose we have a plane given and put our point charge on one side of this plane. Then, placing a negative charge at the mirror image with respect to this plane, they will cancel each other on the plane. Explicitly, choosing the half-space

$$\mathbb{R}^n_+ := \{x \in \mathbb{R}^n | x_n > 0\}, \tag{5.54}$$

the corresponding Green function is

$$G(x,y) = \Phi(x-y) - \Phi(\tilde{x}-y), \qquad \tilde{x} = (x_1, \dots, x_{n-1}, -x_n). \tag{5.55}$$

By construction $G(x,y) = 0$ for $(x,y) \in \mathbb{R}^n_+ \times \partial\mathbb{R}^n_+$ since $|x-y| = |\tilde{x}-y|$ if $y_n = 0$. Also note that the second singularity at $y = \tilde{x}$ is not within our domain $\mathbb{R}^n_+$ and hence $\Phi(\tilde{x} - .)$ is harmonic in $\mathbb{R}^n_+$.

While $\mathbb{R}^n_+$ is unbounded and thus does not quite fit into the framework of the previous section, one can still verify directly that G provides a solution of the associated Dirichlet problem. To this end note that (recall (5.25))

$$K(x,y) = -\frac{\partial G(x,y)}{\partial \nu} = G_{y_n}(x,y) = -2\Phi_{x_n}(x-y) = \frac{2x_n}{nV_n}\frac{1}{|x-y|^n}. \tag{5.56}$$

We will leave the details as an exercise (Problem 5.30):

Theorem 5.24. *Suppose $g \in C_b(\mathbb{R}^{n-1})$. Then*

$$u(x) = \frac{2x_n}{nV_n}\int_{\partial\mathbb{R}^n_+} \frac{g(y)}{|x-y|^n} d^{n-1}y \tag{5.57}$$

is a bounded harmonic function on $\mathbb{R}^n_+$ which is continuous up to the boundary and satisfies $u(y) = g(y)$ for $y \in \partial\mathbb{R}^n_+$.

The case of the unit ball is a bit more tricky. It turns out that we need to invert through the unit sphere:

$$\tilde{x} = \frac{x}{|x|^2}. \tag{5.58}$$

Then the Green function for the unit ball is given by

$$G(x,y) = \begin{cases} \Phi(y-x) - \Phi(|x|(y-\tilde{x})), & x \neq 0, \\ \Phi(y) - \Phi(1), & x = 0. \end{cases} \tag{5.59}$$

Indeed, one easily checks

$$|y-x|^2 = 1 - 2x\cdot y + |x|^2 = |x|^2|y-\tilde{x}|^2, \qquad y \in \partial B_1(0). \tag{5.60}$$

Moreover, another straightforward calculation verifies

$$\begin{aligned} K(x,y) &= -\frac{\partial G(x,y)}{\partial \nu} = -y\cdot\nabla\Phi(y-x) + y\cdot\nabla\Phi(|x|(y-\tilde{x}))|x| \\ &= \frac{1}{nV_n}\frac{1-|x|^2}{|x-y|^n}. \end{aligned} \tag{5.61}$$

Theorem 5.25 (Poisson integral)**.** *Suppose $g \in C(\partial B_1(0))$. Then*

$$u(x) = \frac{1-|x|^2}{nV_n}\int_{\partial B_1(0)} \frac{g(y)}{|x-y|^n} dS(y), \qquad x \in B_1(0), \tag{5.62}$$

is a bounded harmonic function which is continuous up to the boundary and satisfies $\lim_{x\to y} u(x) = g(y)$ for $y \in \partial B_1(0)$.

Proof. By Lemma 5.23 the Poisson kernel satisfies $\int_{\partial B_1(0)} K(x,y)dS(y) = 1$. Moreover, since $K(.,y)$ is harmonic (it is the derivative of a harmonic function) and uniformly bounded for $x \in B_r(0)$ with $r<1$, u is harmonic by Lemma 5.17. Hence it remains to verify continuity. To this end fix some $y_0 \in \partial B_1(0)$ and $\varepsilon > 0$. Choose a corresponding δ such that $|g(y)-g(y_0)| \le \varepsilon$ for $|y-y_0| \le \delta$ and set $M := \max_{\partial B_1(0)} |g|$. Moreover, note that for $|x-y_0| < \delta$ and $|y-y_0| \ge \delta$ the reverse triangle inequality implies

$$0 < K(x,y) = \frac{1}{nV_n}\frac{(|y_0|+|x|)(|y_0|-|x|)}{|(x-y_0)-(y-y_0)|^n} \le \frac{1}{nV_n}\frac{2|x-y_0|}{(\delta - |x-y_0|)^n}.$$

Hence using

$$u(x) - g(y_0) = \frac{1-|x|^2}{nV_n}\int_{\partial B_1(0)} \frac{g(y)-g(y_0)}{|x-y|^n} dS(y)$$

and splitting the integral into two regions according to $|y-y_0| \le \delta$ and $|y-y_0| \ge \delta$ shows

$$|u(x) - g(y_0)| \le \varepsilon + \frac{4M|x-y_0|}{(\delta - |x-y_0|)^n}$$

for $|x-y_0| < \delta$. Consequently we have $|u(x)-g(y_0)| \le 2\varepsilon$ for $|x-y_0| \le \min(\frac{\delta}{2}, (\frac{\delta}{2})^n\frac{\varepsilon}{4M})$ and the claim follows. □

By a simple scaling we see that the Poisson kernel for a ball of radius r is

$$K(x,y) = \frac{1}{nV_n}\frac{r^2-|x|^2}{r|x-y|^n}. \tag{5.63}$$

Problem* 5.30. *Show Theorem 5.24. (Hint: Mimic the proof of Theorem 5.25.)*

Problem 5.31. *Let G be the Green function of the unit ball. Compute $\int_{B_1(0)} G(x,y)d^n y$. (Hint: There is no need to do the integral; use Example 5.4.)*

Problem 5.32. *Establish the following qualitative version of Harnack's inequality for positive harmonic functions u on a ball $B_r(0)$:*

$$r^{n-2}\frac{r-|x|}{(r+|x|)^{n-1}}u(0) \le u(x) \le r^{n-2}\frac{r+|x|}{(r-|x|)^{n-1}}u(0).$$

5.7. The Perron method for solving the Dirichlet problem

Now we turn to the Dirichlet problem

$$\Delta u = 0, \qquad u|_{\partial U} = g, \tag{5.64}$$

for given $g \in C(\partial U)$ in a general bounded domain. The first complete solution for a rather general class of domains was given by Poincaré in 1890. Here we look at a more streamlined version due to Perron[18].

The starting point of the Perron method is the characterization of subharmonic functions as subsolutions from Lemma 5.16. In fact, consider the Perron family

$$\mathcal{S}(g) := \{v \in C(\overline{U}) | v \text{ is subharmonic and satisfies } v \le g \text{ on } \partial U\}. \tag{5.65}$$

Then, if our Dirichlet problem has a solution u, we necessarily have $v \le u$ for all $v \in \mathcal{S}(g)$ with equality attained since $u \in S(g)$. This suggests to define the Perron candidate

$$P(g)(x) := \sup_{v\in\mathcal{S}(g)} v(x), \qquad x \in \overline{U}. \tag{5.66}$$

If our Dirichlet problem has a solution, it will be necessarily given by $P(g)$. In particular, if U is a ball, $P(g)$ will be given by the Poisson integral. Moreover, note that $\mathcal{S}(g)$ is nonempty (it contains the constant function $\min g$) and that $P(g)$ is bounded since $\min g \le P(g)(x) \le \max g$. Our first aim is to show that $P(g)$ is harmonic. To this end we will need the following technique:

Lemma 5.26 (Harmonic lifting)**.** *Suppose v is subharmonic on U and let $\bar{B}_r(x_0) \subset U$. Then*

$$\tilde{v}(x) := \begin{cases} v(x), & x \in U \setminus B_r(x_0), \\ P(v|_{\partial B_r(x_0)})(x), & x \in B_r(x_0), \end{cases} \tag{5.67}$$

is again subharmonic with $v \le \tilde{v}$.

[18]Oskar Perron (1880–1975), German mathematician

Proof. Clearly $\tilde{v}$ is continuous and satisfies $v \le \tilde{v}$ by Lemma 5.16. To show that $\tilde{v}$ satisfies the submean property we distinguish three cases: If $x \in U \setminus \bar{B}_r(x_0)$ or $x \in B_r(x_0)$ there is nothing to do. In the case $x \in \partial B_r(x_0)$ it follows from $\tilde{v}(x) = v(x)$ and $v(y) \le \tilde{v}(y)$ on $\partial B_r(x)$. □

Now we are ready to show

Theorem 5.27 (Perron)**.** *$P(g)$ is harmonic on U.*

Proof. It suffices to show that $P(g)$ is harmonic on every ball within U. So fix a ball $\bar{B}_r(x_0) \subset U$. Choose a sequence $v_k \in \mathcal{S}(g)$ such that $v_k(x_0) \to P(g)(x_0)$. Replacing v_k by the harmonic lifting of $\max\{v_1, \dots, v_k\}$ we get an increasing sequence which is harmonic on $B_r(x_0) \subset U$. Of course we still have $v_k \in \mathcal{S}(g)$ and $v_k(x_0) \to P(g)(x_0)$. By Harnack's principle, Theorem 5.13, it converges uniformly on compact subsets of $B_r(x_0)$ to a harmonic function u on $B_r(x_0)$. Hence it remains to show $u = P(g)$ on $B_r(x_0)$.

One direction, $u \le P(g)$, comes for free. To see the other direction choose some $w \in \mathcal{S}(g)$. We need to show $w \le u$ on $B_r(x_0)$. To this end let w_k be the harmonic lifting of $\max\{w, v_k\}$ on $B_r(x_0)$ and note that $w_k \in \mathcal{S}(g)$ as well as $w_k(x_0) \le P(g)(x_0) = u(x_0)$. Hence for any $0 < s < r$ we have

$$\begin{aligned}\int_{S^{n-1}} u(x_0 + s\omega) d\sigma^{n-1}(\omega) &= nV_n u(x_0) \ge nV_n w_k(x_0) \\ &= \int_{S^{n-1}} w_k(x_0 + s\omega) d\sigma^{n-1}(\omega) \\ &\ge \int_{S^{n-1}} \max(w, v_k)(x_0 + s\omega) d\sigma^{n-1}(\omega)\end{aligned}$$

and letting $k \to \infty$ we conclude

$$\int_{S^{n-1}} u(x_0 + s\omega) d\sigma^{n-1}(\omega) \ge \int_{S^{n-1}} \max(w, u)(x_0 + s\omega) d\sigma^{n-1}(\omega).$$

But this implies $\max(w, u) = u$, that is $w \le u$ on $\partial B_s(x_0)$. Since $0 < s < r$ is arbitrary, we get $w \le u$ on $B_r(x_0)$ as required. □

So the remaining question is if $P(g)$ is continuous on $\overline{U}$ and satisfies the given boundary values.

We call a function $w \in C(\overline{U})$ a **barrier function** for U at $x_0 \in \partial U$ provided that w is subharmonic in U, $w(x) < 0$ for $x \in \overline{U} \setminus \{x_0\}$, and $w(x_0) = 0$.

Example 5.5. A point $x_0 \in \partial U$ is said to satisfy the **exterior ball condition** if there is a ball $B_r(x_1)$ such that $B_r(x_1) \cap \overline{U} = \{x_0\}$, that is, if $B_r(x_1)$ touches $\overline{U}$ at the single point x_0 (cf. Figure 5.2). Then

$$w(x) = \Phi(|x - x_1|) - \Phi(r)$$

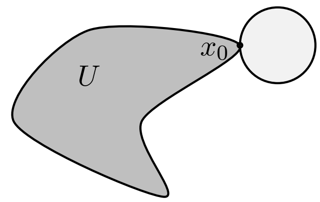


Figure 5.2. Exterior ball condition

is a barrier for U at x_0.

Clearly every convex domain will satisfy this condition. Furthermore, if the boundary is C^2, then we can bound its graph by a quadratic function near each point. Since we can always fit a ball into a parabola, such domains satisfy this condition. If the boundary is merely C^1, this might fail and the exterior ball condition can fail in such a case (Problem 5.33). ⋄

Note that this is a local property of the boundary. In fact let $x_0 \in \partial U$ and suppose we can find a *local barrier* $\tilde{w}$ for $\tilde{U} := U \cap N(x_0)$ at $x_0 \in \partial\tilde{U}$, where $N(x_0)$ is some open neighborhood of x_0. Then we can choose a ball $\bar{B}_\varepsilon(x_0) \subset N(x_0)$ and set $M := \max_{\overline{U \cap N(x_0)} \setminus B_\varepsilon(x_0)} \tilde{w} < 0$ such that

$$w(x) := \begin{cases} \max(\tilde{w}(x), M), & x \in \overline{U} \cap B_\varepsilon(x_0), \\ M, & x \in \overline{U} \setminus B_\varepsilon(x_0), \end{cases} \tag{5.68}$$

will be a barrier for U at $x_0 \in \partial U$.

Theorem 5.28 (Perron)**.** *If U has a barrier at $x_0 \in \partial U$, then $P(g)(x) \to g(x_0)$ whenever $x \to x_0$ from within $\overline{U}$.*

Proof. Let w be a barrier at x_0 and fix $\varepsilon > 0$. Choose a ball $B_r(x_0)$ such that $|g(x) - g(x_0)| < \varepsilon$ for $x \in \partial U \cap B_r(x_0)$. Moreover, if we subtract a suitable multiple $\alpha \geq 0$ of w, we can get a bound $|g(x) - g(x_0)| < \varepsilon - \alpha w(x)$ which holds for all $x \in \partial U$ (e.g. $\alpha = \frac{\max_{\partial U} |g - g(x_0)|}{\min_{\partial U \setminus B_r(x_0)}(-w)}$). We claim that this bound extends to all of $\overline{U}$ in the sense that

$$g(x_0) - \varepsilon + \alpha w(x) \leq P(g)(x) \leq g(x_0) + \varepsilon - \alpha w(x), \qquad x \in \overline{U}.$$

The first inequality comes for free since the function on the left-hand side is in $\mathcal{S}(g)$. For the other inequality choose $v \in \mathcal{S}(g)$ such that $v + \alpha w \leq g + \alpha w < g(x_0) + \varepsilon$ on ∂U. Consequently, Lemma 5.16 shows $v + \alpha w \leq g(x_0) + \varepsilon$ on all of $\overline{U}$ and implies the required inequality.

Since $\varepsilon > 0$ is arbitrary and w is continuous with $w(x_0) = 0$ this establishes the claim. □

A boundary point admitting a barrier is also called **regular** in this context and the boundary ∂U is called regular if all of its points are regular. As a consequence we hence get

Theorem 5.29. *Let U be a bounded domain and $g \in C(\partial U)$. The Dirichlet problem* (5.64) *for U is solvable if and only if ∂U is regular.*

Proof. It remains to show the converse. To this end just note that the solution with boundary data $g(x) = -|x - x_0|$ will be a barrier at $x_0 \in \partial U$. □

Example 5.6. An example where the Dirichlet problem is not solvable is the punctured ball $B_1(0)\setminus\{0\}$: Require u to vanish on $\partial B_1(0)$ and satisfy $u(0) = 1$. Since any rotation of u would be again a solution, uniqueness implies that a potential solution is radial. But the only bounded radial functions which are harmonic on $B_1(0) \setminus \{0\}$ are the constants. A contradiction. ◇

Of course this also settles the Poisson problem:

Theorem 5.30. *Let U be a bounded domain with a regular boundary. Let $g \in C(\partial U)$ and $f \in C(U)$ be bounded and locally Hölder continuous. Then the Poisson problem* (5.29) *has a unique solution $u \in C^2(U) \cap C(\overline{U})$.*

Proof. Subtracting the Newton potential for f from u reduces it to the Dirichlet problem. □

Recall that Example 5.3 shows that the extra assumption that f is locally Hölder continuous cannot be dropped in general.

This shows that if the boundary is regular, then we can solve (5.33) and obtain a Green function for U. The corresponding Green potential

$$u(x) = \int_U G(x,y) f(y) d^n y \tag{5.69}$$

will solve $-\Delta u = f$ and vanish on the boundary. However, this does not answer the question whether the normal derivative exists, that is, whether there is a corresponding Poisson kernel K. Conversely, note that if there is a Green's function then the boundary is regular since one can verify that $w(x) = -\frac{|x-x_0|^2}{2n} - \int_U G(x,y) d^n y$ is a barrier at $x_0 \in \partial U$.

We end this section with the remark that a much weaker condition for the existence of a barrier at $x_0 \in \partial U$ is the **exterior cone condition**. Namely, if x_0 is the vertex of a (truncated) cone contained in the complement of U (cf. Figure 5.3).

To be more specific we set

$$\mathcal{C}^\alpha := \{x \in B_1(0) | x_1^2 + \cdots + x_{n-1}^2 < \alpha^2 x_n^2,\ 0 < x_n\} \subset \mathbb{R}^n. \tag{5.70}$$

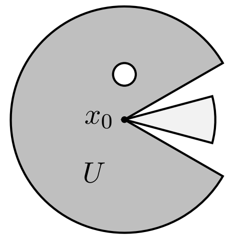


Figure 5.3. Pac-Man crunching an exterior cone

A general (truncated) cone can be obtained from $\mathcal{C}^\alpha$ via translation, scaling, and rotation.

We first note that the maximum principle still holds in a situation, where the harmonic function fails to be continuous at a single point of the boundary.

Lemma 5.31. *Let U be bounded and $x_0 \in \partial U$. Suppose $u \in C(\overline{U} \setminus \{x_0\})$ is bounded and harmonic on U. Then*

$$\sup_U u \le \sup_{\partial U \setminus \{x_0\}} u. \tag{5.71}$$

Proof. Abbreviate $M = \sup_{\partial U \setminus \{x_0\}} u$ and choose r such that $U \subseteq B_r(x_0)$. Since u is bounded, Corollary 5.7 applied to $u(x) - M - \varepsilon(\Phi(x-x_0) - \Phi(r))$ shows $u(x) - M - \varepsilon(\Phi(x-x_0) - \Phi(r)) \le 0$ for all $\varepsilon > 0$. Letting $\varepsilon \downarrow 0$ establishes the claim □

Lemma 5.32. *Let $U := B_1(0) \setminus \mathcal{C}^\alpha$, $g(x) := |x|$ and set $u := P(g)$. Then $-u$ is a barrier for U at 0.*

Proof. By construction u is harmonic on U and satisfies $0 \le u \le 1$. Moreover, since the external ball condition holds on $\partial U \setminus \{0\}$, $u \in C(\overline{U} \setminus \{0\})$. Hence it remains to show $\limsup_{x\to 0} u(x) = 0$.

The idea is to shrink our domain by a factor $r \in (0,1)$ and consider $U_r := rU$. By the strong maximum principle we have $0 \le u \le c < 1$ on $\partial U_r \setminus \{0\}$. Note $c \ge r$. Setting

$$v(x) := u(x) - c\, u(\frac{x}{r})$$

we have $v \le 0$ on $\partial U_r \setminus \{0\}$ and hence on all of U_r by Lemma 5.31. Consequently

$$\limsup_{x\to 0} u(x) \le c \limsup_{x\to 0} u(\frac{x}{r}) = c \limsup_{x\to 0} u(x).$$

As $c < 1$ this establishes $\limsup_{x\to 0} u(x) = 0$ as desired. □

Hence if a point $x_0 \in \partial U$ satisfies the exterior cone condition, a suitable translation, scaling, and rotation of w from the previous lemma will be a barrier.

Problem* 5.33. *Give an example of a C^1 domain which fails the exterior ball condition.*

Problem* 5.34 (Removable singularity theorem)**.** *Suppose $x_0 \in U$ and u is harmonic in $U \setminus \{x_0\}$ and bounded near x_0. Then u can be extended to a function which is harmonic in u. (Hint: Apply Lemma 5.31 to a punctured ball.)*

5.8. General elliptic equations

In this section we briefly discuss the extension to general linear operators of the form

$$Lu := -\sum_{j,k=1}^{n} A_{jk}(x) u_{x_j x_k} + \sum_{j=1}^{n} b_j(x) u_{x_j} + c(x) u, \tag{5.72}$$

where the coefficients are assumed to be real-valued, continuous, and bounded, $A_{jk}, b_j, c \in C_b(U)$ throughout this section. Since the second order partial derivatives are symmetric by the Schwarz theorem, we can assume that A is symmetric $A_{jk} = A_{kj}$. The operator L is sometimes also assumed to be in **divergence form**

$$\tilde{L}u := -\sum_{j,k=1}^{n} \big(A_{jk}(x) u_{x_j}\big)_{x_k} + \sum_{j=1}^{n} \tilde{b}_j(x) u_{x_j} + c(x) u. \tag{5.73}$$

If we assume $A_{jk} \in C^1(U)$, then this is equivalent to the first form with $b_j = \tilde{b}_j - \sum_{k=1}^{n} (A_{jk})_{x_k}$.

The operator L is called **uniformly elliptic** if the matrix A_{jk} is strictly positive definite in the sense that

$$\xi \cdot A(x)\xi = \sum_{j,k=1}^{n} A_{jk}(x) \xi_j \xi_k \geq \theta |\xi|^2 \tag{5.74}$$

for all $x \in U$ and all $\xi \in \mathbb{R}^n$. In other words, the lowest eigenvalue of $A(x)$ is bounded from below by some number $\theta > 0$ independent of $x \in U$.

The purpose of this outlook is not to give a detailed treatment of elliptic equations but only to demonstrate how some of the ideas from the Laplace equation can be extended to this more general case. Specifically, we will only look at the maximum principle. Of course in this case we do not have a mean value property at our disposal from which we can start developing our theory. However, the proof for the maximum principle used in Theorem 3.4 still applies.

Theorem 5.33 (**Maximum principle**). *Suppose L is uniformly elliptic with $c = 0$ on a bounded domain $U \subset \mathbb{R}^n$. Then, if $v \in C(\overline{U}) \cap C^2(U)$ satisfies $Lv \le 0$ we have*

$$\max_{\overline{U}} v \le \max_{\partial U} v. \tag{5.75}$$

Proof. We first assume that $Lv < 0$. Then, if v attains a maximum at $x_0 \in U$, the gradient must vanish $\nabla v(x_0) = 0$ and the Hesse matrix $H_{jk} := v_{x_j x_k}(x_0)$ must be negative definite. Moreover, let O be an orthogonal matrix which diagonalizes $A(x_0)$. That is, such that $OA(x_0)O^T$ is a diagonal matrix which has the eigenvalues $\lambda_1(x_0), \dots, \lambda_n(x_0)$ of $A(x_0)$ as diagonal entries. Then

$$Lv(x_0) = -\sum_{j,k=1}^{n} A_{jk}(x_0) H_{jk} = -\sum_{j=1}^{n} \lambda_j (O^T H O)_{jj} \ge 0$$

gives a contradiction (recall that for a negative definite matrix, all diagonal elements are nonpositive). Thus the maximum principle holds in this case.

If we only have $Lv \le 0$, then we set $v^\varepsilon(x) := v(x) + \varepsilon \mathrm{e}^{\lambda x_1}$ such that

$$Lv^\varepsilon(x) = Lv(x) - \varepsilon\lambda\big(\lambda A_{11}(x) - b_1(x)\big)\mathrm{e}^{\lambda x_1}.$$

Since L is uniformly elliptic we have $A_{11}(x) \ge \theta$ and since b is bounded we can choose $\lambda > \theta^{-1} \sup b_1$ such that $Lv^\varepsilon < 0$. By the first part we have

$$v(x) \le v^\varepsilon(x) \le \max_{x\in\partial U} v(x) + \varepsilon \max_{x\in\partial U} \mathrm{e}^{\lambda x_1}$$

and letting $\varepsilon \to 0$ establishes the claim. □

In the case where $c \ge 0$ we get:

Corollary 5.34. *Let $v \in C(\overline{U}) \cap C^2(U)$ and $c \ge 0$. Then if $Lv \le 0$ we have*

$$\max_{\overline{U}} v \le \max_{\partial U} v^+. \tag{5.76}$$

Here $v^\pm = \max(\pm v, 0)$ is the positive, negative part of v, respectively.

Proof. Following literally the argument of the previous theorem shows that there can be no interior maximum at which v is nonnegative. Hence the claim follows. □

If $c \ge 0$ and $Lu = 0$ we can combine the estimates for u and $-u$ to obtain an priori bound

$$-\max_{\partial U} u^- \le u(x) \le \max_{\partial U} u^+, \tag{5.77}$$

for solutions of the corresponding Dirichlet problem

$$Lu = 0, \qquad u|_{\partial U} = g. \tag{5.78}$$

In particular, there can be at most one solution if $c \ge 0$.

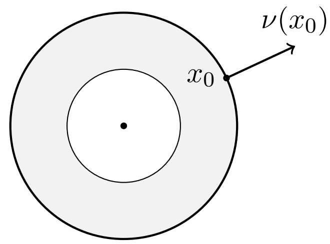


Figure 5.4. Proof of the Hopf–Oleinik lemma

Example 5.7. Note that some condition on c is clearly necessary. Indeed, we have already seen that the Laplace operator $L = -\Delta$ with Dirichlet boundary conditions will have positive eigenvalues λ on a bounded domain. Hence we don't have uniqueness, contradicting a maximum principle for $L = -\Delta - \lambda$. $\diamond$

Using subsolutions one can also establish an a priori bound for solutions of the inhomogeneous problem (Problem 5.35).

Theorem 5.35. *Let U be a bounded domain and suppose L is uniformly elliptic with $c \geq 0$. Then the problem*

$$Lu = f, \qquad u|_{\partial U} = g. \tag{5.79}$$

has at most one solution $u \in C^2(U) \cap C(\overline{U})$ for given $g \in C(\partial U)$, $f \in C(U)$. Moreover, there is a constant C depending only on U and L such that a solution satisfies

$$\max_{\overline{U}} |u| \leq \max_{\partial U} |g| + C \sup_{U} |f|. \tag{5.80}$$

To get the strong maximum principle we need to work a bit harder. First of all note that if there is a maximum at the boundary, then the normal derivative cannot be negative. The following lemma established independently by Hopf[19] and Oleinik[20] says that it is in fact positive.

Lemma 5.36 (Hopf, Oleinik)**.** *Let $v \in C^1(\overline{U}) \cap C^2(U)$ with $U = B_r(0)$ be a subsolution, $Lv \leq 0$. Suppose v attains a strict maximum at some point $x_0 \in \partial B_r(0)$. Then, if either $c = 0$ or $c \geq 0$ and $v(x_0) \geq 0$ we have*

$$\frac{\partial v}{\partial \nu}(x_0) > 0. \tag{5.81}$$

[19] Eberhard Hopf (1902 –1983), Austro-American mathematician and astronomer
[20] Olga Oleinik (1925–2001), Soviet mathematician

Proof. Set $w(x) := \mathrm{e}^{-\lambda|x|^2} - \mathrm{e}^{-\lambda r^2}$ with λ to be determined. Then one computes

$$\begin{aligned} Lw(x) &= \mathrm{e}^{-\lambda|x|^2}\big(-4\lambda^2 x\cdot A(x)x + 2\lambda\operatorname{tr}(A(x)) - 2\lambda b(x)\cdot x + c(x)\big) \\ &\quad - c(x)\mathrm{e}^{-\lambda r^2} \\ &\le \mathrm{e}^{-\lambda|x|^2}\big(-4\lambda^2\theta|x|^2 + 2\lambda\operatorname{tr}(A(x)) + 2\lambda|b(x)|r + c(x)\big). \end{aligned}$$

Since our coefficients are bounded, we can choose λ sufficiently large such that $Lw(x) \le 0$ for $|x| \ge \frac{r}{2}$.

Since by assumption $v(x) < v(x_0)$ for $x \in B_r(0)$, we can find some $\varepsilon > 0$ such that $v(x) + \varepsilon w(x) \le v(x_0)$ for $x \in \partial B_{r/2}(0)$. Moreover, since w vanishes on the boundary, we also have $v(x) + \varepsilon w(x) = v(x) \le v(x_0)$ for $x \in \partial B_r(0)$. So we can apply Corollary 5.34 to $v(x)+\varepsilon w(x)-v(x_0)$ showing $v(x) + \varepsilon w(x) - v(x_0) \le 0$ within the annulus $\bar{B}_r(0) \setminus B_{r/2}(0)$. But then the function $v(x) + \varepsilon w(x) - v(x_0)$ attains its maximum at x_0 implying that the normal derivative is nonnegative:

$$\frac{\partial v}{\partial \nu}(x_0) + \varepsilon\frac{\partial w}{\partial \nu}(x_0) \ge 0.$$

But this establishes the claim

$$\frac{\partial v}{\partial \nu}(x_0) \ge -\varepsilon\frac{\partial w}{\partial \nu}(x_0) = -\frac{\varepsilon}{r}(\nabla w)(x_0)\cdot x_0 = 2\varepsilon\lambda r\mathrm{e}^{-\lambda r^2}. \qquad \square$$

Example 5.8. As an application of the Hopf–Oleinik lemma, note that the Poisson kernel $K(x, y)$ (provided it exists) must be positive at every point $y_0 \in \partial U$ which satisfies an **interior ball condition**. Indeed, by making the ball smaller we can assume that the ball touches the boundary only at y_0 and that the ball does not contain $x \in U$. Then we can apply the Hopf–Oleinik lemma to $u(y) := -G(x, y)$ on this ball to conclude that $K(x, y_0) = -\frac{\partial G}{\partial \nu}(x, y_0) > 0$. ◇

Example 5.9. Here is an example which shows that the Hopf–Oleinik lemma does not hold if the interior ball condition is dropped. Consider the harmonic function ($z = x + \mathrm{i}y$)

$$u(x, y) := \operatorname{Re}\Big(\frac{z}{-\log(z)}\Big), \quad U := \{(x, y) \in \mathbb{R}^2 | 0 < x < 1,\ u(x, y) > 0\}.$$

Note that the boundary of U is given in polar coordinates $z = r\mathrm{e}^{\mathrm{i}\varphi}$ by

$$r = \mathrm{e}^{-\varphi\tan(\varphi)}, \qquad -\frac{\pi}{2} \le \varphi \le \frac{\pi}{2},$$

and hence r starts at $r(-\frac{\pi}{2}) = 0$, increases until it attains its maximum at $r(0) = 1$, and then decreases until it returns again to $r(\frac{\pi}{2}) = 0$. Thus U is a nice symmetric (with respect to the x axis) blob whose boundary is smooth except possible at the origin. With a little effort one can show that the normalized tangent vector of this curve has matching limits and hence the boundary is even C^1. Moreover, u explicitly reads

$$u(x,y) = -\frac{x\log(r) + y\varphi}{\varphi^2 + \log(r)^2}, \qquad r = \sqrt{x^2+y^2}, \quad \varphi = \arcsin\big(\frac{y}{r}\big),$$

and the derivatives are given by

$$u_x(x,y) = -\frac{\varphi^2(1+\log(r)) + \log(r)^2(\log(r)-1)}{(\varphi^2+\log(r)^2)^2},$$
$$u_y(x,y) = -\frac{\varphi(\varphi^2 - 2\log(r) + \log(r)^2)}{(\varphi^2+\log(r)^2)^2}.$$

In particular, $u \in C^1(\overline{U} \setminus \{(1,0)\})$ with

$$\frac{\partial u}{\partial \nu}(0,0) = -u_x(0,0) = 0$$

and hence U cannot satisfy the interior ball condition at 0 (indeed one can verify that the curvature of ∂U is unbounded). ⋄

Theorem 5.37 (**Strong maximum principle**). *Let U be a bounded and connected domain and suppose L is uniformly elliptic with $c = 0$. Then if $v \in C^2(U)$ satisfies $Lv \le 0$ and v attains a maximum at an interior point, it must be constant.*

If $c \ge 0$ the same conclusion holds if v attains a nonnegative maximum at an interior point.

Proof. Set $M := \sup_U v$ and set $K := \{x \in U | v(x) = M\}$, $V := \{x \in U | v(x) < M\}$. By assumption there is some $\tilde{x}_0 \in K$. Suppose we can find some $\tilde{y}_0 \in V$. Since U is path connected, there is some path from $\tilde{y}_0$ to $\tilde{x}_0$. Moreover, we can choose some point $y_0 \in V$ on this path and a corresponding $r > 0$ such that $\bar{B}_r(y_0) \subset U$ and $B_r(y_0)$ contains points from K. Since K is compact, there is some $x_0 \in K$ with $\operatorname{dist}(y_0, K) = |y_0 - x_0| =: r_0 < r$ (in particular, $x_0 \in U$). Consequently $B_{r_0}(y_0) \subseteq V$ and $x_0 \in \partial B_{r_0}(y_0)$. By moving y_0 closer to x_0 and reducing r_0 we can even assume $\bar{B}_{r_0}(y_0) \setminus \{x_0\} \subseteq V$. Now the fact that v attains a strict maximum at x_0 implies $\frac{\partial v}{\partial \nu}(x_0) = 0$ while the Hopf lemma implies $\frac{\partial v}{\partial \nu}(x_0) > 0$. This contradiction shows that V is empty. □

Problem* 5.35. *Prove Theorem 5.35. (Hint: Assume that U is within a strip $0 < x_1 < r$ and construct a supersolution using $\mathrm{e}^{\lambda x_1}$.)*

Problem 5.36. *Derive a Dirichlet principle for the elliptic operator $\tilde{L}$ in divergence form with $A \in C^1$ and $b = 0$.*

Chapter 6

The heat equation

6.1. The Fourier transform

We have seen that the Fourier transform is a useful tool for solving partial differential equation in one spatial dimension. In this section we want to extend these results to the case of $\mathbb{R}^n$.

For an integrable function $f : \mathbb{R}^n \to \mathbb{C}$ we define its **Fourier transform** via

$$\mathcal{F}(f)(k) \equiv \hat{f}(k) := \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \mathrm{e}^{-\mathrm{i}k\cdot x} f(x) d^n x. \tag{6.1}$$

Here $k \cdot x = k_1 x_1 + \cdots + k_n x_n$ is the usual scalar product in $\mathbb{R}^n$ and we will use $|x| = \sqrt{x_1^2 + \cdots + x_n^2}$ for the Euclidean norm.

By Lemma A.6 it follows that $\hat{f}$ is a bounded continuous function and we have the explicit bound

$$|\hat{f}(k)| \le (2\pi)^{-n/2} \int_{\mathbb{R}^n} |f(x)| d^n x. \tag{6.2}$$

The following simple properties are left as an exercise.

Lemma 6.1. *Let f be integrable. Then*

$$(f(x+a))^\wedge(k) = \mathrm{e}^{\mathrm{i}a\cdot k} \hat{f}(k), \qquad a \in \mathbb{R}^n, \tag{6.3}$$

$$(f(Mx))^\wedge(k) = |\det(M)|^{-1} \hat{f}((M^{-1})^T k), \qquad M \in \mathrm{GL}(\mathbb{R}^n), \tag{6.4}$$

$$(\mathrm{e}^{\mathrm{i}x\cdot a} f(x))^\wedge(k) = \hat{f}(k-a), \qquad a \in \mathbb{R}^n, \tag{6.5}$$

$$(f(\lambda x))^\wedge(k) = \frac{1}{\lambda^n} \hat{f}(\frac{k}{\lambda}), \qquad \lambda > 0, \tag{6.6}$$

$$(f(-x))^\wedge(k) = \hat{f}(-k). \tag{6.7}$$

Also the connection with differentiation can be established by literally following the proof of the one-dimensional case from Lemma 4.1.

Lemma 6.2. *Suppose $f \in C^1(\mathbb{R}^n)$ such that $\lim_{|x|\to\infty} f(x) = 0$ and $f, \partial_j f$ integrable for some $1 \le j \le n$. Then*

$$(\partial_j f)^\wedge(k) = \mathrm{i} k_j \hat{f}(k). \tag{6.8}$$

Similarly, if $f(x), x_j f(x)$ are integrable for some $1 \le j \le n$, then $\hat{f}(k)$ is differentiable with respect to k_j and

$$(x_j f(x))^\wedge(k) = \mathrm{i}\partial_j \hat{f}(k). \tag{6.9}$$

Again this result immediately extends to higher derivatives and, roughly speaking shows, that the decay of a function is related to the smoothness of its Fourier transform and the smoothness of a function is related to the decay of its Fourier transform.

Next recall the multi-index notation (Section A.1) and the **Schwartz space**[1]

$$\mathcal{S}(\mathbb{R}^n) := \{f \in C^\infty(\mathbb{R}^n) |\sup_x |x^\alpha (\partial_\beta f)(x)| < \infty,\ \forall \alpha, \beta \in \mathbb{N}_0^n\}. \tag{6.10}$$

Note that if $f \in \mathcal{S}(\mathbb{R}^n)$, then the same is true for $x^\alpha f(x)$ and $(\partial_\alpha f)(x)$ for every multi-index α. Also, by Leibniz' rule[2], the product of two Schwartz functions is again a Schwartz function.

Example 6.1. The prototypical Schwartz function is the **Gaussian** $f(x) := \mathrm{e}^{-c|x|^2}$ for $c > 0$. Clearly f is smooth and its derivatives have the form of a polynomial times f. Since the exponential function grows faster than any polynomial we conclude that f is Schwartz. ⋄

Lemma 6.3. *The Fourier transform satisfies $\mathcal{F} : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$. Furthermore, for every multi-index $\alpha \in \mathbb{N}_0^n$ and every $f \in \mathcal{S}(\mathbb{R}^n)$ we have*

$$(\partial_\alpha f)^\wedge(k) = (\mathrm{i}k)^\alpha \hat{f}(k), \qquad (x^\alpha f(x))^\wedge(k) = \mathrm{i}^{|\alpha|}\partial_\alpha \hat{f}(k). \tag{6.11}$$

Proof. The formulas are immediate from the previous lemma. To see that $\hat{f} \in \mathcal{S}(\mathbb{R}^n)$ if $f \in \mathcal{S}(\mathbb{R}^n)$, we begin with the observation that $\hat{f}$ is bounded by (6.2). But then $k^\alpha(\partial_\beta \hat{f})(k) = \mathrm{i}^{-|\alpha|-|\beta|}(\partial_\alpha x^\beta f(x))^\wedge(k)$ is bounded since $\partial_\alpha x^\beta f(x) \in \mathcal{S}(\mathbb{R}^n)$ if $f \in \mathcal{S}(\mathbb{R}^n)$. □

The Fourier transform of a Gaussian follows immediately from the one -dimensional case.

[1] Laurent Schwartz (1915–2002), French mathematician

[2] Gottfried Wilhelm Leibniz (1646–1716), German logician, mathematician, and natural philosopher

Lemma 6.4. *We have* $\mathrm{e}^{-t|x|^2/2} \in \mathcal{S}(\mathbb{R}^n)$ *for* $t > 0$ *and*

$$\mathcal{F}(\mathrm{e}^{-t|x|^2/2})(k) = \frac{1}{t^{n/2}}\mathrm{e}^{-|k|^2/(2t)}. \tag{6.12}$$

Proof. Due to the product structure of the exponential, one can use Fubini and treat each coordinate separately, reducing the problem to Lemma 4.2. □

Now we can show

Theorem 6.5. *Suppose* f *is continuous such that* f*,* $\hat{f}$ *are integrable. Then*

$$(\hat{f})^\vee = f, \tag{6.13}$$

where

$$\check{f}(k) := \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \mathrm{e}^{\mathrm{i}k\cdot x} f(x) d^n x = \hat{f}(-k). \tag{6.14}$$

In particular, $\mathcal{F} : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$ *is a bijection.*

Proof. Abbreviate $\phi_\varepsilon(x) := (2\pi)^{-n/2} \exp(-\varepsilon|x|^2/2)$. Then

$$\int_{\mathbb{R}^n} \phi_\varepsilon(k)\mathrm{e}^{\mathrm{i}k\cdot x}\hat{f}(k)d^n k = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \phi_\varepsilon(k)\mathrm{e}^{\mathrm{i}k\cdot x} f(y)\mathrm{e}^{-\mathrm{i}k\cdot y} d^n y d^n k$$

and, invoking Fubini and Lemma 6.1, we further see that this is equal to

$$= \int_{\mathbb{R}^n} (\phi_\varepsilon(k)\mathrm{e}^{\mathrm{i}k\cdot x})^\wedge(y) f(y) d^n y = \int_{\mathbb{R}^n} \frac{1}{\varepsilon^{n/2}}\phi_{1/\varepsilon}(y-x) f(y) d^n y.$$

Letting $\varepsilon \to 0$ the integral we have started with converges to $(\hat{f})^\vee(x)$ while the last one converges to $f(x)$ by Lemma A.12. □

Another fundamental property is the fact that the Fourier transform preserves square integrability.

Lemma 6.6 (**Plancherel identity**)**.** *Suppose* $f, \hat{f}$ *are both integrable. Then* $f, \hat{f}$ *are square integrable and*

$$\int_{\mathbb{R}^n} |f(x)|^2 d^n x = \int_{\mathbb{R}^n} |\hat{f}(k)|^2 d^n k \le \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} |f(x)| d^n x \int_{\mathbb{R}^n} |\hat{f}(k)| d^n k \tag{6.15}$$

holds.

Proof. This follows from Fubini's theorem since

$$\begin{aligned} \int_{\mathbb{R}^n} |\hat{f}(k)|^2 d^n k &= \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} f(x)^* \hat{f}(k)\mathrm{e}^{\mathrm{i}k\cdot x} d^n k\, d^n x \\ &= \int_{\mathbb{R}^n} |f(x)|^2 d^n x \end{aligned}$$

for $f, \hat{f}$ integrable. □

Another key property is the convolution formula.

Lemma 6.7. *The* ***convolution***

$$(f * g)(x) = \int_{\mathbb{R}^n} f(y)g(x-y)d^n y = \int_{\mathbb{R}^n} f(x-y)g(y)d^n y \tag{6.16}$$

of two integrable functions f, g is again integrable and we have ***Young's inequality***[3]

$$\int_{\mathbb{R}^n} |(f*g)(x)| d^n x \le \int_{\mathbb{R}^n} |f(x)| d^n x \int_{\mathbb{R}^n} |g(x)| d^n x. \tag{6.17}$$

Moreover, its Fourier transform is given by

$$(f * g)^\wedge = (2\pi)^{n/2} \hat{f} \hat{g}. \tag{6.18}$$

Proof. The fact that $f * g$ is in integrable together with Young's inequality follows by applying Fubini's theorem to $h(x,y) = f(x-y)g(y)$. For the last claim we compute

$$\begin{aligned}
(f*g)^\wedge(k) &= \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \mathrm{e}^{-\mathrm{i}k\cdot x} \left(\int_{\mathbb{R}^n} f(y)g(x-y)d^n y \right) d^n x \\
&= \int_{\mathbb{R}^n} \mathrm{e}^{-\mathrm{i}k\cdot y} f(y) \left(\frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \mathrm{e}^{-\mathrm{i}k\cdot(x-y)} g(x-y) d^n x \right) d^n y \\
&= \int_{\mathbb{R}^n} \mathrm{e}^{-\mathrm{i}k\cdot y} f(y) \hat{g}(k) d^n y = (2\pi)^{n/2} \hat{f}(k)\hat{g}(k),
\end{aligned}$$

where we have again used Fubini's theorem. □

Example 6.2. The image of the integrable functions under the Fourier transform is known as the **Wiener algebra**

$$\mathcal{A}(\mathbb{R}^n) := \{\hat{f} | f \text{ integrable}\}.$$

By construction, this is just the range of the Fourier transform and hence a subset of $C_b(\mathbb{R}^n)$. Moreover, Lemma 6.7 shows that the product of two functions $f, g \in \mathcal{A}(\mathbb{R}^n)$ is again in the Wiener algebra. ⋄

As a consequence we can also deal with the case of convolution on $\mathcal{S}(\mathbb{R}^n)$.

Corollary 6.8. *The convolution of two $\mathcal{S}(\mathbb{R}^n)$ functions as well as their product is in $\mathcal{S}(\mathbb{R}^n)$ and*

$$(f * g)^\wedge = (2\pi)^{n/2} \hat{f}\hat{g}, \qquad (fg)^\wedge = (2\pi)^{-n/2} \hat{f} * \hat{g} \tag{6.19}$$

in this case.

[3] William Henry Young (1863–1942), English mathematician

Proof. Clearly the product of two functions in $\mathcal{S}(\mathbb{R}^n)$ is again in $\mathcal{S}(\mathbb{R}^n)$ (show this!). Since functions in $\mathcal{S}(\mathbb{R}^n)$ are integrable, the previous lemma implies $(f*g)^\wedge = (2\pi)^{n/2}\hat{f}\hat{g} \in \mathcal{S}(\mathbb{R}^n)$. Moreover, since the Fourier transform is injective on $\mathcal{S}(\mathbb{R}^n)$ we conclude $f*g = (2\pi)^{n/2}(\hat{f}\hat{g})^\vee \in \mathcal{S}(\mathbb{R}^n)$. Replacing f, g by $\check{f}, \check{g}$ in the last formula finally shows $\check{f}*\check{g} = (2\pi)^{n/2}(fg)^\vee$ and the claim follows by a simple change of variables using $\check{f}(k) = \hat{f}(-k)$. □

Finally let us comment on how to use the Fourier transform to solve linear partial differential equations with constant coefficients. By virtue of Lemma 6.3 the Fourier transform will map such equations to algebraic equations, thereby providing a mean of solving them. To illustrate this procedure we look at the Poisson equation

$$-\Delta u = f. \tag{6.20}$$

For simplicity, let us investigating this problem in the space of Schwartz functions $\mathcal{S}(\mathbb{R}^n)$. Assuming there is a solution we can take the Fourier transform on both sides to obtain

$$|k|^2\hat{u}(k) = \hat{f}(k) \qquad \Rightarrow \qquad \hat{u}(k) = |k|^{-2}\hat{f}(k). \tag{6.21}$$

Since the right-hand side is integrable for $n \ge 3$ we obtain that our solution is necessarily given by

$$u(x) = (|k|^{-2}\hat{f}(k))^\vee(x). \tag{6.22}$$

Moreover, since $|k|^2\hat{u}(k) = \hat{f}(k) \in \mathcal{S}(\mathbb{R}^n)$ Lemma 6.2 implies that $u \in C^2(\mathbb{R}^n)$ as well as that it is indeed a solution (in fact, we even get $u \in C^\infty(\mathbb{R}^n)$). Hence we clearly expect that u should be given by a convolution (in fact, we of course already know this from Corollary 5.20). However, since $|k|^{-2}$ is not integrable, Lemma 6.7 does not apply. It is still possible to compute the inverse Fourier transform of $|k|^{-2}$ upon using a suitable regularization procedure (cf. [**34**, Section 8.3]) and this will of course again give the fundamental solution Φ of the Laplace equation which we have already found in Section 5.3.

The situation improves slightly if one looks at the Helmholtz equation

$$-\Delta u + u = f. \tag{6.23}$$

As before we obtain

$$u(x) = ((1+|k|^2)^{-1}\hat{f}(k))^\vee(x) \tag{6.24}$$

and at least there now is no singularity at the origin. However, $(1+|k|^2)^{-1}$ is still not integrable and we cannot apply Lemma 6.7 directly. The trick is to use the elementary integral

$$\hat{B}_\varepsilon(k) := (2\pi)^{-n/2}\frac{\mathrm{e}^{-(1+|k|^2)\varepsilon}}{1+|k|^2} = (2\pi)^{-n/2}\int_\varepsilon^\infty \mathrm{e}^{-r(1+|k|^2)}dr. \tag{6.25}$$

Clearly the left-hand side is integrable for $\varepsilon > 0$ and reduces to the desired function in the limit $\varepsilon \to 0$. Hence Lemma 6.7 implies

$$u_\varepsilon(x) = (2\pi)^{n/2}(\hat{B}_\varepsilon \hat{f})^\vee(x) = (B_\varepsilon * f)(x), \tag{6.26}$$

where

$$\begin{aligned} B_\varepsilon(x) &= \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} \mathrm{e}^{\mathrm{i}k\cdot x}\int_\varepsilon^\infty \mathrm{e}^{-r(1+|k|^2)} dr\, d^n k \\ &= \frac{1}{(2\pi)^n}\int_\varepsilon^\infty \mathrm{e}^{-r}\int_{\mathbb{R}^n} \mathrm{e}^{\mathrm{i}k\cdot x}\mathrm{e}^{-r|k|^2} d^n k\, dr \\ &= \frac{1}{(4\pi)^{n/2}}\int_\varepsilon^\infty r^{-n/2}\mathrm{e}^{-r-\frac{|x|^2}{4r}}\, dr. \end{aligned} \tag{6.27}$$

Here we have used Fubini and Lemma 6.4. To perform the limit $\varepsilon \to 0$ we observe

$$|B_\varepsilon(x)| \le \frac{1}{(4\pi)^{n/2}}\int_0^\infty r^{-n/2}\mathrm{e}^{-\frac{|x|^2}{4r}}\, dr = \frac{\Gamma(n/2-1)}{4\pi^{n/2}} x^{2-n}. \tag{6.28}$$

Hence for $n > 2$ and $\hat{f}$, f integrable we get that the solution is given by the **Bessel potential**

$$u(x) = (B * f)(x), \tag{6.29}$$

where

$$B(x) := \frac{1}{(4\pi)^{n/2}}\int_0^\infty r^{-n/2}\mathrm{e}^{-r-\frac{|x|^2}{4r}}\, dr. \tag{6.30}$$

Using [**25**, (10.32.10)] B can be expressed as

$$B(x) = \frac{|x|^{1-n/2}}{(2\pi)^{n/2}} K_{n/2-1}(|x|), \tag{6.31}$$

where $K_\nu(r)$ is the **modified Bessel function** of the second kind. While the fundamental solution of the Helmholtz equation B has the same singularity as the fundamental solution of the Laplace equation Φ, it has much nicer decay properties in accordance with the fact that its Fourier transform is smooth (cf. [**25**, (10.30.2), (10.40.2)], respectively).

Problem 6.1. *Compute the Fourier transform of*

$$|x|^2 \mathrm{e}^{-|x|^2/2}$$

in $\mathbb{R}^n$. (Hint: There is no need to compute integrals.)

Problem 6.2. *Find a function f such that $\int_{\mathbb{R}} f(y)f(x-y)dy = \mathrm{e}^{-x^2}$.*

Problem 6.3. *Let $F : \mathbb{R} \to \mathbb{C}$ be an even function and set $f_n(x) = F(|x|)$ for $x \in \mathbb{R}^n$. Establish the following formulas for the Fourier transform of* *radial functions:*

$$\hat{f}_n(k) = \begin{cases} \sqrt{\frac{2}{\pi}} \int_0^\infty \cos(kr)F(r)dr, & n = 1, \\ \int_0^\infty J_0(|k|r)F(r)r\,dr, & n = 2, \\ \sqrt{\frac{2}{\pi}} \int_0^\infty \frac{\sin(|k|r)}{|k|r} F(r)r^2 dr, & n = 3. \end{cases}$$

Note that $\hat{f}_3(k) = \frac{1}{r}\frac{\partial}{\partial r}\hat{f}_1(r)\big|_{r=|k|}$ *provided* f_1, f_3 *are integrable. This is only the tip of the iceberg and the case of arbitrary dimensions is given in* **[34]**. *(Hint: Spherical coordinates and Problem 3.24 for* $n = 2$*. Use the fact that you can rotate the coordinate system appropriately.)*

Problem* 6.4. *Compute the Fourier transform of*

$$\varphi_t(x) = \sqrt{\frac{\pi}{2}} \frac{\chi_{[0,t]}(|x|)}{|x|}, \qquad t \geq 0,$$

in $\mathbb{R}^3$.

Problem 6.5. *Show that the Fourier transform of a function with compact support is real analytic. (Hint: Find an appropriate bound on the derivatives.)*

Problem 6.6. *Solve the* ***Schrödinger equation***

$$-\mathrm{i}u_t = \Delta u, \qquad u(0,x) = u_0(x) \in \mathcal{S}(\mathbb{R}^n),$$

using the Fourier transform:

$$u(t,x) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \mathrm{e}^{\mathrm{i}k\cdot x - \mathrm{i}|k|^2 t} \hat{u}_0(k) d^n k.$$

Conclude that

$$\int_{\mathbb{R}^n} |u(t,x)|^2 d^n x = \int_{\mathbb{R}^n} |u_0(x)|^2 d^n x.$$

6.2. The fundamental solution

In this section we want to solve the heat equation

$$u_t = \Delta u, \qquad u(0,x) = g(x), \tag{6.32}$$

on $\mathbb{R}^n$. We assume that $g \in \mathcal{S}(\mathbb{R}^n)$ and that there is a solution which is in $\mathcal{S}(\mathbb{R}^n)$ for all $t > 0$. Then, taking the Fourier transform we obtain

$$\hat{u}_t(t,k) = -|k|^2 \hat{u}(t,k), \qquad \hat{u}(0,k) = \hat{g}(k). \tag{6.33}$$

Solving this differential equation shows

$$\hat{u}(t,k) = \hat{g}(k)\mathrm{e}^{-t|k|^2} \tag{6.34}$$

and taking the inverse Fourier transform (using Lemma 6.8) gives

$$u(t,x) = (\hat{g}(k)\mathrm{e}^{-t|k|^2})^\vee(x) = (\Phi_t * g)(x) \tag{6.35}$$

where (by Lemma 6.4)

$$\Phi_t(x) \equiv \Phi(t,x) := \frac{1}{(2\pi)^{n/2}}(\mathrm{e}^{-t|k|^2})^\vee(x) = \frac{1}{(4\pi t)^{n/2}}\mathrm{e}^{-\frac{|x|^2}{4t}}, \quad t>0, \tag{6.36}$$

is the **fundamental solution of the heat equation**. Hence we obtain the analogue of Theorem 4.4 with literally the same proof.

Theorem 6.9. *Suppose g is bounded. Then*

$$u(t,x) := \begin{cases} \int_{\mathbb{R}^n} \Phi(t,x-y)g(y)d^ny, & t>0, \\ g(x), & t=0, \end{cases} \tag{6.37}$$

defines a solution of the heat equation which satisfies $u \in C^\infty((0,\infty)\times\mathbb{R}^n) \cap C([0,\infty)\times\mathbb{R}^n)$. Moreover,

$$\inf g \le u(t,x) \le \sup g. \tag{6.38}$$

Note that since $\Phi_t > 0$ the inequality in (6.38) is strict for $t>0$ unless g is constant implying infinite propagation speed.

Corollary 6.10. *If g is integrable the solution* (6.37) *has the following properties:*

(i) *(Mass conservation)*

$$\int_{\mathbb{R}^n} u(t,y)d^ny = \int_{\mathbb{R}^n} g(y)d^ny. \tag{6.39}$$

(ii)

$$|u(t,x)| \le \frac{1}{(4\pi t)^{n/2}}\int_{\mathbb{R}^n}|g(y)|d^ny. \tag{6.40}$$

Of course the inhomogeneous heat equation

$$u_t = \Delta u + f \tag{6.41}$$

can be solved in the same way and we obtain the **Duhamel principle**

$$\hat{u}(t,k) = \hat{g}(k)\mathrm{e}^{-t|k|^2} + \int_0^t \mathrm{e}^{-(t-s)|k|^2}\hat{f}(s,k)ds. \tag{6.42}$$

Assuming

$$|\hat{g}| + |\hat{f}(s,k)| \le M(k), \qquad \int_{\mathbb{R}^n} M(k)d^nk < \infty, \tag{6.43}$$

we can take the inverse Fourier transform and conclude that

$$u(t,x) = \int_{\mathbb{R}^n}\Phi(t,x-y)g(y)d^ny + \int_0^t\int_{\mathbb{R}^n}\Phi(t-s,x-y)f(s,y)d^ny\,ds \tag{6.44}$$

is in $C^{1;2}((0,\infty)\times\mathbb{R}^n)\cap C([0,\infty)\times\mathbb{R}^n)$, solves the inhomogeneous heat equation and satisfies the initial condition $u(0,x) = g(x)$. Note that existence of the required derivatives follows from Lemma 6.2 since

$$\left|\int_0^t \mathrm{e}^{-(t-s)|k|^2} \hat{f}(s,k)ds\right| \le M(k)\int_0^t \mathrm{e}^{-(t-s)|k|^2}ds = M(k)\frac{1-\mathrm{e}^{-|k|^2 t}}{|k|^2}. \quad (6.45)$$

Uniqueness will be investigated in the next section.

Since checking (6.43) might be nontrivial, we remark that as in the one-dimensional case (Theorem 4.6) one can show:

Theorem 6.11. *Suppose $f \in C([0,\infty)\times\mathbb{R}^n)$ is bounded and uniformly Hölder continuous with respect to the second argument on compact sets with respect to the first argument:*

$$|f(t,x) - f(t,y)| \le C_T|x-y|^\gamma, \qquad 0 \le t \le T. \quad (6.46)$$

Then

$$u(t,x) := \int_0^t \int_{\mathbb{R}^n} \Phi(t-s,x-y) f(s,y) d^n y\, ds, \qquad t \ge 0, \quad (6.47)$$

is in $C^{1;2}((0,\infty)\times\mathbb{R}^n)\cap C([0,\infty)\times\mathbb{R}^n)$ and solves the inhomogeneous heat equation with initial condition $u(0,x)=0$.

Finally we remark that, if the inhomogeneous term f does not depend on t, then one could subtract the corresponding solution of the Poisson problem (i.e. the Newton potential of f) to reduce the problem to a homogenous one.

Problem 6.7. *Suppose $g \in \mathcal{S}(\mathbb{R}^n)$. Show that the solutions (6.37) of the heat equation preserve the first moment:*

$$\int_{\mathbb{R}^n} x\, u(t,x) d^n x = \int_{\mathbb{R}^n} x\, g(x) d^n x.$$

(Hint: How can you read off mass conservation from (6.34)?)

Problem 6.8. *Suppose g is integrable. Show that solutions of the heat equation satisfy*

$$\int_{\mathbb{R}^n} |u(t,x)| d^n x \le \int_{\mathbb{R}^n} |g(x)| d^n x$$

with the inequality being strict unless g is of one sign.

Problem 6.9. *Use the **Hopf–Cole transformation**[4]*

$$v := \mathrm{e}^{-bu/a}$$

to find a solution of the nonlinear equation

$$u_t - a\Delta u + b|\nabla u|^2 = 0, \qquad u(0,x) = g(x),$$

on $\mathbb{R}^n$.

[4] Julian Cole (1925–1999), America mathematician

Problem 6.10. *Use the transformation*

$$v(t,x) := \int_{-\infty}^{x} u(t,y)dy$$

to find a solution of the vicious Burgers' equation equation

$$u_t - au_{xx} + uu_x = 0, \qquad u(0,x) = g(x),$$

on $\mathbb{R}$. *(Hint: Problem 6.9.)*

6.3. The heat equation on a bounded domain and the maximum principle

Our next aim is to derive a representation formula analog to (5.36) for the inhomogeneous heat equation

$$u_t = \Delta u + f \tag{6.48}$$

on a bounded domain U (which we assume sufficiently smooth such that we can apply the Gauss–Green theorem). Recall the notation $U_T := (0,T] \times U$ and $\Gamma_T := \overline{U_T} \setminus U_T$ from Section 3.1. Assume that we have a solution $u \in C^{1;2}(\overline{U_T})$ and let $v \in C^{1;2}(\overline{U_T})$, to be chosen later. Then integration by parts with respect to t and applying Green's second identity with respect to x we obtain

$$\begin{aligned}
\int_{U_t} vf\, d^n y\, ds &= \int_{U_t} v(u_s - \Delta u)\, d^n y\, ds \\
&= \int_U \left(\int_0^t v u_s ds \right) d^n y - \int_0^t \left(\int_U v \Delta u\, d^n y \right) ds \\
&= \int_U \left(v(t)u(t) - v(0)u(0) - \int_0^t v_s u\, ds \right) d^n y \\
&\quad - \int_0^t \left(\int_U u \Delta v\, d^n y + \int_{\partial U} \left(v \frac{\partial u}{\partial \nu} - u \frac{\partial v}{\partial \nu} \right) dS \right) ds \\
&= \int_U (v(t)u(t) - v(0)u(0))\, d^n y - \int_{U_t} u(v_s + \Delta v) d^n y\, ds \\
&\quad - \int_0^t \int_{\partial U} \left(v \frac{\partial u}{\partial \nu} - u \frac{\partial v}{\partial \nu} \right) dS\, ds.
\end{aligned} \tag{6.49}$$

Choosing $v(s,y) := \Phi(t-s+\varepsilon, x-y)$ and taking $\varepsilon \to 0$ we obtain

$$\begin{aligned}
u(t,x) &= \int_U \Phi(t, x-y) u(0,y) d^n y + \int_0^t \int_U \Phi(t-s, x-y) f(s,y) d^n y\, ds \\
&+ \int_0^t \int_{\partial U} \left(\Phi(t-s, x-y) \frac{\partial u(s,y)}{\partial \nu} - u(s,y) \frac{\partial \Phi(t-s,x-y)}{\partial \nu} \right) dS(y) ds
\end{aligned} \tag{6.50}$$

for $x \in U$. Now this already looks quite promising, except for the fact that this formula involves both the values of u and the values of its normal derivative on the boundary. The way to fix this is of course by adding a correction term to Φ such that the result vanishes on the boundary. Explicitly, let ϕ^x be the solution of the homogenous heat equation

$$\phi^x_t = \Delta\phi^x, \quad \phi^x(0,y) = 0,\ y \in U, \quad \phi^x(t,y) = \Phi(t, x-y),\ y \in \partial U, t > 0, \tag{6.51}$$

and define the **heat kernel** of U as

$$K(t,x,y) := \Phi(t,x-y) - \phi^x(t,y), \qquad t > 0,\ x,y \in \overline{U}. \tag{6.52}$$

We will say that the heat kernel exists if the above problem has a solution $\phi^x \in C^{1;2}([0,\infty) \times \overline{U})$ for all $x \in U$. It is quite tedious to establish existence of K, see for example [**16**, Section 4.3]. Of course if K exists, then our representation formula simplifies to

$$\begin{aligned} u(t,x) = &\int_U K(t,x,y)u(0,y)d^n y + \int_0^t \int_U K(t-s,x,y)f(s,y)d^n y\, ds \\ &- \int_0^t \int_{\partial U} \frac{\partial K(t-s,x,y)}{\partial\nu} u(s,y)dS(y)\, ds. \end{aligned} \tag{6.53}$$

Unfortunately, in contradistinction to the Laplace equation, it does not seem possible to write down simple expressions even when U is the unit ball (except in one dimension, where we have found an expression in terms of theta functions (3.19)). If one can compute eigenvalues and eigenfunctions for the Laplacian with Dirichlet boundary conditions (cf. e.g. Section 3.4) one can use (3.51), but this does not give a simple form either.

So while it looks like we are stuck here, there is still something we can do. Indeed, observe that the mean value formula for harmonic functions can be derived from the analog formula (5.32) for the Laplace equation by choosing a domain such that the fundamental solution is constant on the boundary. Of course in the case of the Laplace equation these are just balls and this leads naturally to the mean value formula – cf. Problem 5.21.

Hence we will replace U_t in the above calculation by a set where Φ is constant on the boundary. Correspondingly we define the **heat ball** as

$$E_r(t,x) := \{(s,y) \in \mathbb{R}^{n+1} | s < t,\ \Phi(t-s, x-y) \ge \frac{1}{r^n}\} \cup \{(t,x)\}. \tag{6.54}$$

Note that the slice of $E_r(t,x)$ for fixed s is a ball centered at x given by

$$|x-y| \le \rho_n(t-s), \qquad \rho_n(t) := \sqrt{2nt\log(\frac{r^2}{4\pi t})}. \tag{6.55}$$

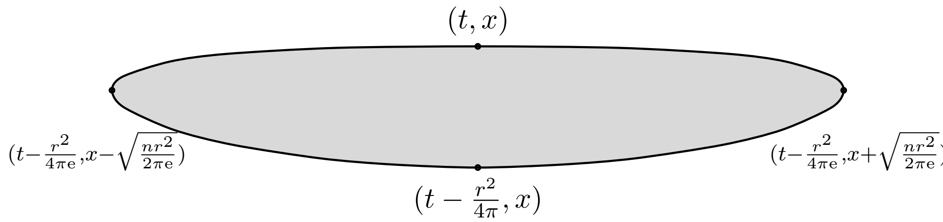


Figure 6.1. Heat ball

The radius vanishes at $s = t$, increases to a maximal value $r\sqrt{\frac{n}{2\pi e}}$ at $s = t - \frac{r^2}{4\pi e}$ and then decreases until it vanishes again at $s = t - \frac{r^2}{4\pi}$ (see Figure 6.1). By construction we have $\Phi(t-s, x-y) = r^{-n}$ for $(s,y) \in \partial E_r(t,x)$.

Lemma 6.12. *A solution $u \in C^{1;2}(U_T)$ of the inhomogeneous heat equation* (6.48) *can be represented as*

$$u(t,x) = \int_{E_r(t,x)} (\Phi(t-s,x-y) - r^{-n}) f(s,y) d^n y\, ds + \frac{1}{2r^n} \int_{t-r^2/4\pi}^{t} \frac{\rho_n(t-s)}{t-s} \int_{|y-x|=\rho_n(t-s)} u(s,y) dS(y) ds \quad (6.56)$$

for any closed heat ball $E_r(t,x) \subset U_T$.

Proof. We repeat the above calculation with U_t replaced by $E_r(t,x)$ and $v(s,y) = \Phi(t-s,x-y) - r^{-n}$. The only difference is that, since E_r does not have a simple product structure like U_t, we have to use (A.44) for the integration by parts with respect to the time variable. We obtain

$$\lim_{\varepsilon\to 0} \int_{\partial E_r(t,x)} (\Phi(t-s+\varepsilon, x-y) - r^{-n}) u(y,s) \nu_0(s,y) dS(s,y) = \int_{E_r(t,x)} (\Phi(t-s,x-y) - r^{-n}) f(s,y) d^n y\, ds - \int_{t-r^2/4\pi}^{t} \int_{|y-x|=\rho_n(t-s)} \frac{\partial \Phi(t-s,x-y)}{\partial \nu} u(s,y) dS(y) ds.$$

Here ν_0 is the first component of the outward pointing unit normal vector of ∂E_r and $\frac{\partial}{\partial \nu}$ is the normal derivative with respect to the sphere $|y-x| = \rho_n(t-s)$. In particular, $\frac{\partial \Phi(t-s,x-y)}{\partial \nu} = -\frac{|x-y|}{2(t-s)} \Phi(t-s,x-y)$ and the right-hand side equals the right-hand side of (6.56).

To evaluate the limit observe that $\lim_{\varepsilon\to 0}(\Phi(t-s+\varepsilon,x-y) - r^{-n}) = 0$ uniformly on $\partial E_r(t,x)$ away from a neighborhood of (t,x). Hence we can replace u and ν_0 by its value at (t,x) and the left-hand side is given by

$L_r(t)u(t,x)$, where

$$L_r(t) := \lim_{\varepsilon\to 0}\int_{\partial E_r(t,0)} (\Phi(t-s+\varepsilon,y) - r^{-n})dS(s,y).$$

This shows that our representation holds with the left-hand side replaced by $L_r(t)u(t,x)$. Evaluating the right-hand side for $u = 1$ (Problem 6.11) finally shows $L_r = 1$ and establishes the claim. □

Theorem 6.13 (Fulks[5]). *A function $u \in C^{1;2}(U_T)$ satisfies the* ***mean value property***

$$u(t,x) = \frac{1}{2r^n}\int_{t-r^2/4\pi}^{t} \frac{\rho_n(t-s)}{t-s}\int_{|y-x|=\rho_n(t-s)} u(s,y)dS(y)ds \tag{6.57}$$

for every closed heat ball contained in U_T if and only if it solves the heat equation.

Proof. The mean value property follows from the previous lemma. Conversely, if u satisfies the mean value property (6.56) shows

$$\int_{E_r(t,x)} \Phi(t-s,x-y)f(s,y)d^n y\, ds = 0$$

for every $E_r(t,x) \subset U_T$. Hence $f = u_t - \Delta u = 0$ and u solves the heat equation. □

In particular, the heat ball determines a domain of influence for the heat equation: The fact that it only includes points from the past, shows that the equation is deterministic, while the fact that it has a horizontal tangent at its top is connected with infinite propagation speed. If it were contained in a cone (with fixed opening angle), we had finite propagation speed.

Like for harmonic functions, this can also be turned into a volume version where we take the integral over $E_r(t,x)$ by averaging with respect to r:

Corollary 6.14. *Let $u \in C^{1;2}(U_T)$ be a solution of the heat equation and $E_r(t,x) \subset U_T$. Then*

$$u(t,x) = \frac{1}{r^n}\int_{E_r(t,x)} \frac{|x-y|^2}{4(t-s)^2} u(s,y)d^n y\, ds. \tag{6.58}$$

Proof. It will be more convenient to work with $\eta := r^n$ and parametrize $\partial E_r(t,x)$ using

$$\frac{|x-y|^2}{4(t-s)} = \alpha(t-s,\eta), \qquad \alpha(\tau,\eta) := \log(\frac{\eta}{(4\pi\tau)^{n/2}}).$$

[5] Watson Fulks (1919–2001), American mathematician

Then, integrating the mean value formula with respect to η, we obtain

$$\begin{aligned} u(t,x) &= \frac{1}{\eta_0}\int_0^{\eta_0} u(t,x)d\eta \\ &= \frac{1}{\eta_0}\int_0^{\eta_0}\int_{t-\eta^{2/n}/4\pi}^{t}\int_{\frac{|x-y|^2}{4(t-s)}=\alpha(t-s,\eta)} \frac{|x-y|}{2(t-s)}u(s,y)dS(y)ds\frac{d\eta}{\eta} \\ &= \frac{1}{\eta_0}\int_{t-\eta_0^{2/n}/4\pi}^{t}\int_{(4\pi(t-s))^{n/2}}^{\eta_0}\int_{\frac{|x-y|^2}{4(t-s)}=\alpha(t-s,\eta)} \frac{|x-y|}{2(t-s)}u(s,y)dS(y)\frac{d\eta}{\eta}ds. \end{aligned}$$

Next, for $(4\pi(t-s))^{n/2} \le \eta \le \eta_0$ we can make a change of variables (observe $\frac{d\eta}{\eta} = d\alpha$) to obtain

$$\begin{aligned} u(t,x) &= \frac{1}{\eta_0}\int_{t-\eta_0^{2/n}/4\pi}^{t}\int_0^{\alpha(t-s,\eta_0)}\int_{\frac{|x-y|^2}{4(t-s)}=\alpha} \frac{|x-y|}{2(t-s)}u(s,y)dS(y)d\alpha\, ds \\ &= \frac{1}{\eta_0}\int_{t-\eta_0^{2/n}/4\pi}^{t}\int_{\frac{|x-y|^2}{4(t-s)}\le\alpha(t-s,\eta_0)} \frac{|x-y|^2}{4(t-s)^2}u(s,y)d^n y\, ds. \end{aligned}$$

Here we have used radial integration in the last step. □

Let me remark, that if one already knows the mean value formula, it can be verified by showing that the derivative with respect to r vanishes, similarly to what we did for the Laplace equation in Lemma 5.1. However, in case of the heat equation this turns out much more tedious (cf. [**10**, Sect. 2.3.2]).

Of course the mean value formula implies the **strong maximum principle**.

Theorem 6.15 (Strong maximum principle). *Let $U \subset \mathbb{R}^n$ be connected. If a subsolution of the heat equation $u \in C^{1;2}(U_T)$ attains its maximum at $(t_0,x_0) \in U_T$, then u is constant on U_{t_0}.*

Proof. Suppose $u(t_0,x_0) = M := \sup_{(t,x)\in U_T} u(t,x)$ for some $(t_0,x_0) \in U_T$. Let $E_r(t_0,x_0) \subset U_T$ and observe that the submean property implies that in fact $u(t,x) = M$ on $E_r(t_0,x_0)$ (if it were strictly smaller at some point, it would be smaller on a neighborhood by continuity and thus the whole integral would be smaller). Now since U is connected, so is U_{t_0} and hence, for any given $(s,y) \in U_{t_0}$, we can find a path connecting (t_0,x_0) and (s,y). Consider the heat balls with top point on this path. By compactness finitely many cover the path and applying the previous argument recursively shows that $u(t,x) = M$ on this entire path. □

Example 6.3. Note that it is important that the maximum is attained at an interior point. Indeed, the function

$$u(t,x) = x^2 + 2t$$

solves the one-dimensional heat equation on $(0,T] \times [0,1]$ and attains its maximum value at $(T,1)$. ⋄

Corollary 6.16 (Maximum principle). *Let $U \subset \mathbb{R}^n$ be bounded. If $u \in C(\overline{U_T}) \cap C^{1;2}(U_T)$ is a subsolution of the heat equation, then*

$$\max_{\overline{U_T}} u \le \max_{\Gamma_T} u. \tag{6.59}$$

As always there are corresponding minimum principles for supersolutions. Moreover, out of this we get the usual a priori estimates together with uniqueness on bounded domains (cf. Problem 3.11).

Theorem 6.17. *Let $U \subset \mathbb{R}^n$ be a bounded domain and $g \in C(\overline{U})$ and $f \in C(\overline{U_T})$. Then the problem*

$$u_t = \Delta u + f, \qquad \begin{cases} u(0,x) = g(x), & x \in U, \\ u(t,y) = a(t,y), & (t,y) \in [0,T] \times \partial U. \end{cases} \tag{6.60}$$

has at most one solution $u \in C(\overline{U_T}) \cap C^{1;2}(U_T)$. Moreover, u satisfies

$$|u| \le \max_{\overline{U}} |g| + \max_{[0,T]\times\partial U} |a| + T \max_{[0,T]\times\overline{U}} |f|. \tag{6.61}$$

To extend the maximum principle to solutions on $\mathbb{R}^n$ requires an additional growth estimate.

Theorem 6.18 (Maximum principle on $\mathbb{R}^n$). *If $u \in C([0,T]\times\mathbb{R}^n) \cap C^{1;2}((0,T]\times\mathbb{R}^n)$ is a subsolution of the heat equation which satisfies the growth estimate*

$$u(t,x) \le A\mathrm{e}^{a|x|^2}, \qquad x \in \mathbb{R}^n, 0 \le t \le T, \tag{6.62}$$

for some constants $A, a > 0$, then

$$\sup_{(t,x)\in[0,T]\times\mathbb{R}^n} u(t,x) \le \sup_{x\in\mathbb{R}^n} u(0,x). \tag{6.63}$$

Proof. Set $M := \sup_{x\in\mathbb{R}^n} u(0,x)$ and note that we can assume that M is finite since otherwise there is nothing to do. Furthermore, note that we can assume $T < \frac{1}{4a}$ without loss of generality. If this is not the case, split $[0,T]$ into a finite number of intervals satisfying this condition and apply the result successively. Choose some $\delta > 0$ such that $b := \frac{1}{4(T+\delta)} > a$ and note that

$$v(t,x) = u(t,x) - \frac{\varepsilon}{(T+\delta-t)^{n/2}} \mathrm{e}^{|x|^2/4(T+\delta-t)}$$

is a subsolution of the heat equation on $(0,T]\times\mathbb{R}^n$. Moreover, there is some sufficiently large R such that

$$v(t,x) \le A\mathrm{e}^{a|x|^2} - \varepsilon(4b)^{n/2}\mathrm{e}^{b|x|^2} \le M, \qquad 0 \le t \le T,$$

for $|x| \ge R$. Since we also have $v(0,x) \le u(0,x) \le M$ the maximum principle shows that we have $v(t,x) \le M$ on $[0,T]\times B_R(0)$. Combined with our above estimate this shows that we have $v(t,x) \le M$ on all of $[0,T]\times\mathbb{R}^n$. Finally, taking $\varepsilon \to 0$ establishes the claim. □

We remark that the above result is a variation of the classical Phragmén–Lindelöf principle from complex analysis. It immediately gives the following uniqueness result.

Corollary 6.19. *Let $g \in C(\mathbb{R}^n)$ and $f \in C([0,T]\times\mathbb{R}^n)$. Then the problem*

$$u_t = \Delta u + f, \qquad u(0,x) = g(x), \tag{6.64}$$

has at most one solution $u \in C([0,T]\times\mathbb{R}^n) \cap C^{1;2}((0,T]\times\mathbb{R}^n)$ satisfying

$$|u(t,x)| \le A\mathrm{e}^{a|x|^2}, \qquad x \in \mathbb{R}^n, 0 \le t \le T. \tag{6.65}$$

Moreover, u satisfies

$$|u| \le \sup_{\mathbb{R}^n} |g| + T \sup_{[0,T]\times\mathbb{R}^n} |f|. \tag{6.66}$$

Example 6.4. Examples of nontrivial solutions of the heat equation on $\mathbb{R}^n$ which vanish for $t = 0$ were first given by Tikhonov[6], who also established uniqueness under the above growth condition. Here is a simple example in one dimension: Let

$$\varphi(t) := \begin{cases} \mathrm{e}^{-1/t^2}, & t > 0, \\ 0, & t \le 0, \end{cases}$$

which is in $C^\infty(\mathbb{R})$. Then one can show (see Problems 2.3 and 2.4) that

$$u(t,x) := \sum_{m=0}^\infty \varphi^{(m)}(t) \frac{x^{2m}}{(2m)!}$$

converges for all $x \in \mathbb{R}$ and solves the heat equation. Clearly $u(0,x) = 0$. ⋄

We conclude with the observation that solutions are smooth.

Theorem 6.20. *Suppose $u \in C^{1;2}(U_T)$ solves the heat equation. Then $u \in C^\infty(U_T)$.*

Proof. We fix $(t_0,x_0) \in U_T$ and choose $C_r := (t_0 - r, t_0] \times \bar{B}_r(x_0) \subset U_T$. Next choose a smooth cut-off function $\zeta \in C^\infty(U_T)$ such that $0 \le \zeta \le 1$, $\zeta = 1$ on $C_{3r/4}$ and $\zeta = 0$ near the parabolic boundary of C_r. We set $\zeta = 0$ on $([0,t_0]\times\mathbb{R}^n)\setminus C_r)$ such that $v := \zeta u \in C^{1;2}([0,t_0]\times\mathbb{R}^n)$. Then we have

$$v_t - \Delta v = \zeta_t u - 2(\nabla\zeta)\cdot(\nabla u) - u\Delta\zeta =: f \in C^{1;1}([0,t_0]\times\mathbb{R}^n).$$

Hence by Corollary 6.19 and Theorem 6.11

$$v(t,x) = \int_0^t \int_{\mathbb{R}^n} \Phi(t-s, x-y) f(s,y) d^n y\, ds.$$

[6] Andrey Nikolayevich Tikhonov (1906–1993), Russian mathematician and geophysicist

Moreover, for $(t,x) \in C_{r/2}$ the left-hand side equals u and the integral on the right-hand side is over a compact region which avoids the singularity of Φ. In particular, $(t,x) \mapsto \Phi(t-s, x-y)$ is in $C^\infty(C_{r/2})$ for $(s,y) \in \operatorname{supp}(f) \subset C_r \setminus C_{3r/4}$ establishing the claim. $\square$

Problem* 6.11. *Show*

$$\frac{1}{2r^n} \int_{t-r^2/4\pi}^{t} \frac{\rho_n(t-s)}{t-s} \int_{|y-x|=\rho_n(t-s)} dS(y) ds = 1.$$

(Hint: Problem A.6.)

Problem 6.12. *Show that the volume of the heat ball is given by* $|E_r(t,x)| = \frac{1}{2\pi n}\left(\frac{n}{n+2}\right)^{n/2+1} r^{n+2}$. *(Hint: Problem A.6.)*

Problem 6.13. *Let U be a bounded domain and $g \in C(\overline{U})$, $a \in C(\partial U)$ with $g|_{\partial U} = a$. Suppose $u \in C^{1;2}(U_\infty) \cap C(\overline{U_\infty})$ is a solution of the heat equation*

$$u_t = \Delta u, \qquad \begin{cases} u(0,x) = g(x), & x \in U, \\ u(t,y) = a(y), & (t,y) \in [0,\infty) \times \partial U. \end{cases}$$

Show that there are constants C and $\varepsilon > 0$ such that

$$|u(t,x) - v(x)| \le \sup_U |g - v| C \mathrm{e}^{-\varepsilon t},$$

where $v \in C^2(U) \cap C(\overline{U})$ is the solution of the Dirichlet problem

$$\Delta v = 0, \qquad v|_{\partial U} = a.$$

(Hint: Use $w_\varepsilon(t,x) := \cos(\varepsilon x_1)\mathrm{e}^{-\varepsilon^2 t}$ and apply the maximum principle.)

6.4. Energy methods

Like in the one-dimensional case one can consider the **energy**

$$E(t) := \frac{1}{2} \int_U u(t,x)^2 d^n x \tag{6.67}$$

of a solution $u \in C^{1;2}(\overline{U_T})$ of the heat equation satisfying Dirichlet (or Neumann) boundary conditions on ∂U, that is, $u(t,x) = 0$ (or $\frac{\partial u}{\partial \nu}(t,x) = 0$) for $x \in \partial U$. Then one computes using Green's first identity

$$\frac{d}{dt} E(t) = \int_U u u_t \, d^n x = \int_U u \Delta u \, d^n x = -\int_U |\nabla u|^2 d^n x \le 0. \tag{6.68}$$

Hence the energy is nonincreasing and we get uniqueness as well as stability of solutions. Moreover, in the case of Dirichlet boundary conditions one can use the Poincaré inequality (cf. Theorem 9.34)

$$\int_U f(x)^2 d^n x \le C^2 \int_U |\nabla f(x)|^2 d^n x \tag{6.69}$$

to conclude exponential decay of the energy

$$E(t) \le \mathrm{e}^{-t/C^2} E(0), \tag{6.70}$$

which is again Newton's law of cooling. Note that if we prescribe a time independent temperature on the boundary ∂U, then we can subtract the corresponding solution of the Dirichlet problem for the Laplace equation to conclude that the solution will converge to this solution of the Dirichlet problem as $t \to \infty$. In this context note that, by virtue of the mean value formula, decay of the energy implies pointwise decay (Problem 6.14).

In the case of Neumann boundary conditions this cannot work since there are constant solutions. However, there is a variant of the Poincaré inequality (cf. Theorem 9.34) which reads

$$\int_U \big(f(x) - \bar{f}\big)^2 d^n x \le C^2 \int_U |\nabla f(x)|^2 d^n x, \tag{6.71}$$

where $\bar{f} := \frac{1}{|U|} \int_U f\, d^n x$ is the average of f over U. To apply this note that in case of Neumann boundary conditions Green's first identity shows

$$\frac{d}{dt}\bar{u}(t) = \frac{1}{|U|}\int_U u_t(t,x) d^n x = \frac{1}{|U|}\int_U \Delta u(t,x) d^n x = 0 \tag{6.72}$$

that the average is constant. Hence we get

$$\frac{1}{2}\int_U \big(u(t,x) - \bar{u}\big)^2 d^n x \le \mathrm{e}^{-C^{-2}t} \frac{1}{2}\int_U \big(u(0,x) - \bar{u}\big)^2 d^n x \tag{6.73}$$

and u tends to $\bar{u}$ as $t \to \infty$.

Another neat application is

Theorem 6.21 (Backwards uniqueness). *Suppose $u_1, u_2 \in C^2(\overline{U_T})$ solve the heat equation with equal boundary values $u_1(t,x) = u_2(t,x)$ for $(t,x) \in \Gamma_T$. Then $u_1 = u_2$ within $\overline{U_T}$ if $u_1(T,x) = u_2(T,x)$ for $x \in U$.*

Proof. Let

$$E(t) := \frac{1}{2}\int_U w(t,x)^2 d^n x$$

be the energy of the difference $w := u_2 - u_1$. Clearly $E(0) = 0$, $E(T) = 0$ and if we had $E(t) = 0$ for all $t \in [0,T]$ we were done. Now suppose $E(t) > 0$ on some interval (t_0, t_1) and choose this interval maximal, that is, $E(t_0) = E(t_1) = 0$.

We first compute

$$\dot{E}(t) = \int_U w(\Delta w) d^n x = -\int_U |\nabla w|^2 d^n x,$$

$$\ddot{E}(t) = -2\int_U (\nabla w)\cdot(\nabla w_t) d^n x = 2\int_U (\Delta w) w_t d^n x = 2\int_U (\Delta w)^2 d^n x.$$

Now by the Cauchy-Schwarz inequality we have

$$\dot{E}(t)^2 = \left(\int_U w(\Delta w) d^n x\right)^2 \le E(t)\ddot{E}(t).$$

and hence the logarithmic energy $L(t) := \log(E(t))$ is convex in (t_0, t_1) since

$$\ddot{L}(t) = \frac{\ddot{E}(t)}{E(t)} - \frac{\dot{E}(t)^2}{E(t)^2}.$$

Consequently $L((1-\tau)s + \tau t) \le (1-\tau)L(s) + \tau L(t)$ and hence $E((1-\tau)s + \tau t) \le E(s)^{1-\tau} E(t)^\tau$ for all $t_0 < s < t < t_1$ and $0 \le \tau \le 1$. Letting $s \downarrow t_0$ and $t \uparrow t_1$ shows $E((1-\tau)t_0 + \tau t_1) \le E(t_0)^{1-\tau} E(t_1)^\tau = 0$ contradicting our assumption. □

Problem 6.14. *Show that decay of the energy implies pointwise decay. (Hint:* (6.58).*)*

6.5. General parabolic equations

Based on the general elliptic operator L defined in (5.72) we define the corresponding parabolic equation as

$$u_t = -Lu. \tag{6.74}$$

Of course the coefficients of L are now allowed to also depend on t, that is, $A_{jk}, b_j, c \in C_b(U_T)$, where $U_T := (0,T] \times U$ and $\Gamma_T := \overline{U_T} \setminus U_T$ with $U \subset \mathbb{R}^n$ a bounded domain as always.

Again we do not intend to give a detailed treatment of parabolic equations but only to demonstrate how some of the ideas from the heat equation can be extended to this more general case. We first note that the proof of the maximum principle from Theorem 3.4 still applies.

Theorem 6.22 (Maximum principle). *Let $U \subset \mathbb{R}^n$ be a bounded domain, $v \in C(\overline{U_T}) \cap C^{1;2}(U_T)$ and suppose $c \ge 0$. If $v_t + Lv \le 0$ then*

$$\max_{\overline{U_T}} v \le \max_{\Gamma_T} v^+. \tag{6.75}$$

If $v_t + Lv \ge 0$ then

$$\min_{\overline{U_T}} v \ge -\max_{\Gamma_T} v^-. \tag{6.76}$$

Here $v^\pm = \max(\pm v, 0)$ is the positive, negative part of v, respectively.

Proof. Following the argument from Theorem 3.4 using the extension to $\mathbb{R}^n$ from the proof of Theorem 5.33 shows that there can be no interior maximum at which u is nonnegative. Hence the first claim follows. For the second claim replace v by $-v$. □

If $u_t + Lu = 0$ we can combine both estimates to obtain

$$-\max_{\Gamma_T} u^- \leq u(t,x) \leq \max_{\Gamma_T} u^+, \tag{6.77}$$

which shows for example $0 \leq u \leq 1$ if this inequality holds for the initial and boundary data.

Corollary 6.23. *Let $U \subset \mathbb{R}^n$ be a bounded domain, $g \in C(\overline{U})$, $f \in C(\overline{U_T})$ and suppose $c \geq 0$. Then the problem*

$$u_t = -Lu + f, \qquad \begin{cases} u(0,x) = g(x), & x \in U, \\ u(t,y) = a(t), & (t,y) \in [0,T] \times \partial U. \end{cases} \tag{6.78}$$

has at most one solution $u \in C(\overline{U_T}) \cap C^{1;2}(U_T)$. Moreover, u satisfies

$$|u| \leq \max_{\overline{U}} |g| + \max_{[0,T]\times\partial U} |a| + T \max_{[0,T]\times\overline{U}} |f|. \tag{6.79}$$

Moreover, we have the following **comparison principle**:

Corollary 6.24. *Let $u, v \in C(\overline{U_T}) \cap C^{1;2}(U_T)$ with $u_t + Lu = 0$ and $v_t + Lv \leq 0$. Then $v \leq u$ on the parabolic boundary Γ_T implies $v \leq u$ on all of $\overline{U_T}$.*

Proof. If $c \geq 0$ the result follows immediately by applying the maximum principle to $w := v - u$. In the general case we set $c_0 := \inf_{\overline{U_T}} c$ and consider $\tilde{w}(t,x) = \mathrm{e}^{c_0 t} w(t,x)$ which satisfies $\tilde{w}_t + (L - c_0)\tilde{w} \leq 0$ and reduces it to the case $c \geq 0$. □

The strong maximum principle is more involved and was first established by Nirenberg[7]. The key is the following lemma:

Lemma 6.25. *Let $u \in C^{1;2}(U_T)$ and suppose $c = 0$. If $v_t + Lv \leq 0$ and v attains a maximum at an interior point $(t_0, x_0) \in U_T$, then v is constant along every line emanating from (t_0, x_0) for as long as this line stays within U_{t_0}.*

If $c \geq 0$ the same conclusion holds if v attains a nonnegative maximum at an interior point.

Proof. Without loss of generality, we can assume $x_0 = 0$. We will first look at the vertical line $t \mapsto (t, 0)$. Denote the maximum by $M := v(t_0, 0)$ and assume we had $v(t_1, 0) < M$ for some $0 < t_1 < t_0$.

In this case we can choose $a > 0$ and $r \in (0,1)$ sufficiently small such that $v(t_1, x) \leq M - a$ for $x \in \bar{B}_r(0) \subset U$. Then, if $cM \geq 0$, we can choose $\lambda > 0$ such that (Problem 6.15)

$$w(t,x) := M - a\mathrm{e}^{-\lambda(t-t_1)}\big(r^2 - |x|^2\big)^2$$

[7] Louis Nirenberg (1925–2020), Canadian-American mathematician

is a supersolution on $[t_1, t_0] \times B_r(0)$. Moreover, we have $w \ge M - a$ on the parabolic boundary of $(t_1, t_0] \times B_r(0)$ and hence $w \ge v$ on $(t_1, t_0] \times B_r(0)$ by the maximum principle, a contradiction.

The case of a general line $t \mapsto (t, \beta(t - t_0))$ can be reduced to the case $\beta = 0$ by virtue of the change of coordinates $y = x - \beta(t - t_0)$. Then the transformed operator satisfies the same assumptions (indeed, we just have $b \to b - \beta$) and the transformed function will be defined on $(t_2, t_1] \times B_r(0)$ for t_2 larger than the first intersection of our line with U_{t_0} and r sufficiently small. □

The strong maximum principle now follows effortless.

Theorem 6.26 (Strong maximum principle). *Let $U \subset \mathbb{R}^n$ be connected and $c = 0$. If a subsolution $v \in C^{1;2}(U_T)$ attains its maximum at $(t_0, x_0) \in U_T$, then v is constant on U_{t_0}. If $c \ge 0$ the same conclusion holds if v attains a nonnegative maximum at an interior point.*

Proof. Just follow the proof of Theorem 6.15 and use that a path connecting two points can be assumed piecewise linear without loss of generality since U is open (cover a path by balls, by compactness finitely many suffice, now connect the centers to get a piecewise linear path). □

Also the energy methods apply if one assumes the divergence form $\tilde{L}$ with $b = 0$ and $c \ge 0$. Then defining

$$E(t) := \frac{1}{2} \int_U u(t,x)^2 d^n x \tag{6.80}$$

of a solution $u \in C^{1;2}(\overline{U_T})$ of the heat equation satisfying Dirichlet (or Neumann) boundary conditions on ∂U, that is, $u(t,x) = 0$ (or $\frac{\partial u}{\partial \nu}(t,x) = 0$) for $x \in \partial U$. Then one computes using integration by parts

$$\frac{d}{dt} E(t) = \int_U u \tilde{L} u \, d^n x = - \int_U \big(\nabla u \cdot (A \nabla u) + c\, u^2 \big) d^n x \le 0. \tag{6.81}$$

Problem* 6.15. *Show that the function w defined in the proof of Lemma 6.25 is a supersolution for λ sufficiently large.*

Problem 6.16. *Suppose $c(t,x) \ge -c_0$ with $c_0 \in \mathbb{R}$. Show that if $v \in C(\overline{U_T}) \cap C^{1;2}(U_T)$ satisfies $v_t + Lv \le 0$, then*

$$v(t,x) \le \mathrm{e}^{c_0 t} \max_{(t,x) \in \Gamma_T} v(t,x)^+, \qquad (x,t) \in \overline{U_T}.$$

Problem 6.17. *Show that solutions $u \in C(\overline{U_T}) \cap C^{1;2}(U_T)$ of the* ***Fisher–Kolmogorov–Petrovsky–Piskunov equation***[8]

$$u_t - \Delta u = ru(1-u), \qquad r \in \mathbb{R},$$

on a bounded domain U satisfy $0 \le u \le 1$ provided this holds for the initial condition. (Hint: Choose c such that u satisfies a linear equation.)

[8]Ronald Fisher (1890–1962), British statistician and biologist
[8]Andrey Kolmogorov (1903–1987), Soviet mathematician
[8]Ivan Petrovsky (1901–1973), Soviet mathematician
[8]Nikolai Piskunov (1908–1977), Soviet mathematician

The wave equation

7.1. Solution via the Fourier transform

Finally we turn to the **wave equation**

$$\Box u := u_{tt} - \Delta u = 0, \qquad u(0) = g, \quad u_t(0) = h, \tag{7.1}$$

where

$$\Box := \frac{\partial^2}{\partial t^2} - \Delta. \tag{7.2}$$

is known as the **d'Alembert operator**.

After applying the Fourier transform this reads

$$\hat{u}_{tt} + |k|^2 \hat{u} = 0, \qquad \hat{u}(0) = \hat{g}, \quad \hat{u}_t(0) = \hat{h}, \tag{7.3}$$

and the solution is given by

$$\hat{u}(t,k) = \cos(t|k|)\hat{g}(k) + \frac{\sin(t|k|)}{|k|}\hat{h}(k). \tag{7.4}$$

Note that the solution for the case $u(0) = g$, $u_t(0) = 0$ can be obtained by differentiating the solution for $u(0) = 0$, $u_t(0) = g$ with respect to t (**Stokes' rule** — Problem 7.6). Moreover, we can assume $t > 0$ without loss of generality since if $u(t,x)$ is a solution, so will be $v(t,x) := u(-t,x)$ (corresponding to the initial conditions $v(0) = g$, $v_t(0) = -h$).

To obtain the analog of d'Alembert's formula in $n = 3$ dimensions we can hence restrict our attention to the case $g = 0$. We will use the fact that the Fourier transform of

$$\varphi_t(x) := \sqrt{\frac{\pi}{2}} \frac{\chi_{[0,|t|]}(|x|)}{|x|}. \tag{7.5}$$

is given by (Problem 6.4)

$$\hat{\varphi}_t(k) = \frac{1 - \cos(t|k|)}{|k|^2}. \tag{7.6}$$

Clearly

$$\frac{\partial}{\partial t}\hat{\varphi}_t(k) = \frac{\sin(t|k|)}{|k|}. \tag{7.7}$$

and as in the one-dimensional case (cf. Lemma 4.7) we can use

Lemma 7.1. *Suppose f, g are integrable. Then*

$$\mathcal{F}^{-1}(\hat{f}g)(x) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} f(x-y)\check{g}(y) d^n y \tag{7.8}$$

to obtain (assuming $h \in \mathcal{S}(\mathbb{R}^3)$ and using Lemma A.11 to compute the derivative pointwise)

$$\begin{aligned} u(t,x) &= \frac{1}{4\pi}\frac{\partial}{\partial t}\int_{B_{|t|}(x)} \frac{1}{|x-y|} h(y) d^3 y \\ &= \frac{1}{4\pi}\frac{\partial}{\partial t}\int_0^{|t|}\int_{S^2} \frac{1}{r} h(x+r\omega) r^2 d\sigma^2(\omega) dr \\ &= \frac{t}{4\pi}\int_{S^2} h(x+t\omega) d\sigma^2(\omega). \end{aligned} \tag{7.9}$$

Thus we arrive at **Kirchhoff's formula**[1]

$$\begin{aligned} u(t,x) &= \frac{\partial}{\partial t}\frac{t}{4\pi}\int_{S^2} g(x+t\omega)d\sigma^2(\omega) + \frac{t}{4\pi}\int_{S^2} h(x+t\omega) d\sigma^2(\omega) \\ &= \frac{1}{4\pi}\int_{S^2}\big(g + t\,\nabla g\cdot\omega + t\,h\big)(x+t\omega)d\sigma^2(\omega) \\ &= \frac{1}{4\pi t^2}\int_{\partial B_{|t|}(x)}\big(g(y) + \nabla g(y)\cdot(y-x) + t\,h(y)\big) dS(y). \end{aligned} \tag{7.10}$$

In fact, it is straightforward to check that Kirchhoff's formula provides the solution under much weaker assumptions:

Theorem 7.2. *Suppose $h \in C^2(\mathbb{R}^3)$ and $g \in C^3(\mathbb{R}^3)$. Then $u(t,x)$ defined via Kirchhoff's formula (7.10) is in $C^2(\mathbb{R}^4)$ and solves the Cauchy problem (7.1) for the wave equation.*

Proof. By Stokes' rule (Problem 7.6) it suffices to establish the case $g = 0$. Interchanging differentiation and integration one computes (using Green's

[1] Gustav Kirchhoff (1824–1887), German physicist

first identity)

$$\begin{aligned}
u_t(t,x) &= \frac{1}{4\pi}\int_{S^2}\big(h(x+t\omega)+t\,\nabla h(x+t\omega)\cdot\omega\big)d\sigma^2(\omega)\\
&= \frac{1}{4\pi}\int_{S^2} h(x+t\omega)d\sigma^2(\omega)+\frac{1}{4\pi t}\int_0^t\int_{S^2}\Delta h(x+r\omega)d\sigma^2(\omega)r^2dr,\\
u_{tt}(t,x) &= \frac{t}{4\pi}\int_{S^2}\Delta h(x+t\omega)d\sigma^2(\omega)
\end{aligned}$$

and

$$\begin{aligned}
u_{x_j}(t,x) &= \frac{t}{4\pi}\int_{S^2} h_{x_j}(x+t\omega)d\sigma^2(\omega),\\
u_{x_jx_k}(t,x) &= \frac{t}{4\pi}\int_{S^2} h_{x_jx_k}(x+t\omega)d\sigma^2(\omega),
\end{aligned}$$

which shows that $u\in C^2(\mathbb{R}^4)$ and

$$\Delta u(t,x)=\frac{t}{4\pi}\int_{S^2}\Delta h(x+t\omega)d\sigma^2(\omega)=u_{tt}(t,x).$$

Since $u(0,x)=0$ and $u_t(0,x)=h(x)$ we conclude that u also satisfies the initial condition. □

It is interesting to observe that we need to require $g\in C^3$ in order to get a solution which is C^2 (in contradistinction to the one-dimensional case, where $g\in C^2$ was sufficient). This is due to the fact that irregularities of the initial conditions could be focused at one point during the time evolution.

Example 7.1. Let us look at the case of radial initial conditions $g(x):=\tilde{g}(|x|)$, $h(x):=0$. Then the solution at 0 is given by

$$u(t,0)=\tilde{g}(t)+\tilde{g}'(t).$$

For the special choice

$$\tilde{g}(r):=\begin{cases}\sqrt{1-r}, & 0\le r\le 1,\\ 0, & r>1,\end{cases}$$

(which is continuous but not C^3) we get

$$u(0,t)=\sqrt{1-t}-\frac{1}{2\sqrt{1-t}},\qquad 0\le t<1,$$

which shows that the irregularity of the initial condition along $|x|=1$ gets focused in the origin at $t=1$, leading to a blow-up of the solution. ⋄

As always, the solution of the inhomogeneous problem follows from the **Duhamel principle**:

Corollary 7.3. *Suppose* $f \in C^{0;2}(\mathbb{R} \times \mathbb{R}^3)$ *then*

$$u(t,x) = \frac{1}{4\pi} \int_0^t (t-s) \int_{S^2} f(s, x + (t-s)\omega) d\sigma^2(\omega) ds \tag{7.11}$$

is in $C^2(\mathbb{R}^4)$ *and solves the inhomogeneous wave equation*

$$u_{tt} = \Delta u + f, \qquad u(0) = u_t(0) = 0. \tag{7.12}$$

Proof. Let us write

$$u(t,x) = \int_0^t U(s,t,x) ds, \quad U(s,t,x) := \frac{(t-s)}{4\pi} \int_{S^2} f(s, x + (t-s)\omega) d\sigma^2(\omega),$$

where $U \in C^{0;2}(\mathbb{R} \times \mathbb{R}^4)$ with $U_{tt} = \Delta U$ and $U(s,s,x) = 0$, $U_t(s,s,x) = f(s,x)$ by Theorem 7.2.

Hence we obtain using Leibniz integral rule (Problem A.12):

$$u_t(t,x) = U(t,t,x) + \int_0^t U_t(s,t,x) ds = \int_0^t U_t(s,t,x) ds,$$
$$u_{tt}(t,x) = U_t(t,t,x) + \int_0^t U_{tt}(s,t,x) ds = f(t,x) + \int_0^t U_{tt}(s,t,x) ds.$$

In particular, $u(0,x) = u_t(0,x) = 0$. Moreover, interchanging integration and differentiation we obtain

$$\Delta u(t,x) = \int_0^t \Delta U(s,t,x) ds,$$

which establishes the claim. □

Note that for $t > 0$ (7.11) can be rewritten as

$$\begin{aligned} u(t,x) &= \frac{1}{4\pi} \int_0^t \frac{1}{t-s} \int_{\partial B_{t-s}(x)} f(s,y) dS(y) ds \\ &= \frac{1}{4\pi} \int_0^t \frac{1}{r} \int_{\partial B_r(x)} f(t-r, y) dS(y) dr \\ &= \frac{1}{4\pi} \int_{B_t(x)} \frac{f(t - |x-y|, y)}{|x-y|} d^3y, \end{aligned} \tag{7.13}$$

where the last integrand is known as **retarded potential** (since it reassembles the Newton potential, but the time t is *retarded* by the factor $|x-y|$).

Finally, to obtain a formula in $n = 2$ dimensions we use Hadamard's **method of descent**: That is we use the fact, that a solution in two dimensions is also a solution in three dimensions which happens to be independent of the third coordinate direction. Indeed, if g and h are independent of the third coordinate, so will be u defined via Kirchhoff's formula. Hence it remains to simplify Kirchhoff's formula in the case h does not depend on x_3: Using spherical coordinates we obtain

$$\begin{aligned}
\frac{t}{4\pi}\int_{S^2} & h(x+t\omega)d\sigma^2(\omega) = \\
&= \frac{t}{2\pi}\int_0^{\pi/2}\int_0^{2\pi} h(x_1+t\sin(\theta)\cos(\varphi), x_2+t\sin(\theta)\sin(\varphi))\sin(\theta)d\varphi d\theta \\
&\overset{\rho=\sin(\theta)}{=} \frac{t}{2\pi}\int_0^1\int_0^{2\pi}\frac{h(x_1+t\rho\cos(\varphi), x_2+t\rho\sin(\varphi))}{\sqrt{1-\rho^2}}d\varphi\,\rho d\rho \\
&= \frac{t}{2\pi}\int_{B_1(0)}\frac{h(x+ty)}{\sqrt{1-|y|^2}}d^2y,
\end{aligned}$$

which gives **Poisson's formula**

$$\begin{aligned}
u(t,x) &= \frac{\partial}{\partial t}\frac{t}{2\pi}\int_{B_1(0)}\frac{g(x+ty)}{\sqrt{1-|y|^2}}d^2y + \frac{t}{2\pi}\int_{B_1(0)}\frac{h(x+ty)}{\sqrt{1-|y|^2}}d^2y \\
&= \frac{1}{2\pi}\int_{B_1(0)}\frac{g(x+ty)+t\nabla g(x+ty)\cdot y + t\,h(x+ty)}{\sqrt{1-|y|^2}}d^2y \\
&= \frac{1}{2\pi|t|}\int_{B_{|t|}(x)}\frac{g(y)+\nabla g(y)\cdot(y-x)+t\,h(x)}{\sqrt{t^2-|x-y|^2}}d^2y. \qquad (7.14)
\end{aligned}$$

Theorem 7.4. *Suppose $h \in C^2(\mathbb{R}^2)$ and $g \in C^3(\mathbb{R}^2)$. Then $u(t,x)$ defined via Poisson's formula* (7.14) *is in $C^2(\mathbb{R}^3)$ and solves the Cauchy problem* (7.1) *for the wave equation.*

Corollary 7.5. *Suppose $f \in C^{0;2}(\mathbb{R}\times\mathbb{R}^2)$ then*

$$u(t,x) = \frac{1}{2\pi}\int_0^t (t-s)\int_{B_1(0)}\frac{f(s,x+(t-s)y)}{\sqrt{1-|y|^2}}d^2y\,ds \qquad (7.15)$$

is in $C^2(\mathbb{R}^3)$ and solves the inhomogeneous wave equation (7.12).

Note that, if the inhomogeneous term f does not depend on t, then one could subtract the corresponding solution of the Poisson problem to reduce the problem to a homogenous one.

Uniqueness will be addressed in Section 7.3.

Finally, observe that there is a striking difference between three and two dimensions: While in three dimensions $u(x,t)$ depends only on the values of the initial data on $\partial B_t(x)$, in two dimensions it depends on the values of the initial data on all of $B_t(x)$. Hence the values of the initial data at x affect the solution only on the boundary of the cone $C(x,t) := \{(t,y) | |x-y| < |t|\}$ in three dimensions, while they affect the solution on the entire cone $C(x,t)$ in two dimensions. Consequently, the effect of a disturbance is only felt at the front of its propagation in three dimensions, while in two dimensions it is felt forever after the front passed. This is known as **Huygens' principle**.

Problem 7.1. *Guess the solution corresponding to the initial conditions* $u(0,x) = g(a \cdot x)$, $u_t(0,x) = g'(a \cdot x)$ *with a given unit vector* $a \in \mathbb{R}^n$ *and* $g \in C^2(\mathbb{R})$.

Problem 7.2. *Find the solution of the wave equation with initial condition* $g(x) = 0$, $h(x) = |x|^2$ *in* $\mathbb{R}^3$.

Problem 7.3. *Suppose* g *and* h *are supported in a ball of radius* r. *If you observe the solution of the wave equation* $u(t,x)$ *at a distance* $|x| = R > r$ *from the origin. What is the first time* t_0 *after which you can notice the effects of the initial conditions in two, three dimensions, respectively? After what time* $t_1 > t_0$ *can you no longer notice any effects?*

Problem 7.4. *Suppose* g *and* h *are supported in a ball of radius* r. *Use Kirchhoff's formula to show*

$$|u(t,x)| \le \frac{Cr}{|t|}\big(\sup_{\mathbb{R}^3} |g| + r \sup_{\mathbb{R}^3} |\partial g| + r \sup_{\mathbb{R}^3} |h|\big).$$

(Hint: What is the area of $\partial B_{|t|}(x) \cap B_r(0)$*?)*

Problem 7.5. *Derive a formula for the Fourier transform* $\hat{u}(t,k)$ *of a solution of the* ***Klein–Gordon equation***

$$u_{tt} = \Delta u - m^2 u, \qquad u(0) = g, \quad u_t(0) = h,$$

with $m > 0$.

Problem* 7.6 (Stokes' rule)**.** *Let* $u \in C^3(\mathbb{R}^n)$ *be a solution of the Klein–Gordon equation with initial condition* $u(0,x) = 0$, $u_t(0,x) = h(x)$. *Then* $v := u_t$ *solves the Klein–Gordon equation with initial condition* $v(0,x) = h(x)$, $v_t(0,x) = 0$.

Problem 7.7. *Show*

$$\sqrt{\frac{\pi}{2}} \frac{m}{(2\pi)^{3/2}} \int_{-1}^{1} \frac{J_1(m\sqrt{1-|x|^2})}{\sqrt{1-|x|^2}} \mathrm{e}^{-\mathrm{i}k\cdot x} d^3x = \frac{\sin(|k|)}{|k|} - \frac{\sin(\sqrt{|k|^2+m^2})}{\sqrt{|k|^2+m^2}},$$

where $J_1(z)$ *is the Bessel function of order* 1. *(Hint: Start with Problem 4.22 and use Problem 6.3. Problem 3.22 will also be helpful.)*

Problem 7.8. *Show that the solution of the Klein–Gordon equation in* $n = 3$ *dimensions with initial condition* $u(0,x) = 0$, $u_t(0,x) = h(x)$ *is given by*

$$u(t,x) = \frac{t}{4\pi} \int_{S^2} h(x + t\omega) d\sigma^2(\omega) - \frac{m \operatorname{sign}(t)}{4\pi} \int_{B_x(|t|)} \frac{J_1(m\sqrt{t^2-|x|^2})}{\sqrt{t^2-|x|^2}} h(x-y) d^3y.$$

The solution with initial condition $u(0,x) = g(x)$, $u_t(0,x) = 0$ of course follows from Stokes' rule (Problem 7.6). (Hint Problems 7.5 and 7.7.)

Problem 7.9. ***Maxwell's equations***[2] *in vacuum for the electric field $E(t,x)$ and the magnetic field $B(t,x)$ are given by*

$$B_t = -\operatorname{curl} E, \quad \mu_0\varepsilon_0 E_t = \operatorname{curl} B, \quad \operatorname{div} E = 0, \quad \operatorname{div} B = 0,$$

where $\mu_0 > 0$ and $\varepsilon_0 > 0$ are the permeability and the permittivity of the vacuum, respectively. Here $\operatorname{curl} f := \nabla \times f$ is the infinitesimal circulation (also known as rotation) of a vector field f in $\mathbb{R}^3$. Find the solution corresponding to the initial conditions $E(0,x) = E_0(x)$ and $B(0,x) = B_0(x)$ (assumed to satisfy $\operatorname{div} B_0 = \operatorname{div} E_0 = 0$). (Hint: Show that the components of both E and B satisfy the wave equation with $c := (\mu_0\varepsilon_0)^{-1/2}$.)

7.2. The solution in arbitrary dimensions

While (7.4) works for arbitrary dimensions, we were only able to find an explicit solution formula for u in $n = 2$ and $n = 3$ dimensions in the previous section. While in principle it is possible to find an explicit formula using the fact that the Fourier transform of radial functions in $n+2$ dimensions can be computed from the n dimensional result (cf. Problem 6.3) by applying $\frac{1}{r}\frac{\partial}{\partial r}$, this requires distribution theory. Hence we will resort to a more elementary approach due to Poisson. We will assume $t \geq 0$ for notational simplicity. In this respect note, that if $u(t,x)$ solves (7.1) for $t \geq 0$ then $v(t,x) := u(-t,x)$ solves the wave equation for $t \leq 0$ but with initial conditions $v(0,x) = g(x)$, $v_t(0,x) = -h(x)$.

The idea is to look at the **spherical means**

$$U(t,r,x) := \frac{1}{nV_n r^{n-1}} \int_{\partial B_r(x)} u(t,y) dS(y). \tag{7.16}$$

If $u \in C^2(\mathbb{R}^{n+1})$ solves the wave equation then one can verify (Problem 7.10) that $U \in C^2(\mathbb{R}^{n+1} \times [0,\infty))$ solves the **Euler–Poisson–Darboux equation**[3]

$$U_{tt} = r^{1-n}\big(r^{n-1}U_r\big)_r = U_{rr} + \frac{n-1}{r}U_r. \tag{7.17}$$

While this almost looks like a one-dimensional wave equation, there is an undesired extra term. If $n = 3$ the Euler–Poisson–Darboux equation can be reduced to the wave equation upon considering $\tilde{U}(t,r,x) := rU(t,r,x)$. In arbitrary odd dimensions the following more complicated transformation works:

[2] James Clerk Maxwell (1831–1879), Scottish mathematical physicist

[3] Jean Gaston Darboux (1842–1917), French mathematician

Lemma 7.6. *For every $f \in C^{k+1}(0,\infty)$ we have*

$$\frac{\partial^2}{\partial r^2} D_r^{k-1} r^{2k-1} f(r) = D_r^{k-1} r^{2k-1} \frac{1}{r^{2k}} \frac{\partial}{\partial r} r^{2k} \frac{\partial}{\partial r} f(r), \quad D_r := \frac{1}{r}\frac{\partial}{\partial r}. \tag{7.18}$$

Proof. We use induction on k. The case $k = 1$ is easy

$$\partial_r^2 r = \partial_r (1 + r\partial_r) = 2\partial_r + r\partial_r^2 = \frac{1}{r}\partial_r r^2 \partial_r.$$

Similarly the induction step follows

$$\begin{aligned} \partial_r^2 D_r^k r^{2k+1} &= \partial_r^2 D_r^{k-1} r^{2k-1} \big((2k+1) + r\partial_r\big) \\ &= D_r^k r^{2k} \partial_r \big((2k+1) + r\partial_r\big) = D_r^k D_r r^{2k+2} \partial_r. \end{aligned}$$ □

Hence if $n = 2k+1$ is odd, we conclude that

$$\tilde{U}(t,r,x) := D_r^{k-1} r^{2k-1} U(t,r,x) \tag{7.19}$$

will satisfy the one-dimensional wave equation provided $U \in C^{k+1}$ satisfies the Euler–Poisson–Darboux equation. Moreover, the above spherical averages satisfy $U_r(t,x,0) = 0$ implying (see Problem 7.11) $\tilde{U}_r(t,x,0) = 0$ and consequently (Problem 4.16)

$$\tilde{U}(t,r,x) = \frac{\tilde{G}(t+r,x) - \tilde{G}(t-r,x)}{2} + \frac{1}{2}\int_{t-r}^{t+r} \tilde{H}(\rho,x) d\rho \tag{7.20}$$

for $0 \le r \le t$. Here $\tilde{G}(r,x) := \tilde{U}(0,r,x)$ and $\tilde{H}(r,x) := \tilde{U}_t(0,r,x)$. Finally, note that we can recover u from

$$u(t,x) = \lim_{r\to 0} U(t,r,x) = \lim_{r\to 0} \frac{\tilde{U}(t,r,x)}{(n-2)!!r} \tag{7.21}$$

(with $n!! = n \cdot (n-2) \cdots 3 \cdot 1$ for n odd) and hence

$$u(t,x) = \frac{1}{(n-2)!!}\big(\tilde{G}_r(t,x) + \tilde{H}(t,x)\big). \tag{7.22}$$

In summary, if u is a sufficiently smooth solution of the wave equation it is given by

$$\begin{aligned} u(t,x) = {} & \frac{\partial}{\partial t} D_t^{(n-3)/2} \frac{1}{n!!V_n t} \int_{\partial B_t(x)} g(y) dS(y) \\ & + D_t^{(n-3)/2} \frac{1}{n!!V_n t} \int_{\partial B_t(x)} h(y) dS(y). \end{aligned} \tag{7.23}$$

Conversely, one can check that the above formula provides a solution of the wave equation provided g and h are sufficiently smooth:

Theorem 7.7. *Let $n = 2k+1$ be odd and suppose $h \in C^{k+1}(\mathbb{R}^n)$ and $g \in C^{k+2}(\mathbb{R}^n))$. Then $u(t,x)$ defined via (7.23) is in $C^2(\mathbb{R}^{n+1})$ and solves the Cauchy problem (7.1) for the wave equation.*

Proof. By Stokes' rule (Problem 7.6) it suffices to establish the case $g = 0$. Abbreviate

$$H(t,x) := \frac{1}{nV_n t^{n-1}} \int_{\partial B_t(x)} h(y) dS(y) = \frac{1}{nV_n} \int_{S^{n-1}} h(x+t\omega) d\sigma^{n-1}(\omega)$$

and note that $H(t,x) \in C^{k+1}(\mathbb{R}^{n+1})$. Using Lemma 7.6 and Problem A.11 we find

$$\begin{aligned} u_{tt} &= \frac{n}{n!!} D_t^k t^{2k} \frac{\partial}{\partial t} H = \frac{1}{n!!V_n} D_t^k \int_{B_t(x)} \Delta h(y) d^n y \\ &= D_t^{k-1} \frac{1}{n!!V_n t} \int_{\partial B_t(x)} \Delta h(y) dS(y). \end{aligned}$$

On the other hand,

$$\Delta H = \frac{1}{nV_n t^{n-1}} \Delta \int_{\partial B_t(0)} h(x+y) dS(y) = \frac{1}{nV_n t^{n-1}} \int_{\partial B_t(x)} \Delta h(y) dS(y),$$

which shows $u_{tt} = \Delta u$.

That u satisfies the initial conditions follows from Problem 7.11. □

The case of even dimensions follows using the method of descent. An analogous computation to the one for $n = 2$ dimensions gives

$$\begin{aligned} u(t,x) = &\frac{\partial}{\partial t} D_t^{(n-2)/2} \frac{1}{n!!V_n} \int_{B_t(x)} \frac{g(y)}{\sqrt{t^2 - |y-x|^2}} d^n y \\ &+ D_t^{(n-2)/2} \frac{1}{n!!V_n} \int_{B_t(x)} \frac{h(y)}{\sqrt{t^2 - |y-x|^2}} d^n y. \end{aligned} \tag{7.24}$$

Theorem 7.8. *Let $n = 2k$ be even and suppose $h \in C^{k+1}(\mathbb{R}^n)$ and $g \in C^{k+2}(\mathbb{R}^n))$. Then $u(t,x)$ defined via (7.24) is in $C^2(\mathbb{R}^{n+1})$ and solves the Cauchy problem (7.1) for the wave equation.*

Finally, the inhomogeneous case follows from the Duhamel principle (literally follow the proof of Corollary 7.3):

Corollary 7.9. *Suppose $f \in C^{0;\lfloor n/2 \rfloor + 1}(\mathbb{R} \times \mathbb{R}^n)$ and let $U(s,t,x) \in C^{0;2}(\mathbb{R} \times \mathbb{R}^{n+1})$ be the solution of $U_{tt} = \Delta U$ with initial conditions $U(s,s,x) = 0$, $U_t(s,s,x) = f(s,x)$ given by Theorem 7.7, Theorem 7.8 in even, odd dimensions, respectively. Then*

$$u(t,x) = \int_0^t U(s,t,x) ds \tag{7.25}$$

is in $C^2(\mathbb{R}^{n+1})$ and solves the inhomogeneous wave equation

$$u_{tt} = \Delta u + f, \qquad u(0) = u_t(0) = 0. \tag{7.26}$$

Problem* 7.10. *Show that* (7.16) *is in* $C^2(\mathbb{R}^{n+1}\times[0,\infty))$ *and solves the Euler–Poisson–Darboux equation* (7.17) *as well as* $U_r(t,x,0)=0$ *if* $u\in C^2(\mathbb{R}^{n+1})$ *solves the wave equation. (Hint: Problem A.11.)*

Problem* 7.11. *For* $k\in\mathbb{N}$ *and* $f\in C^{k-1}(0,\infty)$ *we have*

$$D_r^{k-1}r^{2k-1}f(r)=\sum_{j=0}^{k-1}\alpha_{k,j}r^{j+1}f^{(j)}(r)$$

with $\alpha_{k,0}=(2k-1)!!$.

Problem 7.12. *Show that that if the initial conditions* f *and* g *are radial, so is the corresponding solution of the wave equation. (Hint: Uniqueness.)*

Problem 7.13. *Show that the Klein–Gordon equation can be solved using a variant of the method of descent: If* $u(t,x)$ *is a solution for* $x\in\mathbb{R}^n$, *find a suitable function* ϕ *such that* $v(t,\bar{x})=u(t,x)\phi(x_{n+1})$ *solves the wave equation for* $\bar{x}=(x,x_{n+1})\in\mathbb{R}^{n+1}$.

Problem 7.14. *Find all nontrivial radial solutions of the wave equation of the form* $u(t,x)=\alpha(r)f(t-\delta(r))$ *with* $|x|=r$, *where* $f\in C^2$ *is an arbitrary function and* α, δ *are fixed functions independent of* f.

7.3. Energy methods

Let U be a bounded domain and $u\in C^2(\overline{U_T})$ a solution of the wave equation. We define its **energy** to be

$$E(t):=\frac{1}{2}\int_U\big(|\nabla u(t,x)|^2+u_t(t,x)^2\big)d^nx. \tag{7.27}$$

A straightforward calculation verifies that if u satisfies Dirichlet (or Neumann) boundary conditions on ∂U, then $\dot{E}(t)=0$ and hence $E(t)=E(0)$ is a constant of motion (Problem 7.15).

Lemma 7.10. *Let* $U\subset\mathbb{R}^n$ *be a bounded domain and* $u\in C^2(\mathbb{R}\times U)\cap C(\mathbb{R}\times\overline{U})$ *a solution of the wave equation satisfying Dirichlet boundary conditions* $u(t,.)|_{\partial U}=0$. *Then the energy* $E(t)$ *is constant. The same result holds if* $u\in C^2(\mathbb{R}\times U)\cap C^1(\mathbb{R}\times\overline{U})$ *satisfies Neumann boundary conditions* $\frac{\partial u}{\partial\nu}(t,.)|_{\partial U}=0$.

Hence by considering the difference of two solutions we get uniqueness as well as stability of solutions.

Theorem 7.11. *There exists at most one solution* $u\in C^2(\mathbb{R}\times U)\cap C(\mathbb{R}\times\overline{U})$ *of the problem*

$$u_{tt}-\Delta u=0,\qquad u(0)=g,\quad u_t(0)=h,\quad u(t)|_{\partial U}=a(t), \tag{7.28}$$

for given data $g,h\in C(U)$, $a\in C(\mathbb{R}\times\partial U)$.

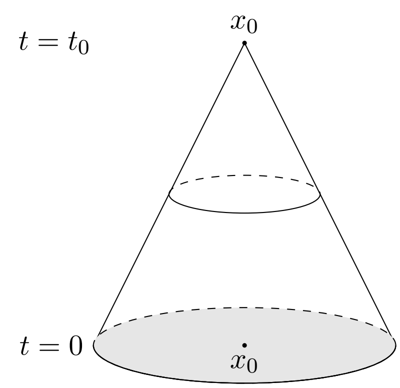


Figure 7.1. Cone of dependence: If u, u_t vanishes on the gray ball at $t = 0$, it will vanish on the entire cone

In fact we can even get a bit more:

Theorem 7.12. *Let $u \in C^2$ be a solution of the wave equation. If $u(0,x) = u_t(0,x) = 0$ for $x \in B_{t_0}(x_0)$, then $u(t,x) = 0$ for all (t,x) in the cone*

$$K(t_0, x_0) := \{(t,x) \in \mathbb{R}^{n+1} | 0 \le t \le t_0,\ |x - x_0| \le t_0 - t\}. \tag{7.29}$$

Proof. We define the local energy

$$E(t) := \frac{1}{2} \int_{B_{t_0 - t}(x_0)} \big(|\nabla u(t,x)|^2 + u_t(t,x)^2\big) d^n x.$$

Then one computes

$$\begin{aligned}
\dot{E}(t) &= \frac{1}{2}\frac{d}{dt} \int_0^{t_0 - t} \int_{\partial B_r(x_0)} \big(|\nabla u|^2 + u_t^2\big) dS\, dr \\
&= \int_{B_{t_0-t}(x_0)} \big((\nabla u)\cdot(\nabla u_t) + u_t u_{tt}\big) d^n x - \frac{1}{2}\int_{\partial B_{t_0-t}(x_0)} \big(|\nabla u|^2 + u_t^2\big) dS \\
&= \int_{B_{t_0-t}(x_0)} \big(u_{tt} - \Delta u\big) u_t\, d^n x + \int_{\partial B_{t_0-t}(x_0)} u_t \frac{\partial u}{\partial \nu} dS \\
&\qquad - \frac{1}{2}\int_{\partial B_{t_0-t}(x_0)} \big(|\nabla u|^2 + u_t^2\big) dS \\
&= \int_{\partial B_{t_0-t}(x_0)} \left(u_t \frac{\partial u}{\partial \nu} - \frac{1}{2} u_t^2 - \frac{1}{2}|\nabla u|^2 \right) dS,
\end{aligned}$$

where we have used integration by parts in the second step. Now

$$\left| u_t \frac{\partial u}{\partial \nu} \right| \le |u_t||\nabla u| \le \frac{1}{2} u_t^2 + \frac{1}{2}|\nabla u|^2$$

implies $\dot{E}(t) \leq 0$. Hence $E(t) \leq E(0) = 0$ and consequently ∇u and u_t vanish within K. Thus u is constant within u and since it vanishes for $t = 0$ the claim follows. □

Of course this result implies uniqueness for the wave equation on $\mathbb{R}^n$ as a special case:

Corollary 7.13. *There exists at most one solution $u \in C^2(\mathbb{R}^{n+1})$ of the wave equation for given data $g, h \in C(\mathbb{R}^n)$.*

Problem 7.15. *Show Lemma 7.10.*

Problem 7.16. *Find a corresponding energy for the Klein–Gordon equation and extend this section to the Klein–Gordon equation (with $m \geq 0$).*

Part 2

Advanced Partial Differential Equations

Chapter 8

A first look at weak derivatives and L^2 based Sobolev spaces

8.1. Motivation

Let us repeat the treatment of the **Poisson equation** in $\mathbb{R}^n$

$$-\Delta u = f \tag{8.1}$$

via the Fourier transform from Section 6.1. But now from a slightly different angle. Formally applying the Fourier transform we have found

$$|k|^2 \hat{u}(k) = \hat{f}(k). \tag{8.2}$$

When we solve this equation for $\hat{u}$,

$$\hat{u}(k) = \frac{\hat{f}(k)}{|k|^2}, \tag{8.3}$$

we notice two things: First of all $\hat{u}$ has better decay properties than $\hat{f}$ and by virtue of Lemma 6.2 this reflects that the solution u, roughly speaking, has two more derivatives than f. Secondly we see that $\hat{u}$ has a singularity at 0 (unless $\hat{f}$ vanishes sufficiently fast at 0) which reflects the fact, that the solution u has in general less decay than f. In fact, even if f has compact support, u will only decay like the fundamental solution unless $\hat{f}(0) = (2\pi)^{-n/2} \int_{\mathbb{R}^n} f(y) d^n y = 0$ (cf. Problem 5.15).

While these considerations give us some important insight, it is not straightforward to turn this intuition into precise mathematical statements.

We could assume that f is integrable and try to look for solutions u for which the function together with its first and second derivatives are integrable. However, since we don't know the image of this space of solutions under the Fourier transform (we don't even know the image of $L^1(\mathbb{R}^n)$, we had to invent a new name for it, the Wiener algebra), it is unclear in what sense (8.1) and (8.2) are equivalent.

This problem could be circumvented if we look for solutions in the Schwartz space $\mathcal{S}(\mathbb{R}^n)$, where the Fourier transform is bijective (Theorem 6.5). Indeed, this works fine and we see that for $f \in \mathcal{S}(\mathbb{R}^n)$ we have a solution $u \in \mathcal{S}(\mathbb{R}^n)$ if and only if $|k|^{-2}\hat{f}(k) \in \mathcal{S}(\mathbb{R}^n)$. Now one could argue that $\mathcal{S}(\mathbb{R}^n)$ is a rather restrictive class and wonder if there is a better one. Indeed, the Plancherel identity (Lemma 6.6) implies that the Fourier transform can be extended to a bijection on $L^2(\mathbb{R}^n)$ (details will follow in the next section). So if $u \in L^2(\mathbb{R}^n)$, then we also have $\hat{u} \in L^2(\mathbb{R}^n)$ and Lemma 6.2 tells us that under appropriate conditions $k_j\hat{u}(k) \in L^2(\mathbb{R}^n)$ will also imply that $\partial_j u \in L^2(\mathbb{R}^n)$ and similarly for higher derivatives. Now these extra *appropriate* conditions will make our approach again inconvenient and hence the trick is to simply drop them! More precisely, we introduce the space

$$H^r(\mathbb{R}^n) := \{u \in L^2(\mathbb{R}^n) | |k|^r \hat{u}(k) \in L^2(\mathbb{R}^n)\}, \qquad r \in \mathbb{N}_0, \tag{8.4}$$

and *define* partial derivatives of $u \in H^r(\mathbb{R}^n)$ via

$$\partial_\alpha u := ((\mathrm{i}k)^\alpha \hat{u}(k))^\vee, \qquad |\alpha| \le r. \tag{8.5}$$

By Lemma 6.2 this will agree with the usual definition under moderate conditions and hence can be viewed as an extension of the usual definition of differentiability. We call (8.5) **weak derivatives**. Please observe that $|k|^\alpha \le |k|^{|\alpha|} \le 1 + |k|^r$ for all $|\alpha| \le r$ which implies that existence of all diagonal derivatives (where all derivatives are taken with respect to the same coordinate direction) implies existence of all off-diagonal as well as all lower order derivatives.

Then, for $f \in L^2(\mathbb{R}^n)$ we have a solution of (8.2) if and only if $|k|^{-2}\hat{f}(k)$ is locally square integrable near the origin. By construction this will also give us the solution of (8.1) provided the derivatives are understood as weak derivatives. Moreover, Lemma 6.2 could be used to work out sufficient conditions when this *weak* solution is a classical solution, that is, when this solution is $C^2(\mathbb{R}^n)$.

Finally, note that if, in addition to the above solvability condition, we have $f \in H^r(\mathbb{R}^n)$, then $u \in H^{r+2}(\mathbb{R}^n)$ and in this sense the solution indeed gains precisely two derivatives.

In summary, the introduction of the weak derivative in (8.5) allowed us to give a transparent discussion of solvability of the Poisson equation. It should be evident that this also works for other constant coefficient partial differential equations on $\mathbb{R}^n$.

Problem 8.1. *Discuss the Helmholz equation* (6.23) *on* $\mathbb{R}^n$.

8.2. The Fourier transform on L^2

The Plancherel identity $\|f\|_2 = \|\hat{f}\|_2$ (Lemma 6.6) allows us to extend the Fourier transform to $L^2(\mathbb{R}^n)$ as follows: First of all, this identity holds on $\mathcal{S}(\mathbb{R}^n)$ which is dense. Now choosing an arbitrary sequence $f_m \in \mathcal{S}(\mathbb{R}^n)$ converging to f in the L^2 norm, we can set

$$\mathcal{F}(f) := \lim_{m\to\infty} \mathcal{F}(f_m). \tag{8.6}$$

The Plancherel identity ensures that this limit exists and is independent of the sequence chosen (you can verify this directly or resort to the extension principle from functional analysis: every densely defined bounded operator can be extendend to the entire space; Theorem 1.16 from [**35**]).

Theorem 8.1 (Plancherel)**.** *The Fourier transform $\mathcal{F}$ extends to a unitary operator $\mathcal{F} : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$.*

Proof. As already noted, $\mathcal{F}$ extends uniquely to a bounded operator on $L^2(\mathbb{R}^n)$. Since Plancherel's identity remains valid by continuity of the norm and since its range is dense, this extension is a unitary operator. □

We also note that this extension is still given by (6.1) whenever the right-hand side is integrable.

Lemma 8.2. *Let $f \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$, then* (6.1) *continues to hold, where $\mathcal{F}$ now denotes the extension of the Fourier transform from $\mathcal{S}(\mathbb{R}^n)$ to $L^2(\mathbb{R}^n)$. Similarly, for $f \in L^2(\mathbb{R}^n)$ with $\hat{f} \in L^1(\mathbb{R}^n)$ Theorem 6.5 continues to hold.*

Proof. If f has compact support, then (by Lemma B.13 and Lemma B.14) its mollification $\phi_\varepsilon * f \in C_c^\infty(\mathbb{R}^n)$ converges to f both in L^1 and L^2. Hence the claim holds for every f with compact support. Finally, for general $f \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ consider $f_m := f\chi_{B_m(0)}$. Then $f_m \to f$ in both $L^1(\mathbb{R}^n)$ and $L^2(\mathbb{R}^n)$ and the claim follows.

For the second claim note that by Theorem 6.5 $(\mathcal{F}^{-1}f)(x) = (\mathcal{F}f)(-x)$ at least for $f \in \mathcal{S}(\mathbb{R}^n)$ which remains true when taking limits. □

In particular,

$$\hat{f}(k) = \lim_{m\to\infty} \frac{1}{(2\pi)^{n/2}} \int_{|x|\le m} \mathrm{e}^{-\mathrm{i}k\cdot x} f(x) d^n x, \tag{8.7}$$

where the limit has to be understood in $L^2(\mathbb{R}^n)$ and can be omitted if $f \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$. Similarly for $\check{f}$.

For later use we also need the fact that the Fourier transform of an integrable function will vanish at ∞. To this end we denote the space of all continuous functions $f : \mathbb{R}^n \to \mathbb{C}$ which vanish at ∞ by $C_0(\mathbb{R}^n)$.

Lemma 8.3 (Riemann-Lebesgue). *The Fourier transform maps $L^1(\mathbb{R}^n)$ into $C_0(\mathbb{R}^n)$.*

Proof. First of all recall that $C_0(\mathbb{R}^n)$ equipped with the sup norm is a Banach[1] space and that $\mathcal{S}(\mathbb{R}^n)$ is dense (Problem B.4). By the previous lemma we have $\hat{f} \in C_0(\mathbb{R}^n)$ if $f \in \mathcal{S}(\mathbb{R}^n)$. Moreover, since $\mathcal{S}(\mathbb{R}^n)$ is dense in $L^1(\mathbb{R}^n)$, the estimate (6.2) shows that the Fourier transform extends to a continuous map from $L^1(\mathbb{R}^n)$ into $C_0(\mathbb{R}^n)$. □

Concerning convolutions we can easily extend Corollary 6.8.

Corollary 8.4. *The convolution of two $L^2(\mathbb{R}^n)$ functions is in $\mathrm{Ran}(\mathcal{F}) \subset C_0(\mathbb{R}^n)$ and we have $\|f * g\|_\infty \le \|f\|_2 \|g\|_2$ as well as*

$$(fg)^\wedge = (2\pi)^{-n/2} \hat{f} * \hat{g}, \qquad (f * g)^\wedge = (2\pi)^{n/2} \hat{f}\hat{g} \tag{8.8}$$

in this case.

Proof. The inequality $\|f * g\|_\infty \le \|f\|_2 \|g\|_2$ is immediate from Cauchy–Schwarz and shows that the convolution is a continuous bilinear form from $L^2(\mathbb{R}^n)$ to $L^\infty(\mathbb{R}^n)$. Now take sequences $f_m, g_m \in \mathcal{S}(\mathbb{R}^n)$ converging to $f, g \in L^2(\mathbb{R}^n)$. Then using Corollary 6.8 together with continuity of the Fourier transform from $L^1(\mathbb{R}^n)$ to $C_0(\mathbb{R}^n)$ and on $L^2(\mathbb{R}^n)$ we obtain

$$(fg)^\wedge = \lim_{m\to\infty} (f_m g_m)^\wedge = (2\pi)^{-n/2} \lim_{m\to\infty} \hat{f}_m * \hat{g}_m = (2\pi)^{-n/2} \hat{f} * \hat{g}.$$

Similarly,

$$(f * g)^\wedge = \lim_{m\to\infty} (f_m * g_m)^\wedge = (2\pi)^{n/2} \lim_{m\to\infty} \hat{f}_m \hat{g}_m = (2\pi)^{n/2} \hat{f}\hat{g}$$

from which that last claim follows since $\mathcal{F} : \mathrm{Ran}(\mathcal{F}) \to L^1(\mathbb{R}^n)$ is closed as it is the inverse of a bounded operator (Lemma B.19). □

Problem 8.2. *Show that* (8.6) *is well defined (i.e., the limit exists and is independent of the sequence). Show that the Plancherel identity continues to hold.*

Problem 8.3. *Show*

$$\int_0^\infty \frac{\sin(x)^2}{x^2} dx = \frac{\pi}{2}.$$

(Hint: Problem 4.1 (i).)

[1] Stefan Banach (1892–1945), Polish mathematician

8.3. The Sobolev spaces $H^r(\mathbb{R}^n)$

Following our previous considerations we introduce the **Sobolev space**[2]

$$H^r(\mathbb{R}^n) := \{f \in L^2(\mathbb{R}^n) | |k|^r \hat{f}(k) \in L^2(\mathbb{R}^n)\}. \tag{8.9}$$

The most important case is when r is an integer, however our definition makes sense for any $r \geq 0$. Moreover, note that $H^r(\mathbb{R}^n)$ becomes a Hilbert space if we introduce the scalar product

$$\langle f, g \rangle_{H^r} := \int_{\mathbb{R}^n} \hat{f}(k)^* \hat{g}(k) (1 + |k|^2)^r d^n k. \tag{8.10}$$

In particular, note that by construction $\mathcal{F}$ maps $H^r(\mathbb{R}^n)$ unitarily onto $L^2(\mathbb{R}^n, \langle k \rangle^r d^n k)$, where (sometimes known as **Japanese bracket**)

$$\langle k \rangle := (1 + |k|^2)^{1/2}, \qquad k \in \mathbb{C}^n. \tag{8.11}$$

Clearly $H^{r+1}(\mathbb{R}^n) \subset H^r(\mathbb{R}^n)$ with the embedding being continuous. Moreover, $\mathcal{S}(\mathbb{R}^n) \subset H^r(\mathbb{R}^n)$ and this subset is dense (since $\mathcal{S}(\mathbb{R}^n)$ is dense in $L^2(\mathbb{R}^n, \langle k \rangle^r d^n k)$).

As already mentioned, the motivation for the definition (8.9) stems from Lemma 6.3 which allows us to extend differentiation to a larger class. In fact, every function in $H^r(\mathbb{R}^n)$ has partial derivatives up to order $\lfloor r \rfloor$, which are defined via

$$\partial_\alpha f := ((\mathrm{i}k)^\alpha \hat{f}(k))^\vee, \qquad f \in H^r(\mathbb{R}^n),\ |\alpha| \leq r. \tag{8.12}$$

By Lemma 6.3 this definition coincides with the usual one for every $f \in \mathcal{S}(\mathbb{R}^n)$.

Example 8.1. Consider $f(x) := (1-|x|)\chi_{[-1,1]}(x)$. Then $\hat{f}(k) = \sqrt{\frac{2}{\pi}} \frac{1-\cos(k)}{k^2}$ and $f \in H^1(\mathbb{R})$. The weak derivative is $f'(x) = -\operatorname{sign}(x)\chi_{[-1,1]}(x)$. ⋄

Of course a natural question to ask is when the weak derivatives are in fact classical derivatives. To this end observe that the Riemann–Lebesgue lemma implies that $\partial_\alpha f(x) \in C_0(\mathbb{R}^n)$ provided $k^\alpha \hat{f}(k) \in L^1(\mathbb{R}^n)$. Moreover, in this situation the derivatives will exist as classical derivatives:

Lemma 8.5. *Suppose $f \in L^1(\mathbb{R}^n)$ or $f \in L^2(\mathbb{R}^n)$ with $(1 + |k|^r)\hat{f}(k) \in L^1(\mathbb{R}^n)$ for some $r \in \mathbb{N}_0$. Then $f \in C_0^r(\mathbb{R}^n)$, the set of functions with continuous partial derivatives of order r all of which vanish at ∞. Moreover, the classical and weak derivatives coincide in this case.*

Proof. We begin by observing that by Lemma 8.2

$$f(x) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \mathrm{e}^{\mathrm{i}k\cdot x} \hat{f}(k) d^n k.$$

[2] Sergei Sobolev (1908–1989), Soviet mathematician

Now the claim follows as in the proof of Lemma 6.3 by differentiating the integral using Lemma A.7. □

Now we are able to prove the following embedding theorem.

Theorem 8.6 (Sobolev embedding). *Suppose $r > s + \frac{n}{2}$ for some $s \in \mathbb{N}_0$. Then $H^r(\mathbb{R}^n)$ is continuously embedded into $C_0^s(\mathbb{R}^n)$ with*

$$\|\partial_\alpha f\|_\infty \le C_{n,r}\|f\|_{H^r}, \qquad |\alpha| \le s. \tag{8.13}$$

Proof. Use $|(\mathrm{i}k)^\alpha \hat{f}(k)| \le \langle k\rangle^{|\alpha|}|\hat{f}(k)| = \langle k\rangle^{-\sigma} \cdot \langle k\rangle^{|\alpha|+\sigma}|\hat{f}(k)|$. Now $\langle k\rangle^{-\sigma} \in L^2(\mathbb{R}^n)$ if $\sigma > \frac{n}{2}$ (see Example A.8) and $\langle k\rangle^{|\alpha|+\sigma}|\hat{f}(k)| \in L^2(\mathbb{R}^n)$ if $\sigma + |\alpha| \le r$. Hence $\langle k\rangle^{|\alpha|}|\hat{f}(k)| \in L^1(\mathbb{R}^n)$ and the claim follows from the previous lemma. □

In fact, we can even do a bit better.

Lemma 8.7 (Morrey[3] inequality). *Suppose $f \in H^{n/2+\gamma}(\mathbb{R}^n)$ for some $\gamma \in (0,1)$. Then $f \in C_0^{0,\gamma}(\mathbb{R}^n)$, the set of functions which are Hölder continuous with exponent γ and vanish at ∞. Moreover,*

$$|f(x) - f(y)| \le C_{n,\gamma}\|\hat{f}\|_{H^{n/2+\gamma}}|x-y|^\gamma \tag{8.14}$$

in this case.

Proof. We begin with

$$f(x+y) - f(x) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \mathrm{e}^{\mathrm{i}k\cdot x}(\mathrm{e}^{\mathrm{i}k\cdot y} - 1)\hat{f}(k) d^n k$$

implying

$$|f(x+y) - f(x)| \le \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \frac{|\mathrm{e}^{\mathrm{i}k\cdot y} - 1|}{\langle k\rangle^{n/2+\gamma}} \langle k\rangle^{n/2+\gamma}|\hat{f}(k)| d^n k.$$

Hence, after applying Cauchy–Schwarz, it remains to estimate (recall (A.29))

$$\begin{aligned}\int_{\mathbb{R}^n} \frac{|\mathrm{e}^{\mathrm{i}k\cdot y} - 1|^2}{\langle k\rangle^{n+2\gamma}} d^n k &\le S_n \int_0^{1/|y|} \frac{(|y|r)^2}{r^{n+2\gamma}} r^{n-1} dr + S_n \int_{1/|y|}^\infty \frac{4}{r^{n+2\gamma}} r^{n-1} dr \\ &= \frac{S_n}{2(1-\gamma)}|y|^{2\gamma} + \frac{2S_n}{\gamma}|y|^{2\gamma} = \frac{S_n(4-3\gamma)}{2\gamma(1-\gamma)}|y|^{2\gamma},\end{aligned}$$

where $S_n = nV_n$ is the surface area of the unit sphere in $\mathbb{R}^n$. □

Using this lemma we immediately obtain:

Corollary 8.8. *Suppose $r \ge s + \gamma + \frac{n}{2}$ for some $s \in \mathbb{N}_0$ and $\gamma \in (0,1)$. Then $H^r(\mathbb{R}^n)$ is continuously embedded into $C_0^{s,\gamma}(\mathbb{R}^n)$, the set of functions in $C_0^s(\mathbb{R}^n)$ whose highest derivatives are Hölder continuous of exponent γ.*

---

[3] Charles B. Morrey (1907–1984), American mathematician

In the case $r > \frac{n}{2}$, when Morrey's inequality holds, we even get that $H^r(\mathbb{R}^n)$ is a Banach algebra:

Lemma 8.9. *Suppose $f, g \in H^r(\mathbb{R}^n)$ and $\hat{f}, \hat{g} \in L^1(\mathbb{R}^n)$ for some $r \geq 0$. Then we have $fg \in H^r(\mathbb{R}^n)$ with*

$$\|fg\|_{H^r} \leq C_{n,r}\big(\|f\|_{H^r}\|\hat{g}\|_1 + \|\hat{f}\|_1\|g\|_{H^r}\big). \tag{8.15}$$

Moreover, if $r > \frac{n}{2}$, then $\|\hat{f}\|_1 \leq c_{n,r}\|f\|_{H^r}$ and we get

$$\|fg\|_{H^r} \leq \tilde{C}_{n,r}\|f\|_{H^r}\|g\|_{H^r}. \tag{8.16}$$

Proof. Note that we have

$$\begin{aligned}\langle k\rangle^r &\leq (1 + 2|k-l|^2 + 2|l|^2)^{r/2} \leq 2^{r/2}(1 + |k-l|^2 + 1 + |l|^2)^{r/2}\\ &\leq c_r\big(\langle k-l\rangle^r + \langle l\rangle^r\big)\end{aligned}$$

for $c_r = \max(2^{r-1}, 2^{r/2})$ and $r \geq 0$. Here we have used $|k|^2 \leq (|k-l|+|l|)^2 \leq 2|k-l|^2 + 2|l|^2$ in the first step and that $x \mapsto x^{r/2}$ is subadditive for $r \leq 2$ and convex for $r \geq 2$ in the last step. Hence by Corollary 8.4 we have

$$\begin{aligned}(2\pi)^{n/2}\langle k\rangle^r|(fg)^\wedge(k)| = \langle k\rangle^r|(\hat{f} * \hat{g})(k)| &\leq \int_{\mathbb{R}^n} \langle k\rangle^r|\hat{f}(k-l)||\hat{g}(l)|d^n l\\ &\leq c_r\big((|\langle.\rangle^r\hat{f}| * |\hat{g}|)(k) + (|\hat{f}| * |\langle.\rangle^r\hat{g}|)(k)\big)\end{aligned}$$

and the first claim with $C_{n,r} = c_r(2\pi)^{-n/2}$ follows from Young's inequality (B.25). The second claim follows from Cauchy–Schwarz $\|\hat{f}\|_1 \leq \|\langle.\rangle^{-r}\|_2\|f\|_{H^r}$ as in the proof of Theorem 8.6. □

To end this section I remark that our definition of a weak derivative is not the usual one. The standard definition says that a function $h \in L^1_{loc}(\mathbb{R}^n)$ satisfying

$$\int_{\mathbb{R}^n} \varphi(x)h(x)d^n x = (-1)^{|\alpha|}\int_{\mathbb{R}^n} (\partial_\alpha\varphi)(x)f(x)d^n x, \quad \forall \varphi \in C_c^\infty(\mathbb{R}^n), \tag{8.17}$$

is called the **weak derivative** or the derivative in the sense of distributions of f (by Lemma B.16 such a function is unique if it exists). This definition is more flexible (it is not tied to $\mathbb{R}^n$, since $\mathbb{R}^n$ can be easily replaced by an arbitrary open set) and will be pursued in the next chapter.

To establish equivalence we first note that

$$\begin{aligned}\int_{\mathbb{R}^n} g(x)(\partial_\alpha f)(x)d^n x &= \langle g^*, (\partial_\alpha f)\rangle = \langle \hat{g}(k)^*, (\mathrm{i}k)^\alpha\hat{f}(k)\rangle\\ &= (-1)^{|\alpha|}\langle(\mathrm{i}k)^\alpha\hat{g}(k)^*, \hat{f}(k)\rangle = (-1)^{|\alpha|}\langle\partial_\alpha g^*, f\rangle\\ &= (-1)^{|\alpha|}\int_{\mathbb{R}^n}(\partial_\alpha g)(x)f(x)d^n x,\end{aligned} \tag{8.18}$$

for $f, g \in H^r(\mathbb{R}^n)$. Hence, choosing $g = \varphi$ in (8.18), we see that functions in $H^r(\mathbb{R}^n)$ have weak derivatives (in the sense of (8.17)) up to order r, which are in $L^2(\mathbb{R}^n)$. Moreover, the weak derivatives coincide with the derivatives defined in (8.12). Conversely, given (8.17) with $f, h \in L^2(\mathbb{R}^n)$ we can use that $\mathcal{F}$ is unitary to conclude $\int_{\mathbb{R}^n} \hat{\varphi}(k)\hat{h}(k) d^n k = \int_{\mathbb{R}^n} (\mathrm{i}k)^\alpha \hat{\varphi}(k)\hat{f}(k) d^n k$ for all $\varphi \in C_c^\infty(\mathbb{R}^n)$. By approximation this follows for $\varphi \in H^r(\mathbb{R}^n)$ with $r \ge |\alpha|$ and hence in particular for $\hat{\varphi} \in C_c^\infty(\mathbb{R}^n)$. Consequently $(\mathrm{i}k)^\alpha \hat{f}(k) = \hat{h}(k)$ a.e. implying that $f \in H^r(\mathbb{R}^n)$ if all weak derivatives exist up to order r and are in $L^2(\mathbb{R}^n)$.

In this connection the following norm for $H^m(\mathbb{R}^n)$ with $m \in \mathbb{N}_0$ is more common:

$$\|f\|_{m,2}^2 := \sum_{|\alpha| \le m} \|\partial_\alpha f\|_2^2. \tag{8.19}$$

By $|k^\alpha| \le |k|^{|\alpha|} \le (1 + |k|^2)^{m/2}$ it follows that this norm is equivalent to (8.10).

Problem 8.4. *Provide the details for Example 8.1.*

Problem 8.5. *Suppose $f \in L^2(\mathbb{R}^n)$ show that $\varepsilon^{-1}(f(x+e_j\varepsilon) - f(x)) \to g_j(x)$ in L^2 if and only if $k_j \hat{f}(k) \in L^2$, where e_j is the unit vector into the j'th coordinate direction. Moreover, show $g_j = \partial_j f$ if $f \in H^1(\mathbb{R}^n)$.*

Problem 8.6. *Show that if $u \in H^2(\mathbb{R}^n)$ solves (8.1) for some $f \in H^r(\mathbb{R}^n)$, then $u \in H^{r+2}(\mathbb{R}^n)$ with $\|u\|_{H^{r+2}} \le C(\|f\|_{H^r} + \|u\|_{L^2})$. Is an estimate of the form $\|u\|_{H^{r+2}} \le C\|f\|_{H^r}$ possible?*

8.4. Evolution problems

Now let us look at the case of equations evolving time, like for example the heat equation

$$u_t = \Delta u, \qquad u(0, x) = g(x), \tag{8.20}$$

on $\mathbb{R}^n$. Taking the Fourier transform we obtained

$$\hat{u}_t(t, k) = -|k|^2 \hat{u}(t, k), \qquad \hat{u}(0, k) = \hat{g}(k), \tag{8.21}$$

and solving this differential equation showed

$$\hat{u}(t, k) = \hat{g}(k) \mathrm{e}^{-t|k|^2}. \tag{8.22}$$

Now the proper way of looking at this problem is to regard $u(t, x)$ as a function $u(t)$ with values in $L^2(\mathbb{R}^n)$. From this point of view, the heat equation now just looks like an ordinary constant coefficient linear differential equation $u_t = Au$ with the only difference being that the coefficient matrix A is now a linear operator acting on the Hilbert space $L^2(\mathbb{R}^n)$. The solution $\hat{u}(t)$ (for fixed initial condition g) is a continuous map from $t \ge 0$ to $L^2(\mathbb{R}^n)$. And this is already the first big difference: while for an ordinary differential equation there is no distinction between positive and negative times, now there is! The exponential function $\mathrm{e}^{-t|k|^2}$ is bounded for $t \geq 0$ and unbounded for $t < 0$. Hence $\hat{u}(t)$ will in general be no longer in $L^2(\mathbb{R}^n)$ for $t < 0$. The second difference arises if we ask about differentiability of $\hat{u}(t)$. Then defining the time derivative as a difference quotient

$$\frac{d}{dt}\hat{u}(t) := \lim_{\varepsilon \to 0} \frac{\hat{u}(t+\varepsilon) - \hat{u}(t)}{\varepsilon}, \tag{8.23}$$

the dominated convergence theorem (Problem 8.7) shows that this limit exists in $L^2(\mathbb{R}^n)$ for $t > 0$. Moreover, note that $\hat{u}(t) \in H^r(\mathbb{R}^n)$ for any $r > 0$ for $t > 0$.

These ideas form the starting point of semigroup theory and it will be discussed in detail in subsequent chapters.

Problem 8.7. *Assume $g \in L^2(\mathbb{R}^n)$. Show that $\hat{u}(t)$ defined in (8.22) is differentiable and solves $\frac{d}{dt}\hat{u}(t)(k) = -|k|^2\hat{u}(t)(k)$ for $t > 0$. (Hint: $|\mathrm{e}^{-\varepsilon|k|^2} - 1| \leq \varepsilon|k|^2$ for $\varepsilon \geq 0$.)*

Chapter 9

General Sobolev spaces

The modern theory of partial differential equations uses powerful results from functional analysis. The classical Banach spaces of continuously differentiable functions are frequently not suitable for this purpose. The weapon of choice in this context are Sobolev spaces. As a preparation we will hence study them in this chapter.

9.1. Basic properties

Throughout this chapter $U \subseteq \mathbb{R}^n$ will be nonempty and open and we will use the notation $V \subset\subset U$ if V is a relatively compact set with $\overline{V} \subset U$.

Our aim is to extend the Lebesgue spaces to include derivatives. To this end, for a locally integrable function $f \in L^1_{loc}(U)$, a locally integrable function $h \in L^1_{loc}(U)$ satisfying

$$\int_U \varphi(x)h(x)d^n x = (-1)^{|\alpha|}\int_U (\partial_\alpha \varphi)(x)f(x)d^n x, \quad \forall \varphi \in C_c^\infty(U), \qquad (9.1)$$

is called the **weak derivative** or the derivative in the sense of distributions of f. Note that by Lemma B.16 such a function is unique if it exists. Moreover, if $f \in C^k(U)$, then integration by parts (A.44) shows that the weak derivative coincides with the usual derivative. Also note that the order in which the partial derivatives are taken is irrelevant for φ (by the classical theorem of Schwarz) and hence the same is true for weak derivatives. This is no contradiction to the classical counterexamples for the theorem of Schwarz since weak derivatives are only defined up to equivalence a.e.

Example 9.1. Consider $f(x) := x^2 \sin(\frac{\pi}{x^2})$ on $U := (-1, 1)$ (here $f(0) := 0$). Then it is straightforward to verify that

$$f'(x) = \begin{cases} 2x\sin(\frac{\pi}{x^2}) - \frac{2\pi}{x}\cos(\frac{\pi}{x^2}), & x \neq 0, \\ 0, & x = 0, \end{cases}$$

that is, f is everywhere differentiable on U. Of course f is weakly differentiable on $(-1, 0)$ as well as on $(0, 1)$ implying that the weak derivative of f on U must equal f' (consider test functions φ supported away from 0). However, one can check that f' is not integrable (Problem 9.2) and hence f is not weakly differentiable. ⋄

Example 9.2. Consider $U := \mathbb{R}$. If $f(x) := |x|$, then $\partial f(x) = \operatorname{sign}(x)$ as a weak derivative. If we try to take a second derivative we are lead to

$$\int_{\mathbb{R}} \varphi(x)h(x)dx = -\int_{\mathbb{R}} \varphi'(x)\operatorname{sign}(x)dx = 2\varphi(0)$$

and it is easy to see that no locally integrable function can satisfy this requirement. ⋄

Example 9.3. In fact, in one dimension the class of weakly differentiable functions can be identified with the class of antiderivatives of integrable functions, that is, the class of **absolutely continuous** functions

$$AC[a, b] := \{f(x) = f(a) + \int_a^x h(y)dy | h \in L^1(a, b)\},$$

where $a < b$ are some real numbers. It is easy to see that every absolutely continuous function is in particular continuous, $AC[a, b] \subset C[a, b]$. Moreover, using Lebesgue's differentiation theorem one can show that an absolutely continuous function is differentiable a.e. with $f'(x) = h(x)$ (and hence h is uniquely defined a.e.). However, not every continuous function is absolutely continuous. The most striking counter example being the Cantor function, which is a continuous monotone function whose derivative exists and vanishes a.e. Hence the Cantor function cannot be reconstructed by integrating its derivative, that is, the fundamental theorem of calculus fails for the Cantor function. We refer to Section 4.4 from [**34**] for further details.

If $f, g \in AC[a, b]$, we have the integration by parts formula (Problem 9.4), which shows that every absolutely continuous function has a weak derivative which equals the a.e. derivative. In fact, with a little more effort one can show (Problem 9.5) that the converse is also true, that is, $W^{1,1}(a, b) = AC[a, b]$ and $W^{1,p}(a, b) = \{f \in AC[a, b] | f' \in L^p(a, b)\}$. Consequently $W^{k,1}(a, b) = \{f \in C^{k-1}[a, b] | f^{(k-1)} \in AC[a, b]\}$ and $W^{k,p}(a, b) = \{f \in C^{k-1}[a, b] | f^{(k-1)} \in AC[a, b],\ f^{(k)} \in L^p(a, b)\}$. ⋄

Example 9.4. One can verify (Problem 9.7) that $f(x) := |x|^{-\gamma}$ is weakly differentiable for $\gamma < \frac{n-p}{p}$. Hence in higher dimensions weakly differentiable functions might not be continuous. ⋄

Now we can define the **Sobolev space** $W^{k,p}(U)$ as the set of all functions in $L^p(U)$ which have weak derivatives up to order k in $L^p(U)$. Clearly $W^{k,p}(U)$ is a linear space since $f, g \in W^{k,p}(U)$ implies $af + bg \in W^{k,p}(U)$ for $a, b \in \mathbb{C}$ and $\partial_\alpha(af + bg) = a\partial_\alpha f + b\partial_\alpha g$ for all $|\alpha| \le k$. Moreover, for $f \in W^{k,p}(U)$ we define its norm

$$\|f\|_{k,p} := \begin{cases} \left(\sum_{|\alpha|\le k} \|\partial_\alpha f\|_p^p\right)^{1/p}, & 1 \le p < \infty, \\ \max_{|\alpha|\le k} \|\partial_\alpha f\|_\infty, & p = \infty. \end{cases} \tag{9.2}$$

We will also use the gradient $\nabla u = (\partial_1 u, \dots, \partial_n u)$ and by $\|\nabla u\|_p$ we will always mean $\||\nabla u|\|_p$, where $|\nabla u|$ denotes the Euclidean norm.

It is easy to check that with this definition $W^{k,p}$ becomes a normed linear space. Of course for $p = 2$ we have a corresponding scalar product

$$\langle f, g\rangle_{W^{k,2}} := \sum_{|\alpha|\le k} \langle \partial_\alpha f, \partial_\alpha g\rangle_{L^2} \tag{9.3}$$

and one reserves the special notation $H^k(U) := W^{k,2}(U)$ for this case. Similarly we define local versions of these spaces $W^{k,p}_{loc}(U)$ as the set of all functions in $L^p_{loc}(U)$ which have weak derivatives up to order k in $L^p_{loc}(U)$.

Theorem 9.1. *For each $k \in \mathbb{N}_0$, $1 \le p \le \infty$ the Sobolev space $W^{k,p}(U)$ is complete, that is, it is a Banach space. It is separable for $1 \le p < \infty$ as well as reflexive and uniformly convex for $1 < p < \infty$.*

Proof. Let f_m be a Cauchy sequence in $W^{k,p}$. Then $\partial_\alpha f_m$ is a Cauchy sequence in L^p for every $|\alpha| \le k$. Consequently $\partial_\alpha f_m \to f_\alpha$ in L^p. Moreover, letting $m \to \infty$ in

$$\begin{aligned} \int_U \varphi f_\alpha d^n x &= \lim_{m\to\infty} \int_U \varphi(\partial_\alpha f_m) d^n x = \lim_{m\to\infty} (-1)^{|\alpha|} \int_U (\partial_\alpha\varphi) f_m d^n x \\ &= (-1)^{|\alpha|} \int_U (\partial_\alpha \varphi) f_0 d^n x, \qquad \varphi \in C_c^\infty(U), \end{aligned}$$

shows $f_0 \in W^{k,p}$ with $\partial_\alpha f_0 = f_\alpha$ for $|\alpha| \le k$. By construction $f_m \to f_0$ in $W^{k,p}$ which implies that $W^{k,p}$ is complete.

Concerning the last claim note that $W^{k,p}(U)$ can be regarded as a subspace of $\bigoplus_{p,|\alpha|\le k} L^p(U)$ which has the claimed properties by Lemma 3.14 from **[34]** and Theorem 3.11 from **[34]** (see also the remark after the proof), Corollary 6.2 from **[34]** (cf. also Problem 4.30 from **[35]**). □

As a consequence of the proof we record:

Corollary 9.2. *Fix some multi-index α. Let $f_n \in L^p(U)$ be a sequence such that $\partial_\alpha f_n \in L^p(U)$ exists. Then $f_n \to f$, $\partial_\alpha f_n \to g$ in $L^p(U)$ implies that $\partial_\alpha f = g$ exists.*

Of course we have the natural embedding $W^{k,p}(U) \hookrightarrow L^p(U)$ and if $V \subseteq U$ is nonempty and open, then $f \in W^{k,p}(U)$ implies $f|_V \in W^{k,p}(V)$ (since $C_c^\infty(V) \subseteq C_c^\infty(U)$). Sometimes it is also useful to look at functions with values in $\mathbb{C}^n$ in which case we define $W^{k,p}(U, \mathbb{C}^n)$ as the corresponding direct sum.

Regarding $W^{k,p}(U)$ as a subspace of $\bigoplus_{p,|\alpha|\le k} L^p(U)$ also provides information on its dual space. Indeed, if $M \subseteq X$ is a closed subset of a Banach space X, then every linear functional on X gives rise to a linear functional on M by restricting its domain. Clearly two functionals give rise to the same restriction if their difference vanishes on M. Conversely, any functional on M can be extended to a functional on X by Hahn–Banach and thus any element from M^* arises in this way. Hence we have $M^* = X^*/M^\perp$, where $M^\perp$ denotes the annihilator of a subspace, that is, the set of all linear functionals which vanish on the subspace (cf. also Theorem 4.21 from **[35]**). Concerning the norm of a functional, note that when extending a functional from M to X, the norm can only increase. Moreover, the extension obtained from Hahn–Banach preserves the norm and hence the norm is given by taking the minimum over all extensions. In a strictly convex space the functional where the norm is attained is even unique (Problem 6.39 from **[35]**). Applied to $W^{k,p}(U)$ this gives

$$W^{k,p}(U)^* = \left(\bigoplus_{q,|\alpha|\le k} L^p(U)^*\right)/W^{k,p}(U)^\perp, \qquad \frac{1}{p}+\frac{1}{q}=1, \tag{9.4}$$

and for $1 \le p < \infty$, such that $L^p(U)^* = L^q(U)$, we even get:

Lemma 9.3. *For $1 \le p < \infty$, every linear functional $\ell \in W^{k,p}(U)^*$ can be represented as*

$$\ell(f) = \sum_{|\alpha|\le k} \int_U g_\alpha(x)(\partial_\alpha f)(x) d^n x, \tag{9.5}$$

with some functions $g_\alpha \in L^q(U)$, $|\alpha| \le k$. Moreover,

$$\|\ell\| = \min\{\|g\|_{\bigoplus_{q,|\alpha|\le k} L^q(U)} | g_\alpha \text{ as in } (9.5)\}, \tag{9.6}$$

where the minimizer is unique if $1 < p < \infty$.

In the case $p = 2$ the Riesz[1] representation theorem (for the dual of a Hilbert space; Theorem 2.10 from **[35]**) tells us that the unique minimizer is given by some $g \in H^k(U)$ such that $g_\alpha = \partial_\alpha g$. As another consequence note

[1]Frigyes Riesz (1880–1956), Hungarian mathematician

that a sequence converges weakly in $W^{k,p}(U)$ if and only if all derivatives converge weakly in $L^p(U)$.

Example 9.5. Consider $W^{1,p}(0,1)$, then functions in this space are absolutely continuous and one can consider the linear functional

$$\ell_{x_0}(f) := f(x_0)$$

for given $x_0 \in [0,1]$. Defining

$$g_{x_0}(x) := \frac{1}{\sinh(1)} \begin{cases} \cosh(1-x_0)\cosh(x), & x \le x_0, \\ \cosh(1-x)\cosh(x_0), & x \ge x_0, \end{cases}$$

one verifies

$$\ell_{x_0}(f) = \int_0^1 g_{x_0}(x) f(x) dx + \int_0^1 g_{x_0}'(x) f'(x) dx.$$

This representation is however not unique! To this end, note that for any $h \in W_0^{1,q}[0,1]$

$$\ell(f) = \int_0^1 h'(x) f(x) dx + \int_0^1 h(x) f'(x) dx = h(1)f(1) - h(0)f(0) = 0$$

represents the zero functional. In fact, any representation of the zero functional is of this form (show this).

Moreover, note that

$$|f(x_0)| \le \|g_{x_0}\|_{W^{1,q}} \|f\|_{W^{1,p}}.$$

Since we have $\|g_{x_0}\|_\infty = g_{x_0}(x_0)$, which attains its maximum at the boundary points, we infer

$$\|f\|_\infty \le \coth(1) \|f\|_{W^{1,p}}.$$

In particular, we have a continuous embedding $W^{1,p}(0,1) \hookrightarrow C[0,1]$. Moreover, for $1 < p \le \infty$, there is even a continuous embedding into the space of Hölder continuous functions $W^{1,p}(0,1) \hookrightarrow C^{0,\gamma}[0,1]$ with exponent $\gamma := 1 - \frac{1}{p}$ (Problem 9.6). ◇

Since functions from $W^{k,p}$ might not even be continuous, it will be convenient to know that they still can be well approximated by nice functions. To this end we next show that smooth functions are dense in $W^{k,p}$. A first naive approach would be to extend $f \in W^{k,p}(U)$ to all of $\mathbb{R}^n$ by setting it 0 outside U and consider $f_\varepsilon := \phi_\varepsilon * f$, where ϕ is the standard Friedrichs[2] mollifier. The problem with this approach is that we generically create a non-differentiable singularity at the boundary and hence this only works as long as we stay away from the boundary.

[2] Kurt Otto Friedrichs (1901–1982), German American mathematician

Lemma 9.4 (Friedrichs). *Let $f \in W^{k,p}(U)$ and set $f_\varepsilon := \phi_\varepsilon * f$, where ϕ is the standard mollifier. Then for every $\varepsilon_0 > 0$ we have $f_\varepsilon \to f$ in $W^{k,p}(U_{\varepsilon_0})$ if $1 \le p < \infty$, where $U_\varepsilon := \{x \in U | \operatorname{dist}(x, \mathbb{R}^n \setminus U) > \varepsilon\}$. If $p = \infty$ we have $\partial_\alpha f_\varepsilon \to \partial_\alpha f$ a.e. for all $|\alpha| \le k$.*

Proof. Just observe that all derivatives converge in L^p for $1 \le p < \infty$ since $\partial_\alpha f_\varepsilon = (\partial_\alpha \phi_\varepsilon) * f = \phi_\varepsilon * (\partial_\alpha f)$. Here the first equality is Lemma B.13 (ii) and the second equality only holds (by definition of the weak derivative) on U_ε since in this case $\operatorname{supp}(\phi_\varepsilon(x - .)) = B_\varepsilon(x) \subseteq U$. So if we fix $\varepsilon_0 > 0$, then $f_\varepsilon \to f$ in $W^{k,p}(U_{\varepsilon_0})$. In the case $p = \infty$ the claim follows since $L^\infty_{loc} \subseteq L^1_{loc}$ after passing to a subsequence. That selecting a subsequence is superfluous follows from Lemma B.14. □

Note that, by Lemma 9.11, if $f \in W^{k,\infty}(U)$ then $\partial_\alpha f$ is locally Lipschitz[3] continuous for all $|\alpha| \le k - 1$. Hence $\partial_\alpha f_\varepsilon \to \partial_\alpha f$ locally uniformly for all $|\alpha| \le k - 1$.

So in particular, we get convergence in $W^{k,p}(U)$ if f has compact support. To adapt this approach to work on all of U we will use a partition of unity.

Theorem 9.5 (Meyers[4]–Serrin[5]). *Let $U \subseteq \mathbb{R}^n$ be open and $1 \le p < \infty$. Then $C^\infty(U) \cap W^{k,p}(U)$ is dense in $W^{k,p}(U)$.*

Proof. The idea is to use a partition of unity to decompose f into pieces which are supported on layers close to the boundary and decrease the mollification parameter ε as we get closer to the boundary. To this end we start with the sets $U_j := \{x \in U | \operatorname{dist}(x, \mathbb{R}^n \setminus U) > \frac{1}{j}\}$ and we set $U_j := \emptyset$ for $j \le 0$. Now consider $\tilde{\zeta}_j := \phi_{\varepsilon_j} * \chi_{V_j}$, where $V_j := U_{j+1} \setminus U_{j-1}$ and ε_j chosen sufficiently small such that $\operatorname{supp}(\tilde{\zeta}_j) \subset U_{j+2} \setminus U_{j-2}$. Since the sets V_j cover U, the function $\tilde{\zeta} := \sum_j \tilde{\zeta}_j$ is positive on U and since for $x \in V_k$ only terms with $|j - k| \le 2$ contribute, we also have $\tilde{\zeta} \in C^\infty(U)$. Hence considering $\zeta_j := \tilde{\zeta}_j / \tilde{\zeta} \in C^\infty_c(U)$ we have $\sum_j \zeta_j = 1$.

Now let $f \in W^{k,p}(U)$ be given and fix $\delta > 0$. By the previous lemma we can choose $\varepsilon_j > 0$ sufficiently small such that $f_j := \phi_{\varepsilon_j} * (\zeta_j f)$ has still support inside $U_{j+2} \setminus U_{j-2}$ and satisfies

$$\|f_j - \zeta_j f\|_{W^{k,p}} \le \frac{\delta}{2^{j+1}}.$$

[3] Rudolf Lipschitz (1832–1903), German mathematician

[4] Norman George Meyers (*1930), American mathematician

[5] James Serrin (1926–2012), American mathematician

Then $f_\delta = \sum_j f_j \in C^\infty(U)$ since again only terms with $|j-k| \le 2$ contribute. Moreover, for every set $V \subset\subset U$ we have

$$\|f_\delta - f\|_{W^{k,p}(V)} = \|\sum_j (f_j - \zeta_j f)\|_{W^{k,p}(V)} \le \sum_j \|f_j - \zeta_j f\|_{W^{k,p}(V)} \le \delta$$

and letting $V \nearrow U$ we get $f_\delta \in W^{k,p}(U)$ as well as $\|f_\delta - f\|_{W^{k,p}(U)} \le \delta$. □

Historically this theorem had a significant impact since it showed that the two competing ways of defining Sobolev spaces, namely as the set of functions which have weak derivatives in L^p on one side and the closure of smooth functions with respect to the $W^{k,p}$ norm on the other side, actually agree.

Example 9.6. The example $f(x) := |x| \in W^{1,\infty}(-1,1)$ shows that the theorem fails in the case $p = \infty$ since $f'(x) = \mathrm{sign}(x)$ cannot be approximated uniformly by smooth functions. ⋄

For L^p we know that smooth functions with compact support are dense. This is no longer true in general for $W^{k,p}$ since convergence of derivatives enforces that the vanishing of boundary values is preserved in the limit. However, making this precise requires some additional effort. So for now we will just give the closure of $C_c^\infty(U)$ in $W^{k,p}(U)$ a special name $W_0^{k,p}(U)$ as well as $H_0^k(U) := W_0^{k,2}(U)$. It is easy to see that $C_c^k(U) \subseteq W_0^{k,p}(U)$ for every $1 \le p \le \infty$ and $W_c^{k,p}(U) \subseteq W_0^{k,p}(U)$ for every $1 \le p < \infty$ (mollify to get a sequence in $C_c^\infty(U)$ which converges in $W^{k,p}(U)$). In the case $p = \infty$ we have $W_0^{k,\infty}(U) \subseteq C_0^k(U)$ (with equality for nice domains, see Problem 9.18). Moreover, note

Lemma 9.6. *We have $W_0^{k,p}(\mathbb{R}^n) = W^{k,p}(\mathbb{R}^n)$ for $1 \le p < \infty$.*

Proof. We choose some cutoff function $\zeta_m \in C_c^\infty(\mathbb{R}^n)$, such that $\zeta_m(x) = 1$ for $|x| \le m$, $\zeta_m(x) = 0$ for $|x| \ge m+1$, and $\|\partial_\alpha \zeta_m\|_\infty \le C_\alpha$. For example, choose a function $h \in C^\infty(\mathbb{R})$ such $h(x) = 1$ for $x \le 0$, $h(x) = 0$ for $x \ge 1$ and let $\zeta_m(x) := h(|x| - m)$. Now note that for $f \in L^p(\mathbb{R}^n)$ dominated convergence implies $\zeta_m f \to f$ in L^p and $(\partial_\alpha \zeta_m) f \to 0$ in L^p for $|\alpha| \ge 1$.

Fix $f \in W^{k,p}(\mathbb{R}^n)$ and consider $f_m := f\zeta_m \in W_c^{k,p}(\mathbb{R}^n)$. Then using Leibniz rule we see that $\partial_\alpha f_m \to \partial_\alpha f$ in $L^p(\mathbb{R}^n)$ for $|\alpha| \le k$ and hence $f_m \to f$ in $W^{k,p}(\mathbb{R}^n)$. Thus $W_c^{k,p}(\mathbb{R}^n)$ is dense in $W^{k,p}(\mathbb{R}^n)$ and by mollification $C_c^\infty(\mathbb{R}^n)$ is dense in $W^{k,p}(\mathbb{R}^n)$. □

Next we collect some basic properties of weak derivatives.

Lemma 9.7. *Let $U \subseteq \mathbb{R}^n$ be open and $1 \le p \le \infty$.*

(i) *The operator* $\partial_\alpha : W^{k,p}(U) \to W^{k-|\alpha|,p}(U)$ *is a bounded linear map and* $\partial_\beta\partial_\alpha f = \partial_\alpha\partial_\beta f = \partial_{\alpha+\beta} f$ *for* $f \in W^{k,p}$ *and all multi-indices* α, β *with* $|\alpha| + |\beta| \le k$.

(ii) *We have*

$$\int_U g(\partial_\alpha f) d^n x = (-1)^{|\alpha|} \int_U (\partial_\alpha g) f d^n x, \quad g \in W_0^{k,q}(U),\ f \in W^{k,p}(U), \tag{9.7}$$

for all $|\alpha| \le k$, $\frac{1}{p} + \frac{1}{q} = 1$.

(iii) *Suppose* $f \in W^{1,p}(U)$ *and* $g \in W^{1,q}(U)$ *with* $\frac{1}{r} := \frac{1}{p} + \frac{1}{q} \le 1$. *Then* $f \cdot g \in W^{1,r}(U)$ *and we have the* ***product rule***

$$\partial_j (f \cdot g) = (\partial_j f) g + f(\partial_j g), \qquad 1 \le j \le n. \tag{9.8}$$

The same claim holds with $q = p = r$ *if* $f \in W^{1,p}(U) \cap L^\infty(U)$.

(iv) *Suppose* $\eta \in C^1(\mathbb{R}^m)$ *has bounded derivatives and satisfies* $\eta(0) = 0$ *if* $|U| = \infty$. *Then the map* $f \mapsto \eta \circ f$ *is a continuous map* $W^{1,p}(U, \mathbb{R}^m) \to W^{1,p}(U)$ *and we have the* ***chain rule*** $\partial_j(\eta \circ f) = \sum_k (\partial_k \eta)(f) \partial_j f_k$. *If* $\eta(0) = 0$, *then composition with* η *will also map* $W_0^{1,p}(U, \mathbb{R}^m) \to W_0^{1,p}(U)$.

(v) *Let* $\psi : U \to V$ *be a* C^1 *diffeomorphism such that both* ψ *and* ψ^{-1} *have bounded derivatives. Then we have a bijective bounded linear map* $W^{1,p}(V) \to W^{1,p}(U)$, $f \mapsto f \circ \psi$ *and we have the* ***change of variables formula*** $\partial_j (f \circ \psi) = \sum_k (\partial_k f)(\psi) \partial_j \psi_k$.

(vi) *Let* U *be connected and suppose* $f \in W^{1,p}(U)$ *satisfies* $\partial_j f = 0$ *for* $1 \le j \le n$. *Then* f *is constant.*

Proof. (i) Problem 9.9.

(ii) Take limits in (9.1) using Hölder's inequality. If $g \in W_c^{k,q}(U)$ only the case $q = \infty$ is of interest which follows from dominated convergence.

(iii) First of all note that if $\phi, \varphi \in C_c^\infty(U)$, then $\phi\varphi \in C_c^\infty(U)$ and hence using the ordinary product rule for smooth functions and rearranging (9.1) with $\varphi \to \phi\varphi$ shows $\phi f \in W_c^{1,p}(U)$. Hence (9.7) with $g \to g\varphi \in W_c^{1,q}$ shows

$$\int_U g f (\partial_j \varphi) d^n x = - \int_U \big((\partial_j f) g + f(\partial_j g)\big) \varphi d^n x,$$

that is, the weak derivatives of $f \cdot g$ are given by the product rule and that they are in $L^r(U)$ follows from the generalized Hölder inequality (B.32).

(iv) Since the claim is trivially true for constant η, we can subtract $\eta(0)$ and can assume $\eta(0) = 0$ without loss of generality. Moreover, by assumption $|\nabla \eta| \le L$ and hence we have $|\eta(x) - \eta(y)| \le L|x - y|$ by the mean value theorem. Hence we see $\|\eta \circ f - \eta \circ g\|_p \le L\|f - g\|_p$ which shows that composition with η is a continuous map $L^p(U,\mathbb{R}^n) \to L^p(U)$. Also note that the proposed derivative will be in $L^p(U)$ provided $f \in W^{1,p}(U,\mathbb{R}^n)$.

Now suppose $f_n \to f$ in $W^{1,p}(U,\mathbb{R}^m)$ for some $V \subset U$. Then

$$\begin{aligned}\|(\nabla\eta)(f)\cdot\partial_j f - (\nabla\eta)(f_n)\cdot\partial_j f_n\|_{L^p} &\le L\|\partial_j f - \partial_j f_n\|_{L^p}\\ &\quad + \|((\nabla\eta)(f) - (\nabla\eta)(f_n))\cdot\partial_j f\|_{L^p},\end{aligned}$$

where the first norm tends to zero by assumption and the second by dominated convergence after passing to a subsequence which converges a.e. at least for $1 \le p < \infty$. In the case $p = \infty$ this holds since $\nabla\eta$ is uniformly continuous on bounded sets.

Now for $p < \infty$ we can choose f_n to be smooth (by Theorem 9.5). In this case the derivative of $\eta \circ f_n$ can be computed using the chain rule and the above argument shows that this formula remains true in the limit, that is, the proposed derivative is indeed the weak derivative. Since $W^{1,\infty}(U) \subset W^{1,1}_{loc}(U)$ this also covers the case $p = \infty$ by restricting to bounded sets.

Finally, the above argument also shows that if $f_n \to f$ in $W^{1,p}(U,\mathbb{R}^n)$ then every subsequence f_{n_j} has another subsequence f_{m_j} for which $\eta\circ f_{m_j} \to \eta \circ f$ in $W^{1,p}(U)$. This implies that $\eta \circ f_n \to \eta \circ f$.

For the last claim observe that composition with η maps $C^1_c(U,\mathbb{R}^n) \subset W^{1,p}_0(U,\mathbb{R}^n) \to C^1_c(U) \subset W^{1,p}_0(U)$ and hence the claim follows by density of these subspaces.

(v) If J_ψ denotes the Jacobi determinant, then using $|J_\psi| \ge C$ the change of variables formula implies

$$\int_U |f\circ\psi|^p d^n x \le \frac{1}{C}\int_U |f\circ\psi|^p |J_\psi| d^n x = \frac{1}{C}\int_V |f|^p d^n y,$$

which shows that composition with ψ is a homeomorphism between $L^p(U)$ and $L^p(V)$ for $1 \le p < \infty$. In the case $p = \infty$ we have $\|f\circ\psi\|_\infty = \|f\|_\infty$ and the claim is also true. To compute the weak derivative recall $L^p \subset L^1_{loc}$. Now let ϕ_ε be the standard mollifier and consider $f_\varepsilon := \phi_\varepsilon * f$. Then, using this fact, one computes

$$\begin{aligned}\int_U (f\circ\psi)\partial_j\varphi\, d^n x &= \lim_{\varepsilon\downarrow 0}\int_U (f_\varepsilon\circ\psi)(\partial_j\varphi) d^n x\\ &= \lim_{\varepsilon\downarrow 0}\int_U \sum_k ((\partial_k f_\varepsilon)\circ\psi)(\partial_j\psi_k)\varphi\, d^n x\\ &= \int_U \sum_k ((\partial_k f)\circ\psi)(\partial_j\psi_k)\varphi\, d^n x.\end{aligned}$$

This establishes the claim.

(vi) This is just a reformulation of Lemma B.17. □

Of course item (iv) can be applied to complex-valued functions upon observing that taking real and imaginary parts is a bounded (real) linear map $W^{1,p}(U) \to W^{1,p}(U, \mathbb{R}^2)$, $f \mapsto (\mathrm{Re}(f), \mathrm{Im}(f))$. However, the important case of taking absolute values is not covered by (iv).

Lemma 9.8. *For $f \in W^{1,p}(U)$, $1 \le p \le \infty$, we have $|f| \in W^{1,p}(U)$ with*

$$\partial_j |f|(x) = \begin{cases} \frac{\mathrm{Re}(f(x))}{|f(x)|} \partial_j \mathrm{Re}(f(x)) + \frac{\mathrm{Im}(f(x))}{|f(x)|} \partial_j \mathrm{Im}(f(x)), & f(x) \neq 0, \\ 0, & f(x) = 0. \end{cases} \tag{9.9}$$

In particular, $\big|\partial_j |f|(x)\big| \le |\partial_j f(x)|$. For $1 \le p < \infty$ this map is continuous on $W^{1,p}(U)$.

Furthermore, if f is real-valued we also have $f_\pm := \max(0, \pm f) \in W^{1,p}(U)$ with

$$\partial_j f_\pm(x) = \begin{cases} \pm\partial_j f(x), & \pm f(x) > 0, \\ 0, & else, \end{cases} \qquad \partial_j |f|(x) = \begin{cases} \partial_j f(x), & f(x) > 0, \\ -\partial_j f(x), & f(x) < 0, \\ 0, & else. \end{cases}$$

Moreover, if $f \in W_0^{1,p}(U)$, then $|f| \in W_0^{1,p}(U)$ for $1 \le p < \infty$.

Proof. In order to reduce it to (iv) from the previous lemma we will take $f_1 := \mathrm{Re}(f)$, $f_2 := \mathrm{Im}(f)$ and approximate the absolute value of f by $\eta_\varepsilon(f_1, f_2)$ with $\eta_\varepsilon(x,y) = \sqrt{x^2 + y^2 + \varepsilon^2} - \varepsilon$.

We start by noting that $\|\nabla \eta_\varepsilon\|_\infty \le 1$ and hence we can apply the chain rule (Lemma 9.7 (iv)) with η_ε to see

$$\int_U \varphi(x) \frac{f_1(x)\partial_j f_1(x) + f_2(x)\partial_j f_2(x))}{\sqrt{|f(x)|^2 + \varepsilon^2}} d^n x = -\int_U \partial_j \varphi(x) \eta_\varepsilon(f_1(x), f_2(x)) d^n x$$

Letting $\varepsilon \to 0$ (using dominated convergence) shows

$$\int_U \varphi(x) \frac{f_1(x)\partial_j f_1(x) + f_2(x)\partial_j f_2(x))}{|f(x)|} d^n x = -\int_U \partial_j \varphi(x) |f(x)| d^n x$$

(with the expression for the derivative understood as being 0 if $f(x) = 0$) and establishes the first part. The estimate for the derivative follows from $\mathrm{Re}(z_1)\mathrm{Re}(z_2) + \mathrm{Im}(z_1)\mathrm{Im}(z_2) = |z_1||z_2| \cos(\arg(z_2/z_1))$ for $z_1, z_2 \in \mathbb{C}$.

The second part follows from $f_\pm(x) = \frac{|f(x)| \pm f(x)}{2}$ and linearity of the weak derivative.

Moreover, using $\nabla f = \nabla f_+ - \nabla f_-$ shows that $\nabla f = 0$ for a.e. x with $f(x) = 0$. Hence if we have a sequence $f_n \to f$ in $W^{1,p}(U)$ we can choose a subsequence such that both f and ∇f converge pointwise a.e. Then, by the above formulas and the preceding remark, the same is true for $|f|$ and $\nabla |f|$. Thus dominated convergence shows $|f_n| \to |f|$ in $W^{1,p}(U)$ for $1 \le p < \infty$.

The claim for $f \in W_0^{1,p}(U)$ follows since if $f \in W_c^{1,p}(U)$ then $|f| \in W_c^{1,p}(U)$ and the claim follows by density. □

As byproduct of the proof we note:

Corollary 9.9. *For every $m \in \mathbb{R}$ and every $f \in W^{1,p}(U)$ we have $\nabla f = 0$ for a.e. x with $f(x) = m$.*

Of course this implies that item (iv) from Lemma 9.7 continues to hold if η is only piecewise C^1 with bounded derivative. To see this observe that by linearity it suffices to consider the case where η has only one kink. But then η can be written as the sum of a C^1 function and a multiple of a translated absolute value.

Example 9.7. The example $f_\varepsilon(x) := x - \varepsilon \in W^{1,\infty}(-1,1)$ shows that taking absolute values is not continuous in $W^{1,\infty}(-1,1)$ since $|f_\varepsilon|'(x) = \operatorname{sign}(x-\varepsilon)$ does not converge uniformly to $|f_0|'(x) = \operatorname{sign}(x)$. ◇

Finally we look at situations where it is not a priori known that the function has a weak derivative. We will offer two variants. The first variant (ii) shows that an estimate on $T_a f - f$ is sufficient, where $T_a f(x) := f(x-a)$ is the translation operator from (B.28). Note that the estimate (ii) below should be thought of as an estimate for the difference quotient

$$D_j^\varepsilon f := \frac{T_{\varepsilon\delta^j} f - f}{-\varepsilon} \tag{9.10}$$

in the direction of the j'th coordinate axis. Our second variant employs duality and requires that the integral in (iii) below gives rise to a bounded functional. Both characterizations fail in the case $p = 1$.

Lemma 9.10. *For $f \in L^p(U)$ consider*

(i) *$f \in W^{1,p}(U)$ with $\|\nabla f\|_p \le C$.*

(ii) *There exists a constant C such that*

$$\|T_a f - f\|_{L^p(V)} \le C|a| \tag{9.11}$$

for every $V \subset\subset U$ and all $a \in \mathbb{R}^n$ with $|a| < \operatorname{dist}(V, \partial U)$.

(iii) *There exists a constant C with*

$$\left| \int_U f \big(\nabla \varphi\big) d^n x \right| \le C \|\varphi\|_{p'}, \qquad \varphi \in C_c^\infty(U), \tag{9.12}$$

where p' is the dual index, $\frac{1}{p} + \frac{1}{p'} = 1$.

Then we have (i) ⇒ (ii) ⇔ (iii) for $1 \le p \le \infty$ and (i) ⇔ (ii) ⇔ (iii) for $1 < p \le \infty$.

Proof. (i) $\Rightarrow$ (ii): By Theorem 9.5 we can assume that f is smooth without loss of generality and hence

$$|f(x-a)-f(x)| \le |a| \int_0^1 |\nabla f(x-t\,a)| dt$$

from which the case $p=\infty$ is immediate. In the case $1 \le p < \infty$ we integrate this inequality over V and employ Jensen's inequality to obtain

$$\begin{aligned}
\|T_a f - f\|^p_{L^p(V)} &\le |a|^p \int_V \left| \int_0^1 |\nabla f(x-t\,a)| dt \right|^p d^n x \\
&\le |a|^p \int_V \int_0^1 |\nabla f(x-t\,a)|^p dt\, d^n x \\
&\le |a|^p \int_0^1 \int_V |\nabla f(x-t\,a)|^p d^n x\, dt \le |a|^p \|\nabla f\|^p_{L^p(U)}.
\end{aligned}$$

(ii) $\Rightarrow$ (iii): Fix φ and choose some $V \subset\subset U$ with $\operatorname{supp}(\varphi) \subset V$. Using

$$\int_U (T_a f - f)\varphi d^n x = \int_U f\,(T_{-a}\varphi - \varphi) d^n x$$

for $|a| < \operatorname{dist}(V, \partial U)$ we obtain from (ii)

$$\left| \int_U f\,(T_{-a}\varphi - \varphi) d^n x \right| = \left| \int_V (T_a f - f)\varphi d^n x \right| \le C|a| \|\varphi\|_{p'}.$$

Choosing $a = \varepsilon \delta^j$ and taking $\varepsilon \to 0$ we get

$$\left| \int_U f\,(\partial_j \varphi) d^n x \right| \le C \|\varphi\|_{p'},$$

which implies (iii) with C replaced by $\sqrt{n}C$.

(iii) $\Rightarrow$ (i) for $p \ne 1$: Item (iii) implies that $\ell_j(\varphi) := \int_U f\big(\partial_j\varphi\big) d^n x$ is a densely defined bounded linear functional on $L^{p'}(U)$. Hence by Theorem B.11 there is some $g_j \in L^p(U)$ (with $\|g_j\|_p \le C$) such that $\ell_j(\varphi) = -\int_U g_j \varphi d^n x$, that is $\partial_j f = g_j$.

This establishes the lemma in the case $p \ne 1$. The direction (iii) $\Rightarrow$ (ii) without the assumption $p \ne 1$ is left as an exercise (Problem 9.14). $\square$

Example 9.8. Consider $f(x) = \operatorname{sign}(x)$ on $U = \mathbb{R}$. We already know from Example 9.2 that f does not have a weak derivative. However, items (ii) and (iii) hold with $C = 2$.

The problem in the case $p = 1$ is that, by the Riesz–Markov[6] representation theorem (Theorem 6.10 from **[34]**), every bounded linear functional on

[6] Andrey Markov Jr. (1903–1979), Soviet mathematician

$C_0(U)$ is given by a complex measure. Hence in this case the weak derivatives of f are in general complex measures rather than functions, that is, there exist Borel measures μ_j such that

$$\int_U f(\partial_j \varphi) d^n x = -\int_U \varphi d\mu_j(x).$$

The optimal constant in (iii)

$$V_U(f) := \sup_{\varphi \in C_c^\infty(U), \|\varphi\|_\infty \le 1} \left| \int_U f(\nabla \varphi) d^n x \right| \tag{9.13}$$

is known as the **total variation** of f. It is a semi-norm since $V_U(f) = 0$ if and only if f is constant on every connected component by Lemma B.17. Accordingly the class of functions satisfying (ii) or (iii),

$$BV(U) := \{f \in L^1(U) | V_U(f) < \infty\}, \tag{9.14}$$

is known as the functions of **bounded variation**. In terms of the measures μ_j it is given by

$$V_U(f)^2 = \sum_{j=1}^n |\mu_j|(U)^2.$$

$\diamond$

In the case $p = \infty$ we can also characterize $W^{1,\infty}$ as follows:

Lemma 9.11. *We have $C_b^{0,1}(U) \subseteq W^{1,\infty}(U)$ with the embedding being continuous. Conversely, if U is convex then we have equality and the embedding is a homeomorphism.*

Proof. If $f \in C_b^{0,1}(U)$, then Lemma 9.10 implies $f \in W^{1,\infty}$ with $\|\nabla f\|_\infty \le [f]_1$.

Conversely, let $f \in W^{1,\infty}(U)$ and $f_\varepsilon = \phi_\varepsilon * f$ with ϕ the standard mollifier. Then by the mean value theorem

$$|f_\varepsilon(x) - f_\varepsilon(y)| \le \|\nabla f_\varepsilon\|_\infty |x - y| \le \|\nabla f\|_\infty |x - y|,$$

where the last inequality holds for ε sufficiently small. In fact, note that for $\varepsilon < \operatorname{dist}(x, \partial U)$ we have $\partial_j f_\varepsilon(x) = (\phi_\varepsilon * \partial_j f)(x)$ and the claim follows from Young's inequality (B.25). Now if x, y are Lebesgue points of f (cf. Lemma B.14), then we can take the limit $\varepsilon \to 0$ to conclude

$$|f(x) - f(y)| \le \|\nabla f\|_\infty |x - y| \qquad \text{a.e. } x, y \in U.$$

In particular, f is uniformly continuous on a dense subset and hence has a (Lipschitz) continuous extension to all of U. $\square$

Problem 9.1. *Consider $f(x) = \sqrt{x}$, $U = (0, 1)$. Compute the weak derivative. For which p is $f \in W^{1,p}(U)$?*

Problem 9.2. *Show that the derivative of the function from Example 9.1 is not integrable.*

Problem 9.3. *Consider the Hilbert space $H^1(0,1)$. Compute the orthogonal complement of the following subspaces:*
a) $H^1_0(0,1)$ *b)* $\{f \in H^1(0,1) | \int_0^1 f(x)dx = 0\}$

Problem* 9.4. *Show that for $f, g \in AC[a,b]$ we have the integration by parts formula*

$$\int_a^b f(x)g'(x)dx = f(b)g(b) - f(a)g(a) - \int_a^b f'(x)g(x)dx.$$

(Hint: Insert the definition on the left and use Fubini.)

Problem* 9.5. *Show that f is weakly differentiable in the interval (a,b) if and only if $f(x) = f(c) + \int_c^x h(t)dt$ is absolutely continuous and $f' = h$ in this case. (Hint: Lemma 9.7 (vi).)*

Problem* 9.6. *Show that if $f \in AC[a,b]$ and $f' \in L^p(a,b)$, $1 < p \le \infty$, then f is Hölder continuous:*

$$|f(x) - f(y)| \le \|f'\|_p |x-y|^{1-\frac{1}{p}}.$$

Show that the claim fails for $p = 1$: The function $f(x) = -\log(x)^{-1}$ is absolutely continuous but not Hölder continuous on $[0, \frac{1}{2}]$.

Problem* 9.7. *Consider $U := B_1(0) \subset \mathbb{R}^n$ and $f(x) = \tilde{f}(|x|)$ with $\tilde{f} \in C^1(0,1]$. Then $f \in W^{1,p}_{loc}(B_1(0) \setminus \{0\})$ and*

$$\partial_j f(x) = \tilde{f}'(|x|)\frac{x_j}{|x|}.$$

Show that if $\limsup_{r\to 0} r^{n-1}|\tilde{f}(r)| < \infty$, then $f \in W^{1,p}(B_1(0))$ if and only if $\tilde{f}, \tilde{f}' \in L^p((0,1), r^{n-1}dr)$.

Conclude that for $f(x) := |x|^{-\gamma}$, $\gamma > 0$, we have $f \in W^{1,p}(B_1(0))$ with

$$\partial_j f(x) = -\frac{\gamma x_j}{|x|^{\gamma+2}}$$

provided $\gamma < \frac{n-p}{p}$. (Hint: Use integration by parts on a domain which excludes $B_\varepsilon(0)$ and let $\varepsilon \to 0$.)

Problem 9.8. *Show that for $\varphi \in C_c^\infty(\mathbb{R}^n)$ we have*

$$\varphi(0) = -\frac{1}{S_n}\int_{\mathbb{R}^n} \frac{x}{|x|^n} \cdot \nabla\varphi(x) d^n x.$$

Hence this weak derivative cannot be interpreted as a function. (Hint: Start with $\varphi(0) = -\int_0^\infty \left(\frac{d}{dr}\varphi(r\omega)\right)dr = -\int_0^\infty \nabla\varphi(r\omega)\cdot\omega\, dr$ and integrate with respect to ω over the unit sphere S^{n-1}; cf. Lemma A.11.)

Problem* 9.9. *Show Lemma 9.7 (i).*

Problem* 9.10. *Suppose $f \in W^{k,p}(U)$ and $h \in C_b^k(U)$. Then $h \cdot f \in W^{k,p}(U)$ and we have* ***Leibniz' rule***

$$\partial_\alpha (h \cdot f) = \sum_{\beta \le \alpha} \binom{\alpha}{\beta} (\partial_\beta h)(\partial_{\alpha-\beta} f), \tag{9.15}$$

where $\binom{\alpha}{\beta} := \frac{\alpha!}{\beta!(\alpha-\beta)!}$, $\alpha! := \prod_{j=1}^m (\alpha_j!)$, and $\beta \le \alpha$ means $\beta_j \le \alpha_j$ for $1 \le j \le m$.

Problem 9.11. *Let $\psi : U \to V$ be a C^k diffeomorphism such that all derivatives of both ψ and ψ^{-1} are bounded. Then we have a bijective bounded liner map $W^{k,p}(V) \to W^{k,p}(U)$, $f \mapsto f \circ \psi$.*

Problem 9.12. *Suppose for each $x \in U$ there is an open neighborhood $V(x) \subseteq U$ such that $f \in W^{k,p}(V(x))$. Then $f \in W^{k,p}_{loc}(U)$. Moreover, if $\|f\|_{W^{k,p}(V)} \le C$ for every $V \subset\subset U$, then $f \in W^{k,p}(U)$.*

Problem 9.13. *Suppose $1 < p \le \infty$. Show that if $f_n \in W^{1,p}(U)$ is a sequence such that $f_n \to f$ in L^p and $\|\nabla f_n\|_p \le C$, then $f \in W^{1,p}(U)$. (Hint: Since $L^p(U)$ is the dual of the separable Banach space $L^{p'}(U)$ with $\frac{1}{p} + \frac{1}{p'} = 1$, we can extract a weak-$*$ convergent subsequence from any bounded sequence (Lemma 4.36 from* **[35]***).)*

Problem 9.14. *Establish the direction (iii) $\Rightarrow$ (ii) in Lemma 9.10 for arbitrary $1 \le p \le \infty$. (Hint:* (B.18)*.)*

Problem 9.15. *Show that $W^{k,p}(U) \cap W^{j,q}(U)$ (with $1 \le p, q \le \infty$, $j, k \in \mathbb{N}_0$) together with the norm $\|f\|_{W^{k,p} \cap W^{j,q}} := \|f\|_{W^{k,p}} + \|f\|_{W^{j,q}}$ is a Banach space.*

9.2. Extension and trace operators

To proceed further we will need to be able to extend a given function beyond its original domain U. As already pointed out before, simply setting it equal to zero on $\mathbb{R}^n \setminus U$ will in general create a non-differentiable singularity along the boundary. Moreover, considering $U = (-1, 0) \cup (0, 1)$ we have $f(x) := \operatorname{sign}(x) \in W^{1,p}(U)$ but it is not possible to extend f to $\mathbb{R}$ such that the extension is in $W^{1,p}(\mathbb{R})$.

Of course such problems do not arise if $f \in W_0^{1,p}(U)$ since we can simply extend f to a function on $\bar{f} \in W^{1,p}(\mathbb{R}^n)$ by setting $\bar{f}(x) = 0$ for $x \in \mathbb{R}^n \setminus U$ (Problem 9.16).

We will say that a domain $U \subseteq \mathbb{R}^n$ has the **extension property** if for all $1 \le p \le \infty$ there is an extension operator $E : W^{1,p}(U) \to W^{1,p}(\mathbb{R}^n)$ such that

- E is bounded, i.e., $\|Ef\|_{W^{1,p}(\mathbb{R}^n)} \le C_{U,p}\|f\|_{W^{1,p}(U)}$ and
- $Ef\big|_U = f$.

We begin by showing that if the boundary is a hyperplane, we can do the extension by a simple reflection. To this end consider the reflection $x^\star := (x_1, \dots, x_{n-1}, -x_n)$ which is an involution on $\mathbb{R}^n$. For a domain U which is symmetric with respect to reflection, that is, $U^\star = U$, write $U_\pm := \{x \in U | \pm x_n > 0\}$ and a function f defined on U_+ can be extended to $U_- \cup U_+$ using

$$f^\star(x) := \begin{cases} f(x), & x \in U_+, \\ f(x^\star), & x \in U_-. \end{cases} \tag{9.16}$$

Note that $f^\star$ extends to a continuous function in $C(\overline{U})$ provided $f \in C(\overline{U_+})$.

Lemma 9.12. *Let $U \subseteq \mathbb{R}^n$ be symmetric with respect to reflection and $1 \le p \le \infty$. If $f \in W^{1,p}(U_+)$ then the symmetric extension $f^\star \in W^{1,p}(U)$ satisfies $\|f^\star\|_{W^{1,p}(U)} = 2^{1/p}\|f\|_{W^{1,p}(U_+)}$. Moreover,*

$$(\partial_j f^\star) = \begin{cases} (\partial_j f)^\star, & 1 \le j < n, \\ \operatorname{sign}(x_n)(\partial_n f)^\star, & j = n. \end{cases} \tag{9.17}$$

Proof. It suffices to compute the weak derivatives. We start with $1 \le j < n$ and

$$\int_U f^\star \partial_j \varphi d^n x = \int_{U_+} f\, \partial_j \varphi^\# d^n x,$$

where $\varphi^\#(x) = \varphi(x) + \varphi(x^\star)$. Since $\varphi^\#$ is not compactly supported in U_+ we use a cutoff function $\eta_\varepsilon(x) = \eta(x_n/\varepsilon)$, where $\eta \in C^\infty(\mathbb{R}, [0,1])$ satisfies $\eta(r) = 0$ for $r \le \frac{1}{2}$ and $\eta(r) = 1$ for $r \ge 1$ (e.g., integrate and shift the standard mollifier to obtain such a function). Then

$$\begin{aligned} \int_U f^\star \partial_j \varphi\, d^n x &= \lim_{\varepsilon\to 0} \int_{U_+} f\, \partial_j(\eta_\varepsilon \varphi^\#) d^n x = -\lim_{\varepsilon\to 0}\int_{U_+} (\partial_j f)\eta_\varepsilon \varphi^\# d^n x \\ &= -\int_{U_+} (\partial_j f)\varphi^\# d^n x = -\int_U (\partial_j f)^\star \varphi\, d^n x \end{aligned}$$

for $1 \le j < n$. For $j = n$ we proceed similarly,

$$\int_U f^\star \partial_n \varphi\, d^n x = \int_{U_+} f\, \partial_n \varphi^\sharp d^n x,$$

where $\varphi^\sharp(x) = \varphi(x) - \varphi(x^\star)$. Note that $\varphi^\sharp(x_1, \dots, x_{n-1}, 0) = 0$ and hence $|\varphi^\sharp(x)| \le L\, x_n$ on U_+. Using this last estimate we have $|(\partial_n \eta_\varepsilon)\varphi^\sharp| \le C$ and hence we obtain as before

$$\int_U f^\star \partial_n \varphi \, d^n x = \lim_{\varepsilon \to 0} \int_{U_+} f \, \partial_n(\eta_\varepsilon \varphi^\sharp) d^n x = -\lim_{\varepsilon \to 0} \int_{U_+} (\partial_n f) \eta_\varepsilon \varphi^\sharp d^n x$$
$$= -\int_{U_+} (\partial_n f) \varphi^\sharp d^n x = -\int_U \operatorname{sign}(x_n) (\partial_n f)^\star \varphi \, d^n x,$$

which finishes the proof. □

Corollary 9.13. *$\mathbb{R}^n_+$ has the extension property. In fact, any rectangle (not necessarily bounded) Q has the extension property.*

Proof. Given a rectangle use the above lemma to extend it along every hyperplane bounding the rectangle. Finally, use a smooth cut-off function (e.g. mollify the characteristic function of a slightly larger rectangle). □

While this already covers some interesting domains, note that it fails if we look for example at the exterior of a rectangle. So our next result shows (maybe not too surprising), that it is the boundary which will play the crucial role. To this end we recall that U is said to have a C^1 boundary if around any point $x^0 \in \partial U$ we can find a C^1 diffeomorphism ψ which straightens out the boundary (cf. Section A.2). As a preparation we note:

Lemma 9.14. *Suppose U has a bounded C^1 boundary. Then there is a finite number of open sets $\{U_j\}_{j=0}^m$ and corresponding functions $\zeta_j \in C^\infty(\mathbb{R}^n)$ with $\operatorname{supp}(\zeta_j) \subset U_j$ such that $\sum_{j=0}^m \zeta_j(x) = 1$ for all $x \in U$, $U_0 \subseteq U$, $\{U_j\}_{j=1}^m$ are bounded and cover ∂U, and for each U_j, $1 \le j \le m$, there is a C^1 diffeomorphism $\psi_j : U_j \to Q_j$, where Q_j is a rectangle which is symmetric with respect to reflection.*

Proof. Since near each $x \in \partial U$ we can straighten out the boundary, there is a corresponding open neighborhood U_x and a C^1 diffeomorphism $\psi_x : U_x \to Q_x$, where Q_x is a rectangle which is symmetric with respect to reflection. Moreover, there is also a corresponding radius $r(x)$ such that $\bar{B}_{r(x)}(x) \subset U_x$. By compactness of ∂U there are finitely many points $\{x_j\}_{j=1}^m$ such that the corresponding balls $B_{r(x_j)}(x_j)$ cover the boundary. Take $U_j := U_{x_j}$. Choose nonnegative functions $\tilde{\zeta}_j \in C_c^\infty(U_j)$ such that $\tilde{\zeta}_j > 0$ on $\bar{B}_{r(x_j)}(x_j)$ (e.g. mollify the characteristic function of $B_{r(x_j)}(x_j)$). Let $V := \bigcup_{j=1}^m B_{r(x_j)}(x_j)$ and $U_0 := U$. Choose a nonnegative function $\tilde{\zeta}_0 \in C^\infty(\mathbb{R}^n)$ supported inside U such that $\tilde{\zeta}_0 > 0$ on $\overline{U} \setminus V$ and a nonnegative function $\tilde{\zeta}_{m+1} \in C^\infty(\mathbb{R}^n)$ supported on $\mathbb{R}^n \setminus \overline{U}$ such that $\tilde{\zeta}_{m+1} > 0$ on $\mathbb{R}^n \setminus (U \cup V)$ (e.g. again by mollification of the corresponding characteristic functions). Then $\zeta := \sum_{j=0}^{m+1} \tilde{\zeta}_j \in C^\infty(\mathbb{R}^n)$ is positive on $\mathbb{R}^n$ and $\zeta_j := \tilde{\zeta}_j / \zeta$ are the functions we are looking for. □

Now we are ready to show:

Lemma 9.15. *Suppose U has a bounded C^1 boundary, then U has the extension property. Moreover, the extension of a continuous function can be chosen continuous and if U is bounded, the extension can be chosen with compact support.*

Proof. Choose functions ζ_j as in Lemma 9.14 and split $f \in W^{1,p}(U)$ according to $\sum_j f_j$, where $f_j := \zeta_j f$. Then f_0 can be extended to $\mathbb{R}^n$ by setting it equal to 0 outside U. Moreover, f_j can be mapped to $Q_{j,+}$ using ψ_j and extended to Q_j using the symmetric extension. Note that this extension has compact support and so has the pull back $\bar{f}_j$ to U_j; in particular, it can be extended to $\mathbb{R}^n$ by setting it equal to 0 outside U_j. By construction we have $\|\bar{f}_j\|_{W^{1,p}(U_j)} \le C_j \|f_j\|_{W^{1,p}(U_j)}$ and the product rule implies $\|f_j\|_{W^{1,p}(U_j)} \le \tilde{C}_j \|f\|_{W^{1,p}(U)}$. Hence $\bar{f} := \sum_j \bar{f}_j$ is the required extension.

The last claim follows since the symmetric extension of a continuous function is continuous. □

As a first application note that by mollifying an extension we see that we can approximate by functions which are smooth up to the boundary. Moreover, using a suitable cutoff function we can also assume that the approximating functions have compact support.

Corollary 9.16. *Suppose U has the extension property, then $C_c^\infty(\mathbb{R}^n)$ is dense in $W^{1,p}(U)$ for $1 \le p < \infty$.*

Proof. Simply mollify an extension. If U is bounded the extension will have compact support and its mollification will be in $C_c^\infty(\mathbb{R}^n)$. If U is unbounded, multiply the mollification with a cutoff function to get a function from $C_c^\infty(\mathbb{R}^n)$ as in the proof of Lemma 9.6. □

Corollary 9.17. *Suppose U has the extension property, then $W^{1,\infty}(U) = C_b^{0,1}(U)$ with equivalent norms.*

Proof. Since U has the extension property, we can extend f to $W^{1,\infty}(\mathbb{R}^n)$ and hence f is Lipschitz continuous by Lemma 9.11. □

Note that it is sometimes also of interest to look at the corresponding extension problem for $W^{k,p}(U)$ with $k \ge 1$. It can be shown that there is an extension operator $E : W^{k,p}(U) \to W^{k,p}(\mathbb{R}^n)$ provided the boundary satisfies a local Lipschitz condition (see Theorem VI.5 in **[28]** for details).

Next we show that functions in $W^{1,p}$ have boundary values in L^p. This might be surprising since a function from $W^{1,p}(U)$ is only defined almost everywhere and the boundary ∂U is a set of measure zero. Please recall that for U with a C^1 boundary there is a corresponding surface measure dS and by $L^p(\partial U)$ we will always understand $L^p(\partial U, dS)$.

Theorem 9.18. *Suppose U has a bounded C^1 boundary, then there exists a bounded* ***trace operator***

$$T : W^{1,p}(U) \to L^p(\partial U) \tag{9.18}$$

which satisfies $Tf = f\big|_{\partial U}$ for $f \in C(\overline{U}) \cap W^{1,p}(U)$. Moreover, we have $|Tf| = T|f|$ and for real-valued functions also $(Tf)_\pm = Tf_\pm$.

Proof. In the case $p = \infty$ functions from $W^{1,\infty}(U)$ are Lipschitz continuous by Corollary 9.17 and hence continuous up to the boundary. So there is nothing to do.

Thus we can focus on the case $1 \le p < \infty$. As a preparation we note that by Corollary 9.16 the set $C(\overline{U}) \cap W^{1,p}(U)$ is dense by the previous lemma and that the Gauss–Green theorem continues to hold for $u \in C(\overline{U}, \mathbb{R}^n) \cap W^{1,1}(U, \mathbb{R}^n)$ if U is bounded. To see this choose $u \in C(\overline{U}, \mathbb{R}^n) \cap W^{1,1}(U, \mathbb{R}^n)$ and extend it to a function $\bar{u} \in C_c(\mathbb{R}^n, \mathbb{R}^n) \cap W^{1,1}(\mathbb{R}^n, \mathbb{R}^n)$. Then the Gauss–Green theorem holds for the mollification $u_\varepsilon := \phi_\varepsilon * \bar{u}$ and since we have $u_\varepsilon \to u$ uniformly on $\overline{U}$ as well as $\partial_j u_\varepsilon \to \partial_j u$ in $L^1(U, \mathbb{R}^n)$ the Gauss–Green theorem remains true in the limit $\varepsilon \to 0$.

Now take $f \in C(\overline{U}) \cap W^{1,p}(U)$. As in the proof of Lemma 9.15, using a partition of unity and straightening out the boundary, we can reduce it to the case where $f \in C^1_c(\mathbb{R}^n)$ has compact support $\operatorname{supp}(f) \subset Q$ such that $\partial U \cap \operatorname{supp}(f) \subset \partial \mathbb{R}^n_+$. Then using the Gauss–Green theorem and assuming f real-valued without loss of generality we have (cf. Problem 9.17)

$$\begin{aligned}\int_{\partial U} |f|^p d^{n-1}x &= -\int_{Q_+} (|f|^p)_{x_n} d^n x = -p \int_{Q_+} \operatorname{sign}(f)|f|^{p-1}(\partial_n f) d^n x \\ &\le p\|f\|_p^{p-1} \|\nabla f\|_p,\end{aligned}$$

where we have used Hölders inequality in the last step. Hence the trace operator defined on $C(\overline{U}) \cap W^{1,p}(U)$ is bounded and since the latter set is dense, there is a unique extension to all of $W^{1,p}(U)$.

To see the last claim observe that if $f_n \in C(\overline{U}) \cap W^{1,p}(U) \to f$, then $|f_n| \in C(\overline{U}) \cap W^{1,p}(U) \to |f|$ by Lemma 9.8. Hence $|Tf| = \lim_{n\to\infty} |Tf_n| = \lim_{n\to\infty} T|f_n| = T|f|$. □

Of course this result can also be applied to derivatives:

Corollary 9.19. *Suppose U has a bounded C^1 boundary, then there exists a bounded trace operator*

$$T : W^{k,p}(U) \to W^{k-1,p}(\partial U) \tag{9.19}$$

which satisfies $Tf = f\big|_{\partial U}$ for $f \in C^{k-1}(\overline{U}) \cap W^{k,p}(U)$.

As an application we can extend the Gauss–Green theorem and integration by parts to $W^{1,p}$ vector fields.

Lemma 9.20. *Let U be a bounded C^1 domain in $\mathbb{R}^n$ and $u \in W^{1,p}(U, \mathbb{R}^n)$ a vector field. Then the Gauss–Green formula* (A.43) *holds if the boundary values of u are understood as traces as in the previous theorem. Moreover, the integration by parts formula* (A.44) *also holds for $f \in W^{1,p}(U)$, $g \in W^{1,q}(U)$ with $\frac{1}{p} + \frac{1}{q} = 1$.*

Proof. Since U has the extension property, we can extend u to $W^{1,p}(\mathbb{R}^n, \mathbb{R}^n)$. Consider the mollification $u_\varepsilon := \phi_\varepsilon * u$ and apply the Gauss–Green theorem to u_ε. Now let $\varepsilon \to 0$ and observe that the left-hand side converges since $\partial_j u_\varepsilon \to \partial_j u$ in $L^p \subset L^1$. Similarly the right-hand side converges by continuity of the trace operator. The integration by parts formula follows from the Gauss–Green theorem applied to the product fg and employing the product rule. □

Finally we identify the kernel of the trace operator.

Lemma 9.21. *If U is a bounded C^1 domain in $\mathbb{R}^n$, then the kernel of the trace operator is given by* $\operatorname{Ker}(T) = W_0^{1,p}(U)$ *for* $1 \le p < \infty$.

Proof. Clearly $W_0^{1,p}(U) \subseteq \operatorname{Ker}(T)$. Conversely it suffices to show that $f \in \operatorname{Ker}(T)$ can be approximated by functions from $C_c^\infty(U)$. Using a partition of unity as in the proof of Lemma 9.15 we can assume $U = Q_+$, where Q is a rectangle which is symmetric with respect to reflection. Setting

$$\bar{f}(x) = \begin{cases} f(x), & x \in Q_+, \\ 0, & x \in Q_-, \end{cases}$$

integration by parts using the previous lemma shows

$$\int_Q \bar{f} \partial_j \varphi \, d^n x = \int_{Q_+} f \partial_j \varphi \, d^n x = -\int_{Q_+} (\partial_j f) \varphi \, d^n x, \quad \varphi \in C_c^\infty(Q),$$

that $\bar{f} \in W^{1,p}(Q)$ with $\partial_j \bar{f}(x) = \partial_j f(x)$ for $x \in Q_+$ and $\partial_j \bar{f}(x) = 0$ else. Now consider $f_\varepsilon(x) = (\phi_{\varepsilon/2} * \bar{f})(x - \varepsilon \delta^n) \in C_c^\infty(Q_+)$ with ϕ the standard mollifier. This is the required sequence by Lemma B.14 and Problem B.15. □

Problem* 9.16. *Show that $f \in W_0^{k,p}(U)$ can be extended to a function $\bar{f} \in W_0^{k,p}(\mathbb{R}^n)$ by setting it equal to zero outside U. In this case the weak derivatives of $\bar{f}$ are obtained by setting the weak derivatives of f equal to zero outside U.*

Problem* 9.17. *Suppose $\gamma \ge 1$. Show that $f \in W^{1,p}(U)$ implies $|f|^\gamma \in W^{1,p/\gamma}(U)$ with $\partial_j |f|^\gamma = \gamma |f|^{\gamma-1} \partial_j |f|$. (Hint: Lemma 9.8.)*

Problem 9.18. *Suppose U has a bounded C^1 boundary. Show that $W_0^{1,\infty}(U) = C_0^1(U)$. (Hint: Use Lemma 9.14 to reduce it to the case of a straight boundary. Near the straight boundary use a cutoff η_ε as in the proof of Lemma 9.12.)*

Problem 9.19. *Let $1 \le p < \infty$ and U bounded. Show that $Tf = f\big|_{\partial U}$ defined on $C(\overline{U}) \subseteq L^p(U) \to L^p(\partial U)$ is unbounded (and hence has no meaningful extension to $L^p(U)$). (Hint: Take a sequence which equals 1 on the boundary and converges to 0 in the interior.)*

Problem 9.20. *Suppose $u \in H_0^1(B_1(0))$ satisfies $u(x) = -u(x^\star)$, where $x^\star := (x_1, \dots, x_{n-1}, -x_n)$. Show that $u \in H_0^1(B_1(0) \cap \mathbb{R}^n_+)$, where $\mathbb{R}^n_+ = \{x \in \mathbb{R}^n | x_n > 0\}$.*

Problem 9.21. *Consider the punctured ball $U := B_1(0) \setminus \{0\}$. Show that $W_0^{1,p}(U) = W_0^{1,p}(B_1(0))$ and $W^{1,p}(U) = W^{1,p}(B_1(0))$ for $p < n$.*

Problem 9.22. *Let $1 \le p < \infty$ and U bounded. Show that $Tf = f\big|_{\partial U}$ defined on $C(\overline{U}) \subseteq L^p(U) \to L^p(\partial U)$ is unbounded (and hence has no meaningful extension to $L^p(U)$). (Hint: Take a sequence which equals 1 on the boundary and converges to 0 in the interior.)*

Problem 9.23. *Show that $C_0(\overline{U}) \cap W^{1,p}(U) \subseteq W_0^{1,p}(U)$, $1 \le p < \infty$. (Hint: Of course, if U has a nice boundary, this is immediate using traces. For the general case use an approximation based on Lemma 9.8.)*

9.3. Embedding theorems

We have already seen that functions in $W^{1,p}$ are not necessarily continuous (unless $n = 1$). This raises the question in what sense a function from $W^{1,p}$ is better than a function from L^p? For example, is it in L^q for some q other than p? In this respect it is instructive to look at an example which should be understood as a benchmark for the results to follow.

Example 9.9. Let $U := B_1(0)$ and consider $f(x) := |x|^{-\gamma}$. Then by Problem 9.7

$$\partial_j f(x) = -\gamma \frac{x_j}{|x|} |x|^{-\gamma-1},$$

where the factor $\frac{|x_j|}{|x|} \le 1$ is bounded. Hence by Example A.8 we have $f \in W^{1,p}(U)$ provided $\gamma < \frac{n}{p} - 1$. Since we have $f \in L^q(U)$ provided $\gamma < \frac{n}{q}$ the optimal index for which $f \in L^{p^*}(U)$ is $p^* := \frac{np}{n-p}$ provided $n > p$. If $n < p$, then we have $-\gamma > 1 - \frac{n}{p} > 0$ and hence f is continuous. In fact it will be Hölder continuous of exponent $1 - \frac{n}{p}$.

Of course we can also take higher derivatives into account. To this end, using induction, it is straightforward to verify that

$$\partial_\alpha f(x) = \frac{P_\alpha(x)}{|x|^{|\alpha|}} |x|^{-\gamma-|\alpha|},$$

where P_α is a homogenous polynomial of degree $|\alpha|$. In particular, note that the factor $P_\alpha(x)|x|^{-|\alpha|}$ is bounded. Hence the optimal index for which $f \in L^{p^*}(U)$ provided $f \in W^{k,p}(U)$ is $p_k^* := \frac{np}{n-kp}$ for $n > pk$. For $n < pk$ we will have $f \in C^{k-l-1}$, where $l \in \lfloor \frac{n}{p} \rfloor$ with the highest derivative being Hölder continuous of exponent $1 - \frac{n}{p} + l$. ⋄

Theorem 9.22 (Gagliardo[7]–Nirenberg–Sobolev)**.** *Suppose $1 \le p < n$ and $U \subseteq \mathbb{R}^n$ is open. Then the natural embedding $W_0^{1,p}(U) \hookrightarrow L^q(U)$ is continuous for all $p \le q \le p^*$, where $\frac{1}{p^*} := \frac{1}{p} - \frac{1}{n}$. Moreover,*

$$\|f\|_{p^*} \le \frac{p(n-1)}{(n-p)} \prod_{j=1}^n \|\partial_j f\|_p^{1/n} \le \frac{p(n-1)}{n(n-p)} \sum_{j=1}^n \|\partial_j f\|_p. \tag{9.20}$$

Proof. It suffices to prove the case $q = p^*$ since the rest follows from interpolation (Problem B.14). Moreover, by density it suffices to prove the inequality for $f \in C_c^\infty(\mathbb{R}^n)$. In this respect note that if you have a sequence $f_n \in C_c^\infty(U)$ which converges to some f in $W_0^{1,p}(U)$, then by (9.20) this sequence will also converge in $L^{p^*}(U)$ and by considering pointwise convergent subsequences both limits agree.

We start with the case $p = 1$ and observe

$$|f(x)| = \left| \int_{-\infty}^{x_1} \partial_1 f(r, \tilde{x}_1) dr \right| \le \int_{-\infty}^{\infty} |\partial_1 f(r, \tilde{x}_1)| dr,$$

where we denote by $\tilde{x}_j := (x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)$ the vector obtained from x with the j'th component dropped. Denote by $f_1(\tilde{x}_1)$ the right-hand side of the above inequality and apply the same reasoning to the other coordinate directions to obtain

$$|f(x)|^n \le \prod_{j=1}^n f_j(\tilde{x}_j).$$

Now we claim that if $f_j \in L^1(\mathbb{R}^{n-1})$, then

$$\Big\| \prod_{j=1}^n f_j(\tilde{x}_j)^{\frac{1}{n-1}} \Big\|_{L^1(\mathbb{R}^n)} \le \prod_{j=1}^n \|f_j\|_{L^1(\mathbb{R}^{n-1})}^{\frac{1}{n-1}}.$$

[7] Emilio Gagliardo (1930–2008), Italian Mathematician

For $n = 2$ this is just Fubini and hence we can use induction. To this end fix the last coordinate x_{n+1} and apply Hölder's inequality and the induction hypothesis to obtain

$$\int_{\mathbb{R}^n} \prod_{j=1}^{n+1} |f_j(\tilde{x}_j)|^{\frac{1}{n}} d^n x \le \|f_{n+1}^{1/n}\|_{L^n(\mathbb{R}^n)} \Big\| \prod_{j=1}^{n} |f_j(\tilde{x}_j)|^{\frac{1}{n}} \Big\|_{L^{n/(n-1)}(\mathbb{R}^n)}$$

$$= \|f_{n+1}\|_{L^1(\mathbb{R}^n)}^{1/n} \Big\| \prod_{j=1}^{n} |f_j(\tilde{x}_j)|^{\frac{1}{n-1}} \Big\|_{L^1(\mathbb{R}^n)}^{1-\frac{1}{n}} \le \|f_{n+1}\|_{L^1(\mathbb{R}^n)}^{1/n} \prod_{j=1}^{n} \|f_j(\tilde{x}_j)\|_{L^1(\mathbb{R}^{n-1})}^{1/n}.$$

Now integrate this inequality with respect to the missing variable x_{n+1} and use the iterated Hölder inequality (Problem B.13 with $r = 1$ and $p_j = n$) to obtain the claim.

Moreover, applying this to our situation where $\|f_j\|_{L^1(\mathbb{R}^{n-1})} = \|\partial_j f\|_1$ we obtain

$$\|f\|_{n/(n-1)}^{n/(n-1)} \le \prod_{j=1}^{n} \|\partial_j f\|_1^{\frac{1}{n-1}}$$

which is precisely (9.20) for the case $p = 1$ (the second inequality in (9.20) is just the inequality of arithmetic and geometric means). To see the case of general p let $f \in C_c^\infty(\mathbb{R}^n)$ and apply the case $p = 1$ to $f \to |f|^\gamma$ for $\gamma > 1$ to be determined and recall Problem 9.17. Then

$$\left(\int_{\mathbb{R}^n} |f|^{\frac{\gamma n}{n-1}} d^n x\right)^{\frac{n-1}{n}} \le \prod_{j=1}^{n} \left(\int_{\mathbb{R}^n} \big|\partial_j |f|^\gamma\big| d^n x\right)^{1/n} \tag{9.21}$$

$$= \gamma \prod_{j=1}^{n} \left(\int_{\mathbb{R}^n} |f|^{\gamma-1} |\partial_j f| d^n x\right)^{1/n} \le \gamma \| |f|^{\gamma-1} \|_{p/(p-1)} \prod_{j=1}^{n} \|\partial_j f\|_p^{1/n},$$

where we have used Hölder in the last step. Now we choose $\gamma := \frac{p(n-1)}{n-p} > 1$ such that $\frac{\gamma n}{n-1} = \frac{(\gamma-1)p}{p-1} = p^*$, which gives the general case. □

Note that a simple scaling argument (Problem 9.25) shows that (9.20) can only hold for p^*. Furthermore, using an extension operator this result also extends to $W^{1,p}(U)$:

Corollary 9.23. *Suppose U has the extension property and $1 \le p < n$, then the natural embedding $W^{1,p}(U) \hookrightarrow L^q(U)$ is continuous for every $p \le q \le p^*$, where $\frac{1}{p^*} = \frac{1}{p} - \frac{1}{n}$.*

Proof. Let $\tilde{f} = Ef \in W^{1,p}(\mathbb{R}^n)$ be an extension of $f \in W^{1,p}(U)$. Then $\|f\|_{L^q(U)} \le \|\tilde{f}\|_{L^q(\mathbb{R}^n)} \le C_p \|\tilde{f}\|_{W^{1,p}(\mathbb{R}^n)} \le C_p C_{U,p} \|f\|_{W^{1,p}(U)}$, where we have used $W_0^{1,p}(\mathbb{R}^n) = W^{1,p}(\mathbb{R}^n)$ (Lemma 9.6). □

Note that involving the extension operator implies that we need the full $W^{1,p}$ norm to bound the L^{p*} norm. A constant function shows that indeed an inequality involving only the derivatives on the right-hand side cannot hold on bounded domains (cf. also Theorem 9.34).

In the borderline case $p = n$ one has $p^* = \infty$, however, the example in Problem 9.26 shows that functions in $W^{1,n}$ can be unbounded if $n > 1$ (for $n = 1$ see the example after the lemma). Nevertheless, we have at least the following result:

Lemma 9.24. *Suppose $p = n$ and $U \subseteq \mathbb{R}^n$ is open. Then the natural embedding $W_0^{1,n}(U) \hookrightarrow L^q(U)$ is continuous for every $n \le q < \infty$.*

Proof. As before it suffices to establish $\|f\|_q \le C\|f\|_{W^{1,n}}$ for $f \in C_c^\infty(\mathbb{R}^n)$. To this end we employ (9.21) with $p = n$ implying

$$\|f\|_{\gamma n/(n-1)}^{\gamma} \le \gamma \|f\|_{(\gamma-1)n/(n-1)}^{\gamma-1} \prod_{j=1}^{n} \|\partial_j f\|_n^{1/n} \le \frac{\gamma}{n} \|f\|_{(\gamma-1)n/(n-1)}^{\gamma-1} \sum_{j=1}^{n} \|\partial_j f\|_n.$$

Using Young's inequality (Problem 9.28), $\alpha^{(\gamma-1)/\gamma}\beta^{1/\gamma} \le \frac{\gamma-1}{\gamma}\alpha + \frac{1}{\gamma}\beta$ for nonnegative numbers $\alpha, \beta \ge 0$, this gives

$$\|f\|_{\gamma n/(n-1)} \le C\Big(\|f\|_{(\gamma-1)n/(n-1)} + \sum_{j=1}^{n} \|\partial_j f\|_n\Big).$$

Now choosing $\gamma = n$ we get $\|f\|_{n^2/(n-1)} \le C\|f\|_{W^{1,n}}$ and by interpolation (Problem B.14) the claim holds for $q \in [n, n\frac{n}{n-1}]$. So we can choose $\gamma = n+1$ to get the claim for $q \in [n, (n+1)\frac{n}{n-1}]$ and iterating this procedure finally gives the claim for $q \in [n, (n+k)\frac{n}{n-1}]$, which establishes the result. □

Corollary 9.25. *Suppose U has the extension property and $p = n$, then the natural embedding $W^{1,n}(U) \hookrightarrow L^q(U)$ is continuous for every $n \le q < \infty$.*

Example 9.10. If $n = 1$ with $U = (a, b)$ an interval we have $W^{1,1}(a, b) \hookrightarrow C_b(a, b)$ in the borderline case $n = p$. In the case of a bounded interval this was shown in Example 9.5. If $f \in W^{1,1}(a, b)$ then it is (locally) absolutely continuous and we can choose some $c \in (a, b)$ and write $f(x) = f(c) + \int_c^x f'(y)dy$ implying $|f(x)| \le |f(c)| + \|f'\|_1$. Averaging this last equation with respect to c shows $\|f\|_\infty \le \frac{\|f\|_1}{b-a} + \|f'\|_1$ (with the first term understood as being zero if (a, b) is unbounded). ◇

In the case $p > n$ functions from $W^{1,p}$ will be continuous (in the sense that there is a continuous representative). In fact, they will even be bounded Hölder continuous functions and hence are continuous up to the boundary (cf. Theorem B.2 and the discussion after this theorem).

Theorem 9.26 (Morrey). *Suppose $n < p \le \infty$ and $U \subseteq \mathbb{R}^n$ is open. The natural embedding $W_0^{1,p}(U) \hookrightarrow C_0^{0,\gamma}(\overline{U})$, where $\gamma = 1 - \frac{n}{p}$, is continuous. Here $C_0^{0,\gamma}(\overline{U}) := C_b^{0,\gamma}(\overline{U}) \cap C_0(U)$ is the space of Hölder continuous functions vanishing at the boundary.*

Proof. In the case $p = \infty$ there is nothing to do, since $W_0^{1,\infty}(U) \subseteq C_0^1(U)$ and $[f]_1 \le \|\nabla f\|_\infty$. Hence we can assume $n < p < \infty$. Moreover, as before, by density we can assume $f \in C_c^\infty(\mathbb{R}^n)$.

We begin by considering a cube Q of side length r containing 0. Then, for $x \in Q$ and $\bar{f} = r^{-n} \int_Q f(x) d^n x$ we have

$$\bar{f} - f(0) = r^{-n} \int_Q \big(f(x) - f(0)\big) d^n x = r^{-n} \int_Q \int_0^1 \frac{d}{dt} f(tx) dt\, d^n x$$

and hence

$$\begin{aligned} |\bar{f} - f(0)| &\le r^{-n} \int_Q \int_0^1 |\nabla f(tx)| |x| dt\, d^n x \le r^{1-n} \int_Q \int_0^1 |\nabla f(tx)| dt\, d^n x \\ &= r^{1-n} \int_0^1 \int_{tQ} |\nabla f(y)| \frac{d^n y}{t^n}\, dt \le r^{1-n} \int_0^1 \|\nabla f\|_{L^p(tQ)} \frac{|tQ|^{1-1/p}}{t^n}\, dt \\ &\le \frac{r^\gamma}{\gamma} \|\nabla f\|_{L^p(Q)} \end{aligned}$$

where we have used Hölder's inequality in the fourth step. By a translation this gives

$$|\bar{f} - f(x)| \le \frac{r^\gamma}{\gamma} \|\nabla f\|_{L^p(Q)}$$

for any cube Q of side length r containing x and combining the corresponding estimates for two points we obtain

$$|f(x) - f(y)| \le \frac{2\|\nabla f\|_{L^p(Q)}}{\gamma} |x - y|^\gamma \tag{9.22}$$

for any cube containing both x and y (note that we can choose the side length of Q to be $r = \max_{1 \le j \le n} |x_j - y_j| \le |x - y|$). Since we can of course replace $L^p(Q)$ by $L^p(\mathbb{R}^n)$ we get Hölder continuity of f. Moreover, taking a cube of side length $r = 1$ containing x we get (using again Hölder)

$$|f(x)| \le |\bar{f}| + \frac{2\|\nabla f\|_{L^p(Q)}}{\gamma} \le \|f\|_{L^p(Q)} + \frac{2\|\nabla f\|_{L^p(Q)}}{\gamma} \le C \|f\|_{W^{1,p}(\mathbb{R}^n)}$$

establishing the theorem. □

Corollary 9.27. *Suppose U has the extension property and $n < p \le \infty$, then there is a continuous embedding $W^{1,p}(U) \hookrightarrow C_b^{0,\gamma}(\overline{U})$, where $\gamma = 1 - \frac{n}{p}$.*

Proof. Let $\tilde{f} = Ef \in W^{1,p}(\mathbb{R}^n)$ be an extension of $f \in W^{1,p}(U)$. Then $\|f\|_{C_b^{0,\gamma}(\overline{U})} \le \|\tilde{f}\|_{C_b^{0,\gamma}(\mathbb{R}^n)} \le C_p\|f\|_{W^{1,p}(\mathbb{R}^n)} \le C_p C_{U,p}\|f\|_{W^{1,p}(U)}$, where we have used $W_0^{1,p}(\mathbb{R}^n) = W^{1,p}(\mathbb{R}^n)$ and Morrey's theorem if $p < \infty$ and Lemma 9.11 if $p = \infty$. □

Example 9.11. The example from Problem 9.27 shows that for a domain with a cusp, functions from $W^{1,p}$ might be unbounded (and hence in particular not in $C^{1,\gamma}$) even for $p > n$. ⋄

Example 9.12. For $p = \infty$ this embedding is surjective in case of a convex domain (Lemma 9.11) or a domain with the extension property (Corollary 9.17). However, for $n < p < \infty$ this is not the case. To see this consider the Takagi function (Problem 4.27 from [**34**]) which is in $C^{1,\gamma}[0,1]$ for every $\gamma < 1$ but not absolutely continuous (not even of bounded variation) and hence not in $W^{1,p}(0,1)$ for any $1 \le p \le \infty$. Note that this example immediately extends to higher dimensions by considering $f(x) = b(x_1)$ on the unit cube. ⋄

As a consequence of the proof we also get that for $n < p$ Sobolev functions are differentiable a.e.

Lemma 9.28. *Suppose $n < p \le \infty$ and $U \subseteq \mathbb{R}^n$ is open. Then $f \in W^{1,p}_{loc}(U)$ is differentiable a.e. and the a.e. derivative equals the weak derivative.*

Proof. Since $W^{1,\infty}_{loc} \subseteq W^{1,p}_{loc}$ for any $p < \infty$ we can assume $n < p < \infty$. Let $x \in U$ be an L^p Lebesgue point of the gradient, that is,

$$\lim_{r\downarrow 0} \frac{1}{|Q_r(x)|} \int_{Q_r(x)} |\nabla f(x) - \nabla f(y)|^p d^n y = 0,$$

where $Q_r(x)$ is a cube of side length r containing x. Now let $y \in Q_r(x)$ and $r = |y - x|$ (by shrinking the cube w.l.o.g.). Then replacing $f(y) \to f(y) - f(x) - \nabla f(x)\cdot(y-x)$ in (9.22) we obtain

$$\big|f(y) - f(x) - \nabla f(x)\cdot(y-x)\big| \le \frac{2}{\gamma}|x-y|^\gamma \left(\int_{Q_r(x)} |\nabla f(x) - \nabla f(z)|^p d^n z\right)^{1/p}$$
$$= \frac{2}{\gamma}|x-y| \left(\frac{1}{|Q_r(x)|}\int_{Q_r(x)} |\nabla f(x) - \nabla f(z)|^p d^n z\right)^{1/p}$$

and, since x is an L^p Lebesgue point of the gradient, the right-hand side is $o(|x-y|)$, that is, f is differentiable at x and its gradient equals its weak gradient. □

Note that since by Lemma 9.11 every locally Lipschitz continuous function is locally $W^{1,\infty}$, we obtain as an immediate consequence:

Theorem 9.29 (Rademacher)**.** *Every locally Lipschitz continuous function is differentiable almost everywhere.*

So far we have only looked at first order derivatives. However, we can also cover the case of higher order derivatives by repeatedly applying the above results to the fact that $\partial_j f \in W^{k-1,p}(U)$ for $f \in W^{k,p}(U)$.

Theorem 9.30. *Suppose $U \subseteq \mathbb{R}^n$ is open and $1 \le p \le \infty$. The following natural embeddings are continuous:*

$$W_0^{k,p}(U) \hookrightarrow L^q(U), \quad q \in [p, p_k^*] \text{ if } \frac{1}{p_k^*} = \frac{1}{p} - \frac{k}{n} > 0,$$

$$W_0^{k,p}(U) \hookrightarrow L^q(U), \quad q \in [p, \infty) \text{ if } \frac{1}{p} = \frac{k}{n},$$

$$W_0^{k,p}(U) \hookrightarrow C_0^{k-l-1,\gamma}(\overline{U}), \quad l = \lfloor \frac{n}{p} \rfloor, \begin{cases} \gamma = 1 - \frac{n}{p} + l, & \frac{n}{p} \notin \mathbb{N}_0, \\ \gamma \in [0,1), & \frac{n}{p} \in \mathbb{N}_0, \end{cases} \text{ if } \frac{1}{p} < \frac{k}{n}.$$

If in addition $U \subseteq \mathbb{R}^n$ has the extension property, then the following natural embeddings are continuous:

$$W^{k,p}(U) \hookrightarrow L^q(U), \quad q \in [p, p_k^*] \text{ if } \frac{1}{p_k^*} = \frac{1}{p} - \frac{k}{n} > 0,$$

$$W^{k,p}(U) \hookrightarrow L^q(U), \quad q \in [p, \infty) \text{ if } \frac{1}{p} = \frac{k}{n},$$

$$W^{k,p}(U) \hookrightarrow C_b^{k-l-1,\gamma}(\overline{U}), \quad l = \lfloor \frac{n}{p} \rfloor, \begin{cases} \gamma = 1 - \frac{n}{p} + l, & \frac{n}{p} \notin \mathbb{N}_0, \\ \gamma \in [0,1), & \frac{n}{p} \in \mathbb{N}_0, \end{cases} \text{ if } \frac{1}{p} < \frac{k}{n}.$$

Proof. If $\frac{1}{p} > \frac{k}{n}$ we apply Theorem 9.22 to successively conclude $\|\partial^\alpha f\|_{L^{p_j^*}} \le C\|f\|_{W_0^{k,p}}$ for $|\alpha| \le k - j$ for $j = 1, \dots, k$. If $\frac{1}{p} = \frac{k}{n}$ we proceed in the same way but use Lemma 9.24 in the last step. If $\frac{1}{p} < \frac{k}{n}$ we first apply Theorem 9.22 l times as before. If $\frac{n}{p}$ is not an integer we then apply Theorem 9.26 to conclude $\|\partial^\alpha f\|_{C_0^{0,\gamma}} \le C\|f\|_{W_0^{k,p}}$ for $|\alpha| \le k - l - 1$. If $\frac{n}{p}$ is an integer, we apply Theorem 9.22 $l-1$ times and then Lemma 9.24 once to conclude $\|\partial^\alpha f\|_{L^q} \le C\|f\|_{W_0^{k,p}}$ for any $q \in [p,\infty)$ for $|\alpha| \le k - l$. Hence we can apply Theorem 9.26 to conclude $\|\partial^\alpha f\|_{C_0^{0,\gamma}} \le C\|f\|_{W_0^{k,p}}$ for any $\gamma \in [0,1)$ for $|\alpha| \le k - l - 1$.

The second part follows analogously using the corresponding results for domains with the extension property. □

Note that for $p = 1$ we have a slightly stronger result $W_0^{n,1}(U) \hookrightarrow C_0(U)$ in the borderline case $k = n$ — see Problem 9.32.

Moreover, for $q \in [p, p^*)$ the embedding is even compact (it fails for $q = p^*$ — see Problem 9.29).

Theorem 9.31 (Rellich–Kondrachov[8])**.** *Suppose $U \subseteq \mathbb{R}^n$ is open and bounded and $1 \le p \le \infty$. Then following natural embeddings are compact:*

$$W_0^{1,p}(U) \hookrightarrow L^q(U), \quad q \in [p, p^*), \quad \frac{1}{p^*} = \frac{1}{p} - \frac{1}{n}, \text{ if } p \le n,$$
$$W_0^{1,p}(U) \hookrightarrow C_0(\overline{U}), \text{ if } p > n.$$

If in addition $U \subseteq \mathbb{R}^n$ has the extension property, then the following natural embeddings are compact:

$$W^{1,p}(U) \hookrightarrow L^q(U), \quad q \in [p, p^*), \quad \frac{1}{p^*} = \frac{1}{p} - \frac{1}{n}, \text{ if } p \le n,$$
$$W^{1,p}(U) \hookrightarrow C(\overline{U}), \text{ if } p > n.$$

Proof. The case $p > n$ follows from $W_0^{1,p}(U) \hookrightarrow C_0^{0,\gamma}(\overline{U}) \hookrightarrow C_0(\overline{U})$, where the first embedding is continuous by Theorem 9.26 and the second is compact by Theorem B.3. Similarly in the second case using $W^{1,p}(U) \hookrightarrow C_b^{0,\gamma}(\overline{U}) \hookrightarrow C(\overline{U})$, where the first embedding is continuous by Corollary 9.27.

Next consider the case $p < n$. Let $F \subset W^{1,p}(U)$ (or $F \subset W_0^{1,p}(U)$) be a bounded subset. Using an extension operator (or Problem 9.16) we can assume that $F \subset W^{1,p}(\mathbb{R}^n)$ with $\operatorname{supp}(f) \subseteq V$ for all $f \in F$ and some fixed set V. By Lemma 9.10 (applied on $\mathbb{R}^n$) we have

$$\|T_a f - f\|_p \le |a| \|\nabla f\|_p$$

and using the interpolation inequality from Problem B.14 and Theorem 9.30 we have

$$\|T_a f - f\|_q \le \|T_a f - f\|_p^{1-\theta} \|T_a f - f\|_{p^*}^{\theta} \le |a|^{1-\theta} \|\nabla f\|_p^{1-\theta} C^{\theta} \|f\|_{1,p}^{\theta},$$

where $\frac{1}{q} = \frac{1-\theta}{p} + \frac{\theta}{p^*}$, $\theta \in [0, 1]$. Hence F is relatively compact by Theorem B.15. In the case $p = n$, we can replace p^* by any value larger than q. □

Since for bounded U the embedding $C(\overline{U}) \hookrightarrow L^p(U)$ is continuous, we obtain:

Corollary 9.32. *Under the assumptions of the above theorem the natural embeddings $W_0^{k+1,p}(U) \hookrightarrow W_0^{k,p}(U)$ and $W^{k+1,p}(U) \hookrightarrow W^{k,p}(U)$ are compact.*

[8] Vladimir Iosifovich Kondrashov (1909–1971), Russian mathematician

Proof. By the Rellich–Kondrachov theorem the embedding $W^{1,p}(U) \hookrightarrow L^p(U)$ is compact. Hence, given a bounded sequence in $W^{k+1,p}(U)$ we can find a subsequence for which all partial derivatives of order up to k converge in $L^p(U)$. Hence this sequence converges in $W^{k,p}(U)$ by Corollary 9.2. $\square$

Example 9.13. Note that the Rellich–Kondrachov theorem fails for $q = p^*$. To see this choose a nonzero function $f \in W_0^{1,p}(U)$ with compact support in some small ball. Now consider $f_\varepsilon(x) := \varepsilon^{-n/p^*} f(x/\varepsilon)$. Then $\|f_\varepsilon\|_{1,p} \le \|f\|_{1,p}$ and $\|f_\varepsilon\|_{p^*} = \|f\|_{p^*}$. If f_ε had a convergent subsequence in $L^{p^*}(U)$, this subsequence must converge to 0 since $f_\varepsilon(x) \to 0$ a.e., a contradiction. $\diamond$

Example 9.14. Note that the Rellich–Kondrachov theorem fails on $\mathbb{R}^n$. To see this choose $\varphi \in C_c^\infty(\mathbb{R}^n)$ with support in $B_{1/2}(0)$ and consider $F = \{\varphi_k := \varphi(. - k\delta^1) | k \in \mathbb{N}\}$. Now note that both the $W^{1,p}$ as well as the L^q norm of φ_k are independent of k and two different functions have disjoint supports. So there is no way to extract a convergent subsequence and an extra condition is needed. $\diamond$

Theorem 9.33. *Let $1 \le p \le n$. A set $F \subseteq W^{1,p}(\mathbb{R}^n)$ is relatively compact in $L^q(\mathbb{R}^n)$ for every $q \in [p, p^*)$ if F is bounded and for every $\varepsilon > 0$ there is some $r > 0$ such that $\|(1 - \chi_{B_r(0)})f\|_p < \varepsilon$ for all $f \in F$.*

Proof. Condition (i) of Theorem B.15 is verified literally as in the previous theorem. Similarly condition (ii) follows from interpolation since $\|(1 - \chi_{B_r(0)})f\|_q \le \|(1-\chi_{B_r(0)})f\|_p^{1-\theta}\|(1-\chi_{B_r(0)})f\|_{p^*}^\theta \le \|(1-\chi_{B_r(0)})f\|_p^{1-\theta}\|f\|_{p^*}^\theta$. $\square$

Note that this extra condition might come for free in case of radial symmetry (Problem 14.24).

As a consequence of Theorem 9.31 we also can get an important inequality.

Theorem 9.34 (Poincaré inequality)**.** *Let $U \subset \mathbb{R}^n$ be open and bounded. Then for $f \in W_0^{1,p}(U)$, $1 \le p \le \infty$, we have*

$$\|f\|_p \le C\|\nabla f\|_p. \tag{9.23}$$

If in addition U is a connected subset with the extension property, then for $f \in W^{1,p}(U)$, $1 \le p \le \infty$, we have

$$\|f - (f)_U\|_p \le C\|\nabla f\|_p, \tag{9.24}$$

where $(f)_U := \frac{1}{|U|}\int_U f\, d^n x$ is the average of f over U.

Proof. We begin with the second case and argue by contradiction. If the claim were wrong we could find a sequence of functions $f_m \in W^{1,p}(U)$ such that $\|f_m - (f_m)_U\|_p > m\|\nabla f_m\|_p$. hence the function $g_m := \|f_m - (f_m)_U\|_p^{-1}(f_m - (f_m)_U)$ satisfies $\|g_m\|_p = 1$, $(g_m)_U = 0$, and $\|\nabla g_m\|_p < \frac{1}{m}$.

In particular, the sequence is bounded and by Corollary 9.32 we can assume $g_m \to g$ in $L^p(U)$ without loss of generality. Moreover, $\|g\|_p = 1$, $(g)_U = 0$, and

$$\int_U g\partial_j\varphi d^n x = \lim_{m\to\infty} \int_U g_m \partial_j \varphi d^n x = -\lim_{m\to\infty} \int_U (\partial_j g_m)\varphi d^n x = 0.$$

That is, $\partial_j g = 0$ and since U is connected, g must be constant on U by Lemma 9.7 (vi). Moreover, $(g)_U = 0$ implies $g = 0$ contradicting $\|g\|_p = 1$.

To see the first case we proceed similarly to find a sequence $g_m := \|f_m\|_p^{-1} f_m$ producing a limit such that $\|g\|_p = 1$ and $\partial_j g = 0$. Now take a ball $B := B_r(0)$ containing U such that $B \setminus U$ has positive Lebesgue measure. Observe that we can extend f_m to $\bar{f}_m \in W^{1,p}(B)$ by setting it equal to 0 outside U which will give a corresponding sequence $\bar{g}_m := \|f_m\|_p^{-1}\bar{f}_m$ and a corresponding limit $\bar{g}$. Since B is connected we again get that $\bar{g}$ is constant on B and since $\bar{g}$ vanishes on $B \setminus U$ it must vanish on all of B contradicting $\|\bar{g}\|_p = 1$. □

Example 9.15. Using the Poincaré inequality we can shed some further light on the case $f \in W^{1,n}(\mathbb{R}^n)$ from Lemma 9.24. First note that a simple scaling shows that the constant C_r for a ball of radius r in the Poincaré inequality is given by $C_r = C_1 r$. Hence using Poincaré's and Hölder's inequalities we obtain

$$\begin{aligned} \frac{1}{|B_r|}\int_{B_r(x)} |f(y) - (f)_{B_r(x)}| d^n y &\le C_1 r \int_{B_r(x)} |\nabla f(y)| \frac{d^n y}{|B_r|} \\ &\le C_1 r \left(\int_{B_r(x)} |\nabla f(y)|^n \frac{d^n y}{|B_r|} \right)^{1/n} \le \frac{C_1}{V_n^{1/n}} \|\nabla f\|_n. \end{aligned}$$

Locally integrable functions for which the left-hand side is bounded are called functions of **bounded mean oscillation** $\mathrm{BMO}(\mathbb{R}^n)$ and one sets

$$\|f\|_{\mathrm{BMO}} = \sup_{x,r} \frac{1}{|B_r|} \int_{B_r(x)} |f(y) - (f)_{B_r(x)}| d^n y.$$

It is straightforward to verify that this is a semi-norm and $\|f\|_{\mathrm{BMO}} = 0$ if and only if f is constant. ◇

Finally it is often important to know when $W^{1,p}(U)$ is an algebra: By the product rule we have $\partial_j(fg) = (\partial_j f)g + f(\partial_j g)$ and for this to be in L^p we need that f, g are bounded which follows from Morrey's inequality (Theorem 9.26) if $n < p$. Working a bit harder one can even show:

Theorem 9.35. *Suppose $U \subseteq \mathbb{R}^n$ is open. If $\frac{1}{p} < \frac{k}{n}$, then $W_0^{k,p}(U)$ is a Banach algebra with*

$$\|fg\|_{k,p} \le C \|f\|_{k,p} \|g\|_{k,p}. \tag{9.25}$$

If U is bounded and has the extension property, the result also holds for $W^{k,p}(U)$.

Proof. First of all it suffices to show the inequality for the case when f and g are $C^\infty \cap W^{k,p}$. Moreover, by Leibniz' rule (Problem 9.10) it suffices to estimate $\|(\partial_\alpha f)(\partial_\beta g)\|_p$ for $|\alpha| + |\beta| \le k$. To this end we will use the generalized Hölder inequality (Problem B.12) and hence we need to find $1 \le q_\alpha, q_\beta \le \infty$ with $\frac{1}{p} = \frac{1}{q_\alpha} + \frac{1}{q_\beta}$ such that $W^{m-|\alpha|,p} \hookrightarrow L^{q_\alpha}$ and $W^{m-|\beta|,p} \hookrightarrow L^{q_\beta}$.

Let l be the largest integer such that $\frac{1}{p} < \frac{k-l}{n}$. Then Theorem 9.30 allows us to choose $q_\alpha = \infty$, $q_\beta = p$ for $|\alpha| \le l$ and similarly $q_\alpha = p$, $q_\beta = \infty$ for $|\beta| \le l$. Otherwise, that is if $\frac{1}{p} \ge \frac{k-|\alpha|}{n}$ and $\frac{1}{p} \ge \frac{k-|\beta|}{n}$ then Theorem 9.30 imposes the restrictions $\frac{1}{q_\alpha} \ge \frac{1}{p} - \frac{k-|\alpha|}{n}$ and $\frac{1}{q_\beta} \ge \frac{1}{p} - \frac{k-|\beta|}{n}$. Hence $\frac{1}{q_\alpha} + \frac{1}{q_\beta} \ge \frac{1}{p} - \left(\frac{k}{n} - \frac{1}{p}\right)$ and we can find the required indices. □

Problem 9.24. *Show that for $f \in H_0^1((a,b))$ we have*

$$\|f\|_\infty^2 \le 2\|f\|_2\|f'\|_2.$$

Show that the inequality continues to hold if $f \in H^1(\mathbb{R})$ or $f \in H^1((0,\infty))$. (Hint: Start by differentiating $|f(x)|^2$.)

Problem* 9.25. *Show that the inequality $\|f\|_q \le C\|\nabla f\|_p$ for $f \in W^{1,p}(\mathbb{R}^n)$ can only hold for $q = \frac{np}{n-p}$. (Hint: Consider $f_\lambda(x) = f(\lambda x)$.)*

Problem* 9.26. *Show that $f(x) := \log\log(1 + \frac{1}{|x|})$ is in $W^{1,n}(B_1(0))$ if $n > 1$. (Hint: Problem 9.7.)*

Problem* 9.27. *Consider $U := \{(x,y) \in \mathbb{R}^2 | 0 < x, y < 1,\ x^\beta < y\}$ and $f(x,y) := y^{-\alpha}$ with $\alpha, \beta > 0$. Show $f \in W^{1,p}(U)$ for $p < \frac{1+\beta}{(1+\alpha)\beta}$. Now observe that for $0 < \beta < 1$ and $\alpha < \frac{1-\beta}{2\beta}$ we have $2 < \frac{1+\beta}{(1+\alpha)\beta}$.*

Problem* 9.28. *Prove* ***Young's inequality***

$$\alpha^{1/p}\beta^{1/q} \le \frac{1}{p}\alpha + \frac{1}{q}\beta, \qquad \frac{1}{p} + \frac{1}{q} = 1, \quad \alpha, \beta \ge 0.$$

Show that equality occurs precisely if $\alpha = \beta$. (Hint: Take logarithms on both sides.)

Problem* 9.29. *Let $U = B_1(0) \subset \mathbb{R}^n$ and consider*

$$u_m(x) = \begin{cases} m^{\frac{n}{p}-1}(1 - m|x|), & |x| < \frac{1}{m}, \\ 0, & \text{else}. \end{cases}$$

Show that u_m is bounded in $W^{1,p}(U)$ for $1 \le p < n$ but has no convergent subsequence in $L^{p^*}(U)$. (Hint: The beta integral from Problem A.8 might be useful.)

Problem 9.30 (Lions' lemma).[9] *Let X, Y, and Z be Banach spaces. Assume X is compactly embedded into Y and Y is continuously embedded into Z. Show that for every $\varepsilon > 0$ there exists some $C(\varepsilon)$ such that*

$$\|x\|_Y \le \varepsilon \|x\|_X + C(\varepsilon)\|x\|_Z.$$

Problem 9.31. *Suppose $U \subseteq \mathbb{R}^n$ is bounded and has the extension property. Show that there exists a constant C such that*

$$\|f\|_{k,p} \le C \left(\sum_{|\alpha|=k} \|\partial_\alpha f\|_p + \|f\|_p \right)$$

(Hint: Problem 9.30 and Corollary 9.32.)

Problem 9.32. *Let $U \subseteq \mathbb{R}^n$. Show that there is a bounded embedding $W_0^{n,1}(U) \hookrightarrow C_0(U)$ satisfying*

$$\|f\|_\infty \le \|\partial_{(1,\dots,1)} f\|_1.$$

Problem 9.33. *Let U be a bounded domain with a C^1 boundary and $1 \le p < \infty$. Show that for every $a > 0$ there is a constant C such that*

$$\int_U |f|^p d^n x \le C \left(\int_U |\nabla f|^p d^n x + a \int_{\partial U} |f|^p dS \right), \qquad f \in W^{1,p}(U).$$

Problem 9.34. *Show that item (v) From Lemma 9.7 holds for bi-Lipschitz maps ψ.*

Note that with this result all results from the present chapter can be extended from C^1 to Lipschitz domains.

(Hint: Use that the change of variables formula for integrals holds for bi-Lipschitz functions and Lemma 9.28.)

[9]Jacques-Louis Lions (1928–2001), French mathematician

Chapter 10

Elliptic equations

10.1. The Poisson equation

As a warmup we will start by looking at the **Poisson equation**

$$\begin{aligned} -\Delta u(x) &= f(x), && x \in U, \\ u(x) &= 0, && x \in \partial U, \end{aligned} \tag{10.1}$$

on a bounded domain $U \subseteq \mathbb{R}^n$ with Dirichlet boundary conditions. Our analysis will apply both in the case of real as well as complex solutions. Hence we will look at complex solutions which contain real solutions as a special case.

If we regard the derivatives as weak derivatives, then our equation reads

$$-\int_U (\Delta\varphi)(x)u(x)d^n x = \int_U \varphi(x)f(x)d^n x, \qquad \varphi \in C_c^\infty(U), \tag{10.2}$$

or, after an integration by parts, we can also write it in the more symmetric form

$$\int_U (\nabla\varphi)\cdot(\nabla u)d^n x = \int_U \varphi(x)f(x)d^n x, \qquad \varphi \in C_c^\infty(U). \tag{10.3}$$

Now recall that by the Poincaré inequality (Theorem 9.34) we have a scalar product

$$\langle v, u\rangle := \int_U (\nabla v)^* \cdot (\nabla u)d^n x \tag{10.4}$$

on $H_0^1(U)$ whose associated norm is equivalent to the usual one. Here '$*$' denotes complex conjugation, which of course can be dropped in case one is only interested in real solutions.

Moreover, using the fact that $C_c^\infty(U) \subset H_0^1(U)$ is dense we see that we can write our last form as

$$\langle v, u\rangle = \langle v, f\rangle_2, \qquad v \in H_0^1(U), \tag{10.5}$$

where $\langle v, u\rangle_2 := \int_U v(x)^* u(x) d^n x$ denotes the scalar product in $L^2(U)$.

We will call a solution $u \in H_0^1(U)$ of (10.5) a **weak solution** of the Dirichlet problem (10.1). If a solution is, in addition, in $H^2(U)$, it is called a **strong solution**. In this case we can undo our integration by parts and conclude that u solves (10.1), where the derivatives are understood as weak derivatives and the boundary condition is understood in the sense of traces (at least for U with sufficiently smooth boundary; see Lemma 9.21).

Finally note that (10.5) should be more precisely written as

$$\langle v, u\rangle = \langle Jv, f\rangle_2, \qquad v \in H_0^1(U), \tag{10.6}$$

where $J : H_0^1(U) \hookrightarrow L^2(U)$ is the natural embedding. Since this embedding is bounded (in fact, even compact; we will come back to this later), we can use the adjoint operator to write this as

$$\langle v, u\rangle = \langle v, J^* f\rangle, \qquad v \in H_0^1(U), \tag{10.7}$$

which shows that the weak problem (10.5) has a unique solution $u = J^* f \in H_0^1(U)$ for every $f \in L^2(U)$. Also note the estimate $\|u\| = \|J^* f\| \le C\|f\|_2$, where C is the optimal constant from the Poincaré inequality (since $\|J^*\| = \|J\| = C$).

Moreover, observe that in the complex case the solution corresponding to f^* is u^* and hence, by uniqueness, the weak solution corresponding to a real-valued f will also be real-valued.

In summary,

Theorem 10.1. *Suppose $U \subseteq \mathbb{R}^n$ is a bounded domain. The Poisson equation (10.1) has a unique weak solution $u \in H_0^1(U)$ for every $f \in L^2(U)$. This solutions satisfies $\|\nabla u\|_2 \le C\|f\|_2$ and will be real-valued if f is real-valued.*

Moreover, note that many principles carry over to weak solutions. For example, let us call a function $v \in L^1_{loc}(U, \mathbb{R})$ **weakly subharmonic** if

$$\int_U v \Delta\varphi \, d^n x \ge 0, \qquad \varphi \in C_c^\infty(U, [0,\infty)). \tag{10.8}$$

Similarly, v is called **weakly superharmonic** if the inequality is reversed and weakly harmonic if equality holds. It can be shown that this definition generalizes the classical definition from Section 5.2 (Problem 10.4). As a motivation for the definition, observe that if $v \in C^2(U)$ then integration by parts and Lemma B.16 shows that (10.8) reduces to the usual condition

$\Delta v \geq 0$ in this case (cf. Lemma 5.14). Moreover, it turns out that a weakly harmonic function is harmonic (see again Problem 10.4).

To formulate the comparison principle in this setting (the analog of Lemma 5.16) we will write $v \leq u$ on ∂U for two real-valued functions $u, v \in H^1(U)$ provided the positive part $(v-u)_+$ is in $H^1_0(U)$. Note that if U is a domain with a bounded C^1 boundary, then $f \leq 0$ on the boundary in the sense of traces if and only if $Tf_+ = (Tf)_+ = 0$ and hence if and only if $f_+ \in H^1_0(U)$ by Lemma 9.21. Thus the condition reduces to the natural one for such domains.

Lemma 10.2 (Subharmonic functions are subsolutions)**.** *Let $u, v \in H^1(U, \mathbb{R})$ with u harmonic and v weakly subharmonic. Then $v \leq u$ on ∂U implies $v \leq u$ on U.*

Proof. By considering $v - u$ we can assume $u = 0$ without loss of generality and hence we have $v_+ \in H^1_0(U, [0, \infty))$. Furthermore, for $v \in H^1(U)$ we can use integration by parts to rewrite (10.8) as

$$\int_U (\nabla\varphi) \cdot (\nabla v) d^n x \leq 0, \qquad \varphi \in C^\infty_c(U, [0, \infty)).$$

and by approximation this even holds for all $\varphi \in H^1_0(U, [0, \infty))$. Choosing $\varphi = v_+$ we get

$$0 \geq \int_U (\nabla v_+) \cdot (\nabla v) d^n x = \int_U |\nabla v_+|^2 d^n x.$$

Hence $\nabla v_+ = 0$ and by Lemma 9.7 (vi) we see that v_+ is constant on every connected component. Since it vanishes on the boundary, it is zero (if a component has no boundary, i.e. if $U = \mathbb{R}^n$, observe that the only square integrable constant is zero). □

Setting

$$\sup_{x \in \partial U} f(x) := \inf\{M \in \mathbb{R} | f \leq M \text{ on } \partial U\} \tag{10.9}$$

the maximum principle follows upon choosing $u = M := \sup_{x \in \partial U} v(x)$.

Corollary 10.3 (Maximum principle)**.** *If $v \in H^1(U, \mathbb{R})$ is weakly subharmonic, then*

$$\sup_{x \in U} v(x) \leq \sup_{x \in \partial U} v(x), \tag{10.10}$$

with the supremum understood as an essential supremum and the boundary values understood as explained above.

Now what about strong solutions? To this end we regard (10.1) as an operator equation

$$Lu = f, \tag{10.11}$$

in $L^2(U)$ where

$$Lu := -\Delta u, \qquad u \in \mathfrak{D}(L) := H_0^1(U,\mathbb{R}) \cap H^2(U,\mathbb{R}). \tag{10.12}$$

On the other hand, weak solutions are associated with the operator equation

$$\bar{L}u = f, \tag{10.13}$$

where

$$\bar{L} := (JJ^*)^{-1}, \qquad \mathfrak{D}(\bar{L}) = \operatorname{Ran}(JJ^*). \tag{10.14}$$

Since every strong solution is also a weak solution and since weak solutions are unique, we see that $\bar{L}$ is an extension of L (in the sense that $\mathfrak{D}(L) \subseteq \mathfrak{D}(\bar{L})$ and both agree on $\mathfrak{D}(L)$). In fact, note that since JJ^* is self-adjoint, so is its inverse $\bar{L}$, which is also known as the **Friedrichs extension** of the Dirichlet Laplacian (see [**32**, Sect. 2.3]) in this context.

To shed some further light on $\mathfrak{D}(\bar{L})$ we observe that we trivially have $\mathfrak{D}(\bar{L}) = \operatorname{Ran}(JJ^*) \subset \operatorname{Ran}(J) = H_0^1(U)$. Moreover, if $u \in \mathfrak{D}(\bar{L})$ there is some $f \in L^2(U)$ such that $u = JJ^*f$, that is, such that (10.5) holds. Choosing $v \in C_c^\infty(U)$ in (10.5) shows that

$$-\int_U (\Delta\varphi) u \, d^n x = \int_U \varphi f \, d^n x, \qquad \varphi \in C_c^\infty(U), \tag{10.15}$$

which can be rephrased as

$$\mathfrak{D}(\bar{L}) = \{u \in H_0^1(U) | \Delta u \in L^2(U)\}, \tag{10.16}$$

where Δu is understood as a weak derivative (this does not mean that the second derivatives exist individually, it is only this particular combination of second derivatives which is required to exist). In this context $H_0^1(U)$ is known as the **form domain** of $\bar{L}$, written as $\mathfrak{Q}(\bar{L}) = H_0^1(U)$.

When we also have $\mathfrak{D}(\bar{L}) \subseteq \mathfrak{D}(L)$, that is, when every weak solution is also a strong solution, is a tricky question which we defer to the next section. However, we note the following local result:

Lemma 10.4. *Suppose $u \in H^1(U)$ is a weak solution of the Poisson equation* (10.3). *Then $u \in H^{r+2}_{loc}(U)$ whenever $f \in H^r_{loc}(U) \cap L^2(U)$.*

Proof. The idea is to reduce it to the case $U = \mathbb{R}^n$ by using a cutoff function. Indeed, given an arbitrary compact subset $K \subset U$, choose $\zeta \in C_c^\infty(U)$ such that $0 \le \zeta \le 1$ with $\zeta = 1$ on K. Now consider $v := \zeta u \in H_c^1(\mathbb{R}^n)$. Then using the product rule (Lemma 9.7 (iii)) and integration by parts one verifies

$$\begin{aligned}\int_{\mathbb{R}^n} (\nabla\varphi) \cdot (\nabla v) d^n x &= \int_{\mathbb{R}^n} \big((\nabla\zeta\varphi) \cdot (\nabla u) - \varphi(\nabla\zeta) \cdot (\nabla u) + u(\nabla\varphi) \cdot (\nabla\zeta)\big) d^n x \\ &= \int_{\mathbb{R}^n} \varphi h \, d^n x,\end{aligned}$$

for all $\varphi \in C_c^\infty(U)$, where

$$h := \zeta f - (\Delta\zeta)u - 2(\nabla\zeta)\cdot(\nabla u) \in L^2(\mathbb{R}^n).$$

Moreover, since both v and h vanish outside the support of ζ, this holds in fact for all $\varphi \in C_c^\infty(\mathbb{R}^n)$. Thus v is a weak solution on $\mathbb{R}^n$ implying $v \in H^2(\mathbb{R}^n)$ (cf. Problem 8.6). This establishes the case $r = 0$.

Now we can bootstrap this argument: If $f \in H^r$, then $h \in H^s$ with $s := \min(1, r)$ implying $v \in H^{s+2}(\mathbb{R}^n)$. Iterating this process shows $v \in H^{r+2}(\mathbb{R}^n)$. Finally, since $u = v$ on K and since K is arbitrary, we get the claimed result. □

Remark: It seems tempting to apply the same argument to a weak solution $u \in H_0^1(U)$ by setting it equal to 0 outside of U. Why does this fail?

Next we return to the observation that the embedding J is not only continuous, but even compact by the Rellich–Kondrachov theorem (Theorem 9.31). Hence we can apply the spectral theorem for compact operators (since JJ^* is self-adjoint, Theorem 3.7 from **[35]** will do; cf. also Theorem B.23):

Theorem 10.5. *Suppose $U \subseteq \mathbb{R}^n$ is a bounded domain. The operator $\bar{L}$ has a sequence of discrete real eigenvalues $0 < E_0 \le E_1 \le \cdots$ converging to ∞. The corresponding normalized eigenfunctions w_j are in $H_0^1(U) \cap C^\infty(U)$ and (can be chosen to) form an orthonormal basis for $L^2(U)$.*

Observe that the inverse of the lowest eigenvalue E_0^{-1} is the optimal constant for the Poincarè inequality. Moreover, the eigenvalues appear in this list according to their (geometric) multiplicity. In this context note that in our case the multiplicity of every eigenvalue is finite.

Moreover, note that the orthonormal eigenfunctions w_j provide a unitary map

$$W : L^2(U) \to \ell^2(\mathbb{N}_0), \qquad f \mapsto f_j := \langle w_j, f\rangle_2 \tag{10.17}$$

which diagonalizes $\bar{L}$ in the sense that $\bar{L}$ is mapped to the multiplication operator

$$W\bar{L}W^{-1}f_j = E_j f_j, \quad W\mathfrak{D}(\bar{L}) = \{f_j \in \ell^2(\mathbb{N}_0) | E_j f_j \in \ell^2(\mathbb{N}_0)\}. \tag{10.18}$$

To see this start with the corresponding formula $W\bar{L}^{-1}W^{-1}f_j = E_j^{-1}f_j$ for the inverse (which is a bounded operator) and then compute its range to get the domain of $\bar{L}$. In particular we have

$$\bar{L}u = \sum_{j\in\mathbb{N}_0} E_j u_j w_j, \qquad u \in \mathfrak{D}(\bar{L}). \tag{10.19}$$

Similarly one has

$$\langle v, u\rangle = \sum_{j\in\mathbb{N}_0} E_j v_j^* u_j, \qquad u, v \in H_0^1(U), \tag{10.20}$$

and

$$WH_0^1(U) = \{f_j \in \ell^2(\mathbb{N}_0) | \sqrt{E_j} f_j \in \ell^2(\mathbb{N}_0)\}. \tag{10.21}$$

Here the special case $u \in \mathfrak{D}(\bar{L})$ follows from $\langle v, u\rangle = \langle v, \bar{L}u\rangle_2$ and the general case since $\mathfrak{D}(\bar{L}) \subset H_0^1(U)$ is dense (by $\mathrm{Ran}(J^*)^\perp = \mathrm{Ker}(J) = \{0\}$). Choosing $v = u$ we get

$$\|u\|^2 = \sum_{j\in\mathbb{N}_0} E_j |u_j|^2 \geq E_0 \|u\|_2^2 \tag{10.22}$$

with equality for $u = w_0$. This can be rephrased as

$$E_0 = \min_{u\in H_0^1(U)\setminus\{0\}} \frac{\|u\|^2}{\|u\|_2^2} = \min_{u\in\mathfrak{D}(\bar{L})\setminus\{0\}} \frac{\langle u, \bar{L}u\rangle_2}{\|u\|_2^2}, \tag{10.23}$$

which is known as the **Rayleigh–Ritz method**.[1] In particular, any choice of trial function $u \in H_0^1(U)$ will give an upper bound for the lowest eigenvalue.

Lemma 10.6. *Suppose $U \subseteq \mathbb{R}^n$ is a bounded connected domain. The lowest eigenvalue E_0 of the operator $\bar{L}$ is simple and the corresponding eigenfunction w_0 can be chosen positive.*

Proof. Let w be an eigenfunction corresponding to E_0. Then $w_\pm \in H_0^1(U, \mathbb{R})$ and both give the minimum in the Rayleigh–Ritz method, $\|w_\pm\|^2 = E_0\|w_\pm\|_2^2$ (check this). Hence both are in the eigenspace corresponding to E_0 (which is finite dimensional) and hence $w_\pm \in \mathfrak{D}(\bar{L})$. In particular, they are smooth and by $-\Delta w_\pm = E_0 w_\pm \geq 0$ we can apply the strong minimum principle (Lemma 5.14 (ii)) to conclude that either $w_+ = 0$ or $w_+ > 0$. In the first case we have $w_- < 0$ and in the latter $w_- = 0$. Hence w can be chosen positive. Since two positive functions cannot be orthogonal, the lowest eigenvalue must be simple. □

Note that if you happen to have found a positive eigenfunction, it must correspond to the lowest eigenvalue since two positive functions cannot be orthogonal. That is, all higher eigenfunctions must change sign. In fact, a look at the case of a rectangle, Figure 3.7, or a disc, Figure 3.9, shows that the number of nodal domains is increasing. Here the **nodal domains** of an eigenfunction w are the connected components of $U \setminus \{x \in U | w(x) = 0\}$. In fact, the following result holds:

[1] Walther Ritz (1878–1909), Swiss theoretical physicist

Theorem 10.7 (Courant[2]'s Nodal Domain Theorem)**.** *Suppose $U \subseteq \mathbb{R}^n$ is a bounded connected domain. Then the j'th eigenfunction of $\bar{L}$ has at most j nodal domains.*

Proof. If w is an eigenfunction, $\bar{L}w = \lambda w$, which has m nodal domains U_j, we can look at the restrictions $w_j := w\chi_{U_j}$. Then this restriction satisfies $w_j \in H^1_0(U)$. To see this assume that (w.l.o.g.) w is positive on U_j and consider $w_\varepsilon := \max(w, \varepsilon) - \varepsilon \in H^1_c(U_j)$ (cf. Lemma 9.8). Then $w_\varepsilon \to w$ (again by Lemma 9.8) and hence $w_j \in H^1_0(U_j)$ as required. Moreover, w_j satisfies $\|\nabla w_j\|_2 = \lambda \|w_j\|_2$ and since these functions are orthogonal, a well-known principle from spectral theory (Theorem 4.12 from **[32]**) implies that there are at least m eigenvalues (counting multiplicity) below or equal to λ. □

Concerning solvability, our considerations imply that the equation

$$\begin{aligned} (-\Delta - \lambda)u(x) &= f(x), && x \in U, \\ u(x) &= 0, && x \in \partial U, \end{aligned} \tag{10.24}$$

has a unique solution whenever λ is not an eigenvalue, that is, when the homogenous problem has only the trivial solutions. This is an instance of the Fredholm alternative for compact operators. Moreover, in this case the solutions can be written by means of the **resolvent operator**

$$u = (\bar{L} - \lambda)^{-1} f = \sum_{j \in \mathbb{N}_0} \frac{f_j}{E_j - \lambda} w_j. \tag{10.25}$$

In addition, we have the estimate

$$\|u\|^2 = \sum_{j \in \mathbb{N}_0} \frac{E_j |f_j|^2}{|E_j - \lambda|^2} \le C_\lambda \|f\|_2^2. \tag{10.26}$$

Finally, we remark that the general Poisson problem

$$\begin{aligned} (-\Delta - \lambda)u(x) &= f(x), && x \in U, \\ u(x) &= g(x), && x \in \partial U, \end{aligned} \tag{10.27}$$

can be reduced to the case $g = 0$ whenever g is the trace of some function $\bar{g} \in H^2(U)$ by considering $v := u - \bar{g}$. It is not obvious for which g such an extension exists. However, if the boundary is C^2, then a sufficient condition for such an extension to exist is $g \in C^2(\partial U)$ (Problem 10.11).

Moreover, for the weak formulation it would suffice if $f \in H^1_0(U)^*$ and in this setting it also suffices to assume that g is the trace of some function $\bar{g} \in H^1(U)$. For this last condition there is an easy criterion: If g is Lipschitz continuous, it has a Lipschitz continuous extension to $\bar{U}$ (Lemma B.29 from

[2] Richard Courant (1888–1972), German American mathematician

[**35**]). Thus $\bar{g} \in W^{1,\infty}(U) \subset H^1(U)$ as required. Note, that in such a situation, where we can extend the boundary values g to a function $\bar{g} \in H^1(U)$, it is not even necessary to assume that the domain admits a trace operator since we could define the boundary condition $u = g$ on ∂U via $u - \bar{g} \in H^1_0(U)$. In this sense, the possible boundary data are $H^1(U) \setminus H^1_0(U)$.

For example, this establishes solvability of the classical Dirichlet problem ($\lambda = 0$ and $f = 0$) for Lipschitz continuous g. However, it is neither guaranteed that the weak solution is continuous up to the boundary (like Theorem 5.29 does) nor that it attains the boundary values everywhere (cf. Problem 10.7).

The corresponding Neumann problem

$$\begin{aligned} -\Delta u(x) &= f(x), && x \in U, \\ \frac{\partial u}{\partial \nu}(x) &= 0, && x \in \partial U, \end{aligned} \tag{10.28}$$

can be handled as follows: We look at the weak formulation

$$\langle v, u \rangle = \langle v, f \rangle_2, \qquad v \in H^1(U), \tag{10.29}$$

where now we do not restrict the functions at the boundary at all, that is, we replace H^1_0 by H^1. Now we run into the problem that $\|.\|$ is no longer a norm on H^1 since the Poincaré inequality fails on H^1. This is not surprising, since it just reflects the fact, that the Neumann problem has the eigenvalue 0. The remedy is to take the orthogonal complement with respect to the corresponding eigenfunction

$$\mathfrak{H}_1 := \{ f \in H^1(U) | \langle 1, f \rangle_{1,2} = 0 \}. \tag{10.30}$$

Then, if we assume U to be connected, the Poincaré inequality continues to hold and we can proceed as before. Of course we also need to assume $f \in \mathfrak{H}_1$, that is,

$$\int_U f(x) d^n x = 0, \tag{10.31}$$

the necessary solvability condition we have already found in (5.43). Of course the only remaining question is to understand that this solution is indeed the solution to the Neumann problem. To understand this we need to look at the domain of the self-adjoint operator L_N associated with our weak formulation in $L^2(U)$. More precisely, we will consider

$$\mathfrak{H}_0 := \{ f \in L^2(U) | \langle 1, f \rangle_2 = 0 \} \tag{10.32}$$

such that we have a natural embedding $J_N : \mathfrak{H}_1 \hookrightarrow \mathfrak{H}_0$ and look at $\bar{L}_N := (J_N J_N^*)^{-1}$. To understand its domain note that, as in the Dirichlet case, we see that functions $u \in \mathfrak{D}(\bar{L}_N)$ must satisfy $u \in \mathfrak{H}_1$ as well as $\Delta u \in L^2(U)$. However, since the weak formulation must hold for all $v \in H^1(U)$ (and not only for all $v \in H^1_0(U)$) this is not everything! Indeed, let U have a nice boundary (the usual C^1 condition), assume $u \in H^2(U)$, and use integration by parts to see

$$\int_{\partial U} v \frac{\partial u}{\partial \nu} dS = \int_U (\Delta u + f) v \, d^n x = 0. \tag{10.33}$$

Of course this is expected to hold precisely if $\frac{\partial u}{\partial \nu}$ vanishes in the sense of traces. Indeed, for a nice domain (with a C^2 boundary) it is not hard to see that a function which is C^2 on the boundary can be extended to a function which is in $C^2_b(\overline{U})$ (Problem 10.11). Since $C^2(\partial U)$ is dense in $L^2(\partial U)$ the claim follows.

Finally, note that we can extend $\bar{L}_N$ to all of $L^2(U)$ by setting $\bar{L}_N 1 := 0$.

Theorem 10.8. *Suppose $U \subseteq \mathbb{R}^n$ is a bounded domain. The operator $\bar{L}_N$ has a sequence of discrete real eigenvalues $0 = E_0 < E_1 \le \cdots$ converging to ∞. The corresponding normalized eigenfunctions w_j are in $H^1(U) \cap C^\infty(U)$ and (can be chosen to) form an orthonormal basis for $L^2(U)$. In particular, $w_0 = 1$.*

Of course, inhomogeneous Neumann conditions can be handled analogous to the case of inhomogeneous Dirichlet conditions.

Problem 10.1. *Compute J^* for $U := (0,1) \subset \mathbb{R}$.*

Problem 10.2. *Show that the eigenfunctions w_j are orthogonal in $H^1_0(U)$. Find the correct normalization.*

Problem 10.3. *Investigate the **Helmholtz equation***

$$\begin{aligned} -\Delta u(x) + u(x) &= f(x), & x &\in U, \\ u(x) &= 0, & x &\in \partial U, \end{aligned}$$

on a domain $U \subseteq \mathbb{R}^n$. (Note, that U is not required to be bounded.)

Problem 10.4. *Show that a weakly subharmonic function satisfies the submean property*

$$v(x) \le \frac{1}{|B_r|} \int_{B_r(x)} v(y) d^n y$$

for every ball with sufficiently small radius $r \le r_0(x) < \mathrm{dist}(x, \partial U)$ at every Lebesgue point of v. Hence if v is continuous, it is subharmonic as defined in Section 5.2. Moreover, conclude that every weakly harmonic function is harmonic. (Hint: Mollify v and use Problem 5.11. For the last claim look at the proof of Lemma 5.2.)

Problem 10.5 (Strong maximum principle)**.** *Show: If U is connected and a subharmonic function v on U assumes an interior maximum at an Lebesgue point, then v is constant. (Hint: Use the submean property from the previous problem.)*

Problem 10.6. *Find a weak formulation of the Poisson problem with Robin boundary conditions*

$$\begin{aligned} -\Delta u(x) + \lambda u(x) &= f(x), & x &\in U, \\ \frac{\partial u}{\partial \nu}(x) + a(x)u(x) &= 0, & x &\in \partial U, \end{aligned}$$

on a bounded domain $U \subseteq \mathbb{R}^n$ with a C^1 boundary. Here $a \in L^\infty(U,\mathbb{R})$. Establish existence of weak solutions for $\lambda > E_0$. Show that if $a \ge 0$ is nonzero and U is bounded and connected, then all eigenvalues of the Laplacian with Robin boundary conditions are positive. (Hint: Green's first identity.)

Problem 10.7. *Consider the Dirichlet problem $-\Delta u = 0$ on the punctured disc $U := B_1(0) \setminus \{0\} \subset \mathbb{R}^n$ with boundary data $g(x) = 0$ for $|x| = 1$ and $g(0) = 1$. Since this domain does not have a trace operator, we understand the boundary condition as $u - \bar{g} \in H_0^1(U)$, where $\bar{g} = 1 - |x|^2$. Find the corresponding weak solution. (Hint: Observe that the weak solution must be radial. In particular, you are looking for a radial harmonic function satisfying the boundary conditions.)*

Problem 10.8. *Consider the punctured disc $B_1(0) \setminus \{0\}$. Show $H_0^1(U) = H_0^1(B_1(0))$ as well as $H^1(U) = H^1(B_1(0))$ for $n \ge 2$. (Hint: Use the previous problem.)*

Problem* 10.9. *Suppose $f \le 0$ on ∂U (in the sense that $f_+ \in H_0^1(U)$) and $g \ge 0$ a.e. in U. Show that $f \le g$ on ∂U. Conclude that if $M := \sup_{\partial U} f$ is finite, then $f \le m$ on ∂U for all $m \in [M, \infty)$.*

Problem 10.10. *Consider a function $F : \mathbb{R} \to \mathbb{R}$ such that $|F(t)| \le |t|^3$ for all $t \in \mathbb{R}$ and let $U \subset \mathbb{R}^3$ be a bounded domain with the extension property. Prove that if $u \in H^1(U)$ is a weak solution of the nonlinear Poisson equation $-\Delta u = F(u)$, then in fact we have $u \in H^2_{loc}(U)$. (Hint: Corollary 9.23.)*

Problem 10.11. *Let $U \subset \mathbb{R}^n$ be an open set with a bounded C^k boundary. Show that a function $f \in C^k(\partial U)$ has an extension $\bar{f} \in C_b^k(\overline{U})$ such that $\bar{f}|_{\partial U} = f$. (Hint: Reduce it to the case of a flat boundary.)*

Problem 10.12. *Let w be an eigenfunction, $\bar{L}w = \lambda w$, and U_0 a nodal domain. Show that $\lambda = E_0(U_0)$ is the lowest Dirichlet eigenvalue of U_0.*

10.2. Elliptic equations

The main strength of the methods from the previous section is that, in contradistinction to methods based on explicit representation formulas for solutions, they can be easily extended to general elliptic operators in divergence form

$$Lu(x) := -\sum_{i,j=1}^{n} \partial_i A_{ij}(x)\partial_j u(x) + \sum_{j=1}^{n} b_j(x)\partial_j u(x) + c(x)u(x), \qquad (10.34)$$

where $A_{ij}, b_j, c \in L^\infty(U,\mathbb{R})$. We will assume that the matrix A_{ij} is symmetric,

$$A_{ij} = A_{ji}. \qquad (10.35)$$

In case of a differentiable matrix, this can be done without loss of generality. We will only look at the real case for notational simplicity and leave the straightforward modifications for the complex case as an exercise (Problem 10.14).

In addition we require L to be uniformly elliptic, that is

$$\xi \cdot A(x)\xi = \sum_{i,j=1}^{n} A_{ij}(x)\xi_i\xi_j \ge \theta|\xi|^2 \qquad (10.36)$$

for a.e. $x \in U$ and all $\xi \in \mathbb{R}^n$. As domain for L we choose

$$\mathfrak{D}(L) = \{u \in H_0^1(U,\mathbb{R}) | A_{ij}\partial_j u \in H^1(U,\mathbb{R}), 1 \le i,j \le n\}. \qquad (10.37)$$

As in the case of the Laplacian, a weak solution of the elliptic problem

$$\begin{aligned} Lu(x) &= f(x), & x &\in U, \\ u(x) &= 0, & x &\in \partial U, \end{aligned} \qquad (10.38)$$

is a function $u \in H_0^1(U)$ satisfying

$$a(v,u) = \langle v, f\rangle_2, \qquad v \in H_0^1(U), \qquad (10.39)$$

where

$$a(v,u) = \int_U \Big(\sum_{i,j} A_{ij}(\partial_i v)(\partial_j u) + \sum_j b_j v \partial_j u + c\, v\, u\Big) d^n x. \qquad (10.40)$$

However, note that now a is no longer symmetric (unless $b \equiv 0$) and hence cannot be used as a scalar product. Nevertheless, a will still be the form of a bounded operator (since we have assumed the coefficients to be bounded) and this operator will be invertible since the form is coercive (here is where the ellipticity is used). In fact, this is the content of the famous **Lax–Milgram theorem**[3]:

Theorem 10.9 (Lax–Milgram). *Let $a : \mathfrak{H} \times \mathfrak{H} \to \mathbb{R}$ be a bilinear form on a Hilbert space $\mathfrak{H}$ which is*

- *bounded, $|a(v,u)| \le C\|v\|\,\|u\|$, and*
- *coercive, $a(u,u) \ge \varepsilon\|u\|^2$ for some $\varepsilon > 0$.*

[3] Peter Lax (*1926), American mathematician of Hungarian origin

[3] Arthur Milgram (1912–1961), American mathematician

Then for every $f \in \mathfrak{H}$ there is a unique $u \in \mathfrak{H}$ such that

$$a(v,u) = \langle v, f\rangle, \qquad \forall v \in \mathfrak{H}. \tag{10.41}$$

Moreover, $\|u\| \le \frac{1}{\varepsilon}\|f\|$.

Proof. See Theorem 2.17 from [**35**]. □

The boundedness assumption in the Lax–Milgram theorem implies (by virtue of the Riesz representation theorem for Hilbert spaces) that there is a bounded operator $A \in \mathscr{L}(\mathfrak{H})$ such that $a(v,u) = \langle v, Au\rangle$ and the coercivity assumptions implies that A has a bounded inverse $A^{-1} \in \mathscr{L}(\mathfrak{H})$.

Note that if a is symmetric, then a can be taken as a new scalar product and the conditions imply that the associated norm is equivalent to the original norm. This is precisely what we did in the previous section. Moreover, in this case the solution can also be obtained via an abstract Dirichlet principle just as we did for the Laplace operator in Section 5.5 (Problem 10.13; see also Problem 5.36) As pointed out before, this road is closed unless we make the additional assumption $b \equiv 0$.

Now we are ready to apply this to our elliptic problem. Let us abbreviate

$$a_0 := \max_{ij} \|A_{ij}\|_\infty, \quad b_0 := \max_j \|b_j\|_\infty, \quad c_0 := \|c\|_\infty. \tag{10.42}$$

By a simple use of the Cauchy–Schwarz inequality we see that the bilinear form $a(u,v)$ is bounded:

$$|a(v,u)| \le a_0\|v\|\|u\| + b_0\|v\|_2\|u\| + c_0\|v\|_2\|u\|_2 \le C\|v\|\|u\|. \tag{10.43}$$

To see coercivity, $a(u,u) \ge c\|u\|^2$, we begin with

$$a(u,u) \ge \int_U \Big(\theta|\nabla u|^2 - b_0|u||\nabla u| + c_1|u|^2\Big) d^n x, \tag{10.44}$$

where

$$c_1 := \inf_{x\in U} c(x). \tag{10.45}$$

Here the inf is understood as an essential infimum, that is, the smallest constant such that $c(x) \ge c_1$ for a.e $x \in U$.

Now we distribute the middle term by means of the elementary inequality

$$|u||\nabla u| \le \frac{\varepsilon}{2}|u|^2 + \frac{1}{2\varepsilon}|\nabla u|^2 \tag{10.46}$$

which gives

$$a(u,u) \ge \int_U \Big((\theta - \frac{b_0}{2\varepsilon})|\nabla u|^2 + (c_1 - \frac{\varepsilon b_0}{2})|u|^2\Big) d^n x. \tag{10.47}$$

This estimate is sometimes known as **Gårding inequality**[4]. To apply the Lax–Milgram theorem we need $\theta - \frac{b_0}{2\varepsilon} > 0$ and $c_1 - \frac{\varepsilon b_0}{2} > 0$. Solving for ε this

[4]Lars Gårding (1919–2014), Swedish mathematician

leads to the conditions $\frac{b_0}{2\theta} < \varepsilon < \frac{2c_1}{b_0}$ and hence we need $4\theta c_1 > b_0^2$. In the case of a bounded domain the norm of the gradient suffices by the Poincaré inequality and we can also admit the borderline case $b_0 = c_1 = 0$.

Theorem 10.10. *Let $U \subseteq \mathbb{R}^n$ be open and L a uniformly elliptic operator* (10.34) *with associated constants θ, b_0, c_1 as defined in* (10.36), (10.42), (10.45), *respectively. Then the elliptic problem* (10.38) *has a unique weak solution in $H_0^1(U,\mathbb{R})$ for every $f \in L^2(U,\mathbb{R})$ provided*

$$4\theta c_1 > b_0^2. \tag{10.48}$$

If U is bounded the claim also holds if $b_0 = c_1 = 0$. Moreover, there is a constant $C > 0$ such that $\|u\| \le C\|f\|_2$.

Proof. Everything follows from the above analysis. For a general domain we take the usual Sobolev norm $\|u\|^2 := \|\nabla u\|_2^2 + \|u\|_2^2$ whereas in the case of a bounded domain we use $\|u\|^2 := \|\nabla u\|_2^2$. □

Of course there is also a corresponding comparison principle. To this end call a function $v \in H^1(U,\mathbb{R})$ a **weak subsolution** if

$$a(\varphi, v) \le 0, \qquad \varphi \in C_c^\infty(U,[0,\infty)). \tag{10.49}$$

Then the analog of Lemma 10.2 reads:

Lemma 10.11. *Suppose $c \ge 0$ and let U be a bounded domain. Let $u, v \in H^1(U,\mathbb{R})$ with u a weak solution and v a weak subsolution. Then $v \le u$ on ∂U implies $v \le u$ on U.*

Proof. We proceed as in the proof of Lemma 10.2. Assume $u = 0$ without loss of generality and consider $v_+ \in H_0^1(U,[0,\infty))$. By approximation $a(\varphi, v) \le 0$ holds for all $\varphi \in H_0^1(U,\mathbb{R})$ and choosing $\varphi = v_+$ we get $0 \ge a(v_+, v) = a(v_+, v_+)$ and we can use Gårding inequality (10.47) to conclude $v_+ = 0$ as in the proof of Lemma 10.2. However, this argument requires the restriction $4\theta c_1 > b_0^2$ or $b_0 = c_1 = 0$ (boundedness of U is not needed in this case).

The following argument avoids this restriction on b_0: For any real number m such that $\sup_{\partial U} v \le m < \sup_U v$ (if there is no such number, there is nothing to do) we have $w := (v - m)_+ \in H_0^1(U)$ by Problem 10.9 and we can repeat the calculation form above to obtain

$$\int_U \theta\Big(|\nabla w|^2 - b_0 w|\nabla w|\Big) d^n x \le -\int_U c\, w^2 d^n x \le 0.$$

Here we have used that $\nabla w(x) = \nabla v(x)$ for $v(x) > k$ and $\nabla w(x) = 0$ otherwise as well as $wv \ge w^2$. Using Cauchy–Schwarz this implies $\|\nabla w\|_2 \le$

$\frac{b_0}{\theta}\|w\|_{L^2(S)}$, where $S := \operatorname{supp}(\nabla w)$. And using Theorem 9.22 as well as the generalized Hölder inequality (Problem B.12) we get

$$\|w\|_p \le C\|\nabla w\|_2 \le \frac{b_0 C}{\theta}\|w\|_{L^2(S)} \le \frac{b_0 C}{\theta}|S|^{1/n}\|w\|_p,$$

where $p := \frac{2n}{n-1}$ and we have assumed $n \ge 3$. Now since w cannot vanish, we conclude $|S| \ge \varepsilon > 0$ where ε does not depend on m. Hence, letting $m \to \sup_U v$ we see that the gradient of v does not vanish on the set where v attains its supremum, contradicting Lemma 9.9. This establishes the claim in the case $n \ge 3$.

In the case $n = 2$ we use (9.21) with $p = \gamma = 2$ to obtain $\|w\|_4^2 \le C\|w\|_2\|\nabla w\|_2 \le C|U|^{1/4}\|w\|_4\|\nabla w\|_2$, where we have again invoked the generalized Hölder inequality to obtain the second inequality. Hence we have $\|w\|_4^2 \le \tilde{C}\|\nabla w\|_2$ provided U is bounded and we can proceed as in the case $n \ge 3$. In the case $n = 1$ we use the inequality from Problem 9.24. □

Corollary 10.12 (Maximum principle)**.** *Suppose $c \ge 0$ and let U be a bounded domain. If $v \in H^1(U,\mathbb{R})$ is a weak subsolution, then*

$$\sup_{x\in U} v(x) \le \sup_{x\in\partial U} v_+(x). \tag{10.50}$$

Proof. Note that a nonegative constant is a supersolution and hence we can apply the previous lemma with $u = 0$ and $v - M$, where $M := \sup_{x\in\partial U} v_+(x)$. □

In particular, there is an operator $A \in \mathscr{L}(H_0^1(U,\mathbb{R}))$ such that $a(v,u) = \langle v, Au\rangle$. Note that A will be self-adjoint if and only if a is symmetric, that is, if $b = 0$. In the case $L = -\Delta$ we have $A = \mathbb{I}$. Consequently, the solution of the weak problem is $u = A^{-1}J^* f$. Moreover, introducing the operator

$$\bar{L} := (JA^{-1}J^*)^{-1}, \quad \mathfrak{D}(\bar{L}) := \operatorname{Ran}(JA^{-1}J^*) \subset H_0^1(U,\mathbb{R}), \tag{10.51}$$

we see that for every $f \in L^2(U,\mathbb{R})$, there is a unique solution $u \in \mathfrak{D}(\bar{L})$ of the operator equation

$$\bar{L}u = f \tag{10.52}$$

such that Ju is the unique solution of our weak problem. Furthermore, the inverse $\bar{L}^{-1} = JA^{-1}J^*$ is compact if U is bounded. As in the case of the Laplacian, we have

$$\mathfrak{D}(\bar{L}) = \{u \in H_0^1(U,\mathbb{R}) | \sum_{i,j} \partial_i A_{ij}(x)\partial_j u(x) \in L^2(U,\mathbb{R})\}, \tag{10.53}$$

where $\sum_{i,j} \partial_i A_{ij}(x)\partial_j u(x)$ is understood as a weak derivative (this does not mean that the second derivatives exist individually, it is only this particular combination of second derivatives which is required to exist). Moreover, note that the restriction (10.48) is not required for these considerations since we can replace L by $L+\gamma$ with some sufficiently large γ such that $4\theta(c_1+\gamma) > b_0^2$. Again we can apply the spectral theorem for compact operators to $(\bar{L}+\gamma)^{-1}$ to obtain:

Theorem 10.13. *Suppose $U \subseteq \mathbb{R}^n$ is a bounded domain. The uniformly elliptic operator $\bar{L}$ has a sequence of discrete eigenvalues E_n converging to ∞. If $b \equiv 0$, then $\bar{L}$ is self-adjoint and the eigenvalues are bounded from below and can be ordered according to*

$$E_0 \le E_1 \le \dots \tag{10.54}$$

The corresponding normalized eigenfunctions w_j (can be chosen to) form an orthonormal basis for $L^2(U,\mathbb{R})$.

Note that in the one-dimensional case our elliptic problem is known as **Sturm–Liouville problem** and the above theorem establishes the result mentioned in Section 3.2 at least for Dirichlet boundary conditions. Neumann or Robin boundary conditions can be handled as outlined in the previous section.

Of course the spectral theorem also implies that, whenever λ is not an eigenvalue, then the equation

$$(\bar{L} - \lambda)u = f \tag{10.55}$$

has a unique solution $u = (\bar{L}-\lambda)^{-1}f$. Moreover, one can even get a solution in case λ is an eigenvalue provided $f \in \operatorname{Ker}(\bar{L}^* - \lambda)^\perp$, where $\bar{L}^*$ is the adjoint operator (computed from the adjoint form $a^*(v,u) := a(u,v)$). This is known as the **Fredholm alternative** (Theorem 7.6 from [**35**]) and we have already encountered this fact in the previous section when looking at the Neumann problem for the Laplacian. Note that the maximum principle shows that 0 cannot be an eigenvalue if $c \ge 0$.

Corollary 10.14. *Let U be bounded, then the condition* (10.48) *in Theorem 10.10 can be replaced by $c \ge 0$.*

Clearly, in the self-adjoint case the orthonormal eigenfunctions w_j provide a unitary map

$$W : L^2(U,\mathbb{R}) \to \ell^2(\mathbb{N}_0), \qquad f \mapsto f_j := \langle w_j, f\rangle_2 \tag{10.56}$$

which diagonalizes $\bar{L}$ and everything from the previous section applies verbatim. In particular, the Rayleigh–Ritz method reads

$$E_0 = \min_{u \in H_0^1(U)\setminus\{0\}} \frac{a(u,u)}{\|u\|_2^2} = \min_{u \in \mathfrak{D}(\bar{L})\setminus\{0\}} \frac{\langle u, \bar{L}u\rangle_2}{\|u\|_2^2}. \tag{10.57}$$

Note that this shows $E_0 > c_1$, where the inequality is strict since 0 cannot be an eigenvalue of $\bar{L} - c_1$ as already observed above.

Also Lemma 10.6 extends to this situation almost verbatim. The only difference is that we need to assume that the coefficients of L are sufficiently smooth such that we can apply the classical strong maximum principle.

Lemma 10.15. *Suppose $U \subseteq \mathbb{R}^n$ is a bounded connected domain and $A \in W^{k+1,\infty}$, $b = 0$, $c \in W^{k,\infty}$ with $k > \frac{n}{2}$. Then the lowest eigenvalue E_0 of the operator $\bar{L}$ is simple and the corresponding eigenfunction w_0 can be chosen positive.*

Proof. By Corollary 10.17 below, the eigenfunctions are in $H^{k+2}_{loc}(U,\mathbb{R}) \subset C^2(U)$. Moreover, considering $\bar{L} - c_1$ we can assume $c \geq 0$ without loss of generality. Hence we can literally follow the proof of Lemma 10.6 now using the strong maximum principle from Theorem 5.37. □

If we assume that a is coercive then we can choose $\langle v, u\rangle := a(v,u)$ as the scalar product for $H^1_0(U)$ and it is interesting to observe that w_j are also orthogonal with respect to this scalar product. Indeed, we have

$$\langle w_j, w_k\rangle = a(w_j, w_k) = \langle w_j, \bar{L} w_k\rangle = E_j^2\, \delta_{jk}. \tag{10.58}$$

In particular, we can write the solution with the help of the eigenfunctions as

$$u = \sum_{j\in\mathbb{N}_0} \frac{\langle w_j, u\rangle}{\|w_j\|^2} w_j = \sum_{j\in\mathbb{N}_0} \frac{\langle w_j, f\rangle_2}{E_j^2} w_j \tag{10.59}$$

which would give a constructive formula for computing the solution if we knew the eigenfunctions. However, note that if we take an arbitrary orthogonal basis φ_j with respect to our adapted scalar product, we can still write

$$u = \sum_{j\in\mathbb{N}_0} \frac{\langle \varphi_j, u\rangle}{\|\varphi_j\|^2} \varphi_j = \sum_{j\in\mathbb{N}_0} \frac{\langle \varphi_j, f\rangle_2}{a(\varphi_j, \varphi_j)} \varphi_j, \tag{10.60}$$

where we have used that (10.39) implies $\langle \varphi_j, u\rangle = a(\varphi_j, u) = \langle \varphi_j, f\rangle_2$. Consequently, one way of computing the solution is to compute an orthogonal basis (e.g. using the Gram–Schmidt[5] procedure). However, computing an orthogonal basis is still too time consuming from a practical point of view and hence one drops the requirement that the φ_j are orthonormal. Moreover, when implementing this idea only finitely many, say $\varphi_1, \dots, \varphi_N$, are chosen and we look for the projection u_N of the solution onto the linear span of the these functions. Note that this projection will be the solution of our elliptic problem restricted to this subspace. Writing

$$u_N = \sum_{j=1}^{N} \alpha_j \varphi_j \tag{10.61}$$

[5] Jørgen Pedersen Gram (1850–1916), Danish actuary and mathematician

we see that the unknown coefficients α_j can be determined by solving the linear system

$$\sum_{j=1}^{N} a(\varphi_k, \varphi_j)\alpha_j = \langle \varphi_k, f\rangle. \tag{10.62}$$

This system will have a unique solution provided the vectors φ_j are linearly independent. By construction we have $u_n \to u$ as $n \to \infty$. Such a scheme of approximating the solution by solving a finite dimensional system is known as **Galerkin method**.[6] For example, in two dimensions one assumes that U is a polygon, which is split into small triangles. For the functions φ_j one chooses functions which are piecewise linear such that φ_j equals one on the j'th inner vertex of the triangulation and vanishes on all other vertices. This has the advantage that $a(\varphi_k, \varphi_j)$ can be easily computed and will vanish unless both vertices are either neighbors or equal. Hence the coefficient matrix of the above linear system will be sparse and hence it can be solved effectively. This is known as **finite element method**.

Problem 10.13 (Abstract Dirichlet principe)**.** *Let a be a coercive and symmetric bilinear form. Show that the solution of* (10.41) *is also the unique minimizer of*

$$v \mapsto \frac{1}{2}a(v,v) - \langle v, f\rangle.$$

(Hint: Compare with Section 5.5.)

Problem 10.14. *Extend the results from this section to the case of complex equations. (Hint: You will need the complex version of Lax–Milgram, see Theorem 2.17 from* [**35**]*.)*

Problem 10.15. *In quantum mechanics the* ***Schrödinger operator***

$$Lu := -\Delta u + cu$$

is a central object. However, in this case the assumption $c \in L^\infty$ is too restrictive. Show that the Schrödinger equation $(L+\gamma)u = f$ has a weak solution in $H^1_0(U)$ for every $f \in L^2(U)$ and

$$\gamma > -\inf_{v \in H^1_0(U), \|v\|_{H^1}=1} \mathrm{Re}(\langle v, cv\rangle_2)$$

provided

$$c \in L^q(U), \qquad \begin{cases} \frac{n}{2} \le q \le \infty, & n > 2, \\ 1 < q \le \infty, & n = 2, \\ 1 \le q \le \infty, & n = 1. \end{cases}$$

(Hint: Look at the sesquilinear form $\langle v, cu\rangle_2$ on H^1_0. When is it bounded?)

[6] Boris Grigoryevich Galerkin (1871–1945), Soviet mathematician and an engineer

Problem 10.16. *Extend the results from this section to the case of a magnetic **Schrödinger operator***

$$Lu := -(\nabla - \mathrm{i}b)^2 u + cu = -\sum_j (\partial_j - \mathrm{i}b_j)(\partial_j + \mathrm{i}b_j)u + cu$$

with $b_j \in L^\infty(U, \mathbb{R})$ and $c \in L^\infty(U, \mathbb{R})$.

10.3. Elliptic regularity

Finally we turn to regularity of weak solutions of the elliptic problem (10.38). For our first result we will not need the solution to vanish at the boundary. In this context we will call a function $u \in H^1(U)$ a weak solution of the the elliptic problem $\bar{L}u = f$ provided it satisfies (10.39).

Lemma 10.16. *Suppose $A_{ij} \in W^{1,\infty}(U)$. If $u \in H^1(U)$ is a weak solution of the elliptic problem $\bar{L}u = f$ for $f \in L^2(U)$, then $u \in H^2_{loc}(U)$. Moreover, for every $V \subset\subset U$ we have an estimate of the form $\|u\|_{H^2(V)} \le C(\|f\|_{L^2(U)} + \|u\|_{L^2(U)})$.*

Proof. The idea is to use the weak formulation $a(v, u) = \langle f, v\rangle_2$. If we formally choose $v = \partial_l^2 u$ we can perform an integration by parts (neglecting boundary terms) such that we can use the ellipticity condition to get an a priori estimate for the second derivatives. To make this idea work we will use a cutoff function and replace derivatives by finite differences (9.10) (invoking Lemma 9.10).

So the first step is to choose some function $\zeta \in C_c^\infty(U)$ and observe that $\tilde{u} := \zeta u$ solves a modified problem

$$\tilde{a}(v, \tilde{u}) = \langle v, \tilde{f}\rangle_2,$$

where $\tilde{A} = A$, $\tilde{b} = 0$, $\tilde{c} = 0$, and

$$\tilde{f} = \zeta\big(f - \sum_j b_j \partial_j u - cu\big) - \sum_{i,j} \big((\partial_i A_{ij} \partial_j \zeta)u + 2(\partial_i \zeta) A_{ij} \partial_j u\big).$$

Since $\tilde{f} \in L^2(U)$ we have reduced it to the case where u has compact support and $b = c = 0$ (and we will drop the tildes for notational simplicity).

Hence suppose $u \in H^1_c(U)$ and choose $v := D_l^{-\varepsilon} D_l^\varepsilon u$ (with ε sufficiently small such that $v \in H^1_c(U)$) in the weak formulation to obtain (cf. Problem 10.17)

$$\begin{aligned}\langle D_l^{-\varepsilon} D_l^\varepsilon u, f\rangle_2 = a(D_l^{-\varepsilon} D_l^\varepsilon u, u) &= \int_U \sum_{i,j} A_{ij} (\partial_i u)(D_l^{-\varepsilon} D_l^\varepsilon \partial_j u) d^n x \\ &= -\int_U \sum_{i,j} (D_l^\varepsilon A_{ij})(T_{\varepsilon\delta^l} \partial_i u)(D_l^\varepsilon \partial_j u) d^n x - a(D_l^\varepsilon u, D_l^\varepsilon u).\end{aligned}$$

Using ellipticity, $\|D_l^\varepsilon A_{ij}\|_\infty \le C$, and Cauchy–Schwarz we obtain

$$\theta\|D_l^\varepsilon \nabla u\|_2^2 \le C\|\nabla u\|_2 \|D_l^\varepsilon \nabla u\|_2 + \|f\|_2 \|D_l^{-\varepsilon} D_l^\varepsilon u\|_2.$$

Invoking Lemma 9.10 we have $\|D_l^{-\varepsilon} D_l^\varepsilon u\|_2 \le \|D_l^\varepsilon \partial_l u\|_2 \le \|D_l^\varepsilon \nabla u\|_2$ implying

$$\theta\|D_l^\varepsilon \nabla u\|_2 \le C\|\nabla u\|_2 + \|f\|_2.$$

Hence invoking again Lemma 9.10 shows $u \in H^2(U)$.

This also establishes an estimate of the form $\|u\|_{H^2(V)} \le C(\|f\|_{L^2(U)} + \|u\|_{H^1(U)})$ for every $V \subset\subset U$ for our original u. The slightly stronger estimate claimed is left as Problem 10.19. □

Observe that the appearance of the L^2 norm of u in the estimate reflects the fact that we do not require u to vanish on the boundary. Indeed, we can add a constant to u without changing f. Also note that once we know $u \in H^2_{loc}$, we can use integration by parts in the weak formulation whenever $v \in C_c^\infty(U)$ to conclude that the differential equation holds (a.e.) provided the derivatives are understood as weak derivatives.

Moreover, by formally applying ∂_l to the partial differential equation we get an elliptic partial differential for $\partial_l u$ (in particular, the second order coefficient A will not change) provided we assume sufficient smoothness for the coefficients. Of course this also works for the weak formulation (where we need one derivative less and the formal arguments gets rigorous) and hence we can apply the previous result recursively to obtain:

Corollary 10.17. *Suppose $A_{ij} \in W^{k+1,\infty}(U)$, $b_j \in W^{k,\infty}(U)$, and $c \in W^{k,\infty}(U)$. If $u \in H^1(U)$ is a weak solution of the elliptic problem $\bar{L}u = f$ for $f \in H^k(U)$, then $u \in H^{k+2}_{loc}(U)$.*

Proof. As in the proof of the previous lemma, we can reduce it to the case where u has compact support by introducing a cutoff function. In particular, by our previous lemma, $u \in H^2_c(U)$ in this case. Moreover, using the product rule and integration by parts one verifies

$$a(v, \partial_l u) = \langle v, \tilde{f}\rangle, \qquad \tilde{f} = \partial_l\big(f - \sum_j b_j \partial_j u - cu\big) - \sum_{i,j} \partial_i(\partial_l A)\partial_j u.$$

Hence the claim follows by induction on k. □

In order to get regularity up to the boundary we need to add extra assumptions on the domain as the following example shows.

Example 10.1. Let U be a domain in $\mathbb{R}^2$ which is given by a sector of angle $\beta \in (0, 2\pi)$, in polar coordinates $U := \{(r\cos(\varphi), r\sin(\varphi)) | 0 < r < 1,\ 0 < \varphi < \beta\}$. Then, it is straightforward to check that $u(x,y) = r^{\pi/\beta}\sin(\frac{\pi\varphi}{\beta})$, where (r,φ) are the polar coordinates of (x,y), is harmonic on U. Moreover,

$u \in H^1(U)$ but $u \notin H^2(U)$ for $\beta > \pi$. One can even multiply u with a smooth radial cutoff function to obtain a smooth solution of the Poisson equation which vanishes on ∂U but exhibits the same behavior. $\diamond$

Lemma 10.18. *Suppose $A_{ij} \in W^{1,\infty}(U)$ and U has a bounded $C^{1,1}$ boundary. If $u \in H_0^1(U)$ is a weak solution of the elliptic problem $\bar{L}u = f$ for $f \in L^2(U)$, then $u \in H^2(U)$. Moreover, we have an estimate of the form $\|u\|_{H^2} \le C(\|f\|_{L^2} + \|u\|_{L^2})$.*

Proof. Using a partition of unity as in Lemma 9.14 it suffices to consider $u_j := \zeta_j u$. The function u_0 has support strictly within U and hence is covered by Lemma 10.16. It remains to look at the case $j > 0$. As in the proof of the previous lemma, u_j will satisfy the corresponding modified problem (there is no need to assume that ζ vanishes near the boundary since we assume $u \in H_0^1(U)$). Since we can straighten out the boundary within the support of ζ_j we can perform a change of variables (Problem 10.18), to reduce it to the situation where $U = B_1(0) \cap \mathbb{R}^n_+$ with $\mathbb{R}^n_+ := \{x \in \mathbb{R}^n | x_n > 0\}$ and u has support inside $\bar{B}_{1/2}(0)$. Note that we need a $C^{1,1}$ boundary such that the transformed equation satisfies again $A \in W^{1,\infty}(U)$.

Now for all tangential coordinate directions a sufficiently small translation will not affect the property of being in $H_0^1(U)$. Consequently we have that $v := D_l^{-\varepsilon} D_l^{\varepsilon} u \in H_0^1(U)$ for $1 \le l < n$ and we can proceed as in the proof of Lemma 10.16 to see that $\partial_l \nabla u \in L^2$ (together with the required estimate). Of course this breaks down if we translate v into the normal direction. To obtain an estimate for the missing derivative $\partial_n^2 u$ we use the differential equation. Indeed, since we already know $u \in H^2_{loc}$ by Lemma 10.16, we know that the differential equation holds a.e. in U. Hence solving the differential equation for $\partial_n^2 u$ and observing that by ellipticity we have $A_{nn} \ge \theta$ gives the required estimate for $\partial_n^2 u$.

Finally, for each piece with $j > 0$ we have an estimate of the form $\|u_j\|_{H^2} \le C_j(\|f\|_{L^2} + \|u\|_{L^2})$. For $j = 0$ this only follows directly from Lemma 10.16 if U is bounded (such that u_0 has compact support). If U is unbounded one can check that the same proof works to provide the corresponding estimate for u_0 (only the fact that ζ and its derivatives are bounded and that it vanishes in a neighborhood of the boundary is used during the proof). $\square$

Note that under the conditions of Theorem 10.10 we have $\|u\|_{H^1} \le C\|f\|_{L^2}$ and hence we have $\|u\|_{H^2} \le C\|f\|_{L^2}$ in this case. The same is true on a bounded domain as long as 0 is not an eigenvalue (which also shows that the L^2 norm cannot be dropped in general).

Corollary 10.19. *Suppose $A_{ij} \in W^{k+1,\infty}(U)$, $b_j \in H^k(U)$, $c \in H^k(U)$ and U has a bounded $C^{k+1,1}$ boundary. If $u \in H^1_0(U)$ is a weak solution of the elliptic problem $\bar{L}u = f$ for $f \in H^k(U)$, then $u \in H^{k+2}(U)$. Moreover, we have an estimate of the form $\|u\|_{H^{k+2}} \le C(\|f\|_{H^k} + \|u\|_{L^2})$.*

Proof. Observe that we cannot apply the previous result recursively since the derivatives of u will not satisfy the Dirichlet boundary conditions. Hence we proceed slightly different. Of course using our partition of unity we can reduce it to the case where $U = B_1(0)\cap\mathbb{R}^n_+$ and u has support inside $\bar{B}_{1/2}(0)$.

Now we observe that as long as we restrict our attention to derivatives which are tangential to the boundary plane, that is, $\partial_l u$ with $1 \le l < n$, we still have $\partial_l u \in H^1_0(U)$ (Problem 10.20). Moreover, as in the proof of Corollary 10.17, $\partial_l u$ is the weak solution of an associated elliptic problem and we can apply the previous lemma to conclude $\partial_l u \in H^2(U)$. Finally, since $\partial_n u$ also satisfies an associated differential equation, we can solve for the missing derivative $\partial_n^3 u$ just as we did in the proof of the previous lemma. Hence the claim for $U = B_1(0) \cap \mathbb{R}^n_+$ follows using induction. □

As already discussed before, these results also allow us to identify the domain of $\bar{L}$ more explicitly. Indeed, under the assumptions of Lemma 10.18 we have

$$\mathfrak{D}(\bar{L}) = H^2(U) \cap H^1_0(U) = \{u \in H^2(U) | u|_{\partial U} = 0\}. \tag{10.63}$$

Furthermore, using the previous corollary we can even identify the domain of powers of $\bar{L}$ which are defined recursively as $\mathfrak{D}(\bar{L}^k) := \{u \in \mathfrak{D}(\bar{L}) | \bar{L}u \in \mathfrak{D}(\bar{L}^{k-1})\}$.

Corollary 10.20. *Suppose $A_{ij} \in W^{k+1,\infty}(U)$, $b_j \in H^k(U)$, $c \in H^k(U)$ and U has a bounded $C^{k+1,1}$ boundary. Then*

$$\mathfrak{D}(\bar{L}^k) = \{u \in H^{2k}(U) | (\bar{L}^j u)|_{\partial U} = 0,\ 1 \le j < k\}. \tag{10.64}$$

Moreover, $\|u\|_{H^{2k}} \le C \sum_{j=0}^k \|\bar{L}^j u\|_{L^2}$ for $u \in \mathfrak{D}(\bar{L}^k)$.

Proof. As already pointed out before the case $k = 1$ is immediate from Lemma 10.18 since $u \in H^2$ implies $\sum_{i,j} \partial_i A_{ij} \partial_j u \in L^2$ as required by (10.53). Now we can use induction on k. If $u \in \mathfrak{D}(\bar{L}^k)$, then by definition $u \in \mathfrak{D}(\bar{L})$ and $\bar{L}u \in \mathfrak{D}(\bar{L}^{k-1})$. Hence by the induction hypothesis $u \in H^2 \cap H^1_0$ and $\bar{L}u \in H^{2(k-1)}$ such that by Corollary 10.19 we have $u \in H^{2k}$ and $\|u\|_{H^{2k}} \le \tilde{C}(\|\bar{L}u\|_{H^{2(k-1)}} + \|u\|_{L^2}) \le C \sum_{j=0}^k \|\bar{L}^j u\|_{L^2}$. □

We remark that the inequality can be improved to $\|u\|_{H^{2k}} \le C(\|\bar{L}^k u\|_{L^2} + \|u\|_{L^2})$. This will follow from the abstract Landau inequality (Problem 11.21) together with the fact that $-\bar{L}$ generates a strongly continuous semigroup, which will be the content of Section 11.5.

Problem* 10.17. *Recall* (B.28) *and* (9.10). *Show*

$$\int_U v(D_l^{-\varepsilon}u)d^nx = -\int_U (D_l^{\varepsilon}v)u\,d^nx$$

as well as

$$D_l^{\varepsilon}(uv) = (T_{\varepsilon\delta^l}v)(D_l^{\varepsilon}u) + (D_l^{\varepsilon}v)u.$$

Problem* 10.18. *Let* $\psi \in C^1(\overline{V}, \overline{U})$ *be a volume preserving diffeomorphism between bounded open sets* U, V. *Show that if* $u \in H^1(U)$ *is a weak solution of the elliptic problem* $Lu = f$, *then* $\tilde{u} = u \circ \psi \in H^1(V)$ *is a weak solution of the elliptic problem* $\tilde{L}\tilde{u} = \tilde{f}$, *where*

$$\tilde{A} = \frac{\partial\psi}{\partial y}(A \circ \psi)\big(\frac{\partial\psi}{\partial y}\big)^T, \quad \tilde{b} = (b \circ \psi)\frac{\partial\psi}{\partial y}, \quad \tilde{c} = c \circ \psi, \quad \tilde{f} = f \circ \psi.$$

Moreover, $\tilde{A}$ *is uniformly elliptic provided* A *is.*

Problem* 10.19. *Show that for a weak solution* $u \in H^1(U)$ *we have*

$$\|\nabla u\|_2 \le \varepsilon\|f\|_2 + C\|u\|_2.$$

(Hint: Use ellipticity and start from $\theta\|\nabla u\|_2^2 \le \dots$*.)*

Problem* 10.20. *Let* $U := B_1(0) \cap \mathbb{R}^n_+$. *Suppose* $u \in H^1_0(U) \cap H^2(U)$ *with* $\operatorname{supp}(u) \subset\subset B_1(0)$. *Show that* $\partial_l u \in H^1_0(U)$ *for* $1 \le l < n$. *(Hint: First consider smooth functions. Then approximate using* $u_\varepsilon = \phi_\varepsilon * T_{-2\varepsilon}u$*.)*

10.4. The Poisson equation in $C(U)$

In this section we want to have a brief look at the Poisson equation in spaces other than L^2. In particular, we want to extend Theorem 5.30 to the case where f is not required to be Hölder continuous. By Example 5.3 we know that in this case the solution u might not be C^2 and hence the Laplacian has to be understood in the sense of distributions. We begin by looking at the homogenous equation (see also Problem 10.4 for a different proof).

Lemma 10.21. *Suppose* $h \in L^1_{loc}(U)$ *satisfies* $\Delta h = 0$ *in the sense of distributions. Then* h *is harmonic in* U.

Proof. Let ϕ_ε be a rotationally symmetric mollifier. Then the Poisson problem $\Delta u = \phi_{\varepsilon_1} - \phi_{\varepsilon_2}$ is solved by the Newton potential which satisfies $u \in C_c^\infty(\mathbb{R}^n)$ since $\int \big(\phi_{\varepsilon_1} - \phi_{\varepsilon_2}\big)d^nx = 1 - 1 = 0$, as shown in Problem 5.17. Consequently

$$h_{\varepsilon_1}(x) - h_{\varepsilon_2}(x) = \big((\phi_{\varepsilon_1} - \phi_{\varepsilon_2}) * h\big)(x) = \int_U (\Delta f)(x-y)h(y)d^ny = 0$$

as long as $\operatorname{dist}(x, \partial U) < \min(\varepsilon_1, \varepsilon_2)$. Hence $h(x) = h_\varepsilon(x)$ for $\operatorname{dist}(x, \partial U) < \varepsilon$ which shows that $h \in C^\infty(U)$. □

Neglecting boundary conditions we expect the solution to be given by the Newton potential (cf. Section 5.3). Hence we will first investigate in what sense the Newton potential solves the Poisson equation.

Lemma 10.22. *Suppose $U \subseteq \mathbb{R}^n$ is bounded and $f \in L^1(U)$. Then the Newton potential $u := \Phi * f$ is in $W^{1,1}(U)$ and satisfies $-\Delta u = f$ in the sense of distributions. Moreover, the first order derivatives are given by*

$$\partial_j u = (\partial_j \Phi) * f. \tag{10.65}$$

In particular, if additionally $f \in L^p(U)$ we have $u \in W^{1,p}(U)$ and if $f \in C_b(U)$ we have $u \in C^1_b(U)$.

Proof. We consider f as a function on $\mathbb{R}^n$ by setting it equal to 0 outside U as usual. First of all note that u is a well-defined function in $L^1_{loc}(\mathbb{R}^n)$. To see this note that we can split $\Phi = \Phi_1 + \Phi_2$ where $\Phi_1 \in W^{1,\infty}(\mathbb{R}^n)$ and $\Phi_2 \in W^{1,1}(\mathbb{R}^n)$. Hence $\Phi * f$, $(\partial_j \Phi) * f \in L^\infty(\mathbb{R}^n) + L^1(\mathbb{R}^n)$.

Next observe that for $\varphi \in C_c^\infty(\mathbb{R}^n)$ we have $\Phi * (-\Delta\varphi) = \varphi$. Indeed both sides satisfy the same Poisson problem and hence their difference must be a harmonic function with compact support, thus zero. Hence by Fubini we get

$$\begin{aligned}\int (-\Delta\varphi)(x) u(x) d^n x &= \int (-\Delta\varphi)(x) (\Phi * f)(x) d^n x \\ &= \int (\Phi * (-\Delta\varphi))(x) f(x) d^n x = \int \varphi(x) f(x) d^n x,\end{aligned}$$

which shows $-\Delta u = f$ in the sense of distributions.

Moreover, note that $\partial_j \Phi \in L^\infty(\mathbb{R}^n) + L^1(\mathbb{R}^n)$ and $\partial_j(\Phi * \phi) = (\partial_j \Phi) * \phi = \Phi * (\partial_j \phi)$ (as shown in Theorem 5.19). Hence we have

$$\begin{aligned}\int (\partial_j\varphi)(x) u(x) d^n x &= \int (\partial_j\varphi)(x) (\Phi * f)(x) d^n x = \int (\Phi * (\partial_j\varphi))(x) f(x) d^n x \\ &= \int ((\partial_j \Phi) * \varphi)(x) f(x) d^n x = -\int \varphi(x) ((\partial_j \Phi) * f)(x) d^n x,\end{aligned}$$

which shows $\partial_j u = (\partial_j \Phi) * f$ in the sense of distributions. The rest follows from Young's inequality. □

Now we have all the ingredients to extend Theorem 5.30:

Theorem 10.23. *Let $U \subseteq \mathbb{R}^n$ be a bounded domain with a regular boundary. Let $g \in C(\partial U)$ and $f \in C_b(U)$. Then the Poisson problem*

$$-\Delta u = f, \qquad u|_{\partial U} = g, \tag{10.66}$$

has a unique solution $u \in C^1(U) \cap C(\overline{U})$ in the sense of distributions.

Proof. Uniqueness follows from Lemma 10.21 since a harmonic function vanishing on the boundary must be zero. To see existence let $u_1 := \Phi * f$ be the corresponding Newton potential and u_0 the harmonic function which satisfies $u_0 = g - u_1$ on ∂U (Theorem 5.29). Then $u := u_1 - u_0$ is the solution we are looking for. □

Note that there is also a corresponding maximum principle.

Lemma 10.24. *Let U be a bounded domain and $\lambda \geq 0$. Suppose $v \in C(\overline{U})$ satisfies $(\lambda - \Delta)v \leq 0$ in the sense of distributions, that is, $\int_U (\lambda\varphi - \Delta\varphi) v d^n y \leq 0$ for all $\varphi \in C_c^\infty(U, [0,\infty))$. Then*

$$\max_{\overline{U}} v \leq \max_{\partial U} v^+. \tag{10.67}$$

Proof. Set $M := \max_{\partial U} v^+$ and consider $v_\varepsilon := \phi_\varepsilon v$ with ϕ_ε the Friedrichs mollifier. For $x \in U$ with $\operatorname{dist}(x, \partial U) > \varepsilon$ we have $\phi_\varepsilon(x - .) \in C_c^\infty(U, [0,\infty))$ and hence

$$(\lambda - \Delta) v_\varepsilon(x) = (\lambda\phi_\varepsilon - \Delta\phi_\varepsilon) * v(x) \leq 0.$$

Moreover, by continuity of v, for every δ there is an ε such that $v(x) \leq M + \delta$ whenever $\operatorname{dist}(x, \partial U) \leq \varepsilon$. Hence we also have $v_\varepsilon \leq M + \delta$ by the classical maximum principle Corollary 5.34. Taking $\varepsilon \downarrow 0$ we get $v \leq M + \delta$ and since $\delta > 0$ is arbitrary, the claim follows. □

Problem 10.21. *Show that that the Newton potential of $f \in C_b^k(U)$ is in $C_b^{k+1}(U)$.*

Chapter 11

Operator semigroups

In this chapter we want to look at (semi)linear ordinary linear differential equations in Banach spaces. We will need a few relevant facts about differentiation and integration for Banach space valued functions to be reviewed first.

11.1. Single variable calculus in Banach spaces

Let X be a Banach space. Let $I \subseteq \mathbb{R}$ be some interval and denote by $C(I, X)$ the set of continuous functions from I to X. Given $t \in I$ we call $f : I \to X$ differentiable at t if the limit

$$\dot{f}(t) := \lim_{\varepsilon \to 0} \frac{f(t + \varepsilon) - f(t)}{\varepsilon} \tag{11.1}$$

exists. If t is a boundary point, the limit/derivative is understood as the corresponding onesided limit/derivative.

The set of functions $f : I \to X$ which are differentiable at all $t \in I$ and for which $\dot{f} \in C(I, X)$ is denoted by $C^1(I, X)$. Clearly $C^1(I, X) \subset C(I, X)$. As usual we set $C^{k+1}(I, X) := \{f \in C^1(I, X) | \dot{f} \in C^k(I, X)\}$. Note that if $A \in \mathscr{L}(X, Y)$ and $f \in C^k(I, X)$, then $Af \in C^k(I, Y)$ and $\frac{d}{dt} Af = A\dot{f}$.

The following version of the mean value theorem will be crucial.

Theorem 11.1 (Mean value theorem)**.** *Suppose $f \in C^1(I, X)$. Then*

$$\|f(t) - f(s)\| \le M|t - s|, \qquad M := \sup_{\tau \in [s,t]} \|\dot{f}(\tau)\|, \tag{11.2}$$

for $s \le t \in I$.

Proof. Fix $\tilde{M} > M$ and consider $d(\tau) := \|f(\tau) - f(s)\| - \tilde{M}(\tau - s)$ for $\tau \in [s,t]$. Suppose τ_0 is the largest τ for which $d(\tau) \le 0$ holds. Then there must be a sequence $\varepsilon_n \downarrow 0$ such that

$$\begin{aligned} 0 < d(\tau_0 + \varepsilon_n) &\le \|f(\tau_0 + \varepsilon_n) - f(\tau_0)\| - \tilde{M}\varepsilon_n + d(\tau_0) \\ &= \|\dot{f}(\tau_0)\varepsilon_n + o(\varepsilon_n)\| - \tilde{M}\varepsilon_n \le (M - \tilde{M} + o(1))\varepsilon_n < 0. \end{aligned}$$

Taking $n \to \infty$ contradicts our assumption. □

In particular,

Corollary 11.2. *For $f \in C^1(I,X)$ we have $\dot{f} = 0$ if and only if f is constant.*

Next we turn to integration. Let $I := [a,b]$ be compact. A function $f : I \to X$ is called a **step function** provided there are numbers

$$t_0 = a < t_1 < t_2 < \cdots < t_{n-1} < t_n = b \tag{11.3}$$

such that $f(t)$ is constant on each of the open intervals (t_{j-1}, t_j). The set of all step functions $S(I,X)$ forms a linear space and can be equipped with the sup norm. The corresponding Banach space obtained after completion is called the set of **regulated functions** $R(I,X)$. In other words, a regulated function is the uniform limit of a step function.

Observe that $C(I,X) \subset R(I,X)$. In fact, consider the functions $f_n := \sum_{j=0}^{n-1} f(t_j)\chi_{[t_j,t_{j+1})} \in S(I,X)$, where $t_j = a + j\frac{b-a}{n}$ and χ is the characteristic function. Since $f \in C(I,X)$ is uniformly continuous, we infer that f_n converges uniformly to f. Slightly more general, note that piecewise continuous functions are regulated since every piecewise continuous function is the sum of a continuous function and a step function.

For a step function $f \in S(I,X)$ we can define a linear map $\int : S(I,X) \to X$ by

$$\int_a^b f(t)dt := \sum_{j=1}^n x_j(t_j - t_{j-1}), \tag{11.4}$$

where x_i is the value of f on (t_{j-1}, t_j). This map satisfies

$$\left\| \int_a^b f(t)dt \right\| \le \|f\|_\infty (b-a). \tag{11.5}$$

and hence it can be extended uniquely to a linear map $\int : R(I,X) \to X$ with the same norm $(b-a)$. We even have

$$\left\| \int_a^b f(t)dt \right\| \le \int_a^b \|f(t)\| dt \tag{11.6}$$

since this holds for simple functions by the triangle inequality and hence for all functions by approximation.

We remark that it is possible to extend the integral to a larger class of functions in various ways. The first generalization is to replace step functions by simple functions (and at the same time one could also replace the Lebesgue measure on I by an arbitrary finite measure). Then the same approach defines the integral for uniform limits of simple functions. However, things only get interesting when you also replace the sup norm by an L^1 type seminorm: $\|f\|_1 := \int \|f(x)\|\, d\mu(x)$. As before the integral can be extended to all functions which can be approximated by simple functions with respect to this seminorm. This is known as the Bochner integral and we refer to Section B.5 for details.

Denote by $\mathscr{L}(X,Y)$ the bounded linear operators from $X \to Y$. If $A \in \mathscr{L}(X,Y)$, then $f \in R(I,X)$ implies $Af \in R(I,Y)$ and

$$A \int_a^b f(t)dt = \int_a^b A\, f(t)dt. \tag{11.7}$$

Again this holds for step functions and thus extends to all regulated functions by continuity. In particular, if $\ell \in X^*$ is a continuous linear functional, then

$$\ell(\int_a^b f(t)dt) = \int_a^b \ell(f(t))dt, \qquad f \in R(I,X). \tag{11.8}$$

Moreover, we will use the usual conventions $\int_{t_1}^{t_2} f(s)ds := \int_I \chi_{(t_1,t_2)}(s) f(s)ds$ and $\int_{t_2}^{t_1} f(s)ds := -\int_{t_1}^{t_2} f(s)ds$. Note that we could replace (t_1,t_2) by a closed or half-open interval with the same endpoints (why?) and hence $\int_{t_1}^{t_3} f(s)ds = \int_{t_1}^{t_2} f(s)ds + \int_{t_2}^{t_3} f(s)ds$.

Theorem 11.3 (Fundamental theorem of calculus)**.** *Suppose $F \in C^1(I,X)$, then*

$$F(t) = F(a) + \int_a^t \dot{F}(s)ds. \tag{11.9}$$

Conversely, if $f \in C(I,X)$, then $F(t) = \int_a^t f(s)ds \in C^1(I,X)$ and $\dot{F}(t) = f(t)$.

Proof. Let $f \in C(I,X)$ and set $G(t) := \int_a^t f(s)ds$. Then $G \in C^1(I,X)$ with $\dot{G}(t) = f(t)$ as can be seen from

$$\begin{aligned} \| \int_a^{t+\varepsilon} f(s)ds - \int_a^t f(s)ds - f(t)\varepsilon\| &= \| \int_t^{t+\varepsilon} (f(s) - f(t))ds\| \\ &\le |\varepsilon| \sup_{s\in[t,t+\varepsilon]} \|f(s) - f(t)\|. \end{aligned}$$

Hence if $F \in C^1(I,X)$ then $G(t) := \int_a^t (\dot{F}(s))ds$ satisfies $\dot{G} = \dot{F}$ and hence $F(t) = C + G(t)$ by Corollary 11.2. Choosing $t = a$ finally shows $F(a) = C$. □

Problem* 11.1 (Product rule). *Let X be a Banach algebra. Show that if $f, g \in C^1(I, X)$ then $fg \in C^1(I, X)$ and $\frac{d}{dt} fg = \dot{f} g + f \dot{g}$.*

Problem* 11.2. *Let $f \in R(I, X)$ and $\tilde{I} := I + t_0$. then $f(t - t_0) \in R(\tilde{I}, X)$ and*

$$\int_I f(t)dt = \int_{\tilde{I}} f(t - t_0)dt.$$

Problem* 11.3. *Let $A : \mathfrak{D}(A) \subseteq X \to X$ be a closed operator. Show that* (11.7) *holds for $f \in C(I, X)$ with* $\operatorname{Ran}(f) \subseteq \mathfrak{D}(A)$ *and $Af \in C(I, X)$.*

Problem 11.4. *Let $I = [a, b]$ and $J = [c, d]$ be two compact intervals. Suppose $f(s, t) : I \times J \to X$ is regulated in the sense that it is a uniform limit of step functions being constant on disjoint open rectangles $(s_{j-1}, s_j) \times (t_{k-1}, t_k)$ whose closure cover $I \times J$. Show that*

$$\int_J \left(\int_I f(s, t)ds \right) dt = \int_I \left(\int_J f(s, t)dt \right) ds.$$

(Hint: One way is to use linear functionals and reduce it to the classical Fubini theorem.)

11.2. Uniformly continuous operator groups

Our aim is to investigate the abstract Cauchy problem

$$\dot{u} = Au, \qquad u(0) = u_0 \tag{11.10}$$

in some Banach space X. Here A is some linear operator and we will assume that $A \in \mathscr{L}(X)$ to begin with. Note that in the simplest case $X = \mathbb{R}^n$ this is just a linear first order system with constant coefficient matrix A. In this case the solution is given by

$$u(t) = T(t)u_0, \tag{11.11}$$

where

$$T(t) := \exp(tA) := \sum_{j=0}^{\infty} \frac{t^j}{j!} A^j \tag{11.12}$$

is the exponential of tA. It is not difficult to see that this also gives the solution in our Banach space setting.

Theorem 11.4. *Let $A \in \mathscr{L}(X)$. Then the series in* (11.12) *converges and defines a* ***uniformly continuous operator group****:*

(i) *The map $t \mapsto T(t)$ is continuous, $T \in C(\mathbb{R}, \mathscr{L}(X))$.*
(ii) *$T(0) = \mathbb{I}$ and $T(t + s) = T(t)T(s)$ for all $t, s \in \mathbb{R}$.*

Moreover, $T \in C^\infty(\mathbb{R}, \mathscr{L}(X))$ is the unique solution of $\dot{T}(t) = AT(t)$ with $T(0) = \mathbb{I}$ and it commutes with A, $AT(t) = T(t)A$.

Warning: The *uniformly* in the name refers to the fact that continuity is required with respect to the operator norm topology, which is (in the older literature) is also known as uniform operator topology.

Proof. Set

$$T_n(t) := \sum_{j=0}^{n} \frac{t^j}{j!} A^j.$$

Then (for $m \le n$)

$$\|T_n(t) - T_m(t)\| = \left\| \sum_{j=m+1}^{n} \frac{t^j}{j!} A^j \right\| \le \sum_{j=m+1}^{n} \frac{|t|^j}{j!} \|A\|^j \le \frac{|t|^{m+1}}{(m+1)!} \|A\|^{m+1} \mathrm{e}^{|t| \|A\|}.$$

In particular,

$$\|T(t)\| \le \mathrm{e}^{|t|\, \|A\|}$$

and $AT(t) = \lim_{n\to\infty} AT_n(t) = \lim_{n\to\infty} T_n(t)A = T(t)A$. Furthermore we have $\dot{T}_{n+1} = AT_n$ and thus

$$T_{n+1}(t) = \mathbb{I} + \int_0^t AT_n(s)ds.$$

Taking limits shows

$$T(t) = \mathbb{I} + \int_0^t AT(s)ds$$

or equivalently $T \in C^1(\mathbb{R}, \mathscr{L}(X))$ and $\dot{T}(t) = AT(t)$, $T(0) = \mathbb{I}$. Differentiating this last equation shows $\left(\frac{d}{dt}\right)^{k+1} T(t) = A\left(\frac{d}{dt}\right)^k T(t)$ and establishes $T \in C^\infty(\mathbb{R}, \mathscr{L}(X))$.

Suppose $S(t)$ is another solution, $\dot{S} = AS$, $S(0) = \mathbb{I}$. Then, by the product rule (Problem 11.1), $\frac{d}{dt} T(-t)S(t) = T(-t)AS(t) - AT(-t)S(t) = 0$ implying $T(-t)S(t) = T(0)S(0) = \mathbb{I}$. In the special case $T = S$ this shows $T(-t) = T^{-1}(t)$ and in the general case it hence proves uniqueness $S = T$. Finally, $T(t+s)$ and $T(t)T(s)$ both satisfy our differential equation and coincide at $t = 0$. Hence they coincide for all t by uniqueness. □

Note that choosing $s = -t$ in (ii) shows

$$T(t)^{-1} = T(-t). \tag{11.13}$$

Clearly A is uniquely determined by $T(t)$ via $A = \dot{T}(0)$. Moreover, from this we also easily get uniqueness for our original Cauchy problem. We will in fact be slightly more general and consider the inhomogeneous problem

$$\dot{u} = Au + g, \qquad u(0) = u_0, \tag{11.14}$$

where $g \in C(I, X)$. A solution necessarily satisfies

$$\frac{d}{dt} T(-t)u(t) = -AT(-t)u(t) + T(-t)\dot{u}(t) = T(-t)g(t)$$

and integrating this equation (fundamental theorem of calculus) shows the **Duhamel formula**

$$u(t) = T(t)\left(u_0 + \int_0^t T(-s)g(s)ds\right) = T(t)u_0 + \int_0^t T(t-s)g(s)ds. \quad (11.15)$$

Lemma 11.5. *Let $A \in \mathscr{L}(X)$ and $g \in C^k(I, X)$ for some $k \in \mathbb{N}_0$. Then* (11.14) *has a unique solution $u \in C^{k+1}(I, X)$ given by* (11.15).

Proof. Using Problem 11.6 it is straightforward to verify that this is indeed a solution for any given $g \in C(I, X)$. This also shows that $u \in C^{k+1}(I, X)$ since $T \in C^\infty(\mathbb{R}, \mathscr{L}(X))$. □

Example 11.1. For example, look at the discrete linear wave equation

$$\ddot{q}_n(t) = k\big(q_{n+1}(t) - 2q_n(t) + q_{n-1}(t)\big), \qquad n \in \mathbb{Z}.$$

Factorizing this equation according to

$$\dot{q}_n(t) = p_n(t), \qquad \dot{p}_n(t) = k\big(q_{n+1}(t) - 2q_n(t) + q_{n-1}(t)\big),$$

we can write this as a first order system

$$\frac{d}{dt}\begin{pmatrix} q_n \\ p_n \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ k\,A_0 & 0 \end{pmatrix}\begin{pmatrix} q_n \\ p_n \end{pmatrix}$$

with the Jacobi operator $(A_0 q)_n = q_{n+1} - 2q_n + q_{n-1}$. Since A_0 is a bounded operator on $X = \ell^p(\mathbb{Z})$, we obtain a well-defined uniformly continuous operator group in $\ell^p(\mathbb{Z}) \oplus \ell^p(\mathbb{Z})$. ⋄

Problem 11.5. *Show that if $A, B \in \mathscr{L}(X)$ commute, $[A, B] := AB - BA = 0$, then so do their associated groups and we have*

$$\exp(sA + tB) = \exp(sA)\exp(tB) = \exp(tB)\exp(sA), \qquad [A, B] = 0.$$

Problem* 11.6 (Product rule). *Suppose $f \in C^1(I, X)$ and $T \in C^1(I, \mathscr{L}(X, Y))$. Show that $Tf \in C^1(I, Y)$ and $\frac{d}{dt}Tf = \dot{T}f + T\dot{f}$.*

Problem 11.7. *Let X be a Hilbert space and $A \in \mathscr{L}(X)$. Show that $T(t)^*$ is a uniformly continuous operator group whose generator is A^*. Conclude that if A is skew adjoint, that is, $A^* = -A$, then T is unitary.*

Problem 11.8. *Discuss the discrete Schrödinger equation*

$$\mathrm{i}\dot{u} = Hu, \qquad (Hu)_n := u_{n+1} + u_{n-1} + q_n u_n,$$

in $\ell^2(\mathbb{Z})$, where $q \in \ell^\infty(\mathbb{Z}, \mathbb{R})$. In particular, show $\|u(t)\| = \|u(0)\|$ and $\langle u(t), Hu(t)\rangle = \langle u(0), Hu(0)\rangle$.

Problem 11.9. *Let $X := C_0(\mathbb{R})$ and consider $(A_\alpha f)(x) := \frac{1}{\alpha}(f(x + \alpha) - f(x))$ for $\alpha > 0$. Show that A_α is bounded and compute its norm. Compute the corresponding group T as well as its norm.*

11.3. Strongly continuous operator semigroups

In the previous section we have found a quite complete solution of the abstract Cauchy problem (11.14) in the case when A is bounded. However, since differential operators are typically unbounded, this assumption is too strong for applications to partial differential equations. Since it is unclear what the conditions on A should be, we will go the other way and impose conditions on T. First of all, even rather simple equations like the heat equation are only solvable for positive times and hence we will only assume that the solutions give rise to a semigroup. Moreover, continuity in the operator topology is too much to ask for (in fact, it is equivalent to boundedness of A — Problem 11.10) and hence we go for the next best option, namely strong continuity. In this sense, our problem is still well-posed.

A **strongly continuous operator semigroup** (also C_0-semigroup[1]) is a family of bounded operators $T(t) \in \mathscr{L}(X)$, $t \ge 0$, such that

(i) $T(t)g \in C([0,\infty), X)$ for every $g \in X$ (strong continuity) and

(ii) $T(0) = \mathbb{I}$, $T(t+s) = T(t)T(s)$ for all $t, s \ge 0$ (semigroup property).

If $T(t) \in \mathscr{L}(X)$ is defined for $t \in \mathbb{R}$ and item (ii) holds for all $t, s \in \mathbb{R}$ it is called a **strongly continuous operator group**.

We first note that $\|T(t)\|$ is uniformly bounded on compact time intervals.

Lemma 11.6. *Let $T(t)$ be a C_0-semigroup. Then there are constants $M \ge 1$, $\omega \ge 0$ such that*

$$\|T(t)\| \le M \mathrm{e}^{\omega t}, \qquad t \ge 0. \tag{11.16}$$

In case of a C_0-group we have $\|T(t)\| \le M \mathrm{e}^{\omega |t|}$, $t \in \mathbb{R}$.

Proof. Since $\|T(.)g\| \in C[0,1]$ for every $g \in X$ we have $\sup_{t\in[0,1]} \|T(t)g\| \le M_g$. Hence by the uniform boundedness principle $\sup_{t\in[0,1]} \|T(t)\| \le M$ for some $M \ge 1$. Setting $\omega = \log(M)$ the claim follows by induction using the semigroup property. For the group case apply the semigroup case to both $T(t)$ and $S(t) := T(-t)$. □

The infimum over all possible ω for which a corresponding M exists such that (11.16) holds is known as the **growth bound** $\omega_0(T)$ of the semigroup. It is given by

$$\omega_0(T) := \limsup_{t\to\infty} \frac{\log(\|T(t)\|)}{t} \tag{11.17}$$

and it can be shown that the limsup is actually a limit (Problem 11.11). However, there will not be a corresponding constant M for ω_0.

[1] This naming might look odd at first, but just reflects the fact that originally strong continuity was just one of several continuity properties, which were considered.

Inspired by the previous section we define the **generator** A of a strongly continuous semigroup as the linear operator

$$Af := \lim_{t\downarrow 0} \frac{1}{t}\big(T(t)f - f\big), \tag{11.18}$$

where the domain $\mathfrak{D}(A)$ is precisely the set of all $f \in X$ for which the above limit exists. By linearity of limits $\mathfrak{D}(A)$ is a linear subspace of X (and A is a linear operator) but at this point it is unclear whether it contains any nontrivial elements. We will however postpone this issue and begin with the observation that a C_0-semigroup is the solution of the abstract Cauchy problem associated with its generator A:

Lemma 11.7. *Let $T(t)$ be a C_0-semigroup with generator A. If $f \in \mathfrak{D}(A)$ then $T(t)f \in \mathfrak{D}(A)$ and $AT(t)f = T(t)Af$. Moreover, suppose $g \in X$ with $u(t) := T(t)g \in \mathfrak{D}(A)$ for $t > 0$. Then $u(t) \in C([0,\infty), X) \cap C^1((0,\infty), X)$ and $u(t)$ is the unique solution of the abstract Cauchy problem*

$$\dot{u}(t) = Au(t), \qquad u(0) = g. \tag{11.19}$$

This is, for example, the case if $g \in \mathfrak{D}(A)$ in which case we even have $u(t) \in C^1([0,\infty), X)$.

Similarly, if $T(t)$ is a C_0-group and $g \in \mathfrak{D}(A)$, then $u(t) := T(t)g \in C^1(\mathbb{R}, X)$ is the unique solution of (11.19) for all $t \in \mathbb{R}$.

Proof. Let $f \in \mathfrak{D}(A)$ and $t > 0$ (respectively $t \in \mathbb{R}$ for a group), then

$$\lim_{\varepsilon\downarrow 0} \frac{1}{\varepsilon}\big(u(t+\varepsilon) - u(t)\big) = \lim_{\varepsilon\downarrow 0} T(t)\frac{1}{\varepsilon}\big(T(\varepsilon)f - f\big) = T(t)Af.$$

This shows the first part. To show that $u(t)$ is differentiable it remains to compute

$$\begin{aligned} \lim_{\varepsilon\downarrow 0} \frac{1}{-\varepsilon}\big(u(t-\varepsilon) - u(t)\big) &= \lim_{\varepsilon\downarrow 0} T(t-\varepsilon)\frac{1}{\varepsilon}\big(T(\varepsilon)f - f\big) \\ &= \lim_{\varepsilon\downarrow 0} T(t-\varepsilon)\big(Af + o(1)\big) = T(t)Af \end{aligned}$$

since $\|T(t)\|$ is bounded on compact t intervals. Hence $u(t) \in C^1([0,\infty), X)$ (respectively $u(t) \in C^1(\mathbb{R}, X)$ for a group) solves (11.19). In the general case $f = T(t_0)g \in \mathfrak{D}(A)$ and $u(t) = T(t)g = T(t-t_0)f$ solves our differential equation for every $t > t_0$. Since $t_0 > 0$ is arbitrary it follows that $u(t)$ solves (11.19) by the first part. To see that it is the only solution, let $v(t)$ be a solution corresponding to the initial condition $v(0) = 0$. For $s < t$ we have

$$\begin{aligned}\frac{d}{ds}T(t-s)v(s) =& \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\big(T(t-s-\varepsilon)v(s+\varepsilon) - T(t-s)v(s)\big)\\ =& \lim_{\varepsilon\to 0} T(t-s-\varepsilon)\frac{1}{\varepsilon}\big(v(s+\varepsilon)-v(s)\big)\\ &- \lim_{\varepsilon\to 0} T(t-s-\varepsilon)\frac{1}{\varepsilon}\big(T(\varepsilon)v(s)-v(s)\big)\\ =& T(t-s)Av(s) - T(t-s)Av(s) = 0.\end{aligned}$$

Whence, $v(t) = T(t-t)v(t) = T(t-s)v(s) = T(t)v(0) = 0$. □

Note that our proof in fact even shows a bit more: If $g \in \mathfrak{D}(A)$ we have $u \in C^1([0,\infty), X)$ and hence not only $u \in C([0,\infty), X)$ but also $Au = \dot{u} \in C([0,\infty), X)$. Hence, if we regard $\mathfrak{D}(A)$ as a normed space equipped with the graph norm $\|f\|_A := \|f\| + \|Af\|$, in which case we will write $[\mathfrak{D}(A)]$, then $g \in \mathfrak{D}(A)$ implies $u \in C([0,\infty), [\mathfrak{D}(A)])$. In particular, T restricted to $[\mathfrak{D}(A)]$ is again a C_0-semigroup and it is straightforward to check that its generator is A restricted to $\mathfrak{D}(A^2) = \{f \in \mathfrak{D}(A) | Af \in \mathfrak{D}(A)\}$.

Similarly, $u(t) = T(t)g \in \mathfrak{D}(A)$ for $t > 0$ implies $u \in C((0,\infty), [\mathfrak{D}(A)])$. Moreover, recall that $[\mathfrak{D}(A)]$ will be a Banach space if and only if A is a closed operator (cf. Section B.3) and the latter fact will be established in Corollary 11.10 below.

Also observe that if one assumes $g \in \mathfrak{D}(A^k)$, one can apply Lemma 11.7 recursively to obtain:

Corollary 11.8. *Let $T(t)$ be a C_0-semigroup with generator A. If $g \in \mathfrak{D}(A^k)$ then $T(t)g \in \mathfrak{D}(A^k)$ and $T(t)g \in C^k([0,\infty), X)$ with*

$$\left(\frac{d}{dt}\right)^k T(t)g = A^j T(t) A^{k-j} g, \qquad t \ge 0, \tag{11.20}$$

for any $j = 0, \dots, k$. In case of a C_0-group we have $T(t)g \in C^k(\mathbb{R}, X)$ and the above formula holds for all $t \in \mathbb{R}$.

If we have $T(t)g \in \mathfrak{D}(A)$ for all $t > 0$, then $T(t)g \in \mathfrak{D}(A^k)$ for all $k \in \mathbb{N}$, $t > 0$ and $T(t)g \in C^\infty((0,\infty), X)$.

Proof. The case $k = 1$ is established in Lemma 11.7. Suppose the claim holds for $k \ge 1$ and $g \in \mathfrak{D}(A^{k+1})$. Then, applying A to $T(t)A^k g = A^k T(t) g \in \mathfrak{D}(A)$ shows $T(t)g \in \mathfrak{D}(A^{k+1})$ and $AT(t)A^k g = A^{k+1}T(t)g$. Finally, $(\frac{d}{dt})^{k+1}T(t)g = \frac{d}{dt}T(t)A^k g = AT(t)A^k g$ finishes the induction step.

To see the second claim we use again induction to show $T(t)g \in \mathfrak{D}(A^k)$. The case $k = 1$ holds by assumption. Now suppose the claim holds for some $k \in \mathbb{N}$, then $AT(t)g = T(t/2)\tilde{g}$, where $\tilde{g} := AT(t/2)g \in \mathfrak{D}(A^{k-1})$ by the induction hypothesis. Hence $AT(t)g \in \mathfrak{D}(A^k)$ by assumption implying $T(t)g \in \mathfrak{D}(A^{k+1})$. □

Extending our remark from above we can use the differential equation to show that for $g \in \mathfrak{D}(A^k)$ we have $T(t)g \in C^{k-j}([0,\infty),[\mathfrak{D}(A^j)])$, $0 \le j \le k$, where we equip $\mathfrak{D}(A^j)$ with the norm $\|f\|_{A^j} := \sum_{i=0}^{j} \|A^i f\|$. Then T restricted to $[\mathfrak{D}(A^j)]$ is a C_0-semigroup whose generator is A restricted to $\mathfrak{D}(A^{j+1})$.

A C_0-semigroup for which we have $T(t)g \in \mathfrak{D}(A)$ for all $g \in X$ and all $t > 0$ is called **differentiable**.

Before turning to some examples, we establish a useful criterion for a semigroup to be strongly continuous.

Lemma 11.9. *A (semi)group of bounded operators is strongly continuous if and only if* $\limsup_{\varepsilon\downarrow 0} \|T(\varepsilon)g\| < \infty$ *for every* $g \in X$ *and* $\lim_{\varepsilon\downarrow 0} T(\varepsilon)f = f$ *for* f *in a dense subset.*

Proof. We first show that $\limsup_{\varepsilon\downarrow 0} \|T(\varepsilon)g\| < \infty$ for every $g \in X$ implies that $T(t)$ is bounded in a small interval $[0,\delta]$. Otherwise there would exist a sequence $\varepsilon_n \downarrow 0$ with $\|T(\varepsilon_n)\| \to \infty$. Hence $\|T(\varepsilon_n)g\| \to \infty$ for some g by the uniform boundedness principle, a contradiction. Thus there exists some M such that $\sup_{t\in[0,\delta]} \|T(t)\| \le M$. Setting $\omega = \frac{\log(M)}{\delta}$ we even obtain (11.16). Moreover, boundedness of $T(t)$ shows that $\lim_{\varepsilon\downarrow 0} T(\varepsilon)f = f$ for all $f \in X$ by a simple approximation argument.

In case of a group this also shows $\|T(-t)\| \le \|T(\delta - t)\|\|T(-\delta)\| \le M\|T(-\delta)\|$ for $0 \le t \le \delta$. Choosing $\tilde{M} = \max(M, M\|T(-\delta)\|)$ we conclude $\|T(t)\| \le \tilde{M}\exp(\tilde{\omega}|t|)$.

Finally, right continuity is implied by the semigroup property: $\lim_{\varepsilon\downarrow 0} T(t+\varepsilon)g = \lim_{\varepsilon\downarrow 0} T(\varepsilon)T(t)g = T(t)g$. Left continuity follows from $\|T(t-\varepsilon)g - T(t)g\| = \|T(t-\varepsilon)(T(\varepsilon)g - g)\| \le \|T(t-\varepsilon)\|\|T(\varepsilon)g - g\|$. □

Example 11.2. Let $X := C_0(\mathbb{R})$ be the continuous functions vanishing as $|x| \to \infty$. Then it is straightforward to check that

$$(T(t)f)(x) := f(x+t)$$

defines a group of continuous operators on X. Since shifting a function does not alter its supremum we have $\|T(t)f\|_\infty = \|f\|_\infty$ and hence $\|T(t)\| = 1$. Moreover, strong continuity is immediate for uniformly continuous functions. Since every function with compact support is uniformly continuous and since such functions are dense, we get that T is strongly continuous. Moreover, for $f \in \mathfrak{D}(A)$ we have

$$\lim_{\varepsilon\to 0} \frac{f(t+\varepsilon) - f(t)}{\varepsilon} = (Af)(t)$$

uniformly. In particular, $f \in C^1(\mathbb{R})$ with $f, f' \in C_0(\mathbb{R})$. Conversely, for $f \in C^1(\mathbb{R})$ with $f, f' \in C_0(\mathbb{R})$ we have

$$\frac{f(t+\varepsilon) - f(t) - \varepsilon f'(t)}{\varepsilon} = \frac{1}{\varepsilon}\int_0^\varepsilon \big(f'(t+s) - f'(t)\big) ds \le \sup_{0\le s\le\varepsilon} \|T(s)f' - f'\|_\infty$$

which converges to zero as $\varepsilon \downarrow 0$ by strong continuity of T. Whence

$$A = \frac{d}{dx}, \qquad \mathfrak{D}(A) = \{f \in C^1(\mathbb{R}) \cap C_0(\mathbb{R}) | f' \in C_0(\mathbb{R})\}.$$

It is not hard to see that T is not uniformly continuous or, equivalently, that A is not bounded (cf. Problem 11.10).

Note that this group is not strongly continuous when considered on $X := C_b(\mathbb{R})$. Indeed for $f(x) = \cos(x^2)$ we can choose $x_n = \sqrt{2\pi n}$ and $t_n = \sqrt{2\pi}(\sqrt{n+\frac{1}{4}} - \sqrt{n}) = \frac{1}{4}\sqrt{\frac{\pi}{2n}} + O(n^{-3/2})$ such that $\|T(t_n)f - f\|_\infty \ge |f(x_n + t_n) - f(x_n)| = 1$. ◇

Next consider

$$u(t) = T(t)g, \qquad v(t) := \int_0^t u(s) ds, \qquad g \in X. \tag{11.21}$$

Then $v \in C^1([0,\infty), X)$ with $\dot{v}(t) = u(t)$ and (Problem 11.2)

$$T(\varepsilon)v(t) = \int_0^t u(\varepsilon + s) ds = \int_\varepsilon^{t+\varepsilon} u(s) ds = v(t+\varepsilon) - v(\varepsilon) \tag{11.22}$$

implying

$$\lim_{\varepsilon\downarrow 0} \frac{1}{\varepsilon}\big(T(\varepsilon)v(t) - v(t)\big) = \lim_{\varepsilon\downarrow 0}\Big(-\frac{1}{\varepsilon}v(\varepsilon) + \frac{1}{\varepsilon}\big(v(t+\varepsilon) - v(t)\big)\Big) = -g + u(t). \tag{11.23}$$

Consequently $v(t) \in \mathfrak{D}(A)$ and $Av(t) = -g + u(t)$ implying that $u(t)$ solves the following integral version of our abstract Cauchy problem

$$u(t) = g + A\int_0^t u(s) ds. \tag{11.24}$$

Note that while in the case of a bounded generator both versions are equivalent, this will not be the case in general. So while $u(t) = T(t)g$ always solves the integral version, it will only solve the differential version if $u(t) \in \mathfrak{D}(A)$ for $t > 0$ (which is clearly also necessary for the differential version to make sense). In the latter case $u(t)$ is sometimes called a **strong solution** (also **classical solution**), while otherwise it is called a **mild solution**.

Two further consequences of these considerations are also worth while noticing:

Corollary 11.10. *Let $T(t)$ be a C_0-semigroup with generator A. Then A is a densely defined and closed operator.*

Proof. Since $v(t) \in \mathfrak{D}(A)$ and $\lim_{t\downarrow 0} \frac{1}{t} v(t) = g$ for arbitrary g, we see that $\mathfrak{D}(A)$ is dense. Moreover, if $f_n \in \mathfrak{D}(A)$ and $f_n \to f$, $Af_n \to g$ then

$$T(t)f_n - f_n = \int_0^t T(s)Af_n ds.$$

Taking $n \to \infty$ and dividing by t we obtain

$$\frac{1}{t}\big(T(t)f - f\big) = \frac{1}{t}\int_0^t T(s)g\, ds.$$

Taking $t \downarrow 0$ finally shows $f \in \mathfrak{D}(A)$ and $Af = g$. □

Note that by the closed graph theorem we have $\mathfrak{D}(A) = X$ if and only if A is bounded. Moreover, since a C_0-semigroup provides the unique solution of the abstract Cauchy problem for A, we obtain

Corollary 11.11. *A C_0-semigroup is uniquely determined by its generator.*

Proof. Suppose T and S have the same generator A. Then by uniqueness for (11.19) we have $T(t)f = S(t)f$ for all $f \in \mathfrak{D}(A)$. Since $\mathfrak{D}(A)$ is dense this implies $T(t) = S(t)$ as both operators are continuous. □

Finally, as in the uniformly continuous case, the inhomogeneous problem can be solved by Duhamel's formula. However, now it is not so clear when this will actually be a solution.

Lemma 11.12. *Let A be the generator of a C_0-semigroup and $f \in C([0,\infty), X)$. If the inhomogeneous problem*

$$\dot{u} = Au + f, \qquad u(0) = g, \tag{11.25}$$

has a solution it is necessarily given by ***Duhamel's formula***

$$u(t) = T(t)g + \int_0^t T(t-s)f(s)ds. \tag{11.26}$$

Conversely, for $g = 0$, the function u given by (11.26) satisfies $u \in C^1([0,\infty), X)$ if and only if $u \in C([0,\infty), [\mathfrak{D}(A)])$. Moreover, in this case it will be a solution.

Specifically, (11.26) gives a solution if either one of the following conditions is satisfied:

- *$g \in \mathfrak{D}(A)$ and $f \in C([0,\infty), [\mathfrak{D}(A)])$.*
- *$g \in \mathfrak{D}(A)$ and $f \in C^1([0,\infty), X)$.*

If A is the generator of a strongly continuous group, we can replace $[0,\infty)$ by $\mathbb{R}$.

Proof. Let $u(t)$ be a solution of (11.25) and set $v(s) := T(t-s)u(s)$, $0 \le s \le t$, then one shows as in the proof of Lemma 11.7 that

$$\begin{aligned}\dot v(s) &= -AT(t-s)u(s) + T(t-s)\dot u(s)\\ &= -AT(t-s)u(s) + T(t-s)(Au(s) + f(s))\\ &= T(t-s)f(s), \qquad 0 < s < t.\end{aligned}$$

Hence the fundamental theorem of calculus (taking limits towards the boundary points) gives (11.26).

For the converse observe that $T(t)g$ is a solution of the homogenous equation if $g \in \mathfrak{D}(A)$. Hence it remains to investigate the integral, which we will denote by $u(t)$. We first note that

$$\frac{1}{\varepsilon}(u(t+\varepsilon) - u(t)) = \frac{1}{\varepsilon}\int_0^\varepsilon T(\varepsilon - s)f(t+s)ds + \frac{1}{\varepsilon}(T(\varepsilon) - \mathbb{I})u(t),$$

where the integral term on the right converges to $f(t)$ thanks to our assumption $f \in C([0,\infty), X)$. Hence, if one of the remaining two expressions has a limit, so has the other. In particular, if $u(t)$ is differentiable, we see that the limit on the right exists implying $u(t) \in \mathfrak{D}(A)$ and $\dot u(t) = f(t) + Au(t)$. Similarly if $u(t) \in \mathfrak{D}(A)$, then the limit on the right exists and we see that $u(t)$ is differentiable.

From this the first case is immediate since $u \in C([0,\infty), [\mathfrak{D}(A)])$ provided $f \in C([0,\infty), [\mathfrak{D}(A)])$ by Problem 11.3.

In case of the second condition we note that

$$u(t) = \int_0^t T(s)f(t-s)ds$$

by a change of variables (Problem 11.2) and hence

$$\begin{aligned}\frac{1}{\varepsilon}(u(t+\varepsilon) - u(t)) &= \frac{1}{\varepsilon}\int_0^t T(s)(f(t+\varepsilon-s) - f(t-s))ds\\ &\quad + \frac{1}{\varepsilon}\int_0^\varepsilon T(t+s)f(\varepsilon - s)ds\\ &\underset{\varepsilon\to 0}{\to} \int_0^t T(s)\dot f(t-s)ds + T(t)f(0)\end{aligned}$$

since $f \in C^1$. □

The function $u(t)$ defined by (11.26) is called the **mild solution** of the inhomogeneous problem. In general a mild solution is not a solution:

Example 11.3. Let $T(t)$ be a strongly continuous group with an unbounded generator A (e.g. the one from Example 11.2). Choose $f_0 \in X \setminus \mathfrak{D}(A)$ and set $g := 0$, $f(t) := T(t)f_0$. Then $f \in C(\mathbb{R}, X)$ and the mild solution is given by

$$u(t) = T(t)\int_0^t T(-s)f(s)ds = T(t)\int_0^t f_0 ds = t\,T(t)f_0.$$

Since $T(t)$ leaves $\mathfrak{D}(A)$ invariant, we have $u(t) \notin \mathfrak{D}(A)$ for all $t \in \mathbb{R}$ and hence $u(t)$ is not a solution. ⋄

Problem* 11.10. *Show that a uniformly continuous semigroup has a bounded generator. (Hint: Write $T(t) = V(t_0)^{-1}V(t_0)T(t) = \dots$ with $V(t) := \int_0^t T(s)ds$ and conclude that it is C^1.)*

Problem 11.11. *Let $f : [0,\infty) \to \mathbb{R}$ be bounded from above on every compact interval and subadditive, that is, $f(t_1 + t_2) \le f(t_1) + f(t_2)$. Then*

$$\lim_{t\to\infty} \frac{f(t)}{t} = \inf_{t\ge 0} \frac{f(t)}{t}.$$

Problem 11.12. *Show that the growth bound of a semigroup is given by*

$$\omega_0(T) = \inf_{t\ge 0} \frac{\log(\|T(t)\|)}{t} = \lim_{t\to\infty} \frac{\log(\|T(t)\|)}{t}.$$

Moreover, show that the spectral radius of $T(t)$ is given by

$$r(T(t)) = \mathrm{e}^{\omega_0(T)t}.$$

(Hint: The spectral radius of an operator $T \in \mathscr{L}(X)$ is defined as $r(T) := \sup_{z\in\sigma(T)} |z| = \lim_{n\to\infty} \|T^n\|^{1/n} \le \|T(t)\|$.)

Problem 11.13. *Let $T(t)$ be a C_0-semigroup. Show that if $T(t_0)$ has a bounded inverse for one $t_0 > 0$ then this holds for all $t > 0$ and it extends to a strongly continuous group via $T(t) := T(-t)^{-1}$ for $t < 0$.*

Problem 11.14. *Consider the translation group $T(t) := T_t$ on $L^p(\mathbb{R})$, $1 \le p < \infty$. Show that this is a strongly continuous group and compute its generator. Show that it is not strongly continuous for $p = \infty$. (Hint: Problem B.15.)*

Problem 11.15. *Consider the translation semigroup $T(t) := T_t$ on $L^p(0,1)$, $1 \le p < \infty$. Show that this is a strongly continuous group and compute its generator. Show that it is nilpotent: $T_t g = 0$ for $t \ge 1$. (Hint: Problem B.15.)*

Problem 11.16. *Let $U \subseteq \mathbb{R}^n$ and let $m : U \to \mathbb{C}$ be a measurable function with $\sup_{x\in U} \mathrm{Re}(m(x)) < \infty$. Consider the multiplication semigroup $T(t)g(x) := \mathrm{e}^{tm(x)}g(x)$ on $L^p(U)$, $1 \le p < \infty$. Show that this is a strongly continuous group and compute its generator.*

Problem 11.17. *Define a semigroup on $L^1(-1,1)$ via*

$$(T(t)f)(s) = \begin{cases} 2f(s-t), & 0 < s \le t, \\ f(s-t), & \text{else}, \end{cases}$$

where we set $f(s) = 0$ for $s < 0$. Show that the estimate from Lemma 11.6 does not hold with $M < 2$.

Problem 11.18. *Let A be the generator of a C_0-semigroup $T(t)$. Show*

$$T(t)f = f + tAf + \int_0^t (t-s)T(s)A^2 f\, ds, \qquad f \in \mathfrak{D}(A^2).$$

Problem 11.19. *Let A be the generator of a C_0-semigroup $T(t)$. Show that $\bigcap_{k\in\mathbb{N}} \mathfrak{D}(A^k)$ is dense. (Hint: Set $g_m := m\int_0^1 \phi(ms)T(s)g\, ds$, where $\phi \in C_c^\infty(0,1)$ with $\int_0^1 \phi(s)ds = 1$.)*

Problem 11.20. *Let $T(t)$ be a differentiable C_0-semigroup with generator A. Show $T \in C^\infty((0,\infty), \mathscr{L}(X))$ with $\frac{d^k}{dt^k}T(t) = A^k T(t)$, $t > 0$. (Hint: Show that $A^k T(t)$ is bounded and use Problem 11.18.)*

Problem 11.21 (Landau[2] inequality)**.** *Let A be the generator of a C_0-semigroup $T(t)$ satisfying $\|T(t)\| \le M$. Derive the* **abstract Landau inequality**

$$\|Af\| \le 2M\|A^2 f\|^{1/2}\|f\|^{1/2}.$$

(Hint: Problem 11.18.)

Problem 11.22. *Let A be the generator of a C_0-semigroup. Consider the integral version of our inhomogeneous problem* (11.25)*:*

$$u(t) = g + A\int_0^t u(s)ds + \int_0^t f(s)ds$$

for given $g \in X$, $f \in C([0,1), X)$. Show that this problem has a unique solution $u \in C([0,1), X)$ such that $\int_0^t u(s)ds \in \mathfrak{D}(A)$ for $t \ge 0$ which is given by Duhamel's formula (11.26)*. (Hint: Problem 11.4.)*

Problem 11.23. *A bounded operator $P \in \mathscr{L}(X)$ is said to commute with a closed operator A if*

$$PA \subseteq AP.$$

That is, if $\mathfrak{D}(PA) = \mathfrak{D}(A) \subseteq \mathfrak{D}(AP) = \{x \in X | Px \in \mathfrak{D}(A)\}$ and both operators agree on the smaller set $\mathfrak{D}(A)$. Note that this in particular requires that P leaves the domain invariant, $P\mathfrak{D}(A) \subseteq \mathfrak{D}(A)$.

Show that if $P \in \mathscr{L}(X)$ commutes with the generator of a (semi)group A, then it also ***commutes*** *with the (semi)group, $PT(t) = T(t)P$. (Hint: Uniqueness of solutions.)*

[2] Lev Landau (1908–1968), Soviet physicist

11.4. Generator theorems

Of course in practice the abstract Cauchy problem, that is the operator A, is given and the question is if A generates a corresponding C_0-semigroup. Corollary 11.10 already gives us some necessary conditions but this alone is not enough.

It turns out that it is crucial to understand the resolvent of A (see Section B.3). Using an operator-valued version of the elementary integral $\int_0^\infty \mathrm{e}^{t(a-z)}dt = -(a-z)^{-1}$ (for $\mathrm{Re}(a-z) < 0$) we can make the connection between the resolvent and the semigroup.

Lemma 11.13. *Let T be a C_0-semigroup with generator A satisfying* (11.16). *Then $\{z|\mathrm{Re}(z) > \omega\} \subseteq \rho(A)$ and*

$$R_A(z) = -\int_0^\infty \mathrm{e}^{-zt} T(t)dt, \quad \mathrm{Re}(z) > \omega, \tag{11.27}$$

where the right-hand side is defined as

$$\left(\int_0^\infty \mathrm{e}^{-zt} T(t)dt\right) g := \lim_{s\to\infty} \int_{1/s}^s \mathrm{e}^{-zt} T(t) g\, dt. \tag{11.28}$$

Moreover,

$$\|R_A(z)\| \le \frac{M}{\mathrm{Re}(z)-\omega}, \quad \mathrm{Re}(z) > \omega. \tag{11.29}$$

Proof. Let us abbreviate $R_s(z)f := -\int_0^s \mathrm{e}^{-zt}T(t)f\,dt$. Then, by virtue of (11.16), $\|\mathrm{e}^{-zt}T(t)f\| \le M\mathrm{e}^{(\omega-\mathrm{Re}(z))t}\|f\|$ shows that $R_s(z)$ is a bounded operator satisfying $\|R_s(z)\| \le M(\mathrm{Re}(z)-\omega)^{-1}$. Moreover, this estimate on the integrand also shows that the limit $R(z) := \lim_{s\to\infty} R_s(z)$ exists (and still satisfies $\|R(z)\| \le M(\mathrm{Re}(z)-\omega)^{-1}$). Next note that $S(t) := \mathrm{e}^{-zt}T(t)$ is a semigroup with generator $A - z$ (Problem 11.24) and hence for $f \in \mathfrak{D}(A)$ we have

$$R_s(z)(A-z)f = -\int_0^s S(t)(A-z)f\,dt = -\int_0^s \dot{S}(t)f\,dt = f - S(s)f.$$

In particular, taking the limit $s \to \infty$, we obtain $R(z)(A-z)f = f$ for $f \in \mathfrak{D}(A)$. Similarly, still for $f \in \mathfrak{D}(A)$, by Problem 11.3

$$(A-z)R_s(z)f = -\int_0^s (A-z)S(t)f\,dt = -\int_0^s \dot{S}(t)f\,dt = f - S(s)f$$

and taking limits, using closedness of A, implies $(A-z)R(z)f = f$ for $f \in \mathfrak{D}(A)$. Finally, if $g \in X$ choose $f_n \in \mathfrak{D}(A)$ with $f_n \to g$. Then $R(z)f_n \to R(z)g$ and $(A-z)R(z)f_n = f_n \to g$ proving $R(z)g \in \mathfrak{D}(A)$ and $(A-z)R(z)g = g$ for $g \in X$. $\square$

The number

$$s(A) := \sup\{\mathrm{Re}(z) | z \in \sigma(A)\} \tag{11.30}$$

is known as the **spectral bound** of A. The above lemma shows $s(A) \le \omega_0(T)$. Moreover, for matrices it is not hard to see that we have equality. However, this is not true in general. In particular, knowledge of $\sigma(A)$ alone is in general not sufficient to estimate the growth of the associated semigroup.

Corollary 11.14. *Let T be a C_0-semigroup with generator A satisfying* (11.16). *Then*

$$R_A(z)^{n+1} = \frac{(-1)^{n+1}}{n!} \int_0^\infty t^n \mathrm{e}^{-zt} T(t) dt, \quad \mathrm{Re}(z) > \omega, \tag{11.31}$$

and

$$\|R_A(z)^n\| \le \frac{M}{(\mathrm{Re}(z) - \omega)^n}, \quad \mathrm{Re}(z) > \omega,\ n \in \mathbb{N}. \tag{11.32}$$

Proof. Abbreviate $R_n(z) := \int_0^\infty t^n \mathrm{e}^{-zt} T(t) dt$ and note that

$$\frac{R_n(z+\varepsilon) - R_n(z)}{\varepsilon} = -R_{n+1}(z) + \varepsilon \int_0^\infty t^{n+2} \phi(\varepsilon t) \mathrm{e}^{-zt} T(t) dt$$

where $|\phi(\varepsilon)| \le \sum_{j=0}^\infty \frac{|\varepsilon|^j}{(j+2)!} \le \frac{1}{2} \mathrm{e}^{|\varepsilon|}$ from which we see $\frac{d}{dz} R_n(z) = -R_{n+1}(z)$ and hence $\frac{d^n}{dz^n} R_A(z) = -\frac{d^n}{dz^n} R_0(z) = (-1)^{n+1} R_n(z)$. Now the first claim follows using $R_A(z)^{n+1} = \frac{1}{n!} \frac{d^n}{dz^n} R_A(z)$ (Problem B.20). Estimating the integral using (11.16) establishes the second claim. □

Given these preparations we can now try to answer the question when A generates a semigroup. In fact, we will be constructive and obtain the corresponding semigroup by approximation. To this end we introduce the **Yosida approximation**[3]

$$A_n := -n A R_A(\omega + n) = -n - n(\omega + n) R_A(\omega + n) \in \mathscr{L}(X). \tag{11.33}$$

Of course this is motivated by the fact that this is a valid approximation for numbers since $\lim_{n\to\infty} \frac{-n}{a-\omega-n} = 1$. That we also get a valid approximation for operators is the content of the next lemma.

Lemma 11.15. *Suppose A is a densely defined closed operator with $(\omega, \infty) \subset \rho(A)$ satisfying*

$$\|R_A(\omega + n)\| \le \frac{M}{n}. \tag{11.34}$$

Then

$$\lim_{n\to\infty} -n R_A(\omega + n) f = f, \quad f \in X, \qquad \lim_{n\to\infty} A_n f = A f, \quad f \in \mathfrak{D}(A). \tag{11.35}$$

[3]Kosaku Yosida (1909–1990), Japanese mathematician

Proof. If $f \in \mathfrak{D}(A)$ we have $-nR_A(\omega+n)f = f - R_A(\omega+n)(A-\omega)f$ which shows $-nR_A(\omega+n)f \to f$ if $f \in \mathfrak{D}(A)$. Since $\mathfrak{D}(A)$ is dense and $\|nR_A(\omega+n)\| \le M$ this even holds for all $f \in X$. Moreover, for $f \in \mathfrak{D}(A)$ we have $A_n f = -nAR_A(\omega+n)f = -nR_A(\omega+n)(Af) \to Af$ by the first part. □

Moreover, A_n can also be used to approximate the corresponding semigroup under suitable assumptions.

Theorem 11.16 (Feller–Miyadera–Phillips[4]). *A linear operator A is the generator of a C_0-semigroup T satisfying* (11.16) *if and only if it is densely defined, closed, $(\omega,\infty) \subseteq \rho(A)$, and*

$$\|R_A(\lambda)^n\| \le \frac{M}{(\lambda-\omega)^n}, \quad \lambda > \omega,\ n \in \mathbb{N}. \tag{11.36}$$

Moreover, if A_n is the Yosida approximation (11.33) *and*

$$T_n(t) := \exp(tA_n) = \mathrm{e}^{-tn}\exp(-tn(\omega+n)R_A(\omega+n)) \tag{11.37}$$

are the corresponding groups, we have

$$T(t)g = \lim_{n\to\infty} T_n(t)g, \qquad g \in X. \tag{11.38}$$

Proof. Necessity has already been established in Corollaries 11.10 and 11.14.

For the converse we note

$$\|T_n(t)\| \le \mathrm{e}^{-tn}\sum_{j=0}^{\infty}\frac{(tn(\omega+n))^j}{j!}\|R_A(\omega+n)^j\| \le M\mathrm{e}^{-tn}\mathrm{e}^{t(\omega+n)} = M\mathrm{e}^{\omega t}.$$

Moreover, since $R_A(\omega+m)$ and $R_A(\omega+n)$ commute by the first resolvent identity (Problem B.20), we conclude that the same is true for A_m, A_n as well as for $T_m(t)$, $T_n(t)$ (by the very definition as a power series). Consequently

$$\begin{aligned}\|T_n(t)f - T_m(t)f\| &= \left\|\int_0^1 \frac{d}{ds}T_n(st)T_m((1-s)t)f\,ds\right\|\\ &\le t\int_0^1 \|T_n(st)T_m((1-s)t)(A_n-A_m)f\|ds\\ &\le tM^2\mathrm{e}^{\omega t}\|(A_n-A_m)f\|.\end{aligned}$$

Thus, for $f \in \mathfrak{D}(A)$ we have a Cauchy sequence and can define a linear operator by $T(t)f := \lim_{n\to\infty} T_n(t)f$. Since $\|T(t)f\| = \lim_{n\to\infty}\|T_n(t)f\| \le$

[4]William Feller (1906–1970), Croatian-American mathematician
[4]Isao Miyadera (1925), Japanese mathematician
[4]Ralph S. Phillips (1913–1998), American mathematician

$Me^{\omega t}\|f\|$, we see that $T(t)$ is bounded and has a unique extension to all of X. Moreover, $T(0) = \mathbb{I}$ and

$$\|T_n(t)T_n(s)f - T(t)T(s)f\| \leq \\ Me^{\omega t}\|T_n(s)f - T(s)f\| + \|(T_n(t) - T(t))T(s)f\|$$

implies $T(t+s)f = \lim_{n\to\infty} T_n(t+s)f = \lim_{n\to\infty} T_n(t)T_n(s)f = T(t)T(s)f$, that is, the semigroup property holds. Finally, by

$$\begin{aligned}\|T(\varepsilon)f - f\| &\leq \|T(\varepsilon)f - T_n(\varepsilon)f\| + \|T_n(\varepsilon)f - f\| \\ &\leq \varepsilon M^2 e^{\omega\varepsilon}\|(A - A_n)f\| + \|T_n(\varepsilon)f - f\|\end{aligned}$$

we see $\lim_{\varepsilon\downarrow 0} T(\varepsilon)f = f$ for $f \in \mathfrak{D}(A)$ and Lemma 11.9 shows that T is a C_0-semigroup. It remains to show that A is its generator. To this end let $f \in \mathfrak{D}(A)$, then

$$\begin{aligned}T(t)f - f &= \lim_{n\to\infty} T_n(t)f - f = \lim_{n\to\infty} \int_0^t T_n(s)A_n f\, ds \\ &= \lim_{n\to\infty} \left(\int_0^t T_n(s)Af\, ds + \int_0^t T_n(s)(A_n - A)f\, ds\right) \\ &= \int_0^t T(s)Af\, ds\end{aligned}$$

which shows $\lim_{t\downarrow 0} \frac{1}{t}(T(t)f - f) = Af$ for $f \in \mathfrak{D}(A)$. Finally, note that the domain of the generator cannot be larger, since $A - \omega - 1$ is bijective and adding a vector to its domain would destroy injectivity. But then $\omega + 1$ would not be in the resolvent set contradicting Lemma 11.13. □

Note that in combination with the following lemma this also answers the question when A generates a C_0-group.

Lemma 11.17. *An operator A generates a C_0-group if and only if both A and $-A$ generate C_0-semigroups.*

Proof. Clearly, if A generates a C_0-group $T(t)$, then $S(t) := T(-t)$ is a C_0-group with generator $-A$. Conversely, let $T(t)$, $S(t)$ be the C_0-semigroups generated by A, $-A$, respectively. Then a short calculation shows

$$\frac{d}{dt}T(t)S(t)g = -T(t)AS(t)g + T(t)AS(t)g = 0, \qquad t \geq 0.$$

Consequently, $T(t)S(t) = T(0)S(0) = \mathbb{I}$ and similarly $S(t)T(t) = \mathbb{I}$, that is, $S(t) = T(t)^{-1}$. Hence it is straightforward to check that T extends to a group via $T(-t) := S(t)$, $t \geq 0$. □

The following examples show that the spectral conditions are indeed crucial. Moreover, they also show that an operator might give rise to a Cauchy problem which is uniquely solvable for a dense set of initial conditions, without generating a strongly continuous semigroup.

Example 11.4. Let

$$A = \begin{pmatrix} 0 & A_0 \\ 0 & 0 \end{pmatrix}, \qquad \mathfrak{D}(A) = X \times \mathfrak{D}(A_0).$$

Then $u(t) = \left(\begin{smallmatrix} 1 & tA_0 \\ 0 & 1 \end{smallmatrix}\right)\left(\begin{smallmatrix} f_0 \\ f_1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} f_0 + tA_0 f_1 \\ f_1 \end{smallmatrix}\right)$ is the unique solution of the corresponding abstract Cauchy problem for given $f \in \mathfrak{D}(A)$. Nevertheless, if A_0 is unbounded, the corresponding semigroup is not strongly continuous.

Note that in this case we have $\sigma(A) = \{0\}$ if A_0 is bounded and $\sigma(A) = \mathbb{C}$ else. In fact, since A is not injective we must have $\{0\} \subseteq \sigma(A)$. For $z \neq 0$ the inverse of $A - z$ is given by

$$(A-z)^{-1} = -\frac{1}{z}\begin{pmatrix} 1 & \frac{1}{z}A_0 \\ 0 & 1 \end{pmatrix}, \quad \mathfrak{D}((A-z)^{-1}) = \operatorname{Ran}(A-z) = X \times \mathfrak{D}(A_0),$$

which is bounded if and only if A is bounded. ◇

Example 11.5. Let $X_0 = C_0(\mathbb{R})$ and $m(x) = \mathrm{i}x$. Then we can regard m as a multiplication operator on X_0 when defined maximally, that is, $f \mapsto m\, f$ with $\mathfrak{D}(m) = \{f \in X_0 | m\, f \in X_0\}$. Note that since $C_c(\mathbb{R}) \subseteq \mathfrak{D}(m)$ we see that m is densely defined. Moreover, it is easy to check that m is closed.

Now consider $X = X_0 \oplus X_0$ with $\|f\| = \max(\|f_0\|, \|f_1\|)$ and note that

$$A = \begin{pmatrix} m & m \\ 0 & m \end{pmatrix}, \qquad \mathfrak{D}(A) = \mathfrak{D}(m) \oplus \mathfrak{D}(m),$$

is closed. Moreover, for $z \notin \mathrm{i}\mathbb{R}$ the resolvent is given by the multiplication operator

$$R_A(z) = \frac{1}{m-z}\begin{pmatrix} 1 & -\frac{m}{m-z} \\ 0 & 1 \end{pmatrix}.$$

For $\lambda > 0$ we compute

$$\|R_A(\lambda) f\| \le \left(\sup_{x\in\mathbb{R}} \frac{1}{|\mathrm{i}x - \lambda|} + \sup_{x\in\mathbb{R}} \frac{|x|}{|\mathrm{i}x-\lambda|^2} \right) \|f\| = \frac{3}{2\lambda}\|f\|$$

and hence A satisfies (11.36) with $M = \frac{3}{2}$, $\omega = 0$ and $n = 1$. However, by

$$\|R_A(\lambda + \mathrm{i}n)\| \ge \|R_A(\lambda+\mathrm{i}n)(0, f_n)\| \ge \left| \frac{\mathrm{i}n f_n(n)}{(\lambda - \mathrm{i}n + \mathrm{i}n)^2} \right| = \frac{n}{\lambda^2},$$

where f_n is chosen such that $f_n(n) = 1$ and $\|f_n\|_\infty = 1$, it does not satisfy (11.32). Hence A does not generate a C_0-semigroup. Indeed, the solution of the corresponding Cauchy problem is

$$T(t) = \mathrm{e}^{t\,m}\begin{pmatrix} 1 & t\,m \\ 0 & 1 \end{pmatrix}, \qquad \mathfrak{D}(T) = X_0 \oplus \mathfrak{D}(m),$$

which is unbounded. ◇

When it comes to applying this theorem, the main difficulty will be establishing the resolvent estimate (11.36). Moreover, while it might be already difficult to estimate the resolvent, it will in general be even more challenging to get estimates on its powers. In this connection note that the trivial estimate $\|R_A(z)^n\| \le \|R_A(z)\|^n$ will do the job if and only if $M = 1$. Hence we finally look at the special case of **contraction semigroups** satisfying

$$\|T(t)\| \le 1. \tag{11.39}$$

By a simple transform the case $M = 1$ in Lemma 11.6 can always be reduced to this case (Problem 11.24). Moreover, as already anticipated, in the case $M = 1$ the estimate (11.29) immediately implies the general estimate (11.32) and it suffices to establish (11.36) for $n = 1$:

Corollary 11.18 (Hille[5]–Yosida)**.** *A linear operator A is the generator of a contraction semigroup if and only if it is densely defined, closed, $(0,\infty) \subseteq \rho(A)$, and*

$$\|R_A(\lambda)\| \le \frac{1}{\lambda}, \quad \lambda > 0. \tag{11.40}$$

Example 11.6. If A is the generator of a contraction, then clearly all eigenvalues z must satisfy $\mathrm{Re}(z) \le 0$. Moreover, for

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

we have

$$R_A(z) = -\frac{1}{z}\begin{pmatrix} 1 & 1/z \\ 0 & 1 \end{pmatrix}, \qquad T(t) = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix},$$

which shows that the bound on the resolvent is crucial. ⋄

Example 11.7. If X is a Hilbert space and A is a self-adjoint operator, then we have the required estimate if and only if A is bounded from above,

$$E := \sup \sigma(A) = \sup_{f \in \mathfrak{D}(A), \|f\|=1} \langle f, Af\rangle < \infty.$$

Indeed in this case we have (cf. [**32**, Theorem 2.19])

$$\|R_A(\lambda)\| \le \frac{1}{\lambda - E}, \qquad \lambda > E,$$

such that $A - E$ generates a contraction semigroup. In particular, note that in this case the growth bound equals the spectral bound. ⋄

However, for a given operator even the simple estimate (11.40) might be difficult to establish directly. Hence we outline another criterion.

[5] Einar Hille (1894–1980), American mathematician

Example 11.8. Let X be a Hilbert space and observe that for a contraction semigroup the expression $\|T(t)f\|$ must be nonincreasing. Consequently, for $f \in \mathfrak{D}(A)$ we must have

$$\frac{d}{dt}\|T(t)f\|^2\Big|_{t=0} = 2\mathrm{Re}\big(\langle f, Af\rangle\big) \le 0.$$

Operators satisfying $\mathrm{Re}(\langle f, Af\rangle) \le 0$ are called dissipative and this clearly suggests to replace the resolvent estimate by dissipativity. ⋄

To formulate this condition for Banach spaces, we first introduce the **duality set**

$$\mathcal{J}(x) := \{x' \in X^* | x'(x) = \|x\|^2 = \|x'\|^2\} \tag{11.41}$$

of a given vector $x \in X$. In other words, the elements from $\mathcal{J}(x)$ are those linear functionals which attain their norm at x and are normalized to have the same norm as x. As a consequence of the Hahn–Banach theorem (Corollary 4.13 from [**35**]) note that $\mathcal{J}(x)$ is nonempty. Moreover, it is also easy to see that $\mathcal{J}(x)$ is convex and weak-$*$ closed.

Example 11.9. Let X be a Hilbert space and identify X with X^* via $x \mapsto \langle x, .\rangle$ as usual. Then $\mathcal{J}(x) = \{x\}$. Indeed since we have equality $\langle x', x\rangle = \|x'\|\|x\|$ in the Cauchy–Schwarz inequality, we must have $x' = \alpha x$ for some $\alpha \in \mathbb{C}$ with $|\alpha| = 1$ and $\alpha^*\|x\|^2 = \langle x', x\rangle = \|x\|^2$ shows $\alpha = 1$. ⋄

Example 11.10. Recall (cf. Problem B.21) that a Banach space X is called strictly convex, if equality in the triangle inequality, $\|x + y\| = \|x\| + \|y\|$, can only occur if the vectors are parallel, $y = \alpha x$ for some $\alpha \ge 0$ or $x = 0$.

If X^* is strictly convex , then the duality set contains only one point. In fact, suppose $x', y' \in \mathcal{J}(x)$, then $z' = \frac{1}{2}(x' + y') \in \mathcal{J}(x)$ and $\frac{\|x\|}{2}\|x' + y'\| = z'(x) = \frac{\|x\|}{2}(\|x'\| + \|y'\|)$ implying $x' = y'$ by strict convexity. Note that the converse is also true: If $x', y' \in \mathcal{J}(x)$ for some $x \in B_1^X(0)$, then $x'(x) + y'(x) = 2$ implies $\|x' + y'\| = 2$ contradicting strict convexity.

This applies for example to $X := \ell^p(\mathbb{N})$ if $1 < p < \infty$ (cf. Problem B.21) in which case $X^* \cong \ell^q(\mathbb{N})$ with $q = \frac{p}{p-1}$. In fact, for $a \in X$ we have $\mathcal{J}(a) = \{a'\}$ with $a'_j = \|a\|_p^{2-p}\,\mathrm{sign}(a_j^*)|a_j|^{p-1}$. ⋄

Example 11.11. Let X be a measurable space with a σ-finite measure μ. The previous example can be generalized to $L^p(X, d\mu)$ if $1 < p < \infty$ (which are strictly, in fact even uniformly, convex by Theorem 3.11 from [**34**]). In this case we have $L^p(X, d\mu)^* \cong L^q(X, d\mu)$ and for $f \in L^p(X, d\mu)$ we have $\mathcal{J}(f) = \{g\}$ with $g = \|f\|_p^{2-p}\,\mathrm{sign}(f^*)|f|^{p-1}$. ⋄

Example 11.12. Let $X := C[0, 1]$ and choose $x \in X$. If t_0 is chosen such that $|x(t_0)| = \|x\|_\infty$, then the functional $y \mapsto x'(y) := x(t_0)^* y(t_0)$ satisfies $x' \in \mathcal{J}(x)$. Clearly $\mathcal{J}(x)$ will contain more than one element in general.

Note that for $X = C_b(\mathbb{R})$ the situation is more complicated since the supremum might not be attained. However, we can choose a sequence $t_n \in \mathbb{R}$ such that $x(t_n) \to x_0$ with $|x_0| = \|x\|_\infty$ and set $x'(y) = x_0^* L(y(t_n))$, where L is the Banach limit from Problem 4.23 from [**35**]. ⋄

Now a given operator $\mathfrak{D}(A) \subseteq X \to X$ is called **dissipative** if

$$\operatorname{Re}\big(x'(Ax)\big) \le 0 \quad \text{for one } x' \in \mathcal{J}(x) \text{ and all } x \in \mathfrak{D}(A). \tag{11.42}$$

Lemma 11.19. *Let $x, y \in X$. Then $\|x\| \le \|x - \alpha y\|$ for all $\alpha > 0$ if and only if there is an $x' \in \mathcal{J}(x)$ such that* $\operatorname{Re}(x'(y)) \le 0$.

Proof. Without loss of generality we can assume $x \ne 0$. If $\operatorname{Re}(x'(y)) \le 0$ for some $x' \in \mathcal{J}(x)$, then for $\alpha > 0$ we have

$$\|x'\|\|x\| = x'(x) \le \operatorname{Re}\big(x'(x - \alpha y)\big) \le \|x'\|\|x - \alpha y\|$$

implying $\|x\| \le \|x - \alpha y\|$.

Conversely, if $\|x\| \le \|x - \alpha y\|$ for all $\alpha > 0$, let $x'_\alpha \in \mathcal{J}(x - \alpha y)$ and set $y'_\alpha = \|x'_\alpha\|^{-1} x'_\alpha$. Then

$$\begin{aligned}\|x\| \le \|x - \alpha y\| &= y'_\alpha(x - \alpha y) = \operatorname{Re}\big(y'_\alpha(x)\big) - \alpha \operatorname{Re}\big(y'_\alpha(y)\big)\\ &\le \|x\| - \alpha \operatorname{Re}\big(y'_\alpha(y)\big).\end{aligned}$$

Now choose a sequence $\alpha_j \to 0$ such that both $y'_{\alpha_j}(x)$ and $y'_{\alpha_j}(y)$ converge. This defines a linear functional on the two dimensional subspace spanned by x and y which can be extended to a functional $y'_0 \in X^*$ using Hahn–Banach. Taking the limit in the above inequality yields $y'_0(x) = \|x\|$. Moreover, the above inequality also shows $\operatorname{Re}\big(y'_\alpha(y)\big) \le 0$ and hence $\operatorname{Re}\big(y'_0(y)\big) \le 0$. Whence $x'_0 = \|x\| y'_0 \in \mathcal{J}(x)$ and $\operatorname{Re}\big(x'_0(y)\big) \le 0$. □

As a straightforward consequence we obtain:

Corollary 11.20. *A linear operator is dissipative if and only if*

$$\|(A - \lambda)x\| \ge \lambda \|x\|, \qquad \lambda > 0,\ x \in \mathfrak{D}(A). \tag{11.43}$$

In particular, for a dissipative operator $A - \lambda$ is injective for $\lambda > 0$ and $(A - \lambda)^{-1}$ is bounded with $\|(A - \lambda)^{-1}\| \le \lambda^{-1}$. However, this does not imply that λ is in the resolvent set of A since $\mathfrak{D}((A - \lambda)^{-1}) = \operatorname{Ran}(A - \lambda)$ might not be all of X.

Now we are ready to show

Theorem 11.21 (Lumer–Phillips[6]). *A linear operator A is the generator of a contraction semigroup if and only if it is densely defined, dissipative, and $A - \lambda_0$ is surjective for one $\lambda_0 > 0$. Moreover, in this case* (11.42) *holds for all $x' \in \mathcal{J}(x)$.*

[6] Gunter Lumer (1929–2005), German born American mathematician

Proof. Let A generate a contraction semigroup $T(t)$ and let $x \in \mathfrak{D}(A)$, $x' \in \mathcal{J}(x)$. Then

$$\mathrm{Re}\big(x'(T(t)x - x)\big) \le |x'(T(t)x)| - \|x\|^2 \le \|x'\|\|x\| - \|x\|^2 = 0$$

and dividing by t and letting $t \downarrow 0$ shows $\mathrm{Re}\big(x'(Ax)\big) \le 0$. Hence A is dissipative and by Corollary 11.18 $(0,\infty) \subseteq \rho(A)$, that is, $A - \lambda$ is bijective for $\lambda > 0$.

Conversely, by Corollary 11.20 $A - \lambda$ has a bounded inverse satisfying $\|(A-\lambda)^{-1}\| \le \lambda^{-1}$ for all $\lambda > 0$. In particular, for λ_0 the inverse is defined on all of X and hence closed. Thus A is also closed and $\lambda_0 \in \rho(A)$. Moreover, from $\|R_A(\lambda_0)\| \le \lambda_0^{-1}$ (cf. Lemma B.21) we even get $(0, 2\lambda_0) \subseteq \rho(A)$ and iterating this argument shows $(0,\infty) \subseteq \rho(A)$ as well as $\|R_A(\lambda)\| \le \lambda^{-1}$, $\lambda > 0$. Hence the requirements from Corollary 11.18 are satisfied. □

Note that generators of contraction semigroups are maximal dissipative in the sense that they do not have any dissipative extensions. In fact, if we extend A to a larger domain we must destroy injectivity of $A - \lambda$ and thus the extension cannot be dissipative.

Example 11.13. Let X be a Hilbert space. An equation of the form

$$\mathrm{i}\dot{u} = Hu$$

with H a self-adjoint operator, is known as abstract **Schrödinger equation**. If H is bounded, it is easy to see that $\exp(-\mathrm{i}tH)$ is a uniformly continuous unitary group (cf. Problem 11.7). The Lumer–Phillips theorem also allows us to handle the unbounded case.

In this context a densely defined operator H is called symmetric if

$$\langle f, Hg\rangle = \langle Hf, g\rangle, \qquad f, g \in \mathfrak{D}(H).$$

In this case $\langle f, Hf\rangle$ is real-valued or equivalently, both $A = -\mathrm{i}H$ and $-A = \mathrm{i}H$ are dissipative. Hence if we assume $\mathrm{Ran}(H+\mathrm{i}) = \mathrm{Ran}(H-\mathrm{i}) = X$, then both A and $-A$ will generate contraction semigroups from which it is not hard to see that $T(t)$ is a strongly continuous group which preserves the norm (cf. Problem 11.26). But an operator preserving the norm is unitary, $T(t)^{-1} = T(t)^*$. Since for a symmetric operator $\mathrm{Ran}(H+\mathrm{i}) = \mathrm{Ran}(H-\mathrm{i}) = X$ is equivalent to self-adjointness ([**32**, Lemma 2.3]), we see that a self-adjoint operator gives rise to a strongly continuous unitary group. This also follows from the spectral theorem. The converse is also true and known as **Stone's theorem**.[7]

In fact, the converse is also true. To see this observe that H is symmetric if and only if $\langle f, Hf\rangle$ is real-valued which in turn is equivalent to both $-\mathrm{i}H$ and $\mathrm{i}H$ being dissipative. ◇

[7] Marshall Harvey Stone (1903–1989), American mathematician

Example 11.14. Let $X := C_0[0,1] = \{f \in C[0,1] | f(0) = f(1) = 0\}$ and consider the one-dimensional **heat equation**

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

on a finite interval $x \in [0,1]$ with Dirichlet boundary conditions $u(0) = u(1) = 0$ and the initial condition $u(0,x) = g(x)$. The corresponding operator is

$$Af = f'', \qquad \mathfrak{D}(A) = \{f \in C_0[0,1] | f \in C^2[0,1]\} \subset X.$$

Note that $\mathfrak{D}(A)$ is dense. For $\ell \in \mathcal{J}(f)$ we can choose $\ell(g) = f(x_0)^* g(x_0)$, where x_0 is chosen such that $|f(x_0)| = \|f\|_\infty$. Then $\mathrm{Re}(f(x_0)^* f(x))$ has a global maximum at $x = x_0$ and if $f \in \mathfrak{D}(A)$ we must have $\mathrm{Re}(f(x_0)^* f''(x_0)) \le 0$ provided this maximum is in the interior of $(0,1)$. If x_0 is at the boundary this holds trivially and consequently A is dissipative. That $A-\lambda$ is surjective follows using the Green's function (cf. Section 3.3 from **[35]**): For $g \in X$ the function

$$f(x) := (R_A(\lambda)g)(x) = \int_0^1 G(\lambda,x,y) g(y) dy,$$

where

$$G(\lambda,x,y) := \frac{-1}{\sqrt{\lambda}\sinh(\lambda)} \begin{cases} \sinh(\sqrt{\lambda}(1-x))\sinh(\sqrt{\lambda}y), & y \le x, \\ \sinh(\sqrt{\lambda}(1-y))\sinh(\sqrt{\lambda}x), & x \le y, \end{cases}$$

is in $\mathfrak{D}(A)$ and satisfies $(A-\lambda)f = g$. Note that alternatively one could compute the norm of the resolvent

$$\|R_A(\lambda)\| = \frac{1}{\lambda}\left(1 - \frac{1}{\cosh(\sqrt{\lambda}/2)}\right)$$

(equality is attained for constant functions; while these are not in X, you can approximate them by choosing functions which are constant on $[\varepsilon, 1-\varepsilon]$). ⋄

Example 11.15. Let us consider the heat equation on $X := C_{buc}(\mathbb{R})$ the bounded uniformly continuous functions. Since the uniform limit of uniformly continuous functions is again uniformly continuous, this is a closed subspace of $C_b(\mathbb{R})$ and hence a Banach space (it will become clear why we do not choose $C_b(\mathbb{R})$ in a moment). In this case we choose $\mathfrak{D}(A) := C^2_{buc}(\mathbb{R}) := \{f \in C^2(\mathbb{R}) | f, f', f'' \in X\}$. Now dissipativity does not follow as in the previous example since the maximum might not be attained. Hence we go directly for the resolvent whose kernel is given by

$$G(\lambda,x,y) := \frac{-1}{2\sqrt{\lambda}} \mathrm{e}^{-\sqrt{\lambda}|x-y|}.$$

One checks that $R_A(\lambda)$ is a bounded map on X whose norm is given by $\|R_A(\lambda)\| = \frac{1}{\lambda}$ with equality for constant functions. Moreover, for given $g \in X$ we have $g := R_A(\lambda) f \in \mathfrak{D}(A)$ with $(A-\lambda)f = g$. Conversely, if

$f \in \mathfrak{D}(A)$ one checks that $R_A(\lambda)(A-\lambda)f = f$ and hence $R_A(\lambda)$ is indeed the resolvent of A. Up to this point everything would work on $C_b(\mathbb{R})$. Moreover, note that the mollification of a function $g \in X$ will be in $\mathfrak{D}(A)$ and converge uniformly to g. Hence $\mathfrak{D}(A)$ is dense in X (but not in $C_b(\mathbb{R})$).

Note that since the heat group is given by mollification with the heat kernel, the same argument also shows directly that the heat group is not strongly continuous on $C_b(\mathbb{R})$. ⋄

Example 11.16. Another neat example is the following linear **delay differential equation**

$$\dot{u}(t) = \int_{t-1}^{t} u(s) d\nu(s),\ t > 0, \qquad u(s) = g(s),\ -1 \le s \le 0,$$

where ν is a complex measure. To this end we introduce the following operator

$$Af := f', \quad \mathfrak{D}(A) := \{f \in C^1[-1,0] | f'(0) = \int_{-1}^{0} f(s) d\nu(s)\} \subset C[0,1].$$

Suppose that we can show that it generates a semigroup T on $X = C[0,1]$ and set $u(t) := (T(t)f)(0)$ for $f \in \mathfrak{D}(A)$. Then, since T leaves $\mathfrak{D}(A)$ invariant, the function $r \mapsto (T(t+r)f)(s-r)$ is differentiable with

$$\frac{d}{dr}(T(t+r)f)(s-r) = (T(t+r)Af)(s-r) - (T(t+r)f')(s-r) = 0$$

and we conclude $(T(t+r)f)(s-r) = (T(t)f)(s)$ for $-1+r \le s \le 0$. In particular, for $r = s$ we obtain $u(t+s) = (T(t)f)(s)$. Hence we obtain

$$\begin{aligned}\dot{u}(t) &= \frac{d}{dt}(T(t)f)(0) = (AT(t)f)(0) = \int_{-1}^{0} (T(t)f)(s) d\nu(s) \\ &= \int_{-1}^{0} u(t+s) d\nu(s)\end{aligned}$$

and u solves our delay differential equation. Now if $g \in C[0,1]$ is given we can approximate it by a sequence $f_n \in \mathfrak{D}(A)$. Then $u_n(t) := (T_n(t)f_n)(0)$ will converge uniformly on compact sets to $u(t) := (T(t)g)(0)$ and taking the limit in the differential equation shows that u is differentiable and satisfies the differential equation.

Hence it remains to show that A generates a semigroup. First of all we claim that $\tilde{A} := A - \|\nu\|$ is dissipative, where $\|\nu\|$ is the total variation of ν. As in the previous example, for $\ell \in \mathcal{J}(f)$ we can choose $\ell(g) = f(x_0)^* g(x_0)$ where x_0 is chosen such that $|f(x_0)| = \|f\|_\infty$. Then $\mathrm{Re}(f(x_0)^* f(x))$ has a global maximum at $x = x_0$ and if $f \in \mathfrak{D}(A)$ we must have $\mathrm{Re}(f(x_0)^* f'(x_0)) = 0$ provided x_0 is in the interior. If $x_0 = -1$ we still must have $\mathrm{Re}(f(x_0)^* f'(x_0)) \le 0$. In both cases $\mathrm{Re}(\ell(\tilde{A}f)) \le -\|\nu\||f(x_0)|^2 \le 0$. If $x_0 = 0$ we compute

$$\mathrm{Re}(\ell(\tilde{A}f)) = \mathrm{Re}\left(f^*(0)\int_{-1}^{0} f(s)d\nu(s)\right) - \|\nu\||f(0)|^2 \le 0$$

since $|f(s)| \le |f(0)|$. Thus $\tilde{A}$ is dissipative. Moreover, it is straightforward to verify that the differential equation $(\tilde{A}-\lambda)f = g$ has a unique solution $f \in \mathfrak{D}(A)$ for $\lambda > 0$ since $|\int_{-1}^{0} \mathrm{e}^{(\lambda+\|\nu\|)s} d\nu(s)| \le \|\nu\|$. ⋄

Finally, we note that the condition that $A - \lambda_0$ is surjective can be weakened to the condition that $\mathrm{Ran}(A-\lambda_0)$ is dense. To this end we need:

Lemma 11.22. *Suppose A is a densely defined dissipative operator. Then A is closable and the closure $\overline{A}$ is again dissipative.*

Proof. Recall that A is closable if and only if for every $x_n \in \mathfrak{D}(A)$ with $x_n \to 0$ and $Ax_n \to y$ we have $y = 0$. So let x_n be such a sequence and choose another sequence $y_n \in \mathfrak{D}(A)$ such that $y_n \to y$ (which is possible since $\mathfrak{D}(A)$ is assumed dense). Then by dissipativity (specifically Corollary 11.20)

$$\|(A-\lambda)(\lambda x_n + y_m)\| \ge \lambda\|\lambda x_n + y_m\|, \quad \lambda > 0$$

and letting $n \to \infty$ and dividing by λ shows

$$\|y + (\lambda^{-1}A - 1)y_m\| \ge \|y_m\|.$$

Finally $\lambda \to \infty$ implies $\|y - y_m\| \ge \|y_m\|$ and $m \to \infty$ yields $0 \ge \|y\|$, that is, $y = 0$ and A is closable. To see that $\overline{A}$ is dissipative choose $x \in \mathfrak{D}(\overline{A})$ and $x_n \in \mathfrak{D}(A)$ with $x_n \to x$ and $Ax_n \to \overline{A}x$. Then (again using Corollary 11.20) taking the limit in $\|(A-\lambda)x_n\| \ge \lambda\|x_n\|$ shows $\|(\overline{A}-\lambda)x\| \ge \lambda\|x\|$ as required. □

Consequently:

Corollary 11.23. *Suppose the linear operator A is densely defined, dissipative, and $\mathrm{Ran}(A-\lambda_0)$ is dense for one $\lambda_0 > 0$. Then A is closable and $\overline{A}$ is the generator of a contraction semigroup.*

Proof. By the previous lemma A is closable with $\overline{A}$ again dissipative. In particular, $\overline{A}$ is injective and by Lemma B.19 we have $(\overline{A}-\lambda_0)^{-1} = \overline{(A-\lambda_0)^{-1}}$. Since $(A-\lambda_0)^{-1}$ is bounded its closure is defined on the closure of its domain, that is, $\mathrm{Ran}(\overline{A}-\lambda_0) = \overline{\mathrm{Ran}(A-\lambda_0)} = X$. The rest follows from the Lumer–Phillips theorem. □

Problem* 11.24. *Let $T(t)$ be a C_0-semigroup and $\alpha > 0$, $\lambda \in \mathbb{C}$. Show that $S(t) := \mathrm{e}^{\lambda t}T(\alpha t)$ is a C_0-semigroup with generator $B = \alpha A + \lambda$, $\mathfrak{D}(B) = \mathfrak{D}(A)$.*

Problem 11.25. *Let T be a bounded C_0-semigroup satisfying $\|T(t)\| \le M$. Then*
$$\|g\|_T := \sup_{t\ge 0} \|T(t)g\|$$
defines an equivalent norm on X satisfying
$$\|g\| \le \|g\|_T \le M\|g\|.$$

Problem 11.26. *Show that A generates a C_0 group of isometries, that is, $\|T(t)g\| = \|g\|$ for all $g \in X$ if and only if both A and $-A$ generate contraction semigroups. That is, both A and $-A$ satisfy the hypothesis of either the Hill–Yosida or the Lumer–Phillips theorem.*

Problem 11.27. *Let $T(t)$ be a C_0-semigroup satisfying $\|T(t)\| \le M\mathrm{e}^{\omega t}$ with generator A and $B \in \mathscr{L}(X)$. Show that $A+B$ generates a C_0-semigroup $S(t)$ satisfying $\|S(t)\| \le M\mathrm{e}^{(\omega+M\|B\|)t}$. (Hint: First use Problem 11.25 to reduce it to the case of a contraction. Then use Problem B.18.)*

Problem 11.28. *Let $X = \ell^2(\mathbb{N})$ and $(Aa)_n := \mathrm{i}n^2 a_n$, $(Ba)_n := n a_n$ both defined maximally. Show that A generates a C_0-semigroup but $A+\varepsilon B$ does not for any $\varepsilon > 0$.*

Problem 11.29. *Consider the heat equation (Example 11.14) on $[0,1]$ with Neumann boundary conditions $u'(0) = u'(1) = 0$.*

Problem 11.30. *Consider the heat equation (Example 11.15) on $C_0(\mathbb{R})$.*

11.5. Applications to parabolic equations

In this section we want to look at parabolic equations. As a warmup we look at the heat equation on a bounded domain U. As always we choose $L^2(U)$ to be our underlying Hilbert space. Moreover, we choose Dirichlet boundary conditions and consider the corresponding operator
$$\bar{L} = -\Delta, \qquad \mathfrak{D}(\bar{L}) = \{f \in H^1_0(U) | \Delta f \in L^2(U)\}. \tag{11.44}$$
Since we have
$$\langle f, \bar{L}f\rangle = \int_U (\nabla f)^* \cdot \nabla f \, d^n x \tag{11.45}$$
we see that $-\bar{L}$ is dissipative. Moreover, since it has a bounded inverse we conclude that it generates a contraction semigroup $T(t)$ by the Lumer–Phillips theorem.

Moreover, since $\bar{L}$ is self-adjoint and has an orthonormal basis of eigenfunctions, we can be even more explicit. First of all note that if we choose one of the eigenfunctions w_j as initial condition, we have
$$T(t)w_j = \mathrm{e}^{-tE_j} w_j. \tag{11.46}$$

Indeed, the right-hand side is obviously a solution of the heat equation, which hence must coincide with the left-hand side by uniqueness. Moreover, by linearity we further get

$$u(t) := T(t)g = \sum_{j=0}^{\infty} \langle w_j, g\rangle T(t) w_j = \sum_{j=0}^{\infty} \mathrm{e}^{-tE_j} \langle w_j, g\rangle w_j, \qquad g \in L^2(U). \tag{11.47}$$

In particular, since we assume the eigenvalues to be ordered, this shows

$$\|T(t)\| \le \mathrm{e}^{-tE_0} \tag{11.48}$$

that solutions decay exponentially at a rate determined by the smallest eigenvalue. Moreover, recall that

$$\mathfrak{D}(\bar{L}) = \{u \in L^2(U) | E_j \langle w_j, u\rangle \in \ell^2(\mathbb{N}_0)\} \tag{11.49}$$

and since $E_j \mathrm{e}^{-tE_j}$ is bounded for every $t > 0$ we conclude that $T(t)g \in \mathfrak{D}(\bar{L})$ for any $t > 0$ and hence Corollary 11.8 applies in this situation. In particular, for $t > 0$ we have $u(t) \in \mathfrak{D}(\bar{L}^n)$ for any $n \in \mathbb{N}$. To shed some further light on this, let us work out the domains $\mathfrak{D}(\bar{L}^n)$ more explicitly. We start with

$$\begin{aligned} \mathfrak{D}(\bar{L}^2) &= \{u \in \mathfrak{D}(\bar{L}) | \bar{L}u \in \mathfrak{D}(\bar{L})\} \\ &= \{u \in H_0^1(U) | \Delta u \in H_0^1(U),\ \Delta^2 u \in L^2(U)\}. \end{aligned} \tag{11.50}$$

Now if we assume that U has a $C^{1,1}$ boundary such that we can apply Lemma 10.18 this simplifies to

$$\mathfrak{D}(\bar{L}^2) = \{u \in H_0^1(U) \cap H^2(U) | \Delta u \in H_0^1(U) \cap H^2(U)\} \tag{11.51}$$

and if we further assume that U has a $C^{3,1}$ boundary Corollary 10.19 even shows

$$\mathfrak{D}(\bar{L}^2) = \{u \in H_0^1(U) \cap H^4(U) | \Delta u \in H_0^1(U)\}. \tag{11.52}$$

Proceeding like this we obtain (cf. Corollary 10.20)

$$\mathfrak{D}(\bar{L}^k) = \{u \in H^{2k}(U) | (\Delta^j u)|_{\partial U} = 0,\ 0 \le j < k\} \tag{11.53}$$

provided U has a $C^{2k-1,1}$ boundary. In particular, for $k > \frac{n}{4}$ we have a classical solution which is continuous up to the boundary. Finally, even without any assumptions on the boundary, Lemma 10.4 implies that $\mathfrak{D}(\bar{L}^k) \subset H^{2k}_{loc}(U)$ and hence $u(t) \in C^\infty(U)$ but we do not know anything about the boundary behavior of $u(t)$.

Moreover, note that we can define a corresponding heat kernel

$$K(t,x,y) := \sum_{j=0}^{\infty} \mathrm{e}^{-tE_j} w_j(y) w_j(x). \tag{11.54}$$

This kernel is in $L^2(U \times U) \cap C^\infty(U \times U)$ for $t > 0$ and hence a Hilbert–Schmidt kernel. In particular, it is compact. Using this kernel the solution of the inhomogeneous problem is given by Duhamel's formula

$$u(t,x) = \int_U K(t,x,y)g(y)d^n y + \int_0^t \int_U K(t-s,x,y)f(s,y)d^n y\, ds \quad (11.55)$$

and is a strong solution provided $g \in L^2(U)$ and $f \in C([0,\infty), L^2(U))$ satisfies one of the conditions from Lemma 11.12.

Lemma 11.24 (Maximum principle). *Suppose we have* $u \in C([0,\infty), L^2(U)) \cap C((0,\infty), [\mathfrak{D}(\bar{L})]) \cap C^1((0,\infty), L^2(U))$ *satisfies* $u(0) = g$ *and*

$$\dot{u}(t) \le \Delta u(t). \quad (11.56)$$

Then

$$u(t,x) \le \sup_{y\in\partial U} g(y). \quad (11.57)$$

Proof. Suppose $M := \sup_{\partial U} g < \infty$ and consider

$$\varphi(t) := \frac{1}{2}\|(u(t) - M)_+\|_2^2.$$

Then using that $\eta(r) = (r-M)_+^2$ satisfies $|\eta(r) - \eta(r_0) - \eta'(r_0)(r-r_0)| \le |r-r_0|^2$ one easily verifies that φ is differentiable and

$$\begin{aligned}\dot{\varphi}(t) &= \langle (u(t)-M)_+, \dot{u}(t)\rangle_2 \le \langle (u(t)-M)_+, \Delta u(t)\rangle_2 \\ &= -\int_U \chi_{\{x\in U | u(t,x)>M\}} |\nabla u(t,x)|^2 d^n x \le 0,\end{aligned}$$

where we have used integration by parts in the last step. Consequently $0 = \varphi(0) \ge \varphi(t) \ge 0$ implying $\varphi(t) = 0$, which establishes the claim. □

To handle (time independent) inhomogeneous boundary conditions

$$u(t,x) = a(x), \qquad x \in \partial U \quad (11.58)$$

one needs to solve the inhomogeneous Dirichlet problem

$$-\Delta u_0(x) = 0, \quad x \in U, \qquad u_0(x) = a(x), \quad x \in \partial U, \quad (11.59)$$

and observe that $u(t,x) - u_0(x)$ will solve the heat equation with homogenous boundary conditions. In particular, note that the solution will approach the equilibrium u_0 exponentially fast:

$$\|u(t) - u_0\|_2 \le \mathrm{e}^{-tE_0}. \quad (11.60)$$

In order to handle Neumann or Robin boundary conditions, replace the Dirichlet Laplacian by the corresponding Neumann or Robin Laplacian, as explained in Section 10.1.

Note that using the results from Section 10.4 one can also establish existence within the class of continuous functions $X := C_0(U)$. To this end

$$Au := \Delta u, \qquad \mathfrak{D}(A) := \{C_0(U) | \Delta u \in C_0(U)\} \subset C_0(U), \tag{11.61}$$

where Δu is understood in the sense of distributions. Then A satisfies the assumptions of the Lumer–Phillips theorem and we get a contraction semigroup on $C_0(U)$:

Lemma 11.25. *Let U be a bounded domain with a regular boundary in the sense of Theorem 5.29. The operator A defined in* (11.61) *is closed, bijective, densely defined, and dissipative.*

Proof. That A is closed is straightforward and that it is bijective follows from Theorem 10.23. Next note that $C_c^\infty(U) \subset \mathfrak{D}(A)$ which shows that A is densely defined. To show dissipativity we need to show that for $\lambda \geq 0$ and $u \in \mathfrak{D}(A)$ with $(\Delta - \lambda)u = f$ we have $\|u\|_\infty \leq \lambda^{-1}\|f\|_\infty$. Let $\theta \in [0, \pi)$ and set $v := \mathrm{Re}(\mathrm{e}^{\mathrm{i}\theta} u) - \lambda^{-1}\|f\|_\infty$. Then Lemma 10.24 implies $v \leq 0$, which is the desired estimate. □

Observe that the above proof shows that we have $\mathfrak{D}(A) \subset C^1(U)$ but Example 5.3 shows that $\mathfrak{D}(A) \not\subset C^2(U)$. Moreover, observe that we cannot take $C_b(U)$ as underlying Banach space, since A would not be densely defined. Finally, since on a bounded domain the sup norm is stronger than the L^2 norm, the solution in $C_0(U)$ will agree with the solution in $L^2(U)$.

Of course it is straightforward to extend these considerations to equations of the form

$$u_t = -\bar{L}u, \tag{11.62}$$

where $\bar{L}$ is the elliptic operator from Section 10.2. There we have seen, that λ is in the resolvent set of $-\bar{L}$ for $\mathrm{Re}(\lambda) > -c_1 + \frac{b_0^2}{4\theta}$. Moreover, for such λ coercivity

$$\mathrm{Re}(\langle u, (\bar{L} + \lambda)u\rangle) = \mathrm{Re}\big(a(u,u) + \lambda\|u\|^2\big) \geq C\|u\|^2 \tag{11.63}$$

implies that $-(\bar{L} + \lambda)$ generates a contraction semigroup by the Lumer–Phillips theorem.

Theorem 11.26. *Let $\bar{L}$ be a uniformly elliptic operator. Then $-\bar{L}$ generates a strongly continuous semigroup T satisfying*

$$\|T(t)\| \leq \mathrm{e}^{\omega t}, \qquad \omega := -c_1 + \frac{b_0^2}{4\theta}. \tag{11.64}$$

Note that in the case where $\bar{L}$ is self-adjoint (i.e., if $b = 0$ and c is real-valued) Theorem 10.13 applies the same considerations as above can be made. In particular, in this case the proof of the maximum principle can be easily adapted:

Lemma 11.27 (Maximum principle)**.** *Let $\bar{L}$ be a uniformly elliptic operator with $b = 0$ and $c \geq 0$. Suppose $u \in C([0,\infty), L^2(U)) \cap C((0,\infty), [\mathfrak{D}(\bar{L})]) \cap C^1((0,\infty), L^2(U))$ satisfies $u(0) = g$ and*

$$\dot{u}(t) \leq -\bar{L}u(t). \tag{11.65}$$

Then

$$u(t,x) \leq \sup_{y \in \partial U} g(y). \tag{11.66}$$

Finally, as in Section 10.2 a **Galerkin method** can be used to compute the solution numerically: Choose some linearly independent vectors $\varphi_1, \dots, \varphi_N$ and look for the projection $u_N(t)$ of the solution onto the linear span of the these functions. Writing

$$u_N(t) = \sum_{j=1}^{N} \alpha_j(t)\varphi_j \tag{11.67}$$

we see that the unknown coefficients α_j can be determined by solving the linear system of ordinary differential equations

$$\sum_{j=1}^{N} \dot{\alpha}_j(t)\langle\varphi_k, \varphi_j\rangle_2 = \sum_{j=1}^{N} a(\varphi_k, \varphi_j)\alpha_j(t). \tag{11.68}$$

Since we assumed the φ_k to be linearly independent, the matrix $\langle\varphi_k, \varphi_j\rangle_2$ is invertible and hence we can solve this system for $\dot{\alpha}_j(t)$ to bring it into the usual form of a first order system.

Problem 11.31. *Show that a solution of the heat equation with Dirichlet boundary conditions satisfies*

$$\|u(t)\|_2^2 + 2\int_0^t \|\nabla u(t)\|^2 ds = \|u(0)\|_2^2.$$

Problem 11.32. *Extend* (11.47) *to cover the inhomogeneous heat equation.*

Problem 11.33. *Show that a solution $u(t) := T_N(t)g$ of the heat equation with Neumann boundary conditions satisfies*

$$\|u(t) - g_0\|^2 \leq \mathrm{e}^{-E_1 t}, \qquad g_0 := \frac{1}{|U|}\int_U g(x)d^n x,$$

for some $E_1 > 0$.

Problem 11.34. *Show that for a self-adjoint elliptic operator $\bar{L}$ on a bounded domain we have the estimates*

$$\|(T(t) - 1)f\| \leq Ct\|\bar{L}f\|, \quad \|\bar{L}T(t)\| \leq \frac{C}{t}, \qquad f \in \mathfrak{D}(\bar{L}),\ 0 < t \leq 1.$$

11.6. Applications to hyperbolic equations

In this section we want to look at hyperbolic equations. As a warmup we look at the wave equation on a bounded domain $U \subseteq \mathbb{R}^n$. As in the previous section we choose $L^2(U)$ to be our underlying Hilbert space. Moreover, we choose Dirichlet boundary conditions and consider the corresponding operator

$$\bar{L} = -\Delta, \qquad \mathfrak{D}(\bar{L}) = \{f \in H_0^1(U) | \Delta f \in L^2(U)\}. \tag{11.69}$$

To be able to apply our theory we first must transform the second order equation

$$u_{tt} = \Delta u \tag{11.70}$$

into a first order system by setting $v := u_t$, which gives

$$v_t = \Delta u, \qquad u_t = v. \tag{11.71}$$

Hence the corresponding abstract Cauchy problem is

$$\dot{\xi} = A\xi, \tag{11.72}$$

where

$$\xi := \begin{pmatrix} u \\ v \end{pmatrix}, \qquad A := \begin{pmatrix} 0 & 1 \\ -\bar{L} & 0 \end{pmatrix}. \tag{11.73}$$

Consequently we will choose $X := H_0^1(U) \oplus L^2(U)$ as our underlying Hilbert space and

$$\mathfrak{D}(A) := \mathfrak{D}(\bar{L}) \oplus H_0^1(U). \tag{11.74}$$

As norm we choose

$$\|\xi\|^2 := \|\nabla u\|_{L^2}^2 + \|v\|_{L^2}^2 = \int_U \big(|\nabla u|^2 + |v|^2\big) d^n x, \quad \xi = (u,v). \tag{11.75}$$

This is suggested by the fact that this norm corresponds to the energy (cf. Section 7.3), which is preserved by the time evolution. Consequently we expect A to be dissipative with respect to this norm, which can be easily verified:

$$\langle \xi, A\xi \rangle = \langle \nabla u, \nabla v \rangle_{L^2} - \langle v, \bar{L}u \rangle_{L^2} = 2\mathrm{i}\mathrm{Im}\big(\langle \nabla u, \nabla v \rangle\big). \tag{11.76}$$

Moreover, it is easy to see that $\pm A+1$ is surjective since the solution of $(\pm A + 1)\xi = \eta$ for $\eta = (g,h)$ is given by $\xi = (u,v)$, where u is the solution of $(\bar{L} + 1)u = g \mp h$ and $v = \pm g \mp u$. Hence A generates a strongly continuous group preserving the norm by the Lumer–Phillips theorem (cf. Problem 11.26). Note that our findings imply that $\mathrm{i}A$ is self-adjoint and hence this is an instance of Stone's theorem discussed in Example 11.13.

Theorem 11.28. *Let U be a bounded domain. The operator A defined in* (11.73), (11.74) *is the generator of a unitary C_0-group in the Hilbert space $H_0^1(U) \oplus L^2(U)$ with norm* (11.75).

We remark that if U is unbounded, then (11.75) no longer is a norm and we need to use the full norm for H^1. In this case we still get a C_0-semigroup, but it will no longer be unitary (cf. Problem 11.35).

Again we can derive a more explicit form of this group by looking at the eigenfunctions w_j of $\bar{L}$. It is easy to check that the solution corresponding to the initial conditions $u(0) = g_j w_j$ and $v(0) = h_j w_j$ is given by

$$u(t) = g_j \cos(\sqrt{E_j} t) w_j + h_j \frac{\sin(\sqrt{E_j} t)}{\sqrt{E_j}} w_j \tag{11.77}$$

and hence the solution corresponding to the initial conditions $u(0) = g$ and $v(0) = h$ is given by

$$u(t) = \sum_{j=0}^{\infty} \left(\langle w_j, g\rangle \cos(\sqrt{E_j} t) + \langle w_j, h\rangle \frac{\sin(\sqrt{E_j} t)}{\sqrt{E_j}} \right) w_j. \tag{11.78}$$

Note that in contradistinction to the heat equation, the time dependent multiplication factors do not provide sufficient decay for the time evolution to improve regularity. This is also in agreement with the fact that we have a group since in this case we have $T(t)X = X$ and hence $T(t)X \subseteq \mathfrak{D}(A)$ (which we had in case of the heat equation) is not possible.

If $f \in C(\mathbb{R}, L^2(U))$ the mild solution of the inhomogeneous problem is given by

$$\begin{aligned} u(t) = \sum_{j=0}^{\infty} \Big(& \langle w_j, g\rangle \cos(\sqrt{E_j} t) + \langle w_j, h\rangle \frac{\sin(\sqrt{E_j} t)}{\sqrt{E_j}} \\ & + \int_0^t \frac{\sin(\sqrt{E_j}(t-s))}{\sqrt{E_j}} \langle w_j, f(s)\rangle ds \Big) w_j. \end{aligned} \tag{11.79}$$

Of course the same argument applies to an elliptic problem (10.34) provided $b = 0$ such that the associated form $a(.,..)$ is symmetric and provided $c_1 > 0$ such that it gives rise to a norm which is equivalent to the H^1 norm on $H^1_0(U)$. Then we can choose

$$\|\xi\|^2 := a(u,u) + \|v\|^2_{L^2} = \int_U \Big(\sum_{i,j} A_{ij} (\partial_i u)^* (\partial_j u) + c|u|^2 + |v|^2 \Big) d^n x \tag{11.80}$$

as our norm on $X := H^1_0(U) \oplus L^2(U)$ and proceed as before. If U is bounded we can even allow $c_1 \ge 0$ and we have an eigenfunction expansion.

Problem 11.35. *Let $U \subseteq \mathbb{R}^n$ be a domain (not necessarily bounded). Consider $H^1_0(U) \oplus L^2(U)$ with norm*

$$\|\xi\|^2 := \|u\|^2_{H^1} + \|v\|^2_{L^2}, \qquad \xi = (u, v).$$

Show that A defined in (11.73), (11.74) *generates a C_0-group.*

Problem 11.36. *Discuss the telegraph equation*

$$u_{tt} + b\, u_t = \Delta u + cu,$$

where $c, b \in L^\infty(U)$*, on a bounded domain with Dirichlet boundary conditions. (Hint: Problem 11.27.)*

Problem 11.37. *Show that the one-dimensional hyperbolic equation*

$$u_{tt} = u_{xx} + bu_x + cu,$$

can be reduced to the case $b = 0$ *provided* $b \in C^1$*:*

$$\tilde{u}_{tt} = \tilde{u}_{xx} + \tilde{c}\tilde{u}.$$

(Hint: Make an ansatz $\tilde{u}(x) = \eta(x)u(x)$*.)*

Chapter 12

Nonlinear evolution equations

12.1. Semilinear equations

Linear problems are often only a first approximation and adding a nonlinear perturbation leads to the following semilinear problem

$$\dot{u} = Au + F(u), \qquad u(0) = g, \tag{12.1}$$

where A is supposed to generate a semigroup $T(t)$ and $F \in C(X, X)$ such that we can recast this problem as

$$u(t) = T(t)g + \int_0^t T(t-s)F(u(s))ds. \tag{12.2}$$

In fact, if we have a solution $u \in C([0, t_+), [\mathfrak{D}(A)]) \cap C^1([0, t_+), X)$ of (12.1), then Duhamel's formula shows that (12.2) holds. In the other direction you need a stronger assumption on F. However, it will be more convenient to work with (12.2) and we will call a solution a mild solution of (12.1). In fact, (12.2) is of fixed point type and hence begs us to apply the contraction principle. As always with nonlinear equations, we expect the solution to be only defined on a finite time interval $[0, t_+)$ in general.

Theorem 12.1. *Suppose $F \in C(X, X)$ is Lipschitz continuous on bounded sets. Then for every $g \in X$ there is a $t_0 = t_0(\|g\|) > 0$, such that there is a unique mild solution $u \in C([0, t_0], X)$. Moreover, the solution map $g \mapsto u(t)$ will be Lipschitz continuous from every ball $\|g\| \le \rho$ to $C([0, t_0(\rho)], X)$.*

Proof. We will consider $0 \le t \le 1$ and set $M := \sup_{0 \le t \le 1} \|T(t)\|$. Let $r := 1 + M\|g\|$ and consider the closed ball $\bar{B}_r(0) \subset X$. Let $L = L(r)$ be the

Lipschitz constant of F on $\bar{B}_r(0)$. Set

$$K(u)(t) := T(t)g + \int_0^t T(t-s)F(u(s))ds$$

and note that

$$\begin{aligned}\|K(u)(t)\| &\le M\|g\| + M\int_0^t \big(\|F(0)\| + L\|u(s)\|\big)ds\\ &\le M\|g\| + M\|F(0)\|t + MLt \sup_{0\le s\le t}\|u(s)\|\end{aligned}$$

and

$$\|K(u)(t) - K(v)(t)\| \le M\int_0^t L\big(\|u(s)-v(s)\|\big)ds \le MLt \sup_{0\le s\le t}\|u(s)-v(s)\|$$

Hence if we choose $t_0 \le 1$ such that

$$M(\|F(0)\| + Lr)t_0 < 1$$

then $\theta := MLt_0 < 1$ and K will be a contraction on $\bar{B}_r(0) \subset C([0,t_0],X)$. In particular, for two solutions u_j corresponding to g_j with $\|g_j\| \le \|g\|$ we will have $\|u_1 - u_2\|_\infty \le \frac{1}{1-\theta}\|g_1 - g_2\|$.

This establishes the theorem except for the fact that it only shows uniqueness for solutions which stay within $\bar{B}_r(0)$. However, since K maps from $\bar{B}_r(0)$ to its interior $B_r(0)$, a potential different solution starting at $g \in B_r(0)$ would need to branch off at the boundary, which is impossible since our solution does not reach the boundary. □

Corollary 12.2. *Suppose that $F \in C([\mathfrak{D}(A)],[\mathfrak{D}(A)])$ is Lipschitz continuous on bounded sets. Then for every $g \in \mathfrak{D}(A)$ there is a $t_1 = t_1(\|g\|_A) > 0$, such that there is a unique strong solution $u \in C^1([0,t_1],X) \cap C([0,t_1],[\mathfrak{D}(A)])$.*

Proof. Since T restricted to $[\mathfrak{D}(A)]$ generates a C_0-semigroup (see the discussion after Lemma 11.7), we can apply the previous result to this semigroup giving a solution $u \in C([0,t_1],[\mathfrak{D}(A)])$. This solution is in $C^1([0,t_1],X)$ by Lemma 11.12. □

Corollary 12.3. *If F is globally Lipschitz, then solutions are global.*

Proof. In this case we can consider K on all of $C([0,t_0],X)$ and set $M := \sup_{0\le t\le t_0}\|T(t)\|$. By induction we get for the iterates

$$\|K^n(u)(t) - K^n(v)(t)\| \le \frac{(MLt)^n}{n!}\sup_{0\le s\le t}\|u(s)-v(s)\|$$

and Weissinger's fixed point theorem (Theorem 9.28 from [**35**]) gives a solution on $C([0,t_0],X)$. Since $t_0 > 0$ is arbitrary, the claim follows. □

If solutions are not global, there is still a unique maximal solution: Fix $g \in X$ and let u_j be two solutions on $[0, t_j)$ with $0 < t_1 < t_2$. By the uniqueness part of our theorem, we will have $u_1(t) = u_2(t)$ for $0 \le t < \tau$ for some $\tau > 0$. Suppose $\tau < t_1$ and τ is chosen maximal. Let $r := \max_{0 \le t \le \tau} \|u_1(t)\|$ and $0 < \varepsilon < \min(\tau, t_0(r)/2)$ with $t_0(r)$ from our theorem. Then there is a solution v starting with initial condition $u_1(\tau - \varepsilon)$ which is defined on $[0, 2\varepsilon]$. Moreover, again by the uniqueness part of our theorem $u_1(t) = v(t - (\tau - \varepsilon)) = u_2(t)$ for $\tau - \varepsilon \le t \le \tau + \varepsilon$ contradicting our assumption that τ is maximal. Hence taking the union (with respect to their domain) over all mild solutions starting at g, we get a unique solution defined on a maximal domain $[0, t_+(g))$. Note that if $t_+(g) < \infty$, then $\|u(t)\|$ must blow up as $t \to t_+(g)$:

Lemma 12.4. *Let $t_+(g)$ be the maximal time of existence for the mild solution starting at g. If $t_+(g) < \infty$, then* $\liminf_{t \to t_+(g)} \|u(t)\| = \infty$.

Proof. Assume that $\rho := \sup_{0 \le t < t_+(g)} \|u(t)\| < \infty$. As above, choose $0 < \varepsilon < \min(t_+(g), t_0(\rho)/2)$ with $t_0(\rho)$ from our theorem. Then the solution v starting with initial condition $u(t_+(g) - \varepsilon)$ extends u to the interval $[0, t_+(g) + \varepsilon)$, contradicting maximality. □

In many applications it will happen that the local Lipschitz constant depends only on a weaker norm. In such a situation also the weaker norm will have to blow up.

Lemma 12.5. *Let $\|.\|_0$ be a norm, which is weaker than the standard norm on X, that is, $\|x\|_0 \le C_0 \|x\|$ for all $x \in X$. Suppose that there is a nondecreasing function $L : [0, \infty) \to [0, \infty)$ and a constant C such that*

$$\|F(x)\| \le C + L(\|x\|_0)\|x\|. \tag{12.3}$$

If $t_+(g) < \infty$, then $\liminf_{t \to t_+(g)} \|u(t)\|_0 = \infty$.

Proof. Starting from (12.2) we obtain

$$\begin{aligned} \|u(t)\| &\le M\mathrm{e}^{\omega t}\|g\| + M \int_0^t \mathrm{e}^{\omega(t-s)} \|F(u(s))\| ds \\ &\le M\mathrm{e}^{\omega t}\|g\| + MC\frac{\mathrm{e}^{\omega t} - 1}{\omega} + M \int_0^t \mathrm{e}^{\omega(t-s)} L(\|u(s)\|_0)\|u(s)\| ds \end{aligned}$$

and hence Gronwall's inequality ([**33**, Lemma 2.7]) implies

$$\|u(t)\| \le M(\|g\| + Ct) \exp\left(\omega t + M \int_0^t L(\|u(s)\|_0) ds\right).$$

This shows that the $\|.\|$ norm cannot blow up before the $\|.\|_0$ norm. □

So the key to proving global existence of solutions is an a priori bound on the norm of the solution. Typically such a bound will come from a conservation law.

Example 12.1. Consider the **discrete nonlinear Schrödinger equation (dNLS)**

$$\mathrm{i}\dot{u}(t) = Hu(t) + f(|u(t)|)u(t), \quad t \in \mathbb{R},$$

in $X := \ell^2(\mathbb{Z})$. Here H could be any bounded self-adjoint operator and f any locally Lipschitz continuous function. In applications $Hu_n := u_{n+1} + u_{n-1} + q_n u_n$ is the Jacobi operator, with $q \in \ell^\infty(\mathbb{Z})$ a real-valued sequence corresponding to an external potential and $q := 0$ (or $q := -2$, depending on your preferences) is the free discrete Schrödinger operator. The function f is typically an even polynomial.

Clearly we have

$$|f(|x|)x - f(|y|)y| \le |f(|x|) - f(|y|)||x| + |f(|y|)|x-y| \le L(\max(|x|,|y|))|x-y|$$

for $x, y \in \mathbb{C}$, where

$$L(r) := r \max_{[0,r]} |f'| + \max_{[0,r]} |f|.$$

Consequently

$$\|f(|u|)u - f(|v|)v\|_2 \le L(\max(\|u\|_\infty, \|v\|_\infty))\|u - v\|_2$$

is the required Lipschitz estimate (recall $\|u\|_\infty \le \|u\|_2$) to apply Theorem 12.1 to conclude existence of local solutions. Note that since our generator is bounded, there is no difference between mild and strong solutions. Moreover, Lemma 12.5 implies that if solutions are not global, then $\|u(t)\|_\infty$ must blow up. In this respect note that while $\ell^2(\mathbb{Z})$ is the most natural choice from a quantum mechanical point of view, our analysis still applies if we replace $\ell^2(\mathbb{Z})$ by $\ell^p(\mathbb{Z})$ for any $1 \le p \le \infty$. Then by $\ell^{p_1}(\mathbb{Z}) \subset \ell^{p_2}(\mathbb{Z})$ for $p_2 \le p_1$ and for a solution starting in $\ell^{p_1}(\mathbb{Z}) \subset \ell^{p_2}(\mathbb{Z})$ the existence interval in $\ell^{p_2}(\mathbb{Z})$ could be larger than in $\ell^{p_1}(\mathbb{Z})$. However, by Lemma 12.5 this is not the case and the solutions does not just loose decay but will always blow up pointwise (if it blows up at all).

Finally, if we assume that f is real-valued then solutions satisfy

$$\frac{d}{dt}\|u(t)\|_2^2 = 2\mathrm{Re}\langle \dot{u}(t), u(t)\rangle = 2\mathrm{Im}\left(\langle Hu, u\rangle + \langle f(|u(t)|)u(t), u(t)\rangle\right) = 0$$

and hence the dNLS equation has a unique global norm preserving solution $u \in C^1(\mathbb{R}, \ell^2(\mathbb{Z}))$. ⋄

Let me close with a few remarks: First of all, it is straightforward to extend these results to the situation where F depends on t or to the case where T is a group. Details are left to the reader. Moreover, if A is bounded, then it is Lipschitz continuous and could be absorbed in F. In fact, in this case our theorem just gives the Picard–Lindelöf theorem for ordinary differential equations in Banach spaces (in particular, in this case the differential equation (12.1) and the integral equation (12.2) are equivalent).

Problem 12.1. *Show that solutions of* (12.2) *are global if* $\|F(x)\| \le C(1 + \|x\|)$ *for some constant* C. *(Hint: Use Gronwall's inequality to bound* $\|u(t)\|$*.)*

12.2. Reaction diffusion equations

In this section we want to look at reaction diffusion equations

$$u_t = \Delta u + F(u)$$

on some bounded domain with (e.g.) Dirichlet boundary conditions, such that we know that Δ generates a contraction semigroup and we can apply our results from Section 12.1.

Example 12.2. Let $X := C[0,1]$ and consider the one-dimensional reaction-diffusion equation

$$\frac{\partial}{\partial t} u(t,x) = \frac{\partial^2}{\partial x^2} u(t,x) + F(u(t,x))$$

on a finite interval $x \in [0,1]$ with the boundary conditions $u(0) = u(1) = 0$ and the initial condition $u(0,x) = u_0(x)$. From Example 11.14 we know that the corresponding linear operator generates a C_0-semigroup. Hence if $F : \mathbb{R} \to \mathbb{R}$ is locally Lipschitz, then our theorem applies and we get existence of local mild solutions. If we even have that $F \in C^2(\mathbb{R}, \mathbb{R})$ with $F(0) = 0$ and the second derivative locally Lipschitz, then $F : [\mathfrak{D}(A)] \to [\mathfrak{D}(A)]$ is Lipschitz on bounded sets and the mild solutions will in fact be a strong solution (note that for $f \in \mathfrak{D}(A)$ we have $f'(x_0) = 0$ for some $x_0 \in (0,1)$ and hence $\|f'\|_\infty \le \|f''\|_\infty$). ⋄

In the last example we had to impose additional assumptions on F to obtain strong solutions. Since the heat equation improves the regularity of solutions, one might suspect, that our mild solutions are in fact strong solutions without further assumptions on F. To show that this is indeed the case we will first improve Lemma 11.12. To this end we will assume that A generates a differentiable C_0-semigroup $T(t)$ satisfying the estimate

$$\|AT(t)\| \le \frac{C}{t}, \qquad 0 < t \le 1. \tag{12.4}$$

Such semigroups are called **analytic** and there are various equivalent ways of characterizing them ([**9**, Theorem 4.6]). Our definition is not the most common one, but it will be convenient for our purpose since (12.4) will be all we need.

Example 12.3. Let $\bar{L}$ be a self-adjoint elliptic operator on a bounded domain. Then $-\bar{L}$ generates an analytic semigroup by Problem 11.34. In fact, this is true for any self-adjoint operator which is bounded from below and has an orthonormal basis of eigenfunction. It is even true for any self-adjoint operator which is bounded from below, but the proof requires the spectral theorem for self-adjoint operators. ⋄

Lemma 12.6. *Suppose A generates an analytic semigroup. Then the mild solution* (11.26) *of the inhomogeneous problem is in $C^1((0,t_0),X)$ provided f is locally Hölder continuous on* $(0,t_0)$.

Proof. We can assume $g=0$ without loss of generality and split

$$u(t) = u_1(t) + u_2(t) := \int_0^t T(t-s)f(t)ds + \int_0^t T(t-s)\big(f(s)-f(t)\big)ds$$

and introduce

$$u_{1,\varepsilon}(t) := \int_0^{t-\varepsilon} T(t-s)f(t)ds, \quad u_{2,\varepsilon}(t) := \int_0^{t-\varepsilon} T(t-s)\big(f(s)-f(t)\big)ds$$

such that $u_j(t) = \lim_{\varepsilon\downarrow 0} u_{j,\varepsilon}(t)$. Since T is differentiable, we have $u_{1,\varepsilon}(t) = T(\varepsilon)\int_0^{t-\varepsilon} T(t-\varepsilon-s)f(t)ds \in \mathfrak{D}(A)$ for $t>\varepsilon$ and thus

$$Au_{1,\varepsilon}(t) = \int_0^{t-\varepsilon} AT(t-s)f(t)ds = T(t)f(t) - T(\varepsilon)f(t)$$

showing $u_1(t) \in \mathfrak{D}(A)$ with $Au_1(t) = (T(t)-1)f(t)$. Moreover, we even have $u_1 \in C((0,t_0),[\mathfrak{D}(A)])$.

Similarly we have $u_{2,\varepsilon}(t) \in \mathfrak{D}(A)$ and

$$\begin{aligned}\left\|\int_{t-\varepsilon}^t AT(t-s)\big(f(s)-f(t)\big)ds\right\| &\le \int_{t-\varepsilon}^t \|AT(t-s)\|\|f(s)-f(t)\|ds\\ &\le CC_1\int_{t-\varepsilon}^t (t-s)^{\gamma-1}ds = \frac{CC_1}{\gamma}\varepsilon^\gamma\end{aligned}$$

thanks to our estimate (12.4) and local Hölder continuity $\|f(s)-f(t)\| \le C_1|t-s|^\gamma$ for $t-\varepsilon<s<t$. This shows $u_2(t)\in\mathfrak{D}(A)$ with

$$Au_2(t) = \int_0^t AT(t-s)\big(f(s)-f(t)\big)ds.$$

To see $u_2 \in C((0,t_0),[\mathfrak{D}(A)])$ we write

$$\begin{aligned}Au_2(t+\varepsilon) - Au_2(t) &= \int_t^{t+\varepsilon} AT(t+\varepsilon-s)\big(f(s)-f(t+\varepsilon)\big)ds\\ &\quad + (T(\varepsilon)-1)Au_2(t)\\ &\quad + \int_0^t AT(t+\varepsilon-s)\big(f(t)-f(t+\varepsilon)\big)ds.\end{aligned}$$

The norm of the first integral can be estimated by $\frac{CC_1}{\gamma}\varepsilon^\gamma$ and the last by $CC_1\varepsilon^\gamma\log(\frac{t+\varepsilon}{\varepsilon})$ which establishes continuity.

In summary, we have $u(t) \in C((0,t_0),[\mathfrak{D}(A)])$ and the claim follows from (the proof of) Lemma 11.12. $\square$

Applying this result to our semilinear problem gives:

Lemma 12.7. *Suppose A generates an analytic semigroup. Then the mild solution of the semilinear problem* (12.2) *from Theorem 12.1 is in* $C^1((0,t_0),X)\cap C((0,t_0),[\mathfrak{D}(A)])$.

Proof. We first show that the mild solution is locally Hölder continuous. Let us abbreviate $f(t) := F(u(t)) \in C([0,t_0),X)$ such that

$$u(t+\varepsilon)-u(t) = (T(\varepsilon)-1)u(t) + \int_0^\varepsilon T(\varepsilon-s)f(t+s)ds.$$

Now the last term is clearly Lipschitz continuous

$$\left\|\int_0^\varepsilon T(\varepsilon-s)F(u(t+s))ds\right\| \le \varepsilon M \max_{t\le s\le t+\varepsilon}\|f(s)\|.$$

The first term consists of two terms. For the first we obtain (using $T(\varepsilon)-1 = \int_0^\varepsilon AT(s)ds$)

$$\|(T(\varepsilon)-1)T(t)g\| \le \int_0^\varepsilon \|AT(s)T(t)g\|ds \le \varepsilon M\|AT(t)g\| \le \varepsilon\frac{MC}{t}$$

and for the second

$$\begin{aligned}\|(T(\varepsilon)-1)\int_0^t T(t-s)f(s)ds\| &\le \int_0^t\int_0^\varepsilon \|AT(t+r-s)f(s)\|dr\,ds\\ &\le C\big((t+\varepsilon)\log(t+\varepsilon)-t\log(t)-\varepsilon\log(\varepsilon)\big)\max_{0\le s\le t}\|f(s)\|\\ &\le 3C\big(-\varepsilon\log(\varepsilon)\big)\max_{0\le s\le t}\|f(s)\|.\end{aligned}$$

Now using Lipschitz continuity of F we see that $F(u(t))$ is locally Hölder continuous and the claim follows from Lemma 12.6. $\square$

Example 12.4. Let $X := C_{buc}(\mathbb{R})$ and consider the one-dimensional reaction-diffusion equation

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

The linear evolution was discussed in Example 11.15 and the necessary bound (12.4) can be easily obtained directly:

$$\|AT(t)g\|_\infty = \|\Phi_t'' * g\|_\infty \le \|\Phi_t''\|_1\|g\|_\infty \le \frac{\|\Phi_1''\|_1}{t}\|g\|_\infty.$$

Hence if $F : \mathbb{R} \to \mathbb{R}$ is locally Lipschitz, then Lemma 12.7 applies and we get existence of local strong solutions in X. Since the domain of the generator is a subset of C^2, these solutions are even classical solutions.

Note that we could also choose $X := C_0(\mathbb{R})$ (cf. Problem 11.30) if we additionally assume $F(0) = 0$ such that F maps X to X. In particular, this subspace is left invariant by the time evolution. ⋄

Finally, in order to understand the dynamics we note the following comparison principle (cf. Problem 6.17).

Theorem 12.8. *Let $U \subset \mathbb{R}^n$ be a bounded domain, $f \in C^1(\overline{U} \times \mathbb{R})$, $u, v \in C(\overline{U_T}) \cap C^{1;2}(U_T)$ and suppose*

$$\Delta u + f(x, u) - u_t \le \Delta v + f(x, v) - v_t. \tag{12.5}$$

Then if $u \ge v$ on Γ_T, we have $u \ge v$ on U_T. The same is true if $u, v \in C([0, T], L^2(U)) \cap C((0, T], [\mathfrak{D}(\bar{L})]) \cap C^1((0, T], L^2(U))$, where $\bar{L}$ is the Dirichlet Laplacian on U.

Proof. In this case $w := v - u$ satisfies

$$\Delta w + c(t, x) w - w_t \le 0,$$

where $c(t, x) := g(x, v(t, x), u(t, x))$ with

$$g(x, \xi, \eta) := \frac{f(x, \xi) - f(x, \eta)}{\xi - \eta} = \int_0^1 f_u(x, \eta + (\xi - \eta)s) ds \in C(\overline{U} \times \mathbb{R}^2).$$

Hence the claim follows from the maximum principle, more precisely the variant from Problem 6.16. Similarly, in the case of strong solutions use the maximum principle from Lemma 11.27. In this respect note that the proof of Lemma 11.27 still applies if c depends on t. □

Example 12.5. Consider again the one-dimensional reaction-diffusion equation from the previous example, but now on $X := C_0[0, 1]$ (cf. Example 11.14). Since the domain of the generator is a subset of C^2, these solutions are even classical solutions and our comparison principle applies. Let us look at two classical examples:

The **Fisher–Kolmogorov–Petrovsky–Piskunov equation** is given by

$$u_t = u_{xx} + r\, u(1 - u), \qquad r \in \mathbb{R},$$

and we can apply the comparison principle with $u_0(t, x) = 0$ and $u_1(t, x) = 1$ to show that solutions with $0 \le g \le 1$ remain in this region and hence are global. Moreover, using the solutions

$$u(t, x) = \frac{g_0}{g_0 + (1 - g_0)\mathrm{e}^{-rt}}$$

we see that if $0 \le g(x) \le M$, then $\limsup_{t \to \infty} u(t, x) \le 1$ for $r > 0$.

Similarly one can discuss the **Chafee–Infante problem**[1]

$$u_t = \Delta u + \lambda u - u^3$$

with Dirichlet boundary conditions by comparing the solution with the solutions of the ordinary differential equation

$$\dot{u} = \lambda u - u^3. \qquad \diamond$$

Problem 12.2. *Let L be self-adjoint with an orthonormal basis of eigenfunctions w_j corresponding to the eigenvalues E_j. For a complex-valued function F define*

$$F(L)g := \sum_{j=0}^{\infty} F(E_j)\langle w_j, g\rangle w_j.$$

Show

$$\|F(L)\| = \sup_{j\in\mathbb{N}_0} |F(E_j)|.$$

Problem 12.3. *Suppose A generates an analytic semigroup satisfying $\|T(t)\| \le M\mathrm{e}^{-\delta t}$ for some $\delta > 0$. Show that if f is globally Hölder continuous and converges to a constant $f_0 := \lim_{t\to\infty} f(t)$. Then the solution (11.26) of the inhomogeneous problem satisfies*

$$\lim_{t\to\infty} \dot{u}(t) = 0, \qquad \lim_{t\to\infty} u(t) = -A^{-1}f_0.$$

(Hint: $\|f(t) - f(s)\|^2 \le \|f(t) - f(s)\|(\|f(t) - f_0\| + \|f(s) - f_0\|)$.)

Problem 12.4. *Show that a differentiable semigroup satisfying*

$$\|AT(t)\| \le \frac{C}{t}, \qquad t > 0,$$

also satisfies

$$\|A^k T(t)\| \le \left(\frac{Ck}{t}\right)^k, \qquad t > 0,$$

and use this to conclude that T can be extended to an analytic function via

$$T(z) := \sum_{k=0}^{\infty} \frac{(z-t)^k}{k!}\frac{d^k}{dt^k}T(t), \qquad |z - t| < \frac{t}{\mathrm{e}C}.$$

Show that this extension still satisfies the semigroup property. (Hint: Problem 11.20.)

[1]Nathaniel Chafee (c. 1940), American mathematician
[1]Ettore Ferrari Infante (*1938), American mathematician

Chapter 13

Calculus of Variations

13.1. Differentiation in Banach spaces

Let X and Y be two Banach spaces and let U be an open subset of X. Denote by $C(U,Y)$ the set of continuous functions from $U \subseteq X$ to Y and by $\mathscr{L}(X,Y) \subset C(X,Y)$ the Banach space of bounded linear functions equipped with the operator norm

$$\|L\| := \sup_{\|u\|=1} \|Lu\|. \tag{13.1}$$

Then a function $F : U \to Y$ is called differentiable at $x \in U$ if there exists a linear function $dF(x) \in \mathscr{L}(X,Y)$ such that

$$F(x+u) = F(x) + dF(x)\, u + o(u), \tag{13.2}$$

where o, O are the **Landau symbols**. Explicitly

$$\lim_{u\to 0} \frac{\|F(x+u) - F(x) - dF(x)\, u\|}{\|u\|} = 0. \tag{13.3}$$

The linear map $dF(x)$ is called the **Fréchet derivative**[1] of F at x. It is uniquely defined since, if $dG(x)$ were another derivative, we had $(dF(x) - dG(x))u = o(u)$ implying that for every $\varepsilon > 0$ we can find a $\delta > 0$ such that $\|(dF(x) - dG(x))u\| \le \varepsilon\|u\|$ whenever $\|u\| \le \delta$. By homogeneity of the norm we conclude $\|dF(x) - dG(x)\| \le \varepsilon$ and since $\varepsilon > 0$ is arbitrary $dF(x) = dG(x)$. Note that for this argument to work it is crucial that we can approach x from arbitrary directions u, which explains our requirement that U should be open.

[1] Maurice Fréchet (1900-1980), French mathematician

If $I \subseteq \mathbb{R}$, we have an isomorphism $\mathscr{L}(I, X) \equiv X$ and if $F : I \to X$ we will write $\dot{F}(t)$ instead of $dF(t)$ if we regard $dF(t)$ as an element of X. Clearly this is consistent with the definition (11.1) from Section 11.1.

Example 13.1. Let X be a Hilbert space and consider $F : X \to \mathbb{R}$ given by $F(x) := \|x\|^2$. Then

$$F(x+u) = \langle x+u, x+u\rangle = \|x\|^2 + 2\mathrm{Re}\langle x, u\rangle + \|u\|^2 = F(x) + 2\mathrm{Re}\langle x, u\rangle + o(u).$$

Hence if X is a real Hilbert space, then F is differentiable with $dF(x)u = 2\langle x, u\rangle$. However, if X is a complex Hilbert space, then F is not differentiable. ◇

The previous example emphasizes that for $F : U \subseteq X \to Y$ it makes a big difference whether X is a real or a complex Banach space. In fact, in case of a complex Banach space X, we obtain a version of complex differentiability which of course is much stronger than real differentiability. Note that in this respect it makes no difference whether Y is real or complex.

Differentiability implies existence of directional derivatives

$$\delta F(x, u) := \lim_{\varepsilon \to 0} \frac{F(x + \varepsilon u) - F(x)}{\varepsilon}, \qquad \varepsilon \in \mathbb{R} \setminus \{0\}, \tag{13.4}$$

which are also known as **Gâteaux derivative**[2] or **variational derivative**. Indeed, if F is differentiable at x, then (13.2) implies

$$\delta F(x, u) = dF(x)u. \tag{13.5}$$

In particular, we call F Gâteaux differentiable at $x \in U$ if the limit on the right-hand side in (13.4) exists for all $u \in X$. However, note that Gâteaux differentiability does not imply differentiability. In fact, the Gâteaux derivative might be unbounded or it might even fail to be linear in u. Some authors require the Gâteaux derivative to be a bounded linear operator and in this case we will write $\delta F(x, u) = \delta F(x)u$. But even this additional requirement does not imply differentiability in general. Note that in any case the Gâteaux derivative is homogenous, that is, if $\delta F(x, u)$ exists, then $\delta F(x, \lambda u)$ exists for every $\lambda \in \mathbb{R}$ and

$$\delta F(x, \lambda u) = \lambda\, \delta F(x, u), \qquad \lambda \in \mathbb{R}. \tag{13.6}$$

Example 13.2. The function $F : \mathbb{R}^2 \to \mathbb{R}$ given by $F(x, y) := \frac{x^3}{x^2+y^2}$ for $(x, y) \neq 0$ and $F(0, 0) = 0$ is Gâteaux differentiable at 0 with Gâteaux derivative

$$\delta F(0, (u, v)) = \lim_{\varepsilon \to 0} \frac{F(\varepsilon u, \varepsilon v)}{\varepsilon} = F(u, v),$$

which is clearly nonlinear.

[2] René Gâteaux (1889–1914), French mathematician

The function $F : \mathbb{R}^2 \to \mathbb{R}$ given by $F(x,y) = x$ for $y = x^2$ and $F(x,y) :=$ 0 else is Gâteaux differentiable at 0 with Gâteaux derivative $\delta F(0) = 0$, which is clearly linear. However, F is not differentiable.

If you take a linear function $L : X \to Y$ which is unbounded, then L is everywhere Gâteaux differentiable with derivative equal to Lu, which is linear but, by construction, not bounded. $\diamond$

Example 13.3. Consider $L^p(U)$, $1 \le p < \infty$, and let $G : \mathbb{C} \to \mathbb{R}$ be (real) differentiable with

$$|G(z)| \le C|z|^p, \qquad \sqrt{|\partial_x G(z)|^2 + |\partial_y G(z)|^2} \le C|z|^{p-1}, \qquad z = x + \mathrm{i}y,$$

or, if U is bounded,

$$|G(z)| \le C(1 + |z|^p), \qquad \sqrt{|\partial_x G(z)|^2 + |\partial_y G(z)|^2} \le C(1 + |z|^{p-1}).$$

Note that the estimate for G (with a possibly larger constant) comes for free from the one for the derivatives in the bounded case and also in the general case if $G(0) = 0$. We only consider the first case and leave the easy adaptions for the second case as an exercise.

Then

$$N(f) := \int_U G(f) d^n x$$

is Gâteaux differentiable and we have

$$\delta N(f) g = \int_U \big((\partial_x G)(f)\mathrm{Re}(g) + (\partial_y G)(f)\mathrm{Im}(g)\big) d^n x.$$

In fact, by the chain rule $h(\varepsilon) := G(f + \varepsilon g)$ is differentiable with $h'(0) = (\partial_x G)(f)\mathrm{Re}(g) + (\partial_y G)(f)\mathrm{Im}(g)$. Moreover, by the mean value theorem

$$\begin{aligned}\left|\frac{h(\varepsilon) - h(0)}{\varepsilon}\right| &\le \sup_{0 \le \tau \le \varepsilon} \sqrt{(\partial_x G)(f + \tau g)^2 + (\partial_y G)(f + \tau g)^2}\,|g| \\ &\le C \sup_{0 \le \tau \le \varepsilon} |f + \tau g|^{p-1}|g| \le C 2^{p-1}(|f|^{p-1} + |g|^{p-1})|g|\end{aligned}$$

and hence we can invoke dominated convergence to interchange differentiation and integration. Note that using Hölder's inequality this last estimate also shows Lipschitz continuity on bounded sets:

$$|N(f) - N(g)| \le C(\|f\|_p + \|g\|_p)^{p-1}\|f - g\|_p.$$

In particular, for $1 < p < \infty$ the norm

$$N(f) := \int_U |f|^p d^n x$$

is Gâteaux differentiable with

$$\delta N(f) g = p \int_U |f|^{p-2}\mathrm{Re}(f g^*) d^n x. \qquad \diamond$$

Problem 13.1. *Let X be a Hilbert space and $A : \mathfrak{D}(A) \subseteq X \to X$ a (densely defined) symmetric operator. Show that if*

$$\lambda_0 := \inf_{u \in \mathfrak{D}(A): \|u\|=1} \langle u, Au \rangle$$

is attained for some $u_0 \in \mathfrak{D}(A)$, then u_0 is an eigenvector corresponding to the eigenvalue λ_0.

13.2. The direct method

We already know that the Dirichlet principle (Section 5.5) allows us to cast certain elliptic partial differential equations as a minimization problem. In this chapter we want to pursue this idea further.

We start by looking at the abstract problem of minimizing a nonlinear functional $F : M \subseteq X \to \mathbb{R}$, where X is some Banach space and M some closed subset. If M is compact and F is continuous, then we can proceed as in the finite-dimensional case to show that there is a minimizer: Start with a sequence x_n such that $F(x_n) \to \inf_M F$. By compactness we can assume that $x_n \to x_0$ after passing to a subsequence and by continuity $F(x_n) \to F(x_0) = \inf_M F$.

In the finite dimensional case compactness will follow from boundedness by the Heine–Borel theorem. In the infinite dimensional case this breakes down and the remedy is to switch to weak convergence and use a variant of the Banach–Alaoglu theorem (Theorem B.25). The only problem with this cure is, that, since there are more weakly than strongly convergent subsequences, weak (sequential) continuity is in fact a stronger property than just continuity!

Example 13.4. It is well-known that in general not even the norm is weakly (sequentially) continuous. In fact, in a Hilbert space any infinite orthonormal set will converge weakly to 0. On the other hand, the norm in a Banach space is at least weakly sequentially lower semicontinuous (Lemma B.24 (ii)) and this is still good enough for the above argument to work.

Finally, note that this problem does not occur for linear maps, since a linear functional is continuous precisely if it is weakly continuous by the very definition of weak convergence. $\diamond$

The previous example shows that weak continuity is too much to hope for and hence we will use lower semicontinuity instead. To this end recall that in a topological space X a function $f : X \to \overline{\mathbb{R}}$ is **sequentially lower semicontinuous** if

$$\liminf_{n\to\infty} f(x_n) \geq f(x_0), \qquad x_n \to x_0,\ x_0 \in X. \tag{13.7}$$

Now we are ready to show what is frequently referred to as the **direct method in the calculus of variations** due to Zaremba[3] and Hilbert:

Theorem 13.1 (Variational principle)**.** *Let X be a reflexive Banach space and let $F : M \subseteq X \to (-\infty, \infty]$. Suppose M is nonempty, weakly sequentially closed and that either F is **weakly coercive**, that is $F(x) \to \infty$ whenever $\|x\| \to \infty$, or that M is bounded. Then, if F is weakly sequentially lower semicontinuous, there exists some $x_0 \in M$ with $F(x_0) = \inf_M F$.*

If F is Gâteaux differentiable, then

$$\delta F(x_0, u) = 0 \tag{13.8}$$

for every $u \in X$ with $x_0 + \varepsilon u \in M$ for sufficiently small ε.

Proof. Without loss of generality we can assume $F(x) < \infty$ for some $x \in M$. As above we start with a sequence $x_n \in M$ such that $F(x_n) \to \inf_M F < \infty$. If M is unbounded, then the fact that F is coercive implies that x_n is bounded. Otherwise, if M is bounded, it is obviously bounded. Hence by Theorem B.25 we can pass to a subsequence such that $x_n \rightharpoonup x_0$ with $x_0 \in M$ since M is assumed sequentially closed. Now, since F is weakly sequentially lower semicontinuous, we finally get $\inf_M F = \lim_{n\to\infty} F(x_n) = \liminf_{n\to\infty} F(x_n) \geq F(x_0)$. $\square$

Note that looking for a maximum of F is the same as looking for a minimum of $-F$. However, in this case lower semicontinuous turns into upper semicontinuous, so the conditions on F are different in this case.

Of course in a metric space the definition of closedness in terms of sequences agrees with the corresponding topological definition. In the present situation sequentially weakly closed implies (sequentially) closed and the converse holds at least for convex sets.

Lemma 13.2. *Suppose $M \subseteq X$ is convex. Then M is closed if and only if it is sequentially weakly closed.*

Proof. Suppose M is closed and let x be in the weak sequential closure of M, that is, there is a sequence $x_n \rightharpoonup x$. If $x \notin M$, then by the geometric Hahn–Banach theorem (cf. Corollary 6.4 from [**35**]) we can find a linear functional ℓ which separates $\{x\}$ and M: $\mathrm{Re}(\ell(x)) < c \leq \mathrm{Re}(\ell(y))$, $y \in M$. But this contradicts $\mathrm{Re}(\ell(x)) < c \leq \mathrm{Re}(\ell(x_n)) \to \mathrm{Re}(\ell(x))$. $\square$

Similarly, the same is true with lower semicontinuity. In fact, a slightly weaker assumption suffices. Let X be a vector space and $M \subseteq X$ a convex subset. A function $F : M \to \overline{\mathbb{R}}$ is called **quasiconvex** if

$$F(\lambda x + (1-\lambda)y) \leq \max\{F(x), F(y)\}, \qquad \lambda \in (0,1), \quad x, y \in M. \tag{13.9}$$

[3]Stanisław Zaremba (1863–1942), Polish mathematician and engineer

It is called **strictly quasiconvex** if the inequality is strict for $x \neq y$. By $\lambda F(x) + (1-\lambda)F(y) \leq \max\{F(x), F(y)\}$ every (strictly) convex function is (strictly) quasiconvex. The converse is not true as the following example shows.

Example 13.5. Every (strictly) monotone function on $\mathbb{R}$ is (strictly) quasiconvex. Moreover, the same is true for symmetric functions which are (strictly) monotone on $[0,\infty)$. Hence the function $F(x) = \sqrt{|x|}$ is strictly quasiconvex. But it is clearly not convex on $M = \mathbb{R}$.

Note however that, in contradistinction to convex functions, the sum of quasiconvex functions is in general not quasiconvex. $\diamond$

Note that we can extend a (quasi-)convex function $F : M \to \overline{\mathbb{R}}$ to all of X by setting $F(x) = \infty$ for $x \in X \setminus M$ and the resulting function will still be (quasi-)convex and will have the same infimum.

Now we are ready for the next

Lemma 13.3. *Suppose $M \subseteq X$ is a closed convex set and suppose $F : M \to \overline{\mathbb{R}}$ is quasiconvex. Then F is weakly sequentially lower semicontinuous if and only if it is (sequentially) lower semicontinuous.*

Proof. Suppose F is lower semicontinuous. If it were not weakly sequentially lower semicontinuous we could find a sequence $x_n \rightharpoonup x_0$ with $F(x_n) \leq a < F(x_0)$. Then $x_n \in F^{-1}((-\infty, a])$ implying $x_0 \in F^{-1}((-\infty, a])$ as this set is convex (Problem 13.4) and closed (Problem 13.2). But this gives the contradiction $a < F(x_0) \leq a$. $\square$

Example 13.6. Let $U \subseteq \mathbb{R}^n$ and $K : U \times \mathbb{C} \to [0,\infty)$ measurable. Suppose $u \mapsto K(x,u)$ is convex and continuous for fixed $x \in U$. Then

$$F(u) := \int_U K(x, u(x)) d^n x$$

is weakly sequentially lower semicontinuous on $L^p(U)$ for $1 \leq p \leq \infty$. Since F is convex, it suffices to show lower semicontinuity. Assume the contrary, then we can find some $u \in L^p$ and a sequence $u_n \to u$ such that $F(u) > \liminf F(u_n)$. After passing to a subsequence we can assume that $u_n(x) \to u(x)$ a.e. and hence $K(x, u_n(x)) \to K(x, u(x))$ a.e. Finally applying Fatou's lemma (Theorem 2.4 from **[34]**) gives the contradiction $F(u) \leq \liminf F(u_n)$.

Moreover, if $u \mapsto K(x,u)$ is strictly convex for a.e. $x \in U$, then F is strictly convex. Indeed, in this case

$$\begin{aligned} F(\lambda u + (1-\lambda)v) &= \int_U K(x, \lambda u(x) + (1-\lambda)v(x)) d^n x \\ &\leq \int_U \big(\lambda K(x, u(x)) + (1-\lambda)K(x, v(x))\big) d^n x = \lambda F(u) + (1-\lambda)F(v) \end{aligned}$$

and equality would imply $K(x, \lambda u(x) + (1-\lambda)v(x)) = K(x, \lambda u(x) + (1-\lambda)v(x))$ for a.e. x and hence $u(x) = v(x)$ for a.e. x.

Note that this result generalizes to $\mathbb{C}^n$-valued functions in a straightforward manner. $\diamond$

Moreover, in this case our variational principle reads as follows:

Corollary 13.4. *Let X be a reflexive Banach space and let M be a nonempty closed convex subset. If $F : M \subseteq X \to \overline{\mathbb{R}}$ is quasiconvex, lower semicontinuous, and, if M is unbounded, weakly coercive, then there exists some $x_0 \in M$ with $F(x_0) = \inf_M F$. If F is strictly quasiconvex then x_0 is unique.*

Proof. It remains to show uniqueness. Let x_0 and x_1 be two different minima. Then $F(\lambda x_0 + (1-\lambda)x_1) < \max\{F(x_0), F(x_1)\} = \inf_M F$, a contradiction. $\square$

Of course the first test for our results will be the Poisson problem.

Example 13.7. By the Dirichlet principle (Section 5.5) the solution of the **Poisson problem**

$$-\Delta u = f$$

in a bounded domain $U \subset \mathbb{R}^n$ attaining given boundary values g on ∂U can be found by minimizing the functional

$$F(u) := \int_U \left(\frac{1}{2} |\nabla u|^2 - u f \right) d^n x$$

on $H^1(U, \mathbb{R})$. To incorporate the boundary values we introduce

$$M := \{v \in H^1(U, \mathbb{R}) | v|_{\partial U} = g\}.$$

Here the equality $v|_{\partial U} = g$ has to be understood in the sense of traces and hence we need to require U to have a C^1 boundary such that the trace operator is well-defined. Moreover, we assume $f \in L^2(U, \mathbb{R})$ and g in the range of the trace operator, such that M is nonempty. In particular, there is some $\bar{g} \in H^1(U, \mathbb{R})$ with $\bar{g}|_{\partial U} = g$. See the discussion after (10.27) for conditions that such an extension exists. By continuity of the trace operator, M is closed and convexity is obvious.

Similarly, convexity of F is obvious since the first integrand is convex and the second is linear. Also F is continuous (note that this would still hold true if we replace f by some element from $H^1(U, \mathbb{R})^*$). Moreover, since taking the square is strictly convex, we see that

$$\int_U \big(\lambda |\nabla u|^2 + (1-\lambda)|\nabla v|^2 - |\nabla(\lambda u + (1-\lambda)v)|^2 \big) d^n x \geq 0$$

and equality would imply $\nabla u(x) = \nabla v(x)$ for a.e. $x \in U$. If $u, v \in M$ we can further conclude $u = v$ a.e. and hence F is strictly convex on M.

To see that F is weakly coercive, let $u = \bar{g} + v$, where $v \in H_0^1(U)$ vanishes on the boundary, then

$$F(u) \geq \frac{1}{2}\|\nabla v\|_2^2 - \|\nabla \bar{g}\|_2 \|\nabla v\|_2 - \|f\|_2 \|v\|_2 - C,$$

with C depending on f and g only. Now the Poincaré inequality (Theorem 9.34) $\|v\|_2 \leq C_0 \|\nabla v\|_2$ implies that $F(u) \to \infty$ if $\|v\|_{1,2} \to \infty$. Finally, since F is convex and continuous, Corollary 13.4 implies existence of a unique minimizer.

Finally, F is Gâteaux differentiable with

$$\delta F(u)v = \int_U (\nabla u \cdot \nabla v - f v)\, d^n x$$

and we have $u + \varepsilon v \in M$ whenever $u \in M$ and $v \in H_0^1(U, \mathbb{R})$. Hence this minimizer solves the weak formulation of our boundary value problem. $\diamond$

While the previous example just reproduces what we were already able to show in Section 10.1, the next example shows how to handle a nonlinear problem, where our linear theory from Section 10.1 fails.

Example 13.8. Let us consider the following nonlinear elliptic problem

$$-\Delta u + u|u| + u = f$$

in $L^2(\mathbb{R}^n, \mathbb{R})$ for a given function $f \in L^2(\mathbb{R}^n, \mathbb{R})$. We are going to look for weak solutions, that is, solutions $u \in H^1(\mathbb{R}^n, \mathbb{R})$ satisfying

$$\int_{\mathbb{R}^n} (\nabla u \cdot \nabla \phi + (|u|u + u - f)\phi)\, d^n x = 0, \qquad \phi \in C_c^\infty(\mathbb{R}^n, \mathbb{R}).$$

We start by introducing the functional

$$F(u) := \int_{\mathbb{R}^n} \left(\frac{1}{2}|\nabla u|^2 + \frac{1}{3}|u|^3 + \frac{1}{2}u^2 - uf \right) d^n x$$

on $X := L^2(\mathbb{R}^n, \mathbb{R})$ and set $F(u) = \infty$ if $u \notin H^1(\mathbb{R}^n, \mathbb{R}) \cap L^3(\mathbb{R}^n, \mathbb{R})$. We also choose $M := X$. One checks that for $u \in H^1(\mathbb{R}^n, \mathbb{R}) \cap L^3(\mathbb{R}^n, \mathbb{R})$ and $\phi \in C_c^\infty(\mathbb{R}^n, \mathbb{R})$ this functional has a variational derivative

$$\delta F(u, \phi) = \int_{\mathbb{R}^n} (\nabla u \cdot \nabla \phi + (|u|u + u - f)\phi)\, d^n x = 0$$

which coincides with the weak formulation of our problem. Hence a minimizer (which is necessarily in $H^1(\mathbb{R}^n, \mathbb{R}) \cap L^3(\mathbb{R}^n, \mathbb{R})$) is a weak solution of our nonlinear elliptic problem and it remains to show existence of a unique minimizer.

First of all note that

$$F(u) \geq \frac{1}{2}\|u\|_2^2 - \|u\|_2 \|f\|_2 \geq \frac{1}{4}\|u\|_2^2 - \|f\|_2^2$$

and hence F is coercive. To see that it is weakly sequentially lower continuous, observe that for the first term this follows from strict convexity (as in the previous example), for the second term this follows from Example 13.6 and the last two are easy. Hence we get existence of a unique minimizer from Corollary 13.4.

It is also interesting to discuss possible extensions of this example: First of all we could replace $X = L^2$ by $X = H^1$. The only difference is that the argument that F is coercive needs to be adapted. In fact, we could replace the linear part by an arbitrary elliptic operator, as long as we make sure it is coercive. Moreover, we can also choose a different nonlinearity as long as it is quasiconvex and nonnegative such that Example 13.6 applies. We could even include first order derivatives in the nonlinearity (cf. Problem 13.5). Also replacing $\mathbb{R}^n$ by U does not impose any principal problems. $\diamond$

Problem 13.2. *Let X be a topological space. A function $f : X \to \bar{\mathbb{R}}$ is called **lower semicontinuous** if $f^{-1}((a,\infty])$ is open for every $a \in \mathbb{R}$. Show that a lower semicontinuous is sequentially lower semicontinuous and the converse holds if X is a metric space.*

Problem 13.3. *Show that $F : M \to \overline{\mathbb{R}}$ is convex if and only if its **epigraph** $\operatorname{epi} F := \{(x,a) \in M \times \overline{\mathbb{R}} | F(x) \le a\} \subset X \times \overline{\mathbb{R}}$ is convex.*

Problem* 13.4. *Show that $F : M \to \overline{\mathbb{R}}$ is quasiconvex if and only if the sublevel sets $F^{-1}((-\infty,a])$ are convex for every $a \in \mathbb{R}$.*

Problem 13.5. *Let $U \subseteq \mathbb{R}^n$ and $K : U \times \mathbb{C} \times \mathbb{C}^n \to [0,\infty)$ measurable. Suppose $(u,v) \mapsto K(x,u,v)$ is convex and continuous for fixed $x \in U$. Show that*

$$F(u) := \int_U K(x, u(x), \nabla u(x)) d^n x$$

is weakly sequentially lower semicontinuous on $W^{1,p}(U)$ for $1 \le p \le \infty$.

Problem 13.6. *Let $U \subseteq \mathbb{R}^n$ be a bounded domain with a C^1 boundary. Let $\tilde{L}$ be an elliptic operator in divergence form with $A, c \in L^\infty$ and $b = 0$, $c \ge 0$. Establish existence of weak solutions in $H^1(U,\mathbb{R})$ for*

$$\bar{L}u = f, \qquad u\big|_{\partial U} = g.$$

(Compare Problem 5.36.)

13.3. Constraints

If we look at Example 13.8 in the case $f = 0$, our approach will only give us the trivial solution. In fact, for a linear problem one has nontrivial solutions for the homogenous problem only at an eigenvalue. Since the Laplace operator has no eigenvalues on $\mathbb{R}^n$ (as is not hard to see using the Fourier transform), we look at a bounded domain U instead. To avoid the trivial solution we will add a constraint. Of course the natural constraint is to require admissible elements to be normalized. However, since the unit sphere is not weakly closed (one can show that its weak closure is the unit ball — see Example 6.10 from **[35]**), we cannot simply add this requirement to M. To overcome this problem we will use that another way of getting weak sequential closedness is via compactness:

Lemma 13.5. *Let X, Y be Banach spaces such that X is compactly embedded into Y and let $N : Y \to \mathbb{R}$ be continuous. Then $M := \{x \in X | N(x) = N_0\} \subseteq X$ is weakly sequentially closed for any $N_0 \in \mathbb{R}$. The same holds for $M := \{x \in X | N(x) \le N_0\}$.*

Proof. This follows from Theorem B.26 since every weakly convergent sequence in X is convergent in Y. □

Theorem 13.6 (Variational principle with constraints)**.** *Let X be a reflexive Banach space and let $F : X \to \mathbb{R}$ be weakly sequentially lower semicontinuous and weakly coercive. Let Y be another Banach space such that X is compactly embedded into Y and let $N : Y \to \mathbb{R}$ be continuous. Fix $N_0 \in \mathbb{R}$ and suppose that $M := \{x \in X | N(x) = N_0\}$ is nonempty. Then there exists some $x_0 \in M$ with $F(x_0) = \inf_M F$.*

*If in addition F and N are Gâteaux differentiable and δN does not vanish on M, then there is a constant $\lambda \in \mathbb{R}$ (the **Lagrange multiplier**) such that*

$$\delta F(x_0) = \lambda\, \delta N(x_0). \tag{13.10}$$

Proof. Existence follows from Theorem 13.1 which is applicable thanks to our previous lemma. Now choose some $x_1 \in X$ such that $\delta N(x_0)x_1 \neq 0$ and $x \in X$ arbitrary. Then the function

$$f(t, s) := N(x_0 + t\,x + x_1 s)$$

is $C^1(\mathbb{R}^2)$ and satisfies

$$\partial_t f(t, s) = \delta N(x_0 + t\,x + s\,x_1)x, \qquad \partial_s f(t, s) = \delta N(x_0 + t\,x + x_1 s)x_1$$

and since $\partial_s f(0, 0) \neq 0$ the implicit function theorem implies existence of a function $\sigma \in C^1(-\varepsilon, \varepsilon)$ such that $\sigma(0) = 0$ and $f(t, \sigma(t)) = f(0, 0)$, that is, $x(t) := x_0 + t\,x + \sigma(t)x_1 \in M$ for $|t| < \varepsilon$. Moreover,

$$\sigma'(0) = -\frac{\partial_t f(0, 0)}{\partial_s f(0, 0)} = -\frac{\delta N(x_0)x}{\delta N(x_0)x_1}.$$

Hence, as before together with the chain rule

$$\frac{d}{dt} F(x_0 + t\,x + \sigma(t)x_1)|_{t=0} = \delta F(x_0)(x + \sigma'(0)x_1) = \delta F(x_0)x - \lambda \delta N(x_0)x = 0,$$

where

$$\lambda := \frac{\delta F(x_0)x_1}{\delta N(x_0)x_1}. \qquad \square$$

Example 13.9. Let $U \subset \mathbb{R}^n$ be a bounded domain and consider

$$F(u) := \frac{1}{2}\int_U |\nabla u|^2 d^n x, \qquad u \in H_0^1(U, \mathbb{R})$$

subject to the constraint

$$N(u) := \int_U G(u) d^n x = N_0,$$

where $G : \mathbb{R} \to \mathbb{R}$ is differentiable and satisfies

$$|G'(x)| \le C(1 + |x|).$$

This condition implies

$$|G(x)| \le \tilde{C}(1 + |x|^2)$$

and ensures that $N(u)$ is well-defined for all $u \in L^2(U, \mathbb{R})$.

In order to apply the theorem we set $X := H_0^1(U, \mathbb{R})$ and $Y := L^2(U, \mathbb{R})$. That X is compactly embedded into Y is the Rellich–Kondrachov theorem (Theorem 9.31). Moreover, by the Poincaré inequality (Theorem 9.34) we can choose $\|x\|^2 := F(x)$ as an equivalent norm on X. In particular, F satisfies the requirements of our theorem and so does N by Example 13.3. Consequently, if N_0 is such that

$$M := \{u \in X | N(u) = N_0\}$$

is nonempty, there is a minimizer u_0. By Example 13.1 and Example 13.3

$$dF(u_0)u = \int_U (\nabla u_0)(\nabla u)\, d^n x, \quad \delta N(u_0)u = \int_U G'(u_0)u\, d^n x$$

and if we can find some $u \in H_0^1(U)$ such that this derivative is nonzero, then u_0 satisfies

$$\int_U \big(\nabla u_0 \cdot \nabla u - \lambda G'(u_0)u\big) d^n x = 0, \qquad u \in H_0^1(U, \mathbb{R}),$$

and hence is a weak solution of the nonlinear eigenvalue problem

$$-\Delta u_0 = \lambda G'(u_0).$$

Note that this last condition is for example satisfied if $G(0) = 0$, $G'(x)x > 0$ for $x \neq 0$, and $N_0 > 0$. Indeed, in this case $\delta N(u_0)u_0 = \int_U G'(u_0)u_0 d^n x > 0$ since otherwise we would have $u_0 = 0$ contradicting $0 < N_0 = N(u_0) = N(0) = 0$.

Of course in the case $G(x) = \frac{1}{2}|x|^2$ and $N_0 = 1$ this gives us the lowest eigenvalue of the Laplacian on U with Dirichlet boundary conditions.

Note that using continuous embeddings $L^2 \hookrightarrow L^p$ with $2 \le p \le \infty$ for $n = 1$, $2 \le p < \infty$ for $n = 2$, and $2 \le p \le \frac{2n}{n-2}$ for $n \ge 3$ one can improve this result to the case

$$|G'(x)| \le C(1 + |x|^{p-1}). \qquad \diamond$$

Problem 13.7. *Extend Example 13.9 to the case*

$$F(u) := \frac{1}{2}\int_U |\nabla u|^2 d^n x + \int_U V(x)|u|^2 d^n x, \qquad u \in H_0^1(U, \mathbb{R}),$$

where $V \in L^q(U)$ is nonnegative with $q > \frac{n}{2}$ and $n \ge 2$ and V. (Hint: Theorem 9.31.)

Chapter 14

The nonlinear Schrödinger equation

The purpose of this chapter is to investigate a prototypical example, the initial value problem for the **nonlinear Schrödinger equation** (NLS)

$$\mathrm{i}u_t + \Delta u = \pm |u|^{\alpha-1} u, \qquad u(0) = g. \tag{14.1}$$

The two cases $-$ and $+$ are known as **focusing** and **defocusing**, respectively. Of particular importance in applications are the cubic ($\alpha = 3$) and quintic ($\alpha = 5$) case. Note that if u is a solution, then so will be $v(t,x) = u(-t,x)^*$ and hence it suffices to look at positive times only.

14.1. Local well-posedness in H^r for $r > \frac{n}{2}$

Equation (14.1) is a semilinear equation of the type considered in Section 12.1 and hence we need to look at the linear **Schrödinger equation**

$$\mathrm{i}u_t + \Delta u = 0, \qquad u(0) = g \tag{14.2}$$

first. We recall that the solution for $g \in H^2(\mathbb{R}^n)$ can be obtained using the Fourier transform and is given by

$$u(t) = T_S(t)g, \qquad T_S(t) = \mathcal{F}^{-1} \mathrm{e}^{-\mathrm{i}|p|^2 t} \mathcal{F}. \tag{14.3}$$

Note that $T_S(t) : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ is a unitary operator (since $|\mathrm{e}^{-\mathrm{i}|p|^2 t}| = 1$):

$$\|u(t)\|_2 = \|g\|_2. \tag{14.4}$$

In fact, we even have that $T_S(t) : H^r(\mathbb{R}^n) \to H^r(\mathbb{R}^n)$ is unitary.

Theorem 14.1. *The family $T_S(t)$ is a C_0-group in $H^r(\mathbb{R}^n)$ whose generator is $\mathrm{i}\Delta$, $\mathfrak{D}(\mathrm{i}\Delta) = H^{r+2}(\mathbb{R}^n)$.*

Note that we have

$$(T_S(t)g)^* = T_S(-t)g^*, \qquad g \in L^2(\mathbb{R}^n), \tag{14.5}$$

and

$$\partial_j T_S(t)g = T_S(t)\partial_j g, \qquad 1 \le j \le n,\ g \in H^1(\mathbb{R}^n). \tag{14.6}$$

Next we turn to the nonlinear Schrödinger equation. If we assume that $u, |u|^{\alpha-1}u \in C([0,T], L^2(\mathbb{R}^n))$ we can use Duhamel's formula to rewrite the nonlinear Schrödinger equation as

$$u(t) = T_S(t)g \mp \mathrm{i} \int_0^t T_S(t-s)|u(s)|^{\alpha-1}u(s)ds \tag{14.7}$$

just as we did in Section 12.1. In order to apply our theory, we need that the nonlinearity $F(u) = \mp\mathrm{i}|u|^{\alpha-1}u$ is Lipschitz on X. Clearly for $X = L^2(\mathbb{R}^n)$ this will not be the case, as the image of a square integrable function will not be square integrable. However, the key observation is that for $r > \frac{n}{2}$ the space $H^r(\mathbb{R}^n)$ is a Banach algebra (Lemma 8.9) and hence, if we assume our nonlinearity to be of the form $F(u) = \mp\mathrm{i}|u|^{\alpha-1}u$ with $\alpha - 1 = 2k$ where $k \in \mathbb{N}$, then $F : H^r(\mathbb{R}^n) \to H^r(\mathbb{R}^n)$ is Lipschitz on bounded sets since

$$F(u) - F(v) = u^{k+1}Q_{k-1}(u^*, v^*)(u-v)^* + (v^*)^k Q_k(u,v)(u-v), \tag{14.8}$$

where $Q_k(x,y) = \mp\mathrm{i}\sum_{j=0}^k x^{k-j}y^j$. Another algebra which is natural in this context is the Wiener algebra.

$$\mathcal{A}(\mathbb{R}^n) := \{\hat{f} | f \in L^1(\mathbb{R}^n)\}, \qquad \|\hat{f}\|_{\mathcal{A}} := \|f\|_1. \tag{14.9}$$

Just as with $H^r(\mathbb{R}^n)$, the Schrödinger group T_S leaves $\mathcal{A}(\mathbb{R}^n)$ invariant and preserves its norm. Note that we have $H^r(\mathbb{R}^n) \subset \mathcal{A}(\mathbb{R}^n)$ for $r > \frac{n}{2}$ since $(1+|p|^2)^{-r} \in L^2(\mathbb{R}^n)$ for such r. The embedding being continuous, $\|f\|_{\mathcal{A}} \le \|(1+|.|^2)^{-r}\|_2 \|f\|_{H^r}$.

Hence Theorem 12.1 applies and we get:

Theorem 14.2. *Let $\alpha = 2k+1$ be an odd integer and $X = H^r(\mathbb{R}^n)$ for $r > \frac{n}{2}$ or $X = \mathcal{A}(\mathbb{R}^n)$. Then for every $g \in X$ there is a $t_0 = t_0(\|g\|) > 0$, such that there is a unique solution $u \in C([-t_0, t_0], X)$ of (14.7). Moreover, the solution map $g \mapsto u(t)$ will be Lipschitz continuous from every ball $\|g\| \le \rho$ to $C([-t_0(\rho), t_0(\rho)], X)$.*

Note that the mild solution will be a strong solution for $g \in H^{r+2}$ since $F : H^{r+2} \to H^{r+2}$ is Lipschitz continuous on bounded sets. Moreover, for each initial condition there is a maximal solution and Lemma 12.4 implies:

Lemma 14.3. *This solution exists on a maximal time interval $(t_-(g), t_+(g))$ and if $|t_\pm(g)| < \infty$ we must have $\liminf_{t \to t_\pm(g)} \|u(t)\| = \infty$.*

An interesting observation is that the maximal existence time does not depend on r. This is known as persistence of regularity:

Lemma 14.4. *Let $g \in H^r(\mathbb{R}^n)$ with $r > \frac{n}{2}$ or $g \in \mathcal{A}(\mathbb{R}^n)$. Let $t_{+,r}(g)$, $t_{+,\mathcal{A}}(g)$ be the maximal existence time of the solution with initial condition g with respect to these cases. Then $t_{+,r}(g) = t_{+,\mathcal{A}}(g)$.*

Proof. Using

$$\|fg\|_{H^r} \le C_{n,r}\big(\|f\|_{H^r}\|g\|_{\mathcal{A}} + \|f\|_{\mathcal{A}}\|g\|_{H^r}\big)$$

from Lemma 8.9 recursively we obtain $\||u|^{\alpha-1}u\|_{H^r} \le C\|u\|_{\mathcal{A}}^{\alpha-1}\|u\|_{H^r}$. Now the claim follows from Lemma 12.5. □

Of course up to this point we can replace the nonlinearity by an arbitrary polynomial in u and u^*. In fact, it is even possible to replace the nonlinearity by a (sufficiently smooth) function, but in this case the required Lipschitz estimate is more tedious to derive since we cannot just simply rely on the algebra structure.

In order to get global solutions the following conservation laws will be crucial: **Mass**

$$M(t) := \frac{1}{2}\|u(t)\|_2^2$$

and **energy**

$$E(t) := \frac{1}{2}\|\nabla u(t)\|_2^2 \pm \frac{1}{\alpha+1}\|u(t)\|_{\alpha+1}^{\alpha+1}.$$

Lemma 14.5. *Let $r > \frac{n}{2}$ and $g \in H^r(\mathbb{R}^n)$. Then $M(t) = M(0)$ for all $t \in (t_-(g), t_+(g))$. If in addition, $r \ge 1$ then also $E(t) = E(0)$ for all $t \in (t_-(g), t_+(g))$.*

Proof. If u is a sufficiently smooth solution this can be verified directly (Problem 14.3). For the general case approximate by smooth solutions (using local Lipschitz continuity of the solution map) and conclude that $M(t)$ is locally constant and hence constant on its interval of existence. Similarly for $E(t)$. □

So in the focusing case we get existence of global solutions in H^1 if $n = 1$ such that our local results holds for $r = 1$. In the defocusing case the energy is not positive and we cannot immediately control the H^1 norm using E and M.

Problem 14.1. *Find a plane wave solution*

$$u(t,x) = A\,\mathrm{e}^{\mathrm{i}(k\cdot x + ct)}, \qquad A \in \mathbb{C},\ k \in \mathbb{R}^n,\ c \in \mathbb{R},$$

of the NLS equation.

Problem 14.2. *Show that if $u(t,x)$ is a strong solution of the NLS equation, then so are*

$$\lambda^{-2/(\alpha-1)} u\big(\frac{t}{\lambda^2}, \frac{x}{\lambda}\big), \qquad \lambda > 0,$$

and

$$\mathrm{e}^{\mathrm{i}v\cdot x - \mathrm{i}t|v|^2/2} u(t, x - vt), \qquad v \in \mathbb{R}^n.$$

Moreover, show that for $\alpha = 1 + \frac{4}{n}$ also

$$(\mathrm{i}t)^{-n/2} \mathrm{e}^{\mathrm{i}|x|^2/(2t)} u\big(\frac{1}{t}, \frac{x}{t}\big)^*, \qquad t \neq 0,$$

is a solution.

Problem 14.3. *Let $u \in C([-t_0,t_0], H^{r+2}(\mathbb{R}^n)) \cap C^1([-t_0,t_0], H^r(\mathbb{R}^n))$ be a strong solution of the NLS equation (with $r > \frac{n}{2}$). Show that mass and energy are independent of $t \in [-t_0, t_0]$.*

14.2. Global solutions and blowup in H^1 in one dimension

In the previous section we have seen that in the defocusing case we get global solutions in H^1. In this section we want to have a closer look at the defocusing case. For simplicity we will only consider the one-dimensional case where we already have a local existence result in H^1. In fact, we will slightly generalize this result by dropping the requirement that α is an odd integer.

Theorem 14.6. *Let $n = 1$ and $\alpha \geq 2$. For every $g \in H^1(\mathbb{R})$ there is a $t_0 = t_0(\|g\|_{1,2}) > 0$, such that there is a unique solution $u \in C([-t_0,t_0], H^1(\mathbb{R}))$ of* (14.7). *Moreover, the solution map $g \mapsto u(t)$ will be Lipschitz continuous from every ball $\|g\|_{1,2} \leq \rho$ to $C([-t_0(\rho), t_0(\rho)], H^1(\mathbb{R}))$. If $\alpha \geq 4$ and $g \in H^2(\mathbb{R})$ the solution will be in $C^1([-t_0,t_0], H^1(\mathbb{R}))$.*

Proof. It suffices to verify that $F : H^1(\mathbb{R}) \to H^1(\mathbb{R})$ is locally Lipschitz on bounded sets. But this follows using Problem 14.4 since

$$\|F(u) - F(v)\|_2 \leq \alpha(\|u\|_\infty^{\alpha-1} + \|v\|_\infty^{\alpha-1})\|u - v\|_2$$

and

$$\begin{aligned}\|\partial(F(u) - F(v))\|_2 \leq{}& (\alpha-1)(\alpha+2)(\|u\|_\infty^{\alpha-2} + \|v\|_\infty^{\alpha-2})\|\partial u\|_2 \|u - v\|_2 \\ &+ \alpha\|v\|_\infty^{\alpha-1}\|\partial(u - v)\|_2\end{aligned}$$

together with $\|f\|_\infty \leq \|f\|_{1,2}$ (Problem 9.24). Similarly, one shows that $F : H^3(\mathbb{R}) \to H^3(\mathbb{R})$ is locally Lipschitz on bounded sets provided $\alpha \geq 4$, which establishes the last claim. □

Corollary 14.7. *The solution exists on a maximal time interval $(t_-(g), t_+(g))$ and if $|t_\pm(g)| < \infty$ we must have $\liminf_{t\to t_\pm(g)} \|u(t)\|_\infty = \infty$.*

Proof. Using the estimates from the proof (for $u = 0$) we obtain $\|F(v)\|_{H^1} \leq \alpha\|v\|_\infty^{\alpha-1}\|v\|_{H^1}$ and the claim follows from Lemma 12.5. $\square$

As already noted one gets global existence in the defocusing case. In the focusing case we need to control the $L^{\alpha+1}$ norm in terms of the H^1 norm.

Corollary 14.8. *In the defocusing case solutions are always global. In the focusing case the maximal solution u is global in $C(\mathbb{R}, H^1(\mathbb{R}^n))$ and preserves both mass and energy if one of the following conditions hold:*

(i) $\alpha < 5$.

(ii) $\alpha = 5$ *and* $\|g\|_2 \leq (\frac{3}{4})^{1/4}$.

(iii) $\alpha > 5$ *and* $\|g\|_{1,2}$ *is sufficiently small such that* $\|g'\|_2 < 1$ *and* $2E(0) + \frac{2^{\frac{\alpha+1}{2}}}{\alpha+1}\|g\|_2^{(\alpha+3)/2} < 1$.

Proof. Using $\|f\|_\infty^2 \leq 2\|f\|_2\|f'\|_2$ (Problem 9.24) one obtains

$$\int_{\mathbb{R}} |f(x)|^{\alpha+1} dx \leq \|f\|_\infty^{\alpha-1}\|f\|_2^2 \leq 2^{\frac{\alpha-1}{2}}\|f\|_2^{\frac{\alpha+3}{2}}\|f'\|_2^{\frac{\alpha-1}{2}}$$

and consequently

$$\|u'(t)\|_2^2 \leq 2E(0) + \frac{2^{\frac{\alpha+1}{2}}}{\alpha+1}\|g\|_2^{\frac{\alpha+3}{2}}\|u'(t)\|_2^{\frac{\alpha-1}{2}}.$$

(i). Now if $\alpha < 5$, then $\frac{(\alpha-1)}{2} < 2$ and $\|u'(t)\|_2$ remains bounded.

(ii). In the case $\alpha = 5$ this remains still true if $\frac{4}{3}\|g\|_2^4 < 1$.

(iii). If $\alpha > 5$ we can choose $\|g\|_{1,2}$ so small such that the given conditions hold. Note that this is possible since our above calculation shows

$$E(0) \leq \frac{1}{2}\|g'\|_2^2 + \frac{2^{\frac{\alpha-1}{2}}}{\alpha+1}\|g\|_2^{(\alpha+3)/2}\|g'\|_2^{(\alpha-1)/2}.$$

Now if we start with $\|u'(0)\|_2^2 < 1$ and assume $\|u'(t)\|_2^2 = 1$ we get the contradiction $1 = \|u'(t)\|_2^2 \leq 2E(0) + \frac{2^{\frac{\alpha+1}{2}}}{\alpha+1}\|g\|_2^{(\alpha+3)/2} < 1$. Hence $\|u'(t)\|_2^2 < 1$ as desired. $\square$

Finally, we want to show that solutions are not always global in the focusing case. To this end we need

Lemma 14.9. *Let $n = 1$ and $\alpha \geq 4$. Suppose $g \in H^1(\mathbb{R})$ satisfies $\|xg(x)\|_2 < \infty$ and let $u \in C((t_-, t_+), H^1(\mathbb{R}))$ be the maximal solution of* (14.7). *Then*

$$M_1(t) := \int_{\mathbb{R}} x^2|u(t,x)|^2 dx \tag{14.10}$$

remains finite as long as u exists. Moreover, we have

$$\dot{M}_1(t) = 4\mathrm{Im} \int_{\mathbb{R}} x\, u(t,x)^* u'(t,x) dx, \tag{14.11}$$

$$\ddot{M}_1(t) = 16E(t) \pm \frac{4(\alpha-5)}{\alpha+1} \int_{\mathbb{R}} |u(t,x)|^{\alpha+1} dx, \tag{14.12}$$

known as ***virial*** *and* ***Morawetz identity***,[1] *respectively.*

Proof. Consider $H^{1,1}(\mathbb{R}) := H^1(\mathbb{R}) \cap L^2(\mathbb{R}, x^2 dx)$ together with the norm $\|f\|^2 = \|f\|_2^2 + \|f'\|_2^2 + \|xf(x)\|_2^2$. Then $T_S(t)$ is a C_0 group satisfying $\|T_S(t)f\| \le (1+2|t|)\|f\|$. Moreover, as in the previous theorem one verifies that $F : H^{1,1}(\mathbb{R}) \to H^{1,1}(\mathbb{R})$ is locally Lipschitz on bounded sets. In fact, note that by

$$\|x(F(u)(x) - F(v)(x))\|_2 \le \alpha(\|u\|_\infty^{\alpha-1} + \|v\|_\infty^{\alpha-1})\|x(u(x)-v(x))\|_2$$

we can proceed as before. In particular, we get existence of local solutions and Lemma 12.5 shows that our norm cannot blow up before the sup norm.

To obtain the virial identity we first assume $g \in H^{3,2}(\mathbb{R})$ such that we get $u \in C(I, H^{1,2}(\mathbb{R}))$ since we can replace x^2 by x^4 in the above argument. Here $I \subseteq \mathbb{R}$ is the maximal existence interval. Moreover, since $g \in H^3$, we also have that u is a strong solution in $H^1(\mathbb{R})$, that is, $u \in C(I, H^3(\mathbb{R})) \cap C^1(I, H^1(\mathbb{R}))$. Hence one computes

$$\begin{aligned}\dot{M}_1(t) &= 2\mathrm{Im} \int_{\mathbb{R}} x^2 u(t,x)^* \mathrm{i} u_t(t,x) dx = -2\mathrm{Im} \int_{\mathbb{R}} x^2 u(t,x)^* u''(t,x) dx \\ &= 4\mathrm{Im} \int_{\mathbb{R}} x\, u(t,x)^* u'(t,x) dx,\end{aligned}$$

where we have used integration by parts in the last step. Furthermore, since $u \in C^1(I, H^1(\mathbb{R}))$ we can take another derivative to obtain (note that $u \in H^{2,2}(\mathbb{R})$ implies $xu'(x) \in L^2$ — Problem 14.5)

$$\begin{aligned}\ddot{M}_1(t) &= 4\mathrm{Im} \int_{\mathbb{R}} x\big(\dot{u}(t,x)^* u'(t,x) + u(t,x)^* \dot{u}'(t,x)\big) dx \\ &= -4\mathrm{Im} \int_{\mathbb{R}} \big(2x\, u'(t,x) + u(t,x)\big)^* \dot{u}(t,x) dx \\ &= 4\mathrm{Re} \int_{\mathbb{R}} \big(2x\, u'(t,x) + u(t,x)\big)^* \big(-u''(t,x) \pm |u(t,x)|^{\alpha-1} u(t,x)\big) dx.\end{aligned}$$

To further simplify this expression we note (dropping the t dependence for notational simplicity)

$$-\int_{\mathbb{R}} \big(2x\, u'(x)\big)^* u''(x) dx = \int_{\mathbb{R}} \big(|u'(x)|^2 + 2\mathrm{i}x \mathrm{Im}(u''(x)^* u'(x))\big) dx$$

---

[1] Cathleen Synge Morawetz (1923–2017), Canadian-American mathematician

as well as (Problem 9.17)

$$\begin{aligned}\mathrm{Re}\int_{\mathbb{R}} x\, u'(x)^* |u(x)|^{\alpha-1} u(x) dx &= \frac{1}{\alpha+1}\int_{\mathbb{R}} x\, \partial_x |u(x)|^{\alpha+1} dx \\ &= -\frac{1}{\alpha+1}\int_{\mathbb{R}} |u(x)|^{\alpha+1} dx.\end{aligned}$$

Combing everything we arrive at

$$\ddot{M}_1(t) = 8\int_{\mathbb{R}} |u'(t,x)|^2 dx \mp 4\frac{\alpha-1}{\alpha+1}\int_{\mathbb{R}} |u(t,x)|^{\alpha+1} dx.$$

For a general $g \in H^{1,1}(\mathbb{R})$ we approximate using continuity of the solution map. This verifies both identities. □

Now we are ready to establish blowup for the focusing NLS equation.

Theorem 14.10. *Consider the one-dimensional focusing NLS equation with $\alpha \ge 5$. Let $g \in H^1(\mathbb{R}) \cap L^2(\mathbb{R}, x^2 dx)$ with negative energy $E < 0$. Then the corresponding maximal mild solution u satisfies $t_+(g) < \infty$.*

Proof. Due to our assumption $\alpha \ge 5$ we obtain $\ddot{M}_1(t) \le 16E$ implying $M_1(t) \le 8Et^2 + \dot{M}_1(0)t + M_1(0)$. Hence

$$t_+(g) < \frac{-1}{16E}\left(\dot{M}_1(0) + \sqrt{\dot{M}_1(0)^2 - 32 E M_1(0)}\right)$$

since $M_1(t)$ must remain positive. Note that this also shows $\dot{M}_1(0) > 0$ since otherwise $M_1(t)$ would be decreasing and hence would remain bounded. □

Notice that there are initial conditions with negative energy, since the two contributions to the energy scale differently. In particular, the energy will become negative if we scale g with a sufficiently large factor.

Problem 14.4. *Show that the real derivative (with respect to the identification $\mathbb{C} \cong \mathbb{R}^2$) of $F(u) = |u|^{\alpha-1}u$ is given by*

$$F'(u)v = |u|^{\alpha-1}v + (\alpha-1)|u|^{\alpha-3}u\mathrm{Re}(u^* v).$$

Conclude in particular,

$$|F'(u)v| \le \alpha |u|^{\alpha-1}|v|, \qquad |F(u)-F(v)| \le \alpha(|u|^{\alpha-1} + |v|^{\alpha-1})|u-v|.$$

Moreover, the second derivative is given by

$$vF''(u)w = (\alpha-1)|u|^{\alpha-5}u\big((\alpha+1)\mathrm{Re}(u^* v)\mathrm{Re}(u^* w) - u^2 v^* w^*\big).$$

and hence

$$|vF''(u)w| \le (\alpha-1)(\alpha+2)|u|^{\alpha-2}|v||w|.$$

Problem 14.5. *Show that if $u \in H^2(\mathbb{R}) \cap L^2(\mathbb{R}, x^4 dx)$, then $xu'(x) \in L^2$.*

14.3. Strichartz estimates

In order to improve the existence results from the previous sections we need a better understanding of the linear Schrödinger equation. Unlike for example the heat equation, the Schrödinger equation does only preserve but not improve the regularity of the initial condition. For example, choosing $f \in L^2 \setminus L^p$ (for some $p \neq 2$) and considering $g = T_S(-t_0)f$ shows that there are initial conditions in L^2 which are not in L^p at a given later time t_0. However, our aim in this section is to show that we still have $T(t)g \in L^p$ most of the time.

To this end we first need an explicit expression for the solution. As in the case of the heat equation, we would like to express our solution as a convolution with the initial condition. However, here we run into the problem that $\mathrm{e}^{-\mathrm{i}|p|^2 t}$ is not integrable. To overcome this problem we consider

$$f_\varepsilon(p) = \mathrm{e}^{-(\mathrm{i}t+\varepsilon)p^2}, \quad \varepsilon > 0. \tag{14.13}$$

Then we note

$$\mathcal{F}(\mathrm{e}^{-z|x|^2/2})(p) = \frac{1}{z^{n/2}}\mathrm{e}^{-|p|^2/(2z)}, \qquad \mathrm{Re}(z) > 0, \tag{14.14}$$

where $z^{n/2}$ is the standard branch with branch cut along the negative real axis. In fact, the case when t is real was shown in Lemma 6.4 and the general case follows from the indentity theorem for analytic functions since both sides are analytic in the inidcated region. Together with the fact that the Fourier transform maps convolutions into products (Corollary 8.4) we obtain

$$(f_\varepsilon \hat{g})^\vee(x) = \frac{1}{(4\pi(\mathrm{i}t+\varepsilon))^{n/2}} \int_{\mathbb{R}^n} \mathrm{e}^{-\frac{|x-y|^2}{4(\mathrm{i}t+\varepsilon)}} g(y) d^n y. \tag{14.15}$$

Taking the limit $\varepsilon \downarrow 0$ we finally arrive at

$$T_S(t)g(x) = \frac{1}{(4\pi\mathrm{i}t)^{n/2}} \int_{\mathbb{R}^n} \mathrm{e}^{\mathrm{i}\frac{|x-y|^2}{4t}} g(y) d^n y \tag{14.16}$$

for $t \neq 0$ and $g \in L^2(\mathbb{R}^n) \cap L^1(\mathbb{R}^n)$. In fact, the left-hand side converges to $T_S(t)g$ in L^2 and the limit of the right-hand side exists pointwise by dominated convergence and its pointwise limit must thus be equal to its L^2 limit.

Using this explicit form, we can again draw some further consequences. For example, if $g \in L^2(\mathbb{R}^n) \cap L^1(\mathbb{R}^n)$, then $u(t) := T_S(t)g \in C_0(\mathbb{R}^n)$ for $t \neq 0$ (Problem 14.7) and satisfies

$$\|u(t)\|_\infty \leq \frac{1}{|4\pi t|^{n/2}} \|g\|_1. \tag{14.17}$$

Moreover, we even have $u \in C(\mathbb{R} \setminus \{0\}, C_0(\mathbb{R}^n))$ (Problem 14.7).

Thus we have spreading of wave functions in this case. In fact, invoking the Riesz–Thorin[2] interpolation theorem (Theorem 9.2 from [**34**]) we even get

$$\|u(t)\|_p \le \frac{1}{|4\pi t|^{n/2-n/p}}\|g\|_{p'}. \tag{14.18}$$

for any $p \in [2,\infty]$ with $\frac{1}{p}+\frac{1}{p'}=1$. This also gives $u \in C(\mathbb{R}\setminus\{0\}, L^p(\mathbb{R}^n))$.

Next we look at average decay in an L^p sense instead of pointwise estimates with respect to t. To this end we will consider functions $f \in L^r(\mathbb{R}, L^p(\mathbb{R}^n))$ and we will denote the corresponding norm by

$$\|f\|_{L^r(L^p)} := \begin{cases} \left(\int_{\mathbb{R}} \|f(t)\|_p^r dt\right)^{1/r}, & r<\infty, \\ \sup_{t\in\mathbb{R}} \|f(t)\|_p, & r=\infty. \end{cases} \tag{14.19}$$

Please recall that $L^r(\mathbb{R}, L^p(\mathbb{R}^n))$ is a Banach space defined with the help of the Bochner integral (cf. Theorem B.28). It consists of (equivalence classes with respect to equality a.e. of) strongly measurable functions $f(t)$ for which $\|f(t)\|_p$ is in L^r. Here strongly measurable means, that $f(t)$ is a limit of simple functions $s_n(t)$. It turns out that a function is strongly measurable if and only if it is measurable and its range is separable. In our situation this latter condition will come for free in the case $p<\infty$ and similarly in the case $p=\infty$ if the range is contained in $C_0(\mathbb{R}^n)$. We will also need the following variational characterization of our space-time norms (Problem B.22) for a given strongly measurable function f:

$$\|f\|_{L^r(L^p)} = \sup_{\|g\|_{L^{r'}(L^{p'})}=1} \left| \int_{\mathbb{R}}\int_{\mathbb{R}^n} f(x,t)g(x,t) d^n x\, dt \right|. \tag{14.20}$$

Moreover, it suffices to take the sup over functions which have support in a compact rectangle.

We call a pair (p,r) **admissible** if

$$\begin{cases} 2 \le p \le \infty, & n=1 \\ 2 \le p < \frac{2n}{n-2}, & n \ge 2 \end{cases}, \qquad \frac{2}{r} = \frac{n}{2} - \frac{n}{p}. \tag{14.21}$$

Note $r \in [4,\infty]$ for $n=1$ and $r \in (\frac{2}{n-1},\infty]$ for $n \ge 2$.

Lemma 14.11. *Let T_S be the Schrödinger group and let (p,r) be admissible with $p>2$. Then we have*

$$\left(\int_{\mathbb{R}} \left(\int_{\mathbb{R}} \|T_S(t-s)g(s)\|_p ds\right)^r dt\right)^{1/r} \le C\|g\|_{L^{r'}(L^{p'})}, \tag{14.22}$$

where a prime denotes the corresponding dual index. Moreover, $s \mapsto T(t-s)g(s) \in L^p(\mathbb{R}^n)$ is integrable for a.e. $t \in \mathbb{R}$.

---

[2]Marcel Riesz (1886–1969), Hungarian mathematician

[2]Olof Thorin (1912–2004), Swedish mathematician

Proof. Of course $T_S(t-s)g(s)$ is measurable. Applying our interpolation estimate we obtain

$$\int \|T_S(t-s)g(s)\|_p ds \le C \int \frac{1}{|t-s|^{1-\alpha}} \|g(s)\|_{p'} ds,$$

where $\alpha = 1 - n(1/2 - 1/p) \in (0,1)$ by our restriction on p.

Furthermore, our choice for r implies $\alpha = 1 - \frac{2}{r} = \frac{1}{r'} - \frac{1}{r}$ with $r' = \frac{2}{1+\alpha} \in (1, \alpha^{-1})$. So taking the $\|.\|_{L^r}$ norm on both sides and using the Hardy–Littlewood–Sobolev inequality[3] (Theorem 9.10 from [**34**]) gives the estimate.

Hence the claim about integrability follows from Minkowski's integral inequality (Theorem B.29). □

Note that the case $p = 2$ (and $r = \infty$) the above lemma holds by unitarity and does not provide much new insight.

Theorem 14.12 (Strichartz[4] estimates)**.** *Let T_S be the Schrödinger group and let (p,r) be admissible. Suppose $g \in L^{r'}(\mathbb{R}, L^{p'}(\mathbb{R}^n))$ and $f \in L^2(\mathbb{R}^n)$. Then we have the following estimates:*

$$\|T_S(t)f\|_{L^r(L^p)} \le C\|f\|_2, \tag{14.23}$$

$$\left\| \int_{\mathbb{R}} T_S(s)g(s)ds \right\|_2 \le C\|g\|_{L^{r'}(L^{p'})}, \tag{14.24}$$

$$\left\| \int_{\mathbb{R}} T_S(t-s)g(s)ds \right\|_{L^r(L^p)} \le C\|g\|_{L^{r'}(L^{p'})}, \tag{14.25}$$

where a prime denotes the corresponding dual index.

Here $s \mapsto T_S(t-s)g(s) \in L^p(\mathbb{R}^n)$ is integrable for a.e. $t \in \mathbb{R}$ and the integral in (14.24) *has to be understood as a limit in L^2 when taking an approximating sequence of functions g with support in compact rectangles.*

Proof. Since the case $p = 2$ follows from unitarity, we can assume $p > 2$. The claims about integrability and the last estimate follow from the lemma.

Using unitarity of T_S and Fubini we get

$$\int_{\mathbb{R}} \int_{\mathbb{R}^n} (T_S(t)f)(x) g(t,x) d^n x\, dt = \int_{\mathbb{R}^n} f(x) \int_{\mathbb{R}} (T_S(t)g(t))(x) dt\, d^n x,$$

for $g \in L^{r'}(\mathbb{R}, L^{p'}(\mathbb{R}^n))$ with support in a compact rectangle. Note that in this case we have $g(t) \in L^2(\mathbb{R}^n)$ since $p' \le 2$. This shows that the first and second estimate are equivalent upon using the above characterization (14.20) as well as the analogous characterization for the L^2 norm.

[3] Godfrey Harold Hardy (1877–1947), English mathematician
[3] John Edensor Littlewood (1885–1977), English mathematician
[4] Robert Strichartz (1943–2021), American mathematician

Similarly, using again unitarity of T_S and Fubini

$$\begin{aligned}
&\left\| \int T_S(t) g(t) dt \right\|_2^2 \\
&\quad = \int_{\mathbb{R}^n} \int_{\mathbb{R}} (T_S(t) g(t))(x) dt \int_{\mathbb{R}} (T_S(s) g(s))(x)^* ds\, d^n x \\
&\quad = \int_{\mathbb{R}^n} \int_{\mathbb{R}} g(t,x) \int_{\mathbb{R}} T_S(t-s) g(s,x)^* ds\, dt\, d^n x,
\end{aligned}$$

which shows that the second and the third estimate are equivalent with a similar argument as before. □

Note that using the scaling $f(x) \to f(\lambda x)$ for $\lambda > 0$ shows that the left-hand side of (14.23) scales like $\lambda^{-n/p-2/r}$ while the right-hand side scales like $\lambda^{-n/2}$. So (14.23) can only hold if $\frac{n}{p} + \frac{2}{r} = \frac{n}{2}$.

In connection with the Duhamel formula the following easy consequence is also worth while noticing:

Corollary 14.13. *We also have*

$$\left\| \int_0^t T_S(t-s) g(s) ds \right\|_2 \le C \|g\|_{L^{r'}(L^{p'})}, \tag{14.26}$$

$$\left\| \int_0^t T_S(t-s) g(s) ds \right\|_{L^r(L^p)} \le C \|g\|_{L^{r'}(L^{p'})}. \tag{14.27}$$

Proof. The second estimate is immediate from the lemma and the first estimate follows from (14.24) upon restricting to functions g supported in $[0,t]$ and using a simple change of variables $\int_0^t T(t-s)g(s)ds = \int_0^t T(s)g(t-s)ds$. □

Note that, apart from unitarity of T_S, only (14.17) was used to derive these estimates. Moreover, since T_S commutes with derivatives, we can also get analogous estimates for derivatives:

Corollary 14.14. *We have the following estimates for $k \in \mathbb{N}_0$:*

$$\|T_S(t) f\|_{L^r(W^{k,p})} \le C \|f\|_{H^k}, \tag{14.28}$$

$$\left\| \int_{\mathbb{R}} T_S(s) g(s) ds \right\|_{H^k} \le C \|g\|_{L^{r'}(W^{k,p'})}, \tag{14.29}$$

$$\left\| \int_{\mathbb{R}} T_S(t-s) g(s) ds \right\|_{L^r(W^{k,p})} \le C \|g\|_{L^{r'}(W^{k,p'})}, \tag{14.30}$$

as well as

$$\left\| \int_0^t T_S(t-s)g(s)ds \right\|_{H^k} \le C\|g\|_{L^{r'}(W^{k,p'})}, \tag{14.31}$$

$$\left\| \int_0^t T_S(t-s)g(s)ds \right\|_{L^r(W^{k,p})} \le C\|g\|_{L^{r'}(W^{k,p'})}. \tag{14.32}$$

Proof. Consider dense sets $f \in \mathcal{S}(\mathbb{R}^n)$ and $g \in C_c(\mathbb{R}, \mathcal{S}(\mathbb{R}^n))$. Then we have for example

$$\|\partial_j T_S(t)f\|_{L^r(L^p)} = \|T_S(t)\partial_j f\|_{L^r(L^p)} \le C\|\partial_j f\|_2$$

by applying (14.23) to $\partial_j f$. Combining the estimates for f and its derivatives gives (14.28). Similarly for the other estimates. □

Problem 14.6. *Does the translation group $T(t)g(x) := g(x-t)$ satisfy (14.17)?*

Problem 14.7. *Let $u(t) := T_S(t)g$ for some $g \in L^1(\mathbb{R}^n)$. Show that $u \in C(\mathbb{R}\setminus\{0\}, C_0(\mathbb{R}^n))$. (Hint: Lemma 4.34 from* **[35]** *(iv).)*

Problem 14.8. *Prove that there is no triple p, q, t with $1 \le q < p < \infty$, $t \in \mathbb{R}$ such that*

$$\|T_S(t)g\|_q \le C\|g\|_p.$$

(Hint: The translation operator $T_a f(x) := f(x-a)$ commutes with $T_S(t)$. Moreover, we have

$$\lim_{|a|\to\infty} \|f + T_a f\|_p = 2^{1/p}\|f\|_p, \qquad 1 \le p < \infty.$$

Now apply this to the claimed estimate.)

14.4. Well-posedness in L^2 and H^1

The main obstacle to proving a local existence result in L^2 is the fact that our nonlinearity does not map L^2 to L^2 (and this was precisely the reason for choosing H^r in the previous section). On the other hand, the time evolution conserves the L^2 norm and hence we expect global solutions in this case.

So let us make two observations: First of all our nonlinearity $F(u) = |u|^{\alpha-1}u$ maps L^p to $L^{p/\alpha}$, so the only chance is that the linear time evolution improves this behavior. Now we know, since our evolution is unitary, there is no hope to get this for fixed t, but this is true in some averaged sense by the Strichartz estimate (14.23). Hence, if we add such a space-time norm to the L^2 norm, we might be able to control our singularity. In fact, the estimates (14.26) and (14.27) allow us to control the Duhamel part in (14.7) both in the L^2 and the space-time norm, respectively (the linear part being taken care of by unitarity and (14.23)). Since the spatial parts of the space-time norms must match up, we need $p'\alpha = p$, that is, $p = 1 + \alpha$. For the time part an inequality $r'\alpha \le r$ is sufficient since in this case $L^{r'\alpha} \subseteq L^r$ by Hölder's inequality. This imposes the restriction $\alpha \le 1 + \frac{4}{n}$. In fact, we will impose a strict inequality since we will use the contribution from Hölder's inequality to get a contraction. Moreover, note that the dependence on the initial condition g is controlled by the L^2 norm alone and this will imply that our contraction is uniform (in fact Lipschitz on bounded domains) with respect to the initial condition in L^2, and so will be the solution.

Theorem 14.15. *Suppose* $1 < \alpha < 1 + \frac{4}{n}$ *and consider the Banach space*

$$X := C([-t_0, t_0], L^2(\mathbb{R}^n)) \cap L^r([-t_0, t_0], L^{\alpha+1}(\mathbb{R}^n)), \quad r = \frac{4(\alpha+1)}{n(\alpha-1)}, \tag{14.33}$$

with norm

$$\|f\| := \sup_{t\in[-t_0,t_0]} \|f(t)\|_2 + \left(\int_{-t_0}^{t_0} \|f(t)\|_{\alpha+1}^r dt\right)^{1/r}. \tag{14.34}$$

Then for every $g \in L^2(\mathbb{R}^n)$ *there is a* $t_0 = t_0(\|g\|_2) > 0$*, such that there is a unique solution* $u \in X$ *of* (14.7)*. Moreover, the solution map* $g \mapsto u(t)$ *will be Lipschitz continuous from every ball* $\|g\|_2 \le \rho$ *to* X *defined with* $t_0(\rho)$*.*

Proof. We take $[0, t_0]$ as an interval for notational simplicity. We will show that (14.7) gives rise to a contraction on the closed ball $\bar{B}_a(0) \subset X$ provided a and t_0 are chosen accordingly. Denote the right-hand side of (14.7) by $K(u) \equiv K_g(u)$. We will fist show that $K : \bar{B}_a(0) \to \bar{B}_a(0)$ for a suitable a depending on $\|g\|_2$. To this end we first invoke (14.23), (14.26), and (14.27) with $p = \alpha + 1$ ($p' = \frac{\alpha+1}{\alpha}$) to obtain

$$\begin{aligned}\|K(u)\| &\le (1+C)\|g\|_2 + 2C\left(\int_0^{t_0} \||u|^\alpha(t)\|_{(\alpha+1)/\alpha}^{r'} dt\right)^{1/r'}\\ &\le (1+C)\|g\|_2 + 2C\left(\int_0^{t_0} \|u(t)\|_{\alpha+1}^{\alpha r'} dt\right)^{1/r'}.\end{aligned}$$

Next, since $\frac{1}{r'} = \theta + \frac{\alpha-1}{r} + \frac{1}{r}$, where $\theta = 1 - \frac{\alpha+1}{r} = 1 - \frac{n(\alpha-1)}{4} > 0$ we can use the generalized Hölder inequality in the form

$$\|1 \cdot f_1^{\alpha-1} f_2\|_{r'} \le \|1\|_{1/\theta} \|f_1^{\alpha-1}\|_{r/(\alpha-1)} \|f_2\|_r = t_0^\theta \|f_1\|_r^{\alpha-1} \|f_2\|_r$$

(with $f_1(t) = f_2(t) = \|u(t)\|_{\alpha+1}$) to obtain

$$\begin{aligned}\|K(u)\| &\le (1+C)\|g\|_2 + 2Ct_0^\theta \left(\int_0^{t_0} \|u(t)\|_{\alpha+1}^r dt\right)^{\alpha/r}\\ &\le (1+C)\|g\|_2 + 2Ct_0^\theta a^\alpha\end{aligned}$$

for $u \in \bar{B}_a(0)$. Now we choose $a = (2+C)\|g\|_2$ and $2C(2+C)t_0^\theta a^{\alpha-1} < 1$ such that

$$\|K(u)\| \le (1+C)\|g\|_2 + 2Ct_0^\theta(2+C)^\alpha \|g\|_2^\alpha < (2+C)\|g\|_2 = a.$$

Similarly we can show that K is a contraction. Invoking (14.26) and (14.27) we have

$$\|K(u) - K(v)\| \le 2C \left(\int_0^{t_0} \| |u(t)|^{\alpha-1}u(t) - |v(t)|^{\alpha-1}v(t)\|_{(\alpha+1)/\alpha}^{r'} dt \right)^{1/r'}$$

Now using (Problem 14.4)

$$\big||u|^{\alpha-1}u - |v|^{\alpha-1}v\big| \le \alpha\big(|u|^{\alpha-1} + |v|^{\alpha-1}\big)|u-v|, \qquad u, v \in \mathbb{C},$$

and invoking the generalized Hölder inequality in the form

$$\||u|^{\alpha-1}|u-v|\|_{(\alpha+1)/\alpha} \le \||u|^{\alpha-1}\|_{(\alpha+1)/(\alpha-1)}\|u-v\|_{\alpha+1} = \|u\|_{\alpha+1}^{\alpha-1}\|u-v\|_{\alpha+1}$$

and then in the previous form with $f_1 = \|u\|_{\alpha+1}$, $f_2 = \|u-v\|_{\alpha+1}$, we obtain

$$\begin{aligned}\|K(u) - K(v)\| &\le 2\alpha C \left(\int_0^{t_0} \big((\|u\|_{\alpha+1}^{\alpha-1} + \|v\|_{\alpha+1}^{\alpha-1})\|u-v\|_{\alpha+1}\big)^{r'} dt \right)^{1/r'} \\ &\le 2\alpha C t_0^\theta 2a^{\alpha-1} \left(\int_0^{t_0} \|u-v\|_{\alpha+1}^r dt \right)^{1/r} \\ &\le 4\alpha C t_0^\theta a^{\alpha-1}\|u-v\|.\end{aligned}$$

Hence, decreasing t_0 further (if necessary), such that we also have $4\alpha C t_0^\theta a^{\alpha-1} < 1$, we get a contraction. Moreover, since $\|K_g(u) - K_f(u)\| = \|K_{g-f}(0)\| \le (1+C)\|g-f\|_2$, the uniform contraction principle establishes the theorem. □

By interpolation (Problem 14.10) we also have:

Corollary 14.16. *The solution u is also in*

$$L^{r/\theta}([-t_0, t_0], L^{2(\alpha+1)/(\alpha+1-\theta(\alpha-1))}(\mathbb{R}^n))$$

for any $\theta \in (0,1)$.

Moreover, as in the previous section we obtain:

Corollary 14.17. *The maximal solution u is global in $C(\mathbb{R}, L^2(\mathbb{R}^n))$ and preserves the L^2 norm: $\|u(t)\|_2 = \|g\|_2$. In addition, it has the properties stated in the theorem for any $t_0 > 0$.*

Let me remark that it is possible to cover the case $\alpha = 1 + \frac{4}{n}$. The main difference is that the Hölder-type estimate in terms of t^θ for the integral in (14.7) is useless since $\theta = 0$. However, the integral still tends to zero as $t \to 0$. This will be true locally in a sufficiently small neighborhood, but we cannot control this neighborhood in terms of $\|g\|_2$.

However, we will turn to the case of initial conditions in H^1 instead.

Theorem 14.18. *Suppose $n \geq 3$ and $2 \leq \alpha < \frac{n+2}{n-2}$. Consider the Banach space*

$$X := C([-t_0, t_0], H^1(\mathbb{R}^n)) \cap L^r([-t_0, t_0], W^{1,p}(\mathbb{R}^n)), \tag{14.35}$$

where

$$p = \frac{n(\alpha+1)}{n+\alpha-1}, \qquad r = \frac{4(\alpha+1)}{(n-2)(\alpha-1)}, \tag{14.36}$$

with norm

$$\|f\| := \sup_{t\in[-t_0,t_0]} \|f(t)\|_{1,2} + \left(\int_{-t_0}^{t_0} \|f(t)\|_{1,p}^r dt\right)^{1/r}. \tag{14.37}$$

Then for every $g \in H^1(\mathbb{R}^n)$ there is a $t_0 = t_0(\|g\|_{1,2}) > 0$, such that there is a unique solution $u \in X$ of (14.7)*. Moreover, the solution map $g \mapsto u(t)$ will be Lipschitz continuous from every ball $\|g\|_{1,2} \leq \rho$ to X defined with $t_0(\rho)$.*

Proof. We begin with estimating the nonlinearity. For $u, v \in W^{1,p}$ and $w \in L^p$ we obtain

$$\||u|^{\alpha-2} vw\|_{p'} \leq \|u\|_q^{\alpha-2} \|v\|_q \|w\|_p \leq C \|\nabla u\|_p^{\alpha-2} \|\nabla v\|_p \|w\|_p,$$

where we have applied the generalized Hölder inequality with $\frac{1}{p'} = \frac{\alpha-2}{q} + \frac{1}{q} + \frac{1}{p}$ in the first step (requiring $\alpha \geq 2$) and the Gagliardo–Nirenberg–Sobolev inequality (Theorem 9.22 – since we need $p < n$, we need to require $n > 2$) with $\frac{1}{q} = \frac{1}{p} - \frac{1}{n}$ in the second step. In particular, this imposes

$$1 - \frac{2}{p} = \frac{\alpha-1}{q} = \frac{\alpha-1}{p} - \frac{\alpha-1}{n}$$

and explains our choice for p. The choice of r is of course dictated by (14.21) such that we can apply our Strichartz estimates. At this point a weaker upper bound (namely $\alpha < \frac{n}{n-4}$ for $n \geq 4$) is still sufficient.

Now using this estimate we see (cf. Problem 14.4)

$$\||u|^{\alpha-1}u\|_{p'} \leq C\|\nabla u\|_p^{\alpha-1} \|u\|_p, \quad \|\nabla |u|^{\alpha-1} u\|_{p'} \leq \alpha \||u|^{\alpha-1} \nabla u\|_{p'} \leq \alpha C \|\nabla u\|_p^\alpha$$

and hence

$$\||u|^{\alpha-1}u\|_{1,p'} \leq \tilde{C} \|u\|_{1,p}^\alpha.$$

Similarly we obtain

$$\begin{aligned} \||u|^{\alpha-1}u - |v|^{\alpha-1}v\|_{p'} &\leq \alpha \|\big(|u|^{\alpha-1} + |v|^{\alpha-1}\big)|u-v|\|_{p'} \\ &\leq \alpha C \big(\|\nabla u\|_p^{\alpha-1} + \|\nabla v\|_p^{\alpha-1}\big) \|u-v\|_p \end{aligned}$$

and

$$\begin{aligned}
&\|\nabla|u|^{\alpha-1}u-\nabla|v|^{\alpha-1}v\|_{p'}\\
&\quad\le(\alpha-1)(\alpha+2)\|\big(|u|^{\alpha-2}+|v|^{\alpha-2}\big)|u-v|\,|\nabla u|\|_{p'}\\
&\qquad+\alpha\||v|^{\alpha-1}|\nabla u-\nabla v|\|_{p'}\\
&\quad\le(\alpha-1)(\alpha+2)C\big(\|\nabla u\|_p^{\alpha-2}+\|\nabla v\|_p^{\alpha-2}\big)\|\nabla(u-v)\|_p\|\nabla u\|_p\\
&\qquad+\alpha C\|\nabla v\|_p^{\alpha-1}\|\nabla u-\nabla v\|_p.
\end{aligned}$$

In summary,

$$\||u|^{\alpha-1}u-|v|^{\alpha-1}v\|_{1,p'}\le\bar C\big(\|u\|_{1,p}^{\alpha-1}+\|v\|_{1,p}^{\alpha-1}\big)\|u-v\|_{1,p}.$$

Now the rest follows as in the proof of Theorem 14.15. Note that in this case $\theta=1-\frac{\alpha+1}{r}=\frac{2+n+(2-n)\alpha}{4}$ explaining our upper limit for α. □

Note that since we have $H^1(\mathbb{R}^n)\subseteq L^{\alpha+1}(\mathbb{R}^n)$ for $n\ge3$ and $\alpha<\frac{n+2}{n-2}$ by the Gagliardo–Nirenberg–Sobolev inequality (Theorem 9.22), both the momentum and the energy are finite and preserved by our solutions. Moreover, in the defocusing case the momentum and the energy control the H^1 norm and hence we obtain:

Corollary 14.19. *In the defocusing case the maximal solution u is global in $C(\mathbb{R},H^1(\mathbb{R}^n))$ and preserves both momentum and energy. In addition, it has the properties stated in the theorem for any $t_0>0$.*

In the focusing case we need to control the $L^{\alpha+1}$ norm in terms of the H^1 norm using the Gagliardo–Nirenberg–Sobolev inequality.

Corollary 14.20. *In the focusing case the maximal solution u is global in $C(\mathbb{R},H^1(\mathbb{R}^n))$ and preserves both momentum and energy if one of the following conditions hold:*

(i) $\alpha<1+\frac{4}{n}$.

(ii) $\alpha=1+\frac{4}{n}$ *and* $\|g\|_2<(\frac{2(n-1)}{(n+2)(n-2)})^{n/4}$.

(iii) $\alpha>1+\frac{4}{n}$ *and* $\|g\|_{1,2}$ *is sufficiently small such that* $\|\nabla g\|_2<1$ *and* $2E(0)+\frac{4(n-1)}{(n(n-2)(\alpha+1)}\|g\|_2^{\alpha+1-n(\alpha-1)/2}<1$.

Proof. Using the Gagliardo–Nirenberg–Sobolev inequality and the Lyapunov inequality (Problem B.14) with $\frac{1}{1+\alpha}=\theta(\frac{1}{2}-\frac{1}{n})+\frac{1-\theta}{2}$ (i.e. $\theta=\frac{n(\alpha-1)}{2(\alpha+1)}$) we obtain

$$\|u(t)\|_{\alpha+1}^{\alpha+1}\le\frac{2(n-1)}{n(n-2)}\|u(t)\|_2^{\alpha+1-n(\alpha-1)/2}\|\nabla u(t)\|_2^{n(\alpha-1)/2}.\tag{14.38}$$

Thus

$$\|\nabla u(t)\|_2^2 = 2E(0) + \frac{2}{\alpha+1}\|u(t)\|_{\alpha+1}^{\alpha+1}$$
$$\leq 2E(0) + 2C\|g\|_2^{\alpha+1-n(\alpha-1)/2}\|\nabla u(t)\|_2^{n(\alpha-1)/2}, \qquad (14.39)$$

where we have set $C := \frac{2(n-1)}{n(n-2)(\alpha+1)}$.

(i). Now if $\alpha < 1 + \frac{4}{n}$, then $\frac{n(\alpha-1)}{2} < 2$ and $\|\nabla u(t)\|_2$ remains bounded.

(ii). In the case $\alpha = 1 + \frac{4}{n}$ this remains still true if $2C\|g\|_2^{4/n} < 1$.

(iii). If $\alpha > 1+\frac{4}{n}$ we can choose $\|g\|_{1,2}$ so small such that the given conditions hold. Note that this is possible since our above calculation shows

$$E(0) \leq \frac{1}{2}\|\nabla g\|_2^2 + C\|g\|_2^{\alpha+1-n(\alpha-1)/2}\|\nabla g\|_2^{n(\alpha-1)/2}.$$

Now if we start with $\|\nabla u(0)\|_2^2 \leq 1$ and assume $\|\nabla u(t)\|_2^2 = 1$ we get the contradiction $1 = \|\nabla u(t)\|_2^2 \leq 2E(0) + 2C\|g\|_2^{\alpha+1-n(\alpha-1)/2} < 1$. Hence $\|\nabla u(t)\|_2^2 < 1$ as desired. □

Problem 14.9. *Show that* (14.33) *is a Banach space. (Hint: Work with test functions from* C_c^∞*.)*

Problem 14.10. *Suppose* $f \in L^{p_0}(I, L^{q_0}(U)) \cap L^{p_1}(I, L^{q_1}(U))$*. Show that* $f \in L^{p_\theta}(I, L^{q_\theta}(U))$ *for* $\theta \in [0,1]$*, where*

$$\frac{1}{p_\theta} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}, \qquad \frac{1}{q_\theta} = \frac{1-\theta}{q_0} + \frac{\theta}{q_1}.$$

(Hint: Lyapunov and generalized Hölder inequality — Problem B.14 and Problem B.12.)

14.5. Standing waves

A solution of the form

$$u(x,t) = \varphi_\omega(x)\mathrm{e}^{\mathrm{i}\omega t}, \qquad \omega > 0, \qquad (14.40)$$

of the focusing NLS equation is called a **standing wave**. Inserting this ansatz into the equation shows that φ_ω must be a solution of the following nonlinear elliptic problem

$$-\Delta\varphi_\omega + \omega\varphi_\omega = |\varphi_\omega|^{\alpha-1}\varphi_\omega. \qquad (14.41)$$

Note that one can choose $\omega = 1$ without loss of generality since if φ is a solution for $\omega = 1$ then

$$\varphi_\omega(x) = \omega^{\frac{1}{\alpha-1}}\varphi(\omega^{1/2}x) \qquad (14.42)$$

is a solution for $\omega > 0$. Moreover, if φ is a solution, so is $\mathrm{e}^{\mathrm{i}\theta}\varphi(.-a)$ for any $\theta \in \mathbb{R}$ and $a \in \mathbb{R}^n$.

If one multiplies (14.41) with a test function $v \in H^1(\mathbb{R}^n)$ and integrates over $\mathbb{R}^n$ one obtains the weak formulation

$$\int_{\mathbb{R}^n} \big(\nabla\varphi \cdot \nabla v + \varphi v - |\varphi|^{\alpha-1}\varphi v\big) d^n x = 0, \qquad v \in H^1(\mathbb{R}^n). \tag{14.43}$$

In particular, choosing $v = \varphi^*$ we obtain

$$\int_{\mathbb{R}^n} \big(|\nabla\varphi|^2 + |\varphi|^2 - |\varphi|^{\alpha+1}\big) d^n x = 0, \tag{14.44}$$

which shows that, if we flip the sign in front of the nonlinearity (defocusing case), there is only the trivial solution.

In one-dimension one has the explicit solution

$$\varphi(x) = \left(\frac{\sqrt{1+\beta}}{\cosh(\beta x)}\right)^{1/\beta}, \qquad \beta = \frac{\alpha-1}{2}. \tag{14.45}$$

In higher dimensions we can apply Theorem 13.6 to get existence of solutions:

Theorem 14.21. *Suppose $n \ge 2$ and $1 < \alpha < \frac{n+2}{n-2}$. Then the nonlinear elliptic problem* (14.41) *has a weak positive radial solution in $H^1(\mathbb{R}^n)$.*

Proof. To apply Theorem 13.6 we choose $X = H^1_{\mathrm{rad}}(\mathbb{R}^n, \mathbb{R})$ and $Y = L^{\alpha+1}_{\mathrm{rad}}(\mathbb{R}^n, \mathbb{R})$ and note that the Strauss inequality (Corollary 14.24) implies compactness of the embedding $X \hookrightarrow Y$ for the range of α under consideration. Hence minimizing

$$F(u) = \frac{1}{2}\int_{\mathbb{R}^n} \big(|\nabla u|^2 + |u|^2\big) d^n x$$

under the constraint (cf. Example 13.3)

$$N(u) = \frac{1}{\alpha+1}\int_{\mathbb{R}^n} |u|^{\alpha+1} d^n x = 1$$

gives a weak radial solution u_0 of the problem

$$-\Delta u + u = \lambda |u|^{\alpha-1} u.$$

In particular, choosing u_0 as a test function for the weak formulation shows $\lambda > 0$. Moreover, by Lemma 9.8 we have $|u_0| \in H^1_{rad}(\mathbb{R}^n)$ with $F(|u_0|) = F(u_0)$ and hence $|u_0|$ is also a minimizer. Rescaling this solution according to $\varphi(x) = \lambda^{1/(\alpha-1)}|u_0(x)|$ establishes the claim. □

Note that for $\alpha \le \frac{n}{n-2}$ we have $|u|^{\alpha-1}u \in L^2(\mathbb{R}^n)$ for $u \in H^1(\mathbb{R}^n)$ and hence $(-\Delta + 1)\varphi \in L^2(\mathbb{R}^n)$ implying $\varphi \in H^2(\mathbb{R}^n)$.

Finally, we remark that the restriction $\alpha < \frac{n+2}{n-2}$ is also necessary for solutions to exist. This follows from the **Pokhozhaev identity**[5] (Problem 14.12)

$$\frac{n-2}{2}\int_{\mathbb{R}^n} |\nabla\varphi|^2 d^n x + \frac{n}{2}\int_{\mathbb{R}^n} |\varphi|^2 d^n x = \frac{n}{\alpha+1}\int_{\mathbb{R}^n} |\varphi|^{\alpha+1} d^n x. \tag{14.46}$$

Combining this equality with (14.44) gives

$$(\alpha - 1 - 2\frac{\alpha+1}{n})\int_{\mathbb{R}^n} |\nabla\varphi|^2 d^n x + (\alpha-1)\int_{\mathbb{R}^n} |\varphi|^2 d^n x = 0. \tag{14.47}$$

Since the first coefficient is nonnegative for $\alpha \geq \frac{n+2}{n-2}$ if $n \geq 3$, we see that there cannot be a nontrivial solution in this case.

Example 14.1. The standing waves (14.40) can also be used to establish blowup in the case $\alpha = 1 + \frac{4}{n}$. Indeed by Problem 14.2

$$u(t,x) := t^{-n/2} \mathrm{e}^{\mathrm{i}|x|^2/(4t) - \mathrm{i}/t} \varphi(\frac{x}{t})$$

will be a solution for $t \neq 0$. Now note that while $\|u(t)\|_2 = \|u(0)\|_2$, we have

$$\begin{aligned}\|\nabla u(t)\|_2^2 &= |t|^{n-2}\int_{\mathbb{R}^n}\left(\frac{|x|^2}{4}\varphi(\frac{x}{t})^2 + |\nabla\varphi(\frac{x}{t})|^2\right) d^n x\\ &= \frac{1}{4}\||x|\varphi(x)\|_2^2 + \frac{1}{t^2}\|\nabla\varphi\|_2^2,\end{aligned}$$

which shows that the gradient will blow up as $t \to 0$. $\diamond$

Problem 14.11. *Let $1 < \alpha < \frac{n+2}{n-2}$ be an odd integer (i.e. $n = 2$ and $\alpha = 3, 5, 6, \dots$ or $n = 3$ and $\alpha = 3$). Show that $\varphi \in H^k(\mathbb{R}^n)$ for any $k \in \mathbb{N}$. (Hint: As already pointed out we have $\varphi \in H^2$.)*

Problem 14.12. *Suppose $\varphi \in H^2(\mathbb{R}^n) \cap L^{\alpha+1}(\mathbb{R}^n)$ is a solution of* (14.41) *with $\omega = 1$ such that $|x|\partial_\beta\varphi(x) \in L^2(\mathbb{R}^n)$ for $|\beta| \leq 2$. Show that φ satisfies* (14.46)*. (Hint: Multiply* (14.41) *with $x \cdot \nabla\varphi(x)^*$ and take the real part; compare the proof of Lemma 14.9.)*

14.6. Appendix: Radial Sobolev spaces

Consider $B := B_R(0) \subseteq \mathbb{R}^n$ (with the case $R = \infty$ allowed). The subset

$$W^{k,p}_{\mathrm{rad}}(B) := \{f \in W^{k,p}(B) | f \text{ is radial}\} \tag{14.48}$$

is closed and hence a Banach space of its own. To see this take a Cauchy sequence of radial functions. Without loss of generality we can assume that it converges pointwise a.e. In particular, any rotated point leads to the same limit and hence the limit function is also radial (a.e.).

[5]Stanislav Ivanovich Pokhozhaev (1935–2014), Soviet mathematician

Of course there is a one to one correspondence between radial functions on B and functions on $[0,R)$. For every radial function f on B let $\tilde{f}$ be the associated function on $(0,R)$ defined such that $f(x)=\tilde{f}(r)$, where $r=|x|$. Note that by Lemma A.11 we have

$$\left(S_n\int_0^\infty|\tilde{f}(r)|^p r^{n-1}dr\right)^{1/p}=\|f\|_p,\quad 1\le p<\infty, \tag{14.49}$$

for radial functions. In particular, we have $f\in L^p_{\mathrm{rad}}(B)$ if and only if $\tilde{f}\in L^p((0,R),r^{n-1}dr)$.

Lemma 14.22. *We have $f\in W^{1,p}_{\mathrm{rad}}(B)$ if and only if $\tilde{f}\in W^{1,p}((0,R),r^{n-1}dr)$ with equivalent norms. The derivatives are connected via*

$$\tilde{f}'(r)=\frac{x}{r}\cdot\nabla f(x),\qquad r=|x|, \tag{14.50}$$

and

$$\nabla f(x)=\tilde{f}'(r)\frac{x}{r}. \tag{14.51}$$

Here $W^{1,p}((0,R),r^{n-1}dr)$ is the set of all functions $f\in AC(0,R)$ for which $f,f'\in L^p((0,R),r^{n-1}dr)$.

Proof. Let $f\in W^{1,p}_{\mathrm{rad}}(B)$. To show that $\tilde{f}$ has a weak derivative we take $\tilde{\phi}\in C_c^\infty(0,R)$ and set $\phi(x)=\tilde{\phi}(|x|)$ (note that $\phi\in C_c^\infty(B)$). Then one checks

$$\tilde{\phi}'(r)=\sum_{j=1}^n\frac{x_j}{r}(\partial_j\phi)(x),\qquad r:=|x|,$$

and thus (using integration by parts, the product rule, and $\operatorname{div}(\frac{x}{r^n})=0$)

$$\begin{aligned}\int_0^R\tilde{f}(r)\tilde{\phi}'(r)dr&=S_n^{-1}\sum_{j=1}^n\int_B f(x)\frac{x_j}{r}(\partial_j\phi)(x)r^{-n+1}d^nx\\&=-S_n^{-1}\sum_{j=1}^n\int_B\left(\partial_j\frac{x_j}{r^n}f(x)\right)\phi(x)d^nx\\&=-S_n^{-1}\sum_{j=1}^n\int_B\frac{x_j}{r}(\partial_jf)(x)\phi(x)r^{-n+1}d^nx\\&=-\int_0^R\tilde{g}(r)\tilde{\phi}(r)dr,\end{aligned}$$

where

$$\tilde{g}(r):=S_n^{-1}\sum_{j=1}^n\int_{S^{n-1}}\frac{x_j}{r}(\partial_jf)(r\omega)d\sigma^{n-1}(\omega).$$

Since the differential operator $\sum_{j=1}^n \frac{x_j}{r}\partial_j$ is invariant under rotations, the spherical average is not necessary and we have in fact

$$\tilde{g}(r) = \sum_{j=1}^n \frac{x_j}{r}(\partial_j f)(x).$$

This shows that $\tilde{f}$ has a weak derivative given by $\tilde{g}$.

Conversely, suppose $\tilde{f}$ is $W^{1,p}((0,R), r^{n-1}dr)$ and let $f(x) := \tilde{f}(r)$. Then $x_j \mapsto f(x)$ is absolutely continuous with derivative given by $\partial_j f(x) = \tilde{f}'(r)\frac{x_j}{r}$ (cf. Problem 4.38 from [**34**]). Hence for $\phi \in C_c^\infty(B)$ we have

$$\begin{aligned}\int_B f(x)\partial_j\phi(x)d^n x &= \int_Q \tilde{f}(r)\partial_j\phi(x)d^n x = -\int_Q (\partial_j \tilde{f}(r))\phi(x)d^n x\\ &= -\int_B \tilde{f}'(r)\frac{x_j}{r}\phi(x)d^n x,\end{aligned}$$

where we have replaced B by a cube $Q \supseteq B$ for the purpose of integration by parts.

The connection between the norms follows from (14.49). □

The crucial observation due to Strauss[6] is that a radial function automatically satisfies a decay estimate:

Lemma 14.23. *For $f \in W^{1,p}_{\mathrm{rad}}(\mathbb{R}^n)$, $1 \le p < \infty$, we have the* ***Strauss inequality***

$$r^{n-1}|\tilde{f}(r)|^p \le \frac{p}{S_n}\|f\|_p^{p-1}\|\nabla f\|_p. \tag{14.52}$$

Proof. By Lemma 14.22 $\tilde{f} \in W^{1,p}((0,\infty), r^{n-1}dr)$ and hence $F(r) := r^{n-1}|\tilde{f}(r)|^p$ is absolutely continuous and integrable. In particular,

$$\begin{aligned}F(r) - F(R) &= -\int_r^R F'(s)ds\\ &= -(n-1)\int_r^R |\tilde{f}(s)|^p s^{n-2}ds - p\,\mathrm{Re}\int_r^R |\tilde{f}(s)|^{p-2}\tilde{f}(s)^*\tilde{f}'(s)s^{n-1}ds\end{aligned}$$

and letting $R \to \infty$ shows that $\lim_{R\to\infty} F(R)$ exists. Since F is integrable it must be zero:

$$F(r) = -(n-1)\int_r^\infty |\tilde{f}(s)|^p s^{n-2}ds - p\,\mathrm{Re}\int_r^\infty |\tilde{f}(s)|^{p-2}\tilde{f}(s)^*\tilde{f}'(s)s^{n-1}ds.$$

[6] Walter Alexander Strauss (* 1937), American mathematician

Hence

$$\begin{aligned} F(r) &\le p \int_0^\infty |\tilde{f}(s)|^{p-1} |\tilde{f}'(s)| s^{n-1} ds \\ &\le p \|\tilde{f}\|^{p-1}_{L^p((0,\infty), r^{n-1}dr)} \|\tilde{f}'\|_{L^p((0,\infty), r^{n-1}dr)} \\ &\le p S_n^{-1} \|f\|_p^{p-1} \|\nabla f\|_p \end{aligned}$$

by Lemma 14.22. □

Corollary 14.24. *For $f \in W^{1,p}_{\mathrm{rad}}(\mathbb{R}^n)$ we have*

$$\|(1-\chi_{B_r(0)}) f\|_q \le \frac{C^{1-p/q}}{r^{(n-1)(1/p-1/q)}} \|f\|_{W^{1,p}}, \quad q \ge p. \tag{14.53}$$

Consequently $W^{1,p}_{\mathrm{rad}}(\mathbb{R}^n)$ is compactly embedded into $L^q(\mathbb{R}^n)$ for $q \in (p, \frac{np}{n-p})$ if $n \ge \max(p,2)$.

Proof. By the Strauss inequality (14.52) we have

$$|\tilde{f}(r)| \le C r^{-(n-1)/p} \|f\|_{W^{1,p}}$$

and hence for $q \ge p$

$$|\tilde{f}(r)|^q r^{n-1} \le C^{q-p} r^{-(n-1)(q-p)/p} \|f\|^{q-p}_{W^{1,p}} |\tilde{f}(r)|^p r^{n-1}$$

Integrating from R to ∞ shows

$$\int_R^\infty |\tilde{f}(r)|^q r^{n-1} dr \le C^{q-p} R^{-(n-1)(q-p)/p} \|f\|^{q-p}_{W^{1,p}} \int_R^\infty |\tilde{f}(r)|^p r^{n-1} dr$$

and hence

$$\|(1-\chi_{B_R(0)}) f\|_q^q \le \frac{C^{q-p}}{R^{(n-1)(q-p)/p}} \|f\|^{q-p}_{W^{1,p}} \|f\|_p^p \le \frac{C^{q-p}}{R^{(n-1)(q-p)/p}} \|f\|^q_{W^{1,p}}.$$

Finally Theorem 9.33 gives compactness of the embedding. □

Example 14.2. Note that Examples 9.13 and 9.14 show that the corollary fails for $q = \frac{np}{n-p}$ and $q = p$, respectively. ◇

Part 3

Appendices

Appendix A

Calculus facts

A.1. Differentiation

Let $U \subseteq \mathbb{R}^n$ be a domain. Then $C^k(U)$ denotes the set of all real-valued functions which have continuous partial derivatives of order up to k. The functions which have partial derivatives of all orders are denoted by $C^\infty(U)$. The **support** of a functions is defined as

$$\operatorname{supp}(f) := \overline{\{x \in U | f(x) \neq 0\}} \tag{A.1}$$

and f is said to be supported in V if $\operatorname{supp}(f) \subseteq V$. The set of C^k functions with compact support is denoted by $C_c^k(U)$.

When handling derivatives of functions of several variables, things simplify considerably once one uses the right notation. The weapon of choice is usually the **multi-index notation** of Laurent Schwartz. So, may the Schwartz be with you:

For $f \in C^k(U)$ and $\alpha \in \mathbb{N}_0^n$ we set

$$\partial_\alpha f := \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}}, \qquad |\alpha| := \alpha_1 + \cdots + \alpha_n, \tag{A.2}$$

for $|\alpha| \leq k$. Also recall that by the classical theorem of Schwarz the order in which these derivatives are performed is irrrelevant. In this context $\alpha \in \mathbb{N}_0^n$ is called a **multi-index** and $|\alpha|$ is called its **order**. Multi-indices are partially ordered via $\beta \leq \alpha$ provided $\beta_j \leq \alpha_j$ for $1 \leq j \leq n$.

We will also set

$$x^\alpha := x_1^{\alpha_1} \cdots x_n^{\alpha_n}, \qquad (\lambda x)^\alpha = \lambda^{|\alpha|} x^\alpha \tag{A.3}$$

for $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$. Furthermore, there are generalizations for the **factorial** and the **binomial coefficients**

$$\binom{\alpha}{\beta} := \frac{\alpha!}{\beta!(\alpha-\beta)!}, \qquad \alpha! := \prod_{j=1}^{n} (\alpha_j!). \tag{A.4}$$

Note that the binomial coefficients can be computed recursively as usual:

$$\binom{\alpha+\delta}{\beta} = \binom{\alpha}{\beta} + \binom{\alpha}{\beta-\delta}, \qquad |\delta| = 1. \tag{A.5}$$

Also note that

$$\alpha! \le |\alpha|! \le n^{|\alpha|} \alpha!. \tag{A.6}$$

As a simple exercise one can verify

$$\partial_\beta x^\alpha = \begin{cases} \frac{\alpha!}{(\alpha-\beta)!} x^{\alpha-\beta}, & \alpha \ge \beta, \\ 0, & \text{else}. \end{cases} \tag{A.7}$$

Some slightly more sophisticated formulas are collected in the next lemma.

Lemma A.1. *With this notation one has*

(i) *For $x, y \in \mathbb{R}^n$ we have the **multi-binomial theorem***

$$(x+y)^\alpha = \sum_{\beta \le \alpha} \binom{\alpha}{\beta} x^\beta y^{\alpha-\beta}. \tag{A.8}$$

(ii) *For $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ and $m \in \mathbb{N}$ we have the **multinomial theorem***

$$(x_1 + \ldots + x_n)^m = \sum_{|\alpha|=m} \frac{m!}{\alpha!} x^\alpha. \tag{A.9}$$

(iii) *For $f, g \in C^k(U)$ we have the **Leibniz rule***

$$\partial_\alpha(f \cdot g) = \sum_{\beta \le \alpha} \binom{\alpha}{\beta} (\partial_\beta f)(\partial_{\alpha-\beta} g), \quad |\alpha| \le k. \tag{A.10}$$

If $f \in C^{k+1}(U)$ we can fix $x, y \in U$ and consider the real function $t \mapsto f(x+ty)$ for $t \in \mathbb{R}$ sufficiently small. One computes the directional derivatives

$$\frac{d^m}{dt^m} f(x+ty) = \sum_{|\alpha|=m} \frac{|\alpha|!}{\alpha!} \left(\partial^\alpha f(x+ty)\right) y^\alpha, \quad m \le k,$$

and applying the classical Taylor theorem to $g(t) := f(x+ty)$ shows:

Theorem A.2 (Taylor). *Let $f \in C^{k+1}(U)$ and $x_0 \in U$. Then for every x in a neighborhood of x_0 we have*

$$f(x) = \sum_{|\alpha|\le k} \frac{\partial^\alpha f(x_0)}{\alpha!}(x-x_0)^\alpha + R_k(x,x_0), \tag{A.11}$$

*where the **remainder** is given by*

$$R_k(x,x_0) := \sum_{|\alpha|=k+1} \frac{\partial^\alpha f((1-t)x_0 + t\,x)}{\alpha!}(x-x_0)^\alpha \tag{A.12}$$

for some $t = t(x,x_0) \in [0,1]$.

A function $f : U \to \mathbb{R}$ is called **real analytic**, if it can be expanded into an absolutely convergent power series

$$f(x) = \sum_{\alpha\in\mathbb{N}_0^n} f_\alpha\,(x-x_0)^\alpha \tag{A.13}$$

in a neighborhood of any point $x_0 \in U$. We will write $C^\omega(U)$ for the set of real analytic function on a domain $U \subseteq \mathbb{R}^n$. A convenient way to establish convergence is comparison with a **majorant**, that is, a real analytic function $F(x) = \sum_{\alpha\in\mathbb{N}_0^n} F_\alpha\,(x-x_0)^\alpha$ with $|f_\alpha| \le F_\alpha$ for all $\alpha \in \mathbb{N}_0^n$.

As always, the geometric series will be a favorite choice.

Example A.1. The multidimensional geometric series

$$\sum_{\alpha\in\mathbb{N}_0^n} x^\alpha = \prod_{j=1}^n \sum_{k=1}^\infty x_j^k = \prod_{j=1}^n \frac{1}{1-x_j}$$

converges absolutely and uniformly on every compact subset of the rectangle $|x_j| < 1$, $1 \le j \le n$. Moreover, the same is true for the series associated with the derivatives

$$\begin{aligned}\sum_{\alpha\in\mathbb{N}_0^n} \partial_\beta x^\alpha &= \sum_{|\alpha|\ge|\beta|} \frac{\alpha!}{(\alpha-\beta)!} x^{\alpha-\beta} = \prod_{j=1}^n \sum_{k=\beta_j}^\infty \frac{k!}{(k-\beta_j)!} x_j^{k-\beta_j} \\ &= \prod_{j=1}^n \frac{1}{\beta_j!(1-x_j)^{\beta_j+1}} = \partial_\beta \sum_{\alpha\in\mathbb{N}_0^n} x^\alpha. \end{aligned}$$ ◇

For every $a \in \mathbb{R}^n$ we introduce the associated rectangle $\mathcal{R}(a) := \{x \in \mathbb{R}^n \,|\, |x_j| < |a_j|, 1 \le j \le n\}$. Then using the multidimensional geometric series as a majorant immediately gives

Lemma A.3. *Suppose $|f_\alpha a^\alpha| \le M$ for some $a \in \mathbb{R}^n$. Then the series*

$$f(x) := \sum_{\alpha\in\mathbb{N}_0^n} f_\alpha x^\alpha \tag{A.14}$$

converges absolutely and uniformly on every compact subset of the rectangle $\mathcal{R}(a)$. *Moreover, the same is true for the series associated with the derivatives and we have*

$$\sum_{\alpha\in\mathbb{N}_0^n} f_\alpha \partial_\beta x^\alpha = \sum_{|\alpha|\geq|\beta|} f_\alpha \frac{\alpha!}{(\alpha-\beta)!} x^{\alpha-\beta}. \tag{A.15}$$

In particular, if $\mathcal{R}(a)$ *is open (i.e.* $|a_j| > 0$, $1 \leq j \leq n$*), then* f *is real analytic on* $\mathcal{R}(a)$. *In this case* f *is infinitely differentiable and*

$$\partial_\beta f(x) = \sum_{\alpha\in\mathbb{N}_0^n} f_\alpha \partial_\beta x^\alpha. \tag{A.16}$$

Conversely, note that if the sum in (A.14) converges for some $x \in \mathbb{R}^n$, then $f_\alpha x^\alpha$ converges to zero and hence is bounded. However, while this implies that if x is in the domain of convergence of f, so is the entire rectangle $\mathcal{R}(x)$, this does not imply that the domain of convergence (apart from the question of boundary behavior) is a rectangle.

Example A.2. The series

$$\frac{1}{1-(x_1+\cdots+x_n)} = \sum_{k=0}^\infty (x_1+\cdots+x_n)^k = \sum_{k=0}^\infty \sum_{|\alpha|=k} \frac{k!}{\alpha!} x^\alpha = \sum_\alpha \frac{|\alpha|!}{\alpha!} x^\alpha$$

is obviously convergent for $|x_1|+\cdots+|x_n| < 1$, while the series

$$\frac{1}{1-|x|^2} = \sum_{k=0}^\infty (x_1^2+\cdots+x_n^2)^k = \sum_\alpha \frac{|\alpha|!}{\alpha!} x^{2\alpha}$$

is obviously convergent for $|x| < 1$. ◇

In particular, if f is real analytic we can choose $a = (r,\dots,r)$ for $r > 0$ sufficiently small to see that we have an estimate of the form $|f_\alpha| \leq M r^{-|\alpha|}$. Hence

$$F(x) = M \prod_{j=1}^n \frac{r}{r-x_j} \quad \text{or} \quad G(x) = \frac{Mr}{r-(x_1+\cdots+x_n)}$$

will be majorants (clearly the second function majorizes the first) and the derivatives $\partial^\alpha f$ will be majorized by $\partial^\alpha F$. Consequently, a real analytic function is smooth and is locally given by its Taylor expansion

$$f(x) = \sum_{\alpha\in\mathbb{N}_0^n} \frac{\partial^\alpha f(x_0)}{\alpha!} (x-x_0)^\alpha. \tag{A.17}$$

Furthermore, the Taylor coefficients at one point uniquely determine f.

Theorem A.4 (Unique continuation property)**.** *Suppose* $f \in C^\omega(U)$ *where* $U \subseteq \mathbb{R}^n$ *is connected. If* $(\partial^\alpha f)(x_0) = 0$ *for one* $x_0 \in U$ *and all* $\alpha \in \mathbb{N}_0^n$*, then* f *vanishes identically on* U.

Proof. Let $V := \{x \in U | (\partial^\alpha f)(x) = 0, \forall \alpha \in \mathbb{N}_0^n\}$. Since $x_0 \in V$ it is nonempty. Moreover, if $x \in V$, then the Taylor series vanishes and hence f vanishes on a neighborhood of x. So V is open. If $x \in \partial V$, then all derivatives of f at x vanish by continuity implying $x \in V$ and hence V is closed. Since U is connected we have $V = U$. □

Note that sums and products of real analytic functions are again real analytic. Also the reciprocal of a nonvanishing real analytic function is real analytic. And finally, the composition of real analytic functions is again real analytic. These items are not particularly difficult but nevertheless quite tedious to check and we refer to [**18**] for this, and much more!

Problem A.1. *Show Leibniz' rule.*

Problem A.2. *Fix $k \in \mathbb{R}^n$. Show that $f(x) := \mathrm{e}^{k\cdot x}$ is real analytic on $\mathbb{R}^n$.*

A.2. Integration

When working with integrals one frequently faces the need to interchange a limit with the integral. The premier tool for justifying this operation is

Theorem A.5 (Dominated convergence; Lebesgue)**.** *Let f_n be a convergent sequence of integrable functions and set $f := \lim_{n\to\infty} f_n$. Suppose there is an integrable function g such that $|f_n| \le g$. Then f is integrable and*

$$\lim_{n\to\infty} \int_U f_n d^n x = \int_U f d^n x. \tag{A.18}$$

Example A.3. Note that the existence of g is crucial: The functions $f_n(x) := \frac{1}{2n}\chi_{[-n,n]}(x)$ on $\mathbb{R}$ converge uniformly to 0 but $\int_{\mathbb{R}} f_n(x) dx = 1$. ⋄

As stated, this result only holds for functions which are integrable in the sense of Lebesgue. The problem is, that the limit of Riemann integrable functions is no longer Riemann integrable in general. Of course, from a technical point of view it is convenient to know that one does not have to worry about integrability of the limit (and this is one of the main virtues of the Lebesgue integral). On the other hand, in many situation the limit is known and the question of integrability of the limit does not even arise. So in the case where you know that the limit is Riemann integrable, you can of course apply this result since the Lebesgue integral is an extension of the Riemann integral. We will only use it in such situations.

Using the dominated convergence theorem we can easily answer the question when an integral depends continuously/differentiable on a parameter:

Lemma A.6. *Let $X \subseteq \mathbb{R}$, $Y \subseteq \mathbb{R}^n$, and $f : X \times Y \to \mathbb{C}$. Suppose $y \mapsto f(x, y)$ is integrable for every x and $x \mapsto f(x, y)$ is continuous for every y.*

Then

$$F(x) := \int_Y f(x,y)\, d^n y \tag{A.19}$$

is continuous if there is an integrable function $g(y)$ such that $|f(x,y)| \le g(y)$.

Proof. If $x_k \to x$ then $f(x_k,y) \to f(x,y)$ for every y and hence $F(x_k) = \int_Y f(x_k,y)\, d^n y \to F(x) = \int_Y f(x,y)\, d^n y$ by dominated convergence. □

Lemma A.7. *Let $X \subseteq \mathbb{R}$, $Y \subseteq \mathbb{R}^n$, and $f : X \times Y \to \mathbb{C}$. Suppose $y \mapsto f(x,y)$ is integrable for all x and $x \mapsto f(x,y)$ is differentiable for almost every y. Then*

$$F(x) := \int_Y f(x,y)\, d^n y \tag{A.20}$$

is differentiable if there is an integrable function $g(y)$ such that $|\frac{\partial}{\partial x} f(x,y)| \le g(y)$. Moreover, $y \mapsto \frac{\partial}{\partial x} f(x,y)$ is integrable and

$$F'(x) = \int_Y \frac{\partial}{\partial x} f(x,y)\, d^n y \tag{A.21}$$

in this case.

Proof. Writing

$$\frac{\partial}{\partial x} f(x,y) = \lim_{n\to\infty} \frac{f(x+n^{-1},y) - f(x,y)}{n^{-1}}$$

we see that it is measurable. Moreover, by the mean value theorem we have

$$\frac{|f(x+\varepsilon,y) - f(x,y)|}{\varepsilon} \le g(y)$$

and hence dominated convergence implies

$$\begin{aligned}\lim_{\varepsilon\to 0} \frac{F(x+\varepsilon) - F(x)}{\varepsilon} &= \lim_{\varepsilon\to 0} \int_Y \frac{f(x+\varepsilon,y) - f(x,y)}{\varepsilon} d^n y \\ &= \int_Y \frac{\partial}{\partial x} f(x,y)\, d^n y\end{aligned}$$

by dominated convergence. □

Another related question is exchanging the order of integration:

Theorem A.8 (Fubini)**.** *Let f be mesurable on $X_1 \times X_2 \subseteq \mathbb{R}^n \times \mathbb{R}^m$.*

(i) *If $f \ge 0$, then $\int f(.,x_2) d^m x_2$ and $\int f(x_1,.) d^n x_1$ are both locally integrable and*

$$\begin{aligned}\iint_{X_1\times X_2} f(x_1,x_2) d^{n+m}(x_1,x_2) &= \int_{X_2} \left(\int_{X_1} f(x_1,x_2) d^n x_1 \right) d^m x_2 \\ &= \int_{X_1} \left(\int_{X_2} f(x_1,x_2) d^m x_2 \right) d^n x_1.\end{aligned} \tag{A.22}$$

(ii) *If f is complex-valued, then f is integrable if and only if either of the marginal integrals*

$$x_2 \mapsto \int_{X_1} |f(x_1, x_2)| d^n x_1 \tag{A.23}$$

or

$$x_1 \mapsto \int_{X_2} |f(x_1, x_2)| d^m x_2 \tag{A.24}$$

is integrable, in which case they both are, and (A.22) *holds.*

In particular, if $f(x_1, x_2)$ is either nonnegative or integrable, then the order of integration can be interchanged. The case of nonnegative functions is also called **Tonelli's theorem**.[1] In the general case the integrability condition is crucial, as the following example shows.

Example A.4. Let $X := [0,1] \times [0,1]$ and consider

$$f(x, y) = \frac{x - y}{(x + y)^3}.$$

Then

$$\int_0^1 \int_0^1 f(x, y) dx\, dy = -\int_0^1 \frac{1}{(1 + y)^2} dy = -\frac{1}{2}$$

but (by symmetry)

$$\int_0^1 \int_0^1 f(x, y) dy\, dx = \int_0^1 \frac{1}{(1 + x)^2} dx = \frac{1}{2}.$$

Consequently f cannot be integrable over X (verify this directly). $\diamond$

Theorem A.9 (Jensen's inequality)**.** *Let $\varphi : (a, b) \to \mathbb{R}$ be convex ($a = -\infty$ or $b = \infty$ being allowed). Suppose $X \subset \mathbb{R}^n$ has finite measure $|X| < \infty$ and $f : X \to (a, b)$ is integrable. Then the negative part of $\varphi \circ f$ is integrable and*

$$\varphi\Big(\frac{1}{|X|} \int_X f\, d^n x\Big) \le \frac{1}{|X|} \int_X (\varphi \circ f)\, d^n x. \tag{A.25}$$

For $f \ge 0$ the requirement that f is integrable can be dropped if $\varphi(b)$ is understood as $\lim_{x \to b} \varphi(x)$. Similarly, if $\varphi(x)$ depends only on the absolute value of x, finiteness of the right-hand side will imply integrability of f.

Theorem A.10 (change of variables)**.** *Let $U, V \subseteq \mathbb{R}^n$ and suppose $f \in C^1(U, V)$ is a diffeomorphism and denote by $J_f = \det(\frac{\partial f}{\partial x})$ the Jacobi determinant of f. Then*

$$\int_U g(f(x)) |J_f(x)| d^n x = \int_V g(y) d^n y \tag{A.26}$$

whenever g is nonnegative or integrable over V.

[1] Leonida Tonelli (1885–1946), Italian mathematician

Example A.5. For example, we can consider **polar coordinates** $T_2 : [0,\infty) \times [0, 2\pi) \to \mathbb{R}^2$ defined by

$$T_2(\rho,\varphi) := (\rho\cos(\varphi), \rho\sin(\varphi)).$$

Then

$$\det \frac{\partial T_2}{\partial(\rho,\varphi)} = \det \begin{vmatrix} \cos(\varphi) & -\rho\sin(\varphi) \\ \sin(\varphi) & \rho\cos(\varphi) \end{vmatrix} = \rho$$

and one has

$$\int_U f(\rho\cos(\varphi), \rho\sin(\varphi))\rho\, d(\rho,\varphi) = \int_{T_2(U)} f(x) d^2x.$$

Note that T_2 is only bijective when restricted to $(0,\infty) \times [0, 2\pi)$. However, since the set $\{0\} \times [0, 2\pi)$ is of measure zero, it does not contribute to the integral on the left. Similarly, its image $T_2(\{0\} \times [0, 2\pi)) = \{0\}$ does not contribute to the integral on the right. ⋄

Example A.6. We can use the previous example to obtain the transformation formula for **spherical coordinates** in $\mathbb{R}^n$ by induction. We illustrate the process for $n = 3$. To this end let $x = (x_1, x_2, x_3)$ and start with spherical coordinates in $\mathbb{R}^2$ (which are just polar coordinates) for the first two components:

$$x = (\rho\cos(\varphi), \rho\sin(\varphi), x_3), \qquad \rho \in [0,\infty),\ \varphi \in [0, 2\pi).$$

Next use polar coordinates for (ρ, x_3):

$$(\rho, x_3) = (r\sin(\theta), r\cos(\theta)), \qquad r \in [0,\infty),\ \theta \in [0,\pi].$$

Note that the range for θ follows since $\rho \geq 0$. Moreover, observe that $r^2 = \rho^2 + x_3^2 = x_1^2 + x_2^2 + x_3^2 = |x|^2$ as already anticipated by our notation. In summary,

$$x = T_3(r,\varphi,\theta) := (r\sin(\theta)\cos(\varphi), r\sin(\theta)\sin(\varphi), r\cos(\theta)).$$

Furthermore, since T_3 is the composition with T_2 acting on the first two coordinates with the last unchanged and polar coordinates P acting on the first and last coordinate, the chain rule implies

$$\det \frac{\partial T_3}{\partial(r,\varphi,\theta)} = \det \frac{\partial T_2}{\partial(\rho,\varphi,x_3)}\Big|_{\substack{\rho=r\sin(\theta) \\ x_3=r\cos(\theta)}} \det \frac{\partial P}{\partial(r,\varphi,\theta)} = r^2\sin(\theta).$$

Hence one has

$$\int_U f(T_3(r,\varphi,\theta)) r^2 \sin(\theta) d(r,\varphi,\theta) = \int_{T_3(U)} f(x) d^3x.$$

Again T_3 is only bijective on $(0,\infty) \times [0, 2\pi) \times (0,\pi)$.

It is left as an exercise to check that the extension to arbitrary dimensions $T_n : [0,\infty) \times [0, 2\pi) \times [0,\pi]^{n-2} \to \mathbb{R}^n$ is given by

$$x = T_n(r,\varphi,\theta_1,\ldots,\theta_{n-2})$$

with

$$\begin{aligned}
x_1 &= r\cos(\varphi)\sin(\theta_1)\sin(\theta_2)\sin(\theta_3)\cdots\sin(\theta_{n-2}),\\
x_2 &= r\sin(\varphi)\sin(\theta_1)\sin(\theta_2)\sin(\theta_3)\cdots\sin(\theta_{n-2}),\\
x_3 &= r\cos(\theta_1)\sin(\theta_2)\sin(\theta_3)\cdots\sin(\theta_{n-2}),\\
x_4 &= r\cos(\theta_2)\sin(\theta_3)\cdots\sin(\theta_{n-2}),\\
&\vdots\\
x_{n-1} &= r\cos(\theta_{n-3})\sin(\theta_{n-2}),\\
x_n &= r\cos(\theta_{n-2}).
\end{aligned}$$

The Jacobi determinant is given by

$$\det\frac{\partial T_n}{\partial(r,\varphi,\theta_1,\dots,\theta_{n-2})} = r^{n-1}\sin(\theta_1)\sin(\theta_2)^2\cdots\sin(\theta_{n-2})^{n-2}. \qquad \diamond$$

We also mention the following rule for integrating radial functions.

Lemma A.11. *There is a measure σ^{n-1} on the* ***unit sphere*** $S^{n-1} := \partial B_1(0) = \{x\in\mathbb{R}^n \mid |x|=1\}$, *which is rotation invariant and satisfies*

$$\int_{\mathbb{R}^n} g(x)d^nx = \int_0^\infty\int_{S^{n-1}} g(r\omega)r^{n-1}d\sigma^{n-1}(\omega)dr, \tag{A.27}$$

for every integrable (or positive) function g.

Moreover, the surface area of S^{n-1} is given by

$$S_n := \sigma^{n-1}(S^{n-1}) = nV_n, \tag{A.28}$$

where $V_n := \lambda^n(B_1(0))$ is the volume of the unit ball in $\mathbb{R}^n$, and if $g(x) = \tilde{g}(|x|)$ is radial we have

$$\int_{\mathbb{R}^n} g(x)d^nx = S_n\int_0^\infty \tilde{g}(r)r^{n-1}dr. \tag{A.29}$$

Clearly in spherical coordinates the surface measure is given by

$$d\sigma^{n-1} = \sin(\theta_1)\sin(\theta_2)^2\cdots\sin(\theta_{n-2})^{n-2}d\varphi\, d\theta_1\cdots d\theta_{n-2}. \tag{A.30}$$

Example A.7. Let us compute the volume of a ball in $\mathbb{R}^n$:

$$V_n(r) := \int_{\mathbb{R}^n}\chi_{B_r(0)}d^nx.$$

By the simple scaling transform $f(x) = rx$ we obtain $V_n(r) = V_n(1)r^n$ and hence it suffices to compute $V_n := V_n(1)$.

To this end we use (Problem A.3)

$$\begin{aligned}
\pi^{n/2} &= \int_{\mathbb{R}^n} \mathrm{e}^{-|x|^2}d^nx = nV_n\int_0^\infty \mathrm{e}^{-r^2}r^{n-1}dr = \frac{nV_n}{2}\int_0^\infty \mathrm{e}^{-s}s^{n/2-1}ds\\
&= \frac{nV_n}{2}\Gamma(\frac{n}{2}) = V_n\Gamma(\frac{n}{2}+1),
\end{aligned}$$

where Γ is the gamma function (Problem A.4). Hence

$$V_n = \frac{\pi^{n/2}}{\Gamma(\frac{n}{2}+1)}. \tag{A.31}$$

Since $\Gamma(z+1) = z\Gamma(z)$ (see Problem A.4), the case $V_1 = 2$ shows

$$\Gamma(\frac{1}{2}) = \sqrt{\pi} \tag{A.32}$$

from which we conclude the well-known values $V_2 = \pi$ and $V_3 = \frac{4\pi}{3}$. ⋄

Example A.8. The above lemma can be used to determine when a radial function is integrable. For example, we obtain

$$\int_{B_1(0)} |x|^{-\gamma} d^n x < \infty \Leftrightarrow \gamma < n, \qquad \int_{\mathbb{R}^n \setminus B_1(0)} |x|^{-\gamma} d^n x < \infty \Leftrightarrow \gamma > n. \quad \diamond$$

Next we turn to approximation of f. The idea is to replace the value $f(x)$ by a suitable average computed from the values in a neighborhood. This is done by choosing a nonnegative bump function ϕ, whose area is normalized to 1, and considering the **convolution**

$$(\phi * f)(x) := \int_{\mathbb{R}^n} \phi(x-y) f(y) d^n y = \int_{\mathbb{R}^n} \phi(y) f(x-y) d^n y. \tag{A.33}$$

For example, if we choose $\phi_r = |B_r(0)|^{-1} \chi_{B_r(0)}$ to be the characteristic function of a ball centered at 0, then $(\phi_r * f)(x)$ will be precisely the average of the values of f in the ball $B_r(x)$. In the general case we can think of $(\phi * f)(x)$ as a weighted average. Moreover, if we choose ϕ differentiable, we can interchange differentiation and integration to conclude that $\phi * f$ will also be differentiable. Iterating this argument shows that $\phi * f$ will have as many derivatives as ϕ. Finally, if the set over which the average is computed (i.e., the support of ϕ) shrinks, we expect $(\phi * f)(x)$ to get closer and closer to $f(x)$.

Lemma A.12. *Let ϕ_ε, $\varepsilon \in (0,1]$, be a family of nonnegative integrable functions satisfying*

(i) $\int_{\mathbb{R}^n} \phi_\varepsilon(x) d^n x = 1$ *for all* $\varepsilon > 0$.

(ii) *For every* $\rho > 0$ *we have* $\lim_{\varepsilon \downarrow 0} \int_{|x| \ge \rho} \phi_\varepsilon(x) d^n x = 0$.

Then for every bounded continuous function f we have

$$\lim_{\varepsilon \downarrow 0} (\phi_\varepsilon * f)(x) = f(x) \tag{A.34}$$

locally uniformly. If f vanishes as $|x| \to \infty$ the limit is even uniform.

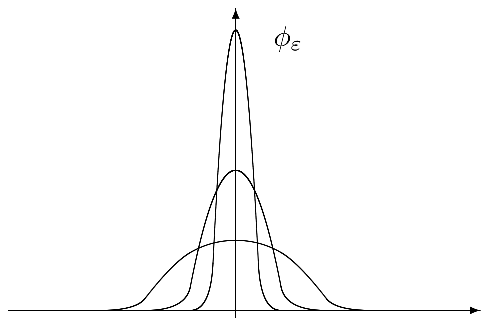


Figure A.1. The Friedrichs mollifier

Proof. Fix $r > 0$ and $\delta > 0$. Let $M := \sup_{x\in\mathbb{R}^n} |g(x)|$. Since f is uniformly continuous on $\bar{B}_{r+1}(0)$ we can choose a $\rho \le 1$ such that $|f(x) - f(y)| \le \delta$ for all $x \in \bar{B}_r(0)$ and y with $|x - y| < \rho$. Then

$$
\begin{aligned}
|(\phi_\varepsilon * f)(x) - f(x)| &\le \int_{\mathbb{R}^n} \phi_\varepsilon(x-y)|f(y) - f(x)|d^n y \\
&= \int_{|y-x|<\rho} \phi_\varepsilon(x-y)|f(y)-f(x)|d^n y + \int_{|y-x|\ge\rho} \phi_\varepsilon(x-y)|f(y)-f(x)|d^n y \\
&\le \delta + 2M \int_{|y|\ge\rho} \phi_\varepsilon(y) d^n y.
\end{aligned}
$$

Hence $\limsup_{\varepsilon\downarrow 0} |(\phi_\varepsilon * f)(x) - f(x)| \le \delta$ uniformly for $x \in \bar{B}_r(0)$. Since δ is arbitrary the claim follows. □

If in the situation of the above lemma $\phi_\varepsilon \in C_c^\infty(\mathbb{R}^n)$, then ϕ_ε is called **mollifier** and in this case Lemma A.7 shows that $\phi_\varepsilon * f \in C^\infty(\mathbb{R}^n)$ with

$$
\partial_\alpha(\phi * f) = (\partial_\alpha \phi) * f. \tag{A.35}
$$

Example A.9. Choosing any nonnegative function ϕ with $\int_{\mathbb{R}^n} \phi\, d^n x = 1$, the family $\phi_\varepsilon(x) = \varepsilon^{-n}\phi(\frac{x}{\varepsilon})$ will satisfy the conditions of the previous lemma and the standard (also Friedrichs) mollifier corresponds to

$$
\phi(x) := \begin{cases} \frac{1}{c}\exp(\frac{1}{|x|^2-1}), & |x| < 1, \\ 0, & |x| \ge 1, \end{cases} \qquad c := \int_{B_1(0)} \exp(\frac{1}{|x|^2-1}) d^n x.
$$

To show that this function is indeed smooth it suffices to show that all right derivatives of $f(r) = \exp(\frac{1}{r})$ at $r = 0$ vanish, which can be done using l'Hôpital's rule. ◇

Finally we come to the Gauss–Green theorem. Here we will need the case of a surface arising as the boundary of some domain $U \subset \mathbb{R}^n$. To this end we recall that $U \subseteq \mathbb{R}^n$ is said to have a C^1 boundary if around any point

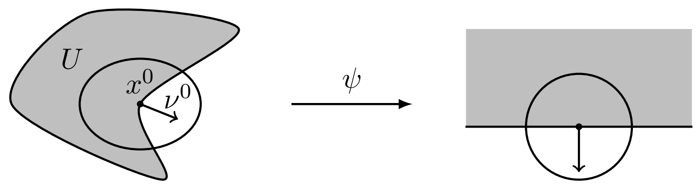


Figure A.2. Straightening out the boundary

$x^0 \in \partial U$ we can find a small neighborhood $O(x^0)$ so that, after a possible permutation of the coordinates, we can write

$$U \cap O(x^0) = \{x \in O(x^0) | x_n > \gamma(x_1, \dots, x_{n-1})\} \tag{A.36}$$

with $\gamma \in C^1$. Similarly we could define C^k or $C^{k,\theta}$ domains. We have

$$\partial U \cap O(x^0) = \{x \in O(x^0) | x_n = \gamma(x_1, \dots, x_{n-1})\} \tag{A.37}$$

and the **outward pointing unit normal vector** of ∂U is defined by

$$\nu := \frac{1}{\sqrt{1 + (\partial_1\gamma)^2 + \cdots + (\partial_{n-1}\gamma)^2}} (\partial_1\gamma, \dots, \partial_{n-1}\gamma, -1). \tag{A.38}$$

The surface integral of a vector field $u : \overline{U} \to \mathbb{R}^n$ is defined as

$$\int_{\partial U} u \cdot \nu \, dS := \int_U u(x_1, \dots, x_{n-1}, \gamma) \cdot (\partial_1\gamma, \dots, \partial_{n-1}\gamma, -1) d^{n-1}x. \tag{A.39}$$

The surface measure dS is obtained by choosing $u = \nu$, that is,

$$dS = \sqrt{1 + (\partial_1\gamma)^2 + \cdots + (\partial_{n-1}\gamma)^2} d^{n-1}x. \tag{A.40}$$

If ∂U cannot be covered by a single neighborhood, the domain has to be split into smaller pieces such that this is possible and the integral has to be defined as a sum.

Moreover, we have a change of coordinates $y = \psi(x)$ such that in these coordinates the boundary is given by (part of) the hyperplane $y_n = 0$. Explicitly we have $\psi \in C^1_b(U \cap O(x^0), V_+(y^0))$ given by

$$\psi(x) = (x_1, \dots, x_{n-1}, x_n - \gamma(x_1, \dots, x_{n-1})) \tag{A.41}$$

with inverse $\psi^{-1} \in C^1_b(V_+(y^0), U \cap O(x^0))$ given by

$$\psi^{-1}(y) = (y_1, \dots, y_{n-1}, y_n + \gamma(y_1, \dots, y_{n-1})). \tag{A.42}$$

Clearly, $\nu = (0, \dots, 0, -1)$ and $dS = d^{n-1}y$ in the new coordinates. This is known as straightening out the boundary (see Figure A.2).

Theorem A.13 (Gauss–Green)**.** *If U is a bounded C^1 domain in $\mathbb{R}^n$ and $u \in C^1(\overline{U}, \mathbb{R}^n)$ is a vector field, then*

$$\int_U (\operatorname{div} u) d^n x = \int_{\partial U} u \cdot \nu \, dS. \tag{A.43}$$

Here $\operatorname{div} u = \sum_{j=1}^n \partial_j u_j$ *is the* ***divergence*** *of a vector field.*

This theorem is also known as the **divergence theorem** or as **Ostrogradski formula**[2]. In the one-dimensional case it is just the fundamental theorem of calculus.

Example A.10. Our main application of the Gauss–Green theorem will be for balls. Hence let us at least verify the case of the unit ball in $\mathbb{R}^3$. By linearity it suffices to consider the case where the vector field u is parallel to one of the coordinate axes, say $u = (0, 0, u_3)$ such that $\operatorname{div} u = \partial_3 u_3$. Abbreviating $\rho := \sqrt{x_1^2 + x_2^2}$ we obtain

$$\begin{aligned}\int_{B_1} (\operatorname{div} u) d^3x &= \int_{\rho\le 1} \int_{-\sqrt{1-\rho^2}}^{\sqrt{1-\rho^2}} \frac{\partial u_3}{\partial x_3}(x) dx_3\, d(x_1, x_2)\\ &= \int_{\rho\le 1} \Big(u_3(x_1, x_2, \sqrt{1-\rho^2}) - u_3(x_1, x_2, -\sqrt{1-\rho^2}) \Big) d(x_1, x_2).\end{aligned}$$

Parametrizing the upper/lower hemisphere $S^2_\pm := \{x | \pm x_3 > 0, |x| = 1\}$ using $x_3 = \pm\sqrt{1-\rho^2}$ we obtain $dS = \frac{1}{\sqrt{1-\rho^2}} d(x_1, x_2)$. Since $\nu = \frac{x}{|x|}$ (remember that ν needs to point outwards) this gives

$$\int_{S^2_\pm} u\cdot\nu\, dS = \pm \int_{\rho\le 1} u_3(x_1, x_2, \pm\sqrt{1-\rho^2}) d(x_1, x_2)$$

and verifies the Gauss–Green theorem for the unit ball in $\mathbb{R}^3$. Of course the calculation easily generalizes to $\mathbb{R}^n$. ⋄

Applying the Gauss–Green theorem to a product fg we obtain

Corollary A.14 (Integration by parts)**.** *We have*

$$\int_U (\partial_j f) g\, d^n x = \int_{\partial U} f g \nu_j dS - \int_U f(\partial_j g) d^n x, \qquad 1 \le j \le n, \tag{A.44}$$

for $f, g \in C^1(\overline{U})$.

Problem* A.3. *Show*

$$I_n := \int_{\mathbb{R}^n} \mathrm{e}^{-|x|^2} d^n x = \pi^{n/2}.$$

(Hint: Use Fubini to show $I_n = I_1^n$ *and compute* I_2 *using polar coordinates.)*

Problem* A.4. *The* ***gamma function*** *is defined via*

$$\Gamma(z) := \int_0^\infty x^{z-1} \mathrm{e}^{-x} dx, \qquad \operatorname{Re}(z) > 0. \tag{A.45}$$

[2]Mikhail Ostrogradsky (1801–1862), Ukrainian mathematician, mechanician and physicist

Verify that the integral converges in the indicated half-plane. Use integration by parts to show

$$\Gamma(z+1) = z\Gamma(z), \qquad \Gamma(1) = 1. \tag{A.46}$$

Conclude $\Gamma(n) = (n-1)!$ *for* $n \in \mathbb{N}$.

Problem* A.5. *Show that* $\Gamma(\frac{1}{2}) = \sqrt{\pi}$. *Moreover, show*

$$\Gamma(n+\frac{1}{2}) = \frac{(2n)!}{4^n n!}\sqrt{\pi}$$

(Hint: Use the change of coordinates $x = t^2$ *and then use Problem A.3.)*

Problem* A.6. *Show*

$$\alpha^{-z}\Gamma(z) = \int_0^1 s^{\alpha-1}\log(1/s)^{z-1} ds, \qquad \mathrm{Re}(z) > 0, \alpha > 0.$$

Problem A.7. *Show*

$$\int_{-\infty}^{\infty} \frac{\mathrm{e}^{zx}}{a+\mathrm{e}^x} dx = \int_0^{\infty} \frac{y^{z-1}}{a+y} dy = \frac{\pi a^{z-1}}{\sin(z\pi)}, \quad 0 < \mathrm{Re}(z) < 1,\ a \in \mathbb{C}\setminus(-\infty,0].$$

(Hint: First reduce it to the case $a = 1$. *Then, use a contour consisting of the straight lines connecting the points* $-R$, R, $R+2\pi\mathrm{i}$, $-R+2\pi\mathrm{i}$. *Evaluate the contour integral using the residue theorem and let* $R \to \infty$. *Show that the contributions from the vertical lines vanish in the limit and relate the integrals along the horizontal lines.)*

Problem A.8. *Show that the* ***Beta function*** *satisfies*

$$B(u,v) := \int_0^1 t^{u-1}(1-t)^{v-1} dt = \frac{\Gamma(u)\Gamma(v)}{\Gamma(u+v)}, \quad \mathrm{Re}(u) > 0,\ \mathrm{Re}(v) > 0.$$

A few other common forms are

$$\begin{aligned} B(u,v) &= 2\int_0^{\pi/2} \sin(\theta)^{2u-1}\cos(\theta)^{2v-1} d\theta \\ &= \int_0^{\infty} \frac{s^{u-1}}{(1+s)^{u+v}} ds = 2^{-u-v+1}\int_{-1}^1 (1+s)^{u-1}(1-s)^{v-1} ds \\ &= n\int_0^1 s^{nu-1}(1-s^n)^{v-1} ds, \qquad n > 0. \end{aligned}$$

Use this to establish ***Euler's reflection formula***

$$\Gamma(z)\Gamma(1-z) = \frac{\pi}{\sin(\pi z)}$$

and ***Legendre's duplication formula***[3]

$$\Gamma(2z) = \frac{2^{2z-1}}{\sqrt{\pi}}\Gamma(z)\Gamma(z+\frac{1}{2}).$$

[3] Adrien-Marie Legendre (1752–1833), French mathematician

Conclude that the Gamma function has no zeros on $\mathbb{C}$.

(Hint: Start with $\Gamma(u)\Gamma(v)$ and make a change of variables $x = ts$, $y = t(1-s)$. For the reflection formula evaluate $B(z, 1-z)$ using Problem A.7. For the duplication formula relate $B(z, z)$ and $B(\frac{1}{2}, z)$.)

Problem A.9. *Verify the Gauss–Green theorem (by computing both integrals) in the case $u(x) = x$ and $U = B_1(0) \subset \mathbb{R}^n$.*

Problem A.10. *Let U be a bounded C^1 domain in $\mathbb{R}^n$ and set $\frac{\partial g}{\partial \nu} := \nu \cdot \nabla g$. Verify* ***Green's first identity***

$$\int_U (f\Delta g + \nabla f \cdot \nabla g) d^n x = \int_{\partial U} f \frac{\partial g}{\partial \nu} dS$$

for $f \in C^1(\overline{U})$, $g \in C^2(\overline{U})$ and ***Green's second identity***

$$\int_U (f\Delta g - g\Delta f) d^n x = \int_{\partial U} \left(f \frac{\partial g}{\partial \nu} - g \frac{\partial f}{\partial \nu} \right) dS$$

for $f, g \in C^2(\overline{U})$.

Problem A.11. *Let f be a locally integrable function and define*

$$\fint_{\partial B_r(x)} f(y) dS(y) := \frac{1}{n V_n r^{n-1}} \int_{\partial B_r(x)} f(y) dS(y),$$
$$\fint_{B_r(x)} f(y) d^n y := \frac{1}{V_n r^n} \int_{B_r(x)} f(y) d^n y.$$

Suppose $f \in C^2(\mathbb{R}^n)$ and show

$$\frac{\partial}{\partial r} \fint_{\partial B_r(x)} f(y) dS(y) = \frac{r}{n} \fint_{B_r(x)} (\Delta f)(y) d^n y,$$
$$\frac{\partial}{\partial r} \fint_{B_r(x)} f(y) d^n y = \frac{n}{r} \left(\fint_{\partial B_r(x)} f(y) dS(y) - \fint_{B_r(x)} f(y) d^n y \right).$$

Problem A.12 (Leibniz integral rule)**.** *Suppose $f \in C(R)$ with $\frac{\partial f}{\partial x}(x, y) \in C(R)$, where $R = [a_1, b_1] \times [a_2, b_2]$ is some rectangle, and $g \in C^1([a_1, b_1], [a_2, b_2])$. Show*

$$\frac{d}{dx} \int_{a_2}^{g(x)} f(x, y) dy = f(x, g(x)) g'(x) + \int_{a_2}^{g(x)} \frac{\partial f}{\partial x}(x, y) dy.$$

A.3. Fourier series

Given an integrable function f on $(-\pi, \pi)$ we can define its **Fourier series**

$$F(f)(x) := \frac{a_0}{2} + \sum_{n \in \mathbb{N}} \big(a_n \cos(nx) + b_n \sin(nx) \big), \tag{A.47}$$

where the corresponding Fourier coefficients are given by

$$a_n := \frac{1}{\pi}\int_{-\pi}^{\pi}\cos(nx)f(x)dx, \qquad b_n := \frac{1}{\pi}\int_{-\pi}^{\pi}\sin(nx)f(x)dx \tag{A.48}$$

for $n \in \mathbb{N}_0$. At this point (A.47) is just a formal expression and it was (and to some extend still is) a fundamental question in mathematics to understand in what sense the above series converges. For example, does it converge at a given point (e.g. at every point of continuity of f) or when does it converge uniformly?

For our purpose the complex form

$$F(f)(x) = \sum_{n\in\mathbb{Z}} \hat{f}_n \mathrm{e}^{\mathrm{i}nx}, \qquad \hat{f}_n := \frac{1}{2\pi}\int_{-\pi}^{\pi}\mathrm{e}^{-\mathrm{i}ny}f(y)dy \tag{A.49}$$

will be more convenient. The connection is given via $\hat{f}_{\pm n} = \frac{a_n \mp \mathrm{i}b_n}{2}$, $n \in \mathbb{N}_0$ (with the convention $b_0 = 0$).

The key observation is the following orthogonality

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\mathrm{e}^{\mathrm{i}my}\mathrm{e}^{\mathrm{i}ny}dy = \begin{cases} 1, & m = n, \\ 0, & m \neq n, \end{cases} \tag{A.50}$$

which shows that if the Fourier series converges uniformly, such that we can interchange summation and integration, the Fourier coefficients are necessarily given by the above formula.

To investigate convergence of the Fourier series let us introduce the associated Fourier polynomial of order n as

$$F_n(f)(x) := \sum_{m=-n}^{n} \hat{f}_m \mathrm{e}^{\mathrm{i}mx}. \tag{A.51}$$

Lemma A.15. *The Fourier coefficients of a square integrable function are square summable and we have* ***Bessel's inequality***

$$\sum_{n\in\mathbb{Z}} |\hat{f}_n|^2 \le \frac{1}{2\pi}\int_{-\pi}^{\pi}|f(y)|^2 dy. \tag{A.52}$$

Proof. First of all note that

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}|F_n(f)(x)|^2dx = \sum_{j,k=-n}^{n}\frac{\hat{f}_j\hat{f}_k^*}{2\pi}\int_{-\pi}^{\pi}\mathrm{e}^{\mathrm{i}(j-k)x}dx = \sum_{m=-n}^{n}|\hat{f}_k|^2.$$

Furthermore, setting $g_n := f - F_n(f)$

$$\begin{aligned}\int_{-\pi}^{\pi}|f(x)|^2dx &= \int_{-\pi}^{\pi}|F_n(f)(x) + g_n(x)|^2dx \\ &= \int_{-\pi}^{\pi}|F_n(f)(x)|^2dx + 2\mathrm{Re}\left(\int_{-\pi}^{\pi}g_n(x)F_n(f)(x)^*dx\right) + \int_{-\pi}^{\pi}|g_n(x)|^2dx\end{aligned}$$

and using

$$\int_{-\pi}^{\pi} g_n(x) \mathrm{e}^{\mathrm{i}mx} dx = 2\pi(\hat{f}_m - \hat{f}_m) = 0, \qquad |m| \le n,$$

shows

$$\int_{-\pi}^{\pi} g_n(x) F_n(f)(x)^* dx = 0,$$

This shows

$$\sum_{k=1}^{n} |\hat{f}_k|^2 + \frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x) - F_n(f)(x)|^2 dx = \frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x)|^2 dx$$

from which the claim follows. □

Observe that our proof in fact shows that we will have equality if the Fourier series converges uniformly. This is known as Parseval's identity[4] and shows that the Fourier transform is unitary when viewed as a map from the Hilbert space of square integrable functions to the Hilbert space of square summable sequences. Making this precise requires further notation from functional analysis as well as from measure theory and will not be addressed here. For now the inequality will be sufficient to establish pointwise convergence (we refer the curious reader to **[17]** for more information).

This lemma implies in particular, that the Fourier coefficients of a square integrable function converge to zero. Since we can approximate integrable functions by square integrable ones, this result extends to integrable functions:

Corollary A.16 (Riemann–Lebesgue lemma). *Suppose f is integrable, then the Fourier coefficients $\hat{f}_k$ converge to zero as $|k| \to \infty$.*

Proof. Consider $f_n := \min(1, \frac{n}{|f|}) f$. Then f_n is bounded ($|f_n| \le n$) and hence square integrable. Moreover since $f_n(x) \to f(x)$ pointwise and $|f_n| \le |f|$, dominated convergence shows that

$$|\hat{f}_{n,k} - \hat{f}_k| \le \int_{-\pi}^{\pi} |f_n(y) - f(y)| dy$$

converges to 0 as $n \to \infty$. Now fix $\varepsilon > 0$ and choose n such that $|\hat{f}_{n,k} - \hat{f}_k| \le \frac{\varepsilon}{2}$ for $k \in \mathbb{Z}$. Since $\hat{f}_{n,k} \to 0$ as $|k| \to \infty$ by Bessel's inequality we can find an index K such that $|\hat{f}_{n,k}| \le \frac{\varepsilon}{2}$ for $|k| \ge K$. In summary we get $|\hat{f}_k| \le |\hat{f}_{n,k} - \hat{f}_k| + |\hat{f}_{n,k}| \le \varepsilon$ for $|k| \ge K$ as desired. □

[4] Marc-Antoine Parseval (1755–1836), French mathematician

Recall that a periodic function $f \in C_{per}[-\pi,\pi]$ is called uniformly **Hölder continuous** with exponent $\gamma \in (0,1]$ if

$$[f]_\gamma := \sup_{x\neq y\in[-\pi,\pi]} \frac{|f(x)-f(y)|}{|x-y|^\gamma} \tag{A.53}$$

is finite. We will denote these functions by $C^{0,\gamma}_{per}[-\pi,\pi]$. Clearly, any Hölder continuous function is uniformly continuous and, in the special case $\gamma = 1$, we obtain the **Lipschitz continuous** functions. Note that for $\gamma = 0$ the Hölder condition boils down to boundedness.

Theorem A.17. *Suppose*

$$\frac{f(x)-f(x_0)}{x-x_0} \tag{A.54}$$

is integrable (e.g. f is Hölder continuous), then

$$\lim_{m,n\to\infty} \sum_{k=-m}^{n} \hat{f}_k \mathrm{e}^{\mathrm{i}kx_0} = f(x_0). \tag{A.55}$$

Proof. Without loss of generality we can assume $x_0 = 0$ (by shifting $x \to x - x_0$ modulo 2π implying $\hat{f}_k \to \mathrm{e}^{-\mathrm{i}kx_0}\hat{f}_k$) and $f(x_0) = 0$ (by linearity since the claim is trivial for constant functions). Then by assumption

$$g(x) := \frac{f(x)}{\mathrm{e}^{\mathrm{i}x}-1}$$

is integrable and $f(x) = (\mathrm{e}^{\mathrm{i}x}-1)g(x)$ implies $\hat{f}_k = \hat{g}_{k-1} - \hat{g}_k$ and hence

$$\sum_{k=-m}^{n} \hat{f}_k = \hat{g}_{-m-1} - \hat{g}_n.$$

Now the claim follows from the Riemann–Lebesgue lemma. □

If we look at symmetric partial sums $F_n(f)$ we can do even better.

Corollary A.18 (Dirichlet–Dini[5] criterion)**.** *Suppose there is some α such that*

$$\frac{f(x_0+x)+f(x_0-x)-2\alpha}{x} \tag{A.56}$$

is integrable. Then $F_n(f)(x_0) \to \alpha$.

Proof. Without loss of generality we can assume $x_0 = 0$. Now observe (since $D_n(-x) = D_n(x)$) $S_n(f)(0) = \alpha + S_n(g)(0)$, where $g(x) := \frac{1}{2}(f(x) + f(-x)) - \alpha$ and apply the previous result. □

[5] Ulisse Dini (1845–1918), Italian mathematician and politician

A consequence of this result is that continuity is not necessary for convergence. If f has a jump at x_0 such that after subtraction of an appropriate step function, the remainder is Hölder continuous, the above result shows that the Fourier series converges to $\frac{f(x_0-)+f(x_0+)}{2}$.

On the other hand, if we want summable Fourier coefficients, then continuity of f is clearly necessary. The space of integrable functions whose Fourier coefficients are summable is known as the **Wiener algebra** and an easy characterization is not known. There are however convenient sufficient conditions:

Theorem A.19 (Bernstein[6]). *Suppose that $f \in C^{0,\gamma}_{per}[-\pi,\pi]$ is Hölder continuous of exponent $\gamma > \frac{1}{2}$, then*

$$\sum_{n\in\mathbb{Z}} |\hat{f}_n| \le C_\gamma(\max|f| + [f]_\gamma).$$

Proof. The proof starts with the observation that the Fourier coefficients of $f_\delta(x) := f(x-\delta)$ are $\hat{f}_k = \mathrm{e}^{-\mathrm{i}k\delta}\hat{f}_k$. Now for $\delta := \frac{2\pi}{3}2^{-m}$ and $2^m \le |k| < 2^{m+1}$ we have $|\mathrm{e}^{\mathrm{i}k\delta}-1|^2 \ge 3$ implying

$$\begin{aligned}\sum_{2^m\le|k|<2^{m+1}} |\hat{f}_k|^2 &\le \frac{1}{3}\sum_k |\mathrm{e}^{\mathrm{i}k\delta}-1|^2|\hat{f}_k|^2 = \frac{1}{6\pi}\int_{-\pi}^{\pi}|f_\delta(x)-f(x)|^2dx\\ &\le \frac{1}{3}[f]_\gamma^2\delta^{2\gamma}\end{aligned}$$

Now the sum on the left has $2\cdot 2^m$ terms and hence Cauchy–Schwarz implies

$$\sum_{2^m\le|k|<2^{m+1}} |\hat{f}_k| \le \frac{2^{(m+1)/2}}{\sqrt{3}}[f]_\gamma\delta^\gamma = \sqrt{\frac{2}{3}}\left(\frac{2\pi}{3}\right)^\gamma 2^{(1/2-\gamma)m}[f]_\gamma.$$

Summing over m shows

$$\sum_{k\ne 0}|\hat{f}_k| \le C_\gamma[f]_\gamma$$

provided $\gamma > \frac{1}{2}$ and establishes the claim since $|\hat{f}_0| \le \max|f|$. □

We can go even further using that for $f \in C^1_{per}[-\pi,\pi]$ integration by parts shows

$$\hat{f}_n = \frac{1}{2\pi\mathrm{i}n}\int_{-\pi}^{\pi}\mathrm{e}^{-\mathrm{i}nx}f'(x)dx. \tag{A.57}$$

So if we assume that f is k times continuously differentiable with the k'th derivative Hölder continuous of exponent $\gamma > \frac{1}{2}$, then

$$\sum_{n\in\mathbb{Z}}|n|^k|\hat{f}_n| < \infty, \qquad f \in C^{k,\gamma}_{per}[-\pi,\pi],\ \gamma > \frac{1}{2}. \tag{A.58}$$

[6]Sergei Bernstein (1880–1913), Russian mathematician

Finally, note that if f is symmetric, $f(x) = f(-x)$, then all Fourier sine coefficients will vanish and we can write it as a **Fourier cosine series**

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n(f)\cos(nx), \qquad a_n = \frac{1}{\pi}\int_0^{\pi} \cos(nx) f(x) dx. \tag{A.59}$$

Similarly, if f is skew symmetric, $f(x) = -f(-x)$, then all Fourier cosine coefficients will vanish and we can write it as a **Fourier sine series**

$$f(x) = \sum_{n=1}^{\infty} b_n \sin(nx), \qquad b_n(f) = \frac{1}{\pi}\int_0^{\pi} \sin(nx) f(x) dx. \tag{A.60}$$

When applying the above results it is important to observe that while a continuous function $f \in C[0,\pi]$ can always extended to a symmetric function in $C_{per}[-\pi,\pi]$, it can be extended to a skew symmetric function in $C_{per}[-\pi,\pi]$ if and only if $f(0) = f(\pi) = 0$. Indeed, if $f \in C_{per}[-\pi,\pi]$ we have $f(0) = -f(0)$ implying $f(0) = 0$ and $f(\pi) = f(-\pi) = -f(\pi)$ implying $f(\pi) = 0$. Similarly, if we want to extend a function $f \in C^k[0,\pi]$ to a symmetric function in $f \in C^k_{per}[0,\pi]$ we need to require that all odd derivatives vanish at 0 and π. If we want to extend a function $f \in C^k[0,\pi]$ to a skew symmetric function in $f \in C^k_{per}[0,\pi]$ we need to require that all even derivatives vanish at 0 and π.

Finally, note that if the interval $[-\pi,\pi]$ is replaced by an arbitrary interval $[-L,L]$, the formulas change according to

$$F(f)(x) := \frac{a_0}{2} + \sum_{n\in\mathbb{N}} \Big(a_n \cos\big(\frac{n\pi}{L}x\big) + b_n \sin\big(\frac{n\pi}{L}x\big)\Big), \tag{A.61}$$

with

$$a_n := \frac{1}{L}\int_{-L}^{L} \cos(nx) f(x) dx, \qquad b_n := \frac{1}{L}\int_{-L}^{L} \sin(nx) f(x) dx \tag{A.62}$$

and

$$F(f)(x) = \sum_{n\in\mathbb{Z}} \hat{f}_n \mathrm{e}^{\mathrm{i}n(\pi/L)x}, \qquad \hat{f}_n := \frac{1}{2L}\int_{-L}^{L} \mathrm{e}^{-\mathrm{i}ny} f(y) dy. \tag{A.63}$$

Problem A.13. *Show that if $f \in C^{0,\gamma}_{per}[-\pi,\pi]$ is Hölder continuous, then*

$$|\hat{f}_n| \le \frac{[f]_\gamma}{2}\left(\frac{\pi}{|n|}\right)^{\gamma}, \qquad n \ne 0.$$

(Hint: What changes if you replace $\mathrm{e}^{-\mathrm{i}ny}$ by $\mathrm{e}^{-\mathrm{i}n(y+\pi/n)}$ in (A.49)? Now make a change of variables $y \to y - \pi/n$ in the integral.)

Appendix B

Real and functional analysis

B.1. Differentiable and Hölder continuous functions

The current section collects some required results for Lebesgue space from [**34**] to which we refer the reader for further details and proofs.

Given $U \subseteq \mathbb{R}^n$ the set of all bounded continuous functions $C_b(U)$ together with the sup norm

$$\|f\|_\infty := \sup_{x \in U} |f(x)| \tag{B.1}$$

is a Banach space (cf. Corollary B.36 from [**35**]). The space of continuous functions with compact support $C_c(U) \subseteq C_b(U)$ is in general not dense and its closure will be denoted by $C_0(U)$. If U is open $C_0(U)$ can be interpreted as the functions in $C_b(U)$ which vanish at the boundary

$$C_0(U) := \{f \in C(U) | \forall \varepsilon > 0, \exists K \subseteq U \text{ compact} : |f(x)| < \varepsilon,\ x \in U \setminus K\}. \tag{B.2}$$

Of course $\mathbb{R}^n$ could be replaced by any topological space up to this point.

Moreover, for U open the above norm can be augmented to handle differentiable functions by considering the space $C_b^1(U)$ of all continuously differentiable functions for which the following norm

$$\|f\|_{1,\infty} := \|f\|_\infty + \sum_{j=1}^{n} \|\partial_j f\|_\infty \tag{B.3}$$

is finite, where $\partial_j = \frac{\partial}{\partial x_j}$. Note that $\|\partial_j f\|$ for one j (or all j) is not sufficient as it is only a seminorm (it vanishes for every constant function). However, since the sum of seminorms is again a seminorm (Problem B.3) the above expression defines indeed a norm. It is also not hard to see that $C_b^1(U)$ is complete. In fact, let f^m be a Cauchy sequence, then $f^m(x)$ converges uniformly to some continuous function $f(x)$ and the same is true for the partial derivatives $\partial_j f^m(x) \to g_j(x)$. Moreover, since $f^m(x) = f^m(c, x_2, \dots, x_n) + \int_c^{x_1} \partial_1 f^m(t, x_2, \dots, x_n) dt \to f(x) = f(c, x_2, \dots, x_n) + \int_c^{x_1} g_1(t, x_2, \dots, x_n) dt$ we obtain $\partial_1 f(x) = g_1(x)$. The remaining derivatives follow analogously and thus $f^m \to f$ in $C_b^1(U)$.

To extend this approach to higher derivatives let $C^k(U)$ be the set of all complex-valued functions and recall the multi-index notation from Section A.1. Also recall that by the classical theorem of Schwarz the order in which partial derivatives are performed is irrelevant. With this notation the above considerations can be easily generalized to higher order derivatives:

Theorem B.1. *Let $U \subseteq \mathbb{R}^n$ be open. The space $C_b^k(U)$ of all functions whose partial derivatives up to order k are bounded and continuous form a Banach space with norm*

$$\|f\|_{k,\infty} := \sum_{|\alpha| \le k} \sup_{x \in U} |\partial_\alpha f(x)|. \tag{B.4}$$

An important subspace is $C_0^k(U)$ which we define as the closure of $C_c^k(U)$:

$$C_0^k(U) := \overline{C_c^k(U)}. \tag{B.5}$$

Note that the space $C_b^k(U)$ could be further refined by requiring the highest derivatives to be Hölder continuous. Recall that a function $f : U \to \mathbb{C}$ is called uniformly **Hölder continuous** with exponent $\gamma \in (0, 1]$ if

$$[f]_\gamma := \sup_{x \ne y \in U} \frac{|f(x) - f(y)|}{|x - y|^\gamma} \tag{B.6}$$

is finite. Clearly, any Hölder continuous function is uniformly continuous and, in the special case $\gamma = 1$, we obtain the **Lipschitz continuous** functions. Note that for $\gamma = 0$ the Hölder condition boils down to boundedness and also the case $\gamma > 1$ is not very interesting (Problem B.2).

Example B.1. By the mean value theorem every function $f \in C_b^1(U)$ is Lipschitz continuous with $[f]_\gamma \le \|\nabla f\|_\infty$, where $\nabla f = (\partial_1 f, \dots, \partial_n f)$ denotes the gradient. ⋄

Example B.2. The prototypical example of a Hölder continuous function is of course $f(x) = x^\gamma$ on $[0, \infty)$ with $\gamma \in (0, 1]$. In fact, without loss of generality we can assume $0 \le x < y$ and set $t = \frac{x}{y} \in [0, 1)$. Then we have

$$\frac{y^\gamma - x^\gamma}{(y - x)^\gamma} \le \frac{1 - t^\gamma}{(1 - t)^\gamma} \le \frac{1 - t}{1 - t} = 1.$$

From this one easily gets further examples since the composition of two Hölder continuous functions is again Hölder continuous (the exponent being the product). ⋄

It is easy to verify that this is a seminorm and that the corresponding space is complete.

Theorem B.2. *Let $U \subseteq \mathbb{R}^n$ be open. The space $C_b^{k,\gamma}(U)$ of all functions whose partial derivatives up to order k are bounded and Hölder continuous with exponent $\gamma \in (0,1]$ form a Banach space with norm*

$$\|f\|_{k,\gamma,\infty} := \|f\|_{k,\infty} + \sum_{|\alpha|=k} [\partial_\alpha f]_\gamma. \tag{B.7}$$

As before, observe that the closure of $C_c^k(U)$ is $C_0^{k,\gamma}(U) := C_b^{k,\gamma}(U) \cap C_0^k(U)$. Moreover, as also already noted before, in the case $\gamma = 0$ we get a norm which is equivalent to $\|f\|_{k,\infty}$ and we will set $C_b^{k,0}(U) := C_b^k(U)$ for notational convenience later on.

Note that by the mean value theorem all derivatives up to order lower than k are automatically Lipschitz continuous if U is convex.

Example B.3. So while locally, differentiability is stronger than Lipschitz continuity, globally the situation depends on the domain: The sign function is in $C_b^1(\mathbb{R}\setminus\{0\})$ but it is not in $C_b^{0,1}(\mathbb{R}\setminus\{0\})$. In fact it is not even uniformly continuous. Also observe that the fact that its derivative is Lipschitz continuous on $\mathbb{R}\setminus\{0\}$ does not help. ⋄

Moreover, every Hölder continuous function is uniformly continuous and hence has a unique extension to the closure $\overline{U}$ (cf. Theorem B.39 from **[35]**). In this sense, the spaces $C_b^{0,\gamma}(U)$ and $C_b^{0,\gamma}(\overline{U})$ are naturally isomorphic. Consequently, we can also understand $C_b^{k,\gamma}(\overline{U})$ in this fashion since for a function from $C_b^{k,\gamma}(U)$ all derivatives have a continuous extension to $\overline{U}$. For a function in $C_b^k(U)$ this will not work in general and hence we define $C_b^k(\overline{U})$ as the functions from $C_b^k(U)$ for which all derivatives have a continuous extensions to $\overline{U}$. Note that with this definition $C_b^k(\overline{U})$ is still a Banach space (since $C_b(\overline{U})$ is a closed subspace of $C_b(U)$). Finally, since Hölder continuous functions on a bounded domain are automatically bounded, we can drop the subscript b in this situation.

Theorem B.3. *Suppose $U \subset \mathbb{R}^n$ is bounded. Then $C^{0,\gamma_2}(\overline{U}) \subseteq C^{0,\gamma_1}(\overline{U}) \subseteq C(\overline{U})$ for $0 < \gamma_1 < \gamma_2 \le 1$ with the embeddings being compact.*

Proof. That we have continuous embeddings follows since $|x-y|^{-\gamma_1} = |x-y|^{-\gamma_2+(\gamma_2-\gamma_1)} \le (2r)^{\gamma_2-\gamma_1}|x-y|^{-\gamma_2}$ if $U \subseteq B_r(0)$. Moreover, that the embedding $C^{0,\gamma_1}(\overline{U}) \subseteq C(\overline{U})$ is compact follows from the Arzelà–Ascoli theorem[1] (Theorem B.40 from **[35]**). To see the remaining claim let f_m be a bounded sequence in $C^{0,\gamma_2}(\overline{U})$, explicitly $\|f_m\|_\infty \le C$ and $[f_m]_{\gamma_2} \le C$. Hence by the Arzelà–Ascoli theorem we can assume that f_m converges uniformly to some $f \in C(\overline{U})$. Moreover, taking the limit in $|f_m(x) - f_m(y)| \le C|x-y|^{\gamma_2}$ we see that we even have $f \in C^{0,\gamma_2}(\overline{U})$. To see that f is the limit of f_m in $C^{0,\gamma_1}(\overline{U})$ we need to show $[g_m]_{\gamma_1} \to 0$, where $g_m := f_m - f$. Now observe that

$$\begin{aligned}[g_m]_{\gamma_1} &\le \sup_{x\ne y\in U:|x-y|\ge\varepsilon} \frac{|g_m(x)-g_m(y)|}{|x-y|^{\gamma_1}} + \sup_{x\ne y\in U:|x-y|<\varepsilon} \frac{|g_m(x)-g_m(y)|}{|x-y|^{\gamma_1}} \\ &\le 2\|g_m\|_\infty \varepsilon^{-\gamma_1} + [g_m]_{\gamma_2}\varepsilon^{\gamma_2-\gamma_1} \le 2\|g_m\|_\infty\varepsilon^{-\gamma_1} + 2C\varepsilon^{\gamma_2-\gamma_1},\end{aligned}$$

implying $\limsup_{m\to\infty}[g_m]_{\gamma_1} \le 2C\varepsilon^{\gamma_2-\gamma_1}$ and since $\varepsilon > 0$ is arbitrary this establishes the claim. □

As pointed out in the example before, the embedding $C_b^1(U) \subseteq C_b^{0,1}(U)$ is continuous and combining this with the previous result immediately gives

Corollary B.4. *Suppose $U \subset \mathbb{R}^n$ is bounded, $k_1, k_2 \in \mathbb{N}_0$, and $0 \le \gamma_1, \gamma_2 \le 1$. Then $C^{k_2,\gamma_2}(\overline{U}) \subseteq C^{k_1,\gamma_1}(\overline{U})$ for $k_1 + \gamma_1 \le k_2 + \gamma_2$ with the embeddings being compact if the inequality is strict.*

Note that in all the above spaces we could replace complex-valued by $\mathbb{C}^n$-valued functions.

Problem B.1. *Show*

$$C_0^k(U) = \overline{C_c^\infty(U)} = \{f \in C_b^k(U) | \partial^\alpha f \in C_0(U),\ 0 \le |\alpha| \le k\}.$$

(Hint: Use mollification and observe that derivatives come for free from Lemma B.13.)

Problem B.2. *Let $U \subseteq \mathbb{R}^n$ be open. Suppose $f : U \to \mathbb{C}$ is Hölder continuous with exponent $\gamma > 1$. Show that f is constant on every connected component of U.*

Problem* B.3. *Suppose X is a vector space and $\|.\|_j$, $1 \le j \le m$, is a finite family of seminorms. Show that $\|x\| := \sum_{j=1}^m \|x\|_j$ is a seminorm. It is a norm if and only if $\|x\|_j = 0$ for all j implies $x = 0$.*

Problem* B.4. *Let $U \subseteq \mathbb{R}^n$. Show that $C_b(U)$ is a Banach space when equipped with the sup norm. Show that $\overline{C_c(U)} = C_0(U)$. (Hint: The function $m_\varepsilon(z) = \operatorname{sign}(z)\max(0, |z| - \varepsilon) \in C(\mathbb{C})$ might be useful.)*

---

[1] Cesare Arzelá (1847–1912), Italian mathematician

[1] Giulio Ascoli (1843–1896), Italian mathematician

Problem B.5. *Let $U \subseteq \mathbb{R}^n$. Show that the product of two bounded Hölder continuous functions is again Hölder continuous with*

$$[fg]_\gamma \le \|f\|_\infty [g]_\gamma + [f]_\gamma \|g\|_\infty.$$

Problem B.6. *Let $\phi \in L^1(\mathbb{R}^n)$ and $f \in C_b^{0,\gamma}(\mathbb{R}^n)$. Show*

$$[\phi * f]_\gamma \le \|\phi\|_1 [f]_\gamma.$$

B.2. Lebesgue spaces

The current section collects some required results about Lebesgue spaces from **[34]** to which we refer the reader for further details and proofs.

We fix some nonempty open subset $U \subseteq \mathbb{R}^n$ and define the L^p norm by

$$\|f\|_p := \left(\int_X |f(x)|^p \, d^n x \right)^{1/p}, \qquad 1 \le p, \tag{B.8}$$

and denote by $\mathcal{L}^p(U)$ the set of all complex-valued measurable functions for which $\|f\|_p$ is finite. First of all note that $\mathcal{L}^p(U)$ is a vector space, since $|f+g|^p \le 2^p \max(|f|,|g|)^p = 2^p \max(|f|^p, |g|^p) \le 2^p(|f|^p + |g|^p)$. Of course our hope is that $\mathcal{L}^p(U)$ is a Banach space. However, there is a small technical problem (recall that a property is said to hold almost everywhere if the set where it fails to hold is contained in a set of measure zero):

Lemma B.5. *Let f be measurable. Then*

$$\int_U |f(x)|^p \, d^n x = 0 \tag{B.9}$$

if and only if $f(x) = 0$ almost everywhere with respect to Lebesgue measure.

Thus $\|f\|_p = 0$ only implies $f(x) = 0$ for almost every x, but not for all! Hence $\|.\|_p$ is not a norm on $\mathcal{L}^p(U)$. The way out of this misery is to identify functions which are equal almost everywhere: Let

$$\mathcal{N}(U) := \{f | f(x) = 0 \text{ almost everywhere}\}. \tag{B.10}$$

Then $\mathcal{N}(U)$ is a linear subspace of $\mathcal{L}^p(U)$ and we can consider the quotient space

$$L^p(U) := \mathcal{L}^p(U)/\mathcal{N}(U). \tag{B.11}$$

If $d\mu$ is the Lebesgue measure on $X \subseteq \mathbb{R}^n$, we simply write $L^p(X)$. Observe that $\|f\|_p$ is well defined on $L^p(U)$.

Even though the elements of $L^p(U)$ are, strictly speaking, equivalence classes of functions, we will still treat them as functions for notational convenience. However, if we do so it is important to ensure that every statement made does not depend on the representative in the equivalence classes. In particular, note that for $f \in L^p(U)$ the value $f(x)$ is not well defined (unless there is a continuous representative and continuous functions with different values are in different equivalence classes, e.g., in the case of Lebesgue measure).

With this modification we are back in business since $L^p(U)$ turns out to be a Banach space. We will show this in the following sections. Moreover, note that $L^2(U)$ is a Hilbert space with scalar product given by

$$\langle f, g\rangle := \int_X f(x)^* g(x) d^n x. \tag{B.12}$$

But before that let us also define $L^\infty(U)$. It should be the set of bounded measurable functions $B(X)$ together with the sup norm. The only problem is that if we want to identify functions equal almost everywhere, the supremum is no longer independent of the representative in the equivalence class. The solution is the **essential supremum**

$$\|f\|_\infty := \inf\{C \,|\, |\{x \in U |\, |f(x)| > C\}| = 0\}, \tag{B.13}$$

where $|V|$ denotes the Lebesgue measure of a set $V \subseteq \mathbb{R}^n$. That is, C is an essential bound if $|f(x)| \le C$ almost everywhere and the essential supremum is the infimum over all essential bounds.

Example B.4. The essential sup of $\chi_{\mathbb{Q}}$ with respect to Lebesgue measure is 0. $\diamond$

As before we set

$$L^\infty(U) := B(U)/\mathcal{N}(U), \tag{B.14}$$

where $B(U)$ are the bounded functions and observe that $\|f\|_\infty$ is independent of the representative from the equivalence class.

If you wonder where the ∞ comes from, have a look at Problem B.8.

Since the support of a *function* in L^p is also not well defined one uses the **essential support** in this case:

$$\mathrm{supp}(f) = X \setminus \bigcup \{O | f = 0 \text{ almost everywhere on } O \subseteq X \text{ open}\}. \tag{B.15}$$

In other words, x is in the essential support if for every neighborhood the set of points where f does not vanish has positive measure. Here we use the same notation as for functions and it should be understood from the context which one is meant. Note that the essential support is always smaller than the support (since we get the latter if we require f to vanish everywhere on O in the above definition).

Example B.5. The support of $\chi_{\mathbb{Q}}$ is $\overline{\mathbb{Q}} = \mathbb{R}$ but the essential support with respect to Lebesgue measure is $\emptyset$ since the function is 0 a.e. $\diamond$

If X is a locally compact Hausdorff space (together with the Borel sigma algebra), a function is called **locally integrable** if it is integrable when restricted to any compact subset $K \subseteq U$. The set of all (equivalence classes of) locally integrable functions will be denoted by $L^1_{loc}(U)$. We will say that $f_n \to f$ in $L^1_{loc}(U)$ if this holds on $L^1(K)$ for all compact subsets $K \subseteq U$. Of course this definition extends to L^p for any $1 \le p \le \infty$.

sometimes we will also replace Lebesgue measure by a weighted Lebesgue measure $f(x)d^n x$ (where f is some nonnegative measurable

Theorem B.6 (Hölder's inequality)**.** *Let p and q be **dual** indices; that is,*

$$\frac{1}{p} + \frac{1}{q} = 1 \tag{B.16}$$

with $1 \le p \le \infty$. If $f \in L^p(U)$ and $g \in L^q(U)$, then $fg \in L^1(U)$ and

$$\|f\, g\|_1 \le \|f\|_p \|g\|_q. \tag{B.17}$$

Moreover, for $f \in L^p(U)$ we have

$$\|f\|_p = \sup_{\varphi \in C_c^\infty(U), \|\varphi\|_q = 1} \left| \int_U f \varphi \, d^n x \right|. \tag{B.18}$$

Corollary B.7 (Minkowski's inequality[2])**.** *Let $f, g \in L^p(U)$, $1 \le p \le \infty$. Then*

$$\|f + g\|_p \le \|f\|_p + \|g\|_p. \tag{B.19}$$

Theorem B.8 (Riesz[3]–Fischer[4])**.** *The space $L^p(U)$, $1 \le p \le \infty$, is a Banach space.*

Corollary B.9. *If $\|f_n - f\|_p \to 0$, $1 \le p \le \infty$, then there is a subsequence f_{n_j} (of representatives) which converges pointwise almost everywhere and a nonnegative function $G \in L^p(U)$ such that $|f_{n_j}(x)| \le G(x)$ almost everywhere.*

Theorem B.10. *The set $C_c(U)$ of continuous functions with compact support is dense in $L^p(U)$, $1 \le p < \infty$.*

Theorem B.11. *Consider $L^p(U)$and let q be the corresponding dual index, $\frac{1}{p} + \frac{1}{q} = 1$. Then the map $g \in L^q \mapsto \ell_g \in (L^p)^*$ given by*

$$\ell_g(f) := \int_U g f \, d^n x \tag{B.20}$$

is an isometric isomorphism for $1 \le p < \infty$. If $p = \infty$ it is at least isometric.

---

[2]Hermann Minkowski (1864–1909), German mathematician
[3]Frigyes Riesz (1880–1956), Hungarian mathematician
[4]Ernst Sigismund Fischer (1875–1954), Austrian mathematician

Note that we will sometimes also consider the case where the Lebesgue measure $d^n x$ by some weighted version $w(x)d^n x$, where w is some nonnegative measurable function. We will write $L^2(U, w(x)d^n x)$ in this case. The norm is given by

$$\|f\|_p := \left(\int_X |f(x)|^p w(x) d^n x\right)^{1/p}, \qquad 1 \le p, \tag{B.21}$$

Everything said so far carries over verbatim to this case.

Theorem B.12 (Lebesgue differentiation theorem)**.** *Let $f \in L^p_{loc}(\mathbb{R}^n)$, $1 \le p < \infty$, then for a.e. $x \in \mathbb{R}^n$ we have*

$$\lim_{r \downarrow 0} \frac{1}{|B_r(x)|} \int_{B_r(x)} |f(y) - f(x)|^p d^n y = 0. \tag{B.22}$$

The points where (B.22) *holds are called L^p* ***Lebesgue points*** *of f and simply Lebesgue points if $p = 1$.*

Note that the balls can be replaced by more general sets. For example we could use cubes instead of balls (show this).

Lemma B.13. *The convolution* (A.33) *has the following properties:*

(i) *$f(x - .)g(.)$ is integrable if and only if $f(.)g(x - .)$ is and*

$$(f * g)(x) = (g * f)(x) \tag{B.23}$$

in this case.

(ii) *Suppose $\phi \in C^k_c(\mathbb{R}^n)$ and $f \in L^1_{loc}(\mathbb{R}^n)$, then $\phi * f \in C^k(\mathbb{R}^n)$ and*

$$\partial_\alpha(\phi * f) = (\partial_\alpha \phi) * f \tag{B.24}$$

for any partial derivative of order at most k.

(iii) *We have* $\operatorname{supp}(f * g) \subseteq \overline{\operatorname{supp}(f) + \operatorname{supp}(g)}$. *In particular, if $\phi \in C^k_c(\mathbb{R}^n)$ and $f \in L^1_c(\mathbb{R}^n)$, then $\phi * f \in C^k_c(\mathbb{R}^n)$.*

(iv) *Suppose $\phi \in L^1(\mathbb{R}^n)$ and $f \in L^p(\mathbb{R}^n)$, $1 \le p \le \infty$, then their convolution is in $L^p(\mathbb{R}^n)$ and satisfies* ***Young's inequality***

$$\|\phi * f\|_p \le \|\phi\|_1 \|f\|_p. \tag{B.25}$$

(v) *Suppose $\phi \ge 0$ with $\|\phi\|_1 = 1$ and $f \in L^\infty(\mathbb{R}^n)$ real-valued, then*

$$\inf_{x \in \mathbb{R}^n} f(x) \le (\phi * f)(x) \le \sup_{x \in \mathbb{R}^n} f(x). \tag{B.26}$$

Lemma B.14. *Let ϕ_ε be an approximate identity. If $f \in L^p(\mathbb{R}^n)$ with $1 \le p < \infty$, then*

$$\lim_{\varepsilon \downarrow 0} \phi_\varepsilon * f = f \tag{B.27}$$

with the limit taken in L^p. Moreover, the convergence will also be pointwise a.e. (in particular at every Lebesgue point). In the case $p = \infty$ the claim holds for $f \in C_0(\mathbb{R}^n)$.

To formulate our result let $U \subseteq \mathbb{R}^n$, $f \in L^p(U)$ and consider the **translation operator**

$$T_a(f)(x) = \begin{cases} f(x-a), & x-a \in U, \\ 0, & \text{else}, \end{cases} \tag{B.28}$$

for fixed $a \in \mathbb{R}^n$. Then one checks $\|T_a\| = 1$ (unless $|(U-a) \cap U| = 0$ in which case $T_a \equiv 0$) and $T_a f \to f$ as $a \to 0$ for $1 \le p < \infty$ (Problem B.15).

Theorem B.15 (Kolmogorov–Riesz–Sudakov). *Let $U \subseteq \mathbb{R}^n$ be open. A subset F of $L^p(U)$, $1 \le p < \infty$, is relatively compact if and only if*

(i) *for every $\varepsilon > 0$ there is some $\delta > 0$ such that $\|T_a f - f\|_p \le \varepsilon$ for all $|a| \le \delta$ and $f \in F$.*

(ii) *for every $\varepsilon > 0$ there is some $r > 0$ such that $\|(1-\chi_{B_r(0)})f\|_p \le \varepsilon$ for all $f \in F$.*

Of course the last condition is void if U is bounded.

Our final result is known as the **fundamental lemma of the calculus of variations**.

Lemma B.16. *Suppose $X \subseteq \mathbb{R}^n$ is open and $f \in L^1_{loc}(X)$. (i) If f is real-valued then*

$$\int_X \varphi(x) f(x) d^n x \ge 0, \qquad \forall \varphi \in C_c^\infty(X),\ \varphi \ge 0, \tag{B.29}$$

if and only if $f(x) \ge 0$ (a.e.). (ii) Moreover,

$$\int_X \varphi(x) f(x) d^n x = 0, \qquad \forall \varphi \in C_c^\infty(X),\ \varphi \ge 0, \tag{B.30}$$

if and only if $f(x) = 0$ (a.e.).

Proof. (i) Choose a compact set $K \subset X$ and some $\varepsilon_0 > 0$ such that $K_{\varepsilon_0} := K + B_{\varepsilon_0}(0) \subseteq X$. Set $\tilde{f} := f \chi_{K_{\varepsilon_0}}$ and let ϕ be the standard mollifier. Then $(\phi_\varepsilon * \tilde{f})(x) = (\phi_\varepsilon * f)(x) \ge 0$ for $x \in K$, $\varepsilon < \varepsilon_0$ and since $\phi_\varepsilon * \tilde{f} \to \tilde{f}$ in $L^1(X)$ we have $(\phi_\varepsilon * \tilde{f})(x) \to f(x) \ge 0$ for a.e. $x \in K$ for an appropriate subsequence. Since $K \subset X$ is arbitrary the first claim follows. (ii) The first part shows that $\mathrm{Re}(f) \ge 0$ as well as $-\mathrm{Re}(f) \ge 0$ and hence $\mathrm{Re}(f) = 0$. Applying the same argument to $\mathrm{Im}(f)$ establishes the claim. □

The following variant is also often useful

Lemma B.17 (du Bois-Reymond[5]). *Suppose $X \subseteq \mathbb{R}^n$ is open and connected. If $f \in L^1_{loc}(X)$ with*

$$\int_X f(x) \partial_j \varphi(x) d^n x = 0, \qquad \forall \varphi \in C_c^\infty(X),\ 1 \le j \le n, \tag{B.31}$$

[5]Paul du Bois-Reymond (1831–1889), German mathematician

then f is constant a.e. on X.

Proof. Choose a ball $B_r \subset\subset U$ and some $\varepsilon_0 > 0$ such that $B_{r+\varepsilon_0} \subseteq U$. Set $\tilde{f} := f\chi_{B_{r+\varepsilon_0}}$ and let ϕ be the standard mollifier. Then by Lemma B.13 (ii)

$$\partial_j(\phi_\varepsilon * \tilde{f})(x) = ((\partial_j\phi_\varepsilon) * \tilde{f})(x) = ((\partial_j\phi_\varepsilon) * f)(x) = 0, \qquad x \in B_r,\ \varepsilon \leq \varepsilon_0.$$

Hence $(\phi_\varepsilon * \tilde{f})(x) = c_\varepsilon$ for $x \in B_r$ and as $\varepsilon \to 0$ there is a subsequence which converges a.e. on B_r. Clearly this limit function must also be constant: $(\phi_\varepsilon * \tilde{f})(x) = c_\varepsilon \to f(x) = c$ for a.e. $x \in B_r$. Now write U as a countable union of open balls whose closure is contained in U. If the corresponding constants for these balls were not all the same, we could find a partition into two union of open balls which were disjoint. This contradicts that U is connected. □

Problem* B.7. *Let $\|.\|$ be a seminorm on a vector space X. Show that $N := \{x \in X \,|\, \|x\| = 0\}$ is a vector space. Show that the quotient space X/N is a normed space with norm $\|x + N\| := \|x\|$.*

Problem* B.8. *Suppose U is bounded. Show that $L^\infty(U) \subseteq L^p(U)$ and*

$$\lim_{p\to\infty} \|f\|_p = \|f\|_\infty, \qquad f \in L^\infty(U).$$

Problem B.9. *Is it true that*

$$\bigcap_{1\leq p<\infty} L^p(0,1) = L^\infty(0,1)?$$

Problem B.10. *Construct a function $f \in L^p(0,1)$ which has a singularity at every rational number in $[0,1]$ (such that the essential supremum is infinite on every open subinterval). (Hint: Start with the function $f_0(x) = |x|^{-\alpha}$ which has a single singularity at 0, then $f_j(x) = f_0(x - x_j)$ has a singularity at x_j.)*

Problem B.11. *Show that for a continuous function on $\mathbb{R}^n$ the support and the essential support with respect to Lebesgue measure coincide.*

Problem* B.12. *Show the* ***generalized Hölder's inequality****:*

$$\|f\,g\|_r \leq \|f\|_p\|g\|_q, \qquad \frac{1}{p} + \frac{1}{q} = \frac{1}{r}. \tag{B.32}$$

Here we can allow $p, q, r \in (0,\infty]$ but of course $\|.\|_p$ will only be a norm for $p \geq 1$.

Problem* B.13. *Show the iterated Hölder's inequality:*

$$\|f_1 \cdots f_m\|_r \leq \prod_{j=1}^m \|f_j\|_{p_j}, \qquad \frac{1}{p_1} + \cdots + \frac{1}{p_m} = \frac{1}{r}. \tag{B.33}$$

Again with $p_j, r \in (0,\infty]$ as in the previous problem.

Problem* B.14. *Show that if $f \in L^{p_0} \cap L^{p_1}$ for some $p_0 < p_1$ then $f \in L^p$ for every $p \in [p_0, p_1]$ and we have the* ***Lyapunov inequality***[6]

$$\|f\|_p \le \|f\|_{p_0}^{1-\theta} \|f\|_{p_1}^{\theta},$$

where $\frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}$, $\theta \in (0,1)$. (Hint: Generalized Hölder inequality from Problem B.12.)

Problem* B.15. *Let $f \in L^p(U)$, $1 \le p < \infty$ and show that $T_a f \to f$ in L^p as $a \to 0$. (Hint: Start with $f \in C_c(U)$ and use Theorem B.10.)*

B.3. Closed operators

The current section collects some required results about closed operators from **[35]** to which we refer the reader for further details and proofs.

The **graph** of an operator $A : \mathfrak{D}(A) \subseteq X \to Y$ between Banach spaces is

$$\Gamma(A) := \{(x, Ax) | x \in \mathfrak{D}(A)\}. \tag{B.34}$$

If A is linear, the graph is a subspace of the Banach space $X \oplus Y$, which is just the Cartesian product together with the norm

$$\|(x,y)\|_{X \oplus Y} := \|x\|_X + \|y\|_Y. \tag{B.35}$$

Note that $(x_n, y_n) \to (x, y)$ if and only if $x_n \to x$ and $y_n \to y$. We call A has a closed graph if $\Gamma(A)$ is a closed subset of $X \oplus Y$.

We say that A has a closed graph if $\Gamma(A)$ is a closed subset of $X \oplus Y$.

Theorem B.18 (Closed graph)**.** *Let $A : X \to Y$ be a linear map from a Banach space X to another Banach space Y. Then A is continuous if and only if its graph is closed.*

Remark: The crucial condition here is that A is defined on *all* of X!

Operators whose graphs are closed are called **closed operators**. Warning: A closed operator will not map closed sets to closed sets in general. In particular, the concept of a closed operator should not be confused with the concept of a closed map in topology!

Being closed is the next option you have once an operator turns out to be unbounded. If A is closed, then $x_n \to x$ does not guarantee you that Ax_n converges (like continuity would), but it at least guarantees that if Ax_n converges, it converges to the right thing, namely Ax:

- A bounded (with $\mathfrak{D}(A) = X$): $x_n \to x$ implies $Ax_n \to Ax$.
- A closed (with $\mathfrak{D}(A) \subseteq X$): $x_n \to x$, $x_n \in \mathfrak{D}(A)$, and $Ax_n \to y$ implies $x \in \mathfrak{D}(A)$ and $y = Ax$.

[6] Aleksandr Lyapunov (1857–1918), Russian mathematician, mechanician and physicist

Please observe that the domain $\mathfrak{D}(A)$ is an intrinsic part of the definition of A and that we cannot assume $\mathfrak{D}(A) = X$ unless A is bounded (which is precisely the content of the closed graph theorem). Hence, if we want an unbounded operator to be closed, we have to live with domains. We will however typically assume that $\mathfrak{D}(A)$ is dense and set

$$\mathscr{C}(X,Y) := \{A : \mathfrak{D}(A) \subseteq X \to Y | A \text{ is densely defined and closed}\}. \tag{B.36}$$

One writes $B \subseteq A$ if $\mathfrak{D}(B) \subseteq \mathfrak{D}(A)$ and $Bx = Ax$ for $x \in \mathfrak{D}(B)$. In this case A is called an **extension** of B.

Example B.6. Two operators having the same prescription but different domains are different. For example

$$\mathfrak{D}(A) = C^1[0,1], \qquad Af = f'$$

and

$$\mathfrak{D}(B) = \{f \in C^1[0,1] | f(0) = f(1) = 0\}, \qquad Bf = f'$$

are two different operators in $X := C[0,1]$. Clearly A is an extension of B. Moreover, both are closed since $f_n \to f$ and $f_n' \to g$ implies that f is differentiable and $f' = g$. Note that A is densely defined while B is not. $\diamond$

Be aware that taking sums or products of unbounded operators is tricky due to the possible different domains. Indeed, if A and B are two operators between Banach spaces X and Y, so is $A + B$ defined on $\mathfrak{D}(A+B) := \mathfrak{D}(A) \cap \mathfrak{D}(B)$. The problem is that $\mathfrak{D}(A+B)$ might contain nothing more than zero. Similarly, if $A : \mathfrak{D}(A) \subseteq X \to Y$ and $B : \mathfrak{D}(B) \subseteq Y \to Z$, then the composition BA is defined on $\mathfrak{D}(BA) := \{x \in \mathfrak{D}(A) | Ax \in \mathfrak{D}(B)\}$.

Example B.7. Consider $X := C[0,1]$. Let M be the the subspace of trigonometric polynomials and N be the subspace of piecewise linear functions. Then both M and N are dense with $M \cap N = \{0\}$. $\diamond$

If an operator is not closed, you can try to take the closure of its graph, to obtain a closed operator. If A is bounded this always works. However, in general, the closure of the graph might not be the graph of an operator as we might pick up points $(x, y_1), (x, y_2) \in \overline{\Gamma(A)}$ with $y_1 \neq y_2$. Since $\overline{\Gamma(A)}$ is a subspace, we also have $(x, y_2) - (x, y_1) = (0, y_2 - y_1) \in \overline{\Gamma(A)}$ in this case and thus $\overline{\Gamma(A)}$ is the graph of some operator if and only if

$$\overline{\Gamma(A)} \cap \{(0,y) | y \in Y\} = \{(0,0)\}. \tag{B.37}$$

If this is the case, A is called **closable** and the operator $\overline{A}$ associated with $\overline{\Gamma(A)}$ is called the **closure** of A. Any linear subset $\mathfrak{D} \subseteq \mathfrak{D}(A)$ with the property that A restricted to $\mathfrak{D}$ has the same closure, $\overline{A|_{\mathfrak{D}}} = \overline{A}$, is called a **core** for A.

In particular, A is closable if and only if $x_n \to 0$ and $Ax_n \to y$ implies $y = 0$. In this case

$$\begin{aligned} \mathfrak{D}(\overline{A}) &= \{x \in X | \exists x_n \in \mathfrak{D}(A),\ y \in Y : x_n \to x \text{ and } Ax_n \to y\}, \\ \overline{A}x &= y. \end{aligned} \tag{B.38}$$

There is yet another way of defining the closure: Define the **graph norm** associated with A by

$$\|x\|_A := \|x\|_X + \|Ax\|_Y, \quad x \in \mathfrak{D}(A). \tag{B.39}$$

Since we have $\|Ax\| \le \|x\|_A$ we see that $A : \mathfrak{D}(A) \to Y$ is bounded with norm at most one. Thus far $(\mathfrak{D}(A), \|.\|_A)$ is a normed space and it suggests itself to consider its completion X_A. Then one can check that X_A can be regarded as a subset of X if and only if A is closable. In this case the completion can be identified with $\mathfrak{D}(\overline{A})$ and the closure of A in X coincides with the extension of A in X_A by virtue of continuity. In particular, A is closed if and only if $(\mathfrak{D}(A), \|.\|_A)$ is complete.

Example B.8 (Sobolev spaces)**.** Let $X := L^p(0,1)$, $1 \le p < \infty$, and consider $Af := f'$ on $\mathfrak{D}(A) := C^1[0,1]$. Then it is not hard to see that A is *not* closed (take a sequence g_n of continuous functions which converges in L^p to a non-continuous function and consider its primitive $f_n(x) = \int_0^x g_n(y)dy$). It is however closable. To see this suppose $f_n \to 0$ and $f_n' \to g$ in L^p. Then $f_n(0) = f_n(x) - \int_0^x f_n'(y)dy \to -\int_0^x g(y)dy$. But a sequence of constant functions can only have a constant function as a limit implying $g \equiv 0$ as required. The domain of the closure is the **Sobolev space** $W^{1,p}(0,1)$ and this is one way of defining Sobolev spaces. In particular, $W^{1,p}(0,1)$ is a Banach space when equipped with the graph norm. In this context one chooses the p-norm for the direct sum $X \oplus_p X$ such that the graph norm reads

$$\|f\|_{1,p} := \left(\|f\|_p^p + \|f'\|_p^p\right)^{1/p}.$$

◇

Example B.9. Another example are point evaluations in $L^p(0,1)$, $1 \le p < \infty$: Let $x_0 \in [0,1]$ and consider $\ell_{x_0} : \mathfrak{D}(\ell_{x_0}) \to \mathbb{C}$, $f \mapsto f(x_0)$ defined on $\mathfrak{D}(\ell_{x_0}) := C[0,1] \subseteq L^p(0,1)$. Then $f_n(x) := \max(0, 1 - n|x - x_0|)$ satisfies $f_n \to 0$ but $\ell_{x_0}(f_n) = 1$. In fact, a linear functional is closable if and only if it is bounded. ◇

For the closure of sums and products see Problem B.16 and Problem B.17, respectively.

Given a subset $\Gamma \subseteq X \oplus Y$ we can define

$$\Gamma^{-1} := \{(y,x) | (x,y) \in \Gamma\} \subseteq Y \oplus X. \tag{B.40}$$

In particular, applying this to the graph of an operator A, we will obtain the graph of its inverse (provided A is invertible). Hence we see that an invertible operator is closed if and only if its inverse is closed. Slightly more general, we have:

Lemma B.19. *Suppose A is closable and $\overline{A}$ is injective. Then $\overline{A}^{-1} = \overline{A^{-1}}$.*

Note that A injective does not imply $\overline{A}$ injective in general.

As a consequence of the closed graph theorem we obtain:

Corollary B.20. *Suppose $A \in \mathscr{C}(X,Y)$ is injective. Then A^{-1} defined on $\mathfrak{D}(A^{-1}) = \mathrm{Ran}(A)$ is closed. Moreover, in this case $\mathrm{Ran}(A)$ is closed if and only if A^{-1} is bounded.*

As in the case of bounded operators we define the **resolvent set** via

$$\rho(A) := \{\alpha \in \mathbb{C} | A - \alpha \text{ is bijective with a bounded inverse}\} \tag{B.41}$$

and call

$$R_A(\alpha) := (A-\alpha)^{-1}, \qquad \alpha \in \rho(A) \tag{B.42}$$

the **resolvent** of A. The complement $\sigma(A) = \mathbb{C} \setminus \rho(A)$ is called the **spectrum** of A. As in the case of Banach algebras it follows that the resolvent is analytic and that the resolvent set is open:

Lemma B.21. *Let A be a closed operator. Then the resolvent set is open and if $\alpha_0 \in \rho(A)$ we have*

$$R_A(\alpha) = \sum_{n=0}^{\infty} (\alpha - \alpha_0)^n R_A(\alpha_0)^{n+1}, \quad |\alpha - \alpha_0| < \|R_A(\alpha_0)\|^{-1}. \tag{B.43}$$

In particular, the resolvent is analytic and

$$\|(A-\alpha)^{-1}\| \ge \frac{1}{\mathrm{dist}(\alpha, \sigma(A))}. \tag{B.44}$$

It is also straightforward to verify the **first resolvent identity**

$$\begin{aligned} R_A(\alpha_0) - R_A(\alpha_1) &= (\alpha_0 - \alpha_1) R_A(\alpha_0) R_A(\alpha_1) \\ &= (\alpha_0 - \alpha_1) R_A(\alpha_1) R_A(\alpha_0), \end{aligned} \tag{B.45}$$

for $\alpha_0, \alpha_1 \in \rho(A)$.

However, note that for unbounded operators the spectrum will no longer be bounded in general and both $\sigma(A) = \emptyset$ as well as $\sigma(A) = \mathbb{C}$ are possible.

Example B.10. Consider $X := C[0,1]$ and $A = \frac{d}{dx}$ with $\mathfrak{D}(A) = C^1[0,1]$. We obtain the eigenvalues by solving the ordinary differential equation $x'(t) = \alpha x(t)$ which gives $x(t) = \mathrm{e}^{\alpha t}$. Hence every $\alpha \in \mathbb{C}$ is an eigenvalue, that is, $\sigma(A) = \mathbb{C}$.

Now let us modify the domain and look at $A_0 = \frac{d}{dx}$ with $\mathfrak{D}(A_0) = \{x \in C^1[0,1] | x(0) = 0\}$ and $X_0 := \{x \in C[0,1] | x(0) = 0\}$. Then the previous eigenfunctions do not satisfy the boundary condition $x(0) = 0$ and hence A_0 has no eigenvalues. Moreover, the solution of the inhomogeneous ordinary differential equation $x'(t) - \alpha x(t) = y(t)$ is given by $x(t) = x(0)\mathrm{e}^{\alpha t} + \int_0^t \mathrm{e}^{\alpha(t-s)} y(s) ds$. Hence $R_{A_0}(\alpha) y(t) = \int_0^t \mathrm{e}^{\alpha(t-s)} y(s) ds$ is the resolvent of A_0. Consequently $\sigma(A_0) = \emptyset$. ⋄

Note that if A is closed, then bijectivity implies boundedness of the inverse (see Corollary B.20). Moreover, by Lemma B.19 an operator with nonempty resolvent set must be closed.

Let us also note the following spectral mapping result.

Lemma B.22. *Suppose $A \in \mathscr{C}(X)$ is injective with $\mathrm{Ran}(A)$ is dense. Then*

$$\sigma(A^{-1}) \setminus \{0\} = (\sigma(A) \setminus \{0\})^{-1} \tag{B.46}$$

and

$$R_{A^{-1}}(\alpha^{-1}) = -\alpha A R_A(\alpha) = -\alpha - \alpha^2 R_A(\alpha), \qquad \alpha \in \rho(A) \setminus \{0\}. \tag{B.47}$$

In addition, for $\alpha \neq 0$ we have $\mathrm{Ker}((A-\alpha))^n = \mathrm{Ker}((A^{-1} - \alpha^{-1})^n)$ as well as $\mathrm{Ran}((A-\alpha))^n = \mathrm{Ran}((A^{-1} - \alpha^{-1})^n)$ for any $n \in \mathbb{N}$.

Concerning $\alpha = 0$ note that $0 \in \sigma(A^{-1})$ if and only if A is unbounded and vice versa.

In particular we can apply this lemma to the resolvent in case $\alpha_0 \in \rho(A)$ which shows

$$\sigma(A) = \alpha_0 + (\sigma(R_A(\alpha_0)) \setminus \{0\})^{-1} \tag{B.48}$$

and $\mathrm{Ker}(R_A(\alpha_0) - \alpha)^n = \mathrm{Ker}(A - \alpha_0 - \frac{1}{\alpha})^n$ as well as $\mathrm{Ran}(R_A(\alpha_0) - \alpha)^n = \mathrm{Ran}(A - \alpha_0 - \frac{1}{\alpha})^n$ for $\alpha \neq 0$ and $n \in \mathbb{N}$.

For example, this can be used to apply the Spectral theorem for compact operators (Theorem 7.7 from **[35]**) to unbounded operators in case they have a compact resolvent. To this end note that if we have $R_A(\alpha) \in \mathscr{K}(X)$ for one $\alpha \in \rho(A)$, then this holds in fact for all $\alpha \in \rho(A)$ by the first resolvent identity (B.45) since compact operators form an ideal.

Theorem B.23. *Suppose $R_A(\alpha) \in \mathscr{K}(X)$ for one $\alpha \in \rho(A)$. Then the spectrum of A consists only of discrete eigenvalues with finite (geometric and algebraic) multiplicity.*

Problem* B.16. *Show that if A, B, and $A+B$ are closable, then $\overline{A} + \overline{B} \subseteq \overline{A+B}$ with equality if A or B is bounded. Moreover, if B is bounded, then $A+B$ is closable if and only if A is.*

Problem* B.17. *Let $A : \mathfrak{D}(A) \subseteq X \to Y$ and $B : \mathfrak{D}(B) \subseteq Y \to Z$ be closable. If $A \in \mathscr{L}(X, Y)$, then BA is closable and $\overline{BA} = \overline{B}A$. Similarly, if $B^{-1} \in \mathscr{L}(Z, Y)$, then BA is closable and $\overline{BA} = B\overline{A}$.*

Problem* B.18. *Suppose A is closed and B satisfies $\mathfrak{D}(A) \subseteq \mathfrak{D}(B)$:*

- *Show that $1 + B$ has a bounded inverse if $\|B\| < 1$.*
- *Suppose A has a bounded inverse. Then so does $A+B$ if $\|BA^{-1}\| < 1$. In this case we have $\|(A+B)^{-1}\| \le \frac{\|A^{-1}\|}{1-\|BA^{-1}\|}$.*

Problem B.19. *Let A be a closed operator. Show that for every $\alpha \in \rho(A)$ the expression $\|f\|_\alpha := \|(A-\alpha)x\|$ defines a norm which is equivalent to the graph norm.*

Problem* B.20. *Let A be a closed operator. Show* (B.45). *Moreover, conclude*

$$\frac{d^n}{d\alpha^n} R_A(\alpha) = n! R_A(\alpha)^{n+1}, \qquad \frac{d}{d\alpha} R_A(\alpha)^n = n R_A(\alpha)^{n+1}.$$

B.4. Weak convergence

The current section collects some required results about weak convergence from **[35]** to which we refer the reader for further details and proofs.

Let X be a Banach space and X^*, X^{**} its dual, bidual space, respectively. Recall that point evaluation defined as $J(x)(\ell) := \ell(x)$ give an isometric map $J : X \to X^{**}$. Then X is called **reflexive** is this map is surjective.

If $\ell(x_n) \to \ell(x)$ for every $\ell \in X^*$ we say that x_n **converges weakly** to x and write

$$\operatorname*{w-lim}_{n\to\infty} x_n = x \quad \text{or} \quad x_n \rightharpoonup x. \tag{B.49}$$

Clearly, $x_n \to x$ implies $x_n \rightharpoonup x$ and hence this notion of convergence is indeed weaker. Moreover, the weak limit is unique, since $\ell(x_n) \to \ell(x)$ and $\ell(x_n) \to \ell(\tilde{x})$ imply $\ell(x - \tilde{x}) = 0$. A sequence x_n is called a **weak Cauchy sequence** if $\ell(x_n)$ is Cauchy (i.e. converges) for every $\ell \in X^*$.

Lemma B.24. *Let X be a Banach space.*

(i) *$x_n \rightharpoonup x$, $y_n \rightharpoonup y$ and $\alpha_n \to \alpha$ implies $x_n + y_n \rightharpoonup x + y$ and $\alpha_n x_n \rightharpoonup \alpha x$.*

(ii) *$x_n \rightharpoonup x$ implies $\|x\| \le \liminf \|x_n\|$.*

(iii) *Every weak Cauchy sequence x_n is bounded: $\|x_n\| \le C$.*

(iv) *If X is reflexive, then every weak Cauchy sequence converges weakly.*

(v) *A sequence x_n is Cauchy if and only if $\ell(x_n)$ is Cauchy, uniformly for $\ell \in X^*$ with $\|\ell\| = 1$.*

One of the most important results is analysis is the Heine–Borel theorem which states that closed bounded sets in $\mathbb{R}^n$ are compact. In the case of Banach spaces, this remains true if and only if the Banach space is finite dimensional (cf. Theorem 4.31 from [**35**]). In the infinite dimensional case one has to replace norm convergence by weak convergence:

Theorem B.25 (Šmulian[7])**.** *Let X be a reflexive Banach space. Then every bounded sequence has a weakly convergent subsequence.*

It is also useful to observe that compact operators will turn weakly convergent into (norm) convergent sequences.

Theorem B.26. *Let $A \in \mathscr{K}(X,Y)$ be compact. Then $x_n \rightharpoonup x$ implies $Ax_n \to Ax$. If X is reflexive the converse is also true.*

Problem* B.21. *Let X be a normed space. Show that the following conditions are equivalent.*

(i) *If $\|x+y\| = \|x\| + \|y\|$ then $y = \alpha x$ for some $\alpha \ge 0$ or $x = 0$.*
(ii) *If $\|x\| = \|y\| = 1$ and $x \neq y$ then $\|\lambda x + (1-\lambda)y\| < 1$ for all $0 < \lambda < 1$.*
(iii) *If $\|x\| = \|y\| = 1$ and $x \neq y$ then $\frac{1}{2}\|x+y\| < 1$.*
(iv) *The function $x \mapsto \|x\|^2$ is strictly convex.*

A norm satisfying one of them is called ***strictly convex****.*

B.5. The Bochner integral

The current section collects some required results about the Bochner integral from [**34**] to which we refer the reader for further details and proofs.

In this section we want to recall how to extend the Lebesgue integral to the case of functions $f : U \subseteq \mathbb{R}^n \to Y$ with values in a normed space Y. This extension is known as **Bochner integral**.[8] Since a normed space has no order we cannot use monotonicity and hence are restricted to finite values for the integral. Other than that, we only need some small adaptions.

The idea simple: Equip Y with the Borel σ-algebra. A measurable function $s : U \to Y$ is call simple if it takes only finitely many values

$$s = \sum_{j=1}^{p} \alpha_j \, \chi_{A_j}, \qquad \operatorname{Ran}(s) =: \{\alpha_j\}_{j=1}^{p}, \quad A_j := s^{-1}(\alpha_j). \tag{B.50}$$

[7]Vitold Shmulyan (1914–1944), Soviet mathematician
[8]Salomon Bochner (1899–1982), Austrian mathematician

We call s integrable if the Lebesgue measure $|A_j|$ is finite unless $\alpha_j = 0$. In this case the integral can be defined via

$$\int s(x) d^n x := \sum_{j=1}^{p} \alpha_j |A_j|. \tag{B.51}$$

Then a measurable function $f : U \to Y$ is call **integrable** if there is a sequence of integrable functions s_n such that

$$\lim_{n\to\infty} \int_U \|f(x) - s_n(x)\| d^n x = 0. \tag{B.52}$$

In this case we can define the **Bochner integral** of f as

$$\int_u f(x) d^n x := \lim_{n\to\infty} \int_U s_n(x) d^n x = 0. \tag{B.53}$$

Functions $f : X \to Y$ which are the pointwise limit of simple functions are also called **strongly measurable** and we have:

Lemma B.27 (Bochner). *A function $f : X \to Y$ is integrable if and only if it is strongly measurable and $\|f\|$ is integrable.*

The Lebesgue spaces $L^p(U, Y)$ can be defined as usual and of course they are complete:

Theorem B.28 (Riesz–Fischer). *The space $L^p(U, Y)$, $1 \le p \le \infty$, is a Banach space.*

Theorem B.29 (Minkowski's integral inequality). *Suppose, $U \subseteq \mathbb{R}^m$ and $V \subseteq \mathbb{R}^n$ open, Y a Banach space, and $f : U \times V \to Y$ is strongly $\mu \otimes \nu$ measurable. Let $1 \le p \le \infty$. Then*

$$\left\| \int_V f(.,y) d^n y \right\|_p \le \int_V \|f(.,y)\|_p d^n y, \tag{B.54}$$

where the p-norm is computed with respect to $L^p(U, Y)$. In particular, this says that $f(x, .)$ is integrable for a.e $x \in U$ and $\int_V f(., y) d^n y \in L^p(U, Y)$ if the integral on the right is finite.

Problem B.22. *Let $U \subseteq \mathbb{R}^n$ be a domain and $I \subseteq \mathbb{R}$ an interval. Show that for a strongly measurable function f and $1 \le p, r < \infty$ we have*

$$\|f\|_{L^r(I, L^p(U))} = \sup_{\|g\|_{L^{r'}(I, L^{p'}(U))} = 1} \int_I \int_U |f(t,x)||g(t,x)| d^n x \, dt,$$

where p', r' are the corresponding dual indices. Moreover, it suffices to take the sup over functions which have support in a compact rectangle.

Bibliography

[1] R.A. Adams and J.J.F. Fournier, *Sobolev Spaces*, 2nd ed., Academic Press, Amsterdam, 2003.

[2] W. Arendt and K. Urban, *Partielle Differenzialgleichungen*, Spektrum, Heidelberg, 2010.

[3] S. Axler, P. Bourdon, and W. Ramey, *Harmonic Function Theory*, 2nd ed., Springer, New York, 2001.

[4] H. Brezis, *Functional Analysis, Sobolev Spaces and Partial Differential Equations*, Springer, New York, 2011.

[5] H. Brézis and F. Browder, *Partial differential equations in the 20th century*, Adv. Math. **135:1**, 76–144 (1998).

[6] J.R. Cannon, *The One-Dimensional Heat Equations*, Addison-Wesley, Menlo Park, 1984.

[7] R. Courant and D. Hilbert, *Methods of Mathematical Physics*, Wiley-VCH, Berlin, 1964.

[8] Y.V. Egorov, *Linear Differential Equations of Principal Type*, Contemporary Soviet Mathematics, Consultants Bureau, New York, 1986.

[9] K.-J. Engel and R. Nagel, *One-Parameter Semigroups for Linear Evolution Equations*, Springer, Berlin, 2000.

[10] L.C. Evans, *Partial Differential Equations*, 2nd ed., AMS, Providence, 2010.

[11] G.B. Folland, *Introduction to Partial Differential Equations*, 2nd ed., Princeton UP, Princeton, 1995.

[12] D. Gilbarg and N.S. Trudinger, *Elliptic Partial Differential Equations of Second Order*, 2nd ed., Springer, Berlin, 2001.

[13] J. Goldstein, *Semigroups of Linear Operators and Appications*, Oxford University Press, New York, 1985.

[14] D. Hundertmark, M. Meyries, L. Machinek, and R. Schnaubelt, *Operator Semigroups and Dispersive Equations*, Lecture Notes (16th Internet Seminar on Evolution Equations), 2013. `https://isem.math.kit.edu/images/b/b3/Isem16_final.pdf`

[15] F. John, *Partial Differential Equations*, 4th ed., Springer, New York, 1982.

[16] J. Jost, *Partial Differential Equations*, Springer, New York, 2002.

[17] Y. Katznelson, *Introduction to Harmonic Analysis*, 3rd ed., Cambridge UP, Cambridge, 2004.

[18] S.G. Krantz and H.R. Parks, *A Primer of Real Analytic Functions*, 2nd ed., Birkäuser, Boston, 2002.

[19] E.M. Landis, *Second Order Equations of Elliptic and Parabolic Type*, AMS, Providence, 1997.

[20] G. Leoni, *A First Course in Sobolev Spaces*, Amer. Math. Soc., Providence, 2009.

[21] E. Lieb and M. Loss, *Analysis*, 2nd ed., Amer. Math. Soc., Providence, 2000.

[22] F. Linares and G. Ponce, *Introduction to Nonlinear Dispersive Equations*, 2nd ed., Springer, New York, 2015.

[23] J.D. Murray, *Asymptotic Anlysis*, Springer, New York, 1984.

[24] Y. Pinchover and J. Rubinstein, *An Introduction to Partial Differential Equations*, Cambridge University Press, Cambridge, 2005.

[25] F.W.J. Olver et al., *NIST Handbook of Mathematical Functions*, Cambridge University Press, Cambridge, 2010.

[26] P.J. Olver, *Applications of Lie Groups to Differential Equations*, 2nd ed., Springer, New York, 1993.

[27] M. Renardy and R.C. Rogers, *An Introduction to Partial Differential Equations*, 2nd ed., Springer, New York, 2004.

[28] E. Stein, *Singular Integrals and Differentiability Properties of Functions*, Princeton UP, Princeton, 1970.

[29] E. Stein and R. Shakarchi, *Complex Analysis*, Princeton UP, Princeton, 2003.

[30] W. Strauss, *Partial Differential Equations. An Introduction*, 2nd ed., Wiley, Hoboken, 2008.

[31] T. Tao, *Nonlinear Dispersive Equations: Local and Global Analysis*, Amer. Math. Soc., Providence, 2006.

[32] G. Teschl, *Mathematical Methods in Quantum Mechanics; With Applications to Schrödinger Operators*, Amer. Math. Soc., Providence, 2009.

[33] G. Teschl, *Ordinary Differential Equations and Dynamical Systems*, Amer. Math. Soc., Providence, 2012.

[34] G. Teschl, *Topics in Real Analysis*, Lecture Notes.

[35] G. Teschl, *Topics in Functional Analysis*, Lecture Notes.

[36] E. Wienerholz, H. Kalf, and T. Kriecherbauer, *Elliptische Differentialgleichungen zweiter Ordnung*, Springer, Berlin, 2009.

Glossary of notation

$\arg(z)$	... argument of $z \in \mathbb{C}$; $\arg(z) \in (-\pi, \pi]$, $\arg(0) = 0$
$B_r(x)$	... open ball of radius r around x
$\mathbb{C}$	... the set of complex numbers
$C(U)$	... set of continuous functions from U to $\mathbb{C}$
$C_0(U)$	... set of continuous functions vanishing on the boundary ∂U
$C^k(U)$	... set of k times continuously differentiable functions
$C_c^\infty(U)$	... set of compactly supported smooth functions
$C(U, Y)$	... set of continuous functions from U to Y
$C^r(U, Y)$	... set of r times continuously differentiable functions
$C^{0,\gamma}(U)$	... Hölder continuous functions of exponent γ
$C^{1;2}(U_T)$	... functions with different degree of differentiability, 54
$\chi_A(.)$	... characteristic function of the set A
$\delta_{n,m}$	... Kronecker delta
det	... determinant
dim	... dimension of a linear space
div	... divergence of a vector filed
$\operatorname{diam}(U)$	$= \sup_{(x,y)\in U^2} d(x,y)$ diameter of a set
$\operatorname{dist}(U, V)$	$= \inf_{(x,y)\in U\times V} d(x,y)$ distance of two sets
e	... Napier's constant, $\mathrm{e}^z = \exp(z)$
$\mathrm{GL}(n)$	... general linear group in n dimensions
$\Gamma(z)$	... gamma function
$\Gamma(f_1, \dots, f_n)$	... Gram determinant,
$\mathfrak{H}$	... a Hilbert space

i	. . . complex unity, $\mathrm{i}^2 = -1$
Im(.)	. . . imaginary part of a complex number
inf	. . . infimum
$J_f(x)$	$= \det df(x)$ Jacobi determinant of f at x
max	. . . maximum
$\mathbb{N}$	. . . the set of positive integers
$\mathbb{N}_0$	$= \mathbb{N} \cup \{0\}$
∇f	$= (\partial_1 f, \dots, \partial_m f)$ gradient in $\mathbb{R}^n$
$\nu(x_0)$	. . . outward pointing unit normal vector, 340
$O(.)$	. . . Landau symbol, $f = O(g)$ iff $\limsup_{x\to x_0} \lvert f(x)/g(x)\rvert < \infty$
$o(.)$	. . . Landau symbol, $f = o(g)$ iff $\lim_{x\to x_0} \lvert f(x)/g(x)\rvert = 0$
$\mathbb{Q}$	. . . the set of rational numbers
$\mathbb{R}$	. . . the set of real numbers
Re(.)	. . . real part of a complex number
S^{n-1}	$= \{x \in \mathbb{R}^n \mid \lvert x\rvert = 1\}$ unit sphere in $\mathbb{R}^n$
sign(z)	$= z/\lvert z\rvert$ for $z \neq 0$ and 1 for $z = 0$; complex sign function
sup	. . . supremum
supp(f)	. . . support of a function f, 329
span(M)	. . . set of finite linear combinations from M
$\mathbb{Z}$	. . . the set of integers
$\mathbb{I}$	. . . identity operator
$\sqrt{z}$	. . . square root of z with branch cut along $(-\infty, 0)$
z^*	. . . complex conjugation
$\hat{f}$	$= \mathcal{F}f$, Fourier coefficients/transform of f
$\check{f}$	$= \mathcal{F}^{-1}f$, inverse Fourier transform of f
$\lvert x\rvert$	$= \sqrt{\sum_{j=1}^n \lvert x_j\rvert^2}$ Euclidean norm in $\mathbb{R}^n$ or $\mathbb{C}^n$
$\lvert A\rvert$	. . . Lebesgue measure of a Borel set A
$\lVert.\rVert$	. . . norm
$\lVert.\rVert_p$	. . . norm in the Banach space L^p
$\langle ., ..\rangle$	. . . scalar product in $\mathfrak{H}$
$\oplus$	. . . direct/orthogonal sum of vector spaces or operators
$\dot\cup$	. . . union of disjoint sets
$\lfloor x\rfloor$	$= \max\{n \in \mathbb{Z} \mid n \le x\}$, floor function
$\lceil x\rceil$	$= \min\{n \in \mathbb{Z} \mid n \ge x\}$, ceiling function
∂_α	. . . partial derivative in multi-index notation
$\partial_x F(x,y)$	. . . partial derivative with respect to x
∂U	$= \overline{U} \setminus U^\circ$ boundary of the set U
$\overline{U}$	. . . closure of the set U
U°	. . . interior of the set U
$V \subset\subset U$	. . . V is relatively compact with $\overline{V} \subset U$
$M^\perp$	. . . orthogonal complement
(λ_1, λ_2)	$= \{\lambda \in \mathbb{R} \mid \lambda_1 < \lambda < \lambda_2\}$, open interval
$[\lambda_1, \lambda_2]$	$= \{\lambda \in \mathbb{R} \mid \lambda_1 \le \lambda \le \lambda_2\}$, closed interval
$x_n \to x$	. . . norm convergence

Index