Delio Mugnolo, Joachim Kerner

Modul 61218

Partielle Differentialgleichungen

LESEPROBE

Fakultät für
Mathematik und Informatik

FernUniversität in Hagen

Liebe Studierende,

herzlich willkommen zum Kurs Partielle Differenzialgleichungen, bei dessen Bearbeitung wir Ihnen viel Spaß und Erfolg wünschen! Wie der Titel des Kurses schon verrät, befassen wir uns mit sogenannten **partiellen Differenzialgleichungen** (dem Englischen folgend schreiben wir auch kurz PDEs, für *partial differential equations*). Aber was sind das für Gleichungen und wozu sind sie gut? Tatsächlich handelt es sich um Gleichungen, deren Unbekannte Funktionen (etwa u) von zwei oder mehr Variablen (etwa $x_0, x_1, \dots, x_N$) sind und in welcher partielle Ableitungen (erster oder höherer Ordnung) dieser gesuchten Funktionen auftauchen. Differenzialgleichungen, welche nur eine Funktion einer Variablen beinhalten sind Ihnen als *gewöhnliche Differenzialgleichungen* (Englisch: *ordinary differential equation*, ODEs) bekannt und werden in diesem Kurs nicht gesondert behandelt, obwohl wir gelegentlich die Analyse partieller Differenzialgleichungen auf Eigenschaften gewöhnlicher Differenzialgleichungen zurückführen. Im Allgemeinen hat eine partielle Differenzialgleichung der m-ten Ordnung die Form

$$F\left(x_0, x_1, \dots, x_N, u, \frac{\partial u}{\partial x_0}, \frac{\partial u}{\partial u}\partial x_1, \dots, \frac{\partial u}{\partial x_N}, \dots, \frac{\partial^m u}{\partial x_0^m}, \frac{\partial^m u}{\partial x_1^m}, \dots, \frac{\partial^m u}{\partial x_N^m}\right) = 0\,,$$

wobei F eine gegebene Funktion ist; diese Funktion muss jedoch nicht explizit von all ihren Argumenten abhängen. Eine solche partielle Differenzialgleichung wird **linear** genannt, wenn sie für geeignete Funktionen $f_0, f_{11}, \dots, f_{1N}, \dots, f_{m1\dots1}, \dots, f_{mN\dots N}$ in der Form

$$f_0(x_0, x_1, \dots, x_N)u + \sum_{k=1}^{N} f_{1k}(x_0, x_1, \dots, x_N)\frac{\partial u}{\partial x_k} + \sum_{j,k=1}^{N} f_{2jk}(x_0, x_1, \dots, x_N)\frac{\partial^2 u}{\partial x_k \partial x_j} + \dots = 0$$

geschrieben werden kann. Wenn es eine Variable gibt – ohne Beschränkung der Allgemeinheit sei es x_0 – die als Zeitvariable interpretiert werden kann, dann schreiben wir oft t für x_0 und nennen eine solche PDE eine **Evolutionsgleichung**, insbesondere, wenn es keine gemischten Terme gibt, die Ableitungen bezüglich t *und* $x_1, \dots, x_N$ beinhalten. Die meisten der Gleichungen, denen wir in diesem Kurs begegnen werden, sind Evolutionsgleichungen.

Typische PDEs sind zum Beispiel die *Transportgleichung*

$$\frac{\partial u}{\partial t}(t,x) = \frac{\partial u}{\partial x}(t,x)\,,$$

die *Wellengleichung*

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

und die *Korteweg–de Vries-Gleichung*

$$\frac{\partial u}{\partial t}(t,x) + \frac{\partial^3 u}{\partial x^3}(t,x) = u(t,x)\frac{\partial u}{\partial x}(t,x). \tag{0.1}$$

In der Tat handelt es sich bei diesen PDEs um Evolutionsgleichungen von erster, zweiter bzw. dritter Ordnung; genauer gesagt ist die Transportgleichung von erster Ordnung bezüglich der Zeit- und Ortsvariablen t, x; die Wellengleichung ist von zweiter Ordnung bezüglich beider Variablen t, x; die Korteweg–de Vries-Gleichung ist von erster Ordnung bezüglich der Zeit t und von dritter Ordnung bezüglich des Ortes x. Die Transport- und Wellengleichungen sind linear, während die Korteweg–de-Vries-Gleichung eine nichtlineare PDE ist. Die Theorie der nichtlinearen Differenzialgleichungen ist in vielen Hinsichten viel komplizierter: wir werden sie in diesem Kurs nur kurz anschneiden.

Die Wichtigkeit vieler partiellen Differentialgleichungen liegt darin, dass sie einen physikalischen Prozess beschreiben, wie etwa Wärmediffusion in einem Festkörper, Schwingungen einer Brücke, Strömungen von Flüssigkeiten, Verteilung elektrischer Ladung auf der Oberfläche einer Metallkugel usw. Im letzten Jahrhundert hat sich jedoch die Modellierung mittels partieller Differenzialgleichungen auch in nichtphysikalischen Disziplinen bewährt, wie im Fall der *Hodgkin–Huxley-Gleichung*

$$\frac{\partial u}{\partial t}(t,x) = \frac{\partial^2 u}{\partial x^2}(t,x) - \frac{1}{2}u(t,x)\Big(u(t,x)-1\Big)\Big(u(t,x)-\alpha\Big)\,,$$

und der *Black–Scholes-Gleichung*

$$\frac{\partial u}{\partial t}(t,x) + \frac{1}{2}\sigma^2 x^2 \frac{\partial^2 u}{\partial x^2}(t,x) + rx\frac{\partial u}{\partial x}(t,x) - ru(t,x) = 0\,,$$

wobei die Parameter α, σ, r entsprechend der Eigenschaften des jeweils betrachteten Modells bestimmt werden müssen. So modelliert die Hodgkin–Huxley-Gleichung die Feuermuster eines Neurons, während die Black–Scholes den Preis einer europäischen Aktienoption beschreibt. In beiden Fällen wurde die zugrundeliegende Arbeit mit einem Nobelpreis gewürdigt: 1961 für Medizin bzw. 1997 für Wirtschaftswissenschaften. Die Bedeutung dieser Gleichungen – und ähnlicher wichtiger Gleichungen in der Physik, Biologie und in anderen angewandten Bereichen – liegt darin, dass Forschende in der Lage waren, die zentralen Mechanismen real existierender Systeme in mathematische Sprache zu übersetzen; vertraut man auf die Genauigkeit dieser mathematischen Übertragung, dann kann man daher erwarten, unbekannte Eigenschaften komplexer Systeme auf mathematische Weise vorherzusagen.

Die Modellierung eines gut verstandenen Prozesses durch ein mathematisches Gesetz ist oft eine anspruchsvolle Aufgabe, die als **Herleitung** einer Gleichung bezeichnet wird. Genau wie im Fall gewöhnlicher Differenzialgleichungen ist eine solche Herleitung einer partiellen Differenzialgleichung mit vielen Vereinfachungen und Annahmen verbunden, deren Plausibilität eingeschätzt werden muss. In einigen Fällen gibt einem die resultierende PDE nur wenig Einblick in das zugrundeliegende Problem, kann aber dennoch auf rein mathematischer Ebene hochinteressant sein. In anderen Fällen geschieht jedoch etwas Magisches: manchmal beschreiben partielle Differenzialgleichung ein reales System so genau, dass durch ihre mathematische Analyse unbekannte Phänomene vorhergesagt werden können. So ist es zum Beispiel immer wieder mit Einsteins Feldgleichungen geschehen, die den Kern der Allgemeinen Relativitätstheorie darstellen: mehrmals sind kosmologische Phänomene zuerst mathematisch vorhergesagt und dann in Experimenten beobachtet worden. Der jüngste derartige Fall ist die Bestätigung der Existenz von den lange vorhergesagten, berüchtigten Gravitationswellen, die schließlich am 14. September 2015 aufgenommen wurden – ein Durchbruch, der 2017 zur Verleihung des Nobelpreises für Physik an Rainer Weiss, Kip Thorne und Barry Barish führte.

Darüber hinaus gibt es technische Prozesse – etwa in der Bildverarbeitung oder im maschinellen Lernen – bei denen man sich die guten Eigenschaften von Lösungen bekannter Gleichungen zu Nutze macht, um ein erwünschtes Verhalten zu erzwingen: etwa um Bilder zu glätten oder Cluster in großen Datenmengen zu erkennen. Eine Gleichung wird somit in diesen Anwendungskontexten nicht hergeleitet, vielmehr eine technische Struktur oder ein Algorithmus so konzipiert, dass ein Verhalten einer Differenzialgleichung nachgeahmt wird.

Nicht immer handelt es sich dabei um partielle Differerenzialgleichungen, aber durchaus in vielen Fällen, wie auch aus den historischen Anmerkungen in dem Kurs hervorgehen wird. Sobald zu einem Phänomen eine partielle Differenzialgleichung hergeleitet wurde, ist es anschließend die Aufgabe, diese Differenzialgleichung zu analysieren. Zunächst stellt sich die grundsätzliche Frage, ob überhaupt die Existenz einer Lösung zu erwarten ist; und wenn ja, ob sie durch geeignete zusätzliche Informationen (Anfangs- und/oder Randbedingungen) eindeutig bestimmt ist. Diese Aufgabe ist in vielen Fällen durchaus nicht trivial.

Einige wichtige Differenzialgleichungen lassen sich sogar mehr oder weniger explizit lösen, meistens in Form von Integralen oder Reihen. Dies gilt aber fast ausschließlich für Gleichungen, deren räumliche Variablen im ganzen Euklidischen Raum definiert sind, oder zumindest auf Gebieten, die starke Symmetrieeigenschaften aufweisen (Halbebenen, Kreisscheiben usw.). In den meisten Fällen ist aber die Suche nach einer expliziten Lösungsformel aussichtslos und die Wege trennen sich: Zum einen kann man etwa versuchen, mit numerischen Methoden eine approximative Lösung zu finden, oder man versucht, *qualitative* Eigenschaften der (einzigen) Lösung zu beweisen: diesem zweiten Weg werden wir in diesem Kurs folgen. Handelt es sich um eine glatte Funktion und hängt sie stetig von den Anfangs- und/oder Randbedingungen ab? Wie verhält sie sich im Falle, dass man ihre Variablen gegen Unendlich schickt? Zu diesen und anderen Zwecken werden wir etliche verschiedene Techniken kennenlernen, die jeweils für eine spezielle Klasse von partiellen Differenzialgleichungen erfolgreich angewendet werden können. Diese Methoden werden wir in der Regel am Beispiel einer oder mehrerer PDEs erläutern.

Der vorliegende Kurs besteht aus sieben Kapiteln (Kurseinheiten), welche Sie am besten nacheinander bearbeiten sollten. Folgende Themengebiete werden dabei behandelt:

- Hyperbolische Gleichungen und die Methode der Charakteristiken

- Die Wärmeleitungsgleichung und Symmetriemethoden
- Die Poisson-Gleichung und Hilbertraum-Methoden
- Unbeschränkte Operatoren und der Spektralsatz
- Halbgruppentheorie und abstrakte Cauchy-Anfangswertprobleme
- Die Telegraphen-Gleichung und der Satz von Noether
- Nicht-lineare Gleichungen und Kompaktheitsmethoden

Die jeweiligen Kurseinheiten bestehen aus formalen mathematischen Aussagen verbunden mit historischen und den Zusammenhang herstellenden Informationen. Auch finden Sie zahlreiche Übungsaufgaben, deren Bearbeitung wir Ihnen sehr ans Herz legen wollen, weil es Ihnen ein tieferes Verständnis des Stoffes ermöglicht. Am Ende der jeweiligen Kurseinheiten finden Sie zudem *Anmerkungen und Empfehlungen* für ein weiterführendes Studium des Stoffes, welches über diesen Kurs hinausgeht. Diese Anmerkungen erlauben es Ihnen, Verbindungen zu anderen Gebieten klarer sehen und den erlernten Stoff besser einordnen zu können. Auch finden Sie darin viele Anregungen für ein Seminar oder eine Abschlussarbeit. Anschließend haben wir einen Appendix beigefügt, der als Repetitorium etlicher nützlicher Begriffe und Resultate – vor allem aus den Kursen [29, 12] – dient.

Da es sich bei dem Gebiet der partiellen Differenzialgleichungen um ein lange gewachsenes Gebiet der Mathematik handelt, gibt es viele verschiedene Zugänge dazu und entsprechend können wir nicht erwarten, alle Facetten innerhalb eines Kurses abzudecken. Deshalb möchten wir Ihnen ans Herz legen, sich auch mit anderen Quellen zum Thema auseinanderzusetzen; eine Auswahl an möglichen Quellen finden Sie am Ende der Kurseinheit aufgeführt. Wie Sie wahrscheinlich schon gehört haben, gibt es derzeit auch ein berühmtes ungelöstes Problem aus dem Gebiet der PDEs – ein sogenanntes Milleniumproblem. Dabei geht es um die Existenz einer sogenannten starken Lösung der (inkompressiblen) Navier–Stokes-Gleichungen, welche die Dynamik einer Flüssigkeit beschreiben. Da wir uns im Haupttext nicht direkt damit befassen, haben wir Ihnen zur Motivation und als Einstieg in das Thema ein kleines Kapitel im Anhang mit weiterführenden Referenzen eingefügt. In jedem Falle sehen Sie daran, dass Sie sich mit einem hochaktuellen Gebiet der Mathematik befassen werden, welches auf vielfältigste Art und Weise für andere Bereiche von großem Nutzen ist.

Abschließend sei noch erwähnt, dass der vorliegende Kurs in englischer Sprache formuliert ist. Das mag für Sie als Studierende an einer deutschen Hochschule zunächst einmal ungewohnt sein, aber birgt auch große Vorteile. Zum einen vertiefen Sie neben der Mathematik auch noch etwas die englische Sprache, viel wichtiger jedoch, Sie rüsten sich für weiterführende Studien. Denn, wie Sie sicher wissen, werden aktuelle Forschungsarbeiten praktisch ausschließlich in Englisch verfasst. Wenn Sie also in naher Zukunft ein Seminar vorbereiten oder eine Abschlussarbeit verfassen, dann werden Sie sich mit wissenschaftlichen Arbeiten in englischer Sprache auseinandersetzen müssen.

Wir haben Ihnen jedoch, zu Beginn einer jeweiligen Kurseinheit, eine Zusammenfassung der wichtigsten Lernziele auf Deutsch bereitgestellt; damit soll Ihnen die Bearbeitung erleichtert werden. Auch steht Ihnen ein umfassender Index zur Verfügung, in dem wir allen (englischen) Einträgen im Index eine deutsche Übersetzung beigefügt haben.

Selbstverständlich werden mündliche Prüfungen zu diesem Kurs in deutscher Sprache abgenommen. Auch die Korrektur der Einsendeaufgaben erfolgt in der Regel auf Deutsch.

Wie wünschen Ihnen eine interessante Lernphase!

PROF. DR. DELIO MUGNOLO

PD DR. JOACHIM KERNER

Dieses Skript wird im Sommersemester 2021 zum ersten Mal eingesetzt. Für ihre wichtigen Anmerkungen und Korrekturen bedanken wir uns ganz herzlich bei Herrn Marvin Plümer und Herrn Dr. Matthias Täufer.

Dennoch ist es sehr wahrscheinlich, dass sich etliche kleine und große mathematische und sprachliche Fehler im Text versteckt haben. Wir wären Ihnen, liebe Studierende, für jeden Hinweis auf Ungenauigkeiten sehr dankbar.

Inhalt

Studierhinweise zu Kurseinheit 1

In dieser Kurseinheit befassen wir uns mit hyperbolischen partiellen Differentialgleichungen. Insbesondere untersuchen wir zwei wichtige Repräsentanten dieser Klasse: die Transportgleichung und die Wellengleichung. Wir werden sehen, dass beide Gleichungen physikalisch motiviert sind und auf Grundlage physikalischer Gesetzmäßigkeiten hergeleitet werden können. Im Zuge der mathematischen Beschreibung beider Gleichungen werden wir eine wichtige Methode, die sogenannte Methode der Charakteristiken, kennenlernen. Diese Methode wird es uns insbesondere erlauben, explizite Formeln für die Lösungen dieser Gleichungen herleiten zu können. Anhand dieser Formeln werden wir dann wichtige Eigenschaften von Lösungen hyperbolischer PDEs untersuchen können. Beispielsweise werden wir sehen, dass die Lösungen sich mit einer *endlichen* Geschwindigkeit ausbreiten.

Lernziele:

Abschnitt: Die Transportgleichung in einer Dimension (The transport equation in one dimension)

Nach dem Lesen des Abschnitts sollten Sie die physikalische Herleitung der Transportgleichung reproduzieren und die allgemeine Lösungsformel des dazugehörigen Anfangswertproblems angeben können. Außerdem sollten sie die Methode der Charakteristiken erklären und ihre Anwendung bei der Herleitung der allgemeinen Lösung aufzeigen können. Sie sollten auch in der Lage sein, wichtige Eigenschaften der Lösung anhand der Lösungsformel nachweisen zu können.

Abschnitte: Die Wellengleichung in einer Dimension (The wave equation in one dimension)

Nach dem Lesen dieses Abschnitts sollten Sie in der Lage sein,

- die Wellengleichung in einer Dimension zu formulieren und physikalisch herzuleiten.
- zu sagen, was man unter einer klassischen Lösung versteht und Ihnen sollten die typischen Randbedingungen bekannt sein.
- die allgemeine Lösungsformel der Wellengleichung auf $\mathbb{R}$ und $[0, \infty)$ anzugeben und herzuleiten.
- die Eindeutigkeit der Lösung mit Energiemethoden nachzuweisen.

Abschnitte: Die Wellengleichung in höheren Dimensionen (The wave equation in higher dimension)

Nach dem Lesen dieses Abschnitts sollten Sie in der Lage sein,

- die Wellengleichung in höheren Dimensionen auf Gebieten $\Omega \subset \mathbb{R}^d$ mit geeigneten Randbedingungen formulieren zu können.
- die expliziten Lösungen des Anfangswertproblems in geraden und ungeraden Raumdimensionen anzugeben. Auch sollten Sie die Herleitung der allgemeinen Lösung nachvollzogen haben.
- die Eindeutigkeit der klassischen Lösung nachzuweisen.

Chapter 1

Hyperbolic equations and the method of characteristics

This chapter is devoted to the study of two classes of partial differential equations; more precisely, of evolution equations: as anticipated in the introduction, an *evolution equation* is a partial differential equation whose unknown is a function of $1 + d$ variables, the first one identified with time and the remaining ones typically (but not necessarily) with space coordinates. Given an open domain $\Omega \subset \mathbb{R}^d$, a rather general evolution equation takes the form

$$\frac{\partial u}{\partial t}(t, x) = F(t, x, u(t, x), \nabla u(t, x), \ldots), \qquad t \geq 0,\ x \in \Omega, \tag{1.1}$$

where F is a function combining the values of u as well as of its gradient ∇u and of its further partial derivatives in a linear or nonlinear fashion: such evolution equations are, accordingly, called **linear** and **nonlinear**, respectively.

We begin by considering *transport equations* on one-dimensional intervals: these are arguably the simplest partial differential equations, as their investigation largely relies upon methods of elementary differential and integral calculus.

We then pass to a class of equations that describe wave-like propagation phenomena in strings or membranes: as you can guess based on your knowledge of the theory of ordinary differential equations, even differential equations whose order with respect to t is two or higher can be reduced to the form (1.1) upon introducing new appropriate vector-valued functions u, thus effectively turning (1.1) into a *system* of partial differential equations.

What these equations have in common is their *finite speed of propagation*: if we consider two initial data that only differ in a bounded open set, then the corresponding solutions will only differ in a bounded neighborhood of this set, even though this neighborhood will actually grow over time and eventually exhaust the whole spacial domain of the partial differential equation. In our approach, this will follow directly from the semi-explicit solution formulae we can derive.

One usually refers to equations displaying this property as *hyperbolic*: it turns out that the above equations – and further hyperbolic equations – can be studied by common methods.

1.1 The transport equation in one dimension

We want to model a transport process of some incompressible fluid, neglecting turbulence (for instance neutrons in a reactor) inside a tube of unit section: in this way the transport is forced to take place along the axis of the tube only. Considering a coordinate system whose x-axis is parallel to the tube, we can therefore regard the domain Ω occupied by the tube as effectively one-dimensional: in practice, unless noted otherwise we represent the tube in mathematical terms as an interval $\Omega = (a, b) \subset \mathbb{R}$, for $-\infty \leq a < b \leq +\infty$; we will turn to higher dimensional settings at the end of this section.

We denote by $u(t,x)$ the **density** of transported matter at point x and time t. The quantity of matter contained in the tube between the points x and $x+\Delta x$ at time t is

$$\int_x^{x+\Delta x} u(t,\xi)\ \mathrm{d}\xi.$$

The difference between the quantity of matter leaving this piece of tube at time $t+\Delta t$ and that entering it at time t is given by

$$\int_x^{x+\Delta x} u(t+\Delta t,\xi)\ \mathrm{d}\xi - \int_x^{x+\Delta x} u(t,\xi)\ \mathrm{d}\xi = \int_x^{x+\Delta x} (u(t+\Delta t,\xi)-u(t,\xi))\,\mathrm{d}\xi.$$

The **flow** of matter (i.e., the quantity of matter crossing a certain section of the tube per unit time) is described by a function ψ. Between times t and $t+\Delta t$ and at the section corresponding to the point x, it is measured by

$$\int_t^{t+\Delta t} \psi(\tau,x)\ \mathrm{d}\tau.$$

Assuming that there are neither sources nor sinks in the tube, the matter has to be conserved: the difference between the matter contained in the piece of tube between times $t+\Delta t$ and t agrees with the difference between the matter that has been flowing through the tube at times t to $t+\Delta t$ between the sections at x and $x+\Delta x$; in other words,

$$\int_x^{x+\Delta x} (u(t+\Delta t,\xi)-u(t,\xi))\ \ \mathrm{d}\xi = \int_t^{t+\Delta t} (\psi(\tau,x)-\psi(\tau,x+\Delta x))\ \ \mathrm{d}\xi.$$

For the sake of simplicity, one may first want impose the – physically not unreasonable – assumption that the density and flow functions u and ψ are continuously differentiable, so that dividing by Δt and passing to the limit $\Delta t \to 0$ we obtain on the one hand by Lebesgue's Differentiation Theorem A.27

$$\begin{aligned}\int_x^{x+\Delta x} \frac{\partial u}{\partial t}(t,\xi)\ \mathrm{d}\xi &= \lim_{\Delta t\to 0}\frac{1}{\Delta t}\int_t^{t+\Delta t} (\psi(\tau,x)-\psi(\tau,x+\Delta x))\ \ \mathrm{d}\xi \\ &= (\psi(t,x)-\psi(t,x+\Delta x))\,.\end{aligned}$$

On the other hand, dividing and passing to the limit – this time with respect to the space variable – again by Lebesgue's Differentiation Theorem we get

$$\frac{\partial u}{\partial t}(t,x) = \lim_{\Delta x\to 0}\frac{1}{\Delta x}\int_x^{x+\Delta x}\frac{\partial u}{\partial t}(t,\xi)\ \mathrm{d}\xi = -\frac{\partial \psi}{\partial x}(t,x). \tag{1.2}$$

If however a diffused source or sink is present, we can model them by a function $f:\mathbb{R}_+\times\Omega\to\mathbb{R}$: this leads us to generalizing the above differential equation to

$$\frac{\partial u}{\partial t}(t,x) = -\frac{\partial \psi}{\partial x}(t,x) + f(t,x).$$

In the special case of transport of matter with small density, it is physically reasonable to assume flow and density to be proportional by a factor c, say,

$$\psi(t,x) = c(t,x)u(t,x),$$

and we finally obtain the partial differential equation

$$\frac{\partial u}{\partial t}(t,x) = -c(t,x)\frac{\partial u}{\partial x}(t,x) + f(t,x), \qquad t\geq 0,\ x\in\Omega. \tag{1.3}$$

This is the **transport equation**, obtained under the assumptions that

- the tube is thin enough that only transport processes in the longitudinal direction take place,
- the functions u,ψ describing density and flow are continuously differentiable, and

- density and flow are proportional by the factor c.

Remark 1.1. Observe that if $f \not\equiv 0$, then, for example,

$$-c(t,x)\frac{\partial(2u)}{\partial x}(t,x) + f(t,x) \neq 2\left(-c(t,x)\frac{\partial u}{\partial x}(t,x) + f(t,x)\right).$$

Thus, for $\not\equiv 0$ the right hand side of (1.3) depends on u nonlinearly, and the transport equation is strictly speaking a nonlinear one. We will soon see that, however, this kind of nonlinearity is easy to deal with: for this reason, one calls *inhomogeneous* those differential equations that can be split in a linear part and a term that only depends on the elementary variables t, x, but *not* on the unknown u. The name *nonlinear* is informally reserved to differential equations, such as the *inviscid Burgers equation*

$$\frac{\partial v}{\partial t}(t,x) = -\frac{\partial v^2}{\partial x}(t,x), \qquad t > 0,\ x \in \mathbb{R},$$

whose nonlinear dependence on the unknown is more substantial.

Generally speaking, the transport equation (1.3) may however not have a unique solution: if, for instance, $f \equiv 0$ and $\Omega = \mathbb{R}$, then it is clear that any constant function (i.e., any function that is constant in both the time and space variables) solves the *homogeneous* transport equation

$$\frac{\partial u}{\partial t}(t,x) = -c(t,x)\frac{\partial u}{\partial x}(t,x), \qquad t \geq 0,\ x \in \Omega; \tag{1.4}$$

indeed it can – and will later, in Theorem 1.3 – be shown that further solutions exist. One way of classifying them – admitting some, ruling out others – is to consider the values attained by a candidate solution at a conventional initial time instant, comparing them with the values that can be expected based on physical principles.

We known from the theory of ordinary differential equations that the solution of a first order equation can be uniquely determined upon imposing one initial condition. In analogy with this situation, we will impose an **initial condition**

$$u(0,x) = u_0(x), \qquad x \in \Omega, \tag{1.5}$$

here, too, for some $u_0 : \Omega \to \mathbb{R}$.

Given open domains[1] $\Omega_1 \subset \mathbb{R}^{d_1}$ and $\Omega_2 \subset \mathbb{R}^{d_2}$, here and in the following we denote by $f \in C^{h,k}(\Omega_1 \times \Omega_2)$ the space of scalar-valued functions supported on $\Omega_1 \times \Omega_2$ that are h-times continuously differentiable with respect to the first d_1 variables and k-times continuously differentiable with respect to the last d_2 variables. We will in general not require joint continuity, or continuous differentiability in both sets of variables.

Definition 1.2. A **classical solution** of the initial value problem for the transport equation is a function $u : \mathbb{R}_+ \times \Omega \to \mathbb{R}$ that

- is $u \in C^{1,1}(\mathbb{R}_+ \times \Omega)$,
- satisfies (1.3), and
- satisfies (1.5).

We next propose a first set of conditions, including $\Omega = \mathbb{R}$, implying existence of a classical solution to the transport equation in the sense of Definition 1.2.

Theorem 1.3. *Let $u_0 \in C^1(\mathbb{R})$ and $f \in C^{0,1}(\mathbb{R}_+ \times \mathbb{R})$. If the velocity function c is constant, then the initial value problem*

$$\begin{cases} \dfrac{\partial u}{\partial t}(t,x) = -c\dfrac{\partial u}{\partial x}(t,x) + f(t,x), & t \geq 0,\ x \in \mathbb{R}, \\ u(0,x) = u_0(x), & x \in \mathbb{R}, \end{cases} \tag{1.6}$$

has a classical solution given by

$$u(t,x) = u_0(x - ct) + \int_0^t f(\tau, x + c(\tau - t))\ \mathrm{d}\tau, \quad t \geq 0,\ x \in \mathbb{R}. \tag{1.7}$$

[1] Throughout this lecture note, any open connected subset of either $\mathbb{R}^d$ or $\mathbb{C}^d$ is called a **domain**. In the literature, domains are usually – but not always! – assumed to be open by definition. To avoid confusion, we prefer to state this explicitly and thus call any open connected subset of $\mathbb{R}^d$ an "open domain".

Before proving this result, let us appreciate the presented expression. The rather explicit formula (1.7) shows that the given classical solution is a wave that propagates to the right (if $c > 0$) or to the left (if $c < 0$) with speed c and whose profile is constantly that of u_0: one sometimes refers to it as a **shift** of u_0.

This means that all the properties of u_0 that only depend on its graph are left invariant under time evolution: smoothness, integrability, convexity/concavity and positivity, to name a few. In particular, for the sake of later reference we note the following.

Corollary 1.4. *Under the assumptions of Theorem 1.3, let $k \geq 2$ and consider the classical solution u given by* (1.7)*. If $u_0 \in C^k(\mathbb{R})$ and $f \in C^{k-1,k}(\mathbb{R}_+ \times \mathbb{R})$, then this classical solution is in fact of class $C^{k,k}(\mathbb{R}_+ \times \mathbb{R})$.*

If $f \equiv 0$, then $u_0(x) \geq 0$ for all $x \in \mathbb{R}$ if and only if $u(t,x) \geq 0$ for all $t \geq 0$ and all $x \in \mathbb{R}$. If $f(t,x) \geq 0$ for all $t \geq 0$ and all $x \in \mathbb{R}$ and $u_0(x) \geq 0$ for all $x \in \mathbb{R}$, then $u(t,x) \geq 0$ for all $t \geq 0$ and all $x \in \mathbb{R}$.

Remark 1.5. A classical solution's positivity is an important property. Physically speaking, observing positivity of the density function provided the initial density is positive is a feature that heuristically confirms the modeling qualities of the transport equation.

Our proof will be based on a common tool in the analysis of transport equations: the so-called **method of characteristics**. Instead of introducing this method formally, we here prefer a learning-by-doing approach: we will see this method in action in a simple case and try to identify the core ideas later on. The method uses a few notions from the analytical theory of curves: we refer to [9, Chapter 6]. With a slight abuse of notation we will in the following, as usual, avoid to distinguish between a curve $\gamma : I \to \mathbb{R}^m$ and its image $\gamma(I) \subset \mathbb{R}^m$, where $I \subset \mathbb{R}$ is an Interval.

Definition 1.6. Let $\Omega \subset \mathbb{R}^d$ be an open domain. A continuously differentiable curve $\gamma : \mathbb{R}_+ \to \Omega$ defines a **characteristic** Γ of the evolution equation (1.1) if each classical solution u of (1.1) is constant along the curve

$$\Gamma := \{(s, \gamma(s)) \in \mathbb{R}_+ \times \Omega\}$$

i.e., $u(s, \gamma(s)) \equiv const$ for all $s \geq 0$ whenever u satisfies identically (1.1).

This means in particular that

$$u(t, \gamma(t)) = u(0, \gamma(0)) \qquad \text{for all } t \geq 0,$$

and, in particular, if an initial condition (1.5) is imposed,

$$u_0(\gamma(0)) = u(t, \gamma(t)) \qquad \text{for all } t \geq 0. \tag{1.8}$$

We stress that while, by definition, $u(t_1, \gamma(t_1)) = u(t_2, \gamma(t_2))$ for all $t_1, t_2 \in \mathbb{R}_+$, we do generally *not* expect that $u(t_1, x_1) = u(t_2, x_2)$ for general $(t_1, x_1), (t_2, x_2) \in \mathbb{R}_+ \times \Omega$. However, we do expect that each point $(t, x) \in \mathbb{R}_+ \times \Omega$, is crossed by a characteristic: indeed, the method we are going to present relies upon the condition that the characteristics of a given partial differential equation form a **family of curves** that fills the space-time $\mathbb{R}_+ \times \Omega$.

Proof of Theorem 1.3. Let first $f = 0$. We will make the ansatz that under our assumption $\Omega = \mathbb{R}$ there exists a characteristic of (1.4) in the sense of Definition 1.6, i.e., we will first assume that there is a continuously differentiable function $\gamma : \mathbb{R}_+ \to \mathbb{R}$ defining the curve Γ such that classical solutions of (1.4) are constant along Γ and then aim at recovering γ: the fact that such a γ actually exists, as we are going to see, will justify a posteriori the ansatz.

To begin with, observe that Γ is a characteristic if and only if

$$0 = \frac{\mathrm{d}u}{\mathrm{d}s}(t,x)_{|(t,x)=(s,\gamma(s))} = \frac{\partial u}{\partial t}(s, \gamma(s)) + \gamma'(s)\frac{\partial u}{\partial x}(s, \gamma(s)) \qquad \text{for all } s \geq 0, \tag{1.9}$$

i.e., comparing with (1.4), if and only if

$$\gamma'(s) \equiv c, \qquad s \geq 0. \tag{1.10}$$

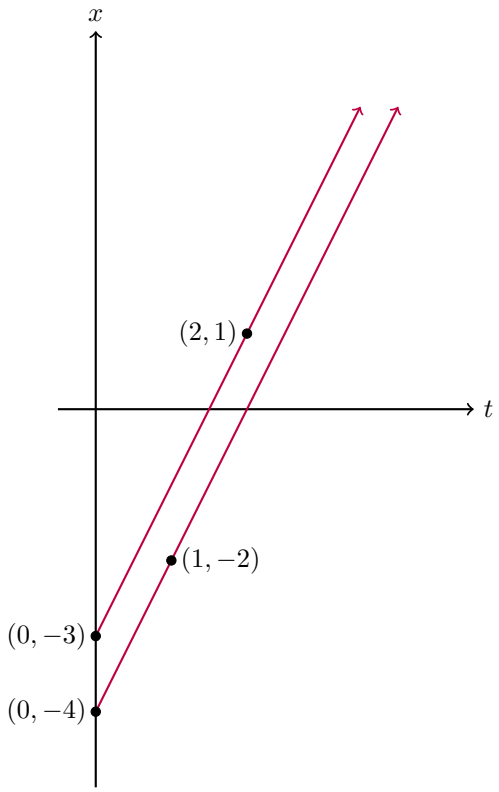

Figure 1.1: Characteristics for the homogeneous transport equation (1.4) on $\Omega = \mathbb{R}$ across the points $(2, 1)$ and $(1, -2)$; the arrows reflect the parametrization of the characteristics. Here, $c \equiv 2$.

This is the ordinary differential equation defining all characteristics of (1.4). Because c is by assumption constant, (1.10) yields

$$\gamma(s) = cs + \gamma(0), \qquad s \geq 0 : \tag{1.11}$$

each characteristic of (1.21) is hence a line of slope c. Now, if (t, x) lies on the same characteristic that crosses $(0, \gamma(0))$, i.e., if

$$x = ct + \gamma(0),$$

we finally conclude by (1.8) that, whenever $f \equiv 0$, a classical solution u is given by

$$u(t, x) = u_0(\gamma(0)) = u_0(x - ct) \qquad \text{for all } t \geq 0,\ x \in \mathbb{R}; \tag{1.12}$$

observe that $x - ct$ is precisely the ordinate of the point where the characteristic reaches the x-axis. The expression (1.12) makes sense, because $x - ct$ lies in the domain of definition of u_0: at this step, we are crucially using the assumption $\Omega = \mathbb{R}$. Furthermore, the computations in (1.9) show that u is actually a classical solution of (1.6) for $f \equiv 0$.

Let us now consider the case of inhomogeneous transport equations, i.e., of (1.3) with $f \neq 0$. A general principle allows to obtain a classical solution, in a way that resembles the variation of constant formula of ordinary differential equations. Namely, the classical solution is given by

$$u(t, x) = u_0(\gamma(0)) + \int_0^t f(\tau, \gamma(\tau))\mathrm{d}\tau,$$

where γ is a parametrization of the characteristic (of the associated homogeneous problem) which crosses

(t, x). This is due to the fact that

$$\begin{aligned}\frac{\mathrm{d}u}{\mathrm{d}s}(s,\gamma(s)) &= \frac{\partial u}{\partial t}(s,\gamma(s)) + \frac{\partial u}{\partial x}(s,\gamma(s))\gamma'(s)\\ &= \frac{\partial u}{\partial t}(s,\gamma(s)) + c(\gamma(s))\frac{\partial u}{\partial x}(s,\gamma(s))\\ &= f(s,\gamma(s)).\end{aligned}$$

Accordingly, by the fundamental theorem of calculus

$$u(t,x) = u(t,\gamma(t)) = u_0(\gamma(0)) + \int_0^t \frac{\mathrm{d}u}{\mathrm{d}s}(s,\gamma(s))\mathrm{d}s = u_0(\gamma(0)) + \int_0^t f(s,\gamma(s))\mathrm{d}s.$$

This completes the proof. □

Exercise 1.7. While Theorem 1.3 has been formulated for classical solutions of the transport equation on the space-time domain $\mathbb{R}_+ \times \mathbb{R}$, the classical solution formula (1.7) can be extended to $\mathbb{R} \times \mathbb{R}$, too. Work out the details: how should the assumptions of the theorem be modified? What partial differential equations is solved by this new expression?

Remark 1.8. (1) The above proof shows that the crucial step in this approach is the possibility of solving the ordinary differential equation (1.10). The same method also works in the general case of non-constant c: if c depends on t and/or on x, then one can mimic the proof and find a classical solution of the homogeneous problem in terms of solutions of the Cauchy problem

$$\begin{cases}\dfrac{\mathrm{d}\gamma}{\mathrm{d}s}(s) = c(s,\gamma(s)), \qquad s \geq 0,\\ \gamma(0) = \gamma_0,\end{cases} \tag{1.13}$$

we have effectively turned the search for classical solutions of a partial differential equation into the analysis of a flow of an ordinary differential equation.

(2) Observe that if the initial value problem (1.3)–(1.5) admits a classical solution u, then by definition u is regular enough that the computation in (1.9) can be repeated along a characteristic curve parametrized by γ. Thus, γ is necessarily a classical solution of the ordinary differential equation (1.13) with initial data $\gamma(0)$. If c satisfies appropriate assumptions that ensure that the Cauchy problem (1.13) has at most one global classical solution, then the classical solution u is uniquely determined by the formula (1.7).

Exercise 1.9. The method of characteristics is also suitable for dealing with transport equations with coefficient c depending on time, like

$$x\frac{\partial u}{\partial t}(t,x) = t\frac{\partial u}{\partial x}(t,x), \qquad t \geq 0,\ x \in \mathbb{R}.$$

Solve this partial differential equation working out the details of the idea sketched in Remark 1.8, with initial value $u(0,x) = 3x^2$, $x \in \mathbb{R}$.

A close analysis of the proof of Theorem 1.3 shows that for the proposed method to work it is crucial that each point (t,x) can be expressed as $(t,\gamma(t))$, i.e., that each (t,x) is crossed by a characteristic – in the case of (1.4) on $\Omega = \mathbb{R}$ and constant c, a line of slope c – connecting it with a point $(0,\gamma(0))$ on the x-axis, at which place the value of u_0 can be read off and "transmitted" to (t,x). This is guaranteed if we study the transport equation on $\Omega = \mathbb{R}$. But what if $\Omega \subsetneq \mathbb{R}$? Given a point (t,x) in the space-time domain, let us follow the line of slope c crossing the point for $t \to \infty$ and for $t \to 0$: things may change in comparison with the simple case of $\Omega = \mathbb{R}$.

- It may happen that this line originates from the x-axis, crosses (t,x), and then leaves the space-time domain;
- or else, following such a line backward, i.e., for times prior to t, we may see that it leaves the space-time domain before reaching the x-axis.

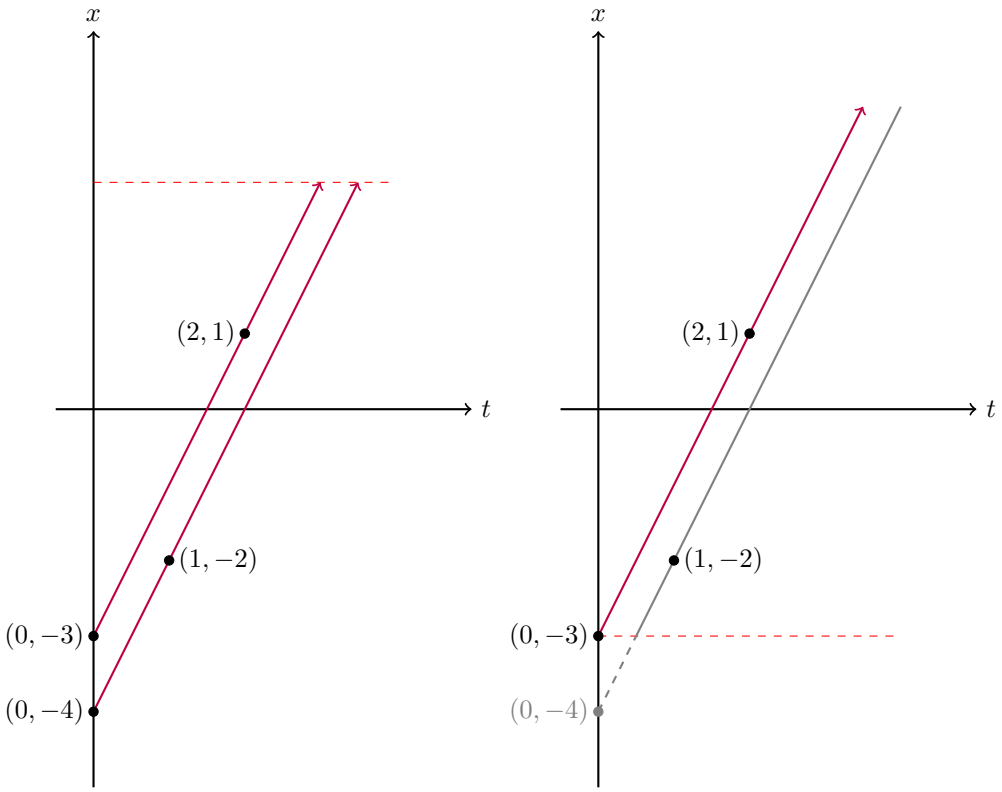

Figure 1.2: Characteristics for the homogeneous transport equation (1.4) across the points $(2,1)$ and $(1,-2)$, here for $\Omega = (-\infty, 3)$ (left) and $\Omega = (3, \infty)$ (right), respectively. The characteristic through $(1,-2)$ does not cross the x-axis at a point where the initial data u_0 is defined. As in Figure 1.2, the plots correspond to the choice $c \equiv 2$.

The former case does not pose any problem; the latter is more delicate and motivates us to turn to a most interesting topic: that of *boundary conditions*, i.e., of conditions that we expect to be satisfied by a solution's restriction to the boundary $\partial\Omega$ of the relevant spacial domain Ω. If we do believe that the transport equation (1.3) is a correct model of the physical reality, than it is reasonable to wish for existence of a uniquely determined classical solution: repeating a physical experiment over and over we do expect to observe the same outcome each time, after all. This behavior of a partial differential equation – namely, existence and uniqueness of classical solution to a given equation, under appropriate conditions at the boundary of the space-time domain – is often referred to as **well-posedness**.

It turns out that not all ways of studying boundary values are appropriate: imposing *wrong conditions*, or imposing unnecessary conditions on *wrong portions of the boundary*, makes the problem unsolvable; whereas, imposing too unrestrictive conditions – too few ones, or on too little portions of the boundary – typically fails to avoid coexistence of several – and then, due to linearity of the equation, infinitely many – classical solutions.

One may naively guess that since the transport equation is also of first order in the space variable, imposing *one* boundary condition is also necessary in order to deduce uniqueness of classical solutions. It may come as a surprise that this is not always true. To see this, let us consider for the sake of simplicity the case of $f = 0$ and focus on the examples depicted in Figure 1.2.

Let us begin with the first one, corresponding to the space domain $\Omega = (-\infty, 3)$. The same computations in the proof of Theorem 1.3 show that (1.7) still yields a classical solution

$$u(t,x) = u_0(x - ct) + \int_0^t f(\tau, x + c(\tau - t))\ \mathrm{d}\tau, \quad t \geq 0,\ x \in (-\infty, 3), \tag{1.14}$$

to

$$\begin{cases} \dfrac{\partial u}{\partial t}(t,x) = -c\dfrac{\partial u}{\partial x}(t,x) + f(t,x), & t \geq 0,\ x \in (-\infty, 3), \\ u(0,x) = u_0(x), & x \in (-\infty, 3), \end{cases} \tag{1.15}$$

and reasoning as in Remark 1.8.(2) one sees that this is the only possible classical solution. Hence, imposing an additional condition at the boundary, i.e., at $x = 3$, like

$$u(t,3) = 0, \qquad t \geq 0,$$

would in general not be compatible with (1.7) and hence simply prevent the otherwise reasonable function in (1.14) to be an acceptable classical solution.

If, on the other hand, the space domain $\Omega = (-3, \infty)$ is considered, then Figure (1.2) shows that there are characteristics that do not reach the x-axis inside the space-time domain $\mathbb{R}_+ \times (-3, \infty)$ and hence cannot be used to define the classical solution by means of the value of the initial data u_0 at a suitable point. In fact, Figure (1.2) shows that this problem is rather fundamental: for each $x \in (-3, \infty)$ there is some t_0 such that the characteristic through (t, x) does not cross the x-axis inside the relevant space-time domain. This means that the classical solution to

$$\begin{cases} \dfrac{\partial u}{\partial t}(t,x) = -c\dfrac{\partial u}{\partial x}(t,x) + f(t,x), & t \geq 0,\ x \in (3, \infty), \\ u(0,x) = u_0(x), & x \in (3, \infty), \end{cases} \tag{1.16}$$

is not unique: *any* extension of u_0 to the whole $\mathbb{R}$ would by means of (1.7) define a function whose restriction to $(3, \infty)$ solves (1.16). So, the question is how to recover a *unique* classical solution. It turns out that imposing a condition like

$$u(t,3) = \phi(t), \qquad t \geq 0,$$

at the boundary $x = -3$ of $\Omega = (-3, \infty)$ is sufficient: the value of ϕ at the point $t_b = t - \frac{3+x}{c}$ (where the characteristic that crosses (t, x) touches the boundary of the space-time domain) allows us to recover the value of the classical solution of u at (t, x).

Summing up, we have obtained the following. We only formulate the result for the homogeneous case of $f \equiv 0$.

Theorem 1.10. *Let the velocity function c be constant and positive. Then the following assertions hold.*

(1) Let $u_0 \in C^1\big((-\infty, b]\big)$. If the velocity function c is constant and positive, then the initial value problem

$$\begin{cases} \dfrac{\partial u}{\partial t}(t,x) = -c\dfrac{\partial u}{\partial x}(t,x), & t \geq 0,\ x \in (-\infty, b], \\ u(0,x) = u_0(x), & x \in (-\infty, b], \end{cases} \tag{1.17}$$

has a unique classical solution given by

$$u(t,x) = u_0(x - ct), \quad t \geq 0,\ x \in (-\infty, b]. \tag{1.18}$$

(2) Let $u_0 \in C^1\big([a, \infty)\big)$ and $\phi \in C^1(\mathbb{R}_+)$ such that $u_0(a) = \phi(0)$. Then the initial-boundary value problem

$$\begin{cases} \dfrac{\partial u}{\partial t}(t,x) = -c\dfrac{\partial u}{\partial x}(t,x), & t \geq 0,\ x \in [a, \infty), \\ u(0,x) = u_0(x), & x \in [a, \infty), \\ u(t,a) = \phi(t), & t \geq 0, \end{cases} \tag{1.19}$$

has a unique classical solution given by

$$u(t,x) = \begin{cases} u_0(x - ct), & t \geq 0,\ x \in [a + ct, \infty), \\ \phi\left(t + \dfrac{a - x}{c}\right), & t \geq 0,\ x \in [a, a + ct]. \end{cases} \tag{1.20}$$

If the velocity c is not positive but *negative*, a simple change of variables (work out the details!) shows that the two cases are switched: no (respectively, one) boundary condition should be imposed at the left (respectively, right) endpoint of the interval.

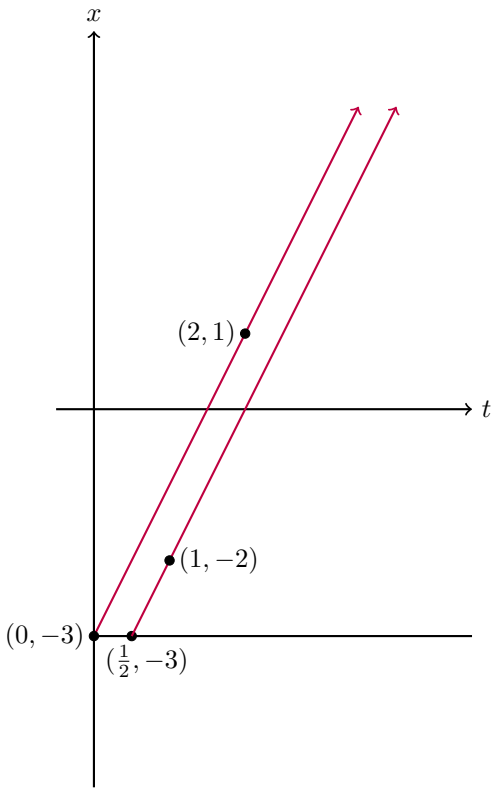

Figure 1.3: Imposing conditions at the portion of the boundary where the characteristics originate is sufficient to recover the values of classical solutions at each (t, x) of the space-time domain; compare with Figure 1.2 (right).

Exercise 1.11. Apply the variation of constant principle like in the proof of Theorem 1.3 to extend the formulae in Theorem 1.10 to the inhomogeneous case $f \not\equiv 0$.

Especially when $f \equiv 0$, and hence the transport equation on $\mathbb{R}$ is linear, it may be interesting to inherit linearity in the case of smaller domains Ω, too. To this aim, the boundary data ϕ have to be chosen in such a way that the condition $u(t, a) = \phi(t)$ is linear. A natural way to do so while satisfying the compatibility condition $u_0(a) = \phi(0)$ (which in turn is crucial to obtain a classical solution in the sense of Definition 1.2) is to choose as initial data

$$\phi \equiv u_0(a).$$

What if Ω is a bounded interval, say, $\Omega = (a, b)$? The above discussion should convince you that nothing really unexpected happens: again, in order to force uniqueness of classical solutions, a boundary condition needs to be imposed at $x = a$; while in order not to jeopardize existence of classical solutions, no boundary condition can be imposed at $x = b$. However, what *is* possible is to impose on $u(t, a)$ a linear condition that depends on $u(t, b)$, like

$$u(t, a) = \kappa u(t, b), \qquad t \geq 0,$$

for some $\kappa \in \mathbb{R}$. The case of $\kappa = 0$ and $\kappa = 1$ are particularly relevant: they are referred to as **zero boundary condition** (or sometimes **Dirichlet boundary conditions**) and **periodic condition**, respectively. The former one describes a model where the mass that gets lost upon touching the boundary point $x = b$ is not replaced; the latter can be thought of as describing a transport process along a closed loop, like a toy model of a proton moving along the Large Hadron Collider at CERN.

Remark 1.12. As we have seen in Exercise 1.7, the formula (1.7) makes perfectly sense also if we extend it to a function defined on $\mathbb{R} \times \mathbb{R}$, provided that also u_0 and f are extended accordingly. This means that the transport equation on $\mathbb{R}$ is uniquely solvable backward as well as forward in time.

This is in sharp contrast to the case of a transport equation on a bounded interval, in which case boundary conditions necessarily have to be imposed: studying the behavior of the classical solution of (1.4) backward in time, i.e., of the classical solution of

$$\frac{\partial u}{\partial t}(t,x) = -c(t,x)\frac{\partial u}{\partial x}(t,x), \qquad t \leq 0,\ x \in \Omega,$$

is equivalent (why?) to studying forward well-posedness of

$$\frac{\partial u}{\partial t}(t,x) = c(t,x)\frac{\partial u}{\partial x}(t,x), \qquad t \geq 0,\ x \in \Omega.$$

However, we know that in order for (1.4) to admit a unique classical solution, boundary conditions have to be imposed on the "right" portion of its boundary; switching from the coefficient $-c$ to c *without changing the boundary conditions* suddenly turns a well-posed problem into an overdetermined problem, as a condition is imposed on the portion of the boundary where the characteristics *end*. Accordingly, a forward well-posed transport equation is generally not backward well-posed. The periodic boundary conditions

$$u(t,a) = u(t,b), \qquad t \geq 0,$$

which are symmetric upon swapping of the endpoints of the interval $[a,b]$, are a noteworthy exception to this rule: indeed, the initial-value problem for the transport equation with such periodic conditions is both forward and backward well-posed.

Remark 1.13. The model studied in this section can be generalized to higher, d-dimensional case, $d \geq 2$, upon considering some open domain $\Omega \subset \mathbb{R}^d$ and prescribing the direction of transport by means of a velocity vector $c \in \mathbb{R}^d$: the transport is then described by

$$\frac{\partial u}{\partial t}(t,x) = -c(x) \cdot \nabla u(t,x) + f(t,x), \qquad t \geq 0,\ x \in \Omega, \tag{1.21}$$

where the vector fields $x \mapsto \nabla u(t,x)$, $t \geq 0$, denote the gradient of u at point x and time t. Observe that, again, we can derive the transport equation from consequence of the continuity equation (1.43) together with the proportionality relation (1.44), which in this case features vector-valued c, j.

One can see that for $f \equiv 0$ and constant c a classical solution of (1.21) is given by

$$u(t,x) = u_0(\gamma(0)) = u_0(x - ct), \tag{1.22}$$

since the d-dimensional characteristic is given by

$$\gamma(s) = x + c(s - t). \tag{1.23}$$

Try repeating the arguments leading to (1.3) in the general d-dimensional case!

Although finding correct boundary conditions is not always easy (all the less so for higher dimensional transport equations!), the formula for the classical solution of the transport problem shows the profile of a function (the initial data) moving (towards the boundary) with finite speed c. Therefore, if the initial data are given by a function concentrated in a ball far from the boundary, this formula keeps its complete validity as long as the traveling initial density function does not reach the boundary.

1.2 The wave equation in one dimension

We are now going to discuss another important class of partial differential equations: this will be made possible, again, by an application of the method of characteristics.

Consider a string of linear density ρ. Assume the string to have horizontally constant tension T, i.e., $\frac{\partial T}{\partial x}(t,x) \equiv 0$. We want to model the vertical oscillations of the string, assuming that they are small enough that the horizontal extension can be neglected. We denote by $u(t,x)$ the vertical extension at each time t and each point x of the string, with respect to a reference level which can be set at 0 without loss of generality. At each point x and time t the row undergoes an acceleration

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

By Newton's second law ("$F = ma$") the force applied to each "infinitesimally small" piece of string of length Δx (which has mass $\rho\Delta x$, by assumption) is

$$\rho(x)\Delta x\frac{\partial^2 u}{\partial t^2}(t,x).$$

Since the tension is assumed to be constant in the horizontal direction, the force acting on the considered piece is given by the difference of the vertical component of the tension at $x+\Delta x$ and x, i.e.,

$$T(t)\sin\theta(t,x+\Delta x) - T(t)\sin\theta(t,x),$$

where $\theta(t,x)$ denotes the angle of the string at each point x at time t. We also ought to add external forces (like the gravity) acting vertically, whose net magnitude we denote by $\Phi(t,x)\Delta x$. All in all the balance equation is

$$\rho(x)\Delta x\frac{\partial^2 u}{\partial t^2}(t,x) = T(t)\sin\theta(t,x) - T(t)\sin\theta(t,x) + \Phi(t,x)\Delta x,$$

i.e.,

$$\rho(x)\frac{\partial^2 u}{\partial t^2}(t,x) = \frac{T(t)\sin\theta(t,x+\Delta x) - T(t)\sin\theta(t,x)}{\Delta x} + \Phi(t,x).$$

Passing to the limit for $\Delta x \to 0$ we obtain

$$\begin{aligned}\rho(x)\frac{\partial^2 u}{\partial t^2}(t,x) &= \frac{\partial}{\partial x}\left(T(t)\sin\theta(t,x)\right) + \Phi(t,x)\\ &= T(t)\frac{\partial}{\partial x}\sin\theta(t,x) + \Phi(t,x)\\ &= T(t)\cos\theta(t,x)\frac{\partial\theta}{\partial x}(t,x) + \Phi(t,x).\end{aligned}$$

Moreover, the vertical extensions are supposed to be small, and accordingly $\cos\theta(t,x) \cong 1$ and $\theta(t,x) \cong \sin\theta(t,x) \cong \frac{\partial u}{\partial x}(t,x)$, whence

$$\frac{\partial\theta}{\partial x}(t,x) \cong \frac{\partial^2 u}{\partial x^2}(t,x),$$

i.e.,

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

which can be (approximately) written as

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

provided T does not depend on time. Here

$$F(t,x) := \frac{1}{\rho(x)}\Phi(t,x), \qquad t \geq 0,\ x \in \Omega.$$

One usually calls

$$c(x) := \sqrt{\frac{T}{\rho(x)}} \tag{1.24}$$

the **wave speed**. We are thus led to

$$\frac{\partial^2 u}{\partial t^2}(t,x) = c^2(x)\frac{\partial^2 u}{\partial x^2}(t,x) + F(t,x), \qquad t \geq 0,\ x \in \Omega, \tag{1.25}$$

which is the most common form of the one-dimensional (homogeneous, if $F = 0$) **wave equation** on a (possibly unbounded) open interval $\Omega \subset \mathbb{R}$: we recall that it has been obtained under the assumptions that

- the horizontal tension T does not depend on space and on time, and
- the vertical extensions are small and the horizontal ones negligible.

If the above argument is repeated in a second or a third (or a d-th) spacial dimension, we formally arrive to the two- or three- (or d-) dimensional wave equation

$$\frac{\partial^2 u}{\partial t^2}(t,x) = c^2(x)\Delta u(t,x) + F(t,x), \qquad t \geq 0,\ x \in \Omega, \tag{1.26}$$

in domains $\Omega \subset \mathbb{R}^d$ representing membranes or elastic bodies. Here we have introduced the common notation

$$\Delta u(t,x) := \sum_{k=1}^{d} \frac{\partial^2 u}{\partial x_k^2}(t,x),$$

for the **Laplacian** Δu of the function u, where the sum is taken over the spacial dimensions.

Since the differential equation is of second order, in analogy with the case of ODEs of second order we are led to introduce two initial conditions

$$u(0,x) = u_0(x), \qquad \frac{\partial u}{\partial t}u(0,x) = u_1(x), \qquad x \in \Omega. \tag{1.27}$$

We will see that these two boundary conditions are indeed necessary in order to solve the problem. Moreover, if Ω is bounded, also a boundary condition is necessary. In the one-dimensional case, common boundary conditions include

$$u(t,a) = u(t,b) \equiv 0, \qquad t \geq 0, \qquad \text{(Dirichlet b.c.);}$$

$$\frac{\partial u}{\partial x}(t,a) = \frac{\partial u}{\partial x}(t,b) \equiv 0, \qquad t \geq 0, \qquad \text{(Neumann b.c.);}$$

$$-\frac{\partial u}{\partial x}(t,a) = p_a u(t,a), \quad \frac{\partial u}{\partial x}(t,b) = p_b u(t,b), \qquad t \geq 0, \qquad \text{(Robin b.c.),}$$

for some $p_a, p_b > 0$; or

$$\frac{\partial^2 u}{\partial x^2}(t,a) = \frac{\partial^2 u}{\partial x^2}(t,b) \equiv 0, \qquad t \geq 0. \qquad \text{(Wentzell b.c.).}$$

(They have been proposed by Gustav Lejeune Dirichlet, Carl Gottfried Neumann, Victor Gustave Robin, and Alexander Wentzell, respectively.)

Observe that the choice of boundary conditions to be imposed is intimately related to the significant model we are going to describe. A system that approximately fulfills our physical conditions under which the wave equation (1.25) has been derived is a vibrating string in a musical instrument – say, a violin. While its tension can be assumed to be homogeneous, neither is a string massless, of course, nor do we want it to stay immobile; but a string *is* typically light enough that we can neglect its mass, and its vertical oscillations are typically small: all in all, it is reasonable to regard (1.25) as an approximate description of the oscillations of a violin string, with $F(t,x)$ describing the strength of a musician's bowing or plucking at time t and point x of the string. In virtually all string instruments, the string is clamped at its endpoints, so it cannot oscillate: a reasonably accurate way of describing this behavior boils down to imposing Dirichlet boundary conditions.

Definition 1.14. Let $\Omega \subset \mathbb{R}^d$. A **classical solution** of the initial value problem associated with the wave equation (1.26) is a function $u : \mathbb{R}_+ \times \Omega \to \mathbb{R}$ such that

- $u \in C^{2,2}(\mathbb{R}_+ \times \Omega)$,
- u satisfies (1.25) ,
- u satisfies (1.27).

Let us first prove that the wave equation (1.26) admits a classical solution at least in the case of $\Omega = \mathbb{R}$.

Theorem 1.15. *Let $\Omega = \mathbb{R}$, $u_0 \in C^2(\mathbb{R})$, and $u_1 \in C^1(\mathbb{R})$. If the wave speed c is constant, then the initial value problem for the homogeneous one-dimensional wave equation*

$$\frac{\partial^2 u}{\partial t^2}(t,x) = c^2 \frac{\partial^2 u}{\partial x^2}(t,x), \qquad t \geq 0,\ x \in \Omega, \tag{1.28}$$

has a classical solution given by

$$u(t,x) = \frac{1}{2}\left(u_0(x+ct) + u_0(x-ct)\right) + \frac{1}{2c}\int_{x-ct}^{x+ct} u_1(y)\mathrm{d}y, \qquad t \geq 0,\ x \in \mathbb{R}. \tag{1.29}$$

Like in the case of the transport equation, a possible and common approach to solve the *homogeneous* one-dimensional wave equations is based on the method of characteristics.

Proof. The crucial idea is to perform the factorization

$$\left(\frac{\partial^2}{\partial t^2} - c^2\frac{\partial^2}{\partial x^2}\right)u = \left(\frac{\partial}{\partial t} + c\frac{\partial}{\partial x}\right)\left(\frac{\partial}{\partial t} - c\frac{\partial}{\partial x}\right)u$$

of the homogeneous wave equation (1.28) (observe that this generally does not hold unless c is a constant), which holds at least for smooth functions, in view of Schwarz' Theorem on symmetry of second derivatives. This suggests to introduce the unknown v defined by

$$v(t,x) := \left(\frac{\partial}{\partial t} - c\frac{\partial}{\partial x}\right)u(t,x), \qquad t \geq 0,\ x \in \mathbb{R},$$

and to study the PDE

$$\left(\frac{\partial}{\partial t} + c\frac{\partial}{\partial x}\right)v(t,x) = 0, \qquad t \geq 0,\ x \in \mathbb{R},$$

in the unknown v. The latter equation is a plain, linear transport equation, as studied in Section 1.1. By Theorem 1.3, the classical solution of this equation is given by

$$v(t,x) = \tilde{\gamma}(x-ct), \qquad t \geq 0\ x \in \mathbb{R},$$

where $\tilde{\gamma}(\cdot)$ is the (not explicitly known) function defining the initial data of v, i.e., $v(0,\cdot)$. In particular, by definition of v we get

$$v(t,x) = \tilde{\gamma}(x-ct) = \frac{\partial u}{\partial t} - c(x)\frac{\partial u}{\partial x}, \qquad t \geq 0,\ x \in \mathbb{R},$$

which is an inhomogeneous transport equation in the unknown u with inhomogeneous term $\tilde{\gamma}$. By Theorem 1.3 we obtain

$$\begin{aligned}
u(t,x) &= u_0(x+ct) + \int_0^t v(\tau, x+c(t-\tau))\ \mathrm{d}\tau \\
&= u_0(x+ct) + \int_0^t \tilde{\gamma}(x+c(t-\tau)-c\tau)\ \mathrm{d}\tau \\
&= u_0(x+ct) + \int_0^t \tilde{\gamma}(x+ct-2c\tau)\ \mathrm{d}\tau \\
&= u_0(x+ct) + \frac{1}{2c}\int_{x-ct}^{x+ct} \tilde{\gamma}(y)\mathrm{d}y.
\end{aligned}$$

We still have to determine $\tilde{\gamma}$: the second boundary condition and the definition of v imply

$$\tilde{\gamma}(x) = v(0,x) = \frac{\partial u}{\partial t}(0,x) - c(x)\frac{\partial u}{\partial x}(0,x) = u_1(x) - cu_0'(x),$$

whence

$$u(t,x) = u_0(x+ct) + \frac{1}{2}\int_{x-ct}^{x+ct}\left(u_1(y) - u_0'(y)\right)\mathrm{d}y.$$

This concludes the proof. □

The classical solution formula (1.29) is known as **D'Alembert formula**, after Jean-Baptiste le Rond d'Alembert who in 1747 in Paris both derived the one-dimensional wave equation and solved it: this is probably the first example of partial differential equation derived from physical principles in history.

We will see that the possibility of writing the classical solution as the sum of two terms, each depending only on one of the initial conditions, is a typical feature of the wave equation. These are often referred to as "cosine" and "sine" term, respectively. Sometimes one writes it as ´

$$u(t,x) := \big(C(t)u_0\big)(x) + \big(S(t)u_1\big)(x),$$

where $C(t)$ and $S(t)$, $t \geq 0$, are operators (i.e., function-valued functions, see Appendix C) – defined in this one-dimensional case by

$$\big(C(t)u_0\big)(x) := \frac{1}{2}\left(u_0(x+ct) + u_0(x-ct)\right), \qquad t \geq 0,\ x \in \mathbb{R},$$

and

$$\big(S(t)u_1\big)(x) := \frac{1}{2c}\int_{x-ct}^{x+ct} u_1(y)\mathrm{d}y, \qquad t \geq 0,\ x \in \mathbb{R}.$$

In other words, we can represent all classical solutions as the sum of classical solutions to the wave equation with $u_1 \equiv 0$ and general initial data u_0, on one hand; and on the other hand, with $u_0 \equiv 0$ and general initial data u_1: they are by $u(t,x) = \big(C(t)u_0\big)(x)$ and $u(t,x) = \big(S(t)u_1\big)(x)$, respectively.

Exercise 1.16. Show that $C(0)f \equiv f$, $S(0) = 0$, and moreover

$$2C(t)\left(C(s)f\right) \equiv \left(C(t+s)f\right) + \left(C(t-s)f\right)$$

as well as

$$\left(S(t+s)f\right) \equiv C(s)(S(t)f) + (S(s)C(t)f)$$

for all $t, s \geq 0$ and all functions $f \in C^1$. Why are C and S called "cosine" and "sine", respectively?

Exercise 1.17. Let $A, k, \omega \in \mathbb{R}$. Show that both functions u, v defined by

$$u(t,x) := A\sin(kx - \omega t) \qquad \text{and} \qquad v(t,x) := A\sin(kx + \omega t), \qquad t, x \in \mathbb{R}$$

are classical solutions to the one-dimensional wave equation on $\mathbb{R}$. They are called **traveling waves**. Due to linearity, their sum is also a classical solution of the wave equation. Show that, however $u + v$ is not a traveling wave, and that in fact it can be written as a product $\eta \cdot \xi$ of two functions, where η only depends on time and ξ only depends on space. Such a classical solution is called a **stationary wave**.

Exercise 1.18. One may try to apply to the wave equation the method of characteristics directly. In this case, one should introduce a change of variables setting $\lambda := x - ct$ and $\mu := x + ct$ and consider those curves along which the classical solutions to the wave equation are not constant (like in the case of the transport equation), but rather satisfy

$$\frac{\partial^2 u}{\partial\lambda\partial\mu} \equiv 0.$$

Work out the details.

Exercise 1.19. Discuss the wave-type partial differential equation

$$\frac{\partial^2 u}{\partial t^2}(t,x) = x^2\frac{\partial^2 u}{\partial x^2}(t,x) - x\frac{\partial u}{\partial x}(t,x), \qquad t \geq 0,\ x \in \mathbb{R}, \tag{1.30}$$

by performing a factorization similar to that crucial in order to solve the standard wave equation.

Remark 1.20. Observe that even if we were only looking for a *forward classical solution* of the wave equation, the classical solution we have derived is also a *backward* one, i.e., it is a function that solves the wave equation also for negative times: knowing the state of a wave at a certain time t_0, it is possible to reconstruct the state of the wave at any given previous time $t < t_0$.

Let us now turn to the issue of *uniqueness*. The starting point of our study is Theorem 1.15, where we have derived an explicit classical solution. Would it be possible to have further classical solutions, at least for some special initial data? No, it would not. This is the first application of the so-called *variational method*, which we will more extensively discuss later on, beginning in Chapter 3.

Theorem 1.21. *Let $\Omega \subset \mathbb{R}$ be a (possibly unbounded) open interval and c be a bounded continuous function with $c(x) > 0$ for all $x \in \Omega$. Let $u_0 \in C^2(\overline{\Omega})$ and $u_1 \in C^1(\overline{\Omega})$. Then the initial value problem for the (possibly inhomogeneous) one-dimensional wave equation* (1.25) *on Ω has at most one classical solution in the class of functions of class in $C^2(\mathbb{R}_+ \times \Omega)$ with (if Ω is unbounded) bounded and integrable partial derivatives up to order* 2*, provided that*

- *either $\Omega = \mathbb{R}$*
- *or the boundary condition is of Dirichlet or Neumann type (possibly inhomogeneous), if $\Omega \neq \mathbb{R}$ and hence it has a boundary.*

Proof. Assume u, v to be two classical solutions to the equation. Due to linearity, also $w := u - v$ is a classical solution of the one-dimensional wave equation (in particular, it fulfills the assumptions of Schwarz's Theorem) but satisfies the initial conditions

$$w(0,x) = 0, \qquad \frac{\partial w}{\partial t}(0,x) = 0, \qquad x \in \Omega.$$

Moreover, if $\Omega \neq \mathbb{R}$, then w also satisfies *homogeneous* Dirichlet or Neumann boundary conditions, since either the boundary values of u, v (if the Dirichlet boundary condition is imposed), or else their normal derivatives (if the Neumann boundary condition is imposed) agree. We introduce a function $E : \mathbb{R}_+ \to \mathbb{R}$ by

$$E(t) := \frac{1}{2}\int_\Omega \left(\left| \frac{\partial w}{\partial t}(t,x) \right|^2 + \left| c(x) \frac{\partial w}{\partial x}(t,x) \right|^2 \right) \mathrm{d}x;$$

observe that these integrals converge, by our assumptions.

Differentiating this expression and integrating by parts we obtain

$$\begin{aligned}
\frac{\mathrm{d}E}{\mathrm{d}t}(t) &= \int_\Omega \frac{\partial w}{\partial t}(t,x)\frac{\partial^2 w}{\partial t^2}(t,x)\mathrm{d}x + \int_\Omega c^2(x)\frac{\partial w}{\partial x}(t,x)\frac{\partial^2 w}{\partial x \partial t}(t,x)\mathrm{d}x \\
&= \int_\Omega \frac{\partial w}{\partial t}(t,x)\frac{\partial^2 w}{\partial t^2}(t,x)\mathrm{d}x + \left[c^2(x)\frac{\partial w}{\partial x}(t,x)\frac{\partial w}{\partial t}(t,x) \right]\Bigg|_{x=0}^{x=1} - \int_\Omega c^2(x)\frac{\partial^2 w}{\partial x^2}(t,x)\frac{\partial w}{\partial t}(t,x)\mathrm{d}x \\
&= \int_\Omega \frac{\partial w}{\partial t}(t,x)\left(\frac{\partial^2 w}{\partial t^2}(t,x) - c^2(x)\frac{\partial^2 w}{\partial x^2}(t,x) \right)\mathrm{d}x + \left[c^2(x)\frac{\partial w}{\partial x}(t,x)\frac{\partial w}{\partial t}(t,x) \right]\Bigg|_{x=0}^{x=1} \\
&= 0.
\end{aligned}$$

The last identity holds because, as observed above, the boundary term vanishes identically at the boundary. We conclude that E is a constant function, i.e., $E(t) = E(0) = 0$, and therefore

$$\left| \frac{\partial w}{\partial t}(t,x) \right| = \left| \frac{\partial w}{\partial x}(t,x) \right| = 0 \qquad \text{for all } t \geq 0 \text{ and all } x \in \Omega.$$

In other words, w is a smooth function with vanishing gradient, hence w is constant in time and in space. Because w vanishes constantly at the boundary and also everywhere for $t = 0$, we deduce that $w \equiv 0$ everywhere and at any time. This concludes the proof. □

The above proof motivates the introduction of the following.

Definition 1.22. The **total energy** of a classical solution u of the homogeneous wave equation (1.32) on an open domain $\Omega \subset \mathbb{R}^d$ is given at any time t by the sum

$$E := E_p + E_k$$

of its **potential energy** [2]

$$E_p(t) := \int_\Omega |\nabla u(t,x)|^2 \mathrm{d}x$$

and its **kinetic energy**

$$E_k(t) := \int_\Omega \left| c(x)\frac{\partial u}{\partial t}(t,x) \right|^2 \mathrm{d}x.$$

[2] We denote by ∇u the gradient of u with respect to its variable x. Accordingly, for each t $x \mapsto \nabla u(t,x)$ is a vector field on $\mathbb{R}^d$ and in particular $\nabla u(t,x) \in \mathbb{R}^d$ for all t: here and in the following we denote by $\nabla u(t,x)$ its Euclidean norm, and more generally by $|\xi|$ the Euclidean norm of any vector ξ in a finite-dimensional space.

Euclidean norm of $\mathbb{R}^m$

We conclude this section by discussing an extension of the D'Alembert formula to the case of the semi-bounded domain $(0,\infty)$. This explains the method of reflections that is typical for the wave equation. It will be used in the next chapter. For the sake of simplicity we consider only the case of unit wave speed, the general one being analogous.

Theorem 1.23. *Let $c = 1$. For $u_0 \in C^2([0,\infty))$ and $u_1 \in C^1([0,\infty))$ the classical solution of the wave equation for $(t,x) \in \mathbb{R}_+ \times (0,\infty)$ with Dirichlet boundary condition*

$$u(t,0) = 0, \qquad t \geq 0,$$

is given by

$$u(t,x) := \begin{cases} \frac{1}{2}\left(u_0(x+t) + u_0(x-t)\right) + \frac{1}{2}\int_{x-t}^{x+t} u_1(y)\mathrm{d}y, & x \geq t \geq 0, \\ \frac{1}{2}\left(u_0(x+t) - u_0(t-x)\right) + \frac{1}{2}\int_{t-x}^{x+t} u_1(y)\mathrm{d}y, & t \geq x \geq 0. \end{cases}$$

Observe that the above formula yields a function u that is twice continuously differentiable everywhere apart from $t = x$. However, if one is interested in the classical notion of classical solution, continuous differentiability in $t = x$ becomes an issue and we have to impose the Wentzell boundary condition on u_0 at 0. In particular, the second derivative (both with respect to space and to time) of u has a jump equal to $2u_0''(0)$.

Proof. Since we have a classical solution formula at hand for the wave equation on $\mathbb{R}$, it is natural to try to extend the wave equation on $(0,\infty)$ to the whole line. This is simply done by introducing the odd extension of u

$$\tilde{u}(t,x) := \begin{cases} u(t,x), & t \geq 0,\ x \geq 0, \\ -u(t,-x), & t \geq 0,\ x \leq 0, \end{cases}$$

i.e.,$\tilde{u}(t,\cdot)$ is for all $t \geq 0$ the unique odd function that agrees with u on $\mathbb{R}$. Similarly, we introduce the odd functions

$$\tilde{u}_0(x) := \begin{cases} u_0(x), & x \geq 0, \\ -u_0(-x), & x \leq 0, \end{cases}$$

and

$$\tilde{u}_1(x) := \begin{cases} u_1(x), & x \geq 0, \\ -u_1(-x), & x \leq 0. \end{cases}$$

Let us prove that if u is a classical solution of the one-dimensional wave equation on $(0,\infty)$ with Dirichlet boundary condition, then $\tilde{u}$ is a classical solution of the one-dimensional wave equation on $\mathbb{R}$ (clearly, *without* boundary conditions). This is due to the fact that $\tilde{u}$ satisfies the differential equation on $(0,\infty)$, where it agrees with u, and that moreover for all $\tilde{x} := -x$, $x \geq 0$, there holds

$$\begin{aligned} \frac{\partial^2 \tilde{u}}{\partial t^2}(t,\tilde{x}) &= -\frac{\partial^2 u}{\partial t^2}(t,x) \\ &= -\frac{\partial^2 u}{\partial x^2}(t,x) \\ &= \frac{\partial^2 u}{\partial x^2}(t,\tilde{x})\frac{\partial \tilde{x}}{\partial x} \\ &= \frac{\partial^2 u}{\partial \tilde{x}^2}(t,\tilde{x}), \end{aligned}$$

i.e., $\tilde{u}$ solves the wave equation on the whole real line. Moreover

$$u(0,\tilde{x}) = -u(0,x) = -u_0(x) = -\tilde{u}_0(\tilde{x})$$

and likewise

$$\frac{\partial u}{\partial t}(0,\tilde{x}) = -u(0,x) = -u_1(x) = -\tilde{u}_1(\tilde{x}).$$

Thus, $\tilde{u}$ is given by the D'Alembert formula, from where the claimed identities hold. □

A relevant consequence of the above D'Alembert formula is that even for positive-valued u_0 and $u_1 \equiv 0$ the classical solution needs not be positive for all $t \geq 0$. This is in sharp contrast to the behavior of the transport equation – and, as we will see in Chapter 2, of further important evolution equations, too.

Exercise 1.24. A similar strategy can also be applied to find the classical solution of the wave equation on a bounded interval, say, on $(0, 1)$, with Dirichlet boundary condition

$$u(t, 0) = u(t, 1) = 0, \qquad t \geq 0.$$

The idea is that the initial conditions have to be reflected *and* extended by periodicity. Work out the details.

Exercise 1.25. While odd extensions are useful to solve the one-dimensional wave equation with the Dirichlet boundary condition, even extensions come into play when the Neumann boundary condition

$$u'(t, 0) = u'(t, 1) = 0, \qquad t \geq 0,$$

is considered. Deduce an explicit formula for the classical solution of this wave equation. Is the classical solution positive-valued, provided the initial data are?

What about the wave equation on bounded one-dimensional domains? Formula (1.29) shows that there are *two* waves traveling to the left and to the right, respectively. Intuitively, the traveling waves induced by the initial data can only spread as described until they impact on the boundary. At that point, one may expect that boundary conditions are needed to describe the further behavior of classical solutions. While this is actually true, we have already seen in Section 1.1 how the issue of boundary conditions might be delicate when dealing with the method of characteristics: roughly speaking, boundary conditions need to be imposed where the wave *leaves* the boundary, rather than when it reaches it. Because however we have two coexisting waves, and one boundary condition per wave is needed, two boundary conditions are actually needed in the case of one-dimensional bounded intervals: this is consistent with the explicit classical solution formulae analogous to (1.29) that can be derived for Dirichlet and Neumann conditions (see Exercises 1.24 and 1.25) based on suitably extending the initial classical solutions to the whole real line. A more general theory of wave-type equations on (possibly higher dimensional) bounded domains that does *not* rely upon explicit solutions and can be hence used if the Laplacian is replaced by more general elliptic operators is however somewhat involved: we will come back to this issue in Section 4.3.

1.3 The wave equation in higher dimension

We have just seen how the one-dimensional wave equation can be explicitly solved on the space domain $\mathbb{R}$ and, under suitable boundary conditions, even on other domains displaying less symmetry. In this section we are turning to the natural issue of studying the higher dimensional version (1.26) of the wave equation, which can be derived in the same ways and describes elastic light systems undergoing small oscillations. While the one-dimensional wave equation can be used to model the vibration of strings, its two dimensional version can – once equipped with appropriate boundary conditions – describe the approximate behavior of musical instruments like gongs and drums (with Neumann and Dirichlet boundary conditions, respectively): here and in the following we will consider the natural generalizations of the boundary conditions already introduced for functions of one real variable, namely

$$u(t, z) \equiv 0, \qquad t \geq 0,\ z \in \partial\Omega \qquad \text{(Dirichlet b.c.);}$$

$$\frac{\partial u}{\partial \nu}(t, z) \equiv 0, \qquad t \geq 0,\ z \in \partial\Omega \qquad \text{(Neumann b.c.),}$$

where $\frac{\partial u}{\partial \nu}$ is the normal derivative of u, see the Appendix B;

$$-\frac{\partial u}{\partial \nu}(t, z) = p(z)u(t, z), \qquad t \geq 0,\ z \in \partial\Omega \qquad \text{(Robin b.c.),}$$

for some $p : \partial\Omega \to \mathbb{R}$ with $p(z) > 0$ for all $z \in \partial\Omega$; or

$$\Delta u(t, z) \equiv 0, \qquad t \geq 0,\ z \in \partial\Omega \qquad \text{(Wentzell b.c.).}$$

Three-dimensional wave equations, in turn, describe for example the propagation of sound through light compressible media, like air. In order to obtain an explicit formula for the classical solutions to such wave equations, our strategy will be the following:

- We show that certain spherical means w of any classical solution u to the d-dimensional (d odd!) wave equation solve a modified one-dimensional wave equation (Lemma 1.30.
- While w only solves the Euler–Poisson–Darboux equation, one can rescale it and introduce a function U that is a classical solution of the one-dimensional wave equation (Lemma 1.33).
- Because the explicit classical solution of the one-dimensional wave equation is given (locally) by the D'Alembert formula, we deduce a formula for U that, in turn, yields a formula for u in terms of spherical means of the initial data (Theorem 1.34).
- The case of d even is solved by the so-called *method of descent*: any classical solution u to the $(d-1)$-dimensional wave equation (d even) is embedded in a classical solution $\tilde{u}$ to the d-dimensional wave equation in a suitable way, and then the above mentioned results yield a formula for $\tilde{u}$ and thus, in turn, for u. The details are carried out for the two-dimensional case only (Theorem 1.36).

In the above strategy there is a hidden assumption, namely that each initial value problem for the wave equation admits at most one classical solution (hence exactly one in the special cases where an explicit formula is known): in the one-dimensional case this assumption is justified by Theorem 1.21, but an analogous assertion also holds in the general d-dimensional case. In the following, we typically assume the relevant domain Ω to have Lipschitz boundary to guarantee that the Gauss–Green Formulae from Appendix B hold; this is especially the case if Ω has C^1-boundary.

Theorem 1.26. *Let $\Omega \subset \mathbb{R}^d$ be a (possibly unbounded) Lipschitz domain. Let $u_0 \in C^2(\overline{\Omega})$ and $u_1 \in C^1(\overline{\Omega})$. Let also*

- *either $\Omega = \mathbb{R}^d$,*
- *or the boundary condition be of Dirichlet or Neumann type (possibly inhomogeneous), if Ω has nonempty boundary.*

Then the initial value problem for the wave equation (1.26) *has at most one classical solution in the class of those classical solutions that are uniformly bounded on compact domains.*

Exercise 1.27. Prove Theorem 1.26.

In order to solve the wave equation in $\mathbb{R}^d$ for $d \geq 2$ we first discuss a modification of the one-dimensional wave equation. Here and in the following we denote by $B_r(x)$ the d-dimensional ball of radius r centered at x, by $\partial B_r(x)$ its surface (i.e., its topological boundary), and by $|B_r| := |B_r(x)|$ and $|\partial B_r| := |\partial B_r(x)|$ their respective measures, i.e.,

$$|B_r| = \frac{\pi^{\frac{d}{2}}}{\Gamma\left(\frac{d}{2}+1\right)} r^d \qquad \text{and} \qquad |\partial B_r| = \frac{d\pi^{\frac{d}{2}}}{\Gamma\left(\frac{d}{2}+1\right)} r^{d-1}, \qquad r > 0,$$

where Γ denotes the gamma-function, see [9, Kapitel 5].

Here and in the following, $\mathrm{d}\sigma$ denotes the $(d-1)$-dimensional Hausdorff (surface) measure. One can regard the above formulae for $|B_r|$ and $|\partial B_r|$ as functions of the variable r: an important feature of these functions is that their derivatives have nice interplays. More precisely,

$$\frac{\mathrm{d}}{\mathrm{d}r}|B_r| = |\partial B_r| \qquad \text{and} \qquad \frac{\mathrm{d}}{\mathrm{d}r}\frac{1}{|\partial B_r|} = \frac{1-d}{d}\frac{1}{|B_r|}, \qquad r > 0. \tag{1.31}$$

(We stress that d on the right hand side of the latter equation stands for the dimension of the Euclidean space, whereas d denotes as usual the differential.)

Lemma 1.28. *Let $u \in C^2(\mathbb{R}^d)$. Then for all $x \in \mathbb{R}^d$ the function $\phi : \mathbb{R} \to \mathbb{R}$ defined by*

$$\phi(r) := \frac{1}{|\partial B_r|} \int_{\partial B_r(x)} u(y)\, \mathrm{d}\sigma(y), \qquad r \in \mathbb{R},$$

is continuously differentiable and

$$\frac{\partial \phi}{\partial r}(r) = \frac{r}{d}\frac{1}{|B_r|}\int_{B_r(x)} \Delta u(y)\mathrm{d}y.$$

Proof. Substituting $y \mapsto \tilde{y} := x + rz$ yields

$$\phi(r) = \frac{1}{|\partial B_1|}\int_{\partial B_1(0)} u(x+rz)\,\mathrm{d}\sigma(z).$$

Now, the chain rule yields

$$\begin{aligned}
\phi'(r) &= \frac{1}{|\partial B_1|}\int_{\partial B_1(0)} \nabla u(x+rz)\cdot z\,\mathrm{d}\sigma(z)\\
&= \frac{1}{|\partial B_r|}\int_{\partial B_r(x)} \nabla u(y)\cdot\frac{y-x}{r}\,\mathrm{d}\sigma(y)\\
&= \frac{1}{|\partial B_r|}\int_{\partial B_r(x)} \frac{\partial u}{\partial \nu}(y)\,\mathrm{d}\sigma(y)\\
&= \frac{r}{d}\frac{1}{|B_r|}\int_{B_r(x)} \Delta u(y)\mathrm{d}y,
\end{aligned}$$

where the last identity follows from the second Gauss–Green formula (B.5). This completes the proof. □

Exercise 1.29. Let $u \in C^\infty(\mathbb{R}^d)$ and $x \in \mathbb{R}^d$. Determine

- $\frac{\partial}{\partial r}\left(\frac{1}{|B_r|}\int_{\partial B_r(x)} u(y)\,\mathrm{d}\sigma(y)\right)$,
- $\frac{\partial}{\partial r}\left(\frac{1}{|\partial B_r|}\int_{B_r(x)} u(y)\mathrm{d}y\right)$, and
- $\frac{\partial}{\partial r}\left(\frac{1}{|B_r|}\int_{B_r(x)} u(y)\mathrm{d}y\right)$,

for $r > 0$. Are you able to find a general formula for

$$\frac{\partial^k}{\partial r^k}\phi(r), \qquad r > 0\,?$$

Lemma 1.30. *Let $u_0 \in C^2(\mathbb{R}^d)$ and $u_1 \in C^1(\mathbb{R}^d)$. Let $u \in C^m([0,\infty)\times\mathbb{R}^d)$ be a classical solution of the d-dimensional wave equation*

$$\frac{\partial^2 u}{\partial t^2}(t,x) = c^2(x)\Delta u(t,x), \qquad t \geq 0,\ x \in \mathbb{R}^d, \tag{1.32}$$

with initial conditions

$$u(0,x) = u_0(x) \qquad \text{and} \qquad \frac{\partial u}{\partial t}(0,x) = u_1(x), \qquad x \in \mathbb{R}^d.$$

Then, for any $x \in \mathbb{R}^d$ the function w defined by

$$w(t,r,x) := \frac{1}{|\partial B_r|}\int_{\partial B_r(x)} u(t,y)\ \mathrm{d}\sigma(y), \qquad t \geq 0,\ r > 0, \tag{1.33}$$

is a classical solution of the one-dimensional wave equation

$$\frac{\partial^2 w}{\partial t^2}(t,r,x) = \frac{\partial^2 w}{\partial r^2}(t,r,x) + \frac{d-1}{r}\frac{\partial w}{\partial r}(t,r,x), \qquad t \geq 0,\ r \geq 0, \tag{1.34}$$

with initial data defined for all $x \in \mathbb{R}$ by

$$w(0,r,x) = w_0(r,x) := \frac{1}{|\partial B_r|}\int_{\partial B_r(x)} u_0(y)\,\mathrm{d}\sigma(y), \qquad r > 0,$$

and

$$\frac{\partial w}{\partial t}(0,r,x) = w_1(r,x) := \frac{1}{|\partial B_r|}\int_{\partial B_r(x)} u_1(y)\,\mathrm{d}\sigma(y), \qquad r > 0.$$

Moreover, $w(\cdot,\cdot,x) \in C^m([0,\infty)\times[0,\infty))$ for all $x \in \mathbb{R}^d$.

The partial differential equation (1.34) – which is, more precisely, a family of partial differential equations indexed in x – is sometimes called **Euler–Poisson–Darboux equation**, after Siméon Poisson, Leonhard Euler, and especially Gaston Darboux, who in the first half of the 1910s in Paris provided an extensive treatment of the theory of this equation.

Proof. The definition of w does not modify the dependence on t of u, and we conclude that w is m-times continuously differentiable in its first variable, provided that u is.

Now, apply Lemma 1.28 and observe that for all $r > 0$

$$\frac{\partial w}{\partial r}(t,r,x) = \frac{\partial}{\partial r}\frac{1}{|\partial B_r|}\int_{\partial B_r(x)} u(t,y)\,\mathrm{d}\sigma(y) = \frac{r}{d}\frac{1}{|B_r|}\int_{B_r(x)}\Delta u(t,y)\mathrm{d}y = \frac{1}{|\partial B_r|}\int_{B_r(x)}\Delta u(t,y)\mathrm{d}y. \tag{1.35}$$

Hence, $w(\cdot,\cdot,x) \in C^1([0,\infty)\times(0,\infty))$. Passing to the limit for $r \to 0$, by Lebesgue's Differentiation Theorem we obtain

$$\lim_{r\to 0}\frac{\partial w}{\partial r}(t,r,x) = 0:$$

thus, $w(\cdot,\cdot,x) \in C^1([0,\infty)\times[0,\infty))$. One can further differentiate and apply (1.35) and (1.31) to obtain

$$\begin{aligned}
\frac{\partial^2 w}{\partial r^2}(t,r,x) &= \frac{\partial^2}{\partial r^2}\frac{1}{|\partial B_r|}\int_{\partial B_r(x)} u(t,y)\,\mathrm{d}\sigma(y)\\
&= \frac{\partial}{\partial r}\left(\frac{1}{|\partial B_r|}\int_{B_r(x)}\Delta u(t,y)\mathrm{d}y\right)\\
&= \frac{1-d}{d}\frac{1}{|B_r|}\int_{B_r(x)}\Delta u(t,y)\mathrm{d}y + \frac{1}{|\partial B_r|}\frac{\partial}{\partial r}\int_0^r\int_{\partial B_s(x)}\Delta u(t,y)\mathrm{d}\sigma(y)\\
&= \frac{1-d}{d}\frac{1}{|B_r|}\int_{B_r(x)}\Delta u(t,y)\mathrm{d}y + \frac{1}{|\partial B_r|}\int_{\partial B_r(x)}\Delta u(t,y)\mathrm{d}\sigma(y).
\end{aligned}$$

Therefore, using (1.35) and the fact that u is a classical solution of the wave equation we deduce

$$\begin{aligned}
\frac{\partial^2 w}{\partial r^2}(t,r,x) &= -\frac{d-1}{d}\frac{d}{r}\frac{1}{|\partial B_r|}\int_{B_r(x)}\Delta u(t,y)\mathrm{d}y + \frac{1}{|\partial B_r|}\int_{\partial B_r(x)}\frac{\partial u}{\partial t^2}(t,y)\,\mathrm{d}\sigma(y)\\
&= -\frac{d-1}{r}\frac{\partial w}{\partial r}(t,r,x) + \frac{\partial^2 w}{\partial t^2}(t,r,x).
\end{aligned}$$

Thus, w is actually a classical solution of the Euler–Poisson–Darboux equation (1.34).

It remains to check that $\frac{\partial^2 w}{\partial r^2}$ is continuous up to the boundary, i.e., up to $r = 0$. Just like before, passing to the limit for $r \to 0$ yields

$$\lim_{r\to 0}\frac{\partial^2 w}{\partial r^2}(t,r,x) = \Delta u(t,x) + \left(\frac{1}{d} - 1\right)\Delta u(t,x) = \frac{1}{d}\Delta u(t,x).$$

One can further differentiate and obtain recursively that in the radial direction w is as regular as u. □

Remark 1.31. Elaborating on the ideas of Remark 1.40 below, one can show that the Euler–Poisson–Darboux equation (1.34) is a wave equation whose energy increases in time, since the damping term $\frac{d-1}{r}\frac{\partial w}{\partial r}$ has a positive-valued coefficient. This rather peculiar behavior suggests that (1.34) does not have an obvious physical meaning. It probably gives more insight to observe that the term $\frac{\partial^2 w}{\partial r^2} + \frac{d-1}{r}\frac{\partial w}{\partial r}$ that appears in the Euler–Poisson–Darboux equation agrees with the radial component of Δw, whenever we write it in spherical coordinates.

In the following we are going to transform the odd- and even-dimensional wave equation into the one-dimensional wave equation, via the formalism introduced to study the Euler–Poisson–Darboux equation. Since a classical solution of the one-dimensional wave equation is known, we will be able to find explicit formulae for general classical solutions in these higher-dimensional cases.

To begin with, we consider the case of d odd, i.e., $d = 2k+1$. In the following, $n!!$ denotes the **double factorial** of a positive integer n, i.e.,

$$n!! := \begin{cases} n\cdot(n-2),\cdot\ldots 2 & \text{if } n \text{ is even,}\\ n\cdot(n-2),\cdot\ldots 1 & \text{if } n \text{ is odd.}\end{cases}$$

Lemma 1.32. *Let $k \in \mathbb{N}$. Then the following identities hold for all $\phi \in C^{k+1}(\mathbb{R})$.*

(1) $\left(\frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}r}\right)^{k-1}\left(r^{2k-1}\phi(r)\right) = (2k-1)!!\; r\phi(r) + \sum_{j=1}^{k-1}\beta_j^k r^{j+1}\frac{\mathrm{d}^j\phi}{\mathrm{d}r^j}(r)$, *for some constants β_j^k independent on ϕ.*

(2) $\frac{d^2}{dr^2}\left(\frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}r}\right)^{k-1}\left(r^{2k-1}\phi(r)\right) = \left(\frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}r}\right)^{k}\left(r^{2k}\frac{\mathrm{d}\phi}{\mathrm{d}r}(r)\right)$.

Here

$$\left(\frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}r}\right)^0 f = f, \quad \left(\frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}r}\right)^1 f = \frac{1}{r}f', \quad \left(\frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}r}\right)^2 f = \frac{1}{r}\left(\frac{1}{r}f'\right)' = -\frac{1}{r^3}f' + \frac{1}{r^2}f''$$

and so on.

Proof. Both identities are proven by induction on k.

(1) The assertion is clearly true for $k = 1$, since both left and right hand sides agree with $r\phi(r)$. If the assertion holds for k, then we see that

$$\begin{aligned}\left(\frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}r}\right)^k\left(r^{2k+1}\phi(r)\right) &= \left(\frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}r}\right)^{k-1}\left(\frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}r}\right)\left(r^{2k+1}\phi(r)\right)\\ &= \left(\frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}r}\right)^{k-1}\left(r^{2k-1}\left((2k+1)\phi(r) + r\frac{\mathrm{d}\phi}{\mathrm{d}r}(r)\right)\right) \qquad \text{for all } r > 0.\\ &:= \left(\frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}r}\right)^{k-1}\left(r^{2k-1}\tilde{\phi}(r)\right)\end{aligned}$$

Applying the induction assumption to $\tilde{\phi}$ we obtain

$$\begin{aligned}\left(\frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}r}\right)^k\left(r^{2k+1}\phi(r)\right) &= (2k-1)!!\;\tilde{\phi}(r) + \sum_{j=1}^{k-1}\beta_j^k r^{j+1}\frac{\mathrm{d}^j}{\mathrm{d}r^j}\tilde{\phi}(r)\\ &= (2k+1)!!\; r\phi(r) + (2k-1)!! r^2\frac{\mathrm{d}\phi}{\mathrm{d}r}(r) + \qquad \text{for all } r > 0.\\ &\quad + \sum_{j=1}^{k-1}(2k+1)\beta_j^k r^{j+1}\frac{\mathrm{d}^j\phi}{\mathrm{d}r^j}(r) + \sum_{j=1}^{k-1}\beta_j^k r^{j+1}\frac{\mathrm{d}^j}{\mathrm{d}r^j}\left(r\frac{\mathrm{d}\phi}{\mathrm{d}r}(r)\right)\end{aligned}$$

Now, it is immediate to prove by induction (try!) that for any function $f \in C^j$ there holds

$$\frac{\mathrm{d}^j}{\mathrm{d}r^j}(rf(r)) = r\frac{d^j f}{dr^j}(r) + j\frac{\mathrm{d}^{j-1}f}{\mathrm{d}r^{j-1}}(r) \qquad \text{for all } j \in \mathbb{N} \text{ and all } r > 0.$$

Therefore, for $f = \frac{\partial\phi}{\partial r}$ we obtain

$$\begin{aligned}\left(\frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}r}\right)^k\left(r^{2k+1}\phi(r)\right) &= (2k+1)!!\; r\phi(r) + (2k-1)!! r^2\frac{\mathrm{d}\phi}{\mathrm{d}r}(r) +\\ &\quad + \sum_{j=1}^{k-1}(2k+1)\beta_j^k r^{j+1}\frac{\mathrm{d}^j\phi}{\mathrm{d}r^j}(r) + \sum_{j=1}^{k-1}\beta_j^k r^{j+1}\left(r\frac{\mathrm{d}^{j+1}\phi}{\mathrm{d}r^{j+1}}(r) + j\frac{d^j\phi}{dr^j}(r)\right)\\ &= (2k+1)!!\; r\phi(r) + (2k-1)!! r^2\frac{\mathrm{d}\phi}{\mathrm{d}r}(r) +\\ &\quad + \sum_{j=1}^{k-1}(2k+1)\beta_j^k r^{j+1}\frac{\mathrm{d}^j\phi}{\mathrm{d}r^j}(r) + \sum_{j=2}^{k}\beta_{j-1}^k r^{j+1}\frac{\mathrm{d}^j\phi}{\mathrm{d}r^j}(r) + \sum_{j=1}^{k-1} j\beta_j^k r^{j+1}\frac{d^j\phi}{dr^j}(r)\\ &= (2k+1)!!\; r\phi(r) + \sum_{j=1}^{k}\beta_j^{k+1} r^{j+1}\frac{\mathrm{d}^j\phi}{\mathrm{d}r^j}(r),\end{aligned}$$

where the explicit definition of the constants β_j^{k+1}, $k, j \in \mathbb{N}$, is left as an easy exercise.

(2) The assertion is true for k=1, because

$$\frac{d^2}{dr^2}\left(r\phi(r)\right) = \frac{\mathrm{d}}{\mathrm{d}r}\left(\phi(r) + r\frac{\mathrm{d}\phi}{\mathrm{d}r}(r)\right) = 2\frac{\mathrm{d}\phi}{\mathrm{d}r}(r) + r\frac{d^2\phi}{dr^2}(r) = \frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}r}\left(r^2\frac{\mathrm{d}\phi}{\mathrm{d}r}(r)\right), \qquad r \in \mathbb{R}.$$

Let now assume that the assertion holds for k. Then

$$\begin{aligned}
\frac{d^2}{dr^2}\left(\frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}r}\right)^k \left(r^{2k+1}\phi(r)\right) &= \frac{d^2}{dr^2}\left(\frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}r}\right)^{k-1}\left(\frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}r}\right)\left(r^{2k+1}\phi(r)\right) \\
&= \frac{d^2}{dr^2}\left(\frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}r}\right)^{k-1}\left((2k+1)r^{2k-1}\phi(r) + r^{2k}\frac{\mathrm{d}\phi}{\mathrm{d}r}(r)\right) \qquad \text{for all } r > 0. \\
&= \dots \\
&= \left(\frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}r}\right)^{k+1}\left(r^{2k+2}\frac{\mathrm{d}\phi}{\mathrm{d}r}(r)\right)
\end{aligned}$$

The missing step is left as an exercise to the reader. □

Lemma 1.33. *If u is a classical solution of the initial value problem*

$$\begin{cases}
\dfrac{\partial^2 u}{\partial t^2}(t,x) = \Delta u(t,x), & t \geq 0,\ x \in \mathbb{R}^{2k+1}, \\
u(0,x) = u_0(x), & x \in \mathbb{R}^{2k+1}, \\
\dfrac{\partial u}{\partial x}(0,x) = u_1(x), & x \in \mathbb{R}^{2k+1},
\end{cases} \tag{1.36}$$

then for all $x \in \mathbb{R}^{2k+1}$ the function U defined by

$$\begin{aligned}
U(t,r,x) &:= \left(\frac{1}{r}\frac{\partial}{\partial r}\right)^{k-1}\left(r^{2k-1}w(t,r,x)\right) \\
&= \left(\frac{1}{r}\frac{\partial}{\partial r}\right)^{k-1}\left(\frac{r^{2k-1}}{|\partial B_r|}\int_{\partial B_r(x)} u(t,y)\,\mathrm{d}\sigma(y)\right), \qquad t \geq 0,\ r > 0,
\end{aligned} \tag{1.37}$$

(where w is the function in (1.33)*) is a classical solution of the homogeneous one-dimensional wave equation* (1.28) *(for $c = 1$) on $\Omega = (0,\infty)$ with Dirichlet boundary condition*

$$U(t,0,x) = 0, \qquad t \geq 0,$$

and with initial data

$$U(0,r,x) = G(r,x), \qquad \frac{\partial U}{\partial t}(0,r,x) = H(r,x), \qquad r \in (0,\infty),$$

where

$$\begin{aligned}
G(r,x) &:= \left(\frac{1}{r}\frac{\partial}{\partial r}\right)^{k-1}\left(r^{2k-1}w_0(r,x)\right) \\
&= \left(\frac{1}{r}\frac{\partial}{\partial r}\right)^{k-1}\left(\frac{r^{2k-1}}{|\partial B_r|}\int_{\partial B_r(x)} u_0(r,y)\,\mathrm{d}\sigma(y)\right)
\end{aligned}$$

and

$$\begin{aligned}
H(r,x) &:= \left(\frac{1}{r}\frac{\partial}{\partial r}\right)^{k-1}\left(r^{2k-1}w_1(r,x)\right) \\
&= \left(\frac{1}{r}\frac{\partial}{\partial r}\right)^{k-1}\left(\frac{r^{2k-1}}{|\partial B_r|}\int_{\partial B_r(x)} u_1(y)\,\mathrm{d}\sigma(y)\right).
\end{aligned}$$

Proof. For fixed $x \in \mathbb{R}^{2k+1}$, let $r > 0$ and $t \geq 0$. Then it is possible to apply Lemma 1.32.(2) and obtain

$$\begin{aligned}
\frac{\partial^2 U}{\partial r^2}(t,r,x) &= \frac{\partial^2}{\partial r^2}\left(\frac{1}{r}\frac{\partial}{\partial r}\right)^{k-1}\left(r^{2k-1}w(t,r,x)\right)\\
&= \left(\frac{1}{r}\frac{\partial}{\partial r}\right)^{k}\left(r^{2k}\frac{\partial w}{\partial r}(t,r,x)\right)\\
&= \left(\frac{1}{r}\frac{\partial}{\partial r}\right)^{k-1}\left(\frac{1}{r}\frac{\partial}{\partial r}\right)\left(r^{2k}\frac{\partial w}{\partial r}(t,r,x)\right)\\
&= \left(\frac{1}{r}\frac{\partial}{\partial r}\right)^{k-1}\left(r^{2k-1}\frac{\partial^2 w}{\partial r^2}(t,r,x) + 2kr^{2k-2}\frac{\partial w}{\partial r}(t,r,x)\right)\\
&= \left(\frac{1}{r}\frac{\partial}{\partial r}\right)^{k-1}\left(r^{2k-1}\left(\frac{\partial^2 w}{\partial r^2}(t,r,x) + \frac{2k}{r}\frac{\partial w}{\partial r}(t,r,x)\right)\right)\\
&= \left(\frac{1}{r}\frac{\partial}{\partial r}\right)^{k-1}\left(r^{2k-1}\frac{\partial^2 w}{\partial t^2}(t,r,x)\right)\\
&= \frac{\partial^2 U}{\partial t^2}(t,r,x),
\end{aligned}$$

where of course the last step is due to the fact that w solves the Euler–Poisson–Darboux equation.

Finally, the fact that U satisfies the initial conditions is a direct consequence of the fact that w satisfies the initial conditions of the Euler–Poisson–Darboux equation, which in turn is just an application of the definition of w. □

We are finally in the position to solve the initial value problem associated with the odd-dimensional wave equation, in a suitable sense.

If for some $k \in \mathbb{N}$ u is a classical solution of the $(2k+1)$-dimensional wave equation on the whole Euclidean space $\mathbb{R}^{2k+1}$, then U defined in (1.37) is a classical solution of the one-dimensional one on $(0,\infty)$. Thus, by the D'Alembert formula (suitably localized in time), we deduce that U satisfies

$$U(t,r,x) = \frac{1}{2}\left(G(t+r,x) - G(t-r,x)\right) + \frac{1}{2}\int_{t-r}^{t+r} H(y,x)\mathrm{d}y \qquad \text{for all } t \geq 0,\ r \in [0,t].$$

It remains to re-write this expression in terms of u, u_0 and u_1. To this aim, we want to consider the above formula in the limit $r \to 0+$. Then, by Lebesgue's Differentiation Theorem A.27, Lemma 1.32.(1) yields

$$\begin{aligned}
u(t,x) &= \lim_{r\to 0+} w(t,r,x)\\
&= \lim_{r\to 0+} w(t,r,x) + \lim_{r\to 0+}\sum_{j=1}^{k-1}\frac{\beta_j^k}{(2k-1)!!}r^j\frac{\partial^j w}{\partial r^j}(t,r,x)\\
&= \lim_{r\to 0+}\frac{1}{(2k-1)!!\,r}\left(\frac{1}{r}\frac{\partial}{\partial r}\right)^{k-1}\left(r^{2k-1}w(t,r,x)\right)\\
&= \lim_{r\to 0+}\frac{U(t,r,x)}{(2k-1)!!\,r}\\
&= \frac{1}{(2k-1)!!}\lim_{r\to 0+}\left(\frac{G(t+r,x)-G(t-r,x)}{2r} + \frac{1}{2r}\int_{t-r}^{t+r} H(y,x)\mathrm{d}y\right)\\
&= \frac{1}{(2k-1)!!}\left(\frac{\partial G}{\partial t}G(t,x) + H(t,x)\right)
\end{aligned} \qquad \text{for all } t \geq 0,\ x \in \mathbb{R}^{2k+1}.$$

Thus, if the $(2k+1)$-dimensional wave equation has a classical solution, it must have the above form. To check that u defined above is actually a classical solution of the equation, we proceed as follows.

Theorem 1.34. *Let $k \in \mathbb{N}$. For $u_0 \in C^{k+2}(\mathbb{R}^{2k+1})$ and $u_1 \in C^{k+1}(\mathbb{R}^{2k+1})$, define a function u by*

$$u(t,x) := \frac{1}{(2k-1)!!}\frac{\partial}{\partial t}\left(\frac{1}{t}\frac{\partial}{\partial t}\right)^{k-1}\left(\frac{t^{2k-1}}{|\partial B_t|}\int_{\partial B_t(x)} u_0(z)\,\mathrm{d}\sigma(z)\right) + \frac{1}{(2k-1)!!}\left(\frac{1}{t}\frac{\partial}{\partial t}\right)^{k-1}\left(\frac{t^{2k-1}}{|\partial B_t|}\int_{\partial B_t(x)} u_1(z)\,\mathrm{d}\sigma(z)\right), \qquad t>0,\ x\in\mathbb{R}^{2k+1}.$$

Then $u \in C^{2,2}(\mathbb{R}_+ \times \mathbb{R}^{2k+1})$. Moreover, u is a classical solution of the $(2k+1)$-dimensional wave equation away from $t=0$ (where u is not defined). However, $\lim_{t\to 0+} u(t,x_0)$ and $\lim_{t\to 0+}\frac{\partial u}{\partial t}(t,x_0)$ exist for all $x_0 \in \mathbb{R}^{2k+1}$, and in fact

$$\lim_{(t,x)\to(0,x_0)} u(t,x) = u_0(x_0) \qquad \text{and} \qquad \lim_{(t,x)\to(0,x_0)} \frac{\partial u}{\partial t}(t,x) = u_1(x_0) \qquad \text{for all } x_0 \in \mathbb{R}^{2k+1}.$$

Proof. We exploit linearity of the (homogeneous) wave equation: in fact, we check separately the two cases

- $u_0 \equiv 0$ and u_1 arbitrary,
- u_0 arbitrary and $u_1 \equiv 0$.

In the former case we obtain that

$$\begin{aligned}
\frac{\partial^2 u}{\partial t^2}(t,x) &= \frac{1}{(2k-1)!!}\frac{\partial^2}{\partial t^2}\left(\frac{1}{t}\frac{\partial}{\partial t}\right)^{k-1}\left(\frac{t^{2k-1}}{|\partial B_t|}\int_{\partial B_t(x)} u_1(z)\,\mathrm{d}\sigma(z)\right)\\
&= \frac{1}{(2k-1)!!}\left(\frac{1}{t}\frac{\partial}{\partial t}\right)^{k}\left(t^{2k}\frac{\partial}{\partial t}\frac{1}{|\partial B_t|}\int_{\partial B_t(x)} u_1(z)\,\mathrm{d}\sigma(z)\right)\\
&= \frac{1}{(2k-1)!!}\left(\frac{1}{t}\frac{\partial}{\partial t}\right)^{k}\left(\frac{t^{2k}}{|\partial B_t|}\int_{B_t(x)} \Delta u_1(y)\mathrm{d}y\right)
\end{aligned}$$

where the second-last equality follows from Lemma 1.32.(2) and the last one from Lemma 1.28. Observe that

$$\alpha_k := \frac{t^{2k}}{|\partial B_t|}$$

only depends on k, but not on t, hence

$$\begin{aligned}
\frac{\partial^2 u}{\partial t^2}(t,x) &= \frac{\alpha_k}{(2k-1)!!}\left(\frac{1}{t}\frac{\partial}{\partial t}\right)^{k}\left(\alpha_k\int_{B_t(x)} \Delta u_1(y)\mathrm{d}y\right)\\
&= \frac{1}{(2k-1)!!}\left(\frac{1}{t}\frac{\partial}{\partial t}\right)^{k-1}\left(\frac{\alpha_k}{t}\frac{\partial}{\partial t}\int_{B_t(x)} \Delta u_1(y)\mathrm{d}y\right)\\
&= \frac{1}{(2k-1)!!}\left(\frac{1}{t}\frac{\partial}{\partial t}\right)^{k-1}\left(\frac{\alpha_k}{t}\frac{\partial}{\partial t}\int_0^t\int_{\partial B_s(x)} \Delta u_1(y)\,\mathrm{d}\sigma(y)\mathrm{d}s\right)\\
&= \frac{1}{(2k-1)!!}\left(\frac{1}{t}\frac{\partial}{\partial t}\right)^{k-1}\left(\frac{\alpha_k}{t}\int_{\partial B_t(x)} \Delta u_1(y)\mathrm{d}y\right)\\
&= \frac{1}{(2k-1)!!}\left(\frac{1}{t}\frac{\partial}{\partial t}\right)^{k-1}\left(\frac{\alpha_k}{t}\int_{\partial B_t(0)} \Delta u_1(x+y)\mathrm{d}y\right)\\
&= \Delta\left(\frac{1}{(2k-1)!!}\left(\frac{1}{t}\frac{\partial}{\partial t}\right)^{k-1}\left(\frac{\alpha_k}{t}\int_{\partial B_t(0)} u_1(x+y)\mathrm{d}y\right)\right)\\
&= \Delta\left(\frac{1}{(2k-1)!!}\left(\frac{1}{t}\frac{\partial}{\partial t}\right)^{k-1}\left(\frac{t^{2k-1}}{|\partial B_t|}\int_{\partial B_t(x)} u_1(y)\mathrm{d}y\right)\right)\\
&= \Delta u(t,x).
\end{aligned}$$

The second case is similar. We left as an exercise to check that the claimed classical solution also satisfies the initial data. □

Exercise 1.35. Check that the function u defined in Theorem 1.34 is actually a classical solution of the $(2k+1)$-dimensional wave equation.

Let us now consider the case of even dimension. The main idea is that if $d = 2k+1$, then the classical solution for the wave equation in $\mathbb{R}^{d-1} = \mathbb{R}^{2k}$ can be obtained by extending each function on $\mathbb{R}^{2k}$ to a function on $\mathbb{R}^{2k+1}$ in a trivial way, i.e., simply prescribing no dependence on x_d. We explain the details only for the case of $d = 2$ and refer to [18, §2.4.d] for the general even-dimensional case.

Theorem 1.36. *Let $k \in \mathbb{N}$. For $u_0 \in C^{k+2}(\mathbb{R}^{2k})$ and $u_1 \in C^{k+1}(\mathbb{R}^{2k})$, define a function u by*

$$\begin{aligned} u(t,x) := \frac{1}{(2k)!!}\frac{\partial}{\partial t}\left(\frac{1}{t}\frac{\partial}{\partial t}\right)^{k-1}\left(\frac{t^{2k}}{|\partial B_t|}\int_{\partial B_t(x)}\frac{u_0(y)}{\sqrt{t^2-|y-x|^2}}\,\mathrm{d}y\right) \\ + \frac{1}{(2k)!!}\left(\frac{1}{t}\frac{\partial}{\partial t}\right)^{k-1}\left(\frac{t^{2k}}{|\partial B_t|}\int_{\partial B_t(x)}\frac{u_1(y)}{\sqrt{t^2-|y-x|^2}}\,\mathrm{d}y\right), \end{aligned} \qquad t>0,\ x\in\mathbb{R}^{2k}.$$

Then $u \in C^{2,2}(\mathbb{R}_+ \times \mathbb{R}^{2k+1})$. Moreover, u is a classical solution of the $(2k)$-dimensional wave equation away from $t = 0$ (where u is not defined). However, $\lim_{t\to 0+} u(t,x_0)$ and $\lim_{t\to 0+}\frac{\partial u}{\partial t}(t,x_0)$ exist for all $x_0 \in \mathbb{R}^{2k+1}$, and in fact

$$\lim_{(t,x)\to(0,x_0)} u(t,x) = u_0(x_0) \qquad \text{and} \qquad \lim_{(t,x)\to(0,x_0)} \frac{\partial u}{\partial t}(t,x) = u_1(x_0)$$

for all $x_0 \in \mathbb{R}^{2k}$.

Proof. As already announced, we are only going to prove this theorem in the case of $k = 1$. Let a function $u : \mathbb{R}_+ \times \mathbb{R}^2 \to \mathbb{R}$ solve the two-dimensional wave equation and define a function $\tilde{u} : \mathbb{R}_+ \times \mathbb{R}^3 \to \mathbb{R}$ by

$$\tilde{u}(t,x_1,x_2,x_3) := u(t,x_1,x_2), \qquad t \geq 0\ , x_1,x_2,x_3 \in \mathbb{R}.$$

Extending similarly u_0 and u_1 to $\tilde{u}_0, \tilde{u}_1 : \mathbb{R}^3 \to \mathbb{R}$ we immediately see that u is a classical solution of the three-dimensional wave equation with initial condition given by $\tilde{u}_0$ and $\tilde{u}_1$. Thus, by Theorem 1.26 the function $\tilde{u}$ necessarily agrees with the classical solution defined in Theorem (1.34), i.e.,

$$\tilde{u}(t,x) = \frac{\partial}{\partial t}\left(\frac{t}{4\pi t^2}\int_{\partial B_t^{(3)}(x)}\tilde{u}_0(z)\,\mathrm{d}\sigma(z)\right) + \frac{t}{4\pi t^2}\int_{\partial B_t^{(3)}(x)}\tilde{u}_1(z)\,\mathrm{d}\sigma(z), \qquad t>0,\ x\in\mathbb{R}^3. \tag{1.38}$$

(Here $B_t^{(3)}(x)$ and $B_t^{(2)}(x)$ denote, for the sake of clarity, the 3- and two-dimensional balls of radius t centered at x, respectively.) By construction, the same formula also holds for u. We want to make it handlier by simplifying the term containing the spherical mean of u_0 by some elementary tools of vector analysis. Observe that

$$\{(y, \sqrt{t^2-|y-x|^2}) : y \in B_t^{(2)}(x)\}$$

is the (upper) (three-dimensional) half-sphere of radius t constructed over the two-dimensional ball $B_t^{(2)}(x)$. Hence, we introduce the parametrising function $\gamma : B_t^{(2)}(x) \to \mathbb{R}$ defined by $\gamma(y) := \sqrt{t^2-|y-x|^2}$, so that the integral of $\tilde{u}_0$ over the sphere $\partial B_t^{(3)}(x)$, hence over the two half-spheres, agrees with

$$\frac{2}{4\pi t^2}\int_{B_t^{(2)}(x)} u_0(y)\sqrt{1+|\nabla\gamma(y)|^2}dy = \frac{1}{2\pi t^2}\int_{B_t^{(2)}(x)}\frac{u_0(y)t}{\sqrt{t^2-|y-x|^2}}dy.$$

Accordingly,

$$\frac{t}{4\pi t^2}\int_{\partial B_t^{(3)}(x)}\tilde{u}_i(z)\,\mathrm{d}\sigma(z) = \frac{1}{2|B_t^{(2)}|}\int_{B_t^{(2)}(x)}\frac{u_i(y)t}{\sqrt{t^2-|y-x|^2}}\mathrm{d}y, \qquad i=0,1. \tag{1.39}$$

We therefore have

$$\begin{aligned} u(t,x) = \tilde{u}(t,x) = \frac{\partial}{\partial t}\left(\frac{1}{2|B_t^{(2)}|}\int_{B_t^{(2)}(x)}\frac{u_0(y)t^2}{\sqrt{t^2-|y-x|^2}}\mathrm{d}y\right) \\ + \frac{1}{2|B_t^{(2)}|}\int_{B_t^{(2)}(x)}\frac{u_1(y)t^2}{\sqrt{t^2-|y-x|^2}}\mathrm{d}y, \qquad t>0,\ x\in\mathbb{R}^3. \end{aligned} \tag{1.40}$$

This concludes the proof. □

In the two-dimensional case it is possible to further simplify the classical solution formula. To reduce the first term on the right hand side we perform a change of variable $y \mapsto x + tz$ and obtain

$$\begin{aligned}\frac{\partial}{\partial t}\left(\frac{t^2}{2|B_t^{(2)}|}\int_{B_t^{(2)}(x)}\frac{u_0(y)}{\sqrt{t^2-|y-x|^2}}\mathrm{d}y\right) &= \frac{\partial}{\partial t}\left(\frac{t}{2|B_1^{(2)}|}\int_{B_1^{(2)}(0)}\frac{u_0(x+tz)}{\sqrt{1-|z|^2}}\mathrm{d}z\right)\\ &= \frac{1}{2|B_1^{(2)}|}\int_{B_1^{(2)}(0)}\frac{u_0(x+tz)}{\sqrt{1-|z|^2}}\mathrm{d}z\\ &\quad + \frac{t}{2|B_1^{(2)}|}\int_{B_1^{(2)}(0)}\frac{\nabla u_0(x+tz)\cdot z}{\sqrt{1-|z|^2}}\mathrm{d}z\end{aligned}$$

Back-substituting finally yields

$$\begin{aligned}\frac{\partial}{\partial t}\left(\frac{t^2}{2|B_t^{(2)}|}\int_{B_t^{(2)}(x)}\frac{u_0(y)}{\sqrt{t^2-|y-x|^2}}\mathrm{d}y\right) &= \frac{t}{2|B_t^{(2)}|}\int_{B_t^{(2)}(x)}\frac{u_0(y)}{\sqrt{t^2-|x-y|^2}}\mathrm{d}y\\ &\quad + \frac{t}{2|B_t^{(2)}|}\int_{B_t^{(2)}(x)}\frac{\nabla u_0(y)\cdot(y-x)}{\sqrt{t^2-|y-x|^2}}dy\end{aligned}$$

Combining this identity with (1.40) we finally arrive at the formula

$$u(t,x) = \frac{1}{2\pi t}\int_{B_t^{(2)}(x)}\frac{u_0(y)+\nabla u_0(y)\cdot(y-x)+tu_1(y)}{\sqrt{t^2-|y-x|^2}}\mathrm{d}y, \qquad t\geq 0,\ x\in\mathbb{R}^2,$$

which is easily seen to be the special case of the general formula for $d = 2$.

Remark 1.37. In other words, if we consider two functions u_0 and $\tilde{u}_0$ that only differ inside a bounded set, say $B_1(0)$, then the classical solutions to the wave equations with initial data u_0 and $\tilde{u}_0$ are seen to differ at points outside the ball only after a certain time. One refers to this behavior by saying that the wave equations enjoys **finite speed of propagation**. Equations – like the wave equations – with finite speed of propagation and such that their classical solutions are not more regular than the initial data are referred to as **hyperbolic equations**.

Remark 1.38. The above formulae for the classical solution of the wave equation are often interpreted as a heuristic explanation for (or rather, recognition of) the fact that the world we are living in is odd-dimensional (with respect to the space). This is due to the fact that, by the formulae just obtained (and their generalizations to arbitrary natural numbers), in the three- (and, more generally, odd-) dimensional case any variation of the initial data at, say, $x \in \mathbb{R}^d$ only affect the classical solution on the *surface* of the **sound cone**

$$\{y\in\mathbb{R}^d : |x-y|\leq t\},$$

in accordance with of the **Huygens' principle** of acoustics, whereas in two- (and, more generally, even-) dimensional case any such variation should also affect the classical solution in the *interior* of the sound cone, against experimental observations (think of a jet fighter's sonic boom).

Let us finally consider again the energy of a wave, which we have introduced in Definition 1.22, a prove a higher-dimensional counterpart of Theorem 1.21.

Theorem 1.39. *Let $\Omega\subset\mathbb{R}^d$ be a domain with C^1-boundary (and in particular such that the Gauss–Green formulae hold). Then the total energy $E_p(t)+E_k(t)$ is constant for all times t, i.e.,*

$$E_p(t)+E_k(t) = \int_\Omega\left|\frac{\partial u}{\partial t}(0,x)\right|^2\mathrm{d}x + \int_\Omega|\nabla u(0,x)|^2dx, \qquad t\geq 0.$$

where u is any classical solution of the homogeneous wave equation (1.32) *on Ω with either Neumann or Dirichlet boundary conditions such that (in the case of Ω unbounded) u is uniformly bounded on any compact domain.*

Proof. The proof essentially reproduces the strategy of the proof in the one-dimensional case, up to replacing integration by parts by the Gauss–Green formula (B.5). Indeed, we find

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\left(E_p(t)+E_k(t)\right) &= \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega \left|\frac{\partial u}{\partial t}(t,x)\right|^2 dx + \frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega |\nabla u(t,x)|^2\,\mathrm{d}x \\
&= \int_\Omega \frac{\mathrm{d}}{\mathrm{d}t}\left\langle \frac{\partial u}{\partial t}(t,x), \frac{\partial u}{\partial t}(t,x)\right\rangle \mathrm{d}x + \int_\Omega \frac{\mathrm{d}}{\mathrm{d}t}\left\langle \nabla u(t,x), \nabla u(t,x)\right\rangle \mathrm{d}x \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega \left\langle \frac{\partial^2 u}{\partial t^2}(t,x), \frac{\partial u}{\partial t}(t,x)\right\rangle \mathrm{d}x + \int_\Omega \left\langle \nabla u(t,x), \nabla\frac{\partial u}{\partial t}(t,x)\right\rangle \mathrm{d}x \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega \left\langle \frac{\partial^2 u}{\partial t^2}(t,x), \frac{\partial u}{\partial t}(t,x)\right\rangle \mathrm{d}x - \int_\Omega \left\langle \Delta u(t,x), \frac{\partial u}{\partial t}(t,x)\right\rangle \mathrm{d}x \\
&\quad + \int_{\partial\Omega} \left\langle \frac{\partial u}{\partial \nu}(t,z), \frac{\partial u}{\partial t}(t,z)\right\rangle \mathrm{d}\sigma(z) \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega \left\langle \frac{\partial^2 u}{\partial t^2}(t,x) - \Delta u(t,x), \frac{\partial u}{\partial t}(t,x)\right\rangle \mathrm{d}x,
\end{aligned}$$

where the boundary term disappears owing to the imposed boundary conditions. However, since u is a classical solution of the wave equation on Ω, the identity $\frac{\partial^2 u}{\partial t^2}(t,x) = \Delta\frac{\partial u}{\partial t}(t,x)$ holds pointwise and for any time t. It follows that

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(E_p(t)+E_k(t)\right) = \frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega \left\langle \frac{\partial^2 u}{\partial t^2}(t,x) - \Delta u(t,x), u(t,x)\right\rangle \mathrm{d}x = 0.$$

This concludes the proof. □

Remark 1.40. The feature described in Theorem 1.39 does not seem very realistic. All wave-type phenomena we can commonly observe are not eternal: due to internal frictions and possibly to the influence of external forces (like the gravity) the oscillation fades away and the system eventually comes to quiet. In other words, the system dissipates energy. This would only be possible if the computations in the proof of Theorem 1.39 could be shown to eventually yield an estimate which is less than 0, say, less than $\int_\Omega |\alpha(x)\frac{\partial u}{\partial t}(t,x)|^2 dx$ for some function $\alpha : \Omega \to (-\infty, 0)$ and any t. In other words, we would like to have

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega \left(\frac{\partial^2 u}{\partial t^2}(t,x) - \Delta u(t,x) - \alpha(x)\frac{\partial u}{\partial t}(t,x)|\frac{\partial u}{\partial t}(t,x)\right)\mathrm{d}x \le 0.$$

This is surely the case if u does not solve the original wave equation, but rather

$$\frac{\partial^2 u}{\partial t^2}(t,x) = \Delta u(t,x) + \alpha(x)\frac{\partial u}{\partial t}(t,x).$$

Since energy is absorbed throughout the process, this is called a **damped wave equation**. If on the other hand $\alpha(x) > 0$, $x \in \Omega$, then the system's energy increases during time (in absence of external force, such a behavior is of course unrealistic).

Exercise 1.41. Another way to control the energy decay of a vibrating system is to apply a so-called **closed feedback**. This means that the system is controlled by some gadget that is able to modify the boundary conditions in real time. Mathematically speaking, this amounts to considering a modified boundary condition, e.g.,

$$\frac{\partial u}{\partial \nu}(t,z) = \frac{\partial u}{\partial t}(t,z), \qquad t \ge 0,\ z \in \partial\Omega.$$

Check whether the wave equation endowed with this boundary condition actually enjoys energy decay.

Exercise 1.42. Certain investigations have led applied mathematicians and theoretical physicists to consider wave equations equipped with **acoustic boundary conditions**, which can be written in the form

$$\begin{cases}
\dfrac{\partial^2 u}{\partial t^2}(t,x) = c^2\Delta u(t,x), & t \ge 0,\ x \in \Omega, \\
m\dfrac{\partial^2 \delta}{\partial t^2}(t,z) = -d\dfrac{\partial \delta}{\partial t}(t,z) - k\delta(t,z) - \rho\dfrac{\partial \phi}{\partial t}(t,z), & t \ge 0,\ z \in \partial\Omega, \\
\dot\delta(t,z) = \dfrac{\partial u}{\partial \nu}(t,z), & t \ge 0,\ z \in \partial\Omega.
\end{cases}$$

Here ϕ is the velocity potential of a fluid filling an open Lipschitz domain $\Omega \subset \mathbb{R}^d$, either bounded or exterior (i.e., the complement of a compact domain); δ is the normal displacement of the boundary $\partial\Omega$ of Ω; m, d, and k are the mass per unit area, the resistivity, and the spring constant of the boundary, respectively; finally, ρ and c are the unperturbed density of, and the speed of sound in the medium, respectively. Assume m, k, d, ρ to be positive constants.

Introduce a suitable energy function and show that this energy is decaying. *(Hint: The term δ on $\partial\Omega$ contributes to the total energy of the system, too.)*

1.4 Notes and suggestions for further study

One may wonder where the name "hyperbolic" for the class of PDEs treated in this section comes from. A classical classification of second order partial differential equations in two variables is based upon writing them as

$$A\frac{\partial^2 u}{\partial x_1^2} + 2B\frac{\partial^2 u}{\partial x_1 \partial x_2} + C\frac{\partial^2 u}{\partial x_2^2} + D\frac{\partial u}{\partial x_1} + E\frac{\partial u}{\partial x_2} + Fu = 0 \tag{1.41}$$

where A, B, C, D, E, F are possibly depending on the variables x_1, x_2, or rather – upon neglecting lower order terms –

$$A\frac{\partial^2 u}{\partial x_1^2} + 2B\frac{\partial^2 u}{\partial x_1 \partial x_2} + C\frac{\partial^2 u}{\partial x_2^2} = 0 : \tag{1.42}$$

clearly, the one-dimensional wave equation takes this form, with $A = c^2$, $B = 0$, $C = 1$. Now, by analogy with the classical classification of conic sections of equation

$$Ax^2 + 2Bxy + Cy^2 = 0$$

into hyperbolic, parabolic, and elliptic depending on the discriminant $B^2 - 4AC$, one calls the partial differential equation (1.42)

- hyperbolic if $B^2 < 4AC$,
- parabolic if $B^2 = 4AC$,
- elliptic if $B^2 > 4AC$.

The motivating feature of this classification is that equations belonging to any of these classes do share specific features with other representatives of the same class. For example, hyperbolic equations can be solved by the method of characteristics and enjoy finite speed of propagation; for this reason, also the transport equation – which is of first order and hence cannot be formally classified within this scheme – is usually considered hyperbolic. We will meet examples of parabolic and elliptic equations in Chapters 2 and 3: the heat equation ($A = c^2$, $E = 1$, $B = C = D = F = 0$) and the Laplace equation ($A = C = 1$, $B = D = E = F = 0$). This classification is incomplete, though, as important equations derived in the last century fail to reasonably fit into this scheme: a prominent example of a partial differential equation that is neither hyperbolic, nor parabolic, nor elliptic is the Schrödinger that we will meet in Chapter 4: indeed, the Schrödinger equations displays a behavior that lies somewhere between those of hyperbolic and parabolic equations.

In (1.2) we have just encountered a fundamental example of **continuity equation** (sometimes also called **scalar conservation law**). If we want to impose the conservation of a physical quantity q (mass, energy or electric charge, for example) or maybe indistinguishable specimen (number of particles in a reactor, individuals in a population; if their number is high enough, we can conveniently decide to regard it as a real number, in order to apply methods of differential calculus, rather than discrete mathematics) one may consider the flow of this quantity through two surfaces of the domain of transmission: if the total mass (or energy, or number of individuals, etc.) passing through the surfaces in the time unit is the same, it means that it has been conserved during the transport process. Imposing this property typically has a physical or modeling meaning. More generally, if a scalar-valued function ρ is the volume density of this quantity (i.e., the amount per unit volume), a vector field j is its flux, and σ is the generation (or loss, if $\sigma < 0$) of q per unit volume and unit time, then the continuity equation can be also expressed, in differential form, as

$$\frac{\partial \rho}{\partial t}(t,x) + \nabla j(t,x) = \sigma(t,x), \qquad t \geq 0,\ x \in \Omega. \tag{1.43}$$

The transport equation in (1.3) arises considering models where the proportionality relation

$$j = c\rho \tag{1.44}$$

is appropriate. Many further equations of mathematical physics share this fundamental structure, as we will see in Chapter 2.

The only reason why we require the initial data u_0 to be continuously differentiable in Theorem 1.3 is that this implies, via the formula (1.7), continuous differentiability of u with respect to time: this is a required property of solutions, according to Definition 1.2. While we derived (1.7) for C^1-data u_0, this formula clearly makes sense for more general u_0; for example, it defines a continuous density function u provided u_0 is merely continuous. More interestingly, if u_0 is continuously differentiable up to a limited number of irregular points (think of the function $u_0(x) := |x - \frac{b-a}{2}|$), then u can be regarded as a solution of the transport equation in a generalized sense. This is e.g. the case if u_0 belongs to a suitable Sobolev space, say $u_0 \in W^{1,2}(a,b)$: multiplying by a suitably smooth function v with compact support in some $\tilde{\Omega} \subset \mathbb{R} \times \Omega$ and then integrating [3], we find by Fubini's Theorem A.21

$$\begin{aligned}\int_0^\infty \int_\Omega f(t,x)v(t,x)\ \mathrm{d}t\,\mathrm{d}x &= \int_0^\infty \int_\Omega \left(\frac{\partial u}{\partial t} + c(x)\frac{\partial u}{\partial x}\right)(t,x)v(t,x)\ \mathrm{d}x\,\mathrm{d}t \\ &= -\int_0^\infty \int_\Omega u\frac{\partial v}{\partial t}(t,x)\ \mathrm{d}x\,\mathrm{d}t - \int_\Omega u(0,x)v(0,x)\ \mathrm{d}x \\ &\quad -\int_0^\infty \int_\Omega u(t,x)\frac{\partial v}{\partial x}(t,x)\ \mathrm{d}x\,\mathrm{d}t.\end{aligned} \tag{1.45}$$

where the second step follows integrating by parts with respect to the time variable. Clearly, each solutions of (1.5) in the sense of Definition 1.2 also verifies (1.45), but the converse fails to hold. We are hence led to the identity

$$\begin{aligned}\int_0^\infty \int_\Omega f(t,x)v(t,x)\mathrm{d}t\,\mathrm{d}x = &-\int_0^\infty \int_\Omega \left(u(t,x)\frac{\partial v}{\partial t}(t,x) - u(t,x)\frac{\partial v}{\partial x}(t,x)\right)\mathrm{d}x\mathrm{d}t \\ &-\int_\Omega u_0(x)v(0,x)\mathrm{d}x.\end{aligned} \tag{1.46}$$

which we regard as integral formulation of (1.3) with initial data (1.5). We want to appreciate that, for any smooth function v with compact support, (1.46) makes sense as soon as u is merely of class $L^1(\mathbb{R}_+ \times \Omega)$.

The introduction of this and further, related generalized notions of solutions is the starting point of the modern theory of continuity equations. [18, Chapter 3] is a good, compact introduction to this topic, whose sophisticated offspring, the modern analytical theory of optimal transport, has led to awarding of two recent Fields Medals: to Cédric Villani (2010) and Alessio Figalli (2018).

The method of characteristics is not almighty: we will meet already in the next chapter an important class of partial differential equations that cannot be tackled by this method. Still, the search for wave-like phenomena in a physical system is an important issue in the analysis of evolution equations and leads to the topic of *traveling waves*, see Example 1.17: these are, roughly speaking, solutions of a given partial differential equation that can be written as in (1.6). We have seen in Section 1.1 that *all* solutions of the one-dimensional transport equation have this structure, but can it be that for other equations *some* initial data u_0 lead to a solution u that evolves in time as a shift of u_0? The initial data 0 induces the trivial solution 0 in all linear, homogeneous partial differential equations (with linear, homogeneous boundary conditions), but the existence of different initial data inducing non-trivial traveling wave solutions is often a fingerprint for some deeper functional organization of a system.

The way sound is perceived by mammalian hearing organ is known to be described by a complicated nonlinear wave equation with boundary conditions that refine the acoustic ones we have encountered in 1.42. Working in his lab at Harvard in the late 1950s, Georg von Békésy experimentally detected patterns similar to traveling waves induced by sound transmission in the cochlea: for this achievement he was awarded the Nobel Prize in Physiology or Medicine in 1961.

[3] One informally refers to this procedure as: *integrating against a test function* v.

1.5 Solutions

Solution 1.43 (Exercise 1.7). A sufficient condition for extending (1.7) to the negative time axis is that $f \in C^{0,1}(\mathbb{R}\times\mathbb{R})$. In that case,

$$(t,x) \mapsto u_0(x-ct) + \int_0^t f(\tau, x+c(\tau-t))\,\mathrm{d}\tau, \qquad t\in\mathbb{R}, x\in\mathbb{R}$$

is a solution of

$$\begin{cases} \dfrac{\partial u}{\partial t}(t,x) = -c\dfrac{\partial u}{\partial x}(t,x) + f(t,x), & t\in\mathbb{R},\ x\in\mathbb{R}, \\ u(0,x) = u_0(x), & x\in\mathbb{R}, \end{cases} \tag{1.47}$$

i.e., $u\in C^{1,1}(\mathbb{R})$ satisfying both the transport equation and the initial condition in (1.47). Observe that we still need a condition at $t=0$, even if our time domain is now the whole real line. Indeed, this procedure delivers a solution for the backward evolution equation, too: this formally corresponds to searching for a function satisfying the PDE for negative t as well as fulfilling a *final* data assignment $u(0,\cdot)=u_0$.

Solution 1.44 (Exercise 1.9). We set $(t,x)=(s,\gamma(s))$ and obtain

$$\frac{\mathrm{d}}{\mathrm{d}s}u(s,\gamma(s)) = \frac{\partial u}{\partial t} + \gamma'(s)\frac{\partial u}{\partial x} \overset{!}{=} 0\,.$$

Comparing with $tu_x = xu_t$ we obtain the relation

$$\gamma'(s) = -\frac{s}{\gamma(s)}$$

which gives $\frac{\mathrm{d}}{\mathrm{d}s}(\gamma(s))^2 = 2s$ and therefore $\gamma^2(s) = -s^2+\gamma^2(0)$. This immediately yields-

$$u(t,x) = 3\left(x^2+t^2\right)\,.$$

Solution 1.45 (Exercise 1.11). We assume that the velocity function c is constant and positive. Assume also that $u_0\in C^1(-\infty,b]$ for some $b\in\mathbb{R}$. Furthermore, let $f\in C^{0,1}(\mathbb{R}_+\times\mathbb{R})$ be given. In this case, we again obtain the formula

$$u(t,x) = u_0(x-ct) + \int_0^t f(\tau, x+c(\tau-t))\,\mathrm{d}\tau\,,$$

for $x\in(-\infty,b]$ and $t\geq 0$.

In the second case, we have $u_0\in C^1[a,\infty)$ and $\phi\in C^1(\mathbb{R}_+)$. For the solution we then obtain

$$u(t,x) = \begin{cases} u_0(x-ct) + \int_0^t f(\tau, x+c(\tau-t))\,\mathrm{d}\tau\,, & t\geq 0,\ x\in[a+ct,\infty)\,, \\ \phi\left(t+\frac{a-x}{c}\right) + \int_0^t f(\tau, x+c(\tau-t))\,\mathrm{d}\tau\,, & t\geq 0,\ x\in[a,a+ct)\,. \end{cases}$$

This implies that $u(0,a)=\phi(0)$ and $u(t,a)=\phi(t)+\int_0^t f(\tau,a+c(\tau-t))\,\mathrm{d}\tau$.

Solution 1.46 (Exercise 1.16). We find for all $t,s,x\in\mathbb{R}$

$$C(0)f(x) = \frac{1}{2}\left(f(x)+f(x)\right) = f(x)$$

and

$$\begin{aligned} 2C(t)\left(C(s)f\right)(x) &= 2C(t)\left(\frac{1}{2}(f(x+cs)+f(x-cs)\right) \\ &= \frac{1}{2}(f(x+cs+ct)+f(x+cs-ct)+f(x-cs+ct)+f(x-cs-ct)) \\ &= C(s+t)f(x)+C(t-s)f(x) \end{aligned}$$

as well as

$$\begin{aligned} \left(C(s)(S(t)f)+S(s)(C(t)f)\right)(x) &= \frac{1}{4}\int_{x-ct+cs}^{x+ct+cs} f(y)\mathrm{d}y \\ &\quad + \frac{1}{4}\int_{x-ct-cs}^{x+ct-cs} f(y)\mathrm{d}y + \frac{1}{4}\int_{x-cs}^{x+cs}\left(f(z+ct)+f(z-ct)\right)\mathrm{d}z \\ &= S(t+s)f(x). \end{aligned}$$

In other words, these operators satisfy

$$\begin{aligned} C(0) &= \mathrm{Id}, \\ 2C(t)C(s) &= C(t+s) + C(t-s), \\ C(s)S(t) + S(s)C(t) &= S(t+s). \end{aligned}$$

These identities shall now be compared with the classical relations

$$\begin{aligned} \cos 0 &= 1, \\ 2\cos t \cos s &= \cos(t+s) + \cos(t-s), \\ \cos s \sin t + \sin s \cos t &= \sin(t+s). \end{aligned}$$

Solution 1.47 (Exercise 1.17). To begin with, observe that

$$\frac{\partial^2}{\partial t^2} u(t,x) = A\frac{\partial^2}{\partial t^2}\sin(kx-\omega t) = -A\omega\frac{\partial}{\partial t}\cos(kx-\omega t) = -A\omega^2\sin(kx-\omega t) = -\omega^2 u(t,x)$$

whereas

$$\frac{\partial^2}{\partial x^2} u(t,x) = A\frac{\partial^2}{\partial x^2}\sin(kx-\omega t) = Ak\frac{\partial}{\partial x}\cos(kx-\omega t) = -Ak^2\sin(kx-\omega t) = -k^2 u(t,x)$$

and likewise

$$\frac{\partial^2}{\partial t^2} v(t,x) = -\omega^2 v(t,x) \quad \text{and} \quad \frac{\partial^2}{\partial x^2} v(t,x) = -k^2 v(t,x).$$

Accordingly, because u, v are clearly smooth functions, both u, v are classical solutions of the wave equation with wave speed $\frac{\omega}{k}$, i.e.,

$$\frac{\partial^2 u}{\partial t^2}(t,x) = \frac{\omega^2}{k^2}\frac{\partial^2 u}{\partial x^2}(t,x), \qquad t \geq 0,\ x \in \mathbb{R}.$$

Because this wave equation is linear, also sums of solutions – especially, $u + v$ – are solutions in their own right. Let us check that this sum cannot be represented as a traveling wave. Indeed, the well-known trigonometric prosthaphaeresis formulae ("product-to-sum identities") can be invoked to write

$$A\sin(kx-\omega t) + A\sin(kx+\omega t) = 2A\sin(kx)\cos(\omega t).$$

However, $2A\sin(kx)\cos(\omega t)$ cannot generally expressed as $\phi(hx+ct)$ for any function $\phi : \mathbb{R} \to \mathbb{R}$ and any $k, \omega \in \mathbb{R}$: indeed, assume for a moment that – for example, if $A = \frac{1}{2}$ and $k = \omega = 1$ – such a ϕ exists: then because $(t,x) \mapsto \phi(hx+ct)$ must solve the wave equation we find

$$0 = \left(\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2}\right)\phi(hx+ct) = (c^2+h^2)\phi(hx+ct), \qquad t \geq 0,\ x \in \mathbb{R}.$$

Now $\phi \not\equiv 0$ because in general $\sin(kx)\cos(\omega t)$ cannot vanish identically, hence either $c = h$ or $c = -h$, i.e.,

$$\sin x \cos t = \phi(h(x+t)) \quad \text{or} \quad \sin x \cos t = \phi(h(x-t)) \qquad \text{for all } t \geq 0,\ x \in \mathbb{R}. \tag{1.48}$$

Evaluating (1.48) in particular along $(t,x) \in \{0\} \times \mathbb{R}$ we find $\phi(h\xi) = \sin(\xi)$, $\xi \in \mathbb{R}$. This leads to the sought-after contradiction since, on the other hand, evaluating (1.48) along $(t,x) \in \mathbb{R}_+ \times \{0\}$ instead, we find $0 = \phi(ht) = \sin(t)$.

Solution 1.48 (Exercise 1.18). Let t, x be defined by

$$t(\lambda,\mu) := \frac{\mu-\lambda}{2}, \quad x(\lambda,\mu) := \frac{\mu+\lambda}{2}.$$

Assuming that u solves the wave equation with wave speed c, we find

$$\frac{1}{4c^2}\frac{\partial^2 u}{\partial t^2} - \frac{1}{4}\frac{\partial^2 u}{\partial x^2} = \frac{\partial^2 v}{\partial\lambda\partial\mu} \overset{!}{=} 0.$$

If the last identity is satisfied, then we can deduce that $\frac{\partial v}{\partial \mu}$ does not depend on μ,

$$\frac{\partial v}{\partial \mu} = f(\lambda),$$

for some function f, and accordingly

$$v(\lambda, \mu) = F(\lambda) + G(\mu)$$

for appropriate questions F, G: therefore, the general question of the wave equation is – as we already know – $(t, x) \mapsto F(x - ct) + G(x - ct)$, for some $F, G \in C^2(\mathbb{R})$.

From the initial conditions $F(x) + G(x) = u_0(x)$ – whence $F'(x) + G'(x) = u_0'(x)$ – and $c(G'(x) - F'(x)) = u_1(x)$ we deduce

$$G'(x)) = \frac{1}{2}u_0'(x) + \frac{1}{2c}u_1(x), \quad F'(x) = \frac{1}{2}u_0'(x) - \frac{1}{2c}u_1(x),$$

and hence

$$G(x + ct) - G(0) = \frac{1}{2c}\int_0^{x+ct} u_1(y)\mathrm{d}y + \frac{1}{2}u_0(x + ct) - \frac{1}{2}u_0(0)$$

along with

$$-F(x - ct) + F(0) = -\frac{1}{2c}\int_{x-ct}^{0} u_1(y)\mathrm{d}y - \frac{1}{2}u_0(x - ct) + \frac{1}{2}u_0(0).$$

This finally delivers the D'Alembert formula (1.29).

Solution 1.49 (Exercise 1.19). To begin with, we observe that

$$\begin{aligned}\left(\frac{\partial}{\partial t} + x\frac{\partial}{\partial x}\right)\left(\frac{\partial}{\partial t} - x\frac{\partial}{\partial x}\right) &= \frac{\partial^2}{\partial t^2} - x\frac{\partial^2}{\partial t \partial x} + x\frac{\partial^2}{\partial t \partial x} - x\frac{\partial}{\partial x}\left(x\frac{\partial}{\partial x}\right)\\ &= \frac{\partial^2}{\partial t^2} - x^2\frac{\partial^2}{\partial x^2} - x\frac{\partial}{\partial x},\end{aligned}$$

by Schartz's Theorem. Let now u be a solution (assuming it exists!) to the given equation and set

$$v(t, x) := \left(\frac{\partial}{\partial t} - x\frac{\partial}{\partial x}\right)u(t, x).$$

Accordingly,

$$v(t, x) = v(0, x\,\mathrm{e}^{-t}) =: \phi(x\,\mathrm{e}^{-t})$$

for some function $\phi \in C^1(\mathbb{R})$. We are thus led to study the equation

$$\left(\frac{\partial}{\partial t} - x\frac{\partial}{\partial x}\right)u(t, x) = \phi(x\,\mathrm{e}^{-t}), \qquad t, x \in \mathbb{R},$$

whose characteristic, for each $(t, x) \in \mathbb{R}^2$, can be easily seen to be given by

$$\gamma(s) = x\,\mathrm{e}^{t-s}, \qquad s \in \mathbb{R}.$$

It follows that

$$u(t, x) = u_0(x\,\mathrm{e}^t) + \int_0^t \phi(x\,\mathrm{e}^{t-s}\,\mathrm{e}^{-s})\mathrm{d}s.$$

Now, because

$$u_1(x) = \frac{\partial u}{\partial t}(0, x) = xu_0'(x) + \phi(x), \quad \text{that is,} \quad \phi(x) = u_1(x) - xu_0'(x),$$

we deduce that

$$\begin{aligned}u(t, x) &= u_0(x\,\mathrm{e}^t) + \int_0^t \left(-x\,\mathrm{e}^{t-2s}\,u_0'(x\,\mathrm{e}^{t-2s}) + u_1(x\,\mathrm{e}^{t-2s})\right)\mathrm{d}s\\ &= \frac{1}{2}\left(u_0(x\,\mathrm{e}^t) + u_0(x\,\mathrm{e}^{-t})\right) + \frac{1}{2}\int_{x\,\mathrm{e}^{-t}}^{x\,\mathrm{e}^t}\frac{1}{y}u_1(y)\mathrm{d}y, \qquad t, x \in \mathbb{R}.\end{aligned} \tag{1.49}$$

Summing up, we have shown that *if* a solution u to (1.30) exists, then it must have the form (1.49). But is this ansatz actually successful? Indeed, a direct computation easily shows that u actually solves (1.30); indeed, if $u_0 \in C^2(\mathbb{R})$ and $u_1 \in C^1(\mathbb{R})$, then $u \in C^{2,2}(\mathbb{R}_+ \times \mathbb{R})$.

Solution 1.50 (Exercise 1.24). In order to solve the wave equation on the interval $[0,1]$ subject to Dirichlet boundary conditions, we assume that $u_0 \in C^2([0,1])$ and $u_1 \in C^1([0,1])$. In addition, u_0 is assumed to fulfill Dirichlet boundary conditions at the end points and we shall also assume that $u_0''(0) = u_0''(1) = 0$. Also, u_1 shall also fulfill Dirichlet boundary conditions at the end points.

In a first step we now extend u_0, u_1 to all of $\mathbb{R}$ via

$$\tilde{u}_{0/1}(x) := \begin{cases} u_{0/1}(x - \lfloor x \rfloor), & \text{if} \quad \lfloor x \rfloor = 2k \,, \\ -u_{0/1}(1 - (x - \lfloor x \rfloor)), & \text{if} \quad \lfloor x \rfloor = 2k+1 \,, k \in \mathbb{Z} \,. \end{cases}$$

Then, $\tilde{u}_{0/1}(n) = 0$ for all $n \in \mathbb{Z}$, $\tilde{u}_{0/1}(-x) = -\tilde{u}_{0/1}(x)$ and $\tilde{u}_{0/1}(1-x) = -\tilde{u}_{0/1}(1+x)$. Consequently, the solution we aim for is given by

$$u(t,x) := \frac{1}{2}\left(\tilde{u}_0(x-t) + \tilde{u}_0(x+t)\right) + \frac{1}{2}\int_{x-t}^{x+t} \tilde{u}_1(y)\,\mathrm{d}y \,.$$

Solution 1.51 (Exercise 1.25). We look at the solution of the wave equation on the interval $[0,1]$ subject to Neumann boundary conditions. As in Solution 1.50 we extend the initial data u_0, u_1 to all of $\mathbb{R}$, assuming the same boundary conditions as in Solution 1.50. More explicitly, we set

$$\tilde{u}_{0/1}(x) := \begin{cases} u_{0/1}(x - \lfloor x \rfloor) & \text{if} \quad \lfloor x \rfloor = 2k \,, \\ u_{0/1}(1 - (x - \lfloor x \rfloor)) & \text{if} \quad \lfloor x \rfloor = 2k+1 \,, k \in \mathbb{Z} \,. \end{cases}$$

Then, $\tilde{u}_{0/1}(-x) = -\tilde{u}_{0/1}(x)$, $\tilde{u}_{0/1}(1-x) = \tilde{u}_{0/1}(1+x)$ from which we obtain $\tilde{u}_0'(-x) = -\tilde{u}_0'(x)$ as well as $\tilde{u}_0'(1-x) = -\tilde{u}_0'(1+x)$. Hence, the solution is given by

$$u(t,x) := \frac{1}{2}\left(\tilde{u}_0(x-t) + \tilde{u}_0(x+t)\right) + \frac{1}{2}\int_{x-t}^{x+t} \tilde{u}_1(y)\,\mathrm{d}y \,.$$

From this formula we directly conclude that $u(t,x) \geq 0$ given that $\tilde{u}_{0/1}(x) \geq 0$ for all $x \in \mathbb{R}$. This, however, is equivalent to assuming $u_0(x), u_1(x) \geq 0$ for $x \in [0,1]$.

Solution 1.52 (Exercise 1.27). We want to prove that the d-dimensional wave equation $u_{tt}(t,x) = \Delta u(t,x)$, $(t,x) \in \mathbb{R}_+ \times \Omega$, with initial conditions $u(0,x) = u_0(x)$ and $u_t(0,x) = u_1(x)$ has at most one solution.

By looking at the difference of two solutions, it is clear that we have to prove uniqueness only for zero initial data. In other words, let w be the solution to zero initial datum. Then, we have to prove that $w \equiv 0$. To do this, hen define the energy function

$$E(t) := \int_\Omega \left(\frac{1}{2}|w_t(t,x)|^2 + \frac{1}{2}|\nabla w(t,x)|^2\right)\,\mathrm{d}x$$

and obtain, $B_R(0)$ denoting the ball of radius $R > 0$ around the origin and $B_R^c(0)$ its complement,

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}E(t) &= \int_{\Omega\cap B_R(0)} \left(w_t(t,x)w_{tt} + \nabla w(t,x)\cdot\nabla w_t(t,x)\right)\,\mathrm{d}x \\ &\quad + \int_{\Omega\cap B_R^c(0)} \left(w_t(t,x)w_{tt} + \nabla w(t,x)\cdot\nabla w_t(t,x)\right)\,\mathrm{d}x \\ &= \int_{\Omega\cap B_R(0)} w_t(t,x)\left(w_{tt} - \Delta w(t,x)\right)\,\mathrm{d}x + \int_{\partial(\Omega\cap B_R(0))} w_t(t,y)\frac{\partial w}{\partial \nu}(t,y)\,\mathrm{d}\sigma(y) \\ &\quad + \int_{\Omega\cap B_R^c(0)} \left(w_t(t,x)w_{tt} + \nabla w(t,x)\cdot\nabla w_t(t,x)\right)\,\mathrm{d}x \\ &= \int_{\partial(\Omega\cap B_R(0))} w_t(t,y)\frac{\partial w}{\partial \nu}(t,y)\,\mathrm{d}\sigma(y) + \int_{\Omega\cap B_R^c(0)} \left(w_t(t,x)w_{tt} + \nabla w(t,x)\cdot\nabla w_t(t,x)\right)\,\mathrm{d}x \,. \end{aligned}$$

Now, since $w_{x_j}(t,\cdot), w_{x_j t}(t,\cdot), w_{tt}(t,\cdot), w_t(t,\cdot)$ are assumed to be square integrable, we conclude that the last term vanishes in the limit $R \to \infty$. Also, the first term vanishes since w fulfills Neumann or Dirichlet boundary conditions (in the case of Dirichlet boundary conditions, just pull the t-derivative in front of the integral). As a consequence, the energy $E(\cdot)$ is constant in time and since $E(0) = 0$ we conclude that $w_t, w_{x_j} = 0$ and hence w is constant and hence the zero function.

Solution 1.53 (Exercise 1.29). We obtain

$$\begin{aligned}\frac{\partial}{\partial r}\left(\frac{1}{|B_r|}\int_{\partial B_r(x)} u(y)\mathrm{d}\sigma(y)\right) &= -\frac{d}{r^2}\phi(r) + \frac{1}{|B_r|}\int_{B_r(x)} \Delta u(y)\,\mathrm{d}y \\ &= -\frac{d}{r^2}\phi(r) + \frac{d}{r}\phi'(r)\,.\end{aligned}$$

Furthermore, realizing that $\frac{\partial}{\partial r}\int_{B_r(x)} u(y)\,\mathrm{d}y = \int_{\partial B_r(x)} u(y)\,\mathrm{d}\sigma(y)$ we obtain

$$\frac{\partial}{\partial r}\left(\frac{1}{|\partial B_r|}\int_{B_r(x)} u(y)\mathrm{d}\sigma(y)\right) = r(d-1)\frac{1}{|\partial B_r|}\int_{B_r(x)} u(y)\,\mathrm{d}y + \phi(r)\,.$$

In the same way,

$$\frac{\partial}{\partial r}\left(\frac{1}{|B_r|}\int_{B_r(x)} u(y)\mathrm{d}\sigma(y)\right) = -\frac{d}{r}\frac{1}{|B_r|}\int_{B_r(x)} u(y)\,\mathrm{d}y + \frac{d}{r}\phi(r)\,.$$

Solution 1.54 (Exercise 1.35). For fixed $x \in \mathbb{R}^{2k+1}$ we set $\phi_{0/1}(t) = \frac{1}{|\partial B_t|}\int_{\partial B_t(x)} u_{0/1}(z)\,\mathrm{d}z$. One obtains

$$\begin{aligned}u(t,x) &= \frac{1}{(2k-1)!!}\frac{\mathrm{d}}{\mathrm{d}t}\left((2k-1)!!t\phi_0(t) + \sum_{j=1}^{k-1}\beta_j^k t^{j+1}\frac{\mathrm{d}^j}{\mathrm{d}t^j}\phi_0(t)\right) \\ &\quad + \frac{1}{(2k-1)!!}\frac{\mathrm{d}}{\mathrm{d}t}\left((2k-1)!!t\phi_1(t) + \sum_{j=1}^{k-1}\beta_j^k t^{j+1}\frac{\mathrm{d}^j}{\mathrm{d}t^j}\phi_1(t)\right) \\ &= \phi_0(t) + t\phi_0'(t) + \frac{1}{(2k-1)!!}\sum_{j=1}^{k-1}\left((j-1)\beta_j^k t^j\frac{\mathrm{d}^j}{\mathrm{d}t^j}\phi_0(t) + \beta_j^k t^{j+1}\frac{\mathrm{d}^{j+1}}{\mathrm{d}t^{j+1}}\phi_0(t)\right) \\ &\quad + t\phi_1(t) + \frac{1}{(2k-1)!!}\left(\sum_{j=1}^{k-1}\beta_j^k t^{j+1}\frac{\mathrm{d}^j}{\mathrm{d}t^j}\phi_1(t)\right)\,.\end{aligned}$$

for some constants β_j^k. This formula then implies

$$\begin{aligned}\frac{\partial}{\partial t}u(t,x) &= 2\phi_0'(t) + t\phi_0''(t) + \phi_1(t) \\ &\quad + \frac{1}{(2k-1)!!}\left(\sum_{j=1}^{k-1} c_1 t^{j+1}\frac{\mathrm{d}^j}{\mathrm{d}t^j}\phi_0(t) + c_2 t^j\frac{\mathrm{d}^{j+1}}{\mathrm{d}t^{j+1}}\phi_0(t) + c_3 t^{j+1}\frac{\mathrm{d}^{j+2}}{\mathrm{d}t^{j+2}}\phi_0(t)\right) \\ &\quad + \frac{1}{(2k-1)!!}\left(\sum_{j=1}^{k-1} c_4 t^j\frac{\mathrm{d}^j}{\mathrm{d}t^j}\phi_1(t) + c_5 t^{j+1}\frac{\mathrm{d}^{j+1}}{\mathrm{d}t^{j+1}}\phi_1(t)\right)\end{aligned}$$

with some constants $c_1, ..., c_5$. From this formula we readily obtain that

$$\lim_{t\to 0}\frac{\partial}{\partial t}u(t,x) = \lim_{t\to 0}\phi_1(t) = u_1(x)\,,$$

as well as

$$\lim_{t\to 0}u(t,x) = \lim_{t\to 0}\phi_0(t) = u_0(x)\,.$$

Solution 1.55 (Exercise 1.41). We recall that the energy is given by

$$\begin{aligned}E(t) :&= \frac{1}{2}\int_\Omega \left(|u_t(t,x)|^2 + |\nabla u(t,x)|^2\right)\,\mathrm{d}x \\ &= \int_\Omega (u_t u_{tt} + \nabla u_t\cdot\nabla u)\,\mathrm{d}x = \int_\Omega u_t(u_{tt} - \Delta u)\,\mathrm{d}x + \int_{\partial\Omega}\frac{\partial u}{\partial\nu}\frac{\partial u}{\partial t}\mathrm{d}\sigma \\ &= \int_{\partial\Omega}\frac{\partial u}{\partial\nu}\frac{\partial u}{\partial t}\mathrm{d}\sigma\,.\end{aligned}$$

Hence, for $-\frac{\partial u}{\partial t} = \frac{\partial u}{\partial \nu}$ the last integral is non-positive which implies energy decay.

Solution 1.56 (Exercise 1.42). The energy we are looking for is given by

$$E(t) := \frac{1}{2}\int_\Omega \left(\frac{\rho}{c^2}|\phi_t(t,x)|^2 + \rho|\nabla\phi(t,x)|^2\right)\,\mathrm{d}x + \frac{1}{2}\int_{\partial\Omega}\left(k\delta^2(t,y) + m\delta_t^2(t,y)\right)\,\mathrm{d}\sigma\,.$$

Indeed, this leads to

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}E(t) &= \int_\Omega \rho\phi_t\left(\frac{1}{c^2}\phi_{tt} - \Delta\phi\right)\,\mathrm{d}x + \int_{\partial\Omega}\left(\rho\frac{\partial\phi}{\partial\nu}\frac{\partial\phi}{\partial t} + k\delta\delta_t + m\delta_t\delta_{tt}\right)\,\mathrm{d}\sigma\\ &= \int_{\partial\Omega}\delta_t\left(\rho\phi_t + k\delta - d\delta_t - k\delta - \rho\phi_t\right)\,\mathrm{d}\sigma\\ &= \int_{\partial\Omega} -d\delta_t^2\,\mathrm{d}\sigma \le 0\,.\end{aligned}$$

Appendix A

A reminder of Lebesgue integration theory

In this section we will repeat and state basic results from measure and integration theory that are necessary for the definition of the *Lebesgue integral* . We restrict to the case of $\mathbb{R}^d$ and subsets thereof, which is sufficient for our purposes. For more details we refer to [15].

A.1 The Lebesgue measure

In order to define the Lebesgue integral, we first need to define the Lebesgue measure in $\mathbb{R}^d$. Intuitively, the Lebesgue measure of a given set should be thought of as its classical "volume". It turned out, however, that a meaningful "volume function" cannot in general be defined on the power set of $\mathbb{R}^d$, leading to the following definition.

Definition A.1. A family Σ of subsets $A \subset \mathbb{R}^d$ is called a **σ-algebra** if the following conditions are fulfilled:

(S1) $\mathbb{R}^d \in \Sigma$;

(S2) if $A \in \Sigma$, then also $\mathbb{R}^d \setminus A \in \Sigma$;

(S3) for any countable family $(A_n)_{n\in\mathbb{N}}$ of elements of Σ, there also holds $\bigcup_{n=1}^{\infty} A_n \in \Sigma$.

We can readily establish the following.

Proposition A.2. *Let Σ be a σ-algebra of $\mathbb{R}^d$. Then the following assertions hold.*

1) $\emptyset \in \Sigma$*;*

2) $A, B \in \Sigma \Rightarrow A \setminus B \in \Sigma$*;*

3) $\forall n \in \mathbb{N} : A_n \in \Sigma \Rightarrow \bigcap_{n=1}^{\infty} A_n \in \Sigma$.

Proof. 1) follows from (S2) since $\mathbb{R}^d \setminus \mathbb{R}^d = \emptyset \in \Sigma$.
2) We write $A\setminus B = \mathbb{R}^d\setminus\{(\mathbb{R}^d\setminus A)\cup B\}$. Hence, by definition of σ-algebra we conclude that $A\setminus B \in \Sigma$.
3) Since

$$\begin{aligned} \mathbb{R}^d \setminus \bigcup_{n=1}^{\infty} (\mathbb{R}^d \setminus A_n) &= \bigcap_{n=1}^{\infty} \mathbb{R}^d \setminus (\mathbb{R}^d \setminus A_n) \\ &= \bigcap_{n=1}^{\infty} A_n \ , \end{aligned} \tag{A.1}$$

the statement follows from (S2) and (S3) in Definition A.1. □

Example A.3. A trivial example of a σ-algebra is given by $\Sigma = \{\emptyset, \mathbb{R}^d\}$.

Definition A.4. Let Σ be a σ-algebra. A function $\mu : \Sigma \to \mathbb{R}^+ \cup \{\infty\}$ is called a (positive, σ-additive) **measure** if the following conditions are fulfilled:

(i) $\mu(\emptyset) = 0$;

(ii) $\forall A \in \Sigma : \mu(A) \geq 0$;

(iii) $A_n \in \Sigma,\ A_l \cap A_k = \emptyset$ if $k \neq l : \mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n)$.

We can now state an important theorem which characterizes the *Lebesgue measure*.

Theorem A.5. *There exists a σ-algebra Σ in $\mathbb{R}^d$ and a measure μ, called **Lebesgue measure**, with the following properties:*

(L1) *if $A \subset \mathbb{R}^d$ is an open set, then $A \in \Sigma$;*

(L2) *for every $A \in \Sigma$ and every $\varepsilon > 0$ there exists an open set $B \in \mathbb{R}^d$ such that*

$$\mu(B \setminus A) \leq \varepsilon;$$

(L3) *if $A \subset B$, $B \in \Sigma$ with $\mu(B) = 0$, then $A \in \Sigma$ and $\mu(A) = 0$;*

(L4) *if $Q := \{x \in \mathbb{R}^d \mid a_k \leq x_k \leq b_k,\ k = 1, \ldots, n\}$, then $Q \in \Sigma$ and*

$$\mu(Q) = \prod_{k=1}^{n} (b_k - a_k);$$

(L5) *μ is invariant under translations, i.e., if $A \in \Sigma$, then $A + x := \{y \in \mathbb{R}^d \mid y - x \in A\} \in \Sigma$ and $\mu(A) = \mu(A + x)$.*

Definition A.6. If $A \in \Sigma$ with Σ being the σ-algebra of Theorem A.5, then A is called **Lebesgue measurable** (or **measurable** for short) and $\mu(A)$ is called the **measure of** A. Furthermore, sets $A \in \Sigma$ with $\mu(A) = 0$ are called **null sets**.

We say that a property depending on $x \in \mathbb{R}^d$ – for example, an identity – is fulfilled for **almost every** (or **a.e.** for short) x if the set of the x which said property is *not* fulfilled has zero Lebesgue measure. Of course, this can be restricted to open subsets of $\mathbb{R}^d$, if needed.

A.2 Measurable functions and the Lebesgue integral

In this paragraph it is sometimes convenient to work with the extended real line $\overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$ with its standard arithmetic operations.

Definition A.7 (Measurable functions)**.** Let $A \in \Sigma$ be a measurable set, and let $f : A \to \overline{\mathbb{R}}$ be a function. Then f is called **measurable** if the set

$$M_a := \{x \in A \mid f(x) > a\}\,, \tag{A.2}$$

is measurable for all $a \in \mathbb{R}$, i.e., $M_a \in \Sigma$.

Furthermore, if $g : A \to \mathbb{C}$ with

$$g(x) = g_1(x) + i g_2(x)\,, \tag{A.3}$$

for real-valued functions g_1, g_2, then g is called measurable if $g_1 : A \to \mathbb{R}$ and $g_2 : A \to \mathbb{R}$ are measurable.

Remark A.8. If $f : A \to \overline{\mathbb{R}}$ is measurable, then so are the sets $\{x \in A \;\big|\; f(x) > a - \frac{1}{k},\ k \in \mathbb{N}\}$. Hence, by Definition A.1 and Proposition A.2, the sets $\{x \in A \mid f(x) \geq a\}$, $\{x \in A \mid f(x) < a\}$, and $\{x \in A \mid f(x) \leq a\}$ are also measurable.

The following theorem is important since it shows that measurability is a robust property. This, on the other hand, will imply that the class of measurable functions is in general very big. To formulate the result we note that if $f : A \to \overline{\mathbb{R}}$ is some function,

$$f_+(x) := \begin{cases} f(x) & \text{if } f(x) \geq 0, \\ 0 & \text{otherwise,} \end{cases} \tag{A.4}$$

is called the positive part of f and

$$f_-(x) := \begin{cases} f(x) & \text{if } f(x) \leq 0, \\ 0 & \text{otherwise,} \end{cases} \tag{A.5}$$

the negative part of f.

Theorem A.9. *(a) If $f : A \to \overline{\mathbb{R}}$ is measurable, then $|f|$, f^+, and f^- are also measurable.*

(b) If $f, g : A \to \mathbb{R}$ are measurable, then $f \cdot g$ and $f + g$ are also measurable.

(c) If $(f_n)_{n\in\mathbb{N}}$ is a sequence of real-valued, measurable functions, then $\sup_n f_n(x)$, $\inf_n f_n(x)$, $\limsup_n f_n(x)$, and $\liminf_n f_n(x)$ are also measurable.

(d) If $f : \mathbb{R} \to \mathbb{R}$ is continuous and $g : \mathbb{R}^d \to \mathbb{R}$ is measurable, then $f \circ g$ is also measurable.

Example A.10. Let $A = [0, 1] \subset \mathbb{R}$ and $f(x) = x^n$, $x \in [0, 1]$. We see that $\limsup_{n\to\infty} f_n(x) = \liminf_{n\to\infty} f_n(x) = 0$ for $x \in [0, 1)$ while $\limsup_{n\to\infty} f_n(x) = \liminf_{n\to\infty} f_n(x) = 1$ for $x = 1$. Hence, by Theorem A.9, the function

$$f(x) = \begin{cases} 0 & \text{for} \quad x \in [0, 1), \\ 1 & \text{for} \quad x = 1, \end{cases} \tag{A.6}$$

is measurable.

As we will see shortly, the Lebesgue integral of a function is defined via a limiting process using a certain class of simple functions.

Definition A.11. Let $A_1, \ldots, A_m \in \mathbb{R}^d$ be measurable sets, and let $a_1, \ldots, a_m \in \mathbb{R}$. Then

$$s(x) = \sum_{k=1}^{m} a_k \chi_{A_k} \tag{A.7}$$

is called a (measurable) step function.

The next theorem is crucial.

Theorem A.12. *Let $f : A \to \overline{\mathbb{R}}$ be measurable. Then there exists a sequence of step functions $(s_k)_{k\in\mathbb{N}}$ such that $s_k \to f$ pointwise. Furthermore, whenever $f : A \to [0, \infty]$, the sequence $(s_k)_{k\in\mathbb{N}}$ can be chosen to be monotonically increasing.*

With Theorem A.12 at hand we can now define the **Lebesgue integral**. For a step function $s(x)$ we define

$$\int_A s(x)\, \mathrm{d}x := \sum_{k=1}^{m} a_k \mu(A_k)\,. \tag{A.8}$$

If $f : A \to [0, \infty]$ is measurable, then we set

$$\int_A f(x)\, \mathrm{d}x := \sup \int_A s(x)\, \mathrm{d}x\,, \tag{A.9}$$

where the supremum is taken over all step functions such that $s(x) = 0$ for $x \notin A$ and $0 \leq s(x) \leq f(x)$ for all $x \in A$. For a general function $f : A \to \overline{\mathbb{R}}$ we set

$$\int_A f(x)\, \mathrm{d}x := \int_A f^+(x)\, \mathrm{d}x - \int_A (-f_-)(x)\, \mathrm{d}x\,, \tag{A.10}$$

and if $g : A \to \mathbb{K}$, then

$$\int_A g(x)\, \mathrm{d}x := \int_A g_1(x)\, \mathrm{d}x + i \int_A g_2(x)\, \mathrm{d}x\,. \tag{A.11}$$

Definition A.13. A measurable function $f : A \to \overline{\mathbb{R}}$ is called **Lebesgue integrable** if

$$\int_A f^+(x)\,\mathrm{d}x + \int_A (-f^-)(x)\,\mathrm{d}x < \infty\,. \tag{A.12}$$

Analogously, a complex-valued measurable function is called Lebesgue integrable if both its real and imaginary parts are integrable.

Theorem A.14. *Let A be a measurable set, and let $f, h : A \to \overline{\mathbb{R}}$ be two measurable functions. Then the following assertions hold.*

1) If $a \leq f(x) \leq b$ for all $x \in A$ and $\mu(A) < \infty$, then

$$a\mu(A) \leq \int_A f(x)\,\mathrm{d}x \leq b\mu(A)\,. \tag{A.13}$$

2) If $f(x) \leq h(x)$ for all $x \in A$ and h is integrable, then f is also integrable and

$$\int_A f(x)\,\mathrm{d}x \leq \int_A h(x)\,\mathrm{d}x\,. \tag{A.14}$$

3) If f, h are integrable, then

$$\int_A (af(x) + bh(x))\,\mathrm{d}x = a\int_A f(x)\,\mathrm{d}x + b\int_A h(x)\,\mathrm{d}x\,. \tag{A.15}$$

4) If $\mu(A) = 0$, then $\int_A f(x)\,\mathrm{d}x = 0$.

Finally, we want to state three important theorems regarding the interchange of integral and limit: the **Monotone Convergence Theorem** A.15, proved by Beppo Levi in 1906 in Turin; **Fatou's Lemma** – our Theorem A.16 –, which goes back to investigations by the mathematician and astronomer Pierre Fatou from 1906; and Lebesgue's **Dominated Convergence Theorem** A.17, proved by Henri Lebesgue in Paris in 1902.

Theorem A.15. *Let $A \subset \mathbb{R}^d$ be measurable, and let $(f_n)_{n\in\mathbb{N}}$ be a sequence of non-negative, monotonically increasing, measurable functions with $f_n : A \to [0, \infty]$. Then*

$$\lim_{n\to\infty} \int_A f_n(x)\,\mathrm{d}x = \int_A \lim_{n\to\infty} f_n(x)\,\mathrm{d}x\,. \tag{A.16}$$

Theorem A.16 (Fatou's lemma)**.** *Let $A \subset \mathbb{R}^d$ be measurable and $(f_n)_{n\in\mathbb{N}}$ a sequence of non-negative, measurable functions with $f_n : A \to [0, \infty]$. Then*

$$\int_A \liminf_{n\to\infty} f_n(x)\,\mathrm{d}x \leq \liminf_{n\to\infty} \int_A f_n(x)\,\mathrm{d}x\,. \tag{A.17}$$

Theorem A.17. *Let $A \subset \mathbb{R}^d$ be measurable and let $(f_n)_{n\in\mathbb{N}}$ be a sequence of measurable functions $f_n : A \to \mathbb{K}$ that converge pointwise. If there exists an integrable function $h : A \to [0, \infty]$ such that $|f_n(x)| \leq |h(x)|$ for all $x \in A$, then*

$$\int_A \lim_{n\to\infty} f_n(x)\,\mathrm{d}x = \lim_{n\to\infty} \int_A f_n(x)\,\mathrm{d}x\,. \tag{A.18}$$

This implies, in particular, that $\lim_{n\to\infty} f_n(x)$ is integrable.

Example A.18. Let $\Omega = [0, 1] \subset \mathbb{R}$ and let $f_n(x) = x^n$, $n \in \mathbb{N}$, $x \in [0, 1]$. We first observe that Ω is a measurable set and that $f_n(x)$ is continuous on Ω. Furthermore, we see that $|f_n(x)| \leq 1$. Since $h(x) = 1$ is an integrable function, we can employ Theorem A.17 to calculate

$$\begin{aligned} \lim_{n\to\infty} \int_0^1 x^n\,\mathrm{d}x &= \int_0^1 \lim_{n\to\infty} x^n\,\mathrm{d}x \\ &= \int_{\Omega\setminus\{1\}} 0\,\mathrm{d}x + \int_{\{1\}} 1\,\mathrm{d}x = 0\,. \end{aligned} \tag{A.19}$$

Here we used that $\lim_{n\to\infty} x^n = 0$ for all $x \in [0,1)$ and the fact that $A = \{1\}$ is a set of measure zero. On the other hand, we have (using the Riemann integral)

$$\begin{aligned} \lim_{n\to\infty} \int_0^1 x^n \, \mathrm{d}x &= \lim_{n\to\infty} \left[\frac{1}{n+1} x^{n+1} \right]_0^1 \\ &= \lim_{n\to\infty} \frac{1}{n+1} = 0 \, . \end{aligned} \tag{A.20}$$

A.3 Some important theorems

We will state a generalization to Lebesgue's integral of the fundamental theorem of Riemann integration. For this, recall that a function $f : [a,b] \to \mathbb{K}$ is called **absolutely continuous** on $[a,b]$ if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that, for any finite family of pairwise disjoint, open subintervals (a_k, b_k), $k = 1, 2, \ldots, n$,

$$\sum_{k=1}^{n} (b_k - a_k) < \delta \quad \text{implies} \quad \sum_{k=1}^{n} |f(b_k) - f(a_k)| < \varepsilon \, . \tag{A.21}$$

Note that every absolutely continuous function is uniformly continuous (the converse statement is false).

Theorem A.19 (Fundamental theorem of calculus)**.** *Let $[a,b] \subset \mathbb{R}$ be some interval, and let $f : [a,b] \to \mathbb{K}$ some integrable function on (a,b), i.e., $f \in L^1(a,b)$. If we define*

$$F(x) := \int_a^x f(t) \, \mathrm{d}t, \tag{A.22}$$

then $F(x)$ is an absolutely continuous function. Furthermore, $F(x)$ is differentiable almost everywhere with $F'(x) = f(x)$ for a.e. $x \in [a,b]$.

On the other hand, if $f \in H^1(a,b)$, then there exists a constant $c \in \mathbb{R}$ such that

$$f(x) = c + \int_a^x f'(y) \, \mathrm{d}y \tag{A.23}$$

for a.e. $x \in [a,b]$.

The following result is sometimes referred to as the *absolute continuity of the Lebesgue integral*.

Theorem A.20. *Let f be an integrable function on the measurable set A. For every $\varepsilon > 0$ there exists $\delta > 0$ which depends on ε and f only, such that for all measurable sets $B \subset A$ with $\mu(B) < \delta$, one has*

$$\int_B |f| \, \mathrm{d}x < \varepsilon \, . \tag{A.24}$$

Another important result is **Fubini's Theorem**, proved in Genua in 1907 by Guido Fubini.

Theorem A.21. *Let f be an integrable function on $\mathbb{R}^{n+m}$, i.e.,*

$$\int_{\mathbb{R}^{n+m}} |f(x,y)| \, \mathrm{d}x\mathrm{d}y < \infty \, . \tag{A.25}$$

Then the following holds:

(a) $f(\cdot, y) \in L^1(\mathbb{R}^d)$ *for almost every* $y \in \mathbb{R}^m$*;*

(b) $f(x, \cdot) \in L^1(\mathbb{R}^m)$ *for almost every* $x \in \mathbb{R}^d$*;*

(c) $\int_{\mathbb{R}^d} f(x,y) \, \mathrm{d}x \in L^1(\mathbb{R}^m)$*;*

(d) $\int_{\mathbb{R}^m} f(x,y) \, \mathrm{d}y \in L^1(\mathbb{R}^d)$*;*

(e) one has

$$\int_{\mathbb{R}^{n+m}} |f(x,y)| \, \mathrm{d}x\mathrm{d}y = \int_{\mathbb{R}^m} \left(\int_{\mathbb{R}^d} |f(x,y)| \, \mathrm{d}x \right) \, \mathrm{d}y$$
$$= \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^m} |f(x,y)| \, \mathrm{d}y \right) \, \mathrm{d}x \, .$$

Employing Fubini's Theorem, one can establish an important integration by parts formula.

Theorem A.22 (Integration by parts)**.** *Let $\Omega \subset \mathbb{R}^d$ be some open set. For $\varphi, \psi \in C_0^\infty(\Omega)$ one has*

$$\int_\Omega \varphi(x)(\partial_{x_j}\psi)(x) \, \mathrm{d}x = -\int_\Omega (\partial_{x_j}\varphi)(x)\psi(x) \, \mathrm{d}x \tag{A.26}$$

for all $j = 1, \ldots, n$.

Another important result which will be used frequently in the text is the so-called Fundamental Lemma of Calculus of Variations, also called the **du Bois-Reymond Lemma**, after Paul du Bois-Reymond, who proved a special case of it in Tübingen in 1879.

For this, recall that a function $f : \Omega \to \mathbb{K}$ is called **locally integrable** if $f\varphi \in L^1(\Omega)$ for all test functions $\varphi \in C_0^\infty(\Omega)$. We define

$$L^1_{loc}(\Omega) := \{ f : \Omega \to \mathbb{K} : f \text{ is locally integrable} \} \, . \tag{A.27}$$

Lemma A.23. *Let $f \in L^1_{loc}(\Omega)$ be some locally integrable function. Then $f = 0$ almost everywhere if and only if*

$$\int_\Omega f\varphi \, \mathrm{d}x = 0 \, , \quad \forall \varphi \in C_0^\infty(\Omega) \, . \tag{A.28}$$

A.4 Lebesgue's differentiation theorem

The well-known fundamental theorem of calculus states that if $-\infty \leq a < b \leq \infty$, and if the function $f : [a,b] \to \mathbb{R}$ is continuous, then

$$F : [a,b] \ni x \mapsto \int_a^x f(s) \, \mathrm{d}s \in \mathbb{R}$$

is differentiable and its derivative $F'(x)$ at each $x \in \mathbb{R}$ is precisely $f(x)$. In other words,

$$f(x) = \lim_{h \to 0} \frac{F(x+h) - F(x)}{h} = \lim_{h \to 0} \frac{\int_x^{x+h} f(s) \, \mathrm{d}s}{h}$$

or, equivalently,

$$f(x) = \lim_{h \to 0} \frac{F(x+h) - F(x-h)}{2h} = \lim_{h \to 0} \frac{\int_{x-h}^{x+h} f(s) \, \mathrm{d}s}{2h}. \tag{A.29}$$

Observe that this formula is a local assertion, which can be formulated and in fact holds as long as f is Riemann integrable on neighborhoods of x: accordingly, (A.29) also holds for continuous functions $f : I \to \mathbb{R}$, for general intervals I, possibly $I = \mathbb{R}$.

In this course we are however mostly dealing with Lebesgue integrals, due to a number of reasons: above all, the possibility offered by Lebesgue theory to integrate over more general domains, and the fact that Lebesgue integrals lie at the core of the theory of Sobolev spaces. Does the fundamental theorem of calculus extend to an assertion for Lebesgue integrable functions? This was among the few fundamental issues left open in Henri Lebesgue's groundbreaking doctoral thesis at the Sorbonne in Paris, in which in 1902 he had introduced the measure theory that is now called after him. However, Lebesgue himself was finally able to show in 1904 that yes, the fundamental theorem of calculus does extend to Lebesgue integrable function, in the sense made precise by the following assertion, the so-called **Lebesgue Differentiation Theorem**.

Theorem A.24. *Let $f \in L^1_{loc}(\mathbb{R})$. Then*

$$\lim_{h\to 0} \frac{\int_{x-h}^{x+h} |f(s) - f(x)| \, \mathrm{d}s}{2h} = 0 \qquad \textit{for a.e. } x \in \mathbb{R}, \tag{A.30}$$

and in particular

$$f(x) = \lim_{h\to 0} \frac{\int_{x-h}^{x+h} f(s) \, \mathrm{d}s}{2h} \qquad \textit{for a.e. } x \in \mathbb{R}. \tag{A.31}$$

We stress that given a function $f \in L^1(I)$, where $I \subsetneq R$ is an interval, then we can identify f with its extension by 0 to the whole $\mathbb{R}$: as this extension is of class $L^1_{loc}(\mathbb{R})$, the Lebesgue Differentiation Theorem holds for such f, too.

While the assertion looks harmless, its proof is not, as it is based on the following deep result. To begin with for every $f \in L^1_{loc}(\mathbb{R}^d)$, we introduce the function Mf defined a.e. by

$$(Mf)(x) := \sup_{\varepsilon>0} \left(\frac{1}{|B_\varepsilon(x)|} \int_{B_\varepsilon(x)} |f(y)| \, \mathrm{d}y \right), \qquad f \in L^1_{loc}(\mathbb{R}^d),\ x \in \mathbb{R}^d. \tag{A.32}$$

Before we proceed, we have to recall a well-known result from measure theory known as **Vitali's covering lemma**. A proof can be found at `https://en.wikipedia.org/wiki/Vitali_covering_lemma`.

Lemma A.25. *Let $B_{r_1}(x_1), \dots, B_{r_n}(x_n)$ be a finite collection of balls with radii $r_1, \dots, r_n > 0$ and centred at the points $x_1, \dots, x_n \in \mathbb{R}^d$. Then there exists a finite sub-collection of disjoint balls $\{B_{r_l}(x_l)\}$ such that*

$$\bigcup_{j=1}^{n} B_{r_j}(x_j) \subset \bigcup_{\textit{finite sub-collection}} 3B_{r_l}(x_l) \, , \tag{A.33}$$

where $3B_r(x) := \{y \in \mathbb{R}^d : \ \frac{1}{3}y \in B_r(x)\}$.

The following goes under the name of **Hardy–Littlewood Maximal Inequality**. Additional information you again find at `https://en.wikipedia.org/wiki/Vitali_covering_lemma`.

Theorem A.26. *Let $f \in L^1(\mathbb{R}^d)$ be given. Then there exists a constant $c > 0$ that only depends on d such that*

$$\mu\left(\{x \in \mathbb{R}^d : (Mf)(x) > t\}\right) \le \frac{c}{t} \|f\|_{L^1(\mathbb{R}^d)}. \tag{A.34}$$

Proof. We set $A_t := \{x \in \mathbb{R}^d : (Mf)(x) > t\}$ and recall that, by the inner regularity of the Lebesgue measure, one has

$$\mu(A_t) = \sup\{\mu(K) : K \subset A_t \text{ with } K \text{ compact}\} \, .$$

It hence remains to prove that (A.34) holds for any such set K. By definition, for any $x \in K$ there exists a ball $B_r(x)$ with some $r > 0$ such that

$$\frac{1}{|B_r(x)|} \int_{B_r(x)} |f(y)| \, \mathrm{d}y > t \, .$$

Since K is compact, there exists a finite number of balls $B_{r_j}(x_j)$ with $j = 1, \dots, N$ that cover K. Employing Lemma A.25, we can pick from those a finite number of *disjoint* balls $B_{r_l}(x_l)$ with $l = 1, \dots, \hat{N} \le N$ such that

$$\begin{aligned}
\mu(K) &\le \sum_{j=1}^{N} \mu(B_{r_j}(x_j)) \\
&\le c \sum_{l=1}^{\hat{N}} \mu(B_{r_l}(x_l)) \\
&\le \frac{c}{t} \sum_{l=1}^{\hat{N}} \int_{B_{r_l}(x_l)} |f(y)| \, \mathrm{d}y \\
&\le \frac{c}{t} \|f\|_{L^1(\mathbb{R}^d)} \, ,
\end{aligned}$$

for some constant $c > 0$ that only depends on the dimension d. □

Theorem A.24 was later generalized by Lebesgue and further mathematicians. In this course we repeatedly use the following extension to higher dimensional domains.

Theorem A.27. *Let $f \in L^1_{loc}(\mathbb{R}^d)$. Then*

$$\lim_{\varepsilon\to 0} \frac{1}{|B_\varepsilon(x)|} \int_{B_\varepsilon(x)} |f(s) - f(x)| \,\mathrm{d}s = 0 \qquad \textit{for a.e. } x \in \mathbb{R}^d, \tag{A.35}$$

and in particular

$$f(x) = \lim_{\varepsilon\to 0} \frac{1}{|B_\varepsilon(x)|} \int_{B_\varepsilon(x)} f(s) \,\mathrm{d}s \qquad \textit{for a.e. } x \in \mathbb{R}^d. \tag{A.36}$$

Proof. We remark that it is enough to prove (A.35) since (A.36) follows readily from this.

We introduce the function

$$f^*(x) := \limsup_{\varepsilon\to 0} \left(\frac{1}{|B_\varepsilon(x)|} \int_{B_\varepsilon(x)} |f(s) - f(x)| \,\mathrm{d}s \right) , \quad x \in \mathbb{R}^d . \tag{A.37}$$

It remains to prove that $f^*(x) = 0$ for almost every $x \in \mathbb{R}^d$. Assume we know that, for every function $f \in L^1_{loc}(\mathbb{R}^d)$,

$$\mu\left(\{x \in \mathbb{R}^d : f^*(x) > t\}\right) \le \frac{c}{t}\|f\|_{L^1(\mathbb{R}^d)} \qquad \text{for all } t > 0 \tag{A.38}$$

holds for some constant $c > 0$ only depending on d. Then, since $C_c(\mathbb{R}^d)$ is dense in $L^1(\mathbb{R}^d)$, we find for every $\epsilon > 0$ a function $h \in C_c(\mathbb{R}^d)$ such that $\|f - h\|_{L^1(\mathbb{R}^d)} < \epsilon$. Furthermore, for all $t > 0$

$$\begin{aligned}
\mu\left(\{x \in \mathbb{R}^d : f^*(x) > t\}\right) &= \mu\left(\{x \in \mathbb{R}^d : (f-h)^*(x) > t\}\right) \\
&\le \frac{c}{t}\|f - h\|_{L^1(\mathbb{R}^d)} \\
&\le \frac{\epsilon c}{t} ,
\end{aligned}$$

and since $\epsilon > 0$ is arbitrary we conclude $\mu\left(\{x \in \mathbb{R}^d : f^*(x) > t\}\right) = 0$ for all $t > 0$; here we used the fact that $h^*(x) = 0$ for all $x \in \mathbb{R}^d$ whenever $h \in C_c(\mathbb{R}^d)$. Consequently, since

$$\left\{x \in \mathbb{R}^d : f^*(x) > 0\right\} = \bigcup_{n=1}^{\infty} \left\{x \in \mathbb{R}^d : f^*(x) > \frac{1}{n}\right\},$$

we finally obtain $\mu(\{x \in \mathbb{R}^d : f^*(x) > 0\}) = 0$, which proves the statment.

It remains to establish (A.38). To this aim, we first observe that

$$f^*(x) \le (Mf)(x) + |f(x)|$$

for almost every $x \in \mathbb{R}^d$, where Mf is the function in (A.32). This yields

$$\begin{aligned}
\{x \in \mathbb{R}^d : f^*(x) > t\} &\subset \{x \in \mathbb{R}^d : (Mf)(x) + |f(x)| > t\} \\
&\subset \{x \in \mathbb{R}^d : (Mf)(x) > t/2\} \cup \{x \in \mathbb{R}^d : |f(x)| > t/2\} .
\end{aligned}$$

Hence, the statement follows using the Hardy–Littlewood Maximal Inequality and

$$\mu(\{x \in \mathbb{R}^d : |f(x)| > t/2\}) \le \frac{2}{t}\|f\|_{L^1(\mathbb{R}^d)} \qquad \text{for all } t > 0.$$

This concludes the proof. □

Clearly, (A.36) is weaker than (A.35): the points x for which the stronger assertion (A.35) hold are called **Lebesgue points** of f.

In analogy with the one-dimensional case, the right hand side of (A.35) is sometimes called the *derivative* of the function

$$A \mapsto \int_A f(x) \,\mathrm{d}x$$

mapping each measurable set of $\mathbb{R}^d$ to the integral of f on it.

Appendix B

Lipschitz domains and the formula of Gauss-Green

In this section we present a very important technical tool we use frequently in the text. This formula of Gauss-Green is nothing else than an n-dimensional generalization of the well-known integration by parts we know from elementary calculus. In addition, it also holds for functions that are not continuously differentiable in the classical sense but in a weak sense; more explicitly, it holds also for Sobolev functions. Note that the presentation follows [41] to which we refer for more details.

Unfortunately, it is not possible to establish an d-dimensional integration by parts formula for functions defined on an arbitrary domain $\Omega \subset \mathbb{R}^d$, at least as long as the functions are not assumed to vanish on the boundary $\partial\Omega$. In other words, we have to assume some regularity for the boundary $\partial\Omega$ which leads us to the important notion of a *Lipschitz domain*, or of C^k-domain. This allows us, in a certain sense, to measure the smoothness of a domain's boundary.

Definition B.1. Let $\Omega \subset \mathbb{R}^d$ be a (not necessarily bounded) domain with boundary $\partial\Omega$. Then we call Ω a **Lipschitz domain** if for every boundary point $y \in \partial\Omega$ there exists an affine hyperplane E through y, parameters $r, h > 0$ and a Lipschitz continuous function $g : B_r(y) \cap E \to \mathbb{R}$ and such that the following holds:

- $\Omega \cap C = \{x + t\nu : \ x \in B_r(y) \cap E \ , \ -h < t < g(x)\}$,
- $\partial\Omega \cap C = \{x + t\nu : \ x \in B_r(y) \cap E \ , \ t = g(x)\}$.

Here $C := \{x + t\nu : \ x \in B_r(y) \cap E \ , \ -h < t < h\}$ and ν is a unit normal vector to E.

Intuitively, this means Lipschitz domains of $\mathbb{R}^d$ have $(d-1)$-dimensional boundary that – *at least locally* – is the graph of a Lipschitz continuous function and lies on only one side of its boundary. Likewise, if Ω is locally the graph of a k-times continuously differentiable function, then Ω is said to have C^k-boundary. Whenever $k = \infty$, we say the boundary is smooth.

Example B.2. Domains with smooth boundary are Lipschitz domains: let us exemplify this with the circle $B_R(0)$ of radius $R > 0$. In that case we can pick, without loss of generality, the point $(0, R) \in \partial\Omega$. Hence, E is given by $y - R = 0$, $g(x, y) := -R + \sqrt{R^2 - x^2}$, $n = (0, -1)^T$, $r = R/2$ and $h = R - R\sqrt{3/4}$.

The same argument shows that the complement of $B_R(0)$ is a Lipschitz domain in its own right. This shows in particular that a Lipschitz domain need not be bounded.

Example B.3. How little smooth can a Lipschitz domain be? Another Lipschitz domain is given by the unit square $\Omega := (0, 1) \times (0, 1) \subset \mathbb{R}^2$. Then Ω is a Lipschitz domain. To see this, assume that you have (for example) the boundary point $(1/2, 0) \in \partial\Omega$. In that case the hyperplane is given by $y = 0$ and we choose $n = (0, -1)^T$, $g(x) \equiv 0$ as well as $h = 1/2$ and $r = 1/4$. On the other hand, if we consider the "corner point" $(0, 0)$, then E is given by $y + x = 0$, $g(x, y) := -\sqrt{x^2 + y^2}$, $n = (-1, -1)^T$ and $h = r = \frac{1}{4}$.

Example B.4. More generally, bounded convex domains of $\mathbb{R}^d$ are Lipschitz. The picture is intuitively clearer for planar domains: each polygon in $\mathbb{R}^2$ is a Lipschitz domain.

An elementary example of a domain which is not a Lipschitz domain is $B_1(0) \setminus \{0\}$ (for the same reason as below). But this is not overly interesting when it comes to integration theory, since the difference between $B_1(0)$ and $B_1(0) \setminus \{0\}$ is only a Lebesgue zero set.

A slightly more refined non-example is

$$B_R(0) \setminus \{(x,y) \in \mathbb{R}^2 : x > 0 \text{ and } y = 0\}$$

which is connected and even star-shaped (but not convex). However, it is not a Lipschitz domain since the domain is on *both* sides of $\{(x,y) \in \mathbb{R}^2 : x > 0 \text{ and } y = 0\}$ which contradicts Definition B.1.

Now, whenever Ω is bounded, the boundary $\partial\Omega$ is a compact set. Also, from Definition B.1 we see that we can cover $\partial\Omega$ with a finite number of "cylinders" U_i, i.e.,

$$\partial\Omega \subset \bigcup_{i=1}^{M} U_i$$

for some number $M \in \mathbb{N}$ and where each U_i can be parametrized as $U_i = V_i \times (-h_i, +h_i)$ with $V_i \in \mathbb{R}^{n-1}$ some open subset. There now exists a so-called *partition of unity* which is a family of C^∞-functions $\{\eta_j\}_{j=1}^M$, $\eta_j \in C_0^\infty(U_j)$, such that

$$\sum_{j=1}^{M} \eta_j(x) = 1 \,, \qquad \forall x \in \partial\Omega \,.$$

Using this, we would like to define a boundary integral of a suitable function $f : \partial\Omega \to \mathbb{R}$ as

$$\int_{\partial\Omega} f \, \mathrm{d}\sigma := \sum_{j=1}^{M} \int_{\partial\Omega \cap U_j} f\eta_j \, \mathrm{d}\sigma \tag{B.1}$$

which means that we have to specify what we understand by $\int_{\partial\Omega\cap U_j} f\eta_j \, \mathrm{d}\sigma$. Now, since $U_i = V_i \times (-h_i, +h_i)$ and since $\partial\Omega$ can be parametrized as the trace of a Lipschitz continuous function $g_i : V_i \to \mathbb{R}$ (compare with Definition B.1), we define

$$\int_{\partial\Omega\cap U_j} f\eta_j \, \mathrm{d}\sigma := \int_{V_i} (f\eta_j)(x)\sqrt{1 + |\nabla g_j(x)|^2} \, \mathrm{d}x \,. \tag{B.2}$$

The right-hand side is now a standard Lebesgue integral which makes sense whenever $f(x) := f(x, g(x))$ is a measurable function on V_i. Note also that, according to Rademacher's Theorem, see [18, § 5.8], $g_i : V_i \to \mathbb{R}$ is indeed differentiable almost everywhere since it is Lipschitz continuous; in addition, Rademacher's Theorem asserts that $\|\nabla g_j\|_{L^\infty(V_j)} \leq L_j$ with $L_j > 0$ being the Lipschitz constant of g_j.

Remark B.5. Note that the factor $\sqrt{1 + |\nabla g_j(x)|^2}$ in (B.2) is important since it guarantees that the boundary integral is independent of the parametrization of the boundary as introduced in Def. B.1.

Also, having introduced a boundary integral, it is now possible to introduce the space $L^2(\partial\Omega)$. More explicitly, $L^2(\partial\Omega)$ shall exist of all measurable functions $f : \partial\Omega \to \mathbb{R}$ (with measurable as defined above) such that

$$\|f\|_{L^2(\partial\Omega)}^2 := \int_{\partial\Omega} |f|^2 \, \mathrm{d}\sigma < \infty \,. \tag{B.3}$$

Of course, regarding $L^2(\partial\Omega)$ we consider two functions $f, g : \partial\Omega \to \mathbb{R}$ as identical if $f\eta_j, g\eta_j : V_j \to \mathbb{R}$ agree almost everywhere for all $j = 1, \ldots, M$. All these ideas can be extended without additional complications to all spaces $L^p(\partial\Omega)$, $1 \leq p < \infty$.

Finally, we observe that

$$\nu(x, g_j(x)) := \frac{(\nabla g_j, -1)}{\sqrt{1 + |\nabla g_j|^2}} \in \mathbb{R}^n$$

is well-defined for almost every $x \in V_j$ and hence at almost every boundary point (due to Rademacher's theorem). The vector field $\nu(z) := \nu(x, g_j(x))$ is the so-called **outward normal** at the boundary point $z := (x, g_j(x)) \in \partial\Omega$. (Observe that ν need not exist at *every* point of the boundary of a general Lipschitz

domain: this is clear if we take the simple case of a polygon in $\mathbb{R}^2$, see Example B.4, and think of its vertices.) It can be shown that if Ω is a Lipschitz domain, then

We are now in position to state the formulae of Gauss–Green in their classical version; an extension to an important class of functions will be presented in the Appendix E.

Theorem B.6. *Let $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain. For all $f, g \in C^1(\overline{\Omega})$ one has*

$$\int_\Omega \partial_j f \cdot g \, \mathrm{d}x = \int_{\partial\Omega} fg \cdot \nu_j \, \mathrm{d}\sigma - \int_\Omega f \cdot \partial_j g \mathrm{d}x. \tag{B.4}$$

[1]

Applying the first Gauss–Green formula to d functions g_i, $i = 1, \ldots, d$, such that $(g_1, \ldots, g_d)^T = \operatorname{grad} h$ and then summing over i yields the following. Here we use the notation

$$\frac{\partial u}{\partial \nu}(z) := \operatorname{grad} u(z) \cdot \nu(z)$$

for the **normal derivative** of u at the boundary point z.

Theorem B.7. *Let $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain. For all $f \in C^2(\overline{\Omega})$ and $h \in C^1(\overline{\Omega})$ one has*

$$\int_\Omega \nabla f \cdot \nabla h \, \mathrm{d}x = \int_{\partial\Omega} f \frac{\partial h}{\partial \nu} \, \mathrm{d}\sigma - \int_\Omega f \Delta h \, \mathrm{d}x. \tag{B.5}$$

Note that, for notational simplicity, one writes $\int_{\partial\Omega} f \, \mathrm{d}\sigma$ instead of $\int_{\partial\Omega} f|_{\partial\Omega} \, \mathrm{d}\sigma$.

Swapping the roles of f, h in Theorem E.22 and adding the corresponding (second) Gauss–Green formulae we obtain the following.

Corollary B.8. *Let $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain. For all $f, h \in C^2(\overline{\Omega})$ one has*

$$\int_\Omega f \Delta h \, \mathrm{d}x - \int_\Omega h \Delta f \, \mathrm{d}x = \int_{\partial\Omega} f \frac{\partial h}{\partial \nu} \, \mathrm{d}\sigma - \int_{\partial\Omega} h \frac{\partial f}{\partial \nu} \, \mathrm{d}\sigma. \tag{B.6}$$

Also, the first Gauss-Green formula (E.6) directly leads to the well-known **Divergence Theorem** of Gauss which relates the divergence of a vector field over a domain with the so-called flow of the field through the surface (meaning boundary) of this domain. Recall that, for a vector field $g := (g_1, \ldots, g_d)^T$ with scalar-valued $g_i \in H^1(\Omega)$ on an open domain $\Omega \subset \mathbb{R}^d$, the divergence $\nabla \cdot g = \operatorname{div} g$ is the scalar-valued function given by

$$(\operatorname{div} g)(x) := \sum_{j=1}^{d} (\partial_j g_j)(x), \qquad x \in \Omega.$$

Theorem B.9. *Let $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain and $g := (g_1, \ldots, g_d)^T$ a vector field with $g_i \in C^1(\overline{\Omega})$. Then,*

$$\int_\Omega \operatorname{div} g \, \mathrm{d}x = \int_{\partial\Omega} g \cdot \vec{\nu} \, \mathrm{d}\sigma \, .$$

Technically speaking, the formula (B.4) was (implicitly) obtained (for $d = 2$) by George Green, back then a Nottinghamshire-based miller without any formal education, in a self-published essay in 1828; because it can be shown to be equivalent to the Divergence Theorem, which (in some special cases) was used by Gauss since 1813, it is now usually called – along with its direct corollaries (E.7) and (E.8) – Gauss–Green Formula.

[1] Here $\nu_j = (\nu)_j$ is the j-th component of the outward normal vector and, for simplicity, we use the notation

$$\partial_j f := \frac{\partial f}{\partial x_j}.$$

Remark B.10. Although (B.1) is a well-defined boundary integral, it is not easy to determine explicitly due to the partition of unity involved. Luckily, useful in most practical applications, it is possible to evaluate the integral in another way. To illustrate this, it is sufficient to assume that $\partial\Omega$ is given as a graph of a Lipschitz continuous function $g : V \subset \mathbb{R}^{n-1} \to \mathbb{R}$, i.e., $\partial\Omega = \{(x, g(x)) : \ x \in V\}$. Now, assume there exists another parametrization of $\partial\Omega$ via a map

$$\Phi : \widetilde{V} \subset \mathbb{R}^{n-1} \to \mathbb{R}^d$$

with $\Phi(\widetilde{V}) = \partial\Omega$ and $\Phi : \widetilde{V} \to \partial\Omega$ bijective. In addition, we assume that the map $\Psi : V \to \widetilde{V}$, $\Psi(x) := \Phi^{-1} \circ (x, g(x))$, is Lipschitz continuous. Then, one has

$$\int_{\partial\Omega} f \, \mathrm{d}\sigma = \int_{\widetilde{V}} (f \circ \Phi)(x) \cdot \sqrt{\det\left(D\Phi^T D\Phi\right)} \, \mathrm{d}x$$

whenever $f : \partial\Omega \to \mathbb{R}$ is measurable and $\int_{\partial\Omega} |f| \, \mathrm{d}\sigma < \infty$. We refer to [41] for the proof of the statement.

Example B.11. Consider the Lipschitz domain $\Omega = (0,1) \times (0,1)$ and the function $f : \Omega \to \mathbb{R}$ defined via $f(x, y) := x$. We note that $f \in H^1(\Omega)$ and since $f \in C^1(\Omega) \cap C(\overline{\Omega})$, the trace of f agrees with the restriction of f to $\partial\Omega$. Using Remark B.10 we directly obtain

$$\int_{\partial\Omega} f\nu_1 \, \mathrm{d}\sigma = \int_0^1 (x \cdot 0) \, \mathrm{d}x + \int_0^1 (x \cdot 0) \, \mathrm{d}x + \int_0^1 (1 \cdot 1) \, \mathrm{d}x = 1 \, .$$

On the other hand, by the formula of Gauss-Green

$$\int_\Omega \partial_x f \, \mathrm{d}x = \int_\Omega \mathrm{d}x = \int_{\partial\Omega} f\nu_1 \, \mathrm{d}\sigma = 1 \, ,$$

which is obviously correct.

In Chapter 6 we make use of an extension of Gauss–Green formulae to *variable domains* that was discovered, for $d = 1$, in Hanover in 1697 by Gottfried Wilhelm Leibniz; and, for $d = 2$ and $d = 3$, in Manchester in 1903 by Osborne Reynolds.

In order to formulate this result, we need the notion of *homeorphism*, which we briefly recall in larger generality than we are actually going to need.

Definition B.12. Let X, Y be two topological spaces.

A function $f : X \to Y$ is called a **homeomorphism** if it is bijective and continuous and if its inverse is continuous, too.

If there is a homeomorphism between X, Y, then they are called **homeomorphic**.

Lemma B.13. *Let $I \subset \mathbb{R}$ be a bounded open interval, and let $\Omega \subset \mathbb{R}^d$ be a bounded domain. Let $(\omega_\epsilon)_{\epsilon \in I}$ be a family of mutually homeomorphic open bounded domains of $\mathbb{R}^d$ with C^1-boundary such that $\omega_\epsilon \subset \Omega$ for all $\epsilon \in I$. Denote by $\partial\omega_\epsilon$ the boundary of ω_ϵ and by $\nu(z)$ the unit normal of the boundary for all $z \in \partial\omega_\epsilon$. If the domains ω_ϵ depend smoothly on ϵ, in the sense that the velocity field $\xi(\cdot, z) : I \to \mathbb{R}^d$ is continuous at each $z \in \partial\omega_\epsilon$. Finally, let $f : I \times \Omega \to \mathbb{R}$ be a continuously differentiable function with bounded values and bounded derivatives. Then*

$$\frac{\mathrm{d}}{\mathrm{d}\epsilon} \int_{\omega_\epsilon} f(\epsilon, x) \mathrm{d}x = \int_{\partial\omega_\epsilon} f(\epsilon, z)\xi(z) \cdot n(z) \mathrm{d}\sigma(z) + \int_{\omega_\epsilon} \frac{\partial f}{\partial \epsilon}(\epsilon, x) \mathrm{d}x \qquad \text{for all } \epsilon \in I. \tag{B.7}$$

The proof is rather technical and heavily relies on ideas of differential geometry and rational mechanics. We refer to [21] for a reasonably self-contained proof.

Observe that if $\omega_\epsilon \equiv \omega$, i.e., if the domain is actually constant, then its velocity field vanishes identically and we recover the usual formula of differentiation under the integral sign, whose exact statement we explicitly formulate for the sake of completeness. In this simpler case, the statement is a direct consequence of Lebesgue's Dominated Convergence Theorem and the assumptions can be slightly relaxed.

Lemma B.14. *Let $I \subset$ be an open interval, and let Ω be a measure space. Let $f : I \times \Omega \to \mathbb{R}$ be a function and assume $f(\epsilon, \cdot)$ to be Lebesgue-integrable for each $\epsilon \in I$; $f(\cdot, x)$ to be differentiable for a.e. $x \in \Omega$; and an integrable function $\theta : \Omega \to \mathbb{R}$ to exist such that $|\frac{\partial f}{\partial \epsilon}(\epsilon, x)| \leq \theta(x)$ for all $\epsilon \in I$ and a.e. $x \in \Omega$.*

Then

$$\frac{\mathrm{d}}{\mathrm{d}\epsilon} \int_\Omega f(\epsilon, x) \mathrm{d}x = \int_\Omega \frac{\partial f}{\partial \epsilon}(\epsilon, x) \mathrm{d}x \qquad \text{for all } \epsilon \in I.$$

Appendix C

Normed spaces and linear operators

Let X, Y be vector spaces over the field $\mathbb{K}$, where $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$. Any mapping $T : X \to Y$ is called an **operator**. We are *not* assuming that the domain of definition – or simply **domain** – $D(T)$ of T is the whole space X. In particular, it could be that $D(T)$ is *not* a vector space; *if*, however, $D(T)$ is a vector space – this is especially the case if $D(T) = X$ – then we call the operator T **linear** if

$$T(x+y) = T(x) + T(y) \quad \text{and} \quad T(\alpha x) = \alpha T(x) \qquad \text{for all } x, y \in D(T) \text{ and all } \alpha \in \mathbb{K}. \tag{C.1}$$

A linear **functional** is a linear operator such that $Y = \mathbb{K}$. . A **non-linear** operator is, accordingly, an operator such the *additivity* condition

$$T(x+y) = T(x) + T(y) \qquad \text{for all } x, y \in D(T)$$

and/or the *homogeneity* condition

$$T(\alpha x) = \alpha T(x) \qquad \text{for all } x \in D(T) \text{ and all } \alpha \in \mathbb{K}$$

are not satisfied.

We usually write Tx instead of $T(x)$, if T is linear.

Example C.1. (1) Strictly speaking, any real-valued function of a real variable is an operator, but historically this notion is reserved for functions between higher-dimensional – or even infinite-dimensional – vectors spaces.

(2) It is a well-known fact of linear algebra that whenever $n, m \in \mathbb{N}$, then each $n \times m$-matrix A can be uniquely associated with a linear operator T_A from $\mathbb{R}^m$ to $\mathbb{R}^d$ via $T_A(x) = Ax$.

(3) Infinite matrices, too, define linear operators in suitable spaces of sequences. For example, if

$$X = Y = \ell^1(\mathbb{N}) := \left\{ x = (x_n)_{n\in\mathbb{N}} : \sum_{n\in\mathbb{N}} |x_n| < \infty \right\}$$

then each diagonal matrix A defines a linear operator.

(4) Let $X = Y = C([0,1])$. and consider the operator

$$T : f \mapsto \frac{\mathrm{d}f}{\mathrm{d}x}$$

in order for Tf to lie in $Y = C([0,1])$, we need $D(T)$ to be a subset of the space of differentiable functions over $[0,1]$. How to take $D(T)$ typically depends on the problem we are considering: let us review some examples.

- If $D(T) = C^1([0,1])$, then the $D(T)$ is a vector space and the usual rules for differentiation of sums and products of functions imply that T is linear.
- If $D(T) = \{f \in C^1([0,1]) : f(0) = \gamma\}$, for some $\gamma \in \mathbb{K}$, then T is linear if and only if $\gamma = 0$.

- If $D(T) = \{f \in C^1([0,1]) : 0 \leq f(x) \leq 1 \text{ for all } x \in [0,1]\}$, then $D(T)$ is not a vector space and hence the conditions for linearity cannot be satisfied, since neither need sums of elements of $D(T)$ lie again in $D(T)$, nor do products of an elements $D(T)$ with a number α.

(5) Consider

$$T_1 : f \mapsto \sin \circ f \quad \text{and} \quad T_2 : f \mapsto f \circ \sin .$$

and pick $X = Y = D(T_1) = D(T_2) = C(\mathbb{R})$. Then T_1, T_2 are well-defined operators: T_2 is linear, but T_1 is not.

In this chapter we will recall some elementary but fundamental notions and results especially concerning *bounded* linear operators between normed spaces: we refer to [9, 29] for further details.

Definition C.2. Let X be a vector space. A **norm** on X is a mapping $\| \cdot \| : X \to \mathbb{R}$ such that for all $x, y \in X$ and all $\lambda \in \mathbb{K}$

(1) $\|x\| \geq 0$, and additionally $\|x\| = 0$ if and only if $x = 0$,

(2) $\|\lambda x\| = |\lambda| \|x\|$, and

(3) $\|x + y\| \leq \|x\| + \|y\|$.

If $\| \cdot \|$ is a norm on X, then $(X, \| \cdot \|)$ is called a **normed space**.

Given a normed space, we recall that it is possible to define a natural notion of convergence by means of an $\epsilon - N$-criterion, just like in $\mathbb{R}$: in particular, a vector subspace Y of a normed space X is said to be **dense in** X if each $x \in X$ is the limit of a suitable sequence $(x_n)_{n \in \mathbb{N}} \subset Y$. We refer to [9, § 1.5] and [29, § 2.1] for details.

We will very often work with different normed spaces: if we want to stress that $\| \cdot \|$ is the norm of X, and this is not sufficiently clear from the context, we sometimes write $\| \cdot \|_X$.

On the other hand, we often drop the reference to the specific norm of a normed space if this is the canonical one in literature: important examples of spaces with a *canonical* norm are

- the space of bounded continuous functions $C_b(\Omega)$ with respect to $\|f\|_\infty := \sup_{x \in \Omega} |f(x)|$, whenever $\Omega \subset \mathbb{R}^d$ is an open domain;
- the space $C(K)$ of continuous functions over a compact metric space K with respect to $\|f\|_\infty := \max_{x \in K} |f(x)|$;
- the Lebesgue spaces $L^p(\Omega)$ over a σ-finite measure space (Ω, μ) with respect to $\|f\|_p := \left(\int_p |f|^p \mathrm{d}\mu \right)^{\frac{1}{p}}$, $p \in [1, \infty)$.

All these normed spaces are also *Banach spaces*, i.e., complete normed spaces, see [9, § 1.5].

Definition C.3. Let X, Y be normed spaces, and let $T : X \to Y$ be a linear operator.

(1) T is called **bounded** if there exists $M > 0$ such that

$$\|Tx\|_Y \leq M \|x\|_X \qquad \text{for all } x \in X.$$

(2) Let additionally $X \subset Y$ and T be an embedding operator, i.e., $Tx = x$ for all $x \in X$. If T is bounded, then X is said to be **continuously embedded** in Y, and we write

$$X \hookrightarrow Y.$$

Then

$$\|T\| := \sup_{\|x\|_X \leq 1} \|Tx\|_Y,$$

defines a norm on the vector space of bounded linear operators from X to Y, which we denote by $\mathcal{L}(X,Y)$, or rather X' if $Y=\mathbb{K}$ and $\mathcal{L}(X)$ if $X=Y$. An invertible bounded linear operator is called an **isomorphism**. A bounded linear operator such that

$$\|Tx\|_Y = \|x\|_X \qquad \text{for all } x\in X$$

is called **isometric**.

Example C.4. (1) All linear operators on finite-dimensional spaces are bounded.

(2) By Hölder's inequality, the linear operator defined on $\ell^1(\mathbb{N})$ by a diagonal matrix defines is bounded if and only if the diagonal entries of A form a bounded sequence.

(3) The linear operator $T : f \mapsto \frac{\mathrm{d}f}{\mathrm{d}x}$ is bounded from $C^1([0,1])$ to $C([0,1])$ but unbounded from $C([0,1])$ to $C([0,1])$.

Remark C.5. In some occasions one has to deal with operators that are not linear, but rather **antilinear**, i.e., such that

$$T(x+y)=Tx+Ty, \qquad T(\alpha x)=\overline{\alpha}Tx, \qquad \alpha\in\mathbb{C} \qquad \text{for all } x,y\in X$$

holds instead of (C.1) (clearly, antilinearity differs from linearity only if X,Y are complex vector spaces).

A bounded antilinear bijection is called an **antiisomorphism**.

Exercise C.6. Let X,Y be normed spaces. Show that any linear or antilinear operator $T : X\to Y$ such that

$$\|Tx\|_X \geq \alpha\|x\|_Y, \qquad x\in X,$$

for some $\alpha>0$, is injective.

Exercise C.7. Show that a linear operator on a normed space is bounded if and only if it is Lipschitz continuous if and only if it is continuous.

Remark C.8. Let X,Y be normed spaces, and let $T\in\mathcal{L}(X,Y)$. It is clear that the sets $\operatorname{Ker}T := \{x\in X : Tx=0\}$ and $\operatorname{Ran}T := \{y\in Y : y=Tx \text{ for some } x\in X\}$ are vector spaces: they are called **null space** and **range** of T, respectively. Since T is continuous, $T^{-1}C$ is a closed subset of X for all closed subsets C of Y. In particular, $\operatorname{Ker}T = T^{-1}\{0\}$ is a closed subspace of X, while Ran$T$ need generally not be closed. Can you find an example?

An important subclass is that of compact operators. Compact *linear* operators are discussed in [29, Chapter 7].

Definition C.9. Let X,Y be normed spaces and let $\Phi : X\to Y$ be a (possibly nonlinear) continuous operator.

(1) Φ is called **compact** if it is continuous and for any bounded sequence $(x_n)_{n\in\mathbb{N}}\subset X$ the sequence of values $(\Phi(x_n))_{n\in\mathbb{N}}\subset Y$ has a convergent subsequence.

(2) Let additionally $X\subset Y$ and Φ be an embedding operator, i.e., Φ is linear and $\Phi x = x$ for all $x\in X$. If Φ is compact, then X is said to be **compactly embedded** in Y.

In other words, by definition a continuous mapping Φ is compact if it maps bounded sets to relatively compact sets, i.e., to sets whose closure is compact.

Remark C.10. It is known that *linear* compact operators on a Banach space *form an ideal*: by this we mean that TS, ST are compact for all T bounded and S compact. Also, in the linear case continuity need not be explicitly assumed, as it is known that it automatically follows from the property of that bounded sets are mapped to relatively compact sets, see [29, Cor. 7.5.3].

Solutions

Solution C.11 (Exercise C.6). We prove it for the linear case only; the proof is similar in the anti-linear case. Assume first that there exists two elements $x_1,x_2\in H_1$ with $T(x_1)=T(x_2)$. Due to linearity one has $T(x_1-x_2)=T(x_1)-T(x_2)=0$ which is in contradiction with the inequality $\|Tx\|_{H_2}\geq\alpha\|x\|_{H_1}$. To see this, note that $\|T(x_1-x_2)\|_{H_2}=0$ whereas $\|x_1-x_2\|_{H_1}\neq 0$

Solution C.12 (Exercise C.7). Let $T : H \to H$ be an operator on a normed space H. Assume that T is bounded. We have

$$\frac{\|T(x+h) - T(x)\|_H}{\|h\|_H} = \frac{\|T(h)\|_H}{\|h\|_H} \leq c$$

where $c > 0$ is some constant. Hence T is indeed Lipschitz continuous. Since every Lipschitz continuous function is continuous, continuity follows directly.

On the other hand, if T is Lipschitz continuous, it is clear from the expression above that T is also bounded (simply set $x = 0$).

Finally, let T be continuous. Then, for $\varepsilon = 1$ there exists $\delta > 0$ such that for all $x \in H$ with $\|x\|_H = \delta$ one has $\|T(x)\|_H \leq 1$. For a general $y \in H$ we set

$$\tilde{x} := \frac{\delta y}{\|y\|_H}$$

and obtain $\|\tilde{x}\|_H = \delta$. Hence, using the properties of the norm,

$$\|T(\tilde{x})\|_H = \frac{\delta}{\|y\|_H}\|T(\tilde{y})\|_H \leq 1 \, .$$

Consequently, $\frac{\|T(\tilde{y})\|_H}{\|y\|_H} \leq \frac{1}{\delta}$.

Appendix D

Hilbert spaces

In this chapter we present some elementary results from the theories of Hilbert spaces and Fourier series. This introduction is not meant to be complete, but only to yield some technical tools that are useful for the study of PDEs. In view of some applications it is useful to develop this theory also in the complex case.

Definition D.1. Let H be a vector space over a field $\mathbb{K}$. An **inner product** $\langle \cdot, \cdot \rangle$ on H is a mapping $H \times H \to \mathbb{K}$ such that for all $x, y \in H$ and all $\lambda \in \mathbb{K}$

(1) $\langle x, x \rangle \geq 0$ and $\langle x, x \rangle = 0 \Leftrightarrow x = 0$,

(2) $\langle \lambda x, y \rangle = \lambda \langle x, y \rangle$,

(3) $\langle x, y + z \rangle = \langle x, y \rangle + \langle x, z \rangle$,

(4) $\langle x, y \rangle = \langle y, x \rangle$ if $\mathbb{K} = \mathbb{R}$, or else $\langle x, y \rangle = \overline{\langle y, x \rangle}$ if $\mathbb{K} = \mathbb{C}$.

Then, $(H, \langle \cdot, \cdot \rangle)$ is called a **pre-Hilbert space**.

If we want to stress that $\langle \cdot, \cdot \rangle$ is the inner product of H, and this is not sufficiently clear from the context, we sometimes write $\langle \cdot, \cdot \rangle_H$.

We usually drop the reference to the specific inner product of a pre-Hilbert space if this is canonical, like

- the Euclidean spaces $\mathbb{R}^d$ with respect to $\langle x, y \rangle := \sum_{k=1}^d x_k \overline{y_k}$,
- the complex space $\ell^2(\mathbb{N})$ of square summable sequences with respect to $\langle x, y \rangle := \sum_{k=1}^\infty x_k y_k$,
- the complex space $C(K)$ of continuous functions over a compact metric space with respect to $\langle f, g \rangle := \int_K f(x) \overline{g(x)} \mathrm{d}x$.

Lemma D.2 (Cauchy–Schwarz inequality)**.** *Let H be a pre-Hilbert space. Then for all $x, y \in H$ one has*

$$|\langle x, y) |^2 \leq \langle x, x \rangle \langle y, y \rangle.$$

Corollary D.3 (Triangle inequality)**.** *Let H be a pre-Hilbert space. Then the mapping $\| \cdot \|$ defined by*

$$\|x\| := \sqrt{\langle x, x \rangle}, \qquad x \in H, \tag{D.1}$$

is a norm on H. Furthermore, we have that

$$\|x + y\| \leq \|x\| + \|y\|, \qquad \text{for all } x, y \in H.$$

The following **Young inequality** is similar to the Cauchy–Schwarz inequality in spirit, but offers the possibility to estimate a product of two vectors in an additive, rather than multiplicative way.

Corollary D.4. *Let H be a pre-Hilbert space. Then for all $x, y \in H$ one has*

$$|\langle x, y\rangle| \leq \frac{1}{2}\|x\|^2 + \frac{1}{2}\|y\|^2.$$

Proof. Apply the Cauchy–Schwarz inequality and the fact that $(a-b)^2 \geq 0$ for all $a, b \in \mathbb{R}$, hence in particular for $\|x\|, \|y\|$. □

Exercise D.5. Prove the following extensions of the Young inequality.

(1) $\langle x, y\rangle \leq \epsilon\|x\|^2 + \frac{1}{4\epsilon}\|y\|^2$ for all $\epsilon > 0$ and all $x, y \in H$.

(2) $\langle x, y\rangle \leq \frac{1}{p}\|x\|^p + \frac{1}{q}\|y\|^q$ for all $p, q \in (1, \infty)$ such that $p^{-1} + q^{-1} = 1$ and all $x, y \in H$.

(3) $\langle x, y\rangle \leq \epsilon\|x\|^p + \frac{1}{q(\epsilon p)^{\frac{q}{p}}}\|y\|^q$ for all $p, q \in (1, \infty)$ such that $p^{-1} + q^{-1} = 1$, all $\epsilon > 0$ and all $x, y \in H$.

(Hint for (2): use convexity of the exponential function.)

Definition D.6. Let H be a pre-Hilbert space. If H is complete, then H is called a **Hilbert space**

We have already introduced the pre-Hilbert spaces $\mathbb{R}^d$ and $\ell^2(\mathbb{N})$: it turns out that they are actually complete, hence Hilbert spaces. On the other hand, given a non-complete normed space, there is a smallest complete space containing the given space, called its **closure**: while for any bounded open domain $\Omega \subset \mathbb{R}^d$ the pre-Hilbert space $C(\overline{\Omega})$ is not complete, its closure is the Hilbert space $L^2(\Omega)$.

In some equations considered in this course it is important to deal with vector-valued functions: all the above notions can be extended to this case. For example, whenever Ω is an open domain of $\mathbb{R}^d$ and $m \in \mathbb{N}$, the space

$$L^2(\Omega; \mathbb{K}^m) := \left\{ f : \Omega \to \mathbb{K}^m : f \text{ is measurable and } \int_\Omega |f(x)|^2 \mathrm{d}x < \infty \right\}$$

is a Hilbert space with respect to the inner product

$$\langle f, g\rangle_{L^2(\Omega;\mathbb{K}^m)} := \int_\Omega \langle f(x), g(x)\rangle_{\mathbb{K}^m} \mathrm{d}x.$$

Here $|f(x)|$ denotes for all $x \in \mathbb{K}^m$ the Euclidean norm of the vector $f(x) \in \mathbb{K}^m$, while $\langle f(x), g(x)\rangle_{\mathbb{K}^m}$ is the Euclidean inner product of the vectors $f(x), g(x) \in \mathbb{K}^m$.

Definition D.7. Let H be a pre-Hilbert space. Two vectors $x, y \in H$ are said to be mutually **orthogonal** if $\langle x, y\rangle_H = 0$, and we write $x \perp y$.

If two subsets A, B of H satisfy $\langle x, y\rangle_H = 0$ for all $x \in A$ and all $y \in B$, also A, B are said to be **mutually orthogonal**. Moreover, the set of all vectors of H that are orthogonal to each vector in A is called **orthogonal complement** of A and is denoted by $A^\perp$.

Definition D.8. Let $(H, \langle\cdot,\cdot\rangle_H)$ be a pre-Hilbert space. Then a family $\{e_n \in H \setminus \{0\} : n \in J\}$, $J \subset \mathbb{N}$, is called **orthogonal** if $\langle e_n, e_m\rangle_H = 0$ for all $n \neq m$, and **orthonormal** if $\langle e_n, e_m\rangle_H = \delta_{mn}$ for all $m, n \in J$, where δ_{mn} denotes the Kronecker delta.

Moreover, $\{e_n \in H : n \in J\}$ is called **total** if its linear span (i.e., the set of all *finite* linear combinations of elements of the family) is dense in H.

An orthonormal and total family is called a Hilbert space orthonormal basis of H, or simply a **basis**.

Exercise D.9. Let H be a pre-Hilbert space. Prove the following assertions.

(1) If x, y are mutually orthogonal, then $\|x\|_H^2 + \|y\|_H^2 = \|x + y\|_H^2$. (This is nothing but the **theorem of Pythagoras** if $H = \mathbb{R}^2$).

(2) More generally, $2\|x\|_H^2 + 2\|y\|_H^2 = \|x + y\|_H^2 + \|x - y\|_H^2$ for all $x, y \in H$.

(3) Also, for all $x, y \in H$ one has

$$4\langle x, y\rangle = \|x + y\|_H^2 - \|x - y\|_H^2 \quad \text{if } \mathbb{K} = \mathbb{R}, \text{ and}$$
$$4\langle x, y\rangle = \|x + y\|_H^2 + i\|x + iy\|_H^2 - \|x - y\|_H^2 - i\|x - iy\|_H^2 \quad \text{if } \mathbb{K} = \mathbb{C}.$$

(4) If A is a subset of H, then $A \subset (A^\perp)^\perp$.

(5) The orthogonal complement $H^\perp$ agrees with $\{0\}$.

The properties of linear functionals count to the most striking differences between Hilbert spaces and general Banach spaces. All bounded linear functional over a normed space X always form a Banach space, but if X is in fact a Hilbert space, much more can be said. The following has been proved in 1907 by Frigyes Riesz and independently also by Maurice René Fréchet: remarkably, both were then high school teachers (in Hungary and France, respectively).

Theorem D.10 (Representation theorem of Riesz–Fréchet)**.** *Let H be a Hilbert space. For each bounded linear functional ϕ on H there exists a unique $y_\phi \in H$ such that*

$$\phi(x) = \langle x, y_\phi \rangle \qquad \text{for all } x \in H. \tag{D.2}$$

Moreover, the mapping $H' \ni \phi \mapsto y_\phi \in H$ is an isometric isomorphism if $\mathbb{K} = \mathbb{R}$, and an isometric antiisomorphism if $\mathbb{K} = \mathbb{C}$.

Remark D.11. One can show likewise that $y \mapsto \phi_y := \overline{\langle \cdot, y \rangle}$ is an isometric isomorphism between H and the vector space of antilinear bounded functionals on H.

We conclude this chapter with a short reminder of abstract Fourier analysis.

Convergence of series in normed spaces has been introduced in [9]: the following is just a special case.

Definition D.12. Let H be a Hilbert space and $(x_n)_{n\in\mathbb{N}} \subset H$ a countable family of vectors. The associated series $\sum_{n=1}^{\infty} x_n$ is called **convergent to** $x \in H$ if for all $\epsilon > 0$ there exists $N_0 \in \mathbb{N}$ such that for all $N \geq N_0$

$$\left\| x - \sum_{n=1}^{N} x_n \right\| < \epsilon$$

(and **convergent** if it is convergent to some $x \in H$).

It is called **absolutely convergent** if $\sum_{n\in\mathbb{N}} \|x_n\|_H < \infty$.

Example D.13. Prove that if H is a Hilbert space, then every absolutely convergent series is convergent. The converse is not true: consider e.g. $H = \ell^2$, the canonical basis $(\delta_{nk})_{n\in\mathbb{N}} : k \in \mathbb{N})$, where δ_{nm} denotes the Kronecker delta, and the sequence defined by $x_n := \frac{1}{n} e_n$, and prove that it is convergent although it is (clearly) not absolutely convergent.

Theorem D.14. *Let H be a pre-Hilbert space. Let $\{e_n \in H : n \in J\}$ be an orthonormal family, $J \subset \mathbb{N}$. Then the following assertions hold.*

(1) $\sum_{n\in J} |\langle x, e_n \rangle|^2 \leq \|x\|^2$ for all $x \in H$.

(2) If H is complete, then the series $\sum_{n\in J} \langle x, e_n \rangle e_n$ converges.

(3) If $x = \sum_{n\in J} a_n e_n$, then $\|x\|^2 = \sum_{n\in J} |a_n|^2$ and $a_n = \langle x, e_n \rangle$ for all $n \in J$.

The assertions in (1) and (3) are usually called **Bessel's inequality** and **Parseval's identity**. The scalars a_n in (3) are called **Fourier coefficients of** x.

Proposition D.15. *Let H be a Hilbert space. An orthonormal family $\{e_n : n \in \mathbb{N}\}$ is in fact total (i.e., a basis of H) if and only if*

$$\langle f, e_n \rangle = 0 \text{ for all } n \in \mathbb{N} \quad \text{implies} \quad f = 0. \tag{D.3}$$

Hence, if $\{e_k : k \in \mathbb{N}\}$ is a basis, then

$$\sum_{n\in\mathbb{N}} \langle f, e_n \rangle e_n$$

actually converges to f, and it is called the **Fourier series** associated with f, while the terms $\langle f, e_n \rangle$ are called **Fourier coefficients** of f.

Exercise D.16. Show that the **Rademacher functions** $\{r_n : n \in \mathbb{N}\}$ define an orthonormal family of $L^2(0,1)$, where $r_n(t) := \operatorname{sgn}\sin(2^n \pi t)$ for a.e. $t \in (0,1)$. Rademacher functions have been proposed by Hans Adolph Rademacher in 1922.

Exercise D.17. Let $\phi \in L^2(\mathbb{R})$. The countable family $\{\phi_{jk} : j,k \in \mathbb{Z}\}$ is called a **wavelet** if it defines a basis of $L^2(\mathbb{R})$, where ϕ_{jk} is defined by

$$\phi_{jk}(t) := 2^{\frac{k}{2}} \phi(2^k t - j) \qquad \text{for a.e. } t \in (0,1).$$

Show that if $\phi(t) = \mathbf{1}_{[0,\frac{1}{2})} - \mathbf{1}_{[\frac{1}{2},1)}$ for a.e. $t \in (0,1)$, then $\{\phi_{jk} : j,k \in \mathbb{Z}\}$ is a wavelet: this is the so-called **Haar wavelet** discovered in 1909 by Alfréd Haar as a doctoral student at the University of Göttingen.

Theorem D.18. *The family*

$$\left\{1, \sqrt{2}\cos\frac{\pi n \cdot}{\ell}, \sqrt{2}\sin\frac{\pi n \cdot}{\ell} : n, m = 1, 2, 3, \dots\right\}$$

is an orthonormal basis of the real Hilbert space $L^2(0,\ell)$.

Finite linear combinations of elements of this family are called **trigonometric polynomials**.

The canonical Fourier series of a function on the bounded interval $[0,1]$ is then introduced with respect to the above basis.

Definition D.19. Let $f : [0,1] \to \mathbb{R}$. Then the Fourier sequence associated with f is

$$\int_0^1 f(x)\,\mathrm{d}x + 2\sum_{k=1}^{n} \int_0^1 f(y)\cos(ky)\,\mathrm{d}x \cos(nx) + 2\sum_{k=1}^{n} \int_0^1 f(y)\sin(ky)\,\mathrm{d}y \sin(nx), \qquad t \in [0,1].$$

Corollary D.20. *The family*

$$\{e_n := \mathrm{e}^{2\pi i n \cdot} : n \in \mathbb{Z}\}$$

is a basis of the complex Hilbert space $L^2(0,1;\mathbb{C})$. Accordingly, for all $f \in L^2(0,1;\mathbb{C})$ the Fourier series associated with it converges to f with respect to the L^2-norm, i.e.,

$$\lim_{n\to\infty} \int_0^1 \left| f(t) - \sum_{|k|\le n} \int_0^1 f(x)\,\mathrm{e}^{2\pi i k(t-x)}\,\mathrm{d}x \right|^2 \mathrm{d}t = 0.$$

Exercise D.21. Deduce Corollary D.20 from Theorem D.18. *(Hint: observe that it suffices to check condition* (D.3) *for any real-valued function f.)*

Appendix E

Sobolev spaces

Sobolev spaces represent the ideal bridge between Functional Analysis and the analysis of Partial Differential Equations. They are Banach spaces with interesting abstract properties, which are discussed thoroughly in [29, Chapter 5]; at the same time, they are a key ingredient when it comes to solving partial differential equations, and in particular they are essential for the purpose of introducing the notion of weak solution.

Everything you always wanted to know about Sobolev spaces, and more, can be found in [1]; a more concise, old but still unsurpassed gentle introduction to Sobolev spaces can be found in [11, Chapters 8–9].

E.1 Weak derivatives

Definition E.1. Let $I \subset \mathbb{R}$ be an open interval. A function $f \in L^2(I)$ is said to be **weakly differentiable** if there exists $g \in L^2(I)$ such that

$$\int_I f(x)\overline{h'(x)}\mathrm{d}x = -\int_I g(x)\overline{h(x)} \qquad \text{for all } h \in C_c^1(I). \tag{E.1}$$

In this case, we call g the **weak derivative** of f and with an abuse of notation we write $f' = g$..

For an open domain $\Omega \subset \mathbb{R}^d$ we denote here and in the following by $C_c^k(\Omega)$ the vector space of k times continuously differentiable functions with compact support, i.e., k-times continuously differentiable functions $f : \Omega \to \mathbb{K}$ such that f and hence its derivative vanish outside some compact subset of Ω.

Example E.2. Let $I = (-1, 1)$. The prototypical case of a weakly differentiable function that does not admit a classical derivative in some point is given by

$$f(x) := \frac{|x| + x}{2} = \begin{cases} 0 & \text{if } x \leq 0, \\ x & \text{if } x > 0. \end{cases}$$

Take some function $h \in C_c^1((-1, 1); \mathbb{R})$ and observe that

$$\begin{aligned} \int_{-1}^1 f(x)h'(x)\mathrm{d}x &= \int_{-1}^0 f(x)h'(x)\mathrm{d}x + \int_0^1 f(x)h'(x)\mathrm{d}x \\ &= \int_0^1 xh'(x)\mathrm{d}x \\ &= \Big[xh(x)\Big]_0^1 - \int_0^1 h(x)\mathrm{d}x \\ &= -\int_0^1 h(x)\mathrm{d}x, \end{aligned}$$

where the last equality follows from compactness of support of h (whence $h(1) = 0$). In other words,

$$\int_{-1}^1 f(x)h'(x)\mathrm{d}x = \int_{-1}^1 H(x)h(x)\mathrm{d}x$$

for all $h \in C_c^1(I)$, where H is defined by

$$H(x) := \begin{cases} 0 & \text{if } x \leq 0, \\ 1 & \text{if } x > 0; \end{cases} \tag{E.2}$$

this shows that f is weakly differentiable with $f' = H$. (Such H is called the **Heaviside function** after Oliver Heaviside, who used it in the 1890s to describe the current flow when an electric circuit is switched on.)

On the other hand, H is not even *weakly* differentiable. Indeed, assume there exists $\delta_0 \in L^2(-1,1)$ such that $H' = \delta$ (by the index 0 in δ_0 we are stressing the point of discontinuity of H): then

$$\int_I H(x)h'(x)\mathrm{d}x = -\int_I \delta_0(x)h(x)\mathrm{d}x \qquad \text{for all } h \in C_c^1(I).$$

Taking into account that by definition of H

$$\int_I H(x)h'(x)\mathrm{d}x = \int_0^1 h'(x)\mathrm{d}x = h(1) - h(0) = -h(0)$$

for all smooth functions h with compact support in $(-1,1)$, we deduce that

$$h(0) = \int_I \delta_0(x)h(x)\,\mathrm{d}x.$$

The restriction of H to $(-1,0)$ and to $(0,1)$ is clearly differentiable – hence weakly differentiable – with identically vanishing derivative; i.e., $\delta_0(x) = 0$ for all $x \neq 0$; but then, as an L^2-function δ_0 would be indistinguishable from 0! Hence, $h(0) = \int_I \delta_0(x)h(x)\mathrm{d}x = 0$ for all $h \in C_c^1(I)$, which is impossible.

Remark E.3. Nowadays we know that δ_0 introduced in Example E.2 cannot exist as an L^2-function, but it is a well-defined *bounded linear functional*, in the sense of Appendix C: it is the so-called **Dirac delta**, after Paul Dirac, who introduced it in Cambridge in 1930. The Dirac delta acts boundedly on $C_c^1(I)$ – in fact, even on the larger space $H_0^1(I)$ – by

$$\delta_0 : h \mapsto h(0).$$

This pathological example is among those that led Laurent Schwartz, then in Grenoble, to develop the *theory of distributions* in 1944.

The definition of Sobolev space can be promptly extended to the higher dimensional case.

Definition E.4. Let $\Omega \subset \mathbb{R}^d$ be an open domain. A function $f \in L^2(\Omega)$ is said to be **weakly differentiable** if there exist $g := (g_1, \ldots, g_d) \in \left(L^2(\Omega)\right)^d$ such that

$$\int_\Omega f(x)\overline{\frac{\partial h}{\partial x_i}}(x)\mathrm{d}x = -\int_\Omega g_i(x)\overline{h(x)}\mathrm{d}x \qquad \text{for all } h \in C_c^1(\Omega) \text{ and all } i = 1, \ldots, d. \tag{E.3}$$

The function g_i is then called the **weak partial derivative** of f in the direction of e_i[1] and with an abuse of notation we write $\frac{\partial f}{\partial x_i} = \partial_i f = g_i$.

Remark E.5. Observe that because $C_c^\infty(I)$ is dense in $C_c^1(I)$ by Weierstrass' Theorem (see [29, § 2.4]), we can equivalently replace C_c^1 by C_c^∞ in Definition E.1 and Definition E.4.

Definition E.6. Let $\Omega \subset \mathbb{R}^d$ be an open domain. The set

$$W^{1,p}(\Omega) := \left\{ f \in L^p(\Omega) : \ f \text{ is weakly differentiable with } \frac{\partial f}{\partial x_i} \in L^p(\Omega) \text{ for all } i = 1, \ldots, d \right\}$$

is called **first Sobolev space** (on Ω, of index p).

[1] e_i denotes as usual the i-th vector of the canonical basis of $\mathbb{R}^d$.

Indeed, is easy to see that if a function is weakly differentiable, then its partial derivative is in each direction unique, and that $W^{1,p}(\Omega)$ is actually a vector space with respect to the usual rules of sum and multiplication of functions. This class of spaces was introduced in 1936 by Sergei Sobolev, during his researches at the Steklov Mathematical Institute in Moscow.

Remark E.7. Since any two continuously differentiable functions f, h satisfy (E.1) (which is nothing but the usual formula of integration by parts), by definition $C^1(\overline{\Omega}) \subset H^1(\Omega)$ provided Ω is bounded: i.e., each function in $C^1(\overline{\Omega})$ is representative of a weakly differentiable L^2-function whose weak derivative is again in L^2.

Moreover, integrating by parts one clearly sees that each continuously differentiable function is also weakly differentiable, i.e., $C^1(\overline{\Omega}) \subset H^1(\Omega)$. In general a function that is merely in $C^1(\Omega)$ need not be in $L^2(\Omega)$, but in fact each $u \in C^1(\Omega) \cap L^2(\Omega)$ such that $u' \in \big(L^2(\Omega)\big)^d$ also belongs to $H^1(\Omega)$.

Higher order Sobolev spaces can be introduced recursively, like in the case of classical derivatives.

Definition E.8. Let $\Omega \subset \mathbb{R}^d$ be an open domain and $k \in \mathbb{N}$. The k^{th} **Sobolev space** of integrability index p is defined recursively by setting

$$W^{k+1,p}(\Omega) := \left\{ f \in W^{k,p}(\Omega) : \frac{\partial f}{\partial x_i} \in W^{k,p}(\Omega) \text{ for all } i = 1, \dots, d \right\}, \qquad k \in \mathbb{N}.$$

Due to historical reasons, in the literature the notation

$$H^k(\Omega) := W^{k,2}(\Omega), \qquad k \in \mathbb{N},$$

is often used.

Lemma E.9. *Let Ω be an open domain of $\mathbb{R}^d$, $k \in \mathbb{N}$, and $p \in [1, \infty]$. Then $W^{k,p}(\Omega)$ is a Banach space with respect to the norm*

$$\|f\|_{W^{k,p}} := \sum_J \|f^J\|_{L^p}, \qquad f \in W^{k,p}(\Omega),$$

where the sum is taken over all multi-indices J of length between 0 and k. We denote the associated norm by $\|\cdot\|_{H^k}$. It is separable for $p \in [1, \infty)$, reflexive for $p \in (1, \infty)$.

Furthermore, $H^k(\Omega) = W^{k,2}(\Omega)$ is a Hilbert space with respect to the inner product

$$\langle f, g \rangle_{H^k} := \sum_{h=1}^{k} \langle f, g \rangle_{L^2} + \langle f', g' \rangle_{(L^2)^d} = \int_\Omega f(x)\overline{g(x)}\mathrm{d}x + \int_\Omega \nabla f(x) \cdot \overline{\nabla g(x)}\mathrm{d}x.$$

Observe that

$$\langle f, g \rangle_{H^1} = \sum_J \langle f^J, g^J \rangle_{L^2}, \qquad f, g \in H^1(\Omega),$$

where the sum is taken over all multi-indices J of length 0 or 1 (recall the notation introduced in (2.18)).

Theorem E.10. *Let Ω be an open domain of $\mathbb{R}^d$. Then the Sobolev space $W^{k,p}(\Omega)$ agrees with the closure of $C^\infty(\Omega) \cap W^{k,p}(\Omega)$ with respect to $\|\cdot\|_{W^{k,p}}$. If Ω is bounded and has C^1-boundary, then $W^{k,p}(\Omega)$ even agrees with the closure of $C^\infty(\overline{\Omega})$ with respect to $\|\cdot\|_{W^{k,p}}$.*

Theorem E.10 war proved by Norman Meyers and James Serrin in Minneapolis in 1964. The main feature of the Meyers–Serrin Theorem is the possibility of extending results that are known to hold for continuous functions to more general functions in certain Sobolev spaces, by simple density arguments. For example, the Gauss–Green formulae hold for functions that are merely in $W^{k,p}$-spaces, instead of C^k.

Corollary E.11. *Let Ω be a bounded open domain of $\mathbb{R}^d$. Then each $u \in H^1(\Omega)$ is limit of a sequence $(u_n)_{n\in\mathbb{N}} \subset C_c^\infty(\Omega)$ with respect to the $H^1(\Omega)$-norm.*

E.2 Sobolev Embedding Theorems

One of the main features of Sobolev spaces is that while they are defined on the top of L^2, their relation to other Lebesgue spaces and to spaces of continuously differentiable functions is quite well-behaved, as the following embedding results show. The first is canonically referred to as **Sobolev Embedding Theorem**.

We take over the statement from [29, Chapter 5], but stress that this result still holds under the weaker assumption that the domain boundary is merely Lipschitz continuous.

Theorem E.12. *Let Ω be an open domain of $\mathbb{R}^d$ with Lipschitz boundary. Then there exists a constant $C > 0$, only depending on Ω, d, p, such that for all $u \in H^1(\Omega)$*

- *if $d = 1$: $\|u\|_{L^\infty} \leq C\|u\|_{W^{1,p}}$, and in particular $W^{1,p}(\Omega) \subset L^\infty(\Omega)$ for all $p \in [1, \infty]$;*
- *if $d = 2$: $\|u\|_{L^q} \leq C\|u\|_{W^{1,p}}$, and in particular $W^{1,p}(\Omega) \subset L^q(\Omega))$, for all $p \in (1, \infty)$ and all $q \in [1, \infty)$;*
- *if $d \geq 3$: $\|u\|_{L^{\frac{2d}{d-2}}} \leq C\|u\|_{H^1}$, and in particular $H^1(\Omega) \subset L^{\frac{2d}{d-2}}(\Omega))$.*

This explains in particular the special role played by $L^6(\mathbb{R}^3)$ in many partial differential equations of physical relevance, where energy functionals are defined on $H^1(\mathbb{R}^3)$. The following generalizes a results obtained in 1966 in Berkeley by Charles B. Morrey and is usually referred to as **Morrey's inequality**.

Theorem E.13. *For all $k \in \mathbb{N}$ and $p \in [1, \infty)$ such that $pk > d$, let h be the largest integer such that $h \leq k - \frac{d}{p}$. Then each $u \in W^{k,p}(\mathbb{R}^d)$ has a representative $u^* \in C^{k-[\frac{d}{p}]-1,\alpha}(\mathbb{R}^d)$, where*

$$\alpha = \begin{cases} [\frac{d}{p}] + 1 - \frac{d}{p}, & \text{if } \frac{d}{p} \notin \mathbb{Z}, \\ 1 - \varepsilon, & \text{otherwise}, \end{cases}$$

for any $\varepsilon > 0$.

Here $C^{h+\alpha}(\mathbb{R}^d)$ denotes the space of **Hölder continuous functions**, i.e., the space of functions such that the Hölder norm

$$\|f\|_{C^{h+\alpha}} := \max_{|k| \leq h} \sup_{x \in \mathbb{R}^d} |D^k f(x)| + \max_{|\beta| = h} \sup_{x \neq y \in \mathbb{R}^d} \frac{|D^\beta f(x) - D^\beta f(y)|}{|x - y|^\alpha}.$$

Because continuity and differentiability are local properties, Morrey inequality extends immediately to Sobolev spaces on domains.

Corollary E.14. *For all $k \in \mathbb{N}$ and $p \in [1, \infty)$ such that $pk > d$, let h be the largest integer such that $h \leq k - \frac{d}{p}$. Let $\Omega \subset \mathbb{R}^d$ be an open domain with Lipschitz boundary. Then each $u \in W^{k,p}(\Omega)$ has a representative $u^* \in C^h(\overline{\Omega})$.*

In particular: for $d = 1$ and for all $p \in [1, \infty]$, each $u \in W^{1,p}(I)$ has a continuous representative. Accordingly, it is common (although a slight abuse of language) to say that *one-dimensional weakly differentiable functions are continuous*. It is worthwhile to emphasize that this important property is exclusive of the 1-dimensional case. In particular, this allows to consider point evaluations of functions in $W^{1,p}(I)$.

The converse is not true, i.e., there exist continuous functions that are not weakly differentiable. A continuous but nowhere differentiable, like the Weierstrass function (see [29, § 3.3.2]), yields a counterexample. In fact, it can be proved that a weakly differentiable function has to be differentiable (in the classical sense) almost everywhere.

Many more estimates relating the Sobolev norms of functions to those in different Lebesgue spaces are known: among the most useful ones we mention the following **Gagliardo–Niremberg inequalities**, proved in 1959 by Emilio Gagliardo in Genua, and independently by Louis Nirenberg in New York.

Lemma E.15. *There exists $C > 0$ such that*

$$\|u\|_{L^\infty} \leq C\|u\|_{H^1}^{\frac{1}{2}} \|u\|_{L^2}^{\frac{1}{2}} \qquad \text{for all } u \in H^1(\mathbb{R}).$$

Also, there exists $C' > 0$ such that for all $p \geq 2$

$$\|u\|_{L^p} \leq C' \|u\|_{H^1}^{1-2/p} \|u\|_{L^2}^{2/p} \qquad \text{for all } u \in H^1(\mathbb{R}^2).$$

Moreover, there exists $\tilde{C} > 0$ such that

$$\|u\|_{L^\infty} \leq C \|u\|_{H^1}^{\frac{1}{2}} \|u\|_{L^2}^{\frac{1}{2}} + \tilde{C} \|u\|_{L^1} \qquad \text{for all } u \in H^1(\Omega)$$

and there exists $C' > 0$ such that for all $p \geq 2$

$$\|u\|_{L^p} \leq C' \|u\|_{H^1}^{1-2/p} \|u\|_{L^2}^{2/p} + \tilde{C}' \|u\|_{L^1} \qquad \text{for all } u \in H^1(\Omega)$$

if $\Omega \subset \mathbb{R}^d$ is a bounded open domain for $d = 1$ or $d = 2$, respectively.

For bounded domains (and – it can be proved – not only for them) the above Sobolev embeddings have actually even better properties: the following result goes under the name of **Rellich–Kondrachov Embedding Theorem**. Its crucial idea was developed in Göttingen in 1930 by Franz Rellich; this was later generalized and translated in the language of Sobolev spaces by Vladimir Kondrachov at the Steklov Institute in Moscow in 1945.

Theorem E.16. *Under the assumptions of Theorem E.12, let additionally Ω be bounded with Lipschitz boundary. Then the embeddings $W^{1,p}(\Omega) \subset L^\infty(\Omega)$ and $W^{1,1}(\Omega) \subset L^q(\Omega)$ are compact for all $p \in (1, \infty]$ and all $q \in [1, \infty)$. Furthermore, the embedding $W^{1,1}(\Omega) \subset L^1(\Omega)$ is compact, too.*

Remark E.17. The proof of the Sobolev Embedding Theorem is based upon the possibility of introducing an operator $\Theta : H^1(\Omega) \to H^1(\mathbb{R}^d)$ that extends functions supported on Ω to the whole $\mathbb{R}^d$ while increasing neither their L^2-norm nor their H^1-norm: this is the content of [29, Thm. 5.3.27], which is stated for domains with C^1 boundary. However, it is possible to prove that this assumption can be much weakened, but the involved geometric conditions are rather technical: we refer to [1] for a very comprehensive treatment. In particular, Theorems E.12-E.13 hold for domains whose boundary is merely Lipschitz, as we have formulated them, while for Theorem E.16 even a continuous boundary suffices, see [45]. These sharpened formulations of embedding results for Sobolev spaces are possibly not very relevant from a functional analytical viewpoint and they were omitted from [29]; but they might be crucial for the analysis of partial differential equations, which often model physical processes taking place in rough domains – think of a room with flat walls like in Example B.3.

E.3 Traces of Sobolev functions

We know from elementary calculus that an integration by parts involves the boundary values of a function. However, since the boundary $\partial\Omega$ of a Lipschitz domain is a set of measure zero, it is now clear a priori how to define boundary values of a function $f \in L^2(\Omega)$. Indeed, although this is impossible for an arbitrary function $f \in L^2(\Omega)$, it is possible for a function $f \in H^1(\Omega)$. More explicitly, to associate boundary values to a function$f \in H^1(\Omega)$ – or, more generally, $f \in W^{1,p}(\Omega)$ – shall mean nothing else than associating to it a function $g \in L^2(\partial\Omega)$ – respectively, $f \in L^p(\partial\Omega)$ – in a meaningful way; we then write $f|_{\partial\Omega} := g$ and call $f|_{\partial\Omega}$ the **trace** of f; the mapping T that associates a function f with its boundary values $Tf = f_{|\partial\Omega}$ is called **trace operator**. The following is therefore referred to as **trace theorem** for Sobolev spaces.

Theorem E.18. *Let $\Omega \in \mathbb{R}^d$ be a bounded Lipschitz domain. Then for all $p \in [1, \infty)$ there exists a bounded linear operator $T : W^{1,p}(\Omega) \to L^p(\partial\Omega)$ such that*

$$(Tu)(z) = u(z) \qquad \text{for all } f \in C^1(\Omega) \cap C(\overline{\Omega}) \text{ and all } x \in \partial\Omega. \tag{E.4}$$

In view of (E.4), with an abuse of notation the function Tu is often denoted by $u_{|\partial\Omega}$.

Definition E.19. Let $\Omega \subset \mathbb{R}^d$ be an open domain. The closure of $C_c^\infty(\overline{\Omega})$ with respect to the $W^{k,p}(\Omega)$-norm is called the **Sobolev space of order k with vanishing boundary values**, which we denote by $W_0^{k,p}(\Omega)$; and by $H_0^k(\Omega)$ if $p = 2$.

If the reason for its name is a little obscure in view of Definition E.4, take into account the theorem of Meyers–Serrin and observe that $W_0^{k,p}(\Omega)$ is constructed in a way that is analogous to $W^{k,p}(\Omega)$, with the only difference that we are taking the $W^{k,p}$-closure of $C_c^\infty(\Omega)$, rather than of the (larger) space $C^\infty(\Omega) \cap W^{k,p}(\Omega)$. Accordingly, $W_0^{k,p}(\Omega)$ is generally smaller than $W^{k,p}(\Omega)$, although they may occasionally agree, cf. Exercise E.32 below. The explanation for the second part of the name of the spaces $W^{k,p}$ is more delicate. Informally speaking, one may e.g. expect functions in $W_0^{1,p}(\Omega)$ to vanish at the boundary. But what does this mean? Theorem E.18 gives us a clue.

Theorem E.20. *Let Ω be an open domain of $\mathbb{R}^d$ with C^1-boundary. If $u \in W_0^{1,p}(\Omega) \cap C(\overline{\Omega})$, then $u(z) = 0$ for all $z \in \partial\Omega$. Conversely, if $u \in W^{1,p}(\Omega) \cap C(\overline{\Omega})$ and $u(z) = 0$ for all $z \in \partial\Omega$, then $u \in W_0^{1,p}(\Omega)$.*

The following result is crucial in that it allows for a convenient re-norming of Sobolev spaces with vanishing boundary values. It goes under the name of **Poincaré inequality**, since a special case was proved by Henri Poincaré in Paris in 1887.

Theorem E.21. *Let Ω be a bounded open domain of $\mathbb{R}^d$ that is contained in a strip of width 2δ, i.e., such that there exist $\delta > 0$ and $i \in \{1, \dots, d\}$ for which*

$$x_i \in (-\delta, \delta) \quad \text{for all } x = (x_1, \dots, x_d) \in \Omega. \tag{E.5}$$

Then for all $f \in H_0^1(\Omega)$ there holds

$$\|f\|_{L^2} \leq 2d \|\nabla f\|_{L^2};$$

more generally, for all $p \in [1, \infty)$ and some $C > 0$ there holds

$$\|f\|_{L^p}^p \leq C \|\nabla f\|_{L^p}$$

for all $u \in W_0^{1,p}(\Omega)$. In particular, $f \mapsto \|\nabla f\|_{L^p}$ defines a norm on $W_0^{1,p}(\Omega)$ which is equivalent to the one induced by $W^{1,p}(\Omega)$.

We conclude this section by observing that all Gauss–Green formulae, which we have already formulated for continuously differentiable functions, can be shown to hold for weakly differentiable functions, too.

Theorem E.22. *[First Gauss–Green formula] Let $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain. For all $f, g \in H^1(\Omega)$ one has*

$$\int_\Omega \partial_j f \cdot g \, \mathrm{d}x = \int_{\partial\Omega} f g \cdot \nu_j \, \mathrm{d}\sigma - \int_\Omega f \cdot \partial_j g \, \mathrm{d}x \tag{E.6}$$

Here, as in Appendix B, $\nu_j = (\nu)_j$ is the j-th component of the outward normal vector.

Theorem E.23. *[Second Gauss–Green formula] Let $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain. For all $f \in H^2(\Omega)$ and $h \in H^1(\Omega)$ one has*

$$\int_\Omega \nabla f \cdot \nabla h \, \mathrm{d}x = \int_{\partial\Omega} f \frac{\partial h}{\partial \nu} \, \mathrm{d}\sigma - \int_\Omega f \Delta h \, \mathrm{d}x. \tag{E.7}$$

Corollary E.24. *[Third Gauss–Green formula] Let $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain. For all $f, h \in H^2(\Omega)$ one has*

$$\int_\Omega f \Delta h \, \mathrm{d}x - \int_\Omega h \Delta f \, \mathrm{d}x = \int_{\partial\Omega} f \frac{\partial h}{\partial \nu} \, \mathrm{d}\sigma - \int_{\partial\Omega} h \frac{\partial f}{\partial \nu} \, \mathrm{d}\sigma. \tag{E.8}$$

Theorem E.25. *Let $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain and $g := (g_1, \dots, g_d)^T$ a vector field with $g_i \in H^1(\Omega)$. Then,*

$$\int_\Omega \operatorname{div} g \, \mathrm{d}x = \int_{\partial\Omega} g \cdot \vec{\nu} \, \mathrm{d}\sigma \, .$$

E.4 Order structure of Sobolev spaces

For a function $u : \Omega \to \mathbb{R}$ on a domain $\Omega \subset \mathbb{R}^d$, we introduce the function $u^+ : \Omega \to \mathbb{R}$

$$u^+(x) := \begin{cases} u(x), & \text{if} \quad u(x) \geq 0, \\ 0\,, & \text{otherwise}, \end{cases}$$

and call it the positive part of u. In the same way, the function $u^- : \Omega \to \mathbb{R}$

$$u^-(x) := \begin{cases} -\,u(x), & \text{if} \quad u(x) \leq 0, \\ 0\,, & \text{otherwise}, \end{cases}$$

is called the negative part of u.

Theorem E.26. *Let $\Omega \subset \mathbb{R}^d$ be a domain. If $f \in H^1_0(\Omega)$ then $x \mapsto |f(x)|$ is weakly differentiable and $|f| \in H^1_0(\Omega)$.*

Proof. For simplicity, we assume that $f \in H^1_0(\Omega)$ is real-valued. We then introduce the function

$$g_\varepsilon(t) := \sqrt{t^2 + \varepsilon^2}\,, \quad t \in \mathbb{R}\,,$$

for some $\varepsilon > 0$. It is clear that $g_\varepsilon \in C^1(\mathbb{R})$; employing the chain rule for Sobolev functions (see, e.g., [29]), we conclude that the function $g_\varepsilon(f)$ is weakly differentiable with, $j = 1, \ldots, d$,

$$\int_\Omega g_\varepsilon(f)\partial_j\varphi \,\mathrm{d}x = \int_\Omega \frac{f\partial_j f}{\sqrt{f^2+\varepsilon^2}}\varphi \,\mathrm{d}x\,, \quad \forall \varphi \in C^\infty_0(\Omega)\,.$$

Dominated convergence then allows us to pull the limit $\varepsilon \to 0$ inside the integral to obtain, for $j = 1, \ldots, d$,

$$\partial_j|f| = \begin{cases} \partial_j f\,, & \text{if} \quad f(x) \geq 0, \\ -\,\partial_j f\,, & \text{if} \quad f(x) \leq 0, \\ 0 & \text{otherwise}\,. \end{cases} \tag{E.9}$$

Since $\partial_j f \in L^2(\Omega)$, the statement follows. □

Corollary E.27. *Let $\Omega \subset \mathbb{R}^d$ be a domain and $f \in H^1(\Omega)$ a real-valued function. Then both its positive and negative part $f_\pm$ are weakly differentiable and, for $j = 1, \ldots, d$ and $x \in \Omega$,*

$$\frac{\partial f^+}{\partial x_j}(x) = \begin{cases} \dfrac{\partial f}{\partial x_j}(x)\,, & \textit{if} \quad f(x) \geq 0, \\ 0 & \textit{otherwise}, \end{cases}$$

as well as

$$\frac{\partial f^-}{\partial x_j}(x) = \begin{cases} -\,\dfrac{\partial f}{\partial x_j}(x)\,, & \textit{if} \quad f(x) \leq 0, \\ 0 & \textit{otherwise}, \end{cases}$$

Proof. Notice that we can write

$$f^+ = \frac{1}{2}\left(|f| + f\right)\,,$$

as well as

$$f^- = \frac{1}{2}\left(|f| - f\right)\,.$$

The statement then readily follows from Theorem E.26. □

Exercise E.28. Let $I \subset \mathbb{R}$ be an open interval. Let $f \in L^2(I)$. Show that if a function g satisfying (E.1) exists, then it is unique.

Exercise E.29. Consider the operator $S : f \mapsto f'$. Show that S is a linear operator that is not bounded on $L^2(0,1)$, but indeed bounded from $H^1(0,1)$ to $L^2(0,1)$.

Exercise E.30. Prove the following special case of Theorem E.13: Let I be an interval.

(1) Let $f \in L^2(I)$ such that $\int_I f(x)\overline{h'(x)}\mathrm{d}x = 0$ for all $h \in C_c^1(\overline{I})$. Show that there exists a constant $c \in \mathbb{K}$ such that $f(x) = c$ for a.e. $x \in I$.

(2) Let $g \in L^2(I)$ and $x_0 \in I$. Define $G : I \ni x \mapsto \int_{x_0}^x g(t)\mathrm{d}t \in \mathbb{K}$. Show that $G \in C(I)$ and moreover $\int_I G(x)\overline{h'(x)}\mathrm{d}x = -\int_I g(x)\overline{h(x)}\mathrm{d}x$ for all $h \in C_c^1(I)$.

(3) Conclude that each $f \in H^1(I)$ has a representative $f^* \in C(\overline{I})$ such that $f(x) = f^*(x)$ for a.e. $x \in I$. Moreover, $\|f^*\|_{C(\overline{I})} \leq \|f\|_{H^1(I)}$ for all $f \in C([0,1])$.

Exercise E.31. Consider $u \in H^1(0,\infty)$ and $\eta \in C^1(0,\infty)$ such that $\eta|_{\{x\geq 1-\epsilon\}} = 0$ for some $\epsilon \in (0,1)$. Prove that then $\eta u \in H^1(0,\infty)$ and the product rule

$$(\eta u)' = \eta' u + \eta u'$$

holds for its weak derivative. *(Hint: After taking a test function, apply the usual rule for products of continuously differentiable functions to its product with η.)*

Exercise E.32. Prove that if $\Omega = \mathbb{R}^d$, then $H^1(\Omega) = H_0^1(\Omega)$.

Exercise E.33. In the literature it is common to find a different formulation of the Poincaré inequality, namely that

$$\left\| u - \frac{1}{|\Omega|}\int_\Omega u(x)\,\mathrm{d}x \right\|_{L^2} \leq 4\delta^2 \|\nabla u\|_{L^2} \qquad \text{for all } u \in H^1(\Omega).$$

Find out what is the connection of this inequality to the one in the above theorem. *(Hint: Take into account the fact that $\frac{1}{|\Omega|}\int_\Omega u(x)\,\mathrm{d}x$ agrees with the best approximation of any function u on the subspace $\langle \mathbf{1}\rangle$ of constant functions, cf. Exercise 3.38, and hence by Exercise 3.36*

$$u - \frac{1}{|\Omega|}\int_\Omega u(x)\,\mathrm{d}x$$

turns out to be the best approximation of u onto the subspace orthogonal to $\langle \mathbf{1}\rangle$.)

E.5 Sobolev spaces and Fourier transform

Let us conclude with a brief discussion aimed to the students who already have a good knowledge of Hilbert spaces and Sobolev spaces from [29]. For the Sobolev spaces $H^m(\mathbb{R}^d)$ there exists an interesting and useful characterization in terms of the Fourier transform. Recall that the Fourier transform is a linear unitary operator from $L^2(\mathbb{R}^d)$ to $L^2(\mathbb{R}^d)$ and which is defined via

$$[\mathcal{F}(f)](k) := \frac{1}{(2\pi)^{d/2}}\int_{\mathbb{R}^d} \mathrm{e}^{-ikx} f(x)\,\mathrm{d}x\,, \qquad f \in L^2(\mathbb{R}^d)\,. \tag{E.10}$$

Now, assume we have a function $f \in H^1(\mathbb{R}^d)$. Then, f has a weak derivative which is also in $L^2(\mathbb{R}^d)$ and using the formula (E.10) one obtains the relation

$$\begin{aligned}\mathcal{F}(f_{x_j}) &= \frac{1}{(2\pi)^{d/2}}\int_{\mathbb{R}^d} \mathrm{e}^{-ikx} f_{x_j}(x)\,\mathrm{d}x \\ &= ik_j \mathcal{F}(f)\,.\end{aligned}$$

As a consequence, $k_j\mathcal{F}(f) \in L^2(\mathbb{R}^d)$ for $j = 1,\ldots,d$. This leads to the following result.

Theorem E.34. *The spaces $H^m(\mathbb{R}^d)$ consist exactly of those functions $f \in L^2(\mathbb{R}^d)$ for which*

$$\int_{\mathbb{R}^d} \left(1 + \|k\|_{\mathbb{C}^d}^{2m}\right) |\mathcal{F}[f](k)|^2\,\mathrm{d}k < \infty\,.$$

We refer to [39] for more details on the connection of Fourier transform and Sobolev spaces.

Appendix F

The Spectral Theorem for bounded linear operators

-

In Chapter 4 an important role is played by the Spectral Theorem for unbounded, self-adjoint operators. Both for historical and didactical reasons it is interesting to present here its predecessor: the Spectral Theorem for general bounded operators. We recall the basic definition.

Proposition F.1. *Let H_1, H_2 be Hilbert spaces and T be a bounded linear operator from H_1 to H_2. Then there exists exactly one bounded linear operator T^* from H_2 to H_1 such that*

$$\langle Tx, y\rangle_{H_2} = \langle x, T^*y\rangle_{H_1} \qquad \text{for all } x \in H_1 \text{ and } y \in H_2. \tag{F.1}$$

holds. Moreover,

$$\|T\| = \|T^*\| \qquad \text{and} \qquad \|T^*T\| = \|T\|^2 \qquad \text{for all } T \in \mathcal{L}(H). \tag{F.2}$$

Definition F.2. A bounded linear operator U on a Hilbert space H is called

- **unitary** if U is invertible and $U^{-1} = U^*$;
- **self-adjoint** if $U = U^*$.
- **normal** if $UU^* = U^*U$.

Example F.3. For any continuous $k : [0,1] \times [0,1] \to \mathbb{R}$ that is symmetric (i.e., such that $k(x,y) = k(y,x)$ for all $x, y \in [0,1]$) the **Fredholm operator** F_k on $L^2(0,1)$ defined by

$$F_k f(x) := \int_0^1 k(x,y) f(y) \mathrm{d}y$$

is self-adjoint. More generally, for any measurable, essentially bounded $k : [0,1] \times [0,1] \to \mathbb{C}$ the adjoint of F_k is given by $F_k^* = F_{k^*}$, where k^* is defined by $k^*(x,y) := \overline{k(y,x)}$ a.e.

Definition F.4. Two linear operators A, B on a Hilbert space H are called **unitarily equivalent** if there exists a unitary operator U such that $A = U^{-1}BU$.

It is apparent that unitary equivalence is an equivalence relation. We can now formulate a fundamental characterization of self-adjoint operators that goes under the name **Spectral Theorem for bounded self-adjoint operators**.

Theorem F.5. *Let A be a bounded, linear, self-adjoint operator on a separable Hilbert space H. Then A is unitarily equivalent to a multiplication operator M_q on $L^2(\Xi)$ for some finite measure space Ξ and some bounded measurable function $q : \Xi \to \mathbb{R}$.*

This beautiful and natural formulation of the Spectral Theorem is one of the main legacies of Paul Halmos, who popularized it in 1963, while appointed at the University of Michigan, in an article boldly entitled *What Does the Spectral Theorem Say?*. The more usual version, which is plagued by obscure notions and technical details both in its statement and its proof, is more common in the mathematical literature and is, almost 70 years after Halmos' article, unfortunately still the most popular one.

One of the highlights in the functional analytical theory of Hilbert spaces is the following special case of the Spectral Theorem: if the relevant operator is not only bounded, but even compact, then it is possible to derive a sharper description of the function q and the measure space Ξ appearing in Theorem F.5.

We mention an important special case of Theorem F.5, the **Spectral Theorem for compact self-adjoint operators**: we refer to [29, Chapter 7] for a much closer look at this topic.

Theorem F.6. *Let H be a Hilbert space and T a compact self-adjoint operator on H. Then there exist*

- *a finite or countably infinite set K,*
- *an orthonormal system $\{e_k : k \in K\}$, and*
- *a sequence $(\lambda_k)_{k\in K}$ of real numbers*

such that the following assertions hold.

(1) $(\lambda_k)_{k\in K}$ has exactly one accumulation point at 0 whenever K is infinite.

(2) The sequence $(\lambda_k)_{k\in K}$ consists of the non-zero eigenvalues of T and $(e_k)_{k\in K}$ are the associated eigenvectors.

(3) At most finitely many of the eigenvalues coincide (i.e.: for all $\lambda \in \mathbb{R}$ there is $r \in \mathbb{N}_0$ such that $\lambda = \lambda_{k_1} = \ldots = \lambda_{k_r}$ and all remaining eigenvalues are different from λ), and we say that all eigenvalues have finite multiplicity.

(4) H can be decomposed as

$$H = \operatorname{Ker} T \oplus \overline{\operatorname{span}\{e_1, e_2, \ldots\}} \,. \tag{F.3}$$

(5) The expansion

$$Tx = \sum_{k\in K} \lambda_k \langle x, e_k\rangle_H e_k \tag{F.4}$$

holds for all $x \in H$.

(Hence, if in particular $\operatorname{Ker} T = \{0\}$ then there exists an orthonormal basis of H consisting of eigenvectors of T only.)

Its proof relies upon the following spectral result for compact operators, see [29, Thm. 7.7.1 and Thm. 7.7.3], which we recapitulate for the sake of self-containedness.

Theorem F.7. *Let X be a Banach space with $\dim X = \infty$. If $T \in \mathcal{K}(X)$ is a compact operator, then the following assertions hold.*

(1) $0 \in \sigma(T)$.

(2) If $\lambda \in \sigma(T) \setminus \{0\}$ then $\lambda \in \sigma_p(T)$, i.e., λ is an eigenvalue.

(3) Every eigenvalue $\lambda \in \sigma(T) \setminus \{0\}$ has finite multiplicity, i.e., the null space of $\lambda - T$ is finite dimensional.

(4) Exactly one of the following assertions holds:

- *$\sigma(T) = \{0\}$.*
- *$\sigma(T) \setminus \{0\}$ is a finite set.*
- *$\sigma(T) \setminus \{0\}$ is a sequence converging to 0.*

Remark F.8. The above theorem can also be extended to self-adjoint operators on non-separable Hilbert spaces, but part (2) of the proof is more delicate, since one also needs an argument based on transfinite induction. Since essentially all Hilbert space that naturally appear in the theory of evolution equations are separable, we omit this extension.

The Spectral Theorem can be extended to deliver a full characterization of bounded operators that are merely normal. While there is no abundance of evolution equations driven by normal operators, it is this formulation of the Spectral Theorem that we invoke when proving the Spectral Theorem 4.40 for unbounded self-adjoint operators.

Theorem F.9. *Let A be a bounded normal operator on a separable Hilbert space H. Then A is unitarily equivalent to a (necessarily bounded) multiplication operator M_q on $L^2(\Xi)$ for some finite measure space Ξ and some measurable function $q : \Xi \to \mathbb{C}$.*

Appendix G

Bochner integrals and Bochner spaces

The Bochner integral forms a natural generalization of the well-known Lebesgue integral to vector-valued functions. Although it can be introduced for general measure spaces, we will focus on intervals $I \subset \mathbb{R}$, which is all we need in this course. We refer to [6, Section 1.1] for a broader introduction.

The ultimate goal of this section is to show one can naively manipulate Bochner integrals pretty much like usual Lebesgue integrals of scalar-valued functions. To begin with, let X denote a Banach space and $f : I \to X$ some vector-valued function. Our goal is to define the integral

$$\int_I f \, \mathrm{d}x \in X \, . \tag{G.1}$$

As in the case of Lebesgue integrals, such an integral is well-defined for *(Bochner-)integrable* functions $f : I \to X$. On the other hand, every integrable function is *(Bochner-)measurable*. To introduce measurability, we consider simple functions $s : I \to X$ defined via

$$s(x) := \sum_{j=1}^{m} \alpha_j \chi_{A_j}(x) \, .$$

Here, $A_j \subset I$ are disjoint Borel-measurable subsets and $\alpha_j \in X$ for every $j = 1, \dots, m \in \mathbb{N}$. For a simple function $s : I \to X$, the Bochner integral is then defined as

$$\int_I s \, \mathrm{d}x := \sum_{j=1}^{m} \alpha_j |A_j| \, ,$$

where $|A_j|$ is the volume of A_j. A simple function is called integrable if $|A_j| < \infty$ whenever $\alpha_j \neq 0$.

Definition G.1. A function $f : I \to X$ is called measurable if there exists a sequence (s_n) of simple functions such that $\lim_{n\to\infty} s_n(x) = f(x)$ for almost every $x \in I$.

Furthermore, a measurable function is called integrable if there exists a sequence of integrable simple functions (s_n) such that

$$\lim_{n\to\infty} \int_I \|f - s_n\| \, \mathrm{d}x = 0 \, .$$

The Bochner integral of an integrable function $f : I \to X$ is then given by

$$\int_I f \, \mathrm{d}x := \lim_{n\to\infty} \int_I s_n \, \mathrm{d}x \, .$$

We have the following useful properties.

Proposition G.2. *A measurable function $f : I \to X$ is integrable if and only if*

$$\int_I \|f\| \, \mathrm{d}x < \infty \, .$$

Furthermore, if $T : X \to X$ is a bounded linear operator and $f : I \to X$ an integrable function, then $Tf : I \to X$ is also integrable and

$$T\left(\int_I f \,\mathrm{d}x\right) := \int_I Tf \,\mathrm{d}x \,.$$

Furthermore, for an integrable function $f : I \to X$ one has the inequality

$$\left\|\int_I f \,\mathrm{d}x\right\| \leq \int_I \|f\| \,\mathrm{d}x \,.$$

Based on the Lebesgue integral one introduced in calculus the well-known and important Lebesgue spaces $L^p(\Omega)$ on domains $\Omega \subset \mathbb{R}^d$ and $1 \leq p \leq \infty$. This construction can now be carried over to Bochner integrals to define the spaces $L^p((0,T),X)$ for $T > 0$. For this, we will consider two measurable functions $f, g : (0,T) \to X$ equal if they agree almost everywhere. Also, the essential supremum of a measurable function $f : I \to X$ is defined as

$$\text{ess sup}_{t\in(0,T)}\|f(t)\| := \inf_{N\subset(0,T):|N|=0} \sup_{x\in(0,T)\setminus N} \|f(t)\| \,.$$

Definition G.3 (Bochner L^p-spaces)**.** The space $L^p((0,T),X)$ consists of all measurable functions $f : (0,T) \to X$ (in the sense of equivalence classes) for which the the norm $\|f\|_{L^p((0,T),X)}$ is finite. Here, for $1 \leq p < \infty$,

$$\|f\|_{L^p((0,T),X)} := \left(\int_0^T \|f(t)\|^p \,\mathrm{d}t\right)^{1/p}$$

and

$$\|f\|_{L^\infty((0,T),X)} := \text{ess sup}_{t\in(0,T)}\|f(t)\| < \infty \,.$$

In a next step one may also introduce Bochner-type Sobolev spaces $W^{m,p}((0,T),X)$ for $m \in \mathbb{N}$. For $m = 2$, one then sets $H^m((0,T),X) := W^{m,2}((0,T),X)$. To do this, we first introduce the notion of a weak derivative of a function $f \in L^1((0,T),X)$.

Definition G.4. A function $f \in L^1((0,T),X)$ is weakly differentiable with derivative $g := \partial_t f \in L^1((0,T),X)$ if

$$\int_0^T \varphi'(t)f(t) \,\mathrm{d}t = -\int_0^T \varphi(t)g(t) \,\mathrm{d}t \,, \quad \forall \varphi \in C_0^\infty((0,T),X) \,.$$

Here $C_0^\infty((0,T),X)$ is the set of all smooth vector-valued functions whose support is compactly contained in $(0,T)$. Note that derivatives are defined, as in standard calculus, via the difference quotient. The only difference is that one uses the norm of X rather than the modulus in the definition of the limit.

Definition G.5 (Bochner $W^{m,p}$-spaces)**.** For $m \in \mathbb{N}$ and $1 \leq p < \infty$, the space $W^{m,p}((0,T),X)$ consists of all measurable functions $f : (0,T) \to X$ (again in the sense of equivalence classes) which are k-times weakly differentiable with weak derivatives $\partial_t^j f : (0,T) \to X$ and such that the norm

$$\|f\|_{W^{m,p}((0,T),X)} := \left(\sum_{j=0}^m \int_0^T \|\partial_t^j f\|^p \,\mathrm{d}t\right)^{1/p}$$

is finite. Here, we set $\partial_t^0 f := f$.

The space $C([0,T],X)$ of vector-valued continuous functions is, when equipped with the supremum norm, a Banach space. Furthermore, one has the following important embedding theorem.

Theorem G.6. *For every function $f \in W^{1,p}((0,T),X)$ one has $f \in C([0,T],X)$. Moreover, there exists a constant $c = c(p,T) > 0$ such that*

$$\|f\|_{L^\infty((0,T),X)} \leq c\|f\|_{W^{1,p}([0,T],X)} \,, \quad \forall f \in W^{1,p}((0,T),X) \,.$$

Furthermore, for every function $f \in W^{1,p}((0,T),X)$ with weak derivative $\partial_t f \in L^p((0,T),X)$ one has the identity

$$f(t) = f(0) + \int_0^t \partial_t f(s)\,\mathrm{d}s\,, \quad t \in [0,T]\,. \tag{G.2}$$

The relation (G.2) implies that the following statement.

Proposition G.7. *Let $f \in W^{1,p}((0,T),X)$ with $1 \le p < \infty$ be a function with zero weak derivative, i.e., $\partial_t f \equiv 0$. Then $f : (0,T) \to X$ is equivalent to a constant function.*

Finally, let us stress that there is a natural connection between Bochner spaces and the "usual" L^2-spaces encountered in the theory of time-dependent partial differential equations. More explicitly, let $\Omega \subset \mathbb{R}^d$ be a domain. Then

$$\begin{aligned} L^2((0,T),L^2(\Omega)) &= L^2((0,T)\times\Omega)\,,\\ L^2((0,T),H^1(\Omega)) &= \{f \in L^2((0,T)\times\Omega) : \partial_{x_j} f \in L^2((0,T)\times\Omega)\,,\ j=1,\dots,d\}\,. \end{aligned} \tag{G.3}$$

For this and more we refer also to [41].

Appendix H

Some remarks on the Navier–Stokes equations

Currently, one of the big open problems in mathematics (the third Millennium problem as announced by the Clay Mathematics Institute) is associated with a system of partial differential equations whose origin is found in fluid dynamics. More explicitly, the so-called Navier-Stokes equation describing an incompressible fluid are given by

$$\begin{cases} \dfrac{\partial v}{\partial t} + (v \cdot \nabla)v + \nabla p = \nu \Delta v \,, \\ \nabla \cdot v = 0 \,. \end{cases} \tag{H.1}$$

These equations are typically defined on a domain $\Omega \subset \mathbb{R}^n$, $n = 2, 3$, or on full space (in case of a bounded domain Ω, one also imposes boundary conditions along $\partial\Omega$). The unknowns are the velocity field $v : \Omega \times (0, T) \to \mathbb{R}^n$ and the pressure field $p : \Omega \times (0, T) \to \mathbb{R}$. In addition, $\nu \in [0, \infty)$ is the so-called viscosity of the fluid an which describes the friction within the fluid. Setting the viscosity equal to zero, i.e., $\nu \equiv 0$, one obtains the equally well-known Euler equations.

Of course, the difficulty in the study of (H.2) stems from the non-linearity of the Navier-Stokes equation coming from the term $(v \cdot \nabla)v$. Neglecting this non-linearity and looking at stationary solutions of (H.2) one obtains the so-called Stokes equations given by

$$\begin{cases} -\nu \Delta v + \nabla p = f \,, \\ \nabla \cdot v = 0 \,, \quad \text{on} \quad \partial\Omega. \\ v = 0, \end{cases} \tag{H.2}$$

Here $f : \Omega \to \mathbb{R}^n$ is an additional force field. Now, we formulate one version of the Millennium problem; it is concerned with the existence of *strong global* solutions of the Navier-Stokes equations. Note that a (C^1-) vector field $u : \mathbb{R}^3 \to \mathbb{R}^3$ is divergence-free iff $\nabla \cdot u = 0$.

Conjecture H.1 (Global regularity Navier-Stokes)**.** *Let $\nu > 0$ be given and $u_0 : \mathbb{R}^3 \to \mathbb{R}^3$ be a divergence-free vector field in the Schwartz class. Then there exists a smooth vector field $v : [0, \infty) \times \mathbb{R}^3 \to \mathbb{R}^3$ and a smooth function $p : [0, \infty) \times \mathbb{R}^3 \to \mathbb{R}^3$ that fulfill* (H.2) *with $v(0, \cdot) = v_0$. In addition, $v : [0, \infty) \times \mathbb{R}^3 \to \mathbb{R}^3$ satisfies the energy bound*

$$\sup_t \int_{\mathbb{R}^3} |v(t, x)|^2 \, \mathrm{d}x < \infty \,.$$

One can show that for any such smooth solution v one has $v(t, \cdot) \in H^{10}(\mathbb{R}^3, \mathbb{R}^3)$, see [43]. Hence, multiplying the first equation in (H.2) with v and integrating with respect to time yields the important *energy identity*, $T > 0$,

$$\frac{1}{2} \int_{\mathbb{R}^3} |v(T, x)|^2 \, \mathrm{d}x + \int_0^T \int_{\mathbb{R}^3} |\nabla v(t, x)|^2 \, \mathrm{d}x \mathrm{d}t = \frac{1}{2} \int_{\mathbb{R}^3} |v_0(x)|^2 \, \mathrm{d}x \,. \tag{H.3}$$

Also, one can show that a strong solution characterized through Conjecture H.1 is unique, see [Theorem 5.4 (iii), [44]]. Hence, the goal is to establish the existence of smooth global solutions to the Navier-Stokes equations which are necessarily unique. At this point we would like to mention that *local existence* of strong solutions can be established for any initial datum in the Schwartz class up to times $T \geq \frac{C}{\|v_0\|_\infty^2}$ where $C > 0$ is some fixed constant. Also, global strong solutions can be established given the initial datum is such that certain norms are small enough (such results were obtained by Fujita and Kato as well as by Kato, Giga, and Miyakawa).

Of course, instead of trying to prove existence of strong solutions one may also look at weak solutions to (H.2). As a matter of fact, this was done already in 1934 by J. Leray; we refer to [36] for a detailed discussion of the work of Leray. A (Leray-Hopf) weak solution to (H.2) is defined to be a vector field $v \in L^\infty((0,\infty), H) \cap L^2((0,\infty), V)$ for which

$$-\int_0^\infty \int_{\mathbb{R}^3} v \cdot \partial_t \varphi \, \mathrm{d}t\mathrm{d}x + \int_0^\infty \int_{\mathbb{R}^3} (v \cdot \nabla) v \cdot \varphi \, \mathrm{d}t\mathrm{d}x \\ + \nu \int_0^\infty \int_{\mathbb{R}^3} \nabla v \cdot \nabla \varphi \, \mathrm{d}t\mathrm{d}x = \int_{\mathbb{R}^3} v_0 \cdot \varphi \, \mathrm{d}x$$

holds for all divergence-free test functions $\varphi \in C_0^\infty([0,\infty) \times \mathbb{R}^3, \mathbb{R}^3)$. In addition, a (Leray-Hopf) weak solution is assumed to fulfill the energy condition (H.3) as an inequality. Also note that $V \subset H_0^1(\mathbb{R}^3, \mathbb{R}^3)$ is the subset of all divergence-free vector fields in $H_0^1(\mathbb{R}^3, \mathbb{R}^3)$ and H is defined as the closure of V in $L^2(\mathbb{R}^3, \mathbb{R}^3)$. Leray then proved the following important statement.

Theorem H.2 (Leray 1934, Existence weak solution)**.** *For every $v_0 \in H$ there exists a weak solution of* (H.2).

Hence, Theorem H.2 establishes existence of a weak global solution for every inital datum $v_0 \in H$! On the other hand, it is still an important *open problem* to prove uniqueness of (Leray-Hopf) weak solutions as described above. However, a recent breakthrough in the study of the Navier-Stokes equations by T. Buckmaster and V. Vicol [13] shows that uniqueness of weak solutions does *not* hold for some related (and a slightly weaker) notion of a weak solution. As a consequence, it seems possible that (Leray-Hopf) weak solutions to the Navier-Stokes equations are actually not unique. Taking into account that uniqueness of weak solutions follows from existence of a strong global solution (since a strong solution is, as soon as it exits, identical to the Leray-Hopf weak solution), it seems possible that the Conjecture H.1 is indeed false.

For further details on the Navier-Stokes equation we refer to [41, Teil VII]. In addition to the papers mentioned above, it might be useful to have a look at the paper [28].

Index

Index of persons

List of symbols

$A^\perp$ – orthogonal of a set A, 56
$B_r(x)$ – open ball of center x and radius r, 20
$C^{h+\alpha}(\Omega)$ – space of k-times differentiable functions on Ω whose k-th derivatives are Hölder continuous of order α, 62
$C^{h,k}(\Omega_1 \times \Omega_2)$ – space of all functions on $\Omega_1 \times \Omega_2$, $\Omega_i \subset \mathbb{R}^{d_i}$, that are h-times differentiable in the direction of the first d_1 coordinates and k-times differentiable in the direction of the last d_2 coordinates, 5
$D(A)$ – domain of an operator A, 51
$H^k(\Omega)$ – Sobolev space of k-times weakly differentiable functions on Ω of integrability index 2, 61
$H_0^k(\Omega)$ – Sobolev space of k-times weakly differentiable functions on Ω of integrability index 2 and vanishing boundary conditions, 64
$L^p(\Omega; \mathbb{K}^m)$ – Lebesgue space of $\mathbb{K}^m$-valued functions, 56
$W^{1,p}(\Omega)$ – Sobolev space of weakly differentiable functions on Ω of integrability order p, 61
$W^{k,p}(\Omega)$ – Sobolev space of k-times weakly differentiable functions on Ω of integrability index p, 61
$W_0^{k,p}(\Omega)$ – Sobolev space of k-times weakly differentiable functions on Ω of integrability index p and vanishing boundary conditions, 64
X' – dual space of a normed space X, 53
$X \hookrightarrow Y$ – X is continuously embedded in Y, 52
Δu – Laplacian of a function u, 14
$\operatorname{Ker} T$ – null space of an operator T, 53
Ω – open connected subset of $\mathbb{R}^d$, 3
$\mathbb{R}^m$ – m-dimensional real Euclidean space, 5
$\operatorname{Ran} T$ – range of an operator T, 53
$\frac{\partial}{\partial \nu}$ – normal derivative of a function u, 49
$\langle \cdot, \cdot \rangle_H$ – inner product in a Hilbert space H, 55
$\nu(z)$ – outward normal vector of Ω at $z \in \partial\Omega$, 49
ν_j – j-th component of the outward normal vector, 49
$\partial B_r(x)$ – sphere of center x and radius r, 20
$\partial\Omega$ – boundary of a set Ω, 19
$\partial_j f$ – partial derivative of f with respect to the j-th variable, 49
$\mathrm{d}\sigma$ – surface measure, 20
$|A|$ – measure of a set $A \subset \mathbb{R}^d$, 20
$x \perp y$ – orthogonality of vectors x, y, 56
$\mathcal{L}(X)$ – space of bounded linear operators from the normed space X to itself, 53
$\mathcal{L}(X, Y)$ – space of bounded linear operators from the normed space X to the normed space Y, 53

Bibliography

[1] Robert A. Adams, John J.F. Fournier, *Sobolev Spaces*, Elsevier, Asterdam, 2003.

[2] Martin Aigner, Günter M. Ziegler, *Proofs from THE BOOK*, Springer-Verlag, Berlin, 2010.

[3] Herbert Amann, Joachim Escher, *Analysis III*, Birkhäuser, Basel, 2001.

[4] Wolfgang Arendt, *Heat Kernels*, Skript zum 9. Internetseminar, `www.uni-ulm.de/fileadmin/website_uni_ulm/mawi.inst.020/arendt/downloads/internetseminar.pdf`, 2006.

[5] Wolfgang Arendt, Karsten Urban, *Partielle Differenzialgleichungen: eine Einführung in analytische und numerische Methoden*, Spektrum, Heidelberg, 2010.

[6] Wolfgang Arendt, Charles J. K. Batty, Matthias Hieber, Frank Neubrander. *Vector-valued Laplace transforms and Cauchy problems*, Birkhäuser, Basel, 2011.

[7] Wolfgang Arendt, Ralph Chill, Christian Seifert, Hendrik Vogt, Jürgen Voigt, *Form Methods for Evolution Equations, and Applications*, Skript zum 18. Internetseminar, `www.mat.tuhh.de/veranstaltungen/isem18/pdf/LectureNotes.pdf`, 2015.

[8] András Bátkai, Marjeta Kramar Fijavž, Abdelaziz Rhandi, *Positive Operator Semigroups – From Finite to Infinite Dimensions*, Birkhäuser, Basel 2017.

[9] Wolfgang Beekmann, *Analysis*, Skript zum Kurs 1144, FernUniversität in Hagen, 2010.

[10] F. A. Berezin, M. A. Shubin, *The Schrödinger Equation*, Kluwer, Dordrecht, 1991.

[11] Haïm Brezis, *Functional Analysis, Sobolev Spaces and Partial Differential Equations*, Springer-Verlag, Berlin, 2010.

[12] J. Brüning, K.W. Wiegmann, *Differentialgeometrie*, Skript zum Kurs 1331, FernUniversität in Hagen, 2016.

[13] Tristan Buckmaster, Vlad Vicol, *Nonuniqueness of weak solutions to the Navier-Stokes equations*, Ann. Math. (189), 101-144, 2019.

[14] Robert Dautray, Jacques-Louis Lions, *Mathematical Analysis and Numerical Methods for Science and Technology, vol. 3*, Springer-Verlag, Berlin, 2000.

[15] Manfred Dobrowolski, *Angewandte Funktionalanalysis*, Springer-Verlag, Berlin, 2010.

[16] Tanja Eisner, *Stability of operators and operator semigroups*, Birkhäuser, Basel, 2010.

[17] Klaus-Jochen Engel, Rainer Nagel, *One-Parameter Semigroups for Linear Evolution Equations*, Springer-Verlag, Berlin, 2000.

[18] Lawrence C. Evans, *Partial Differential Equations*, American Mathematical Society, Providence (RI), 2009.

[19] Gregory Eskin, *Lectures on Linear Partial Differential Equations*, American Mathematical Society, Providence (RI), 2011.

[20] Helmut Fischer, Helmut Kaul, *Mathematik für Physiker*, Teubner, Wiesbaden, 2008.

[21] Harley Flanders, *Differentiation under the integral sign*, American Mathematical Monthly 80 (1973), 615–627.

[22] Markus Haase, *Functional Calculus*, Skript zum 21. Internetseminar, `www.math.uni-kiel.de/isem21/en/course/phase1`, 2018.

[23] Paul R. Halmos, *What Does the Spectral Theorem Say?*, American Mathematical Monthly 70 (1963), 241–247.

[24] Horst Herrlich and Walter Tholen, *Topologische Räume*, Skript zum Kurs 1354, FernUniversität in Hagen, 2003.

[25] Dirk Hundertmark, Lars Machinek, Martin Meyries, Roland Schnaubelt, *Operator Semigroups and Dispersive Equations*, Skript zum 16. Internetseminar, `https://www.math.kit.edu/iana3/~schnaubelt/media/isem16-skript.pdf`, 2013.

[26] John K. Hunter, *Lecture notes on measure theory*, `https://www.math.ucdavis.edu/~hunter/measure_theory/measure_theory.html`.

[27] Jürgen Jost, *Partial Differential Equations*, Springer-Verlag, Berlin, 2002.

[28] Tianwen Luo, Edriss S. Titi, *Non-uniqueness of weak solutions to hyperviscuous Navier-Stokes equations - On sharpness of J. L. Lions exponent*, Calculus of Variations and Partial Differential Equations (59), 2020.

[29] Delio Mugnolo, Joachim Kerner, Hafida Laasri, *Funktionalanalysis*, Skript zum Kurs 1245, FernUniversität in Hagen, 2017.

[30] Tosio Kato, *Perturbation Theory for Linear Operators*, Springer-Verlag, Berlin, 1980.

[31] Alessandra Lunardi, *Analytic Semigroups and Optimal Regularity in Parabolic Problems*, Springer Basel, 1995.

[32] Rainer Nagel (Hrsg.), *One-Parameter Semigroups of Positive Operators*, Springer-Verlag, Berlin, 1986.

[33] Edward Nelson, *A proof of Liouville's Theorem*, Proc. American Math. Society 12 (1961), 995.

[34] Peter J. Olver, *Applications of Lie Groups to Differential Equations*, Springer-Verlag, New York, 1993.

[35] Elmaati Ouhabaz, *Analysis of heat equations on domains*, Princeton University Press, Princeton (NJ), 2005.

[36] Wojciech S. Ożański, Benjamin Pooley, *Leray's fundamental work on the Navier–Stokes equations: a modern review of "Sur le mouvement d'un liquide visqueux emplissant l'espace"*, in: Charles L. Fefferman, James C. Robinson, José L. Rodrigo (Hrsg.), *Partial Differential Equations in Fluid Mechanics*, Cambridge University Press, Cambridge, 2019.

[37] Konstantin Pankrashkin, *Introduction to the spectral theory*, `https://www.imo.universite-paris-saclay.fr/~pankrashkin/2012spec/notes29.pdf`.

[38] Michael Reed, Barry Simon, *Methods of Modern Mathematical Physics. Vol. II*, Academic Press, New York, 1975.

[39] Michael Renardy, Robert C. Rogers, *An Introduction to Partial Differential Equations*, Springer-Verlag, Berlin, 2004.

[40] Carlo Rovelli, *The Order of Time*, Penguin Random House, New York, 2018.

[41] Ben Schweizer, *Partielle Differentialgleichungen: Eine anwendungsorientierte Einführung*, Spektrum Verlag, 2018.

[42] Hans Stephani, *Differential Equations – Their solution using symmetries*, Cambridge University Press, Cambridge, 1989.

[43] Terence Tao, *Finite time blowup for an averaged three-dimensional Navier-Stokes equation*, J. Amer. Math. Soc. (29), 601–674, 2016.

[44] Terence Tao, *Localisation and compactness properties of the Navier-Stokes global regularity problem*, Analysis & PDE (6), 25–107, 2013.

[45] Luc Tartar, *An Introduction to Sobolev Spaces and Interpolation Spaces*, Springer-Verlag, Berlin, 2007.

[46] Bernd Thaller, *The Dirac Equation*, Springer-Verlag, Berlin, 1992.

[47] Jan Van Neerven, *The asymptotic behaviour of semigroups of linear operators*, Springer-Verlag, Berlin, 1996

[48] Dirk Werner, *Funktionalanalysis*, Springer-Verlag, Berlin, 2005.