Introduction to Partial Differential Equations

Course module NWI-WB046B

Radboud University Nijmegen
Fall semester 2020/21

Annegret Burtscher

Contents

Preface

These lecture notes have been composed for the 3rd-year Bachelor course "Introduction to Partial Differential Equations" (NWI-WB046B) at Radboud University Nijmegen, which is taught by myself and Stefanie Sonner alternatingly since the winter semester 2018/19. These notes do not replace the lecture but contain most of the material covered, sometimes more. We cover, in particular, the classical linear partial differential equations of second order, their solutions and properties. We conclude with some results about the method of characteristics for nonlinear first order equations. Besides, we want to convey that there is no general theory like for ordinary differential equations, but most partial differential equations have to be considered independently using some general qualitative properties due to their particular type. Most importantly, although this is different in this course, solutions generally cannot be computed explicitly and analyzing partial differential equations means essentially proving well-posedness and stability results.

Besides these notes, the books by Evans [**4**], John [**6**] and Strauss [**9**] contain much of the same (and generally a lot more) material and may be used as an additional resource. The book by Craig [**2**] is a bit shorter and provides a softer introduction. The book by Vasy [**10**] is a good resource for weak solutions and Fourier methods, most of which unfortunately cannot be covered in this short course. Further references are provided throughout the text and can be provided upon request.

If you continue with a Master, a good follow-up course will be the MasterMath course on partial differential equations, and several courses related to particular types of partial differential equations (e.g., nonlinear wave equations, reaction diffusion equations, calculus of variations), as well as numerics for partial differential equations.

If you encounter any mistakes or typos please send me an email to `burtscher@math.ru.nl`. Comments are also welcome. Future generations of students will thank you!

Nijmegen, August 2020
Annegret Burtscher

CHAPTER 1

Introduction

Partial differential equations (for short: PDEs) are equations containing a function of several variables as well as their partial derivatives. partial differential equations arise in many fields of science, in particular, in physics, engineering, chemistry, biology and finance and are fundamental for the mathematical formulation of continuum models. They also play an important role in "pure mathematics", in particular, in geometry and analysis.

1.1. Basic definitions

In what follows we recall the basic definitions and fix some notation.

1.1.1. Partial derivatives.

DEFINITION 1.1. Let $\Omega \subseteq \mathbb{R}^n$ be open and $u\colon \Omega \to \mathbb{R}$ be a real valued function. The *i-th partial derivative of u at $x \in \Omega$* is given by (if the limit exists)

$$\frac{\partial u}{\partial x_i}(x) := \lim_{h\to 0} \frac{u(x+he_i)-u(x)}{h}, \qquad i \in \{1,\dots,n\},$$

where e_i denotes the i-th standard unit vector of $\mathbb{R}^n$.

We may also write $\partial_{x_i} u, \partial_i u, D_i u, u_{x_i}, u_i$ instead of $\frac{\partial u}{\partial x_i}$. Similarly, we write *partial derivatives of second order* of u as $\frac{\partial^2 u}{\partial x_i \partial x_j}, \partial^2_{x_i x_j} u, \partial^2_{ij} u, u_{x_i x_j}, u_{ij}$ for $i, j \in \{1,\dots,n\}$.

For higher order we make use of the multiindex notation. Note that $\cdot$ denotes the inner product and $|\cdot|$ the norm in $\mathbb{R}^n$.

DEFINITION 1.2. For a multiindex $\alpha = (\alpha_1,\dots,\alpha_n) \in \mathbb{N}_0^n$ with order $|\alpha| = \sum_{i=1}^n \alpha_i$ the *$|\alpha|$-th partial derivative of u at x* is (if the limit exists)

$$D^\alpha u(x) = \partial^\alpha u(x) := \partial^{\alpha_1}_{x_1} \dots \partial^{\alpha_n}_{x_n} u(x).$$

In a similar fashion, one writes $\alpha! = \prod_{i=1}^n \alpha_i!$ and $x^\alpha = \prod_{i=1}^n x_i^{\alpha_i}$. For some $k \in \mathbb{N}$ the term $D^k u$ denotes the set of all derivatives of order k, i.e., all derivatives $D^\alpha u$ where $|\alpha| = k$.

Important special cases are the *(total) first derivative*

$$Du = (\partial_{x_1} u, \dots, \partial_{x_n} u),$$

the *gradient*[1]

$$\nabla u = \begin{pmatrix} \partial_{x_1} u \\ \vdots \\ \partial_{x_n} u \end{pmatrix},$$

[1]Note that in the literature Du is also often used to denote the (spatial) gradient, e.g., in Evans [**4**]. We may use ∇u and Du interchangeably.

and the *Hessian*

$$D^2u = \begin{pmatrix} \partial^2_{x_1x_1}u & \dots & \partial^2_{x_1x_n}u \\ \vdots & \ddots & \vdots \\ \partial^2_{x_nx_1}u & \dots & \partial^2_{x_nx_n}u \end{pmatrix}.$$

By a Theorem of Schwarz D^2u is symmetric if u is twice continuously differentiable.

DEFINITION 1.3. Let $\Omega \subseteq \mathbb{R}^n$, $\overline{\Omega}$ denote its closure and $\partial\Omega$ its boundary. The spaces of continuous functions on Ω are denoted by

$$C(\Omega) := \{u\colon \Omega \to \mathbb{R} \,:\, u \text{ continuous}\},$$
$$C(\overline{\Omega}) := \{u \in C(\Omega) \,:\, u \text{ can be continuously extended to } \partial\Omega\}.$$

For $\Omega \subseteq \mathbb{R}^n$ open and $k \in \mathbb{N}$ we denote the spaces of k-times continuously differentiable functions by

$$C^k(\Omega) := \{u\colon \Omega \to \mathbb{R} \,:\, u \text{ is } k\text{-times continuously differentiable}\},$$
$$C^k(\overline{\Omega}) := \{u \in C^k(\Omega) \,:\, D^\alpha u \text{ can be continuously extended to } \partial\Omega \text{ for } |\alpha| \le k\}.$$

We say that a function $u\colon \Omega \to \mathbb{R}$ is *smooth* if all of its partial derivatives exist and are continuous on Ω. The analogous sets of smooth functions are denoted by $C^\infty(\Omega)$ and $C^\infty(\overline{\Omega})$.

1.1.2. Partial differential equations.

DEFINITION 1.4. Let Ω be an open subset of $\mathbb{R}^n$ and $k \in \mathbb{N}$. An expression of the form

$$F(D^k u(x), D^{k-1}u(x), \dots, u(x), x) = 0, \qquad x \in \Omega, \tag{1.1}$$

is called a *k-th order partial differential equation*, where $F\colon \mathbb{R}^{n^k} \times \mathbb{R}^{n^{k-1}} \times \dots \times \mathbb{R}^n \times \mathbb{R} \times \Omega \to \mathbb{R}$ is a given function and $u\colon \Omega \to \mathbb{R}$ is unknown.

We say that we *solve* a partial differential equation if we find all functions u that satisfy (1.1).

DEFINITION 1.5. A *classical solution* of the partial differential equation (1.1) is a C^k function u that satisfies (1.1).

Although it can be useful to obtain explicit solutions, this will often not be possible and a qualitative result is what we look for instead.

Instead of a scalar equation we can also consider a higher dimensional *system of partial differential equations* (1.1) where $F\colon \mathbb{R}^{mn^k} \times \mathbb{R}^{mn^{k-1}} \times \dots \times \mathbb{R}^{mn} \times \mathbb{R}^m \times \Omega \to \mathbb{R}^m$ is given and $u\colon \Omega \to \mathbb{R}^m$ is unknown. Here, $D^\alpha u = (D^\alpha u_1, \dots, D^\alpha u_m)$ and $D^k u = \{D^\alpha u \,:\, |\alpha| \le k\}$.

1.2. Classification of partial differential equations

Depending on the properties of F in (1.1) we distinguish between linear, semilinear, quasilinear and fully nonlinear partial differential equations.

If the operator F in (1.1) is linear in u and its derivatives, i.e., if (1.1) is of the form

$$\begin{aligned}\mathcal{F}(u)(x) &= F(D^k u(x), D^{k-1}u(x), \dots, u(x), x) \\ &= \sum_{i=0}^{k} \sum_{|\alpha|=i} c_\alpha(x)\partial^\alpha u(x) + f(x) = 0,\end{aligned}$$

then the partial differential equation (1.1) is called *linear*. If $f \equiv 0$ then the partial differential equation is *homogeneous* and otherwise *inhomogeneous*. The advantage of linear homogeneous equations in that the superposition principle[2] holds. That is, if u and v are both solutions to (1.1), then so is $au + bv$ for all $a, b \in \mathbb{R}$ since

$$\mathcal{F}(au + bv) = a\mathcal{F}(u) + b\mathcal{F}(v) = 0.$$

Moreover, adding a homogeneous solution to particular inhomogeneous solution will produce another inhomogeneous solution.

In this course we'll study almost exclusively linear systems with constant coefficients.

A partial differential equation is called *semilinear* if it is linear in the highest order, i.e., if (1.1) is of the form

$$\sum_{|\alpha|=k} c_\alpha(x)\partial^\alpha u(x) + c_0(D^{k-1}u(x), \dots, u(x), x) = 0.$$

The partial differential equation is *quasilinear* if it is of the form

$$\sum_{|\alpha|=k} c_\alpha(D^{k-1}u(x), \dots, u(x), x)\partial^\alpha u(x) + c_0(D^{k-1}u(x), \dots, u(x), x) = 0.$$

Otherwise, that is, if F in (1.1) depends nonlinearly on $D^k u$, the partial differential equation is called *fully nonlinear*. Typically, the difficulty of the analysis of a partial differential equation increases with each degree of nonlinearity.

1.2.1. Classification of second order partial differential equations. In this course we mainly focus on linear partial differential equations of second order, i.e., equations of the form

$$\sum_{i,j=1}^{n} a_{ij}(x)u_{x_i x_j}(x) + \sum_{i=1}^{n} a_i(x)u_{x_i}(x) + a_0(x)u(x) = f(x), \qquad x \in \Omega. \tag{1.2}$$

If we assume that $u \in C^2(\Omega)$, then by Schwarz D^2u is symmetric and the coefficients a_{ij} form a symmetric matrix

$$A(x) = \begin{pmatrix} a_{11}(x) & \dots & a_{1n}(x) \\ \vdots & \ddots & \vdots \\ a_{n1}(x) & \dots & a_{nn}(x) \end{pmatrix}.$$

Based on the definiteness of A we can classify partial differential equations of the form (1.2). Mixed types can occur if the character changes depending on $x \in \Omega$.

Definition 1.6. We call a second order partial differential equation (1.2)

(i) *elliptic* if A is positive (or negative) definite,
(ii) *parabolic* if A is singular ($\det A = 0$), and
(iii) *hyperbolic* if one eigenvalue (with multiplicity one) of A has a different sign than all others.

The following three examples are the archetypes of elliptic, parabolic and hyperbolic linear equations. We will study them in more detail in Chapter 1.

[2]This may sound familiar from the theory of linear ordinary differential equations.

EXAMPLE 1.7. (i) For the *Laplace equation* $\Delta u = u_{x_1x_1} + \ldots + u_{x_nx_n} = 0$ we have $A(x) = I$ and thus the Laplace equation is elliptic.

(ii) The *heat equation* $u_t - \Delta u = 0$ yields a singular matrix $A(t,x) = \begin{pmatrix} 0 & 0 \\ 0 & -I \end{pmatrix}$, and thus the heat equation is parabolic.

(iii) For the *wave equation* $u_{tt} - \Delta u = 0$ we obtain $A(t,x) = \begin{pmatrix} 1 & 0 \\ 0 & -I \end{pmatrix}$, which makes the equation hyperbolic.

For nonlinear partial differential equations of second order one may still use this nomenclature based on the linearized form of the equation.

1.3. Important examples of partial differential equations

We briefly discuss examples of partial differential equations and their origin to illustrate the variety of applications in which partial differential equations occur.

1.3.1. Mathematics. In *complex analysis* the **Cauchy–Riemann equations** consist of two partial differential equations which together describe when a function is holomorphic. More precisely, a function $f\colon \Omega \subseteq \mathbb{C} \to \mathbb{C}$, $f(z) = f(x,y) = u(x,y) + iv(x,y)$ with $u, v \in C^1$ is complex differentiable at $z_0 = (x_0, y_0)$ if and only if

$$\begin{aligned} \partial_x u(x_0,y_0) &= \partial_y v(x_0,y_0), \\ \partial_y u(x_0,y_0) &= -\partial_x v(x_0,y_0). \end{aligned}$$

Several partial differential equations also occur in the field of *differential geometry*. For example, if we search for minimal surfaces with prescribed boundary we may use functions $u\colon \overline{\Omega} \to \mathbb{R}$ such that $u|_{\partial\Omega} = f$ is given. Then the surface area of the graph of u, given by $\int_\Omega \sqrt{1 + |\nabla u(x)|^2} dx$, is minimal if and only if the Euler–Lagrange equations

$$\nabla \cdot \left(\frac{\nabla u}{\sqrt{1 + |\nabla|^2}} \right) = 0$$

are satisfied. This **minimal surface equation** is a quasilinear partial differential equation of second order. The minimal surface problem is also known as the *Plateau problem*, named after the Belgian physicist J.A.F. Plateau (1801 - 1883). He conducted experiments with soap films by dipping wire contours in a solution of soapy water.

Several partial differential equations in geometry are also closely related to physics, such as symplectic geometry and Hamiltonian mechanics (more below).

1.3.2. Physics. In *Hamiltonian mechanics* a classical physical system is described using a set of canonical coordinates (q,p), with generalized coordinates q_i and conjugate momenta $p_i = \frac{\partial L}{\partial \dot{q}_i}$. Here $L = T - U$ is the Lagrangian with kinetic energy T and potential energy U. The Hamiltonian/total energy is $H = T + U$, which via the Legendre transform is $H = \sum_i p_i \dot{q}_i - L$ and the equations of motion (for example, for planets) are

$$\begin{aligned} \dot{q}_i &= \frac{\partial H}{\partial p_i}, \\ \dot{p}_i &= -\frac{\partial H}{\partial q_i}, \end{aligned}$$

consisting of $2n$ first order partial differential equations.

A canonical transformation with a generating function u turns the above equations into the **Hamilton–Jacobi equation**, a system of partial differential equations of the form

$$\partial_t u + H(Du, q) = 0.$$

For the concept of viscosity solutions (some kind of weak solutions) Pierre-Louis Lions received the Fields medal in 1994.

In *hydrodynamics* one is interested in modeling the motion of fluids in $\mathbb{R}^n$. Let I be an open interval. A fluid is described by its velocity vector $u(x,t) = (u_1(x,t), \dots, u_n(x,t)) \colon I \times \mathbb{R}^n \to \mathbb{R}^n$, the pressure $p \colon I \times \mathbb{R}^n \to \mathbb{R}$, and viscosity $\nu > 0$. For incompressible fluids with external force $f \colon \mathbb{R}^n \to \mathbb{R}^n$ the flow is described by the **Navier–Stokes equations**

$$\begin{aligned} \partial_t u + (u \cdot \nabla)u &= \nu \Delta u - \nabla p + f, \\ \operatorname{div} u = \nabla \cdot u &= 0. \end{aligned}$$

on $I \times \mathbb{R}^n$. The Navier–Stokes equations model the weather, ocean currents, water flow in a pipe and air flow around a wing etc. and are very important for applications. In spite of that they are also of great mathematical interest. As of now it still has not been proven whether solutions exist in three dimension and, if so, whether they are smooth. The Clay Mathematics Institute calls this one of the seven most important open problems in mathematics and offers an award of one million US dollars for a solution or counterexample.

Note that inviscid fluids are more easily modeled by the **Euler equations**. These quasi-linear hyperbolic equations are simpler because they have a conservation law form, i.e., are of the form

$$\partial_t u + \operatorname{div} F(u) = 0,$$

for a flux function $F \in C^1(\mathbb{R}, \mathbb{R}^n)$.

The **Korteweg–de Vries equation**

$$\partial_t u(t,x) - u(t,x) u_x(t,x) + u_{xxx}(t,x) = 0,$$

describes shallow water waves in narrow channels and can predict the formation of *solitons*, i.e., wave packets that maintain its shape and travel with a constant speed. It is a semilinear partial differential equation of third order.

The history of the Korteweg–de Vries equation goes back to observations and experiments by J.S. Russell in 1834. He discovered the phenomenon of solitons when observing a boat that was first drawn along a narrow channel and then suddenly stopped. The mass of water which the boat had put in motion accumulated and rolled forward, forming a rounded, well-defined heap. Russel followed this heap on his horse for several kilometers and noticed that it seemed to travel along the channel without changing its form or speed.

The **Maxwell equations** form the basis for the theory of *electromagnetism* (and hence also applications in electrical engineering), given by

$$\begin{aligned}\nabla \cdot \vec{E} &= 4\pi\rho,\\ \nabla \times \vec{E} &= -\frac{1}{c}\partial_t \vec{B},\\ \nabla \cdot \vec{B} &= 0,\\ \nabla \times \vec{B} &= \frac{4\pi}{c}\vec{j} + \frac{1}{c}\partial_t \vec{E},\end{aligned}$$

where $\vec{E}$ denotes the electrical field, $\vec{B}$ the magnetic field, ρ the total electric charge density, $\vec{j}$ the total electric current density and c the speed of light. The symbol ∇ is the 3-dimensional gradient operator (pronounced "del"), $\nabla\cdot$ is the divergence operator ("del dot") and $\nabla\times$ is the curl operator ("del cross").

The Maxwell equations form a system of linear partial differential equations, and hence the solutions are superposable. In some cases, the equations simplify further. If, for a potential u, $\vec{E} = \nabla u$, then the first equation reduces to the *Poisson equation* $\Delta u = 4\pi\rho$. In particular, if $\rho = 0$ then this is the *Laplace equation* $\Delta u = 0$. (More about the Laplace and Poisson equation later.)

In regions with no currents ($\vec{j} = 0$) and constant charges ($\rho = 0$) this leads to *electromagnetic waves*. This is due to the fact that, on one hand

$$\nabla \times (\nabla \times \vec{E}) = -\frac{1}{c}\nabla \times \vec{B} = -\frac{1}{c}\partial_{tt}\vec{E},$$

by the Maxwell equations, and on the other hand

$$\nabla \times (\nabla \times \vec{E}) = \nabla(\nabla \vec{E}) - \nabla^2 \vec{E} = 0 - \Delta\vec{E},$$

so that

$$-\partial_{tt}\vec{E} + c^2\Delta\vec{E} = 0,$$

which is the *wave equation*. (More about the wave equation later.)

In *quantum mechanics*, the **Schrödinger equation** is a linear partial differential equation that describes the changes over time of a physical system in which quantum effects are significant. The time-dependent version for a single (nonrelativistic) particle reads

$$i\hbar\frac{\partial\Psi}{\partial t} = -\frac{\hbar^2}{2\mu}\Delta\Psi + V(x)\Psi,$$

where $\hbar = \frac{h}{2\pi}$ is the reduced Planck constant, μ is the reduced mass, V the potential energy and the solution Ψ is the so-called wave function of the quantum particle. For the formulation of this equation Erwin Schrödinger received the Nobel price in physics in 1933.

In *general relativity* the **Einstein equations** relate the geometry of spacetime to the mass and energy distribution of the universe. More precisely, the Einstein tensor $\mathbf{G}$ reflecting the curvature of spacetime (a combination of second derivatives of the metric) should satisfy

$$\mathbf{G} = \frac{8\pi G}{c^4}\mathbf{T},$$

where $\mathbf{T}$ is the energy-momentum tensor, G Newton's gravitational constant and c the speed of light. In local coordinates these are 10 coupled, nonlinear partial differential equations.

Certain exact solutions are known, but in general one is interested in studying the time evolution of arbitrary initial data.

1.3.3. Chemistry. Reaction-diffusion equations are widely used to model phenomena in chemistry, physics and biology. They describe the changes in space and time of concentrations of chemical substances or densities of populations.

Let I be an open interval, $U \subseteq \mathbb{R}^n$ be open and $\Omega = I \times U$. Moreover, $u\colon \Omega \to \mathbb{R}$ is a function of time $t \in I$ and the spatial position $x \in U$. A *reaction-diffusion equation* is of the form

$$\partial_t u = d\Delta u + f(u) \qquad \text{in } I \times U,$$

where $\Delta u = \Delta_x u = \sum_{i=1}^n u_{x_i x_i}$ denotes the *Laplace operator* (or simply *Laplacian*) with respect to x and $d > 0$ is the diffusion coefficient. The first term on the right hand side of the equation models the diffusion (particles or individuals move from regions with high concentrations to regions of low concentrations) and the given function $f\colon \mathbb{R} \to \mathbb{R}$ describes local reactions.

It is a semilinear partial differential equation of second order.

More generally, we can consider *reaction diffusion systems*,

$$\partial_t u = D\Delta u + f(u) \qquad \text{in } I \times U,$$

where $u = (u_1, \dots, u_m)$, $D \in \mathbb{R}^{m\times m}$ is a diagonal matrix with positive coefficients and $f\colon \mathbb{R}^m \to \mathbb{R}^m$ a given function. They are used to model, e.g., ecological invasions, the spread of epidemics, tumor growth or reactions between several different chemical substances.

1.3.4. Economics. Another equation for which the Nobel price was awarded in 1997 is the *Black–Scholes–Merton equation* which models the dynamics of a financial market containing derivative investment instruments. If V is the price of the derivative, S the price of the stock, ρ the volatility and r the interest rate, then the the change of V can be described by the partial differential equation

$$\partial_t V + \frac{1}{2}\rho^2 S^2 \frac{\partial^2 V}{\partial S^2} - rS\frac{\partial V}{\partial S} - rV = 0.$$

This model is widely used to "eliminate risk". However, in practice also a lot of restrictions apply, especially longterm.

1.4. Strategies for solving/analyzing partial differential equations

In Section 1.3 we will see some examples that illustrate how broad the field of partial differential equations is. Unlike for ordinardy differential equations, a general theory for solving all partial differential equations does not exist. Already the question what a *solution* is cannot be answered in a uniform way.

1.4.1. Solutions and well-posedness. In principle, we distinguish between classical and weak solutions. If the equation is of order k, then a *classical* solution must be of order C^k. Such solutions exist in some cases but the occurrence of shocks (essentially the appearance of discontinuities in derivatives at some point of time), for example, requires the notion of *weak* solutions. Often it can also be easier to prove the existence of weak solutions first and only later show that the solutions actually have a higher regularity.

Ideally, we find explicit solutions for a given partial differential equation, but this is only possible in particular, simple cases and rather the exception than the norm. This classical approach to partial differential equations that dominated the 19th century was to develop methods for deriving explicit representation formulas for solutions. If such formulas cannot be found, we aim at proving the existence and qualitative properties of solutions. Instead, we are interested in **well-posedness**, which involves the following qualitative properties:

(i) There is a (classical or weak) solution.
(ii) The solution is unique.
(iii) The solution depends continuously on the data.

The last property is particularly important in applications, since the solution should change only slightly if we vary the data specifying the problem only a little.

1.4.2. Initial and boundary data. By data in (iii) we mean initial or boundary *data* which are often prescribed on parts (or all) of Ω or $\partial\Omega$.

If data is prescribed on $\partial\Omega$ we call it a *boundary value problem*. Boundary data is given in a way that the solution itself is prescribed on $\partial\Omega$ we talk about a *Dirichlet problem*. If only the derivatives of a solution is prescribed on $\partial\Omega$ it is called a *von Neumann problem*.

A *Cauchy problem* or *initial value problem* is a partial differential equation for which data are prescribed on all of Ω at some initial time, say $t = 0$. For a partial differential equation of order k we need $k - 1$ initial conditions.

Several combinations also exit.

1.4.3. Analysis of partial differential equations. As already mentioned, different from ordinary differential equations there is no general theory or approach for the solvability of partial differential equations. Except for few specific cases, research in PDEs focuses on various, particular partial differential equations that are relevant in applications and on the development of specific methods for the problem at hand.

Typically, the difficulty of the analysis of a partial differential equation increases with the degree of nonlinearity, with the order k of the partial differential equation, with the number of variables n and with the number m of equations (i.e., systems of partial differential equations are generally more difficult to analyze than scalar equations).

In this course we mainly focus on simple prototypes of linear second order partial differential equations (Laplace, heat and wave equation) and nonlinear partial differential equations of first order. Typical questions we address are

- existence and uniqueness of solutions,
- qualitative properties of solutions (e.g., regularity, dependence on data),
- explicit representation formulas for solutions,
- limitations of classical solutions.

Part 1

Elementary linear partial differential equations

In this part we discuss the basic properties of the linear transport equation and the three archetypes of second order linear partial differential equations (Laplace, heat and wave equation). We focus on the representation of classical solutions and their properties. The presentation closely follows that of Evans [**4**]. In the case of second order elliptic equations, Gilbarg and Trudinger [**5**] is also a standard reference.

CHAPTER 2

The transport equation

2.1. Motivation

Consider a (one-dimensional) river with a pollutant moving along it (one may also consider cars driving on a road, and other examples).

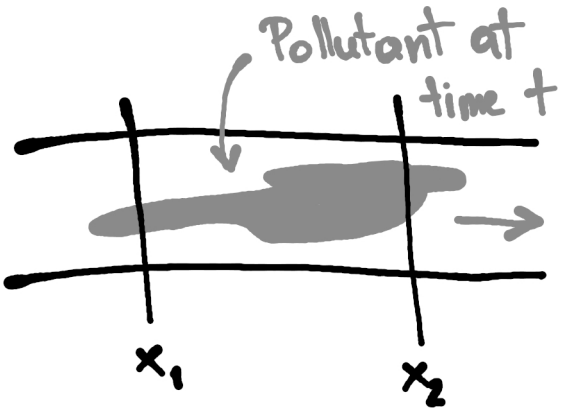

FIGURE 1. A pollutant moving along a river. The mass is conserved.

Let $u = u(t, x)$ denote the concentration of this pollutant at time t and position x. Assume that the total amount of pollutant u remains constant and we want to calculate the change in substance in some spatial interval (x_1, x_2). At time t the total amount of u in (x_1, x_2) is

$$\int_{x_1}^{x_2} u(t, x)\, dx,$$

and the change at time t is (assuming u is sufficiently smooth)

$$\frac{d}{dt}\int_{x_1}^{x_2} u(t, x)\, dx = \int_{x_1}^{x_2} u_t(t, x)\, dx$$

If $q = q(t, x)$ denotes the flux of the substance, i.e., the change of the substance per time unit at the point x, then the same change at time t is given by

$$q(t, x_1) - q(t, x_2),$$

because $q(t, x_1)$ tells us how much enters at time t at x_1 and how much substance leaves at x_2. Due to the conservation of the mass of the pollutant in (x_1, x_2) we can equate

$$\int_{x_1}^{x_2} u_t(t, x)\, dx = q(t, x_1) - q(t, x_2) = -\int_{x_1}^{x_2} q_x(t, x)\, dx,$$

where the last equality is due to the fundamental theorem of calculus. Since the interval was arbitrary, we conclude[1] that the pollutant must satisfy the equation

$$u_t + q_x = 0.$$

It remains to relate u and q. In the simplest case,

$$q(t,x) = cu(t,x),$$

for some constant $c \neq 0$. This leads to the *one-dimensional linear transport equation*

$$u_t + cu_x = 0.$$

Since this is a physical problem, one would also need some *initial condition*

$$u(0,x) = g(x),$$

for some given function g representing the initial density at time 0.

Note that this model can easily be extended to two and more dimensions by considering that the quantity is conserved along a flow of a vector field (which can be represented by integral curves), i.e., q_x would be replaced by $X(u)$ for some vector field X. In the simplest linear and constant case, this means replacing the constant c by a given (constant) vector field $b \in \mathbb{R}^n$ and u_x by the gradient ∇u so that the *linear transport equation with constant coefficients* reads

$$u_t + b \cdot \nabla u = 0. \tag{2.1}$$

For example, in fluid dynamics, the *continuity equation* expresses the law of mass conservation, it is of the form

$$\rho_t + \nabla \cdot (\rho v) = 0, \qquad (0,\infty) \times \mathbb{R}^3,$$

where $\rho\colon (0,\infty) \times \mathbb{R}^3 \to \mathbb{R}^3$ denotes the density of the fluid, $v\colon (0,\infty) \times \mathbb{R}^3 \to \mathbb{R}^3$ the velocity field and ∇ the gradient with respect to x. If the velocity is constant we are back in the simple case (2.1).

2.2. The homogeneous case

The linear transport equation with constant coefficients is the simplest partial differential equation. As described in Section 2.1, for a fixed vector $b \in \mathbb{R}^n$ we consider the partial differential equation

$$u_t + b \cdot \nabla u = 0, \qquad \text{in } (0,\infty) \times \mathbb{R}^n. \tag{2.2}$$

A solution $u = u(t,x)$ is a function $(0,\infty) \times \mathbb{R}^n \to \mathbb{R}$ with spatial gradient $\nabla u = \nabla_x u = (u_{x_1}, \dots, u_{x_n})$. Here $t \geq 0$ should represent the time and $x \in \mathbb{R}^n$ the position in space.

Note that if u is a C^1 solution of (2.2) then the directional derivative of u in direction $(1,b)$ vanishes. More precisely, for

$$z(s) = u(t+s, x+sb),$$

we have that

$$\dot{z}(s) = u_t(t+s, x+sb) + \nabla u(t+s, x+sb) \cdot b = 0.$$

[1] See problem 1 of Assignment 1.

This shows that z is constant, and thus $u(t,x)$ is constant along the line

$$s \mapsto \begin{pmatrix} t \\ x \end{pmatrix} + s \begin{pmatrix} 1 \\ b \end{pmatrix}.$$

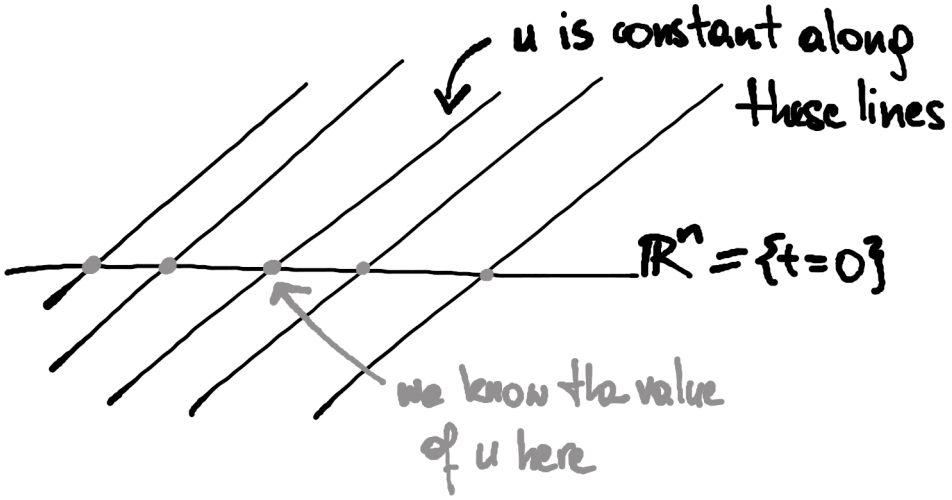

FIGURE 2. The solution is u is constant along the lines. Since we know the value at one point of these lines (at the initial time $t = 0$), hence everywhere.

For a given $g\colon \mathbb{R}^n \to \mathbb{R}$ we consider the *initial value problem* (for short, IVP)

$$\begin{aligned} u_t + b \cdot \nabla u &= 0, \qquad \text{in } (0,\infty) \times \mathbb{R}^n, \\ u(0,.) &= g, \qquad \text{on } \{t = 0\} \times \mathbb{R}^n. \end{aligned} \tag{2.3}$$

By the above (conservation along the line), we must have that every solution u satisfies

$$u(t,x) = u(0, x - tb) = g(x - tb), \qquad t \geq 0, x \in \mathbb{R}^n. \tag{2.4}$$

Conversely, if g is C^1, by differentiating (2.4) one can easily show that it is always a solution to the initial value problem (2.3).

In the beginning we noted that we are looking for a *classical* solution of (2.3). One can show that if g is not C^1, then such a C^1 solution does not exist. Therefore, we only consider this case at the moment. We summarize what we have shown.

THEOREM 2.1 (Solving the linear transport equation). *Consider the linear transport equation with constant coefficients, i.e., the partial differential equation*

$$u_t + b \cdot \nabla u = 0, \qquad \text{in } (0,\infty) \times \mathbb{R}^n. \tag{2.5}$$

for some fixed $b \in \mathbb{R}^n$. Then the following holds:

(i) *If u is a classical (C^1) solution of (2.5), then u satisfies*

$$u(t + s, x + sb) \equiv \text{const.},$$

for all $(t,x) \in (0,\infty) \times \mathbb{R}^n$, $s > -t$.

(ii) *Let $g \in C^1(\mathbb{R}^n)$ be given. Then the initial value problem*

$$\begin{aligned} u_t + b \cdot \nabla u &= 0, \qquad \text{in } (0,\infty) \times \mathbb{R}^n, \\ u(0,.) &= g, \qquad \text{on } \{t = 0\} \times \mathbb{R}^n, \end{aligned}$$

has a unique solution $u \in C^1((0,\infty) \times \mathbb{R}^n) \cap C([0,\infty) \times \mathbb{R}^n)$, which is given by

$$u(t,x) = g(x - tb),$$

for all $(t,x) \in [0,\infty) \times \mathbb{R}^n$. □

2.3. The inhomogeneous case

More generally, we can consider the linear transport equation with a nonvanishing right hand side, i.e., the inhomogeneous initial value problem

$$\begin{aligned} u_t + b \cdot \nabla u &= f, \qquad \text{in } (0,\infty) \times \mathbb{R}^n, \\ u(0,.) &= g, \qquad \text{on } \mathbb{R}^n, \end{aligned} \tag{2.6}$$

for some functions $f\colon (0,\infty) \times \mathbb{R}^n \to \mathbb{R}$ and $g\colon \mathbb{R}^n \to \mathbb{R}$. As before, the function

$$z(s) = u(t + s, x + sb),$$

satisfies

$$\dot{z}(s) = u_t(t + s, x + sb) + \nabla u(t + s, x + sb) \cdot b = f(t + s, x + sb).$$

Integration from $-t$ to 0 yields

$$\begin{aligned} u(t,x) - g(x - tb) = z(0) - z(-t) &= \int_{-t}^{0} \dot{z}(s)\, ds \\ &= \int_{-t}^{0} f(t + s, x + sb)\, ds = \int_0^t f(r, x + (r - t)b)\, dr, \end{aligned}$$

and we therefore obtain the following result.

THEOREM 2.2 (Solving the inhomogeneous linear transport equation). *Let $b \in \mathbb{R}^n$, $f \in C^1((0,\infty) \times \mathbb{R}^n)$ and $g \in C^1(\mathbb{R}^n)$. Then the initial value problem* (2.6) *has a unique classical solution $u \in C^1((0,\infty) \times \mathbb{R}^n) \cap C(\overline{(0,\infty) \times \mathbb{R}^n})$, which is given by*

$$u(t,x) = g(x - tb) + \int_0^t f(s, x + (s - t)b)\, ds \tag{2.7}$$

for all $(t,x) \in [0,\infty) \times \mathbb{R}^n$.

SKETCH OF PROOF. Existence follows by differentiation of (2.7). Uniqueness is due to the derivation before the statement of the theorem (any potential solution must satisfy the formula (2.7), which is unique). □

REMARK 2.3. We have essentially solved the transport equation by converting it to a family of ordinary differential equations along specific curves (here, straight lines). This technique is called the *method of characteristics* and can also be applied successfully to more complicated partial differential equations. More about this in Chapter **??**.

REMARK 2.4. The solution formula (2.7) also makes sense for nondifferentiable functions f and g (they must not even be continuous). Such a definition of *weak solutions* extends the concept of classical solutions. This idea is essential for any satisfying theory on partial differential equations and will be formalized further later (though it is not part of this course to cover existence results for weak solutions).

CHAPTER 3

The Laplace equation

In this chapter we study the Laplace equation and its inhomogeneous version, the Poisson equation. We start with a derivation of these equations.

3.1. Motivation

Let $\Omega \subseteq \mathbb{R}^n$. Suppose $u\colon \Omega \to \mathbb{R}$ describes the density of a physical quantity in equilibrium. For an arbitrary open set $V \subseteq \Omega$ (with C^1 boundary), since the amount $\int_V u$ of the quantity in V does not change, the total flow of the quantity through the boundary ∂V must vanish. That is, for some flux function given be a vector field $F\colon \Omega \to \mathbb{R}^n$ and $\nu\colon \partial V \to \mathbb{R}^n$ the outer

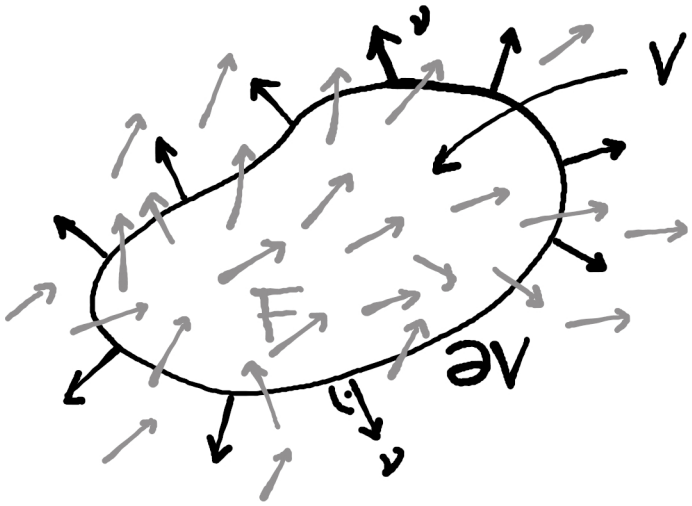

FIGURE 1. A physical quantity (say, temperature) in equilibrium.

unit normal vector field we must have that the total flux through the surface vanishes, i.e.,

$$\int_{\partial V} F \cdot \nu \, dS = 0.$$

The Divergence Theorem[1] then implies

$$\int_V \operatorname{div} F = \int_{\partial V} F \cdot \nu \, dS = 0.$$

In many cases the flux function satisfies

$$F(x) = -a \nabla u(x),$$

[1]See prerequisites on integration theory on Brightspace.

for some $a > 0$, i.e., the flux is parallel to ∇u and points from high to low density (Fick's law of diffusion). Then

$$\operatorname{div} F = -a \operatorname{div} \nabla u = -a(u_{x_1x_1} + \ldots + u_{x_nx_n}) = -a\Delta u,$$

and thus

$$-\int_V a\Delta u = 0.$$

If $u \in C^2(\Omega)$ then the integrand is continuous and it follows that[2] $-a\Delta u = 0$, and therefore

$$\Delta u = 0.$$

This is the *Laplace equation*.

In many cases a physical system has an additional source Q. Then the flow through the boundary does not vanish but equals the amount generated by Q, i.e.,

$$\int_{\partial V} F \cdot \nu \, dS = \int_V Q.$$

In a similar fashion one obtains

$$-a\Delta u = Q,$$

and by setting $f = \frac{Q}{a}$ we obtain the *Poisson equation*

$$-\Delta u = f.$$

Typical applications of the Poisson equation are the steady-state temperature in a solid (u is temperature, f is the heat source strength), static deflection of a thin membrane in 2D (u is deflection, f is the pressure), electrostatics (u is the electrostatic potential, f is the charge per unit volume), and Newtonian gravity (u is the gravitational potential, f is the mass density).

3.2. Harmonic functions

We first introduce some terminology.

Definition 3.1. Let $\Omega \subseteq \mathbb{R}^n$ be open, $u\colon \Omega \to \mathbb{R}$ be the unknown, and $f\colon \Omega \to \mathbb{R}$ be given. The *Laplace equation* reads

$$\Delta u = \sum_{i=1}^{n} u_{x_ix_i} = 0, \tag{3.1}$$

the *Poisson equation* is

$$-\Delta u = f. \tag{3.2}$$

Definition 3.2. Let $\Omega \subseteq \mathbb{R}^n$. A function $u \in C^2(\Omega)$ that satisfies the Laplace equation (3.1) is called *harmonic*. It is called *subharmonic* if $-\Delta u \leq 0$, and *superharmonic* if $-\Delta u \geq 0$ on Ω.

Example 3.3. The Cauchy–Riemann equations for analytic functions imply that their real and imaginary part are harmonic functions (see Section 1.3.1 and problem 2 in Assignment 2).

[2] Essentially problem 1 in Assignment 1, but here for higher dimensions.

3.3. Fundamental solution

We first search for a special explicit solution of the Laplace equation. Since this partial differential equation is linear, solutions can be superposed, leading to more complicated solutions. For such purposes it is often useful to apply certain symmetry properties of the equation. In this case we use that the Laplace equation is *rotation invariant*[3]. That is, we look for radial solutions u of (3.1) that are of the form

$$u(x) = v(r),$$

where $r = |x| = \sqrt{\sum_{i=1}^n x_i^2}$. Since $\frac{\partial r}{\partial x_i}(x) = \frac{x_i}{r}$ for $x \neq 0$ we obtain that the partial derivatives of u are of the form

$$u_{x_i} = v'(r)\frac{x_i}{r}, \qquad u_{x_i x_i} = v''(r)\frac{x_i^2}{r^2} + v'(r)\left(\frac{1}{r} - \frac{x_i^2}{r^3}\right),$$

so that (3.1) is equivalent to

$$\Delta u = v''(r) + \frac{n-1}{r}v'(r) = 0.$$

If $v' \neq 0$ then division yields

$$(\log v')' = \frac{v''}{v'} = \frac{1-n}{r} = (1-n)(\log r)'.$$

Hence for some constant a,

$$\log v' = (1-n)\log r + a = \log r^{1-n} + a,$$

which means for $b = e^a$

$$v' = \frac{e^a}{r^{n-1}}.$$

Thus for some constants $b, c \in \mathbb{R}$, and $r > 0$, integration yields

$$v(r) = \begin{cases} b\log r + c & \text{if } n = 2, \\ \frac{b}{(2-n)r^{n-2}} + c & \text{if } n \geq 3. \end{cases}$$

For the particular choice

$$b = -\frac{1}{|\partial B_1(0)|} = -\frac{1}{\omega_n}, \qquad c = 0,$$

where ω_n denotes the surface area of the unit sphere[4], we obtain the so-called fundamental solution of the Laplace equation (we will see later why this particular choice of constants is useful).

Definition 3.4. The function $\Phi\colon \mathbb{R}^n \setminus \{0\} \to \mathbb{R}$, defined by

$$\Phi(x) = \begin{cases} -\frac{1}{2\pi}\log|x| & \text{if } n = 2, \\ \frac{1}{(n-2)\omega_n}\frac{1}{|x|^{n-2}} & \text{if } n \geq 3, \end{cases} \tag{3.3}$$

is called the *fundamental solution of the Laplace equation*.

[3]See Assignment 2.

[4]Note that $\Omega_n = \frac{\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2})+1}$ is the volume of the unit ball and $\omega_n = n\Omega_n$ the surface area of the unit sphere in n dimensions. You will find a derivation of this formula in some Analysis books (or try it as an exercise).

We often write $\Phi(x) = \Phi(|x|)$, since Φ only depends on the norm of x (this is a slight abuse of notation though).

REMARK 3.5. By construction $\Delta\Phi = 0$ on $\mathbb{R}^n \setminus \{0\}$, but note that Φ has a singularity at the origin. Moreover, for $x \neq 0$ the total derivatives satisfy

$$|D\Phi(x)| \leq \frac{c}{|x|^{n-1}}, \qquad |D^2\Phi| \leq \frac{c}{|x|^n}. \tag{3.4}$$

for some constant $c > 0$[5]. In particular, $\Delta\Phi$ is not integrable at the origin. It is precisely this property that allows us to construct solutions to the Dirichlet problem on Ω later.

To show (3.4) is a short calculation. For $n \geq 3$, we have $\Phi(x) = \frac{1}{(n-2)\omega_n} \frac{1}{r^{n-2}}$ and thus

$$\Phi_{x_i}(x) = -\frac{1}{\omega_n} r^{1-n} \frac{x_i}{r} = -\frac{1}{\omega_n} \frac{x_i}{r^n}$$

which implies $|\Phi_{x_i}| \leq \frac{c}{r^{n-1}}$ and

$$D\Phi(x) = -\frac{1}{\omega_n} \frac{x}{|x|^n},$$

as well as

$$\Phi_{x_i x_j}(x) = -\frac{1}{\omega_n} \left(\frac{\delta_{ij}}{|x|^n} - n \frac{x_i x_j}{|x|^{n+2}} \right),$$

which yields $|\Phi_{x_i x_j}| \leq \frac{c}{r^n}$. For $n = 2$ it follows in a similar fashion.

The function $x \mapsto \Phi(x)$ is harmonic for $x \neq 0$, and similarly, by shifting the origin, for any $y \in \mathbb{R}^n$ the function $x \mapsto \Phi(x - y)$ is harmonic for $x \neq y$. Moreover, taking a function $f \colon \mathbb{R}^n \to \mathbb{R}$, then $x \mapsto f(y)\Phi(x - y)$ is harmonic for every $y \in \mathbb{R}^n$, $x \neq y$, and thus, the same applies to the sum of finitely many such expressions. This might suggest that the *convolution*

$$u(x) = \int_{\mathbb{R}^n} \Phi(x - y) f(y)\, dy$$

is a solution of the Laplace equation (3.1). However, this is wrong since $\Delta\Phi$ is not integrable near the singularity at $x = y$, and thus, interchanging differentiation and integration is not possible. In fact, the function u is not harmonic, but yields a solution of the Poisson equation in $\Omega = \mathbb{R}^n$.

THEOREM 3.6 (Solving Poisson's equation). *Suppose $f \in C_c^2(\mathbb{R}^n)$, i.e., twice continuously differentiable with compact support. Then*

$$u(x) = \int_{\mathbb{R}^n} \Phi(x - y) f(y)\, dy$$

satisfies

(i) *$u \in C^2(\mathbb{R}^n)$, and*
(ii) *$-\Delta u = f$ on $\mathbb{R}^n$.*

[5]As in Analysis it often does not matter how this constant $c > 0$ exactly looks like and therefore we often even use the same letter for different constants. What is important is the decay/integrability with respect to x, which is not affected by the actual value of c.

SKETCH OF PROOF. Roughly speaking, (i) follows from differentation (involves the Mean Value Theorem!) and for (ii) we have to split up the integral into two regions $B_\varepsilon(0)$ and $\mathbb{R}^n \setminus B_\varepsilon(0)$, do some partial integration etc. and take the limit $\varepsilon \to 0$. Of course we also make use of the fact that Φ is harmonic on $\mathbb{R}^n \setminus B_\varepsilon(0)$. (See the book by Evans [**4**, Theorem 1, p. 23] for all the details.) □

REMARK 3.7. Formally we write

$$-\Delta\Phi = \delta_0,$$

where δ is the *delta distribution* (or Dirac delta function). Heuristically, one can calculate

$$-\Delta u(x) = \int_{\mathbb{R}^n} -\Delta_x \Phi(x-y) f(y)\, dy = \int_{\mathbb{R}^n} \delta_x(y) f(y)\, dy = f(x).$$

To make this exact one needs to use the theory of distributions (that is, more functional analysis).

In Section 3.5 we will consider the Poisson equation in bounded domains Ω and use the fundamental solution Φ in (3.3) to construct a representation formula for the solution of the Dirichlet problem.

3.4. Properties of harmonic functions

We prove some crucial properties of solutions of the Laplace and Poisson equations.

3.4.1. Mean value formulas. Let $\Omega \subseteq \mathbb{R}^n$ and consider balls satisfying $\overline{B_r(x)} \subseteq \Omega$. For an integrable function u, the average value of u over balls and spheres (or more general regions) is defined by

$$\fint_{B_r(x)} u(y)\, dy := \frac{1}{|B_r(x)|} \int_{B_r(x)} u(y)\, dy = \frac{1}{\Omega_n r^n} \int_{B_r(x)} u(y)\, dy,$$
$$\fint_{\partial B_r(x)} u(y)\, dS(y) := \frac{1}{|\partial B_r(x)|} \int_{\partial B_r(x)} u(y)\, dS(y) = \frac{1}{\omega_n r^{n-1}} \int_{\partial B_r(x)} u(y)\, dS(y),$$

where Ω_n and $\omega_n = n\Omega_n$ denote the volume of the unit ball $B_1(0)$ and sphere $\partial B_1(0)$ in $\mathbb{R}^n$, respectively. If u is continuous then one can show that[6]

$$\lim_{r\to 0} \fint_{B_r(x)} u(y)\, dy = \lim_{r\to 0} \fint_{\partial B_r(x)} u(y)\, dS(y) = u(x) \tag{3.5}$$

We show that in the case of harmonic functions one does not even have to pass to the limit to obtain such an identity.

THEOREM 3.8 (Mean value property). *Let $\Omega \subseteq \mathbb{R}^n$ be an open set and $u \in C^2(\Omega)$. If u is harmonic, then*

$$u(x) = \fint_{B_r(x)} u(y)\, dy = \fint_{\partial B_r(x)} u(y)\, dS(y) \tag{3.6}$$

for each ball $\overline{B_r(x)} \subseteq \Omega$.

[6]See problem 6(b) on Assignment 2.

PROOF. For a fixed $x \in \Omega$ consider the auxilliary function

$$\begin{aligned}\varphi(r) :&= \fint_{\partial B_r(x)} u(y)\, dS(y) = \frac{1}{\omega_n r^{n-1}} \int_{\partial B_r(x)} u(y)\, dS(y) \\ &= \frac{1}{\omega_n} \int_{\partial B_1(0)} u(x+rz)\, dS(z) = \fint_{\partial B_1(0)} u(x+rz)\, dS(z),\end{aligned}$$

where we performed a change of coordinates $y = x + rz$ in the second line. Differentiation

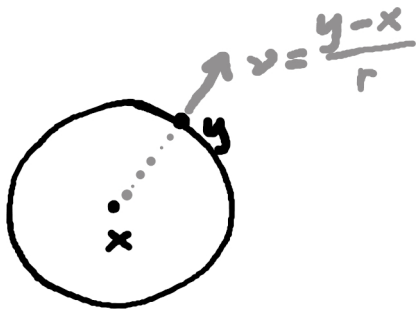

FIGURE 2. The sphere $\partial B_r(x)$ with outward point unit normal vector ν at the point y.

implies, since $\nu = \frac{y-x}{r}$ is the outer unit normal vector of $\partial B_r(x)$ at the point y,

$$\begin{aligned}\varphi'(r) &= \fint_{\partial B_1(0)} Du(x+rz) \cdot z\, dS(z) = \fint_{\partial B_r(x)} Du(y) \cdot \frac{y-x}{r}\, dS(y) \\ &= \fint_{\partial B_r(x)} \frac{\partial u}{\partial \nu}(y)\, dS(y) = \frac{1}{|\partial B_r(x)|} \int_{\partial B_r(x)} \frac{\partial u}{\partial \nu}(y)\, dS(y),\end{aligned}$$

which by Green's formula[7] and the fact that $|\partial B_r(x)| = \frac{n}{r}|B_r(x)|$ yields

$$\varphi'(r) = \frac{r}{n} \frac{1}{|B_r(x)|} \int_{B_r(x)} \Delta u(y)\, dy = 0,$$

since u is harmonic by assumption. Therefore φ is constant, and thus

$$\varphi(r) = \lim_{s \to 0} \varphi(s) = \lim_{s \to 0} \fint_{\partial B_s(x)} u(y)\, dS(y),$$

which by continuity (see (3.5)) implies that the expression on the right hand side equals $u(x)$. Hence

$$\fint_{\partial B_r(x)} u(y)\, dS(y) = u(x),$$

[7]See problem 2b(i) of Assignment 1.

as desired. Moreover, according to problem 6(a) in Assignment 2, this is statement is equivalent to

$$\begin{aligned}\int_{B_r(x)} u(y)\,dy &= \int_0^r \int_{\partial B_s(x)} u(y)\,dS(y)ds \\ &= u(x)\int_0^r \omega_n s^{n-1} ds \\ &= \Omega_n r^n u(x),\end{aligned}$$

which is the second equality. □

REMARK 3.9. Mean value formulas can also be formulated for sub- und superharmonic functions. If u is subharmonic on Ω then it follows that

$$\begin{aligned}u(x) &\leq \fint_{B_r(x)} u(y)\,dy, \\ u(x) &\leq \fint_{\partial B_r(x)} u(y)\,dS(y).\end{aligned} \tag{3.7}$$

If u is superharmonic the sign is reversed, i.e., $\geq$. This is easy to see since $-u$ is then subharmonic. The properties (3.7) for subharmonic functions therefore imply mean value formula (3.6) for harmonic functions.

The converse of Theorem 3.8 also holds.

THEOREM 3.10. *Let $\Omega \subseteq \mathbb{R}^n$ be open. Suppose $u \in C^2(\Omega)$ satisfies*

$$u(x) = \fint_{\partial B_r(x)} u(y)\,dS(y)$$

for all $\overline{B_r(x)} \subseteq \Omega$. Then u is harmonic on Ω.

PROOF. See Assignment 3. □

3.4.2. Maximum principle and uniqueness for boundary value problems. The maximum principle is an important property of solutions of elliptic and parabolic partial differential equations. Roughly speaking, it states that the maximum of a function in a domain is attained on its boundary. In the weak version the maximum may also reoccur in the center of the domain. In the strong version, if that is the case then the function must already be constant. We will also see how this property can be used to prove uniqueness of solutions to the Laplace equation.

THEOREM 3.11. *Let Ω be a bounded domain*[8] *in $\mathbb{R}^n$. If $u \in C^2(\Omega) \cap C(\overline{\Omega})$ is subharmonic on Ω then the following holds:*

(i) Maximum principle:

$$\max_{\overline{\Omega}} u = \max_{\partial\Omega} u$$

(ii) Strong maximum principle: *If there exists an $x_0 \in \Omega$ such that*

$$u(x_0) = \max_{\overline{\Omega}} u,$$

then u is constant on Ω.

[8]We use the word *domain* in the topological sense, that is, a domain is an open and connected set.

In short, the proof of (ii) follows from the mean value formulas, and (ii) implies (i). Note that the connecteness of Ω is not important for (i) because it holds on each connected component, but (ii) only holds for connected Ω.

PROOF. Since Ω is bounded, the sets $\overline{\Omega}$ and $\partial\Omega$ are compact. Thus, by the continuity of u, the maxima exist.

We first observe that (i) is a consequence of (ii). In fact, applying (ii) we conclude that

$$\max_{\overline{\Omega}} u \leq \max_{\partial\Omega} u.$$

Since $\partial\Omega \subseteq \overline{\Omega}$ the converse inequality $\max_{\partial\Omega} u \leq \max_{\overline{\Omega}} u$ always holds, which proves (i).

To show (ii) let $x_0 \in \Omega$ be such that

$$M = u(x_0) = \max_{\overline{\Omega}} u$$

and let $A := \{x \in \Omega : u(x) = M\}$. Then, $A \subseteq \Omega$ is closed since it is the preimage of the closed set $\{M\}$ under the continuous mapping u. On the other hand, since Ω is open, for any $x \in A$ there exists $r > 0$ such that $\overline{B_r(x)} \subseteq \Omega$. By the mean value property (see Theorem 3.8 for harmonic functions and Remark 3.9 for subharmonic functions) we conclude that

$$M = u(x) \leq \fint_{B_r(x)} u(y)dy \leq M, \tag{3.8}$$

where we used that $u(y) \leq M$ in the last inequality. Thus $u \equiv M$ on $B_r(x)$ and proves that A is also open. Consequently, $A = \Omega$ since it is open and closed (and nonempty by the first sentence of this proof), which shows (ii). □

An analogous result to Theorem 3.11 holds true for max and subharmonic replaced by min and superharmonic. This immediately implies the following result for harmonic functions.

COROLLARY 3.12 (Maximum principle for harmonic functions). *Let Ω be a bounded domain of $\mathbb{R}^n$ and $u \in C^2(\Omega) \cap C(\overline{\Omega})$ be harmonic on Ω. Then*

$$\min_{\partial\Omega} u \leq u(x) \leq \max_{\partial\Omega} u, \qquad x \in \Omega.$$

Moreover, since Ω is connected, either the strict inequalities hold or u is constant.

PROOF. Since u and $-u$ are both subharmonic (and superharmonic) the statements follow immediately from Theorem 3.11. □

REMARK 3.13 (Liouville's Theorem). The mean value formulas can be used to estimate the partial derivatives of a harmonic function. On can show that this implies that every harmonic and bounded function $u\colon \mathbb{R}^n \to \mathbb{R}$ satisfies $Du \equiv 0$, hence such u are *constant*. (For more details see Evans [**4**, Thm. 8 on p. 30].)

REMARK 3.14 (Harnack's inequality). Another way to see how a Laplace operator averages a function is Harnack's inequality: Suppose V is a domain (open and connected) with $V \subset\subset \Omega$, i.e., $\overline{V} \subseteq \Omega$ with $\overline{V}$ compact. Then there exists a constant C_V such that

$$\sup_V u \leq C_V \inf_V u$$

for all harmonic functions $u \geq 0$ on Ω.

This implies, in particular, that for all $x, y \in V$,

$$\frac{1}{C_V} u(y) \leq u(x) \leq C_V u(y),$$

which means that the values of all nonnegative harmonic functions on Ω are comparable (as long as V is sufficiently far from $\partial\Omega$). (For more details see Evans [**4**, Thm. 11 on p. 32].)

An important implication of the maximum principle is the uniquness of classical solutions for the *Dirichlet problem* of the Poisson equation.

COROLLARY 3.15 (Uniqueness of solutions to the Poisson equation). *Let Ω be a bounded domain in $\mathbb{R}^n$, $g \in C(\partial\Omega)$, and $f \in C(\Omega)$. Then there exists at most one function $u \in C^2(\Omega) \cap C(\overline{\Omega})$ that solves the boundary value problem*

$$\begin{aligned} -\Delta u &= f, \qquad \text{in } \Omega, \\ u &= g, \qquad \text{on } \partial\Omega. \end{aligned}$$

PROOF. If u and v are solutions of the boundary value problem, then $u - v$ is harmonic with $u - v|_{\partial\Omega} = 0$. Hence by the maximum principle $u = v$ on all of Ω. □

REMARK 3.16. Note that the above results guarantees uniqueness only in the realm of *classical* solutions. There may still exist different *weak* solutions.

3.4.3. Regularity. Elliptic and parabolic partial differential equations have some nice smoothing effects, while hyperbolic partial differential equations generally destroy regularity. We will use convolution by a mollifier to prove that (classical) solutions of the Laplace equation are, in fact, smooth.

REMARK 3.17 (Smoothing by convolution). A standard *mollifier* η is a smooth function[9] on $\mathbb{R}^n$ with support contained in $\overline{B_1(0)}$ and such that $\int_{\mathbb{R}^n} \eta = 1$. For any $\varepsilon > 0$ define

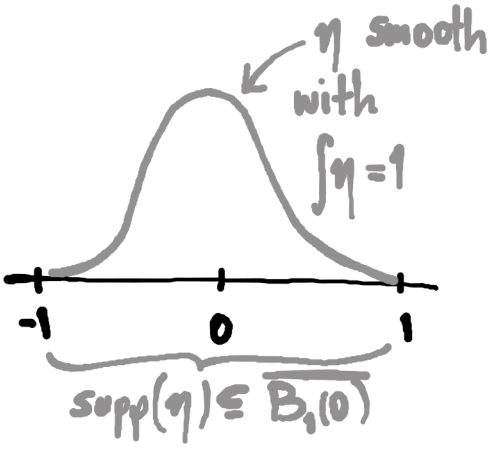

FIGURE 3. A mollifier η in $\mathbb{R}$.

[9]One can show that

$$\eta(x) = \begin{cases} Ce^{-\frac{1}{1-|x|^2}} & \text{if } |x| < 1, \\ 0 & \text{if } |x| \geq 1, \end{cases}$$

with $C > 0$ such that $\int \eta = 1$ is such a function, hence it exists.

$\eta_\varepsilon(x) := \frac{1}{\varepsilon^n}\eta(\frac{x}{\varepsilon})$. In particular, $\int_{\mathbb{R}^n}\eta_\varepsilon = 1$ and $\operatorname{supp}(\eta_\varepsilon) \subseteq \overline{B_\varepsilon(0)}$.

For any open set $\Omega \subseteq \mathbb{R}^n$ and $f \in C(\Omega)$ we set

$$\Omega_\varepsilon = \{x \in \Omega : \operatorname{dist}(x, \partial\Omega) > \varepsilon\},$$

and define the function f^ε on Ω_ε by convolution, i.e., for $x \in \Omega_\varepsilon$

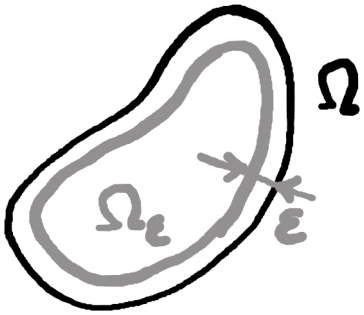

FIGURE 4. The domain Ω_ε.

$$f^\varepsilon(x) := (\eta_\varepsilon * f)(x) = \int_\Omega \eta_\varepsilon(x-y)f(y)\,dy = \int_{B_\varepsilon(0)} \eta_\varepsilon(y)f(x-y)\,dy.$$

One can show[10] that $f^\varepsilon \in C^\infty(\Omega)$ and $f^\varepsilon \to f$ uniformly on compact subsets of Ω as $\varepsilon \to 0$.

THEOREM 3.18. *Suppose $u \in C(\Omega)$ satisfies the mean value property on each set $B_r(x) \subseteq \Omega$. Then $u \in C^\infty(\Omega)$ and u is harmonic.*

PROOF. For $u^\varepsilon = \eta_\varepsilon * u$ defined by convolution with a mollifier as in Remark 3.17 (in particular, we choose the radially symmetric mollifier $\eta(x) = \eta(|x|)$ in the footnote) we show that $u^\varepsilon = u$ on Ω_ε. Let $x \in \Omega_\varepsilon$, then

$$\begin{aligned} u^\varepsilon(x) &= \int_\Omega \eta_\varepsilon(x-y)u(y)\,dy = \frac{1}{\varepsilon^n}\int_{B_\varepsilon(x)} \eta\left(\frac{|x-y|}{\varepsilon}\right)u(y)\,dy \\ &= \frac{1}{\varepsilon^n}\int_0^\varepsilon \eta\left(\frac{r}{\varepsilon}\right)\int_{\partial B_r(x)} u(y)\,dS(y)dr. \end{aligned}$$

Thus by the mean value property of u,

$$u^\varepsilon(x) = \frac{1}{\varepsilon^n}u(x)\int_0^\varepsilon \eta\left(\frac{r}{\varepsilon}\right)\omega_n r^{n-1}\,dr = u(x)\int_{B_\varepsilon(0)} \eta_\varepsilon(y)\,dy = u(x).$$

Since u^ε is smooth by Remark 3.17, this proves that u is smooth as well. Finally, by Theorem 3.10 u is harmonic. □

REMARK 3.19 (Analyticity). If u is harmonic on Ω, one can show that the multivariable Taylor expansion[11] converges (by estimating $\|D^\alpha u\|_{L^\infty}$ locally). Thus u is locally given by a convergent power series, hence even *analytic*. (For further details see Evans [**4**, Thm. 10 on p. 31].)

[10]This was problem 5 on Assignment 2.

[11]See problem 4 of Assignment 1.

3.5. Green's function and representation formula

In what follows we again assume that $\Omega \subseteq \mathbb{R}^n$ is open and bounded with C^1 boundary $\partial\Omega$. By using the fundamental solution Φ in (3.3) we now derive a representation formula for solutions of the *Dirichlet problem*

$$\begin{aligned} -\Delta u &= f, \qquad \text{in } \Omega, \\ u &= g, \qquad \text{on } \partial\Omega. \end{aligned} \tag{3.9}$$

Before analyzing (3.9) we derive an integral representation formula for an *arbitrary* function $u \in C^2(\overline{\Omega})$ in terms of Δu and its behavior at the boundary.

PROPOSITION 3.20. *Let $\Omega \subseteq \mathbb{R}^n$ be open and bounded with C^1 boundary $\partial\Omega$, and let Φ be the fundamental solution* (3.3) *of the Laplace equation. For any $u \in C^2(\overline{\Omega})$ we have for $x \in \Omega$*

$$u(x) = \int_{\partial\Omega} \left[\Phi(y-x) \frac{\partial u}{\partial \nu}(y) - u(y) \frac{\partial \Phi}{\partial \nu}(y-x) \right] dS(y) - \int_\Omega \Phi(y-x) \Delta u(y)\, dy, \tag{3.10}$$

where $\frac{\partial \Phi}{\partial \nu} = D\Phi \cdot \nu$ denotes the normal derivative with respect to $\partial\Omega$.

PROOF. Let $u \in C^2(\overline{\Omega})$ and fix $x \in \Omega, \varepsilon > 0$ such that $B_\varepsilon(x) \subseteq \Omega$. We apply Green's formula[12] to $u(x)$ and $\Phi(y-x)$ on $V_\varepsilon = \Omega \setminus B_\varepsilon(x)$ to obtain

$$\begin{aligned} &\int_{V_\varepsilon} u(y) \Delta\Phi(y-x) - \Phi(y-x)\Delta u(y)\, dy \\ &\qquad = \int_{\partial V_\varepsilon} u(y) \frac{\partial \Phi}{\partial \nu_\varepsilon}(y-x) - \Phi(y-x) \frac{\partial u}{\partial \nu_\varepsilon}(y)\, dS(y), \end{aligned} \tag{3.11}$$

where ν_ε denotes the outward unit normal vector field of the boundary ∂V_ε (we also use $\nu = \nu_\varepsilon$ for the normal vector related to the $\partial\Omega$ part of ∂V_ε that remains unchanged).

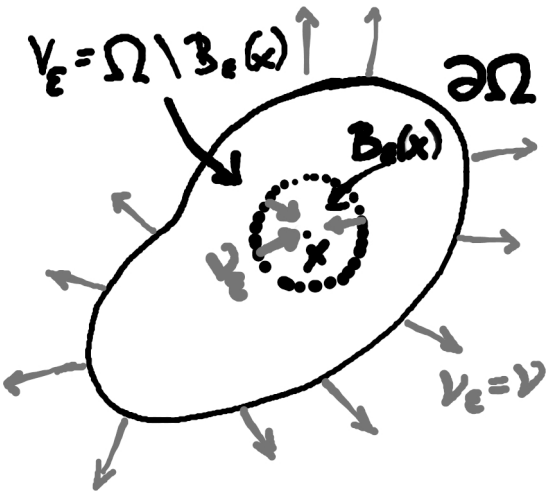

FIGURE 5. The ball $B_\varepsilon(x)$ and its complement V_ε in Ω, together with the outward pointing unit normal vectors ν_ε to $\partial V_\varepsilon = \partial\Omega \cup \partial B_\varepsilon(x)$.

[12]See problem 2b(i) on Assigment 1.

Since $\Delta_y \Phi(x-y) = 0$ as long as $x \neq y$ and $\partial V_\varepsilon = \partial\Omega \cup \partial B_\varepsilon(x)$ we can break up the boundary integral in (3.11) into two regions. More precisely, we consider the two regions (and in total four summands) in

$$\int_{\partial V_\varepsilon} = \int_{\partial B_\varepsilon(x)} + \int_{\partial\Omega}$$

separately.

Firstly, to determine the contribution of $\partial B_\varepsilon(x)$ observe that

$$\left| \int_{\partial B_\varepsilon(x)} \Phi(y-x) \frac{\partial u}{\partial \nu_\varepsilon}(y)\, dS(y) \right| \leq C\varepsilon^{n-1} \max_{\partial B_\varepsilon(x)} |\Phi| \to 0$$

as $\varepsilon \to 0$. Secondly, by Remark 3.5 we know that

$$\frac{\partial \Phi(y-x)}{\partial \nu_\varepsilon} = \nabla\Phi(y-x) \cdot \frac{x-y}{\varepsilon} = -\frac{1}{\omega_n} \frac{y-x}{|y-x|^n} \cdot \frac{x-y}{\varepsilon} = \frac{1}{\omega_n \varepsilon^{n-1}} = \frac{1}{|\partial B_\varepsilon(x)|},$$

thus the second integral involving $\partial B_\varepsilon(x)$ satisfies

$$\int_{\partial B_\varepsilon(x)} u(y) \frac{\partial \Phi}{\partial \nu_\varepsilon}(y-x)\, dS(y) = \fint_{\partial B_\varepsilon(x)} u(y)\, dS(y) \to u(x)$$

as $\varepsilon \to 0$ by (3.5) since u is, in particular, continuous.

Finally, by (3.11) and the fact that $\Delta\Phi(y-x) = 0$ since $y \neq x$,

$$\begin{aligned} -\int_{V_\varepsilon} \Phi(y-x)\Delta u(y)\, dy = \int_{\partial\Omega} u(y) \frac{\partial \Phi}{\partial \nu}(y-x)\, dS(y) + \underbrace{\int_{\partial B_\varepsilon(x)} u(y) \frac{\partial \Phi}{\partial \nu_\varepsilon}(y-x)\, dS(y)}_{\to u(x) \text{ as } \varepsilon \to 0} \\ - \int_{\partial\Omega} \Phi(y-x) \frac{\partial u}{\partial \nu}(y)\, dS(y) - \underbrace{\int_{\partial B_\varepsilon(x)} \Phi(y-x) \frac{\partial u}{\partial \nu_\varepsilon}(y)\, dS(y)}_{\to 0 \text{ as } \varepsilon \to 0}. \end{aligned}$$

Note that the term on the left hand side has the asymptotic behavior

$$\int_{V_\varepsilon} \Phi(y-x)\Delta u(y)\, dy \to \int_\Omega \Phi(y-x)\Delta u(y)\, dy$$

since $\Phi \in L^1_{\text{loc}}$ and Ω is bounded. More precisely, the contribution of the complement vanishes, since $|\Delta u|$ is bounded on Ω and thus (for $n \geq 3$, and similarly for $n = 2$)

$$\begin{aligned} \left| \int_{B_\varepsilon(x)} \Phi(y-x)\Delta u(y)\, dy \right| &\leq \sup_{B_\varepsilon(x)} |\Delta u| \int_0^\varepsilon \int_{\partial B_r(x)} |\Phi(r)|\, dS(y) dr \\ &\leq C \int_0^\varepsilon r\, dr \leq C\varepsilon^2 \to 0 \text{ as } \varepsilon \to 0. \end{aligned}$$

Hence

$$u(x) = \int_{\partial\Omega} \left[\Phi(y-x) \frac{\partial u}{\partial \nu}(y) - u(y) \frac{\partial \Phi}{\partial \nu}(y-x) \right] dS(y) - \int_\Omega \Phi(y-x)\Delta u(y)\, dy,$$

which is the desired representation formula (3.10). □

REMARK 3.21. The representation formula (3.10) can be used to obtain another proof of the smoothness for classical harmonic functions (see Assignment 4).

With the representation formula (3.10) can we express u using Δu on Ω and u, $\frac{\partial u}{\partial \nu}$ on $\partial\Omega$. While the first two quantities are specified by the Dirichlet problem (3.9) on Ω, the normal derivative $\frac{\partial u}{\partial \nu}$ is not. In order to eliminate this term we introduce the *corrector function* ϕ^x which is the solution (if it exists!) of the boundary value problem

$$\begin{aligned} \Delta \phi^x &= 0, && \text{in } \Omega, \\ \phi^x(y) &= \Phi(y-x), && \text{on } \partial\Omega. \end{aligned} \tag{3.12}$$

Using Green's formula and the assumption $\Delta\phi^x = 0$ we obtain

$$\begin{aligned} -\int_\Omega \phi^x(y)\Delta u(y)\,dy &= \int_{\partial\Omega} u(y)\frac{\partial \phi^x}{\partial \nu}(y) - \phi^x(y)\frac{\partial u}{\partial \nu}(y)\,dS(y) \\ &= \int_{\partial\Omega} u(y)\frac{\partial \phi^x}{\partial \nu}(y) - \Phi(y-x)\frac{\partial u}{\partial \nu}(y)\,dS(y). \end{aligned} \tag{3.13}$$

We can thus use ϕ^x to eliminate the term $\frac{\partial u}{\partial \nu}$ in (3.10) in Proposition **??**. We make this more precise by defining Green's function.

Definition 3.22. Let $\Omega \subseteq \mathbb{R}^n$ be open and bounded with C^1 boundary. Then

$$G(x,y) := \Phi(y-x) - \phi^x(y),$$

for $x, y \in \Omega, x \neq y$, is called *Green's function* (if it exists) on Ω.

With this terminology, (3.10) and (3.13) imply that for $x \in \Omega$

$$u(x) = -\int_{\partial\Omega} u(y)\frac{\partial G}{\partial \nu}(x,y)\,dS(y) - \int_\Omega G(x,y)\Delta u(y)\,dy. \tag{3.14}$$

Note that none of the above computations assumed any particular properties of u other than $u \in C^2(\overline{\Omega})$. If u is a classical solution of the Dirichlet problem (and if the Green's function for Ω exists), however, we can derive an integral representation.

Theorem 3.23 (Representation formula using Green's function). *Let $\Omega \subseteq \mathbb{R}^n$ be open and bounded with C^1 boundary $\partial\Omega$. Suppose $u \in C^2(\overline{\Omega})$ is a solution of the Dirichlet problem* (3.9)*, i.e.,*

$$\begin{aligned} -\Delta u &= f, && \text{in } \Omega, \\ u &= g, && \text{on } \partial\Omega, \end{aligned}$$

with $f \in C(\Omega)$ and $g \in C(\partial\Omega)$. Assume furthermore that G is Green's function for Ω. Then

$$u(x) = -\int_{\partial\Omega} g(y)\frac{\partial G}{\partial \nu}(x,y)\,dS(y) + \int_\Omega f(y)G(x,y)\,dy, \qquad x \in \Omega. \tag{3.15}$$

Proof. This follows immediately from Proposition 3.20 and the definition of Green's function. □

Remark 3.24. Note that, formally, G solves the boundary value problem

$$\begin{aligned} -\Delta G &= \delta_x, && \text{in } \Omega, \\ G &= 0, && \text{on } \partial\Omega. \end{aligned}$$

Furthermore, one can prove that G is symmetric[13], i.e.,

$$G(y,x) = G(x,y), \qquad \text{for all } x, y \in \Omega, x \neq y.$$

[13] See Assignment 5.

Although the representation formula (3.15) looks very nice and simply, there is one major issue: The explicit contruction of Green's function for a given Ω can be difficult (or even impossible), since it requires to solution of the auxilliary Dirichet problem (3.12) for the corrector function. In fact, G can only be computed for geometrically simple domains Ω. Without proofs but with some physical interpretation we will have a look at some simple domains and their Green's function.

3.5.1. Green's function for a half space. In this special case we consider the unbounded domain

$$\mathbb{R}^n_+ := \{x = (x_1, \dots, x_n) \in \mathbb{R}^n : x_n > 0\},$$

though one cannot directly apply Theorem 3.23 then. We search for a Green function that satisfies

$$\begin{aligned} -\Delta G &= \delta_x, \\ G|_{\partial\mathbb{R}^n_+} &= 0. \end{aligned}$$

In other words, we can say that we search for an electric potential G on all of $\mathbb{R}^n$ where we assume that there is a point charge $q = 1$ in x and the potential vanishes on $\partial\mathbb{R}^n_+$. The idea is

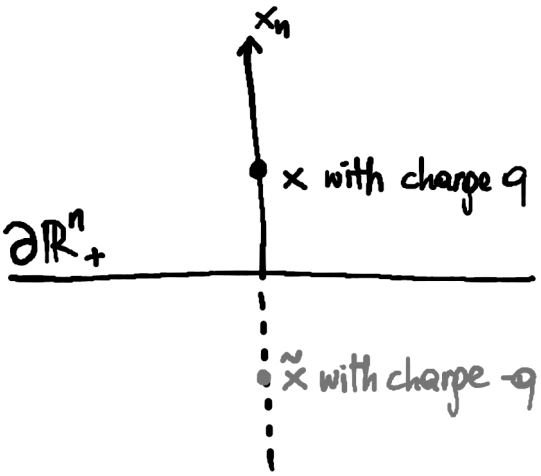

FIGURE 6. Point $x \in \mathbb{R}^n_+$ with charge q and its mirror point $\tilde{x} = -x$ with opposite charge $-q$.

assume a mirroring point charge $-q$ in $\tilde{x} = (x_1, \dots, x_{n-1}, -x_n)$ that neutralizes q in x along $\partial\mathbb{R}^n_+$. Since the potential of the point charge at x is $\Phi(y - x)$ the expected full potential is

$$G(x, y) := \Phi(y - x) - \Phi(y - \tilde{x}), \tag{3.16}$$

where the minus sign appears due to the negative charge $-q$. In other words, we expect that for $x, y \in \mathbb{R}^n_+$ the corrector is

$$\phi^x(y) := \Phi(y_1 - x_1, \dots, y_{n-1} - x_{n-1}, y_n + x_n).$$

One can verify that, indeed, this is the case and that G for $\mathbb{R}^n_+$ is of the form (3.16)[14]. Since $\nu = (0, \dots, 0, -1)$, computing $\frac{\partial G}{\partial \nu}$ using Remark 3.5 yields

$$\begin{aligned}\frac{\partial G}{\partial \nu}(x,y) &= -\frac{\partial G}{\partial y_n}(x,y) = -\frac{\partial \Phi}{\partial y_n}(y-x) + \frac{\partial \Phi}{\partial y_n}(y-\tilde{x}) \\ &= \frac{1}{\omega_n}\left[\frac{y_n - x_n}{|y-x|^n} - \frac{y_n + x_n}{|y-\tilde{x}|^n}\right] = -\frac{2x_n}{\omega_n}\frac{1}{|x-y|^n}\end{aligned}$$

for $y \in \partial\mathbb{R}^n_+$.

If we assume that u is a solution to the boundary value problem

$$\begin{aligned}\Delta u &= 0, \qquad \text{in } \mathbb{R}^n_+, \\ u &= g, \qquad \text{on } \partial\mathbb{R}^n_+,\end{aligned}$$

then in analogy to Theorem 3.23 we expect that u is of the form

$$u(x) := \frac{2x_n}{\omega_n}\int_{\partial\mathbb{R}^n_+} \frac{g(y)}{|x-y|^n}\,dy, \qquad x \in \mathbb{R}^n_+. \tag{3.17}$$

We call

$$K(x,y) := \frac{2x_n}{\omega_n}\frac{1}{|x-y|^n}, \qquad x \in \mathbb{R}^n_+,\ y \in \partial\mathbb{R}^n_+,$$

the *Poisson kernel* for $\mathbb{R}^n_+$ and (3.17) the *Poisson formula*. Note that we can *not* apply Theorem 3.23 directly since we deal with an unbounded domain. One can, however, prove it separately.

THEOREM 3.25 (Poisson's formula for a half space). *Let $g \in C(\mathbb{R}^{n-1}) \cap L^\infty(\mathbb{R}^{n-1})$ and define u as in* (3.17)*. Then*

(i) $u \in C^\infty(\mathbb{R}^n_+) \cap L^\infty(\mathbb{R}^n_+)$,
(ii) $\Delta u = 0$ *in* $\mathbb{R}^n_+$,
(iii) $\lim_{\substack{x \to x^0 \\ x \in \mathbb{R}^n_+}} u(x) = g(x^0)$ *for all* $x^0 \in \partial\mathbb{R}^n_+$.

SKETCH OF PROOF. The regularity of u in (i) follows from the regularity and boundedness of g and K. One makes use of the fact that $x \mapsto G(x,y)$ is harmonic on the set $\{x \neq y\}$ to show that $x \mapsto K(x,y)$ is, which ultimately implies (ii). The limit in (iii) requires that g is continuous and that $\int_{\partial\mathbb{R}^n_+} K(x,y)\,dy = 1$: Continuity of g implies that for every $\varepsilon > 0$ exists a $\delta > 0$ such that $|g(y) - g(x^0)| < \varepsilon$ for all $|y - x^0| < \delta$. Then one estimates $|u(x) - g(x^0)|$ thereby considering the regions $\partial\mathbb{R}^n_+ \cap B_\delta(x^0)$ and $\partial\mathbb{R}^n_+ \setminus B_\delta(x^0)$ separately. (Further details may be found in Evans [**4**, Thm. 14 on p. 37].) □

3.5.2. Green's function for a ball. We now consider the Dirichlet problem (3.9) for the unit ball $\Omega = B_1(0)$. As before we can employ some kind of reflection idea to obtain a candidate for a Green's function on $B_1(0)$. To obtain a mirror image of a point $x \neq 0$ one uses inversion on the sphere, i.e., inversion through $\partial B_1(0)$. The dual point is $\tilde{x} = \lambda\frac{x}{|x|}$ for some $\lambda \in \mathbb{R}$. By the Cathetus Theorem we require that $|x||\tilde{x}| = 1$, hence $\lambda = |\tilde{x}| = \frac{1}{|x|}$. Hence the dual point satisfies $\tilde{x} = \frac{x}{|x|^2}$.

[14]This is essentially problem 6 of Assignment 4.

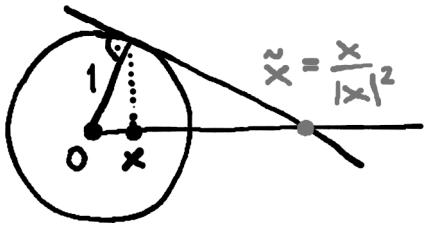

FIGURE 7. The point $x \in B_1(0)$ is reflected on the boundary $\partial B_1(0)$ using the Cathetus Theorem, that is, $1 = |x||\tilde{x}|$.

In order to obtain a Green's function we again want to reflect at the point charge (singularity). For $n \geq 3$, the function $y \mapsto \Phi(y - \tilde{x})$ is harmonic on $\{y \neq \tilde{x}\}$. Thus the map

$$\phi^x \colon y \mapsto |x|^{-(n-2)}\Phi(y - \tilde{x}) = \Phi(|x|(y - \tilde{x}))$$

is harmonic on Ω.

For $y \in \partial B_1(0)$ and $x \neq 0$ we thus have that

$$|x|^2|y - \tilde{x}|^2 = |x|^2(\underbrace{|y|^2}_{=1} - \frac{2y \cdot x}{|x|^2} + \frac{1}{|x|^2}) = |x|^2 - 2y \cdot x + 1 = |x - y|^2.$$

Therefore

$$(|x||y - \tilde{x}|)^{-(n-2)} = |x - y|^{-(n-2)},$$

so that for all $y \in \partial B_1(0)$

$$\phi^x(y) := \Phi(y - x),$$

that is, ϕ^x is the desired corrector function for $x \neq 0$.

DEFINITION 3.26. The *Green's function for the unit ball* $B_1(0)$ is

$$G(x, y) := \Phi(y - x) - \Phi(|x|(y - \tilde{x})),$$

for $x, y \in B_1(0)$, $x \neq y$.

REMARK 3.27. Strictly speaking, we have not computed the corrector function for $x = 0$. In this case the corrector function reduces to

$$\phi^0(y) := \Phi(1),$$

since $\Phi(|x|(y - \tilde{x})) = \Phi(||x|(y - \tilde{x})|)$ only depends on the norm of the point and is smooth, and

$$\lim_{x \to 0} ||x|(y - \tilde{x})| = \lim_{x \to 0} \left| |x|y - |x|\frac{x}{|x|^2} \right| = 1.$$

For the integral representation formula it is sufficient to know Green's function almost everywhere, so we don't actually need the cases $x = 0$ and $x = y$ in the above definition of G.

One can show that ϕ^x is also the desired corrector if $n = 2$.

After computing

$$\frac{\partial G}{\partial \nu}(x, y) = -\frac{1}{\omega_n}\frac{1 - |x|^2}{|x - y|^n}, \qquad x \in B_1(0),\ y \in \partial B_1(0),$$

Theorem 3.23 implies Poisson's formula for the unit ball (see problem 6 on Assignment 5) for the boundary value problem

$$\begin{aligned} \Delta u &= f, \qquad \text{in } B_1(0), \\ u|_{\partial B_1(0)} &= g. \end{aligned}$$

A coordinate transformation implies the analogous result for a ball with radius r. Note that this establishes a representation formula only under the assumption that a classical solution of the Dirichlet problem actually *exists*. The following result shows that this in fact gives a solution.

THEOREM 3.28 (Poisson's formula for a ball). *Suppose $g \in C(\partial B_r(0))$ and u is given by*

$$u(x) = \frac{r^2 - |x|^2}{r\omega_n} \int_{\partial B_r(0)} \frac{g(y)}{|x-y|^n}\, dS(y).$$

Then

(i) $u \in C^\infty(B_r(0))$,
(ii) $\Delta u = 0$ *in* $B_r(0)$,
(iii) $\lim\limits_{\substack{x \to x^0 \\ x \in B_r(0)}} u(x) = g(x^0)$ *for all* $x^0 \in \partial B_r(0)$.

SKETCH OF PROOF. The kernel $K(x,y) := \frac{r^2-|x|^2}{r\omega_n} \frac{1}{|x-y|^n}$ is analytic on $B_r(0)$, which immediately implies (i). Since $x \mapsto G(x,y)$ is harmonic (general property of Green's function), so is $x \mapsto \frac{\partial G}{\partial \nu}$, and hence $x \mapsto K(x,y)$ is harmonic. Moreover, one can show that $\int_{\partial B_r(0)} K(x,y)\, dS(y) = 1$. Thus (ii) holds, since

$$\Delta u = \int_{\partial B_r(0)} g(y) \Delta_x K(x,y)\, dS(y) = 0.$$

Finally, (iii) follows from an $\varepsilon - \delta$ argument by estimating the integrals involved. □

3.6. Energy methods

So far we use explicit representation formulas and the mean value property to show existence, uniqueness and certain properties of solutions. In what follows we apply techniques related to the L^2-norm of certain expressions and the calculus of variations to derive results about solutions of the Laplace equation. Such techniques are heavily used in the study of weak solutions, which we at least want to provide a glimpse of in this course. See [**5**] for more advanced techniques for (linear and nonlinear) elliptic equations of second order.

DEFINITION 3.29. Let $\Omega \subseteq \mathbb{R}^n$ be open and $u \in C(\Omega)$. For $1 \le p < \infty$ the *L^p-norm* of u is defined by

$$\|u\|_{L^p(\Omega)} := \left(\int_\Omega |u|^p \right)^{\frac{1}{p}}.$$

We call a function $u \in C(\Omega)$ *integrable* if $\|u\|_{L^1(\Omega)} < \infty$, and *square-integrable* if $\|u\|_{L^2(\Omega)} < \infty$.

3.6.1. Uniqueness. In Corollary 3.15 we proved uniqueness of solutions for the Dirichlet problem (3.9) using the maximum principle. We now present an alternative proof using *energy methods*.

THEOREM 3.30. *Let $\Omega \subseteq \mathbb{R}^n$ be open and bounded with C^1 boundary $\partial\Omega$. Suppose $f \in C(\Omega)$ and $g \in C(\partial\Omega)$. Then there exists at most one solution $u \in C^2(\overline{\Omega})$ of the boundary value problem*

$$\begin{aligned} -\Delta u &= f, \qquad \text{in } \Omega, \\ u &= g, \qquad \text{on } \partial\Omega. \end{aligned} \tag{3.18}$$

PROOF. Suppose v is another solution. Consider the difference $w := u - v$. Then $\Delta w = 0$ in Ω and $w|_{\partial\Omega} = 0$. Hence multiplication by w and integration implies

$$0 = -\int_\Omega w\Delta w\, dx = \int_\Omega |\nabla w|^2\, dx,$$

where the second equality follows from integration by parts. Since ∇w is continuous, this implies[15] that $\nabla w \equiv 0$ in Ω. Therefore, w must be constant and since $w|_{\partial\Omega} = 0$ it follows that $w \equiv 0$, i.e., $u \equiv v$. □

3.6.2. The Dirichlet principle. The Poisson equation describes, amongst others, steady state deflections of a thin membrane or steady state distributions of a chemical substrate. It is therefore natural to expect that the solution of the Dirichlet problem (3.18) corresponds to a minimum of some *energy functional*. More precisely, we will see that the solution of (3.18) with right hand side $f \in C(\Omega)$ and boundary value $g \in C(\partial\Omega)$ minimizes the energy functional

$$I(w) := \int_\Omega \frac{1}{2}|\nabla w|^2 - wf\, dx,$$

where w is in the admissible set

$$\mathcal{A} = \{w \in C^2(\overline{\Omega}) \,|\, w = g \text{ on } \partial\Omega\}.$$

THEOREM 3.31 (Dirichlet principle). *Let $\Omega \subseteq \mathbb{R}^n$ be open and bounded with C^1 boundary $\partial\Omega$. Suppose $u \in C^2(\overline{\Omega})$ is a solution of the boundary value problem* (3.18) *with $f \in C(\Omega)$ and $g \in C(\partial\Omega)$. Then*

$$I(u) = \min_{w\in\mathcal{A}} I(w). \tag{3.19}$$

Conversely, if $u \in \mathcal{A}$ solves (3.19)*, then u is a solution of the Dirichlet problem* (3.18) *on Ω.*

PROOF. Assume that $u \in C^2(\overline{\Omega})$ is a solution of the Dirichlet problem (3.18). In particular, $u \in \mathcal{A}$. We consider $w \in \mathcal{A}$ and multiply the Poissen equation $-\Delta u = f$ by $u - w$. Integration over Ω and integration by parts (no boundary terms occur since both $u, w \in \mathcal{A}$) thus yields

$$0 = \int_\Omega (-\Delta u - f)(u - w)\, dx = \int_\Omega |\nabla u|^2 - \nabla u \cdot \nabla w - fu + fw\, dx. \tag{3.20}$$

The Cauchy–Schwarz inequality it follows that

$$|\nabla u \cdot \nabla w| \le |\nabla u||\nabla w| \le \frac{1}{2}|\nabla u|^2 + \frac{1}{2}|\nabla w|^2,$$

[15] See problem 1 on Assignment 1.

where we used the inequality $a^2 + b^2 - 2ab = (a-b)^2 \geq 0$, for $a, b \in \mathbb{R}$, in the second step. Using this estimate in the equality (3.20) implies

$$0 \geq \int_\Omega \frac{1}{2}|\nabla u|^2 - uf\,dx - \int_\Omega \frac{1}{2}|\nabla w|^2 - wf\,dx,$$

i.e., $I(w) \geq I(u)$. Since $w \in \mathcal{A}$ was arbitrary, this implies (3.19).

Conversely, let $u \in \mathcal{A}$ satisfy (3.19). For arbitrary test functions $v \in C_c^\infty(\Omega)$ with compact support consider the function

$$i(s) := I(u + sv), \qquad s \in \mathbb{R}.$$

Then, since $u + sv \in \mathcal{A}$, $i\colon \mathbb{R} \to \mathbb{R}$ is well-defined with a minimum in $s = 0$. Moreover, i is continuously differentiable and

$$i'(s) = \int_\Omega (\nabla u + s\nabla v) \cdot \nabla v - fv\,dx,$$

since Ω is bounded and the integrand of I and its partial derivatives with respect to s are continuous for $s \in \mathbb{R}, x \in \overline{\Omega}$. Therefore, since u is a minimizer,

$$0 = i'(0) = \int_\Omega \nabla u \cdot \nabla v - vf\,dx = \int_\Omega (-\Delta u - f)v\,dx,$$

where we used integration by parts in the last step. Since $v \in C_c^\infty(\Omega)$ was arbitrary and $(-\Delta u - f) \in C(\Omega)$, it follows that $-\Delta u = f$ in Ω by the fundamental lemma of the calculus of variations (see [**1**, **3**]). The boundary conditions (3.18) are trivially satisfied, since $u \in \mathcal{A}$. □

If the data f and g or the boundary $\partial\Omega$ are less regular, it is not guaranteed that $\mathcal{A} \neq \emptyset$ or that I attains a minimum in $\mathcal{A}$. In such cases it is natural to enlarge the admissible set $\mathcal{A}$ by allowing less regular classes of functions in order to ensure the existence of a minimizer. Such a minimizer is then a natural candidate for a weak solution of Poisson's equation.

CHAPTER 4

The heat equation

4.1. Motivation

The heat equation (or diffusion equation) describes the time evolution of a physical quantity (e.g., heat, chemical concentration). Let $\Omega \subseteq \mathbb{R}^n$ be open and $V \subseteq \Omega$ be any subregion with C^1 boundary and unit normal vector $\nu\colon \partial V \to \mathbb{R}^n$. Let u be the density of the physical

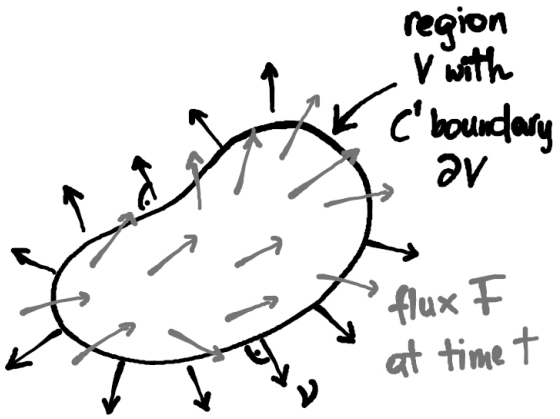

FIGURE 1. A physical quantity (say, temperature) in a given region changes in time due to a flux through the boundary (no equilibrium).

quantity. The rate of change of the physical quantity within V coincides with the negative flux through the boundary ∂V, that is,

$$\frac{d}{dt}\int_V u(t,x)\,dx = -\int_{\partial V} F(t,x)\cdot\nu(x)\,dS(x),$$

where $F\colon [0,\infty)\times\Omega \to \mathbb{R}^n$ is some flux function.[1] By the Divergence Theorem,

$$\int_{\partial V} F(t,x)\cdot\nu(x)\,dS(x) = \int_V \operatorname{div} F(t,x)\,dx,$$

so that

$$\int_V u_t = -\int_V \operatorname{div} F.$$

Note that $\operatorname{div} = \operatorname{div}_x$ denotes the divergence operator with respect to x only. In general, the flux density F is proportional to the (spatial) gradient of u (pointing from regions of higher to lower concentration), so that

$$F(t,x) = -a\nabla u(t,x),$$

[1]Compare this to the Laplace equation, where the total flux was 0 because the quantity is in equilibrium.

for some constant $a > 0$. As in the derivation of the Laplace equation it follows that

$$\int_V u_t = \int_V a\Delta u,$$

and since $V \subseteq \Omega$ was arbitrary we obtain the partial differential equation

$$u_t - a\Delta u = 0$$

on Ω. Rescaling of the time variable (so that $a = 1$) yields the heat equation.

If the physical quantity is generated by a source Q, then we obtain the inhomogeneous partial differential equation

$$u_t - a\Delta u = Q$$

(cf. the derivation of the Poisson equation in Section 3.1).

4.2. Heat equation

DEFINITION 4.1. Let $\Omega \subseteq \mathbb{R}^n$ be open. The *heat equation* reads

$$u_t - \Delta u = 0, \qquad \text{on } (0,\infty) \times \Omega,$$

and the *inhomogeneous heat equation* reads

$$u_t - \Delta u = f, \qquad \text{on } (0,\infty) \times \Omega,$$

where $f\colon [0,\infty) \times \Omega \to \mathbb{R}$ is given and $u\colon [0,\infty) \times \overline{\Omega} \to \mathbb{R}$ is the unknown.

Here, t denotes time, x a point in space and $\Delta = \Delta_x$ the Laplacian with respect to the space variables.

In what follows we will see that many results about the Laplace equation carry over to the heat equation, albeit in a more complicated setting.

4.3. Fundamental solution

Recall that we derived the fundamental solution Φ for the Laplace equation, i.e.,

$$-\Delta\Phi = \delta_0,$$

by making use of the rotation invariance of the equation. To construct a fundamental solution for the heat equation, we proceed in a similar fashion. Note that the the heat equation also satisfies a scaling law: for one time derivative there are two spatial derivatives. Hence whenever $u(t,x)$ is a solution, we expect that $u(\lambda^2 t, \lambda x)$, for every $\lambda \in \mathbb{R}$, also solves the heat equation[2]. Together with the rotational invariance of the Laplace operator, this suggests to look for solutions of the form $u(t,x) = v(\frac{|x|^2}{t})$ for some function v. More generally, we make the ansatz

$$u(t,x) = \frac{1}{t^\alpha} v\left(\frac{x}{t^\beta}\right), \tag{4.1}$$

where the constants α, β and the function $v\colon \mathbb{R}^n \to \mathbb{R}$ have to be found. Inserting (4.1) into the heat equation implies that

$$\begin{aligned} 0 &= u_t(t,x) - \Delta u(t,x) \\ &= -\alpha t^{-(\alpha+1)} v\left(\frac{x}{t^\beta}\right) - \beta t^{-(\alpha+\beta+1)} x \cdot \nabla v\left(\frac{x}{t^\beta}\right) - t^{-(\alpha+2\beta)} \Delta v\left(\frac{x}{t^\beta}\right). \end{aligned}$$

[2]See Assignment 6.

We set $y = t^{-\beta}x$ with $\beta = \frac{1}{2}$ and multiply the equation by $t^{\alpha+1}$ to obtain

$$\alpha v(y) + \frac{1}{2}\nabla v(y) \cdot y + \Delta v(y) = 0. \tag{4.2}$$

We expect that v is radial, that is, we can assume that $v(y) = w(r)$ for $r = |y|$. Thus instead of (4.2) we can deal with the ordinary differential equation[3] for w

$$\alpha w(r) + \frac{r}{2}w'(r) + w''(r) + \frac{n-1}{r}w'(r) = 0. \tag{4.3}$$

Note that

$$(r^n w)' = nr^{n-1}w + r^n w', \qquad (r^{n-1}w')' = (n-1)r^{n-2}w' + r^{n-1}w'',$$

so that (4.3) with $\alpha = \frac{n}{2}$ and divided by r^{n-1} turns into

$$\frac{1}{2}(r^n w)' + (r^{n-1}w')' = 0,$$

which can be integrated to obtain

$$\frac{1}{2}r^n w + r^{n-1}w' = c,$$

for some $c \in \mathbb{R}$. If we require that $w(r), w'(r) \to 0$ as $r \to \infty$, then $c = 0$ and

$$w'(r) = -\frac{r}{2}w(r),$$

which yields that

$$w(r) = be^{-\frac{r^2}{4}}$$

for some constant $b \in \mathbb{R}$. Since we set $r = |y| = t^{-\beta}|x| = \frac{|x|}{\sqrt{t}}$ and $\alpha = \frac{n}{2}$ we obtain a solution of the form

$$u(t,x) = \frac{b}{t^{\frac{n}{2}}}e^{-\frac{|x|^2}{4t}}, \qquad t > 0.$$

For the particular choice of constant $b = (4\pi)^{-\frac{n}{2}}$, the function $u(t,.)$ is the density of the n-dimensional normal distribution $\mathcal{N}(0, 2tI)$ and we obtain the fundamental solution of the heat equation.

Definition 4.2. The function $\Phi\colon (\mathbb{R}\setminus\{0\}) \times \mathbb{R}^n \to \mathbb{R}$, defined by

$$\Phi(t,x) = \begin{cases} \frac{1}{(4\pi t)^{\frac{n}{2}}}e^{-\frac{|x|^2}{4t}}, & t > 0, x \in \mathbb{R}^n, \\ 0, & t \le 0, x \in \mathbb{R}^n, \end{cases} \tag{4.4}$$

is called the *fundamental solution of the heat equation* (or *heat kernel*).

The choice of the normalizing constant $(4\pi)^{-\frac{n}{2}}$ is due to the following result.

Lemma 4.3. *The fundamental solution* (4.4) *satisfies* $\Phi(t,.) > 0$ *and*

$$\int_{\mathbb{R}^n} \Phi(t,x)\,dx = 1$$

for all $t > 0$.

[3]Compare this to the Laplacian in polar coordinates, cf. problem 4(a) of Assignment 4 (for $n = 2$).

PROOF. The first statement is clear. The second one is a short calculation,

$$\begin{aligned}\int_{\mathbb{R}^n} \Phi(t,x)\,dx &= \frac{1}{(4\pi t)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{-\frac{|x|^2}{4t}}\,dx \\ &= \frac{1}{\pi^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{-|z|^2}\,dz = \frac{1}{\pi^{\frac{n}{2}}} \prod_{i=1}^{n} \underbrace{\int_{-\infty}^{\infty} e^{-z_i^2}\,dz_i}_{=\Gamma(\frac{1}{2})=\sqrt{\pi}} = 1. \end{aligned}$$ □

REMARK 4.4. One can prove that, for any compact interval $I \subseteq (0,\infty)$ and any $\alpha \in \mathbb{N}_0^{n+1}$, there exists an integrable function F_α with

$$|D^\alpha_{(t,x)}\Phi(t,x)| \le F_\alpha(x) \qquad \text{on } I \times \mathbb{R}^n.$$

(See Assignment 6 for the details.)

4.4. The initial value problem

We now employ the fundamental solution Φ in (4.4) to construct solutions of the initial value problem via convolution (in the spatial variable).

4.4.1. Homogeneous case. Consider the initial value problem

$$\begin{aligned} u_t - \Delta u &= 0, \qquad \text{in } (0,\infty)\times\mathbb{R}^n, \\ u(0,.) &= g, \qquad \text{on } \{t=0\}\times\mathbb{R}^n, \end{aligned} \tag{4.5}$$

where $g \in C(\mathbb{R}^n)$ is a bounded function, i.e., $\|g\|_{L^\infty} = \sup_{x\in\mathbb{R}^n} |g(x)| < \infty$.

DEFINITION 4.5. A function $u \in C(\overline{(0,\infty)\times\mathbb{R}^n}) \cap C^{1,2}((0,\infty)\times\mathbb{R}^n)$, with

$$\begin{aligned} C^{1,2}((0,\infty)\times\mathbb{R}^n) := \{v \in C^1((0,\infty)\times\mathbb{R}^n)) : D_x^2 v \text{ exists and} \\ D_x^2 v \in C((0,\infty)\times\mathbb{R}^n; \mathbb{R}^{n\times n})\} \end{aligned}$$

that satisfies (4.5) is called *classical solution* for (4.5).

Since Φ solves the heat equation away from the singularity $(0,0) \in (-\infty,\infty)\times\mathbb{R}^n$, so does $(t,x) \to \Phi(t,x-y)$ for fixed $y \in \mathbb{R}^n$. We also expect that the convolution with g should be a solution.

THEOREM 4.6. *Let $g \in C(\mathbb{R}^n)$ be a bounded function. Then the function u, defined by*

$$u(t,x) = \int_{\mathbb{R}^n} \Phi(t,x-y)g(y)\,dy$$

satisfies

(i) $u \in C^\infty((0,\infty)\times\mathbb{R}^n)$,
(ii) $u_t - \Delta u = 0$ *on* $(0,\infty)\times\mathbb{R}^n$, *and*
(iii) $\lim_{\substack{(t,x)\to(0,x^0)\\ x\in\mathbb{R}^n,\, t>0}} u(t,x) = g(x^0)$.

In particular, u is a classical solution of the initial value problem (4.5).

PROOF. (i) Since g is bounded and Φ is integrable by Lemma 4.3, u is well-defined. Moreover, the integrand $(t,x,y) \mapsto h(t,x,y) := \Phi(t,x-y)g(y)$ satisfies

- For every fixed y, the function $h(.,.,y)$ is smooth,
- For every fixed (t,x) the function $h(t,x,.)$ is integrable,

- For every compact set $I \times K \subseteq (0,\infty) \times \mathbb{R}^n$ and $\alpha \in \mathbb{N}_0^{n+1}$ there exists an integrable function G_α that bounds the derivatives, more precisely,
$$|D^\alpha_{(t,x)} h(t,x,y)| \le \|g\|_{L^\infty} \sup_{x \in K} F_\alpha(x-y) =: G_\alpha(y)$$
by Remark 4.4 for all $(t,x,y) \in I \times K \times \mathbb{R}^n$.

This implies (see, e.g., [**7**, Theorem 16.11]) that u is smooth on $(0,\infty) \times \mathbb{R}^n$ and derivatives can be computed under the integral. Hence (i) holds, and moreover,
$$u_t(t,x) - \Delta u(t,x) = \int_{\mathbb{R}^n} \underbrace{(\Phi_t - \Delta\Phi)(t,x-y)}_{=0} g(y)\,dy = 0,$$
which shows (ii).

(iii) It remains to be shown that u assumes the initial data. Suppose $x^0 \in \mathbb{R}^n$ and let $\varepsilon > 0$. Since g is continuous, there exists a $\delta > 0$ such that
$$|g(y) - g(x^0)| < \varepsilon \tag{4.6}$$
if $|y - x^0| < \delta$ for $y \in \mathbb{R}^n$. Consider now $x \in \mathbb{R}^n$ such that $|x - x^0| < \frac{\delta}{2}$. Hence by Lemma 4.3 and splitting up the integral,
$$\begin{aligned} |u(t,x) - g(x^0)| &= \left| \int_{\mathbb{R}^n} \Phi(t,x-y)(g(y) - g(x^0))\,dy \right| \\ &\le \underbrace{\int_{B_\delta(x^0)} \Phi(t,x-y)|g(y) - g(x^0)|\,dy}_{=:I} \\ &\quad + \underbrace{\int_{\mathbb{R}^n \setminus B_\delta(x^0)} \Phi(t,x-y)|g(y) - g(x^0)|\,dy}_{=:J}. \end{aligned}$$
By (4.6), we see that the first integral tends to zero as
$$I \le \varepsilon \underbrace{\int_{\mathbb{R}^n} \Phi(t,x-y)\,dy}_{=1} = \varepsilon.$$
For $|x - x^0| \le \frac{\delta}{2}$ and $|y - x^0| \ge \delta$ we have that

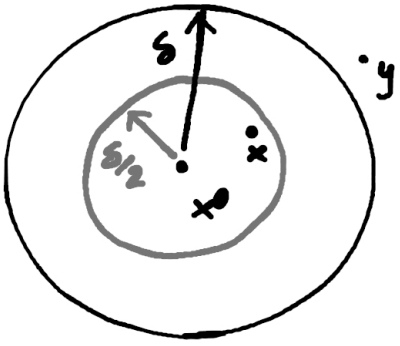

FIGURE 2. Estimating $|x - y|$ do remove x-dependence in the integral J.

$$|y-x^0| \le |y-x| + \frac{\delta}{2} \le |y-x| + \frac{1}{2}|y-x^0|,$$

hence

$$|y-x| \ge \frac{1}{2}|y-x^0|.$$

For the second integral, we then obtain

$$\begin{aligned} J &\le 2\|g\|_\infty \int_{\mathbb{R}^n \setminus B_\delta(x^0)} \Phi(x-y,t)\,dy \le \frac{C}{t^{\frac{n}{2}}} \int_{\mathbb{R}^n \setminus B_\delta(x^0)} e^{-\frac{|y-x|^2}{4t}}\,dy \\ &\le \frac{C}{t^{\frac{n}{2}}} \int_{\mathbb{R}^n \setminus B_\delta(x^0)} e^{-\frac{|y-x^0|^2}{16t}}\,dy, \end{aligned}$$

and by switching to polar coordinates, and setting $s = \frac{r}{\sqrt{t}}$,

$$J \le \frac{C}{t^{\frac{n}{2}}} \int_\delta^\infty e^{-\frac{r^2}{16t}} r^{n-1}\,dr = C \int_{\frac{\delta}{\sqrt{t}}}^\infty e^{-\frac{s^2}{4}} s^{n-1}\,ds \to 0$$

as $t \to 0+$ (or by dominated convergence). Hence $|u(t,x) - g(x^0)| < 2\varepsilon$ for all $x \in B_\delta(x^0)$ and $t > 0$ sufficiently small. □

Lemma 4.3 implies a bound on the solution u (see problem 1 on Assignment 7 for details).

COROLLARY 4.7. *Let $g \in C(\mathbb{R}^n)$ be a bounded function and u as in Theorem 4.6. Then*

$$\|u(t,.)\|_{L^\infty} \le \|g\|_{L^\infty}, \qquad t \ge 0. \quad \square$$

REMARK 4.8. In view of Theorem 4.6, the fundamental solution Φ formally satisfies the initial value problem

$$\begin{aligned} \Phi_t - \Delta\Phi &= 0, && \text{in } (0,\infty)\times\mathbb{R}^n, \\ \Phi(0,.) &= \delta_0, && \text{on } \{t=0\}\times\mathbb{R}^n, \end{aligned}$$

where δ_0 is the Dirac delta function on $\mathbb{R}^n$ centered at $x = 0$.

REMARK 4.9 (Infinite speed of progapation). If g is as in Theorem 4.6 and satisfies $g \ge 0$ and $g \not\equiv 0$, then for all times $t > 0$ and *all* $x \in \mathbb{R}^n$ we have that

$$u(t,x) = \frac{1}{(4\pi t)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{-\frac{|x-y|^2}{4t}} g(y)\,dy > 0.$$

In physical terms that means if the initial temperature is nonnegative and positive somewhere, then at any later time (no matter how small), the temperature is strictly positive. This *infinite speed of propagation of disturbances* is an important property of the heat equation. It is completely different for the wave equation which only supports a finite speed of propagation (if a solution is locally zero, then it will remain zero for some time).

4.4.2. Inhomogeneous case. We first consider the inhomogeneous initial value problem, i.e.,

$$\begin{aligned} u_t - \Delta u &= f, && \text{in } (0,\infty)\times\mathbb{R}^n, \\ u(0,.) &= 0, && \text{on } \{t=0\}\times\mathbb{R}^n, \end{aligned} \tag{4.7}$$

and later combine the solution with (4.5) to include nontrivial initial data (remember, the heat equation is linear).

Note that by Theorem 4.6, for any fixed $s \in (0,t)$, the time-translated function

$$u(t,x;s) := \int_{\mathbb{R}^n} \Phi(t-s,x-y) f(s,y)\,dy$$

solves the initial value problem

$$\begin{aligned} u_t(.;s) - \Delta u(.;s) &= 0, && \text{in } (s,\infty)\times\mathbb{R}^n, \\ u(s,.;s) &= f(s,.), && \text{on } \{t=s\}\times\mathbb{R}^n. \end{aligned} \tag{4.8}$$

This is a homogeneous initial value problem of the form (4.5) with initial time $t = s$ an initial data $g = f(s,.)$. To derive a solution to the inhomogeneous problem (4.7) we apply the so-called *Duhamel principle*. Namely, integrating $u(t,x;s)$ with respect to s leads to

$$u(t,x) := \int_0^t u(x,t;s)\,ds = \int_0^t \int_{\mathbb{R}^n} \Phi(t-s,x-y) f(s,y)\,dy\,ds, \tag{4.9}$$

for all $x \in \mathbb{R}^n$ and $t > 0$. By (4.8), the formal computation

$$\begin{aligned} (u_t - \Delta u)(t,x) &= u(t,x;t) + \int_0^t u_t(t,x;s)\,ds - \int_0^t \Delta_x u(t,x;s)\,ds \\ &= f(t,x) + 0 = f(t,x) \end{aligned}$$

shows that this is indeed the right choice. Due to the singularity at $(t,0)$, however, we have to rigoriously prove that (4.9) is really a (classical) solution of (4.7).

THEOREM 4.10. *Let $f \in C_c^{1,2}([0,\infty)\times\mathbb{R}^n)$, i.e., $f \in C^{1,2}$ with compact support. Then*

$$u(t,x) = \int_0^t \int_{\mathbb{R}^n} \Phi(t-s,x-y) f(s,y)\,dy\,ds,$$

satisfies

(i) $u \in C^{1,2}((0,\infty)\times\mathbb{R}^n)$,
(ii) $u_t - \Delta u = f$ *on* $(0,\infty)\times\mathbb{R}^n$,
(iii) $\lim_{\substack{(t,x)\to(0,x^0)\\ x\in\mathbb{R}^n, t>0}} u(t,x) = 0$.

SKETCH OF PROOF. (i) The regularity of u follows from the regularity of f (the integral is well-defined due to the compact support of f).

(ii) As in the proof of Theorem 4.6 we show that the derivatives can be obtained by differentiation of f under the integral. To move the derivatives onto Φ using integration by parts, one needs to consider integrations on the intervals $(0,\varepsilon)$ and (ε,t) separately due to the singularity of Φ at 0.

(iii) Boundedness of f together with Lemma 4.3 implies that the initial data are assumed.

(Instructions are provided in Assignment 7 so that you can complete the proof.) □

Linearity of the heat equation implies a solution for the combination of (4.5) and (4.7), that is, a general solution to the inhomogeneous initial value problem is obtained both by convolution and the Duhamel principle.

COROLLARY 4.11. *Let $g \in C(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)$ and $f \in C_c^{1,2}([0,\infty)\times\mathbb{R}^n)$. Then*

$$u(t,x) = \int_{\mathbb{R}^n} \Phi(t,x-y) g(y)\,dy + \int_0^t \int_{\mathbb{R}^n} \Phi(t-s,x-y) f(s,y)\,dy\,ds$$

is a classical solution of the initial value problem

$$\begin{aligned} u_t - \Delta u &= f, \qquad \text{in } (0,\infty)\times\mathbb{R}^n,\\ u &= g, \qquad \text{on } \{t=0\}\times\mathbb{R}^n. \quad \square \end{aligned}$$

4.5. Properties of solutions

4.5.1. The maximum principle. We prove a maximum principle for classical solutions of the heat equation on the so-called *parabolic cylinder*

$$\Omega_T := (0,T]\times\Omega,$$

where $T>0$ and $\Omega\subseteq\mathbb{R}^n$ is open and bounded. The *parabolic boundary* is

$$\Gamma_T := \overline{\Omega_T}\setminus\Omega_T = ([0,T]\times\partial\Omega)\cup(\{0\}\times\Omega).$$

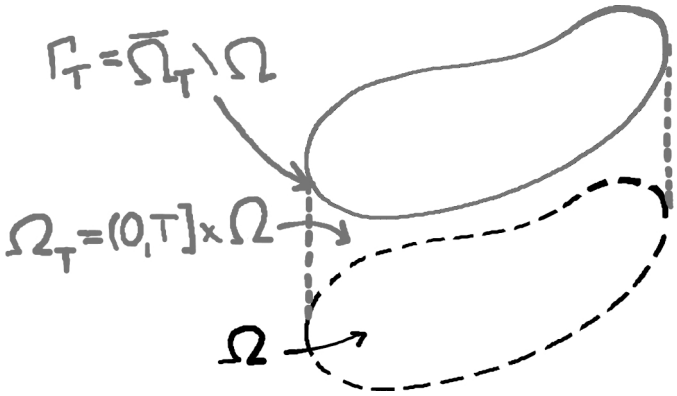

FIGURE 3. Parabolic cylinder and boundary over Ω.

THEOREM 4.12 (Maximum principle for the heat equation). *Let $\Omega\subseteq\mathbb{R}^n$ be open and bounded. Assume that $u\in C^{1,2}(\Omega_T)\cap C(\overline{\Omega_T})$ satisfies $u_t-\Delta u\le 0$ on Ω_T. Then:*

(i) Maximum principle:

$$\max_{\overline{\Omega_T}} u = \max_{\Gamma_T} u.$$

(ii) Strong maximum principle: *If Ω is a domain (i.e., open and connected) and if there exists a point $(t_0,x_0)\in\Omega_T$ with*

$$u(t_0,x_0) = \max_{\overline{\Omega_T}} u,$$

then u is constant on $\overline{\Omega_{t_0}}$.

We will only prove the weak version of the maximum principle (i). The proof of the strong version (ii) requires also the use of a mean value formula (see, for example, [**4**] for the details). With similar assertions we can replace max by min.

REMARK 4.13. Note that (ii) tells us that if a solution u of the heat equations attains a maximum (or minimum) at an interior point (x_0,t_0), then u is constant at all *earlier* times $t\le t_0$. However, the solution may change for $t>t_0$ if the boundary conditions change at a later time.

PROOF OF (I). Consider $Lu := u_t - \Delta u$ and first assume that $Lu < 0$ on Ω_T. Suppose u assumes a maximum in an interior point $(t_0, x_0) \in (0,T) \times \Omega$. Then $u_t(t_0,x_0) = 0$ and $D^2u(t_0,x_0)$ is negative semidefinite, in particular $u_{x_ix_i}(t_0,x_0) \le 0$. Hence $(u_t - \Delta u)(t_0,x_0) \ge 0$, a contradiction. So we have shown that

$$\max_{\overline{\Omega_T}} u = \max_{\partial\Omega_T} u. \tag{4.10}$$

We show next that the same holds true if $Lu \le 0$. In this case consider the perturbed function $u_\varepsilon := u + \varepsilon e^{x_1}$, with $\varepsilon > 0$. This implies that

$$Lu_\varepsilon = Lu - \varepsilon e^{x_1} < 0 \qquad \text{on } \Omega_T.$$

By the above, this implies that

$$\max_{\overline{\Omega_T}} u_\varepsilon = \max_{\partial\Omega_T} u_\varepsilon.$$

Taking the limit $\varepsilon \to 0$ thus yields the same result (4.10) for u.

It remains to be shown that u cannot assume a maximum on the points $\partial\Omega_T \backslash \Gamma_T = \{T\} \times \Omega$. Therefore, first again assume that $Lu < 0$ on Ω_T and u assumes a maximum at a point (T, x_0) with $x_0 \in \Omega$. Then $D^2u(T,x_0)$ is negative semidefinite, thus $-\Delta u(T,x_0) \ge 0$, and hence

$$0 > (u_t - \Delta u)(T,x_0) \ge u_t(T,x_0).$$

However, $u_t(T,x_0) < 0$ contradicts the original assumption that $u(T,x_0)$ is a maximum.

Finally, the general case follows by considering $\tilde{u}_\varepsilon := u + \varepsilon e^{-t}$ for $\varepsilon > 0$. Then

$$L\tilde{u}_\varepsilon = Lu - \varepsilon e^{-t} < 0 \qquad \text{on } \Omega_T.$$

Hence $\tilde{u}_\varepsilon$ has no maximum on $\{T\} \times \Omega$, and $\varepsilon \to 0$ implies the same result for u. □

In order to obtain a maximum principle for the *unbounded* domain $\Omega = \mathbb{R}^n$, an additional growth condition is needed.

THEOREM 4.14 (Maximum principle for the Cauchy problem). *Let $u \in C^{1,2}((0,T] \times \mathbb{R}^n) \cap C([0,T] \times \mathbb{R}^n)$ be a classical solution of the initial value problem (4.5) of the heat equation, that is, of*

$$\begin{aligned} u_t - \Delta u &= 0, \qquad && \text{in } (0,\infty) \times \mathbb{R}^n, \\ u(0,.) &= g, && \text{on } \{t=0\} \times \mathbb{R}^n, \end{aligned}$$

with $g \in C(\mathbb{R}^n)$ a bounded function, that satisfies the growth estimate

$$u(t,x) \le Ae^{a|x|^2}, \qquad \text{for all } (t,x) \in [0,T] \times \mathbb{R}^n,$$

for some $a, A > 0$. Then

$$\sup_{(t,x) \in [0,T] \times \mathbb{R}^n} u(t,x) = \sup_{x \in \mathbb{R}^n} g(x).$$

PROOF. Without loss of generality we assume that $4aT < 1$ (otherwise we can repeat the argument on time slabs of size T). Then there is an $\varepsilon > 0$ such that

$$4a(T+\varepsilon) < 1. \tag{4.11}$$

For fixed $y \in \mathbb{R}^n$ and fixed $\delta > 0$ we consider the function

$$u_\delta(t,x) := u(t,x) - \frac{\delta}{(T+\varepsilon-t)^{\frac{n}{2}}} e^{\frac{|x-y|^2}{4(T+\varepsilon-t)}}$$

If we consider the fundamental solution Φ of the heat equation on $\mathbb{C}^n$ (instead of $\mathbb{R}^n$), then we see that[4]

$$u_\delta(t,x) = u(t,x) - (4\pi)^{\frac{n}{2}}\Phi(T+\varepsilon-t, i(x-y))$$

and hence u_δ also satisfies the heat equation, i.e.,

$$(u_\delta)_t - \Delta u_\delta = 0 \qquad \text{on } (0,T]\times\mathbb{R}^n.$$

In addition to y now c any $r > 0$. The standard maximum principle in Theorem 4.12(i) implies for $\Omega = B_r(y)$ that

$$\max_{\overline{\Omega_T}} u_\delta = \max_{\Gamma_T} u_\delta.$$

On the subset $\{0\}\times\Omega \subseteq \Gamma_T$ we have that

$$u_\delta(0,x) \le u(0,x) = g(x),$$

and on $[0,T]\times\partial B_r(y)$ we obtain

$$u_\delta(t,x) = u(t,x) - \frac{\delta}{(T+\varepsilon-t)^{\frac{n}{2}}} e^{\frac{r^2}{4(T+\varepsilon-t)}} \le Ae^{a(|y|+r)^2} - \frac{\delta}{(T+\varepsilon)^{\frac{n}{2}}} e^{\frac{r^2}{4(T+\varepsilon)}}.$$

Due to (4.11) we have that $\frac{1}{4(T+\varepsilon)} = a+\gamma$ for some $\gamma > 0$, and thus

$$u_\delta(t,x) \le Ae^{a(|y|+r)^2} - \delta(4(a+\gamma))^{\frac{n}{2}} e^{(a+\gamma)r^2} \le \sup_{x\in\mathbb{R}^n} g(x)$$

if we select r large enough. Thus

$$u_\delta(t,x) \le \sup_{\mathbb{R}^n} g \qquad \text{for all } (t,x)\in[0,T]\times\mathbb{R}^n,$$

and for $\delta \searrow 0$, the same holds true for u. □

4.5.2. Uniqueness for the initial and boundary value problem. We can again use the maximum principle to verify uniqueness. The above result, Theorem 4.14, already indicates that a uniqueness result on $\mathbb{R}^n$ does not hold without additional decay assumptions on the solution. In fact, one can show that the initial value problem

$$\begin{aligned} u_t - \Delta u &= 0, && \text{in } (0,T)\times\mathbb{R}^n,\\ u &= 0, && \text{on } \{t=0\}\times\mathbb{R}^n, \end{aligned}$$

has *infinitely many* solutions [**6**, Chapter 7]. Each of the solutions grows very rapidly, except for the solution $u \equiv 0$ which is the only physical solution.

THEOREM 4.15. *Let $\Omega \subseteq \mathbb{R}^n$ be a bounded open domain. The initial-boundary value problem*

$$\begin{aligned} u_t - \Delta u &= f, && \text{in } \Omega_T,\\ u &= h, && \text{on } [0,T]\times\partial\Omega,\\ u(0,.) &= g, && \text{on } \Omega, \end{aligned} \tag{4.12}$$

has at most one classical solution $u \in C^{1,2}(\Omega_T)\cap C(\overline{\Omega_T})$.

[4]One can verify that $(t,x)\mapsto\Phi(T+\varepsilon-t, i(x-y))$ satisfies the heat equation on $(0,T]\times\mathbb{R}^n$. See Assignment 8.

The initial value problem

$$\begin{aligned} u_t - \Delta u &= f, \qquad \text{on } (0,T)\times\mathbb{R}^n, \\ u(0,.) &= g, \qquad \text{on } \mathbb{R}^n, \end{aligned} \tag{4.13}$$

has at most one classical solution satisfying the growth condition

$$|u(t,x)| \le Ae^{a|x|^2}, \qquad (t,x)\in[0,T]\times\mathbb{R}^n, \tag{4.14}$$

for some constants $a, A > 0$.

PROOF. This follows from the weak maximum principle, and maximum principle for the Cauchy problem, respectively.

(i) Let u and v be two classical solutions of the initial-boundary value problem. Their difference $w := u - v$ then satisfies

$$\begin{aligned} w_t - \Delta w &= 0, \qquad \text{in } \Omega_T, \\ w &= 0, \qquad \text{on } \Gamma_T. \end{aligned}$$

The weak maximum principle in Theorem 4.12(i) applied to w and $-w$ implies that $w \le 0$ and $w \ge 0$ in $\overline{\Omega_T}$, and thus $w \equiv 0$.

For (ii) see Assignment 8. □

4.5.3. Long-time behavior. We show that, as time proceeds, the solution of the heat equation approaches the solution of the corresponding Laplace equation. Here we make use of the maximum principle.

THEOREM 4.16. *Let $\Omega \subseteq \mathbb{R}^n$ be open and bounded with C^1 boundary $\partial\Omega$, and let $h \in C(\partial\Omega)$. Then every solution $u \in C^{1,2}((0,\infty)\times\Omega)\cap C([0,\infty)\times\overline{\Omega})$ of*

$$\begin{aligned} u_t - \Delta u &= 0, \qquad \text{on } (0,\infty)\times\Omega, \\ u(t,.) &= h, \qquad \text{on } [0,\infty)\times\partial\Omega, \end{aligned}$$

satisfies

$$\lim_{t\to\infty} u(t,.) = v,$$

uniformly on Ω, where $v \in C^2(\Omega)\cap C(\overline{\Omega})$ is the solution of

$$\begin{aligned} \Delta v &= 0, \qquad \text{on } \Omega, \\ v &= h, \qquad \text{on } \partial\Omega. \end{aligned}$$

PROOF. In order to estimate the difference $|u-v|$ we introduce an auxilliary fuction. Let $\varepsilon > 0$ and define $w_\varepsilon\colon \mathbb{R}^{n+1} \to \mathbb{R}$ by

$$w_\varepsilon(t,x) := \cos(\varepsilon x_1)e^{-\varepsilon^2 t}.$$

Then $(\partial_t - \Delta)w_\varepsilon = 0$ on $\mathbb{R}^{n+1}$, $w_\varepsilon(0,x) > 0$ for all $x \in [-\frac{1}{\varepsilon},\frac{1}{\varepsilon}]^n$ and $w_\varepsilon(t,.) \to 0$ uniformly as $t\to\infty$.[5]

Choose $\varepsilon > 0$ sufficiently small such that $\overline{\Omega} \subseteq [-\frac{1}{\varepsilon},\frac{1}{\varepsilon}]^n$ and define

$$M := \max_{x\in\overline{\Omega}} \frac{|u(0,x)-v(x)|}{|w_\varepsilon(0,x)|}.$$

Then, by definition of u, v, w_ε,

$$(\partial_t - \Delta)(u - v - Mw_\varepsilon) = 0$$

[5]See Assignment 8.

and $u-v-Mw_\varepsilon \le 0$ on Γ_∞. Hence the (weak) maximum principle in Theorem 4.12(i) implies that

$$u \le v + Mw_\varepsilon \qquad \text{on } \Omega_\infty = (0,\infty)\times\Omega.$$

Similarly, we can show that $u - v + Mw_\varepsilon \ge 0$ on Γ_∞ implies that

$$u \ge v - Mw_\varepsilon \qquad \text{on } \Omega_\infty = (0,\infty)\times\Omega.$$

Therefore, $|u-v| \le Mw_\varepsilon$, which implies that

$$\lim_{t\to\infty} \sup_{x\in\overline{\Omega}} |u(t,x) - v(x)| = 0,$$

and therefore yields the desired result. □

4.5.4. Regularity. Classical solutions of the heat equation are automatically smooth. See [**4**, Theorem 2.3.8] for a proof.

THEOREM 4.17 (Smoothness). *Suppose $u \in C^{1,2}(\Omega_T)$ solves the heat equation in Ω_T. Then $u \in C^\infty(\Omega_T)$.*

4.6. Energy methods

As in the case of the Laplace equation, we again make use of energy methods to obtain an alternative uniqueness proof. In order to treat the general initial-boundary value problem we look at the homogeneous Dirichlet and von Neumann boundary problems separately and apply Gronwall's inequality to the energy.

Recall that the Gronwall inequality is an important inequality that turns an implicit inequality into an explicit one.

LEMMA 4.18 (Gronwall's inequality). *Suppose $v\colon [0,T] \to [0,\infty)$ is an integrable function and $C_1, C_2 \ge 0$ are constants such that*

$$v(t) \le C_1 \int_0^t v(s)\,ds + C_2, \qquad t \in [0,T].$$

Then

$$v(t) \le C_2 e^{C_1 t} \le C_2(1 + C_1 t e^{C_1 t}).$$

PROOF. See your ODE course or [**11**, Chapter VII]. □

THEOREM 4.19. *Let $\Omega \subseteq \mathbb{R}^n$ open and bounded with C^1 boundary $\partial\Omega$. Suppose $f \in C(\overline{\Omega_T})$ and $g \in C(\overline{\Omega})$. Suppose $u \in C^{1,2}(\overline{\Omega_T})$ is a solution of the homogeneous initial-boundary value problem, that is the initial value problem*

$$\begin{aligned} u_t - \Delta u &= f, && \text{on } \Omega_T, \\ u(0,.) &= g, && \text{on } \{t=0\}\times\Omega, \end{aligned} \tag{4.15}$$

together with either the homogeneous Dirichlet boundary condition

$$u = 0, \qquad \text{on } [0,T]\times\partial\Omega, \tag{4.16}$$

or the homogeneous von Neumann boundary condition ´

$$\frac{\partial u}{\partial \nu} = 0, \qquad \text{on } [0,T]\times\partial\Omega,$$

where $\frac{\partial u}{\partial \nu}(t,x) = D_x u(t,x)\nu(x)$. *Then* u *satisfies the estimate*

$$\|u(t,.)\|^2_{L^2(\Omega)} + 2\|\nabla u\|^2_{L^2(\Omega_t)} \le e^t(\|g\|^2_{L^2(\Omega)} + \|f\|^2_{L^2(\Omega_t)}), \tag{4.17}$$

for all $t \in (0,T]$.

PROOF. We multiply the inhomogeneous heat equation (4.15) by $2u$ and integrate over Ω_t for $t \in (0,T]$. This yields

$$\int_0^t \int_\Omega (2uu_s - 2u\Delta u)(s,x)\,dx\,ds = \int_0^t \int_\Omega (2uf)(s,x)\,dx\,ds.$$

Since $2uu_s = (u^2)_s$, integration by parts and Green's formula[6] on the left hand side gives us

$$\begin{aligned}&\int_\Omega u^2(t,x)\,dx - \int_\Omega g^2(x)\,dx \\ &\quad - \int_0^t \int_{\partial\Omega} \underbrace{(2u\frac{\partial u}{\partial \nu})(s,x)}_{\substack{=0 \text{ by boundary} \\ \text{conditions}}}\,dS(x)\,ds + \int_0^t \int_\Omega (2\nabla u \cdot \nabla u)(s,x)\,dx\,ds \\ &= \int_0^t \int_\Omega (2uf)(s,x)\,dx\,ds.\end{aligned}$$

Since $2uf \le u^2 + f^2$, this implies

$$\|u(t,.)\|^2_{L^2(\Omega)} + 2\|\nabla u\|^2_{L^2(\Omega_t)} \le \|g\|^2_{L^2(\Omega)} + \|f\|^2_{L^2(\Omega_t)} + \int_0^t \|u(s,.)\|^2_{L^2(\Omega)}\,ds.$$

By Gronwall's inequality, applied to $v(t) = \|u(t,.)\|^2_{L^2(\Omega)} + 2\|\nabla u\|^2_{L^2(\Omega_t)}$, we obtain the explicit inequality

$$\|u(t,.)\|^2_{L^2(\Omega)} + 2\|\nabla u\|^2_{L^2(\Omega_t)} \le e^t(\|g\|^2_{L^2(\Omega)} + \|f\|^2_{L^2(\Omega_t)})$$

for all $t \in [0,T]$. □

An immediate consequence is the uniqueness of initial-boundary value problems.

COROLLARY 4.20. *Let* f, g *be as in Theorem 4.19 and* $h \in C([0,T] \times \partial\Omega)$. *Then the initial-boundary value problem*

$$\begin{aligned} u_t - \Delta u &= f, && \text{in } \Omega_T, \\ u &= h, && \text{on } [0,T] \times \partial\Omega, \\ u(0,.) &= g, && \text{on } \Omega. \end{aligned} \tag{4.18}$$

has at most one classical solution $u \in C^{1,2}(\overline{\Omega_T})$.

PROOF. Consider the difference of two solutions u and v of (4.18). Then $w := u - v$ satisfies (4.15)–(4.16) with $g \equiv 0$ and $f \equiv 0$. Thus the energy estimate (4.17) yields

$$\|w(s,.)\|_{L^2(\Omega)} = 0, \qquad s \in (0,T].$$

Consquently, $w \equiv 0$. □

[6] See Assignment 1, problem 2(b)(ii).

REMARK 4.21. Based on energy methods one can also prove *backward uniqueness* of the heat equation for sufficiently nice domains Ω. That is, if two solutions u and v of the boundary value problem for the heat equation exist on $(0, T] \times \Omega$ and coincide on some time step t_*, i.e., $u(t_*, x) = v(t_*, x)$ for all $x \in \Omega$, then $u = v$ everywhere. (See Assignment 9 for more details.)

CHAPTER 5

The wave equation

5.1. Motivation

The wave equation describes waves in classical physics. For example, it is a (simplified) model for a vibrating string ($n = 1$), membrane ($n = 2$) and elastic solid ($n = 3$). The displacement is denoted by $u(t, x)$ in the point x at time t.

Let $\Omega \subseteq \mathbb{R}^n$ be an open set with C^1 boundary. If the string/membrane/elastic solid has unit mass density, then mass times accelaration within Ω is for sufficiently smooth u given by

$$\frac{d^2}{dt^2} \int_\Omega u(t, x)\, dx = \int_\Omega u_{tt}(t, x)\, dx.$$

The net contact force is

$$-\int_{\partial\Omega} F(t, x) \cdot \nu(x)\, dS(x),$$

where F is the force acting through the surface $\partial\Omega$ on Ω. By Newton's law the net force

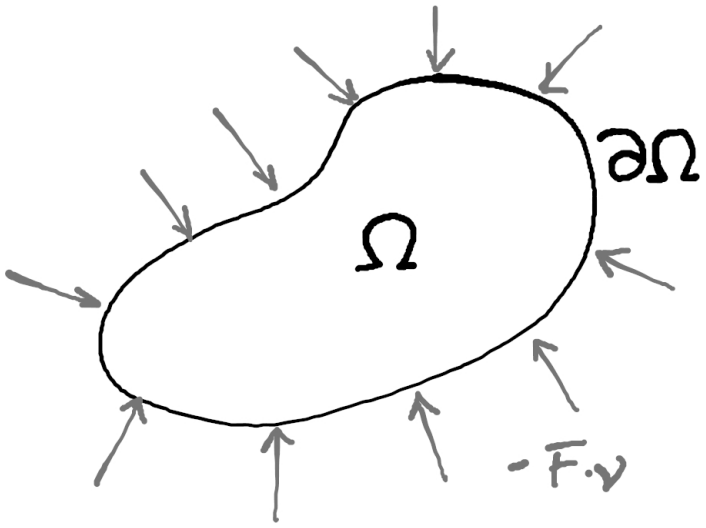

FIGURE 1. Force acting through the boundary $\partial\Omega$.

equals mass times acceleration, hence

$$\int_\Omega u_{tt}(t, x)\, dx = -\int_{\partial\Omega} F(t, x) \cdot \nu(x)\, dS(x),$$

and with the Gauß–Green Theorem we obtain that

$$\int_\Omega u_{tt}(t, x)\, dx = -\int_\Omega \operatorname{div}_x F(t, x)\, dx.$$

For elastic bodies, F is a function of ∇u, and for small displacments the linearization

$$F(t,x) \approx -a\nabla u(t,x),$$

for some $a > 0$, is often appropriate. This leads to the integral equation

$$\int_\Omega u_{tt}(t,x)\,dx = \int_\Omega a\Delta u(t,x)\,dx.$$

If $u \in C^2((0,\infty)\times\Omega)$ then it follows that

$$u_{tt} - a\Delta u = 0.$$

If there is an additional acting volume force Q then we would obtain the inhomogeneous equation

$$u_{tt} - a\Delta u = Q.$$

Rescaling the time variable as in the case of the heat equation leads to the classical wave equation.

5.2. Wave equation

DEFINITION 5.1. Let $\Omega \subseteq \mathbb{R}^n$ be open, $n \geq 1$. The *wave equation* reads

$$u_{tt} - \Delta u = 0, \qquad \text{on } (0,\infty)\times\Omega,$$

and the *inhomogeneous wave equation* reads

$$u_{tt} - \Delta u = f, \qquad \text{on } (0,\infty)\times\Omega,$$

where $f\colon [0,\infty)\times\Omega \to \mathbb{R}$ is given and $u\colon [0,\infty)\times\overline{\Omega} \to \mathbb{R}$ is unknown. We call $\Box = \partial_{tt} - \Delta u$ the *wave operator* (or d'Alembertian, or box operator).

As before, $\Delta u = \Delta_x u$ is the Laplacian with respect to the x-variable only.

For simplicity we will only consider the case $\Omega = \mathbb{R}^n$. The initial value problem should be considered with initial conditions for $u(0,.)$ (initial displacement) and $u_t(0,.)$ (initial velocity), more precisely,

$$\begin{aligned} u_{tt} - \Delta u = 0, &\qquad \text{in } (0,\infty)\times\mathbb{R}^n, \\ u(0,.) = u_0,\ u_t(0,.) = u_1, &\qquad \text{on } \mathbb{R}^n, \end{aligned} \tag{5.1}$$

where $u_0 \in C^2(\mathbb{R}^n)$ and $u_1 \in C^1(\mathbb{R}^n)$.

DEFINITION 5.2. A function $u \in C^2([0,\infty)\times\mathbb{R}^n)$ that satisfies (5.1) is called a *classical solution* to the wave equation.

Solution formulas for the initial value problem (5.1) of the linear wave equation can be derived explicitely, however, their structure and the properties of solutions depend on the spatial dimension n (in particular, whether it is even or odd).

5.3. Representation formulas

5.3.1. D'Alembert's formula in 1D. In one spatial dimension the wave operator factors nicely. That is, if u is a solution to the wave equation then

$$0 = u_{tt} - u_{xx} = (\partial_t - \partial_x)(\partial_t + \partial_x)u. \tag{5.2}$$

Starting with a solution u we compute $u_t + u_x$, call it v, and then $v_t - v_x$ must vanish. Therefore, we have rewritten (5.2) as two equivalent first order equations,

$$v_t - v_x = 0, \tag{5.3}$$

and

$$u_t + u_x = v. \tag{5.4}$$

By Theorem 2.1 for the homogeneous transport equation, we know that the solution to (5.3) is of the form

$$v(t, x) = h(x + t),$$

for any sufficiently regular function h. On the other hand, by Theorem 2.2 for the inhomogeneous transport equation, a general solution to (5.4) is of the form

$$u(t, x) = f(x + t),$$

where $f'(s) = \frac{h(s)}{2}$ (check this by differentiation). Since the equation is linear, we can add $g(x - t)$ to get another solution. This corresponds to adding a particular solution to the homogeneous solution. The most general solution of (5.2) is therefore of the form

$$u(t, x) = f(x + t) + g(x - t). \tag{5.5}$$

(See Assignment 9 for all the details.)

Remark 5.3. Another approach to obtain (5.5) is via *characteristic coordinates*. For coordinates

$$\xi = x + t, \qquad \eta = x - t,$$

equation (5.2) takes the form

$$u_{\xi\eta} = 0,$$

and the solutions of this equation are of the form

$$u = f(\xi) + g(\eta) = f(x + t) + g(x - t).$$

(See Assignment 9 for all the details.)

This general solution shows the simple geometry of the wave equation. There two families of characteristic lines $x \pm t = \text{const}$. The solution is a combination of two waves, $f(x + t)$ is travelling to the left with speed 1, and $g(x - t)$ is travelling to the right with speed 1.

Next we derive solutions to the 1D initial value problem for the wave equation.

Theorem 5.4 (D'Alembert's formula (1746)). *Let $u_0 \in C^2(\mathbb{R})$ and $u_1 \in C^1(\mathbb{R})$. Then* d'Alembert's formula,

$$u(t, x) = \frac{1}{2}(u_0(x - t) + u_0(x + t)) + \frac{1}{2}\int_{x-t}^{x+t} u_1(y)dy, \tag{5.6}$$

provides the unique classical solution $u \in C^2([0, \infty) \times \mathbb{R})$ of the wave equation (5.1) *in one dimension.*

PROOF. We find the solution to (5.1) from the general formula (5.5). Setting $t = 0$ we obtain

$$u_0(x) = f(x) + g(x), \tag{5.7}$$

and by differentiation at $t = 0$ we obtain

$$u_1(x) = f'(x) - g'(x).$$

Hence by addition and substration

$$f' = \frac{1}{2}(u_0' + u_1), \qquad g' = \frac{1}{2}(u_0' - u_1).$$

Integration implies that

$$f(s) = \frac{1}{2}u_0(s) + \frac{1}{2}\int_0^s u_1 + A,$$

and

$$g(s) = \frac{1}{2}u_0(s) - \frac{1}{2}\int_0^s u_1 + B,$$

for some constants A and B. By the initial condition (5.7) we must have that $A + B = 0$. Substituting $s = x \pm t$ into the formulas for f and g we obtain

$$u(x,t) = \frac{1}{2}(u_0(x+t) + u_0(x-t)) + \frac{1}{2}\int_{x-t}^{x+t} u_1(s)\,ds.$$

The regularity of u follows immediately from the regularity of u_0 and u_1. Furthermore, the above calculations show that any classical solution of u is the initial value problem satisfies (5.6). But the formula defines u uniquely. □

5.3.2. Spherical means. The higher space dimensions are handled by deriving a partial differential equation for the *spherical means*, that is, for

$$U(x;t,r) := \fint_{\partial B_r(x)} u(t,y)\,dS(y) = \frac{1}{|\partial B_r(x)|}\int_{\partial B_r(x)} u(t,y)\,dS(y). \tag{5.8}$$

For odd space dimensions, this partial differential equation can be converted to a 1-dimensional wave equation in r. The d'Alembert formula than implies a solution formula. For even dimension, the so-called *method of descent* is applied by generating a solution to a higher dimensional wave equation.

First note that, by using a transformation we already applied for harmonic functions,

$$U(x;t,r) = \frac{1}{|\partial B_r(0)|}\int_{\partial B_r(0)} u(t,x+y)\,dS(y) = \frac{1}{\omega_n}\int_{\partial B_1(0)} u(t,x+ry)\,dS(y).$$

where $\omega_n = |\partial B_1(0)|$ is the surface area of the unit sphere. By setting

$$U(x;t,r) := U(x;t,-r)$$

we get an extension to all $r \in \mathbb{R}$. Whenever u is C^k, the extension $U(x;.)$ is C^k. For the initial data u_0 and u_1 we define the extended spherical means

$$\begin{aligned} U_0(x;r) &:= \frac{1}{\omega_n}\int_{\partial B_1(0)} u_0(x+ry)\,dS(y), \\ U_1(x;r) &:= \frac{1}{\omega_n}\int_{\partial B_1(0)} u_1(x+ry)\,dS(y). \end{aligned} \tag{5.9}$$

Clearly, u can be recovered from U by computing the limit $r \to 0$, i.e.,

$$u(t,x) = \lim_{r\to 0} U(x;t,r) = U(x;t,0).$$

As in the proof of the mean value property (Theorem 3.8), we deduce for $r > 0$ that

$$U_r(x;t,r) = \frac{1}{|\partial B_r(x)|}\int_{B_r(x)} \Delta u(t,y)\,dy = \frac{1}{|\partial B_r(x)|}\int_{B_r(0)} \Delta_x u(t,x+y)\,dy.$$

The right hand side can be expressed in terms of U by using that $|\partial B_r(0)| = \left(\frac{r}{\rho}\right)^{n-1} |\partial B_\rho(0)|$. More precisely,

$$\begin{aligned} U_r(x;t,r) &= \int_0^r \frac{\left(\frac{r}{\rho}\right)^{1-n}}{|\partial B_\rho(0)|}\int_{\partial B_\rho(0)} \Delta_x u(t,x+y)\,dS(y)\,d\rho \\ &= r^{1-n}\Delta_x \int_0^r \rho^{n-1} U(x;t,\rho)\,d\rho. \end{aligned}$$

Multiplication by r^{n-1} and differentiation with respect to r implies that

$$\partial_r(r^{n-1}U_r) = r^{n-1}\Delta_x U.$$

Thus the spherical means satisfy for $r > 0$ the so-called *Darboux equation*

$$U_{rr} + \frac{n-1}{r}U_r = \Delta_x U.$$

We now apply the wave eqation $u_{tt} - \Delta u = 0$ to obtain

$$\begin{aligned} \Delta_x U(x;t,r) &= \Delta_x \frac{1}{|\partial B_r(0)|}\int_{\partial B_r(0)} u(t,x+y)\,dS(y) \\ &= \frac{1}{|\partial B_r(0)|}\int_{\partial B_r(0)} \Delta_x u(t,x+y)\,dS(y) \\ &= \frac{1}{|\partial B_r(0)|}\int_{\partial B_r(0)} u_{tt}(t,x+y)\,dS(y) = U_{tt}(x;t,r), \end{aligned}$$

that is, U satisfies a wave equation in r and t. By definition, our extension of U is an even function in r and hence also

$$\begin{aligned} U_r(x;t,-r) &= -U_r(x;t,r), \\ U_{rr}(x;t,-r) &= U_{rr}(x;t,r). \end{aligned}$$

Therefore, the partial differential equation for U holds for all $r \neq 0$ and we obtain the following result.

THEOREM 5.5. *Let u be a classical solution of the initial value problem* (5.1) *of the wave equation. Then for all $x \in \mathbb{R}^n$ the spherical means $U(x;.)$ with initial data $U_0(x;.)$ and $U_1(x;.)$ as defined in* (5.8)–(5.9) *satisfy the* Euler–Poisson–Darboux equation

$$\begin{aligned} U_{tt}(x;t,r) - U_{rr}(x;t,r) - \frac{n-1}{r}U_r(x;t,r) = 0, &\qquad \text{in } (0,\infty)\times\mathbb{R}, \\ U(x;0,.) = U_0(x;.), \quad U_t(x;0,.) = U_1(x;.), &\qquad \text{on } \mathbb{R}. \end{aligned} \tag{5.10}$$

PROOF. That the spherical means satisfy the Euler–Poisson–Darboux equation was shown above. Moreover, we note that for $x \in \mathbb{R}^n$ and $t \geq 0$ the function $h(r) := U(x;t,x)$ satisfies $h \in C^2(\mathbb{R})$. Since h is even, the derivative h' is odd, and hence, $h(0) = 0$. By L'Hospital's rule we conclude that

$$\lim_{r \to 0} \frac{h'(r)}{r}$$

exists and therefore, the Euler–Poisson–Darboux equation can be considered for all $r \in \mathbb{R}$. □

5.3.3. Kirchhoff's formula in 3D. For $n = 3$ the Euler–Poisson–Darboux formula reads

$$U_{tt}(x;t,r) - U_{rr}(x;t,r) - \frac{2}{r} U_r(x;t,r) = 0 \qquad \text{in } (0,\infty) \times \mathbb{R}. \tag{5.11}$$

We convert this partial differential equation into a wave equation in one dimension and apply d'Alembert's formula.

THEOREM 5.6. *Let u be a classical solution of the intial value problem* (5.1) *of the wave equation in three dimensions. Then for all $x \in \mathbb{R}^3$ the function*

$$\tilde{U}(x;t,r) := rU(x;t,r)$$

with the spherical mean U defined in (5.8) *is a classical solution of the one-dimensional wave equation*

$$\begin{aligned} \tilde{U}_{tt}(x;.) - \tilde{U}_{rr}(x;.) &= 0, && \text{in } (0,\infty) \times \mathbb{R}, \\ \tilde{U}(x;0,r) &= rU_0(x;r), \\ \tilde{U}_t(x;0,r) &= rU_1(x;r), && \text{for } r \in \mathbb{R}. \end{aligned} \tag{5.12}$$

Moreover, if $u_0 \in C^3(\mathbb{R}^3)$ and $u_1 \in C^2(\mathbb{R}^3)$, then u is given by Kirchhoff's formula

$$\begin{aligned} u(t,x) &= \tilde{U}_r(x;t,0) \\ &= \frac{1}{4\pi t^2} \int_{\partial B_t(x)} (u_0(y) + \nabla u_0(y) \cdot (y-x) + t u_1(y))\, dS(y). \end{aligned} \tag{5.13}$$

PROOF. Differentiation implies that

$$\tilde{U}_{tt} = rU_{tt}, \quad \tilde{U}_r = rU_r + U, \quad \tilde{U}_{rr} = 2U_r + rU_{rr}.$$

Hence by the Euler–Poisson–Darboux equation (5.11) we have that

$$\tilde{U}_{tt} - \tilde{U}_{rr} = r\left(U_{tt} - \frac{2}{r}U_r - U_{rr}\right) = 0, \qquad \text{in } (0,\infty) \times \mathbb{R} \setminus \{0\}.$$

Since $\tilde{U} \in C^2([0,\infty) \times \mathbb{R})$ this equation is valid for all $r \in \mathbb{R}$. The initial conditions follow directly from (5.10) from those of U. Hence $\tilde{U}$ is a classical solution of (5.12) and by d'Alembert's formula

$$\tilde{U}(x;t,r) = \frac{1}{2}((r-t)U_0(x;r-t) + (r+t)U_0(x;,r+t)) + \frac{1}{2}\int_{r-t}^{r+t} yU_1(x;y)\, dy.$$

Differentation in r at $r = 0$ yields

$$\tilde{U}_r(x;t,0) = 0 \cdot U_r(x;t,0) + U(x;t,0) = u(t,x)$$

and thus

$$\begin{aligned} u(t,x) &= \tilde{U}_r(x;t,0) \\ &= \frac{1}{2}\left(U_0(x;-t) + U_0(x;t) - t(U_0'(x;-t) - U_0'(x;t)) + t(U_1(x;t) + U_1(x;-t))\right). \end{aligned}$$

To differentiate the last term we used the Leibniz formula

$$\partial_r \left(\int_{a(r)}^{b(r)} f(s,r)\, ds \right) = f(b(r),r)\cdot b'(r) - f(a(r),r)\cdot a'(r) + \int_{a(r)}^{b(r)} \partial_r f(s,r)\, ds$$

for $f(s,r) = sU_1(x;s)$. Since U_0 and U_1 are even in r (by definition),

$$u(t,x) = U_0(x;t) + tU_0'(x;t) + tU_1(x;t). \tag{5.14}$$

Here,

$$\begin{aligned} U_0'(x;t) &= \frac{1}{|\partial B_1(0)|}\int_{\partial B_1(0)} \nabla u_0(x+ty)\cdot y\, dS(y) \\ &= \frac{1}{|\partial B_t(x)|}\int_{\partial B_t(x)} \nabla u_0(z)\cdot \frac{z-x}{t}\, dS(z). \end{aligned}$$

Therefore, since $|\partial B_t(x)| = 4\pi t^2$, we obtain Kirchhoff's formula

$$u(t,x) = \frac{1}{4\pi t^2}\int_{\partial B_t(x)} (u_0(y) + \nabla u_0(y)\cdot(y-x) + tu_1(y))\, dS(y). \qquad \square$$

REMARK 5.7. Kirchhoff's formula can be extended to any odd space dimension. For $n = 2k+1 \geq 3$ a reduction of Euler–Poisson–Darboux equation (5.10) to the wave equation in one dimension can be achieved via the function

$$\tilde{U}(x;t,r) := \left(\tfrac{1}{r}\partial_r\right)^{k-1}(r^{2k-1}U(x;t,r)).$$

For more details see [**4**, Section 2.4.1.d].

REMARK 5.8. By comparing Kirchhoff's formula to d'Alembert's formula we observe that the latter does not involve derivatives of the initial data. Hence, the solution of the wave equation at $t > 0$ for $n > 1$ may be less regular than the initial data.

In Theorem 5.6 have shown that every classical solution of the initial value problem (5.1) of the wave equation in three space dimensions satisfies Kirchhoff's formula. This establishes the uniqueness of such a solution, however, it remains to prove *existence* of classical solutions. We will do so in Section 5.4 after having derived a formula for even dimensions.

5.3.4. Method of descent and Poisson's formula in 2D. For even dimensions a reduction of the Euler–Poisson–Darboux equation to a 1D wave equation is not possible. Instead we view the 2D wave equation in three dimensions. More precisely, assuming that $u \in C^2([0,\infty)\times\mathbb{R}^2)$ is a solution the initial value problem (5.1) with initial data (u_0,u_1) in 2D we consider

$$\bar{u}(t,x,x_3) := u(t,x), \qquad (t,x,x_3) \in [0,\infty)\times\mathbb{R}^2\times\mathbb{R},$$

which then satisfies the 3D wave equation for the initial data given by

$$\bar{u}_0(x,x_3) := u_0(x), \quad \bar{u}_1(x,x_3) := u_1(x).$$

This is called the *method of descent*.

We can now apply Kirchhoff's formula (5.13). Since the integrand should not depend on x_3, we can simplify the integral. In what follows we therefore denote by B^k a ball in $\mathbb{R}^k$. Then we can use the parametrization $(x_1, x_2, \pm\gamma(x_1, x_2))$ with $\gamma(x_1, x_2) = \sqrt{r^2 - x_1^2 - x_2^2}$ for $\partial B_r^3(0)$ and an integral of an integrable function v independent of x_3 over $\partial B_r^3(0)$ is of the form[1]

$$\begin{aligned}\int_{\partial B_r^3(0)} v(y)\, dS(y, y_3) &= 2\int_{B_r^2(0)} v(y)\sqrt{1 + |\nabla\gamma(y)|^2}\, dy \\ &= 2\int_{B_r^2(0)} \frac{r}{\sqrt{r^2 - |y|^2}}\, v(y)\, dy.\end{aligned}$$

Thus for initial data $\bar{u}_0$ and $\bar{u}_1$ Kirchhoff's formula yields *Poisson's formula*

$$\begin{aligned}u(t, x) &= \frac{1}{4\pi t^2}\int_{\partial B_t^3(x)} (\bar{u}_0(y) + \nabla\bar{u}_0(x)\cdot(y - x) + t\bar{u}_1(y))\, dS(y) \\ &= \frac{1}{2\pi t}\int_{B_t(x)} \frac{u_0(y) + \nabla u_0(y)\cdot(y - x) + tu_1(y)}{\sqrt{t^2 - |y - x|^2}}\, dy \qquad (5.15)\end{aligned}$$

for the solution of the initial value problem (5.1) for $n = 2$.

REMARK 5.9. For general even $n = 2k \geq 2$ the method of descent can be applied by considering solutions $\bar{u}(t, x, x_{n+1}) = u(t, x)$ of the $n + 1$ dimensional wave equation.

5.4. Existence and properties of solutions

Previously in Section 5.3 we have shown that *if* classical solutions to the 3D and 2D wave equations exist, *then* they are given by Kirchhoff's and Poisson's formula, respectively. This immediately implies that such solutions are unique. It remains to be shown that classical solutions indeed exist for sufficiently regular initial data.

In the following, for $a \in \mathbb{R}$, $\lfloor a \rfloor$ denotes the largest integer $\leq a$.

THEOREM 5.10. *Let $n \geq 2$, $m = \lfloor \frac{n}{2} \rfloor + 1$ and $u_0 \in C^{m+1}(\mathbb{R}^n)$, $u_1 \in C^m(\mathbb{R}^n)$. Then the initial value problem for the wave equation* (5.1) *has a classical solution $u \in C^2([0, \infty \times \mathbb{R}^n)$. For $n = 2$ it is given by Poisson's formula* (5.15) *and for $n = 3$ by Kirchhoff's formula* (5.13). *Moreover, $u(t, x)$ only depends on the initial data on $\partial B_t(x)$ if n is odd, and on $\overline{B_t(x)}$ is n is even.*

PROOF. We prove the Theorem for $n = 2, 3$. Then $m = 2$ and thus $u_0 \in C^3$ and $u_1 \in C^2$. Other cases are also considered in [**4**, Section 2.4.1.d].

For $n = 3$ we show that Kirchhoff's formula (5.13) is a a classical solution. We can rewrite the formula as

$$\begin{aligned}u(t, x) &= \frac{1}{|\partial B_t(x)|}\int_{\partial B_t(x)} (u_0(y) + t\nabla u_0(y)\cdot(y - x) + tu_1(y))\, dS(y) \\ &= \frac{1}{|\partial B_1(0)|}\int_{\partial B_1(0)} (\underbrace{u_0(x + ty) + t\nabla u_0(x + ty)\cdot y}_{= \frac{d}{dt}(tu_0(x+ty))} + tu_1(x + ty))\, dS(y). \qquad (5.16)\end{aligned}$$

Due to the regularity of the initial data, the integrand is C^2 with respect to t, x, y and $\partial B_1(0)$ is compact. Therefore, the right hand side is $C^2([0, \infty) \times \mathbb{R}^n)$.

[1] See the prerequisites on surface integrals.

We next show that Kirchhoff's formula satisfies the wave equation. Let us consider the case $u_0 \equiv 0$ first. Then (5.13) reads

$$u(t,x) = \frac{t}{|\partial B_t(0)|} \int_{\partial B_t(0)} u_1(x+y)\, dS(y) =: tU_1(x;t).$$

This implies that

$$\Delta u(t,x) = \frac{t}{|\partial B_t(0)|} \int_{\partial B_t(0)} \Delta u_1(x+y)\, dS(y).$$

Similarly we have that

$$u_t = (tU_1)_t = U_1 + t(U_1)_t, \qquad u_{tt} = (tU_1)_{tt} = 2(U_1)_t + t(U_1)_{tt}$$

and by using essentially Green's formula (as earlier) we have that

$$(U_1)_t(x;t) = \frac{1}{|\partial B_t(0)|} \int_{B_t(0)} \Delta_x u_1(x+y)\, dy.$$

Note that $|\partial B_t(0)| = t^2 |\partial B_1(0)| = 4\pi t^2$ yields $\frac{d}{dt}\frac{1}{|\partial B_t(0)|} = -\frac{2}{t|\partial B_t(0)|}$ and thus

$$(U_1)_{tt}(x;t) = -\frac{2}{t}(U_1)_t(x;t) + \frac{1}{|\partial B_t(0)|} \int_{\partial B_t(0)} \Delta_x u_1(x+y)\, dS(y).$$

Thus

$$u_{tt} = 2(U_1)_t + t(U_1)_{tt} = \frac{t}{|\partial B_t(0)|} \int_{\partial B_t(0)} \Delta_x u_1(x+y)\, dS(y) = \Delta u(t,x).$$

Now assume that $u_0 \not\equiv 0$. First note that the C^3 function

$$v(t,x) = \frac{1}{|\partial B_1(0)|} \int_{\partial B_1(0)} t u_0(x+ty)\, dS(y)$$

solves the wave equation as well (we can just replace u_1 by u_0 in the previous step). Due to regularity, we have that $\partial_t \Box v = \Box \partial_t v = 0$, hence also the derivative

$$v_t(t,x) = \frac{1}{|\partial B_1(0)|} \int_{\partial B_1(0)} \frac{d}{dt}(t u_0(x+ty))\, dS(y)$$

solves the wave equation. This is, however, just the missing first component in the formula (5.16) for $u(t,x)$. So both cases together imply the existence of the unique solution, which is given by Kirchhoff's formula.

It remains to be shown that (5.16) is consistant with the initial data u_0 and u_1. (This is part of Assignment 10.)

For $n = 2$ we have seen that Poisson's formula ist just Kirchhoff's formula for initial data that are independent of x_3. Therefore, we obtain a solution to the 2D wave equation.

The statements about the dependence of the solution on the initial data follow directly from the integral expressions used in Kirchhoff's formula (5.13) and Poisson's formula (5.15). □

Theorem 5.10 and the representation formulas reveal several important properties about solutions of the scalar wave equation that also carry over to nonlinear wave equations. We summarize these observations and principles in the following remarks.

REMARK 5.11 (Domain of dependence/influence). The region that determines the value of $u(t,x)$ is called the *domain of dependence*. This set forms a cone. If we change the initial conditions outside this region, the value of $u(t,x)$ will not change. Similarly, one can define a *region of influence* for a given subset of $\mathbb{R}^n$, which consists of all points that are influenced by the values of the initial data on this particular domain. This set forms an inverted truncated cone.

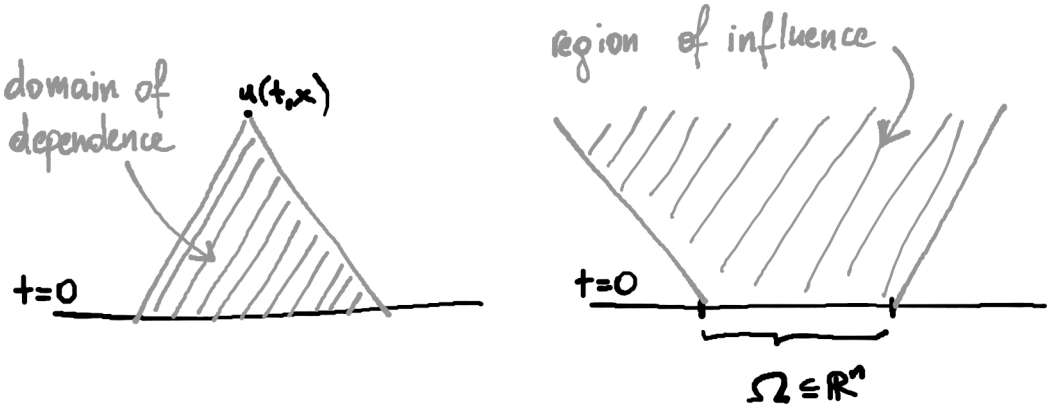

FIGURE 2. Domain of dependence of $u(t,x)$ (on left) and domain of influence of $\Omega \subseteq \mathbb{R}^n$ (on right).

REMARK 5.12 (Regularity and speed of propagation). Theorem 5.10 reveals two aspects about solutions to the wave equation.

(i) For $n \geq 2$ the solution can be less regular than the initial data. This is caused by a *focusing effect*, which means that irregularities in u_0 may focus at a later time and cause u to be less regular. This is in stark contrast to the Laplace and heat operators which have a smoothing effect meaning that, for instance, solutions of the heat equation are smooth for $t > 0$ if u_0 is bounded and continuous.
(ii) The wave equation exhibits a *finite speed of propagation*. This is, $u(t,x)$ depends only on u_0 and u_1 restricted to $\overline{B_t(x)}$ (or even just $\partial B_t(x)$). In constrast, a solution of the heat equation depends on the initial data u_0 on all of $\mathbb{R}^n$.

REMARK 5.13 (Huygens' principle). In optics, *Huygens' principle* is a statement about wave propagation. From solution formulas derived in Section 5.3 we observe the following, which is also exemplary for higher odd/even dimensions.

(i) In dimension $n = 3$, by Kirchhoff's formula, the initial data in a given point $x \in \mathbb{R}^3$ only affect the solution on the boundary $\{(t,y) : t > 0, |y-x| = t\}$ of the cone $C = \{(t,y) : t > 0, |y-x| < t\}$. That is, a disturbance in the point x propagates along a sharp wave front.
(ii) On the other hand, in dimension $n = 2$, by Poisson's formula, the initial data in a given point $x \in \mathbb{R}^n$ affect the solution in the solid cone $\overline{C}$. That is, a disturbance in the point x continues to have an effect even after the leading edge of the wavefront has passed.

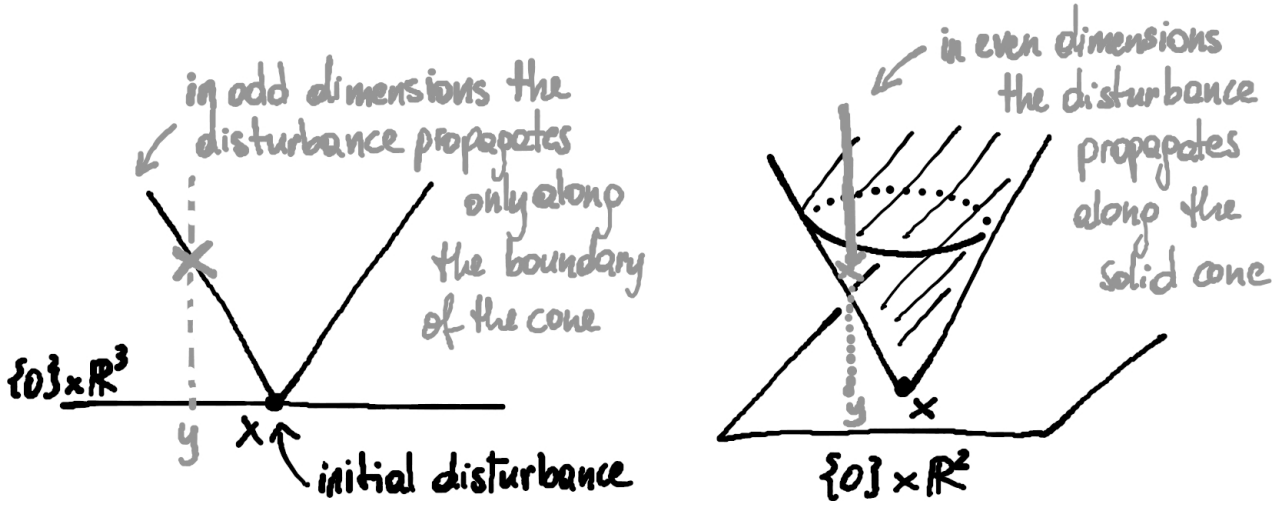

FIGURE 3. Propagation of wave fronts in dimension 3 along the cone (on left, two dimensions surpressed) and dimension 2 along the solid cone (or right).

5.5. The inhomogeneous case

As in the case of the heat equation we apply *Duhamel's principle*. That is, we set

$$u(t,x) := \int_0^t u(t,x;s)\,ds \tag{5.17}$$

for the solution $u(t,x;s)$ of the homogeneous wave equation

$$\begin{aligned} u_{tt}(.;s) - \Delta u(.;s) &= 0, && \text{in } (s,\infty)\times\mathbb{R}^n,\\ u(s,.;s) = 0, \quad u_t(s,.;s) &= f(s,.) && \text{on } \mathbb{R}^n, \end{aligned} \tag{5.18}$$

given by Theorem 5.10.

THEOREM 5.14. *Let $n \geq 1$, $m = \lfloor \frac{n}{2} \rfloor + 1$ and $f \in C^m([0,\infty)\times\mathbb{R}^n)$. Then u defined by* (5.17) *is a classical solution of the initial value problem for the inhomogeneous wave equation*

$$\begin{aligned} u_{tt} - \Delta u &= f, && \text{in } (0,\infty)\times\mathbb{R}^n,\\ u(0,.) = 0, \quad u_t(0,.) &= 0, && \text{on } \mathbb{R}^n. \end{aligned} \tag{5.19}$$

PROOF. By Theorem 5.10 we know that $u(.;s)$ is in $C^2([s,\infty)\times\mathbb{R}^n)$. Hence by (5.17) it follows directly that $u \in C^2([0,\infty)\times\mathbb{R}^n)$. Furthermore we have that

$$\begin{aligned} u_t(t,x) &= \int_0^t u_t(t,x;s)\,ds + u(t,x;t) = \int_0^t u_t(t,x;s)\,ds,\\ u_{tt}(t,x) &= \int_0^t u_{tt}(t,x;s)\,ds + u_t(t,x;t) = \int_0^t u_{tt}(t,x;s)\,ds + f(t,x),\\ \Delta u(t,x) &= \int_0^t \Delta u(t,x;s)\,ds. \end{aligned}$$

Therefore, since $u(.;s)$ satisfies the wave equation,

$$u_{tt}(t,x) - \Delta u(t,x) = \int_0^t (u_{tt} - \Delta u)(t,x;s)\,ds + f(t,x) = f(t,x).$$

Moreover, $u(0,x) = 0$ and $u_t(0,x) = 0$ by the above. Hence the initial conditions are satisfied as well. □

REMARK 5.15. An explicit representation formula for $n = 1, 2, 3$ can be obtained from (5.17) (see Assignment 11).

Due to the linearity of the wave equation, the solution of the general problem follows by adding up the solution of (5.17) and the solution of the initial value problem (5.1) for the homogeneous case.

COROLLARY 5.16. *Let $u_0 \in C^{\lfloor \frac{n}{2} \rfloor + 2}(\mathbb{R}^n)$, $u_1 \in C^{\lfloor \frac{n}{2} \rfloor + 1}(\mathbb{R}^n)$ and $f \in C^{\lfloor \frac{n}{2} \rfloor + 1}([0,\infty) \times \mathbb{R}^n)$. Then there exists a classical solution of the general inhomogeneous initial value problem*

$$\begin{aligned} u_{tt} - \Delta u &= f, && \text{in } (0,\infty) \times \mathbb{R}^n, \\ u(0,.) = u_0, \quad u_t(0,.) &= u_1, && \text{on } \mathbb{R}^n. \quad \square \end{aligned}$$

5.6. Energy methods

The explicit solution formulas for the wave equation show that with increasing space dimension n higher and higher regularity assumptions are required for the initial data u_0, u_1 in order to obtain a classical C^2 solution. Energy "norms" present an alternative to measure the size and regularity of solutions. In this section we use energy methods to prove the uniqueness and to examine the domain of dependence of solutions.

We derive an energy inequality for solutions of the initial value problem

$$\begin{aligned} u_{tt} - \Delta u &= f && \text{in } (0,\infty) \times \mathbb{R}^n, \\ u(0,\cdot) = u_0, \quad u_t(0,\cdot) &= u_1 && \text{on } \mathbb{R}^n. \end{aligned} \tag{5.20}$$

To this end for a given point $(t_0, x_0) \in (0,\infty) \times \mathbb{R}^n$ we define the *backward light cone* (domain of dependence) by

$$C(t_0, x_0) = \{(s,x) : 0 \le s \le t_0,\ |x - x_0| \le t_0 - s\},$$

and for $t \in (0, t_0)$ the *truncated backward light cone* by

$$C(t; t_0, x_0) = \{(s,x) : 0 \le s \le t,\ |x - x_0| \le t_0 - s\}.$$

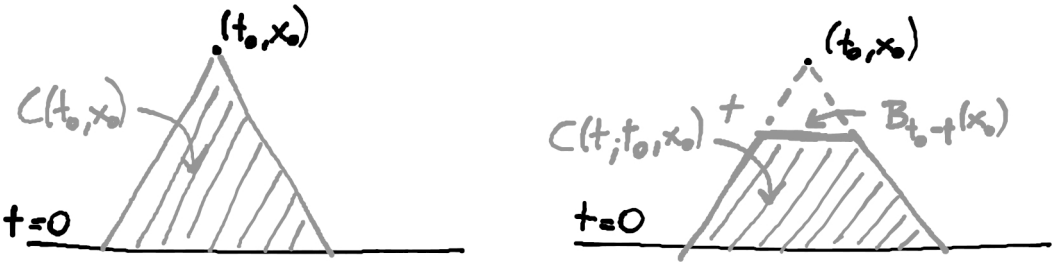

FIGURE 4. The backward light cone $C(t_0, x_0)$ (on left) and truncated backward light cone $C(t; t_0, x_0)$ (on right).

THEOREM 5.17. *Let $u_0 \in C^2(\mathbb{R}^n)$, $u_1 \in C^1(\mathbb{R}^n)$ and $f \in C([0,\infty)\times\mathbb{R}^n)$. Then, any classical solution $u \in C^2([0,\infty)\times\mathbb{R}^n)$ of (5.20) satisfies for all $(t_0,x_0) \in (0,\infty)\times\mathbb{R}^n$ and $\varepsilon > 0$ the energy estimate*

$$\begin{aligned}\|u_t(t,\cdot)\|^2_{L^2(B_{t_0-t}(x_0))} + \|\nabla u(t,\cdot)\|^2_{L^2(B_{t_0-t}(x_0))} \\ \leq e^{\varepsilon t}\left(\|u_1\|^2_{L^2(B_{t_0}(x_0))} + \|\nabla u_0\|^2_{L^2(B_{t_0}(x_0))} + \frac{1}{\varepsilon}\|f\|^2_{L^2(C(t;t_0,x_0))}\right),\end{aligned}$$

where $t \in (0,t_0)$. If $f \equiv 0$, then $\varepsilon = 0$ is allowed.

PROOF. Let $t_0 > 0, x_0 \in \mathbb{R}^n$ and $\varepsilon > 0$. Defining the energy

$$e(t) := \int_{B_{t_0-t}(x_0)} \left(u_t^2(t,\cdot) + |\nabla u(t,\cdot)|^2\right)\, dx, \qquad t > 0,$$

we observe that

$$e'(t) = \int_{B_{t_0-t}(x_0)} (2u_t u_{tt} + 2\nabla u_t\cdot\nabla u) - \int_{\partial B_{t_0-t}(x_0)} \left(u_t^2 + |\nabla u|^2\right)\, dS(x),$$

since $\int_{B_{t_0-t}(x_0)} g(y)\,dy = \int_0^{t_0-t}\int_{\partial B_\rho(x_0)} g(y)\,dS(y)d\rho$ for a continuous function g. Using integration by parts it follows that

$$\begin{aligned}\int_{B_{t_0-t}(x_0)} (2u_t u_{tt} + 2\nabla u_t\cdot\nabla u)\,dx &= \int_{B_{t_0-t}(x_0)} 2u_t(u_{tt}-\Delta u)\,dx + \int_{\partial B_{t_0-t}(x_0)} 2u_t\nabla u\cdot\nu\, dS(x) \\ &= \int_{B_{t_0-t}(x_0)} 2u_t f + \int_{\partial B_{t_0-t}(x_0)} 2u_t\nabla u\cdot\nu\, dS(x),\end{aligned}$$

where we used that u is a solution of the inhomogeneous wave equation (5.20). We further estimate the right hand side using the inequalities $2ab \leq a^2+b^2$ and $2ab \leq \varepsilon a^2 + \frac{b^2}{\varepsilon}$, $a,b\in\mathbb{R}$,

$$\begin{aligned}\int_{B_{t_0-t}(x_0)} 2u_t f\,dx + \int_{\partial B_{t_0-t}(x_0)} 2u_t\nabla u\cdot\nu\, dS(x) \leq{}& \varepsilon\|u_t\|^2_{L^2(B_{t_0-t}(x_0))} + \frac{1}{\varepsilon}\|f\|^2_{L^2(B_{t_0-t}(x_0))} \\ &+ \int_{\partial B_{t_0-t}(x_0)} \left(u_t^2 + |\nabla u|^2\right)\, dS(x).\end{aligned}$$

Consequently, this estimate implies that

$$e'(t) \leq \varepsilon\|u_t\|^2_{L^2(B_t)} + \frac{1}{\varepsilon}\|f\|^2_{L^2(B_{t_0-t}(x_0))} \leq \varepsilon e(t) + \frac{1}{\varepsilon}\|f\|^2_{L^2(B_{t_0-t}(x_0))},$$

and integrating the inequality from 0 to t we obtain

$$e(t) \leq e(0) + \frac{1}{\varepsilon}\|f\|^2_{L^2(C(t;t_0,x_0))} + \varepsilon\int_0^t e(s)\,ds.$$

Finally, the energy inequality follows by Gronwall's lemma applied to the function e. □

An immediate consequence of the energy estimate in Theorem 5.17 is the finite speed of propagation and uniqueness of solutions.

COROLLARY 5.18 (Finite speed of propagation). *Let $u_0 \in C^2(\mathbb{R}^n)$, $u_1 \in C^1(\mathbb{R}^n)$ and u be a classical solution of the homogeneous initial value problem (5.1).*

If $t_0 > 0, x_0 \in \mathbb{R}^n$ and $u \equiv u_t \equiv 0$ on $\{t=0\}\times B_{t_0}(x_0)$ then $u \equiv 0$ within the cone $C(t_0,x_0)$.

PROOF. By the energy inequality for $f = 0$ in Theorem 5.17 it follows that $e(t) = 0$ for all $0 \le t \le t_0$. Hence, we conclude that $u_t \equiv \nabla u \equiv 0$ on $C(t_0, x_0)$, which implies that $u \equiv 0$ on $C(t_0, x_0)$ as $u(0, \cdot) \equiv 0$. □

We notice that any disturbance originating outside of $B_{t_0}(x_0)$ has no effect on the solution within the cone $C(t_0, x_0)$, and consequently, has a finite speed of propagation. We had already observed this property based on the representation formulas for solutions in dimensions $n = 1, 2, 3$. Energy methods provide a much simpler proof and do not require the knowledge of explicit solution formulas.

This immediately also implies uniqueness of solutions.

COROLLARY 5.19 (Uniqueness). *Let u_0, u_1 and f be as in Theorem 5.17. Then, there exists at most one classical solution of the initial value problem* (5.20).

PROOF. Let u and v be two classical solution. Their difference $w = u - v$ satisfies the initial value problem (5.20) with $w_0 = w_1 = f \equiv 0$. Hence by Corollary 5.18 we conclude that $w \equiv 0$ in $C(t_0, x_0)$ for all $t_0 > 0$ and all $x_0 \in \mathbb{R}^n$, hence everywhere. □

REMARK 5.20. A uniqueness result can also be obtained for the initial boundary value problem on open and bounded domains $\Omega \subseteq \mathbb{R}^n$ with C^1 boundary $\partial\Omega$, i.e., for the problem

$$\begin{aligned} u_{tt} - \Delta u &= f, && \text{in } \Omega_T := \Omega \times (0, T], \\ u &= g, && \text{on } \Gamma_T := \overline{\Omega_T} \setminus \Omega_T, \\ u_t &= h, && \text{on } \{t = 0\} \times \Omega. \end{aligned} \tag{5.21}$$

for any $T > 0$. It follows by considering the same energy as in the proof of Theorem 5.17 (but on all of Ω), that is,

$$e(t) := \frac{1}{2} \int_\Omega \underbrace{w_t^2(t, x)}_{E_{\text{kin}}} + \underbrace{|\nabla w(t, x)|^2}_{E_{\text{pot}}} \, dx, \qquad t \in [0, T].$$

for the difference of $w = u - v$ of two solutions.

One can also consider this problem on all of $\mathbb{R}^n$ but for compactly supported initial data. (See Assignment 12 for more details.)

5.7. The wave equation and relativistic geometry (supplementary material)

In 1915 Einstein developed a model of the universe as a 4-dimensional so-called spacetime, a manifold with Lorentzian metric $\mathbf{g}$. General relativity is the study of solutions $\mathbf{g}$ to the so-called Einstein equations,

$$\mathbf{G}[\mathbf{g}] = \frac{8\pi G}{c^4} \mathbf{T},$$

where $\mathbf{G}$ is the Einstein tensor based on the curvature of the manifold (containing second derivatives of $\mathbf{g}$), $\mathbf{T}$ is the energy-momentum tensor of the matter model used, G is Newton's gravitational constant and c is the speed of light. In the case of vacuum, the Einstein equations simplify to

$$\mathbf{Ric}[\mathbf{g}] = 0,$$

where $\mathbf{Ric}$ denotes the Ricci curvature with respect to $\mathbf{g}$.

The Einstein equations are a system of 10 (diffeomorphism invariant) nonlinear partial differential equations. They can transformed to nonlinear elliptic constraint equations and a system of quasilinear hyperbolic equations of the form

$$\Box_{\mathbf{g}}\mathbf{g} + F(\mathbf{g}, \partial\mathbf{g}) = 0.$$

In 1952 Yvonne Choquet-Bruhat proved for the first time a local existence result for the Einstein equations in vacuum (see Ringström [**8**]). Much of the current research is related to global problems, that is, the understanding the long-time behavior of solutions and the occurrence of singularities and black holes.

Some features of the geometry of spacetimes, and hence the way things move in the universe, can be understood already by looking at the linear wave equation in special relativity (see [**9**, Section 9.3]).

DEFINITION 5.21 (Causal structure). Consider the point $o = (0, 0, 0, 0)$. The *past*[2] of o is the set $\{ct < -|x|\}$, its *future* is $\{ct > |x|\}$ and its *present* is $\{-|x| < ct < |x|\}$.

A 4-dimensional vector (v^0, v) is called

(i) *timelike* if $|v^0| > c|v|$,
(ii) *spacelike* if $|v^0| < c|v|$,
(iii) *null* (or lightlike, characteristic) if $|v^0| = c|v|$.

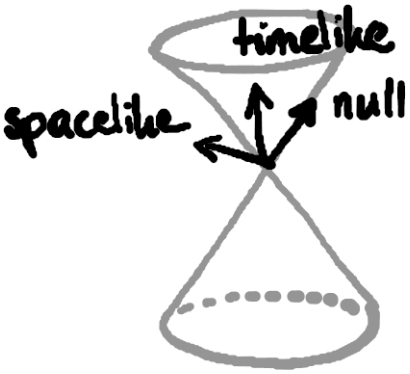

FIGURE 5. Future-directed timelike, null and spacelike vectors with light cone ($c = 1$).

Timelike vectors can be either future- or past-directed. A straight line is called *ray* if its tangent vector is null. A surface in spacetime is called *characteristic* if its 4-dimensional normal vector is a null vector. For example, if a surface $S = \{t = \gamma(x)\}$, then a normal vector is $(-1, \nabla\gamma(x))$. This vector is null if $1 = |v^0| = c|v| = c|\nabla\gamma(x)|$

A surface is called *spacelike* if all its normal vectors are timelike, i.e., $|\nabla\gamma(x)| < \frac{1}{c}$. For instance, the initial surface $\{t = 0\}$ is a spacelike surface since $\gamma \equiv 0$. Spacelike surfaces are naturally the ones that carry initial data.

[2]Note that $|.|$ denotes the absolute value in $\mathbb{R}$ as well as the Euclidean norm in $\mathbb{R}^n$ here (usually the former for t and v^0 and the latter for x or v).

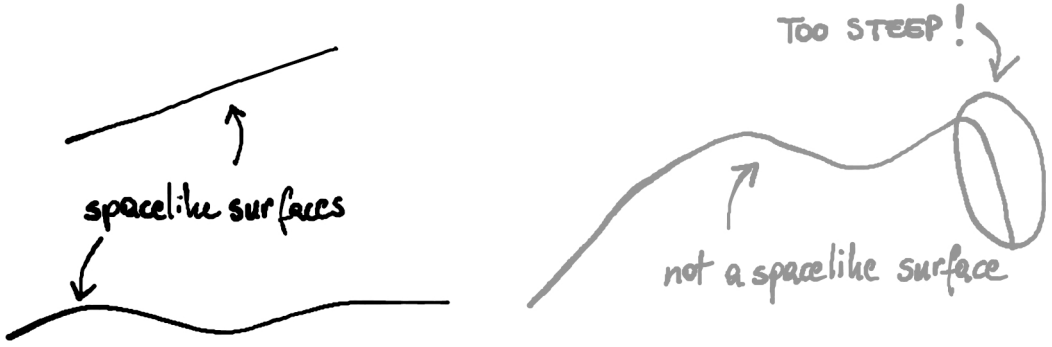

FIGURE 6. Spacelike and not spacelike surfaces.

THEOREM 5.22. *If S is a spacelike surface, then one can uniquely solve the initial value problem*

$$u_{tt} = c^2 \Delta u, \qquad \text{in all of spacetime},$$
$$u = \phi, \ \frac{\partial u}{\partial \nu} = \psi, \qquad \text{on } S,$$

where $\frac{\partial}{\partial \nu}$ indicates the derivative in the direction normal to S. If S is represented as $\{t = \gamma(x)\}$, the second initial condition reads

$$u_t - \nabla \gamma \cdot \nabla u = \sqrt{1 + |\nabla \gamma|^2}\, \psi, \qquad \text{for } t = \gamma(x).$$

We will not prove this result here, but note that if $S = \{t = 0\}$, then $u_t = \psi$, and we are back to the setting for the usual initial value problem.

CHAPTER 6

Limitations of classical solutions

So far we only considered solutions of partial differential equations that possess continuous (partial) derivatives up to the order of the partial differential equation, and therefore satisfy the partial differential equation *pointwise*. This is problematic, because in order to ensure such a regularity we needed

- certain regularity of the boundary of the region, and
- certain regularity of the initial and/or boundary data.

Modern PDE theory takes another approach, consisting of the following two steps:

(i) **Weak solutions:** The concept of a "solution" for a partial differential equation is weakened in a way that the existence of solutions is easier to obtain.

(ii) **Regularity results:** One then shows that under certain regularity assumptions on the boundary of the region and the initial/boundary data the weak solution is actually more regular.

EXAMPLE 6.1. We explain this for the Dirichlet problem for Poisson's equation

$$-\Delta u = f, \qquad \text{in } \Omega, \tag{6.1}$$

$$u = 0, \qquad \text{on } \partial\Omega, \tag{6.2}$$

where $\Omega \subseteq \mathbb{R}^n$ is open and bounded. If u is a classical solution then (6.1) is equivalent to

$$-\int_\Omega \Delta u(x)\phi(x)\,dx = \int_\Omega f(x)\phi(x)\,dx, \qquad \text{for all } \phi \in C_c^\infty(\Omega).$$

Integration by parts over Ω yields the equivalent formulation

$$\int_\Omega \nabla u(x)\cdot\nabla\phi(x)\,dx = \int_\Omega f(x)\phi(x)\,dx, \qquad \text{for all } \phi \in C_c^\infty(\Omega). \tag{6.3}$$

Note that both sides are already defined if, for instance, $\|\nabla u\|_{L^2(\Omega)} < \infty$ and $\|f\|_{L^2(\Omega)} < \infty$. We thus call u a *weak solution* of (6.1)–(6.2) if

(i) u is in the Sobolev space $H_0^1(\Omega)$, which is defined as the closure of $C_c^2(\Omega)$ under the norm $\|v\|_{L^2(\Omega)} + \|\nabla v\|_{L^2(\Omega)}$, and

(ii) u satisfies the integral equation (6.3).

One can show that the definition of $H_0^1(\Omega)$ already ensures that the boundary data, here (6.2), are assumed in an appropriate (weak) sense.

This modern treatment of partial differential equations is based on functional analytic concepts and makes heavy use of function spaces other than $C^k(\Omega)$, more precisely, Hölder spaces and Sobolev spaces (in particular, Sobolev inequalities as they are related to energies). This approach is covered in Master courses on partial differential equations and of course in the literature (see, for example, [**4**, Part 2]).

Part 2

Nonlinear partial differential equations of first order

In the previous chapters we only considered linear partial differential equations, mainly of second order. In what follows we introduce a solution technique called the *method of characteristics* for nonlinear partial differential equations of first order given as a boundary value problem, that generalizes the way we solved the linear transport equation in Chapter 2. Although we only apply this method to first order partial differential equations, it can be applied more generally to hyperbolic differential equations (remember that we could use our insights about the transport equation to solve the wave equation). In practice, characteristics can also provide valuable insight in qualitative properties of solutions, e.g., about the occurance of shock waves when characteristics cross in compressible fluids (and more generally, systems of conservation laws).

CHAPTER 7

Method of characteristics

The idea of the method of characteristics is to reduce a first-order partial differential equation to a family of simpler ordinary differential equations (called the system of characteristics) along which suitable initial data can be integrated. We have already used the linear version of this idea for the transport equation in Chapter 2.

7.1. System of characteristic equations

In what follows, we consider for $\Omega \subseteq \mathbb{R}^n$ open and $\Gamma \subseteq \partial\Omega$ a C^1 hypersurface in $\mathbb{R}^n$ the general first-order nonlinear boundary value problem

$$F(Du, u, x) = 0, \qquad \text{in } \Omega, \tag{7.1}$$

$$u = g, \qquad \text{on } \Gamma \subseteq \partial\Omega, \tag{7.2}$$

for $F\colon \mathbb{R}^n \times \mathbb{R} \times \overline{\Omega} \to \mathbb{R}$ and $g\colon \Gamma \to \mathbb{R}$ given and $u\colon \overline{\Omega} \to \mathbb{R}$ unkown. Let us assume that F and g are smooth functions. We use the differential Du and the gradient ∇u interchangeably.

Our strategy is as follows. We are going to search for *characteristics*, that is curves along which u can be integrated as a solution of a corresponding ordinary differential equation. If those curves intersect Γ, where u is known, we can thus compute u. See Figure 6.

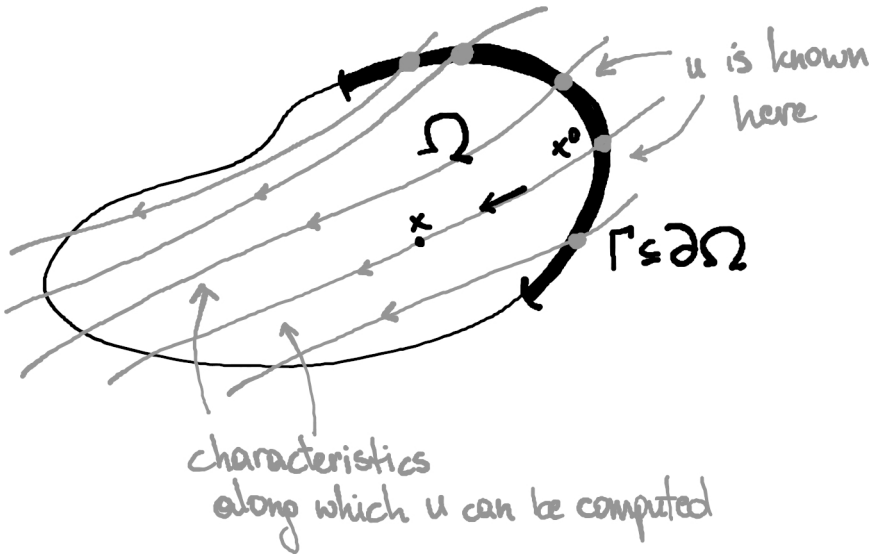

FIGURE 1. Using the method of characteristics we want to calculate u by finding a curve lying within Ω, connecting a point x in the interior with a point $x^0 \in \Gamma$ where the value of $u = g$ is known.

Let us search for suitable characteristics $x(s) = (x_1(s), \dots, x_n(s))$. Suppose u is a C^2 solution of (7.1), and define

$$z(s) := u(x(s)), \tag{7.3}$$

$$p(s) := \nabla u(x(s)). \tag{7.4}$$

Note that z is a scalar-valued while p is a vector with components $p_i(s) = u_{x_i}(x(s))$.

We want to derive x so that we can subsequently compute z and p, and thus recover the solution u. Differentiating the components $p_i = u_{x_i}$ of p we obtain that

$$\dot{p}_i(s) = \sum_{j=1}^{n} u_{x_i x_j}(x(s))\dot{x}_j(s), \qquad i = 1, \dots, n. \tag{7.5}$$

In this system, we want to get rid of the second derivatives of u as we know nothing about them. Considering $F(p, z, x)$, differentiation of (7.1) with respect to x_i yields

$$\sum_{j=1}^{n} F_{p_j}(\nabla u, u, x)u_{x_j x_i} + F_z(\nabla u, u, x)u_{x_i} + F_{x_i}(\nabla u, u, x) = 0. \tag{7.6}$$

In order to get rid of the second derivatives of u in (7.5) we use this identity to set

$$\dot{x}_j(s) = F_{p_j}(p(s), z(s), x(s)), \qquad j = 1, \dots, n. \tag{7.7}$$

Setting $x = x(s)$ in (7.6) we use z and p from (7.3)–(7.4) to obtain

$$\sum_{j=1}^{n} F_{p_j}(p(s), z(s), x(s))u_{x_i x_j}(x(s)) + F_z(p(s), z(s), x(s))p_i(s) + F_{x_i}(p(s), z(s), x(s)) = 0.$$

Plugging this and the condition (7.7) in the system (7.5) for p yields the system of differential equations

$$\dot{p}_i(s) = -F_{x_i}(p(s), z(s), x(s)) - F_z(p(s), z(s), x(s))p_i(s), \qquad i = 1, \dots, n. \tag{7.8}$$

Moreover, using (7.3)–(7.4) together with (7.7) we obtain the differential equation

$$\begin{aligned}\dot{z}(s) &= \sum_{j=1}^{n} u_{x_j}(x(s))\dot{x}_j(s) \\ &= \sum_{j=1}^{n} p_j(s)F_{p_j}(p(s), z(s), x(s)).\end{aligned}$$

That is, the partial differential equation (7.1) implies the following *system of characteristic equations*:

$$\dot{p}(s) = -\nabla_x F(p(s), z(s), x(s)) - F_z(p(s), z(s), x(s))\, p(s) \tag{7.9a}$$

$$\dot{z}(s) = \nabla_p F(p(s), z(s), x(s)) \cdot p(s) \tag{7.9b}$$

$$\dot{x}(s) = \nabla_p F(p(s), z(s), x(s)) \tag{7.9c}$$

This system consists of $2n + 1$ ordinary differential equations of first order. We call p, z and x the *characteristics* of the equation (7.1). Since x is the projection of the full characterstics (p, z, x) in $\mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n = \mathbb{R}^{2n+1}$ onto $\Omega \subseteq \mathbb{R}^n$ we call x also the *projected characteristic*.

We have shown the following result.

THEOREM 7.1. *Suppose $\Omega \subseteq \mathbb{R}^n$ is open and $u \in C^2(\Omega)$ is a solution of* (7.1) *on Ω. If x is a solution of* (7.9c)*, where $p = \nabla u \circ x$ and $z = u \circ x$, then p solves* (7.9a) *and z solves* (7.9b) *for all s with $x(s) \in \Omega$.* □

REMARK 7.2.
(i) At this point we have not yet identified correct initial data for the system (7.9) in order to derive solutions for (7.1). We will specify suitable initial data in Section 7.3.
(ii) If u is a C^2 of (7.1), then (7.9) is a closed system of ordinary differential equations for x, $z = u \circ x$ and $p = \nabla u \circ x$. The reason why no higher derivatives of u are involved is the key idea that we set $\dot{x} = \nabla_p F$ and thereby eliminiated the second derivatives of u.

7.2. Special cases and examples

We discuss several examples for the method of characteristics to compare what changes if F is linear, quasilinear of fully nonlinear.

7.2.1. F is linear. Suppose the equation is of the form

$$F(\nabla u, u, x) = b(x) \cdot \nabla u(x) + c(x)u(x) = 0. \tag{7.10}$$

In terms of the characteristics we have that $F(p, z, x) = b(x) \cdot p + c(x)\, z$, and hence that $\nabla_p F = b(x)$. This implies that (7.9c) reads

$$\dot{x}(s) = b(x(s)),$$

(independent of p and z!) and (7.9b) is

$$\dot{z}(s) = b(x(s)) \cdot p(s).$$

Since $p = \nabla u \circ x$, equation (7.10) implies that

$$\dot{z}(s) = -c(x(s))u(x(s)) = -c(x(s))z(s).$$

We thus obtain the following system of characteristics (we will later see why an equation for p is not necessary):

$$\begin{aligned} \dot{x}(s) &= b(x(s)), \\ \dot{z}(s) &= -c(x(s))z(s). \end{aligned}$$

EXAMPLE 7.3. Let $\Omega = \{x \in \mathbb{R}^2 : x_1 > 0, x_2 > 0\}$ and $\Gamma = \{x_1 > 0, x_2 = 0\} \subseteq \partial\Omega$. We consider the first order boundary value problem

$$\begin{aligned} x_1 u_{x_2} - x_2 u_{x_1} &= u, && \text{in } \Omega, \\ u &= g, && \text{on } \Gamma. \end{aligned}$$

This equation is of the form (7.10) with $b = (-x_2, x_1)$ and $c = -1$. Hence the system of characteristics reads

$$\begin{aligned} \dot{x}_1(s) = -x_2(s), \quad \dot{x}_2(s) = x_1(s), \\ \dot{z}(s) = z(s). \end{aligned}$$

The first two equations imply that $\frac{dx_1}{dx_2} = -\frac{x_2}{x_1}$ and hence by separation of variables $x_1 dx_1 = -x_2 dx_2$, and therefore $x_1^2 + x_2^2 = \text{const}$. For any $x^0 \geq 0$ and $s \in [0, \frac{\pi}{2}]$ we can therefore parametrize x by quarter circles such that

$$x_1(s) = x^0 \cos s, \quad x_2(s) = x^0 \sin s,$$
$$z(s) = z^0 e^s = g(x^0) e^s.$$

For any fixed $(x_1, x_2) \in \Omega$ we have $s > 0$, $x^0 > 0$ such that $(x_1(s), x_2(s)) = (x^0 \cos s, x^0 \sin s)$.

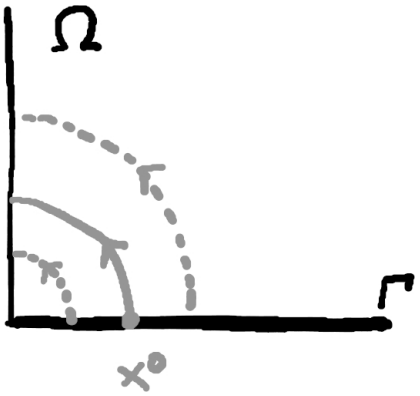

FIGURE 2. Characteristic curves for Example 7.3.

That is, $x^0 = \sqrt{x_1^2 + x_2^2}$ and $s = \arctan \frac{x_2}{x_1}$. Hence u is of the form

$$u(x_1, x_2) = u(x_1(s), x_2(s)) = z(s) = g(x^0) e^s = g(\sqrt{x_1^2 + x_2^2}) e^{\arctan(\frac{x_2}{x_1})}.$$

(Check that u indeed solves the original partial differential equation we started with!)

7.2.2. F is quasilinear. Next we consider the partial differential equation

$$F(\nabla u, u, x) = b(x, u(x)) \cdot \nabla u(x) + c(x, u(x)) = 0. \tag{7.11}$$

In this case $F(p, z, x) = b(x, z) \cdot p + c(x, z)$, and hence $\nabla_p F = b(x, z)$. We thus have that (7.9c) reads

$$\dot{x}(s) = b(x(s), z(s)),$$

and (7.9b) reads

$$\dot{z}(s) = b(x(s), z(s)) \cdot p(s) = -c(x(s), z(s)).$$

EXAMPLE 7.4. Let us consider the special semilinear partial differential equation

$$u_{x_1} + u_{x_2} = u^2, \qquad \text{in } \Omega,$$
$$u = g, \qquad \text{on } \Gamma,$$

where $\Omega = \{x_2 > 0\} \subseteq \mathbb{R}^2$ and $\Gamma = \{x_2 = 0\} = \partial\Omega$. In terms of (7.11) we have that $b = (1, 1)$ and $c = -z^2$. The system of characteristics therefore reads

$$\dot{x}_1 = 1, \quad \dot{x}_2 = 1,$$
$$\dot{z} = z^2.$$

Hence $x_1(s) = x^0 + s$ and $x_2(s) = s$. The last equation implies that $\frac{dz}{z^2} = ds$ and hence $-\frac{1}{z} = s + \text{const}$. Since $-\frac{1}{z^0} = \text{const}$ we have that

$$z(s) = \frac{1}{-s + \frac{1}{z^0}} = \frac{z^0}{1 - sz^0} = \frac{g(x^0)}{1 - sg(x^0)}$$

for $x^0 \in \mathbb{R}$ and $s \geq 0$ (as long as the denominator is nonzero). For a fixed $(x_1, x_2) \in \Omega$ we can find $s > 0$ and $x^0 \in \mathbb{R}$ such that $(x_1, x_2) = (x^0 + s, s)$, which implies that $x^0 = x_1 - x_2$

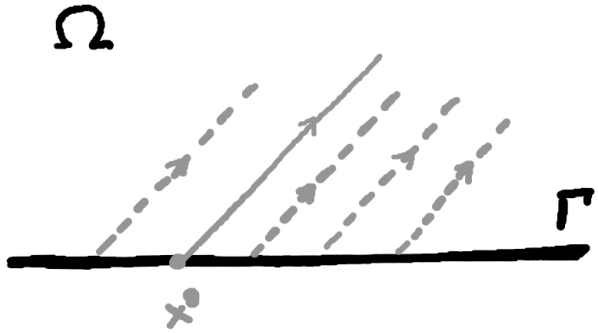

FIGURE 3. Characteristic curves for Example 7.4.

and $s = x_2$. Hence

$$u(x_1, x_2) = u(x_1(s), x_2(s)) = z(s) = \frac{g(x^0)}{1 - sg(x^0)} = \frac{g(x_1 - x_2)}{1 - x_2 g(x_1 - x_2)},$$

whenever the denominator is nonzero.

7.2.3. F is fully nonlinear. In this case we are just left with the system (7.9) of characteristics which cannot be reduced to a simpler form without knowing the particular partial differential equation.

EXAMPLE 7.5. On $\Omega = \{x_1 > 0\} \subseteq \mathbb{R}^2$ with $\Gamma = \{x_1 = 0\} = \partial\Omega$ we want to solve the problem

$$\begin{aligned} u_{x_1} u_{x_2} &= u, && \text{in } \Omega, \\ u &= x_2^2, && \text{on } \Gamma. \end{aligned}$$

Thus $F(p, z, x) = p_1 p_2 - z$, which means that (7.9) reads

$$\begin{aligned} \dot{p}_1 &= p_1, \quad \dot{p}_2 = p_2, \\ \dot{z} &= 2p_1 p_2, \\ \dot{x}_1 &= p_2, \quad \dot{x}_2 = p_1, \end{aligned}$$

which implies that[1]

$$\begin{aligned} x_1(s) &= p_2^0(e^s - 1), \quad x_2(s) = x^0 + p_1^0(e^s - 1), \\ z(s) &= z^0 + p_1^0 p_2^0(e^{2s} - 1), \\ p_1(s) &= p_1^0 e^s, \quad p_2(s) = p_2^0 e^s, \end{aligned}$$

[1]This and other details are part of Assignment 13.

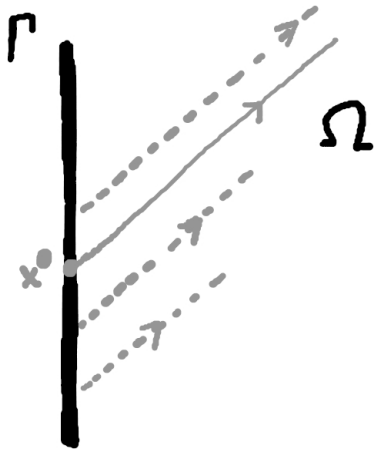

FIGURE 4. Characteristic curves for Example 7.5.

where $x^0 \in \mathbb{R}$, $s \in \mathbb{R}$ and $z^0 = (x^0)^2$. It remains to compute $p^0 = (p_1^0, p_2^0)$. One can show that $p_2^0 = 2x^0$ and $p_1^0 = \frac{x_0}{2}$. For $(x_1, x_2) \in \Omega$ we choose s and x^0 such that $(x_1, x_2) = (x_1(s), x_2(s)) = (2x^0(e^s - 1), \frac{x^0}{2}(e^s + 1))$, hence

$$x^0 = \frac{4x_2 - x_1}{4}, \quad e^s = \frac{x_1 + 4x_2}{4x_2 - x_1},$$

and therefore

$$u(x_1, x_2) = u(x_1(s), x_2(s)) = z(s) = (x^0)^2 e^{2s} = \frac{(x_1 + 4x_2)^2}{16}.$$

7.3. Boundary conditions

To solve the boundary value problem (7.1)–(7.2) we will make use the system (7.9) of characteristic equations and solve it for suitable initial data. Suitable here means, in particular, that the projected characteristic x must intersect the boundary Γ nontrivially so that $\dot{x}(0) \notin T_{x^0}\Gamma$. As in the previous examples, one can assume without loss of generality that the boundary Γ is (locally) flat. This first step showing that one can "flatten out" the boundary $\partial\Omega$ of Ω is carried out in Section 7.3.1. In the next step in Section 7.3.2 we then come up with suitable compatibility conditions, and finally we find local solutions for noncharacteristic boundary data in Section 7.4.

7.3.1. Straightening the boundary.

DEFINITION 7.6. Let $\Omega \subseteq \mathbb{R}^n$ be open and bounded, $k \in \mathbb{N}$. We say that the boundary $\partial\Omega$ is C^k if for each $x^0 \in \partial\Omega$ there is an $r > 0$ and a C^k function $\gamma\colon \mathbb{R}^{n-1} \to \mathbb{R}$ such that (possibly after a change of coordinate axes)

$$\Omega \cap B_r(x^0) = \{x \in B_r(x^0) \,|\, x_n > \gamma(x_1, \ldots, x_{n-1})\}.$$

We say that $\partial\Omega$ is C^∞ if it is C^k for all $k \in \mathbb{N}$, and call it analytic if γ is analytic.

Define

$$f(x_1, \ldots, x_n) := \gamma(x_1, \ldots, x_{n-1}) - x_n.$$

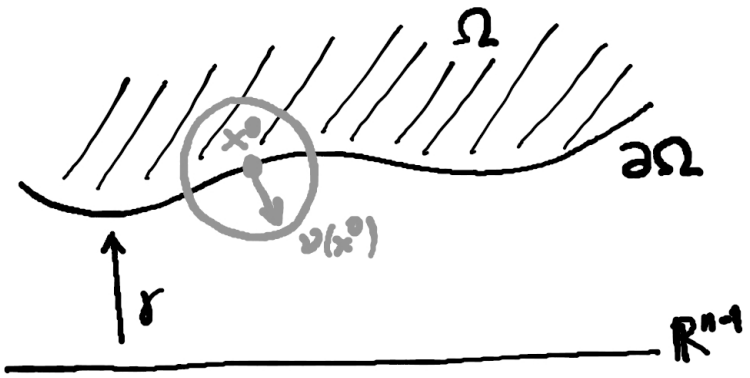

FIGURE 5. Description of boundary $\partial\Omega$.

Then locally $\partial\Omega = \{f = 0\}$. The outward-pointing unit normal vector field ν is given by

$$\nu(x) = \frac{\nabla f(x)}{|\nabla f(x)|} = \frac{1}{\sqrt{1+\gamma_{x_1}^2 + \dots \gamma_{x_n}^2}} \begin{pmatrix} \gamma_{x_1} \\ \vdots \\ \gamma_{x_n} \\ -1 \end{pmatrix}$$

(see Assignment 13).

The aim is to now make a change of coordinates near x^0 such that $\partial\Omega$ is flat. Suppose to this end that $x^0 \in \partial\Omega$ is fixed and consider r and γ as defined earlier. We define $\Phi\colon B_r(x^0) \to \mathbb{R}^n$, $y = \Phi(x)$, by

$$\begin{aligned} \Phi^i(x) &= y_i := x_i, \qquad \text{for } i = 1, \dots, n-1, \\ \Phi^n(x) &= y_n := x_n - \gamma(x_1, \dots, x_{n-1}). \end{aligned}$$

The inverse map $\Psi = \Phi^{-1}$, $x = \Psi(y)$ is given by

$$\begin{aligned} \Psi^i(y) &= x_i := y_i, \qquad \text{for } i = 1, \dots, n-1, \\ \Psi^n(y) &= y_n + \gamma(y_1, \dots, y_{n-1}). \end{aligned}$$

Since $x \in \partial\Omega$ if and only if $x_n = \gamma(x_1, \dots, x_{n-1})$, and this means that $\Phi^n(x) = 0$, we have that

$$x \in \partial\Omega \iff \Phi(x) \in \mathbb{R}^{n-1} \times \{0\},$$

i.e., Φ has the desired "flattening out" effect.

We need reformulate the boundary value problem accordingly. Suppose $u\colon \Omega \to \mathbb{R}$ is an arbitrary function. Define $V := \Phi(\Omega)$ and $v(y) := u(\Psi(y))$ for all $y \in V$, i.e.,

$$u(x) = v(\Phi(x)), \qquad x \in \Omega.$$

If u is a C^1 solution of (7.1)–(7.2), then v satisfies a related boundary value problem. Since

$$\nabla u(x) = \nabla v(\underbrace{\Phi(x)}_{=y}) D\Phi(x)$$

we obtain that

$$0 = F(\nabla u(x), u(x), x) = F(\nabla v(y) D\Phi(\Psi(y)), v(y), \Psi(y)).$$

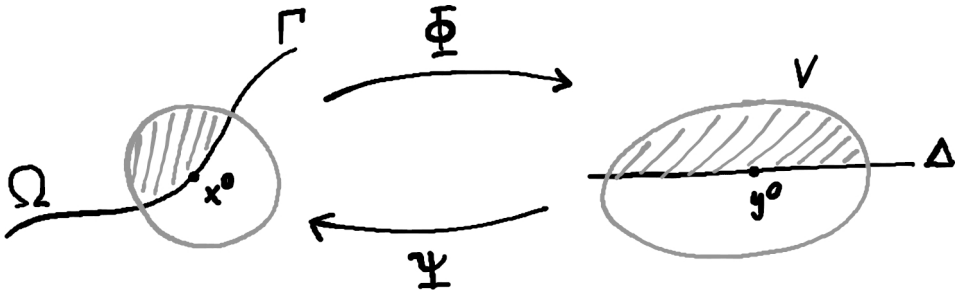

FIGURE 6. Flattening out of boundary.

This equation is of the form

$$G(\nabla v(y), v(y), y) = 0 \qquad \text{in } V.$$

Moreover, if $v = h$ on $\Delta := \Phi(\Gamma)$ with $h(y) := g(\Psi(y))$, then in the new coordinates (7.1)–(7.2) takes the form

$$\begin{aligned} G(\nabla v, v, y) &= 0, \qquad \text{in } V, \\ v &= h, \qquad \text{on } \Delta. \end{aligned}$$

This system is of the same form as (7.1)–(7.2), but with a "flat" boundary. Hence for the general theory we can simply assume that $\Gamma \subseteq \{x_n = 0\}$.

7.3.2. Compatibility conditions on boundary data. We use the assumption of the previous Section 7.3.1, that is, without loss of generality $\Gamma \subseteq \{x_n = 0\}$. Next we look for suitable initial conditions

$$p(0) = p^0, \quad z(0) = z^0, \quad x(0) = x^0,$$

for the characteristics in order to construct a solution to (7.1)–(7.2). Since $x(0) = x^0$ we must have that $z^0 = z(0) = u(x(0)) = u(x^0) = g(x^0)$, hence

$$z^0 = g(x^0). \tag{7.12}$$

Regarding the initial value $p(0) = p^0$ we first note that $u(x_1, \dots, x_{n-1}, 0) = g(x_1, \dots, x_{n-1}, 0)$ near x^0, hence

$$p_i(0) = u_{x_i}(x^0) = g_{x_i}(x^0),$$

for all $i = 1, \dots, n-1$. Moreover, (7.1) should hold, hence we obtain the following n equations for $p^0 = (p_1^0, \dots, p_n^0)$,

$$\begin{aligned} &p_i^0 = g_{x_i}(x^0), \qquad i = 1, \dots, n-1, \\ &F(p^0, z^0, x^0) = 0. \end{aligned} \tag{7.13}$$

DEFINITION 7.7. The conditions (7.12)–(7.13) are called *compatibility conditions*. A triple $(p^0, z^0, x^0) \in \mathbb{R}^{2n+1}$ which satisfies (7.12)–(7.13) is called *admissible*.

Note that, although z^0 is uniquely deteremind by g and x, the vector p^0 as specified in (7.13) may neither exist nor be unique.

7.3.3. Noncharacteristic boundary data. Suppose (p^0, z^0, x^0) is an admissible triple. Then $x(0) = x^0$, $z(0) = z^0$ and $p(0) = p^0$ are possible initial data for the system (7.9) of characteristics. Since we want to solve this system "locally" for nearby values, we must ensure that the triples remain admissible there. Hence we also want to be able to solve (7.9) for

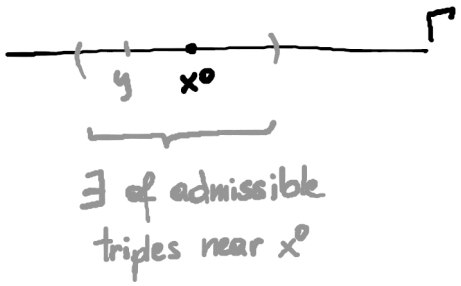

FIGURE 7. We need that initial data are not only admissible at x^0 but also nearby.

initial data

$$p(0) = q(y), \quad z(0) = g(y), \quad x(0) = y,$$

where $y = (y_1, \dots, y_{n-1}, 0) \in \Gamma$ near x^0 and $q = (q_1, \dots, q_n)$ are such that

$$q(x^0) = p^0 \tag{7.14}$$

and $(q(y), g(y), y)$ is admissible. Hence we require that

$$\begin{aligned} &q_i(y) = g_{x_i}(y), \qquad i = 1, \dots, n-1, \\ &F(q(y), g(y), y) = 0 \end{aligned} \tag{7.15}$$

holds for all $q \in \Gamma$ near x^0. The following lemma establishes a sufficient condition for the existence of such a q.

LEMMA 7.8 (Noncharacteristic boundary data). *Let (p^0, z^0, x^0) be an admissible triple. If*

$$F_{p_n}(p^0, z^0, x^0) \neq 0, \tag{7.16}$$

then there exists a unique solution q of (7.15) for all $y \in \Gamma$ near x^0. If F is C^2, then so is q.

DEFINITION 7.9. A triple (p^0, z^0, x^0) that satisfies (7.16) is called *noncharacteristic*.

PROOF. Consider $G\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$, $G(p, y) = (G^1(p, y), \dots, G^n(p, y))$, given by

$$\begin{aligned} G^i(p, y) &:= p_i - g_{x_i}(y_1, \dots, y_{n-1}, 0), \qquad i = 1, \dots, n-1, \\ G^n(p, y) &:= F(p, g(y_1, \dots, y_{n-1}, 0), y). \end{aligned}$$

Then $G(p^0, x^0) = 0$ by the compatibility condition (7.13). Moreover,

$$D_p G(p^0, x^0) = \begin{pmatrix} 1 & 0 & & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & & \ddots & 0 & 0 \\ 0 & \dots & 0 & 1 & 0 \\ F_{p_1}(p^0, z^0, x^0) & \dots & \dots & \dots & F_{p_n}(p^0, z^0, x^0) \end{pmatrix},$$

hence $\det D_p G((p^0, x^0) = F_{p_n}(p^0, z^0, x^0) \neq 0$ by assumption (7.16). The Implicit Function Theorem thus implies that the equation

$$G(p, y) = 0$$

is uniquely solvable for y near x^0, i.e., $p = q(y)$. This q is the desired C^1 solution of (7.15). □

REMARK 7.10. Note that, if Γ is not flat near x^0 then (7.16) reads

$$\nabla_p F(p^0, z^0, x^0) \cdot \nu(x^0) \neq 0,$$

where $\nu(x^0)$ denotes the outward-pointing unit normal to $\partial\Omega$ at x^0.

7.4. Local solution

Thanks to all preparations for suitable boundary conditions in Section 7.3.1 we are now in a position to solve the system (7.9) of characteristics for admissible initial data (p^0, z^0, x^0) and $x^0 \in \Gamma \subseteq \{x_n = 0\}$. By Lemma 7.8 there exists a function q such that $p^0 = q(x^0)$ and $(q(y), g(y), y)$ is admissible for all y near x^0.

For fixed $y = (y', 0) \in \Gamma \subseteq \mathbb{R}^{n-1} \times \{0\}$, with $y' = (y_1, \dots, y_{n-1})$, let us denote by

$$\begin{aligned} p(s) &= p(y', s) = p(y_1, \dots, y_{n-1}, s), \\ z(s) &= z(y', s) = z(y_1, \dots, y_{n-1}, s), \\ x(s) &= x(y', s) = x(y_1, \dots, y_{n-1}, s), \end{aligned} \tag{7.17}$$

the solution of (7.9) with initial data

$$p(0) = q(y), \quad z(0) = g(y), \quad x(0) = y. \tag{7.18}$$

LEMMA 7.11 (Local invertibility). *Suppose (p^0, z^0, x^0) is an admissible noncharacteristic triple and F is C^2. Then there exists an open interval $I \subseteq \mathbb{R}$ containing 0, a neighborhood W of $x^{0\prime}$ in $\Gamma \subseteq \mathbb{R}^{n-1}$ and a neighborhood V of x^0 in $\mathbb{R}^n$ such that for $x \in V$ there is a unique $(y', s) \in W \times I$ with*

$$x = x(y', s).$$

Moreover, the map $x \mapsto (y', s)$ is C^2.

PROOF. Since F is C^2 the solution (7.17) of the system (7.9) of characteristic equations is also C^2 (by Lemma 7.8). We have that

$$x(x^{0\prime}, 0) = x^0.$$

Hence the claim follows from the Inverse Function Theorem if $\det Dx(x^{0\prime}, 0) \neq 0$. By (7.17) and (7.18) we have that

$$x(y', 0) = (y', 0) = y \qquad \text{for } y \in \Gamma.$$

Thus for $i = 1, \dots, n-1$ we have that

$$\frac{\partial x^j}{\partial y_i}(x^{0\prime}, 0) = \begin{cases} \delta_{ij} & j = 1, \dots, n-1, \\ 0 & j = n, \end{cases}$$

since $x^i(x^{0\prime}, 0) = x_i^0$ for $1 \leq i \leq n$, in particular, $x^n(x^{0\prime}, 0) = 0$. Recall that the characteristic equation (7.9c) for x reads

$$\frac{\partial x_j}{\partial s}(x^{0\prime}, 0) = F_{p_j}(p^0, z^0, x^0) \qquad \text{for } j = 1, \dots, n,$$

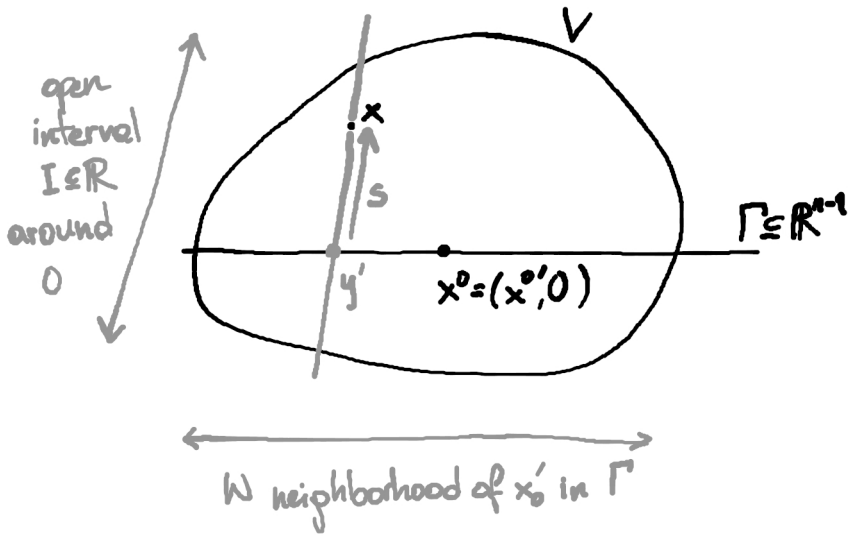

FIGURE 8. Locally around admissible noncharacteristic data we can solve the system of characteristics according to Lemma 7.11.

since $p(x^{0\prime}, 0) = q(x^0) = p^0$ by (7.18) and (7.14). Therefore, we obtain that

$$Dx(x^{0\prime}, 0) = \begin{pmatrix} 1 & \dots & 0 & F_{p_1}(p^0, z^0, x^0) \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & 1 & \vdots \\ 0 & \dots & 0 & F_{p_n}(p^0, z^0, x^0) \end{pmatrix},$$

Altogether,

$$\det Dx(x^{0\prime}, 0) = F_{p_n}(p^0, z^0, x^0) \neq 0,$$

which is nonzero by the noncharacterstic assumption 7.16. □

By Lemma 7.11, for every $x \in V$ the equation

$$x = x(y', s)$$

can be solved uniquely for $y' = y'(x)$ and $s = s(x)$. In order to prove an existence result for (nonlinear) first order partial differential equations, indeed the main result in this Section, it remains to compose x with z and p and thereby translate the ODE perspective of the characterstics back into the PDE setting. Define

$$\begin{aligned} u(x) &:= z(y'(x), s(x)), \\ p(x) &:= p(y'(x), s(x)). \end{aligned} \tag{7.19}$$

THEOREM 7.12. *Let $\Omega \subseteq \mathbb{R}^n$ be open and bounded with C^1 boundary, $\Gamma \subseteq \partial\Omega$. Let $F \in C^2(\mathbb{R}^n \times \mathbb{R} \times \overline{\Omega})$ and $g \in C^2(\Gamma)$ be given, and assume that*[2] *(p^0, z^0, x^0) are admissible and noncharacteristic boundary data, in particular, $F_{p_n}(p^0, z^0, x^0) \neq 0$. Then $u\colon V \to \mathbb{R}$ with $V \subseteq \overline{\Omega}$ as defined in (7.19) is a C^2 function that is the unique solution of the partial differential equation*

$$F(\nabla u(x), u(x), x) = 0, \qquad x \in V,$$

with initial condition

$$u(x) = g(x), \qquad x \in \Gamma \cap V,$$

and $\nabla u(x^0) = p^0$.

PROOF. Suppose that I and W are as in Lemma 7.11, $y' \in W$ and the characteristics given by (7.17)–(7.18) are

$$p(s) = p(y', s), \quad z(s) = z(y', s), \quad x(s) = x(y', s).$$

We procced in several steps.

Step 1. F **vanishes along the characteristics** $(p, z, x)(y', s)$**.** The above implies that

$$f(y', s) := F(p(y', s), z(y', s), x(y', s)) = 0, \qquad s \in I, x(y', s) \in V, \tag{7.20}$$

since

$$f(y', 0) = F(p(y', 0), z(y', 0), x(y', 0)) \overset{(7.18)}{=} F(q(y), g(y), y) \overset{(7.15)}{=} 0$$

and

$$\begin{aligned}\frac{\partial f}{\partial s}(y', s) &= \sum_{j=1}^{n} \frac{\partial F}{\partial p_j}\dot{p}_j + \frac{\partial F}{\partial z}\dot{z} + \sum_{j=1}^{n} \frac{\partial F}{\partial x_j}\dot{x}_j \\ &\overset{(7.9)}{=} \sum_{j=1}^{n} \frac{\partial F}{\partial p_j}\left(-\frac{\partial F}{\partial x_j} - \frac{\partial F}{\partial z}p_j\right) + \frac{\partial F}{\partial z}\left(\sum_{j=1}^{n} \frac{\partial F}{\partial p_j}p_j\right) + \sum_{j=1}^{n} \frac{\partial F}{\partial x_j}\frac{\partial F}{\partial p_j} = 0.\end{aligned}$$

Using (7.19) and Lemma 7.11 we thus have that

$$F(p(x), u(x), x) = 0, \qquad x \in V.$$

Step 2. $p(x) = \nabla u(x)$ **on** V**.** One first shows that for $(y', s) \in W \times I$ we have

$$\frac{\partial z}{\partial s}(y', s) = \sum_{j=1}^{n} p_j(y', s)\frac{\partial x_j}{\partial s}(y', s), \tag{7.21}$$

$$\frac{\partial z}{\partial y_i}(y', s) = \sum_{j=1}^{n} p_j(y', s)\frac{\partial x_j}{\partial y_i}(y', s), \qquad i = 1, \ldots, n-1. \tag{7.22}$$

The first condition follows immediately from the system of characteristics, more preciely, (7.9b) and (7.9c). The second equation is based on solving a linear ordinary differential

[2]These are the criteria of Lemma 7.11.

equation (see Assignment 14 for the details). Together they imply that for $j = 1, \dots, n$ that

$$\begin{aligned}
\frac{\partial u}{\partial x_j} &\overset{(7.19)}{=} \sum_{i=1}^{n-1} \frac{\partial z}{\partial y_i}\frac{\partial y_i}{\partial x_j} + \frac{\partial z}{\partial s}\frac{\partial s}{\partial x_j} \\
&\overset{(7.21),(7.22)}{=} \sum_{i=1}^{n-1} \left(\sum_{k=1}^{n} p_k \frac{\partial x_k}{\partial y_i}\right)\frac{\partial y_i}{\partial x_j} + \left(\sum_{k=1}^{n} p_k \frac{\partial x_k}{\partial s}\right)\frac{\partial s}{\partial x_j} \\
&= \sum_{k=1}^{n} p_k \left(\sum_{i=1}^{n-1} \frac{\partial x_k}{\partial y_j}\frac{\partial y_i}{\partial x_j} + \frac{\partial x_k}{\partial s}\frac{\partial s}{\partial x_j}\right) \\
&= \sum_{k=1}^{n} p_k \underbrace{\frac{\partial x_k}{\partial x_j}}_{=\delta_{kj}} = p_j,
\end{aligned}$$

where in the last two equations one has to remember that x can be written as a map $y \mapsto x(y'(x), s(x))$ and hence $\frac{\partial}{\partial x_j}(x(y'(x), s(x)) = \frac{\partial x_k}{\partial x_j} = \delta_{kj}$.

Step 3. Uniqueness. By Lemma 7.11 every $x \in V$ lies on exactly one characteristic $x(y', s)$, and the solution is uniquely determined along this characterstic by the system (7.9) of characteristics. □

REMARK 7.13. In the quasilinear case, the compatibility conditions are a system of n linear equations, and the noncharacteristic condition ensures that there exists a unique solution.

For fully nonlinear equations the compatibility conditions are a system of nonlinear equations and the noncharacteristic condition is not sufficient to ensure that a solution exists.

7.5. Some discussion of global existence

So far we have analyzed local existence of solutions near the boundary $\Gamma \subseteq \partial\Omega$. We briefly reconsider certain examples that we discussed before, and some new global phenomena. Clearly, this is just the tip of the ice berg, but it should give some insight into what remains to be understand.

7.5.1. F is linear. If

$$F(\nabla u, u, x) = b(x) \cdot \nabla u(x) + c(x)u(x) = 0, \qquad x \in \Omega,$$

then the noncharacteristic initial condition reads by Remark 7.10

$$b(x^0) \cdot \nu(x^0) \neq 0.$$

Note that this condition is *independent* of z^0 and p^0. Thus if the initial condition is given by

$$u = g \qquad \text{on } \Gamma,$$

then the equation (7.15) for q can be uniquely solved to obtain $q(y)$ for $y \in \Gamma$ near x^0 (by Lemma 7.8).

In particular, if Γ is flat then the noncharacteristic condition translates to $b_n(x^0) \neq 0$ and (7.15) is equivalent to

$$b \cdot q + c\, g = 0,$$

hence q_n can be computed from $b_n \neq 0$ and the given $q_i = g_{x_i}$, $i = 1, \dots, n-1$. Theorem 7.12 can be applied to determine the unique solution of the initial value problem in a neighborhood of x^0.

REMARK 7.14. The proof of Theorem 7.12 makes use of the full sytem of characteristics (7.9). As soon as the existence and uniqueness of the solution is known, however, it is sufficient to use the equations for x and z in Section 7.2.1 to compute it (we don't need p). Morever, the characteristics cannot intersect near x^0 since this (reduced) characteristic system is uniquely solvable.

EXAMPLE 7.15. Consider the initial value problem

$$b(x) \cdot \nabla u(x) = 0, \qquad \text{in } \Omega, \tag{7.23}$$

$$u = g, \qquad \text{on } \Gamma, \tag{7.24}$$

and suppose the solutions of the characteristic equation

$$\dot{x}(s) = b(x(s))$$

are all inwardpointing towards the origin. This happens if b inside of Ω is 0 only at the origin and $b \cdot \nu < 0$ on $\Gamma = \partial\Omega$. By Theorem 7.12 there exists a unique solution u in a neighborhood of $\partial\Omega$ and, moreover,

$$\frac{d}{ds} u(x(s)) = \nabla u(x(s)) \cdot \dot{x}(s) = \nabla u(x(s)) \cdot b(x(s)) = 0$$

by the above equations. Then

$$u(x(s)) = u(x(0)) = g(x^0)$$

for each solution with $x(0) = x^0$, hence u is *constant* along the characteristics. This shows that unless g is constant the *solution u is in general not be extendible to all of Ω.*

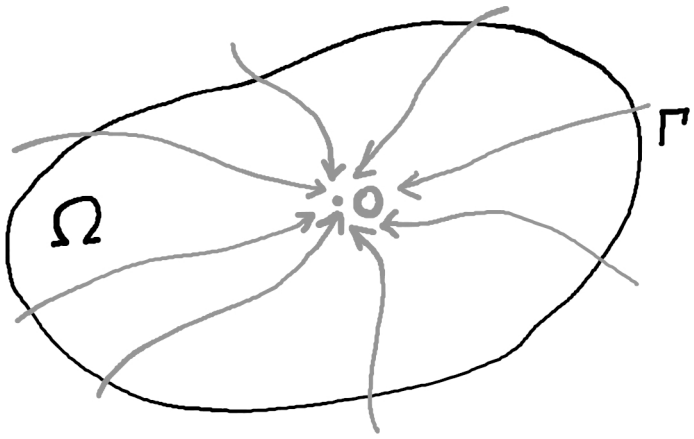

FIGURE 9. While the solution to Example 7.15 exists locally near Γ, problems even with continuity can occur at a fixpoint.

Further possibilities that can occur:

- The solution of the ordinary differential equation (except at the "end" points, here A and B) enters and exits at exactly one point of $\Gamma = \{x \in \partial\Omega : b(x) \cdot \nu(x) < 0\}$ each. Then the solution u can be extended to all of Ω (and is constant along the characteristics).

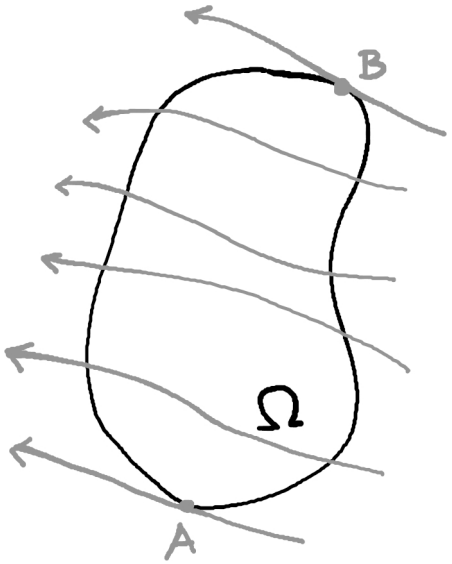

FIGURE 10. The characteristics touch the domain tangentially at A and B, thus the noncharacteristic condition is violated for Γ. At other points, when the curves start on the right boundary, they values on the left are uniquely determined (we can only specify boundary data on one side).

- If the characterstics intersect the boundary more than twice, then the solution can be constantly extended along the characteristics, but u is discontinuous if the values of the boundary points do not coincide. In the characteristic points Theorem 7.12 is not applicable.

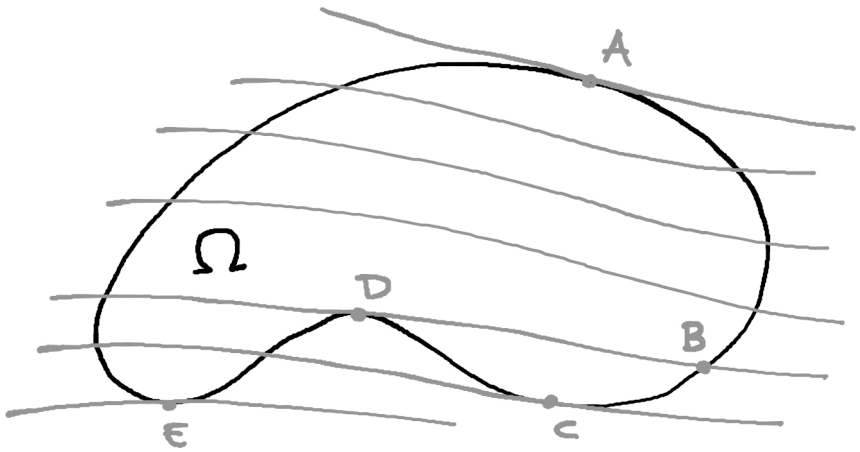

FIGURE 11. We can extend u continuously along the characterstics, but it is discontinuous unless $g(B) \neq g(D)$. Note also that the points A, C, D, E are characteristic, thus the local existence result (Theorem 7.12) cannot be used there.

7.5.2. F **is quasilinear or fully nonlinear.** These cases are much more complicated and are used, for example, for (quasilinear) conservation laws of the form

$$G(u_t, \nabla u, u, t, x) = u_t + \operatorname{div} F(u) = u_t + F'(u) \cdot \nabla u = 0$$

for $F\colon \mathbb{R} \to \mathbb{R}^n$ (and t interpreted as x_{n+1}) or the general (nonlinear) Hamilton–Jacobi equation

$$G(u_t, \nabla u, u, t, x) = u_t + H(\nabla u, x) = 0.$$

For more details see [**4**, Section 3.2.5]. In general, characteristics can cross, in which case classical solutions cease to exist (see also the Burger's equation in Assignment 14).

Bibliography

[1] Annegret Burtscher, *Optimization problems in geometry and physics*, Lecture notes, 2019.

[2] Walter Craig, *A course on partial differential equations*, Graduate Studies in Mathematics, vol. 197, American Mathematical Society, Providence, RI, 2018.

[3] Bernard Dacorogna, *Introduction to the calculus of variations*, 3rd ed., Imperial College Press, London, 2015.

[4] Lawrence C. Evans, *Partial differential equations*, 2nd ed., Graduate Studies in Mathematics, vol. 19, American Mathematical Society, Providence, RI, 2010.

[5] David Gilbarg and Neil S. Trudinger, *Elliptic partial differential equations of second order*, Classics in Mathematics, Springer-Verlag, Berlin, 2001. Reprint of the 1998 edition.

[6] Fritz John, *Partial differential equations*, 4th ed., Applied Mathematical Sciences, vol. 1, Springer-Verlag, New York, 1991.

[7] Jürgen Jost, *Postmodern analysis*, 3rd ed., Universitext, Springer-Verlag, Berlin, 2005.

[8] Hans Ringström, *The Cauchy problem in general relativity*, ESI Lectures in Mathematics and Physics, European Mathematical Society (EMS), Zürich, 2009.

[9] Walter A. Strauss, *Partial differential equations*, 2nd ed., John Wiley & Sons, Ltd., Chichester, 2008. An introduction.

[10] András Vasy, *Partial differential equations*: *An accessible route through theory and applications*, Graduate Studies in Mathematics, vol. 169, American Mathematical Society, Providence, RI, 2015.

[11] Wolfgang Walter, *Ordinary differential equations*, Graduate Texts in Mathematics, vol. 182, Springer-Verlag, New York, 1998.

Index