Essentially nothing found here is original except for a few mistakes and misprints here and there. These lecture notes are based on material from the following books: L. Evans “Partial Differential Equations”, Y. Pinchover and J. Rubinstein “An Introduction to Partial Differential Equations”

1 Some important PDEs

This material is based on Evans Chapter 2, often following it literally.

1.1 A triviality

The very simplest partial differential equation is, probably,

$$\frac{\partial u}{\partial t} = 0,$$

(1.1)

for a function $u(t, x)$ that depends on two variables (but $x$ does not appear in (1.1), of course). However, there is almost nothing interesting to say about (1.1) apart from observing that $u(t, x)$ is an arbitrary function of $x$ and does not depend on $t$. Still, we can draw a couple of conclusions from (1.1), that we will soon generalize to somewhat more interesting equations, before abandoning (1.1) for good. First, (1.1) without any extra conditions, has infinitely many solutions – a function of the form $u(t, x) = g(x)$ solves (1.1) for an arbitrary function $g(x)$. Therefore, to get a unique solution we need to impose some condition apart from (1.1) on the function $u(t, x)$. One such constraint is to prescribe the initial data:

$$u(0, x) = f(x).$$

(1.2)

It is easy to see that solution of (1.1)-(1.2) is unique: $u(t, x) = f(x)$. Another possibility is to add a boundary condition to (1.1) (usually if $t$ and $x$ play the role of “time” and “space”, respectively, then we talk about “initial conditions” if the function is prescribed at $t = 0$, and about “boundary conditions” if the function is prescribed at $x = 0$, or some other fixed point in $x$):

$$u(t, 0) = g(t).$$

(1.3)
However, as one can check immediately, (1.1) with the boundary condition (1.3) has a solution only if \( g(t) = g_0 \), that is, \( g(t) \) is a constant independent of \( t \). Moreover, in that case solution is not unique – any function \( u(t, x) = r(x) \) with \( r(0) = g_0 \) would solve (1.1), (1.3) provided that \( g(t) \) is a constant. Another example of a “good” curve to impose some condition is the line \( x = t \): if we add the constraint

\[
u(t, t) = v(t),
\]

with a prescribed function \( v(t) \) to (1.1) then solution is unique: \( u(t, x) = v(x) \). A quick lesson to remember is that adding a boundary, or initial, or mixed condition even to an absurdly simple PDE like (1.1) may lead to existence of a unique solution, no solutions, or infinitely many solutions.

Another, slightly more obscure lesson to learn is that whether we add (1.2) or (1.4) as an extra condition to (1.1), we have \( \|u(t, x)\| \leq \max_y \|f(y)\| \) and \( \|u(t, x)\| \leq \max_s \|v(s)\| \), respectively, for all \( t \) and \( x \) – solution everywhere does not exceed its maximum on the surface where it is prescribed. Moreover, this equation preserves positivity: if \( f(x) \geq 0 \) then \( u(t, x) \geq 0 \) for all \( t \) and \( x \), as well. These properties are versions of the maximum principle and we will encounter them soon in much less trivial settings.

### 1.2 The linear transport equation

The next simplest PDE after the triviality of (1.1) is the linear transport equation

\[
\frac{\partial \phi}{\partial t} + u \cdot \nabla_x \phi = 0, \quad x \in \mathbb{R}^n, \ t \geq 0.
\]

(1.5)

Here \( \phi(t, x) \) is the unknown function, and \( u = (u_1, \ldots, u_n) \) is a constant vector in \( \mathbb{R}^n \), known as a “drift” – the terminology will become clear later. The notation \( u \cdot v \), with \( u, v \in \mathbb{R}^n \), denotes the standard inner product in \( \mathbb{R}^n \):

\[
u \cdot v = u_1v_1 + \ldots + u_nv_n,
\]

so that (1.5) is a short-hand for

\[
\frac{\partial \phi}{\partial t} + \sum_{j=1}^{n} u_j \frac{\partial \phi}{\partial x_j} = 0, \quad x \in \mathbb{R}^n, \ t \geq 0.
\]

**The initial value problem**

Which functions solve (1.5)? Let us first look at this from a bureaucratic point of view: (1.5) expresses the fact that the directional derivative of \( \phi(t, x) \), understood as a function of \((n+1)\)-variables \((t, x)\), vanishes in the direction \((1, u)\). This means that the function

\[
z(s) = \phi(t + s, x + su)
\]

is constant in \( s \), for any \( x \) and \( t \) fixed, and, indeed, using (1.5) we see that if \( \phi(t, x) \) solves (1.5) then

\[
\frac{dz}{ds} = \frac{\partial \phi}{\partial t}(t + s, x + su) + u \cdot \nabla \phi(t + s, x + su) = 0.
\]
Therefore, if we take any point \((t, x) \in \mathbb{R}^{n+1}\) and draw a line \(L_{t,x} = \{(t + s, x + us), \ s \in \mathbb{R}\}\) (known as a characteristic) in \(\mathbb{R}^{n+1}\), then the function \(\phi(t, x)\) is constant along \(L_{t,x}\). This gives a hint of what kind of initial or boundary value problems we can solve for (1.5). Consider the initial value problem

\[
\frac{\partial \phi}{\partial t} + u \cdot \nabla x \phi = 0, \quad x \in \mathbb{R}^n, \quad t \geq 0.
\]

(1.6)

\[\phi(0, x) = g(x), \quad \text{for } x \in \mathbb{R}^n,\]

with a prescribed function \(g(x)\). Then for a given \(t \geq 0\) and \(x \in \mathbb{R}^n\) we look at the line \(L_{t,x}\).

It intersects the hyper-plane \(\{t = 0\}\) where the solution is prescribed at the point \((0, x - ut)\) (we take \(s = -t\) in the definition of \(L_{t,x}\)). Therefore, \(\phi(t, x) = \phi(0, x - ut) = g(x - ut)\). This is the unique solution to the initial value problem (1.6).

Our choice of the initial conditions as the additional constraint is not arbitrary – the variable \(t\) plays the role of time, and physically, the initial value problem with the data prescribed at \(t = 0\) comes up most often. However, for the sake of completeness and to see what can go wrong, consider the boundary value problem prescribed along the plane \(\{x_1 = 0\}\):

\[
\frac{\partial \phi}{\partial t} + u \cdot \nabla x \phi = 0, \quad x \in \mathbb{R}^n, \quad t \geq 0.
\]

(1.7)

\[\phi(t, x_1 = 0, x_2, \ldots, x_n) = f(t, x_2, \ldots, x_n), \quad \text{for all } t \in \mathbb{R}, \quad \text{and } (x_2, \ldots, x_n) \in \mathbb{R}^{n-1},\]

with a given function \(f\). Again, given a point \((t, x)\), the line \(L_{t,x}\) intersects the plane \(\{x_1 = 0\}\) at the point corresponding to \(s = -x_1/u_1\): the intersection point is

\[(t - \frac{x_1}{u_1}, 0, x_2 - \frac{x_1}{u_1}u_2, \ldots, x_n - \frac{x_1}{u_1}u_n).\]

It follows that the solution of (1.7) is

\[\phi(t, x) = f(t - \frac{x_1}{u_1}, x_2 - \frac{x_1}{u_1}u_2, \ldots, x_n - \frac{x_1}{u_1}u_n).\]

It is unique provided that \(u_1 \neq 0\). If \(u_1 = 0\) then the plane \(\{x_1 = 0\}\) is parallel to the lines \(L_{t,x}\), thus the value of the solution at \((t, x)\) is not determined by the data prescribed along that plane. This is a general lesson: the data should be prescribed along surfaces that are not tangent to characteristics, or one would get non-uniqueness and non-existence (depending on the data prescribed on such surfaces). We will discuss this in a greater detail later.

Another point to make is that the maximum principle still holds: solution of the initial value problem (1.6) satisfies

\[|\phi(t, x)| \leq \sup_y |g(y)|,\]

(1.8)

and, in addition, it preserves positivity: if \(g(x) \geq 0\), then \(\phi(t, x) \geq 0\).

Now, let us discuss the above properties from a slightly different, more physical point of view, with a clear separation of what is space and what is time: consider a family of particles that move along trajectories \(X(s; t, x)\) which solve the ODE

\[
\frac{dX(s; t, x)}{ds} = u, \quad X(s = t; t, x) = x,
\]

(1.9)
that is, we are given a “starting” time \( t \in \mathbb{R} \) and a “starting” position \( x \in \mathbb{R}^n \), and we look at the trajectory parametrized by the parameter \( s \in \mathbb{R} \) that at the time \( s = t \) passes through the point \( x \). Note that the trajectories \( X(s; t, x) \) lie in \( \mathbb{R}^n \), unlike the lines \( L_{t,x} \) that lived in \( \mathbb{R}^{n+1} \), the space that also involved \( t \). Solutions of (1.9) are explicit:

\[
X(s; t, x) = x + u(s - t). \tag{1.10}
\]

Define then a function

\[
\psi(t, x) = g(X(0; t, x)) = g(x - ut).
\]

This is the value of the function \( g \) at the point where the trajectory passes at \( s = 0 \). For a small time increment \( \Delta t \) we have

\[
\psi(t + \Delta t, x + (\Delta t)u) = \psi(t, x), \tag{1.11}
\]

simply because

\[
X(0; t + \Delta t, x + (\Delta t)u) = x + (\Delta t)u - u(t + \Delta t) = x - ut = X(0; t, x).
\]

On the other hand, (1.11) implies immediately that \( \psi(t, x) \) solves

\[
\frac{\partial \psi}{\partial t} + u \cdot \nabla x \psi = 0, \tag{1.12}
\]

with the initial data \( \psi(0, x) = g(X; 0, x) = g(x) \), which is (1.6). The maximum principle (1.8) is an immediate consequence of this interpretation, as is preservation of positivity – the range of \( \psi(t, x) \) lies inside the range of \( g(x) \). Here, this “dynamical systems” approach is essentially indistinguishable from what we did before, but it will become very useful when we turn to the elliptic and parabolic equations. In the case of first order equations it is known as the method of characteristics, while for the second order equations it relates to random processes, Brownian motion and diffusions.

**Variable drift**

Let us now consider the transport equation with a drift \( u(x) = (u_1(x), u_2(x), \ldots, u_n(x)) \) that varies in space:

\[
\frac{\partial \phi}{\partial t} + u(x) \cdot \nabla \phi = 0, \quad x \in \mathbb{R}^n, \quad t \geq 0, \tag{1.13}
\]

with a prescribed initial data \( \phi(0, x) = f(x) \). Recall that when \( u(x) \) was constant in space, we used the trajectories \( X(s; t, x) \) (that happened to be straight lines) to construct the solution. Let us look for the analog of these lines in the case that \( u(x) \) varies in space: let \( X(s; t, x) \) be a curve in \( \mathbb{R}^n \) parametrized by \( s \in \mathbb{R} \) (\( t \) and \( x \) are fixed here) such that \( X(s = t; t, x) = x \) and define the function \( z(s) = \phi(s, X(s; t, x)) \), which is the restriction of the function \( \phi(t, x) \) to our curve. We compute

\[
\frac{dz}{ds} = \frac{\partial \phi(s, X(s; t, x))}{\partial s} + \sum_{j=1}^{n} \frac{\partial \phi(s, X(s; t, x))}{\partial X_j} \frac{dX_j}{ds}
\]

\[
= \frac{\partial \phi(s, X(s; t, x))}{\partial s} + \frac{dX(s; t, x)}{ds} \cdot \nabla \phi(s, X(s; t, x)).
\]
Therefore, we have
\[ \frac{dz}{ds} = 0, \]
or, equivalently, the function \( z(s) \) is constant along the curve \( X(s; t, x) \), if we choose \( X(s; t, x) \) to be the solution of the system of ODE's
\[ \frac{dX}{ds} = u(X). \tag{1.14} \]
If we supplement (1.14) by an initial condition at \( s = t \):
\[ X(s = t; t, x) = x, \tag{1.15} \]
then \( z(t) = \phi(t, x) \). Moreover, if we do choose the curve \( X(s; t, x) \) as in (1.14), so that \( z(s) \) is constant in \( s \), we would have \( z(t) = z(0) \), which, equivalently, means
\[ \phi(t, x) = \phi(t, X(t; t, x)) = \phi(0, X(0; t, x)) = f(X(0; t, x)). \tag{1.16} \]
Therefore, solution of the initial value problem for (1.13) can be found as follows: fix \( t \in \mathbb{R} \) and \( x \in \mathbb{R}^n \) and solve the ODE system (1.14)-(1.15) to find \( X(0; t, x) \). Then use (1.16) to find the value of \( \phi(t, x) \).

It follows, in particular, from (1.16) that the maximum principle still applies:
\[ \sup_{x \in \mathbb{R}^n} \phi(t, x) \leq \sup_{x \in \mathbb{R}^n} f(x), \]
and positivity is preserved as well: \( \phi(t, x) \geq 0 \) for all \( x \in \mathbb{R}^n \) if \( f(x) \geq 0 \) for all \( x \in \mathbb{R}^n \).

**Example 1.** Consider the problem
\[ \frac{\partial \phi}{\partial t} + x \frac{\partial \phi}{\partial x} = 0, \quad t \in \mathbb{R}, \quad x \in \mathbb{R}, \]
with the initial data \( \phi(0, x) = f(x) \). The corresponding ODE is
\[ \frac{dX}{ds} = X, \quad X(t) = x, \]
and its solution is \( X(s) = xe^{s-t} \). Therefore, \( X(0) = xe^{-t} \), hence \( \phi(t, x) = f(xe^{-t}) \) - solution at a positive time \( t > 0 \) has the same profile as \( f(x) \) but is stretched out in the \( x \)-direction by the factor \( e^t \).

**Example 2.** Consider the problem with the opposite sign of the drift
\[ \frac{\partial \phi}{\partial t} - x \frac{\partial \phi}{\partial x} = 0, \quad t \in \mathbb{R}, \quad x \in \mathbb{R}, \]
with the initial data \( \phi(0, x) = f(x) \). The corresponding ODE is
\[ \frac{dX}{ds} = -X, \quad X(t) = x, \]
and its solution is \( X(s) = xe^{-s+t} \). Therefore, \( X(0) = xe^t \), hence \( \phi(t, x) = f(xe^t) \) - solution at a positive time \( t > 0 \) has the same profile as \( f(x) \) but is squished in the \( x \)-direction by the factor \( e^t \).
The inhomogeneous problem

Consider now the initial value problem with a force:

\[ \frac{\partial \phi}{\partial t} + u \cdot \nabla_x \phi = f(t,x), \quad x \in \mathbb{R}^n, \quad t \geq 0. \]  
\[ \phi(0,x) = g(x), \quad \text{for } x \in \mathbb{R}^n, \]  

with prescribed initial data \( g(x) \) and force \( f(t,x) \). Here we assume again, for simplicity, that the vector \( u \) is constant and does not depend on \( x \). Consider what happens now with the function \( z(s) = \phi(t+s,x+us) \):

\[ \frac{dz}{ds} = \frac{\partial \phi}{\partial t}(t+s,x+su) + u \cdot \nabla \phi(t+s,x+su) = f(t+s,x+su). \]

Choosing again \( s = -t \) gives

\[ z(0) - z(-t) = \int_{-t}^{0} \int_{-t}^{0} f(t+\tau,x+\tau u)d\tau = \int_{0}^{t} f(\tau,x-(t-\tau)u)d\tau. \]

This is nothing but

\[ \phi(t,x) = \phi(0,x-ut) + \int_{0}^{t} f(\tau,x-(t-\tau)u)d\tau, \]  
\[ \phi(t,x) = g(x-ut) + \int_{0}^{t} \psi(\tau,x;\tau)d\tau, \]  

or

In order to interpret the formula (1.19) let us define \( \psi(t,x;\tau) = f(\tau,x-(t-\tau)u) \). This function satisfies the following initial value problem starting at time \( \tau \):

\[ \frac{\partial \psi}{\partial t} + u \cdot \nabla_x \psi = 0, \quad x \in \mathbb{R}^n, \quad t \geq \tau. \]  
\[ \psi(t = \tau,x;\tau) = f(\tau,x), \quad \text{for } x \in \mathbb{R}^n, \]

with \( \tau \) playing the role of a parameter, \( 0 \leq \tau \leq t \). We may rephrase (1.19) now as

\[ \phi(t,x) = g(x-ut) + \int_{0}^{t} \psi(\tau,x;\tau)d\tau, \]

that is, we decomposed solution of an initial value problem with a force as a sum of the solution of the initial value problem with zero force, and a time integral of solutions of the initial value problems with zero force, and with the initial data at an intermediate time \( \tau \) given by the force \( f(\tau,x) \). Such decompositions are known as the Duhamel principle, and appear in all sorts of linear time-dependent problems we will encounter later.
1.3 The Laplace and Poisson equations

One of the most frequently encountered PDEs are the Laplace equation

\[ \Delta \phi = 0, \quad (1.22) \]

and its inhomogeneous counterpart

\[ -\Delta \psi = f(x), \quad (1.23) \]

known as the Poisson equation. Recall that the Laplacian is

\[ \Delta \phi = \frac{\partial^2 \phi}{\partial x_1^2} + \ldots + \frac{\partial^2 \phi}{\partial x_n^2}. \]

Why do we put a minus in the left side of (1.23)? We will see that then (1.23) preserves positivity under many of the common boundary conditions, that is, \( \psi(x) \geq 0 \) if \( f(x) \geq 0 \) for all \( x \) in the domain \( U \) where the Poisson equation is posed. Without the minus sign, the function \( \psi \) would be positive if \( f \) is negative, which would be inconvenient in some qualitative considerations.

The Laplace equation is usually derived as follows. Consider a physical quantity \( \Phi \) such as mass or heat whose total amount in any given volume \( U \subset \mathbb{R}^n \) does not change in time – there are no sources or sinks anywhere, and everything is in an equilibrium. The conservation of \( \Phi \) can be expressed as

\[ \int_{\partial U} (F \cdot \nu) dS = 0. \quad (1.24) \]

Here \( F \) is the flux of \( \Phi \), and \( \nu \) is the outward normal to \( U \). The flux is often described by the Fourier law:

\[ F = -k \nabla \Phi. \quad (1.25) \]

The constant \( k > 0 \) is usually called the diffusivity. The meaning of (1.25) is clear: heat (or mass) flows from the regions of high \( \Phi \) to regions of low \( \Phi \). Using the Fourier law in (1.24) gives

\[ \int_{\partial U} (k \nabla \Phi \cdot \nu) dS = 0. \quad (1.26) \]

Using Green’s formula, this can be equivalently written as the volume integral

\[ \int_U \nabla \cdot (k \nabla \Phi) dx = 0. \quad (1.27) \]

We use here the notation \( \nabla \cdot g \) for the divergence of a vector-valued function \( g = (g_1, \ldots, g_n) \):

\[ \nabla \cdot g = \frac{\partial g_1}{\partial x_1} + \ldots + \frac{\partial g_n}{\partial x_n}. \]

As \( U \) is an arbitrary volume element, we conclude from (1.27) that we have

\[ \nabla \cdot (k \nabla \Phi) = 0. \quad (1.28) \]

When \( k = \text{const} \), this equation reduces to the Laplace equation (1.22). Otherwise, if \( k(x) \) varies in space, we get an inhomogeneous equation

\[ \nabla \cdot (k(x) \nabla \Phi) = 0, \quad (1.29) \]

which, as we will see, has many similar properties to the Laplace equation.
A probabilistic connection interlude

Another nice way to understand how the Laplace equation comes about, as well as many of its properties is in terms of the discrete equations. For the sake of simplicity of notation, we describe it in two dimensions. Let \( U \) be a bounded sub-domain of the two-dimensional square lattice \( \mathbb{Z}^2 \), and let \( u(x) \) solve the difference equation

\[
    u(x + 1, y) + u(x - 1, y) + u(x, y + 1) + u(x, y - 1) - 4u(x, y) = 0,
\]

which is a discrete analog of (1.22). We also impose the boundary condition \( u(x, y) = g(x, y) \) on the boundary \( \partial U \). Here \( g(x, y) \) is a prescribed non-negative function, which is positive somewhere.

We claim that the solution of this problem has the following probabilistic interpretation. Let \( (X(t), Y(t)) \) be the standard random walk on the lattice \( \mathbb{Z}^2 \) – the probability to go up down, left or right is equal to 1/4, and let it start at the point \( (x, y) \): \( X(0) = x, Y(0) = y \). Let \( (\bar{x}, \bar{y}) \) be the first point where \( (X(t), Y(t)) \) reaches the boundary \( \partial U \) of the domain. The point \( (\bar{x}, \bar{y}) \) is, of course, random. The beautiful observation is that the function \( v(x, y) = \mathbb{E}(g(\bar{x}, \bar{y})) \) gives a solution of (1.30), connecting this discrete equation to the random walk. Why? First, it is immediate that if the starting point \( (x, y) \) is on the boundary of \( U \) then, of course, the exit point from \( U \) is simply the starting point: \( \bar{x} = x \) and \( \bar{y} = y \), so \( v(x, y) = g(x, y) \) in that case. On the other hand, if \( (x, y) \) is inside \( U \) then the probabilities for the random walk to go up, down, left right are all equal to 1/4, meaning that \( v(x, y) \) can be written as

\[
    v(x, y) = \frac{1}{4}(v(x + 1, y) + v(x - 1, y) + v(x, y + 1) + v(x, y - 1)).
\]

This identity simply uses the definition of the random walk, the definition of \( v(x, y) \) and very elementary probability considerations.

Now, if we let the mesh size be not 1 but \( h > 0 \), the discrete equation (1.30) becomes

\[
    u(x + h, y) + u(x - h, y) + u(x, y + h) + u(x, y - h) - 4u(x, y) = 0.
\]

If we now expand the function \( u(x) \) in the Taylor series:

\[
    u(x + h, y) = u(x, y) + h \frac{\partial u(x, y)}{\partial x} + \frac{h^2}{2} \frac{\partial^2 u(x, y)}{\partial x^2} + ...
\]

and similarly for the other three terms, and let \( h \downarrow 0 \), the discrete equation (1.31) becomes the Laplace equation:

\[
    u_{xx} + u_{yy} = 0,
\]

while the random walk becomes the Brownian motion. More precisely, solution of the Laplace equation (1.22) in \( n \) dimensions has the following probabilistic interpretation: let \( U \) be a domain in \( \mathbb{R}^n \) and let \( g(x) \) be a continuous function on the boundary \( \partial U \). Consider a Brownian motion \( B(t; x) \) that starts at a point \( x \in U \) and let \( \bar{x} \) be a (random) point where \( B(t; x) \) hits the boundary \( \partial U \) for the first time. Then solution of the Laplace equation

\[
    \Delta u = 0 \text{ in } U \quad (1.33)
\]
with the boundary condition $u(x) = g(x)$ for $x \in \partial U$, is $u(x) = \mathbb{E}(g(\bar{x}))$. The reader unfamiliar with the notion of the Brownian motion should not worry – we will not rely on this connection in any way other than provide some intuition and motivation.

From the heuristic point of view, now, if $g(x)$ is continuous and non-negative everywhere on $\partial U$, and positive at some point $x_0 \in \partial U$ (and thus in a neighborhood $V$ of $x_0$ as well) then with a positive probability the exit point $\bar{x}$ lies in $V$, so that we have $g(\bar{x}) > 0$, which means that $u(x) = \mathbb{E}(g(\bar{x})) > 0$ as well.

The maximum principle is also a simple consequence of the probabilistic interpretation: it is easy to see that $\mathbb{E}(g(\bar{x})) \leq \sup_{z \in \partial U} g(z)$ – expected value of a function can not exceed its maximum.

**Radial solutions of the Laplace equation**

Let us first look for explicit radially symmetric solutions to the Laplace equation – they depend only on $r = |x| = \sqrt{x_1^2 + \ldots + x_n^2}$. If $\phi(x) = v(r)$ depends only on the variable $r$, then

$$\frac{\partial \phi}{\partial x_i} = v'(r) \frac{\partial r}{\partial x_i} = v'(r) \frac{x_i}{r}, \quad 1 \leq i \leq n,$$

and

$$\frac{\partial^2 \phi}{\partial x_i^2} = \frac{\partial}{\partial x_i} \left( v'(r) \frac{x_i}{r} \right) = v''(r) \frac{x_i^2}{r^2} + \frac{v'(r)}{r} - v'(r) \frac{x_i^2}{r^3}.$$

Summing over $i$, from $i = 1$ to $i = n$, keeping in mind that $\sum_{i=1}^n x_i^2 = r^2$, gives:

$$\Delta \phi = v''(r) + n v'(r) - v'(r) \frac{r^2}{r^3} = v''(r) + \frac{n-1}{r} v'(r).$$

Therefore, for a radial function $\phi(x) = v(r)$ to satisfy the Laplace equation, the function $v(r)$ has to be the solution of the ODE

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

Dividing by $v'(r)$ ($v(r) = \text{const}$ would be a not very exciting solution) gives

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

Therefore, if $n = 1$ then $v(r) = cr$, that is,

$$u(x) = c|x|.$$

When $n \geq 2$ we get

$$\ln(v'(r)) = -(n-1) \ln r + C,$$

so that when $n = 2$, we get $v'(r) = C/r$, and

$$v(r) = C \ln r + B,$$
with some constants $C$ and $B$. Finally, for $n \geq 3$ we obtain

$$v'(r) = \frac{C}{r^{n-1}},$$

whence $v(r) = -C/r^{n-2} + B$.

You may notice immediately that in all three cases, $n = 1$, $n = 2$ and $n \geq 3$ the radial solutions that we have obtained above are not twice differentiable at $r = 0$, and, moreover, for $n \geq 2$ they are not even bounded. So, do they satisfy the Laplace equation? They certainly do away from $x = 0$ but what happens there? In order to appreciate this point, observe that if a smooth function $K(x)$ satisfies

$$\Delta K(x) = 0,$$

for all $x \in \mathbb{R}$ and a function $f(x)$ is smooth and vanishes outside of a bounded set, then the function

$$v(x) = \int K(x - y)f(y)dy$$

also satisfies $\Delta v(x) = 0$.

Let us now see what happens if the smooth kernel $K(x)$ is replaced by one of the singular solutions we have just constructed. Consider for simplicity the one-dimensional case. Let us define a function

$$\phi(x) = \int_{-\infty}^{\infty} |x - y|f(y)dy, \quad (1.34)$$

with a smooth function $f(y)$ which vanishes outside of some interval $[-L, L]$ so that all differentiations under the integral sign in the following computation are justified:

$$\phi'(x) = \frac{d}{dx} \int_{-\infty}^{\infty} |x - y|f(y)dy = \frac{d}{dx} \int_{-\infty}^{x} (x - y)f(y)dy + \frac{d}{dx} \int_{x}^{\infty} (y - x)f(y)dy$$

$$= \int_{-\infty}^{x} f(y)dy - \int_{x}^{\infty} f(y)dy.$$  

It follows that

$$\phi''(x) = 2f(x). \quad (1.35)$$

Therefore, the function $\phi(x)$ is not a solution of the Laplace equation but rather of the Poisson equation with the right side given by the function $(-2f(x))$. In order to get rid of the pesky $(-2)$ factor we introduce

$$\Phi_1(x) = -\frac{1}{2}|x|, \quad x \in \mathbb{R} \quad (1.36)$$

and observe that for any "nice" function $f$ the function

$$\phi(x) = \int_{-\infty}^{\infty} \Phi_1(x - y)f(y)dy, \quad x \in \mathbb{R} \quad (1.37)$$

is the solution of the Poisson equation

$$-\phi''(x) = f(x), \quad x \in \mathbb{R}. \quad (1.38)$$
1.3.1 The fundamental solution of the Laplace equation

The property that any function of the form (1.37) is the solution of the Poisson equation (1.38) means that the function \( \Phi_1(x) = -|x|/2 \) is the fundamental solution for the Laplace equation in one dimension. In higher dimensions, the fundamental solutions of the Laplace equation are given by

\[
\Phi(x) = \frac{1}{2\pi} \log |x|, \quad n = 2, \tag{1.39}
\]

and

\[
\Phi(x) = \frac{1}{n(n-2)\alpha(n)} \frac{1}{|x|^{n-2}}, \quad n \geq 3. \tag{1.40}
\]

Here \( \alpha(n) \) is the volume of the unit sphere in \( n \) dimensions. When we say that \( \Phi(x) \) is the fundamental solution of the Laplace equation, we mean the following.

**Theorem 1.1** Let \( f \in C^2_c(\mathbb{R}^n) \) (that is, \( f \) is twice continuously differentiable and has compact support), \( n \geq 2 \), and set

\[
\phi(x) = \int_{\mathbb{R}^n} \Phi(x - y)f(y)dy, \tag{1.41}
\]

with \( \Phi(x) \) given by (1.39) and (1.40) for \( n = 2 \) and \( n \geq 3 \), respectively. Then \( \phi(x) \) is twice continuously differentiable and satisfies the Poisson equation

\[
-\Delta \phi = f(x), \quad x \in \mathbb{R}^n. \tag{1.42}
\]

Before we go into the proof, let us recall Lebesgue dominated convergence theorem from real analysis.

**Theorem 1.2 (Lebesgue Dominated Convergence Theorem)** Let \( g_k(x) \) be a sequence of functions such that \( g_k(x) \to g(x) \) as \( k \to \infty \) for all \( x \in \mathbb{R}^n \) and assume that there exists a function \( Q(x) \) such that

\[
\int_{\mathbb{R}^n} |Q(x)|dx < +\infty,
\]

and \( |g_k(x)| \leq Q(x) \) for all \( k \) and all \( x \in \mathbb{R}^n \). Then we have

\[
\lim_{n \to +\infty} \int_{\mathbb{R}^n} g_n(x)dx = \int_{\mathbb{R}^n} g(x)dx.
\]

We will not prove this theorem here, its proof can be found in essentially any textbook on measure theory and real analysis.

**Proof of Theorem 1.1. Step 1. Differentiability of \( \phi(x) \).** Let us first show that \( \phi(x) \) defined by (1.41) is differentiable. Assume that \( f(x) \) vanishes outside of a ball of radius \( R \). Note that

\[
\phi(x) = \int_{\mathbb{R}^n} \Phi(x - y)f(y)dy = \int_{\mathbb{R}^n} \Phi(y)f(x - y)dy = \int_{B_R(x)} \Phi(y)f(x - y)dy,
\]

where \( B_R(x) \) is the ball of radius \( R \) centered at the point \( x \). Therefore, for \( |h| < 1 \) we have

\[
\frac{\phi(x + he_i) - \phi(x)}{h} = \int_{B_{R+1}(x)} \Phi(y)\frac{f(x + he_i - y) - f(x - y)}{h}dy, \tag{1.43}
\]
where \(e_i = (0, \ldots, 1, 0, \ldots, 0)\) is the unit vector in the direction of \(x_i\). However, we have
\[
g_h(y) = \frac{f(x + he_i - y) - f(x - y)}{h} \to \frac{\partial f(x - y)}{\partial x_i} \text{ as } h \to 0,
\]
uniformly in \(y \in \mathbb{R}^n\) (remember that \(f\) is compactly supported) – we consider \(x\) here to be fixed. Moreover, the functions \(g_h(y)\) are uniformly bounded: there exists a point \(\xi\) on the interval connecting \(x - y\) and \(x - y + he_i\) such that
\[
f(x + he_i - y) - f(x - y) = he_i \cdot \nabla f(\xi),
\]
so that
\[
|g_h(y)| \leq \frac{|h(e_i \cdot \nabla f(\xi))|}{|h|} \leq M_0 = \sup_{z \in \mathbb{R}^n} |\nabla f(z)|.
\]
The integrand in (1.43) can be bounded then as
\[
|\Phi(y)g_h(y)| \leq M_0|\Phi(y)|.
\]
Note that while the function \(\Phi(y)\) is not integrable over all \(\mathbb{R}^n\), its integral over any ball is finite – in particular, over the ball \(B_{R+1}(x)\). Hence, we may apply the Lebesgue dominated convergence theorem and pass to the limit \(h \to 0\) in (1.43) to get
\[
\frac{\partial \phi(x)}{\partial x_i} = \int_{\mathbb{R}^n} \Phi(y) \frac{\partial f(x - y)}{\partial x_i} dy.
\]
A very similar argument shows that
\[
\frac{\partial^2 \phi}{\partial x_i \partial x_j} = \int_{\mathbb{R}^n} \Phi(y) \frac{\partial^2 f(x - y)}{\partial x_i \partial x_j} dy,
\]
hence \(\phi\) is twice differentiable (you need also to argue why the second derivatives are continuous but the argument for that is also very similar to what we just did, except without dividing by any \(h\)).

**Step 2. Derivation of the Poisson equation.** Now, we show that \(\phi\) satisfies the Poisson equation. We know from the above that \(\phi(x)\) is twice continuously differentiable and
\[
\Delta \phi(x) = \int_{\mathbb{R}^n} \Phi(y) \Delta f(x - y) dy.
\]
We need to check that the right side equals \(f(x)\). Since \(\Phi(y)\) is singular at \(y = 0\) we can not simply integrate by parts in the right side but rather have to be more careful. To this end, we take a small \(\varepsilon > 0\) (that we will send to zero at the end of the proof), and split the integral above into the integral over the ball \(B(0, \varepsilon)\) of radius \(\varepsilon\) centered at \(y = 0\) and its complement:
\[
\Delta \phi(x) = I_\varepsilon(x) + J_\varepsilon(x), \quad (1.44)
\]
with
\[
I_\varepsilon(x) = \int_{|y| \leq \varepsilon} \Phi(y) \Delta f(x - y) dy, \quad J_\varepsilon(x) = \int_{|y| \geq \varepsilon} \Phi(y) \Delta f(x - y) dy.
\]
Decomposition (1.44) holds, of course, for any \( \varepsilon > 0 \). Therefore, we also have, trivially:

\[
\Delta \phi(x) = \lim_{\varepsilon \downarrow 0} (I_\varepsilon(x) + J_\varepsilon(x)).
\]  

(1.45)

Our strategy will be to compute the limit in the right side of (1.45) in order to verify that the Poisson equation (1.42) holds.

The contribution of \( I_\varepsilon(x) \) as \( \varepsilon \downarrow 0 \) is small:

\[
|I_\varepsilon(x)| \leq C_f \int_{|y| \leq \varepsilon} |\Phi(y)| dy,
\]  

(1.46)

where \( C_f = \sup_{z \in \mathbb{R}^n} |\Delta f(z)| \). The right side of (1.46) vanishes as \( \varepsilon \to 0 \): when \( n = 2 \) we have

\[
\int_{|y| \leq \varepsilon} |\Phi(y)| dy \leq \frac{1}{2\pi} \int_0^\varepsilon \int_0^{2\pi} |\log r| r dr d\omega \leq C \varepsilon^2 |\log \varepsilon|,
\]

and when \( n > 2 \) we have

\[
\int_{|y| \leq \varepsilon} |\Phi(y)| dy \leq \frac{1}{n(n-2)\alpha(n)} \int_0^\varepsilon \int_{S_{n-1}} \frac{r^{n-2}}{r^{n-1}} dr d\omega \leq C \varepsilon^2.
\]

Hence, in both cases we conclude that

\[
\lim_{\varepsilon \downarrow 0} I_\varepsilon(x) = 0,
\]  

(1.47)

uniformly in \( x \in \mathbb{R}^n \). Therefore, the main contribution to \( \phi(x) \) comes from \( J_\varepsilon(x) \):

\[
\Delta \phi(x) = \lim_{\varepsilon \to 0} J_\varepsilon(x).
\]

Let us now look at \( J_\varepsilon \). First, we recall Green’s formula: given a vector-valued function \( v(x) \) and a scalar valued function \( f(x) \) we have, over a nice domain \( U \):

\[
\int_U v(x) \cdot \nabla f(x) dx = \int_{\partial U} (v(x) \cdot \nu(x)) f(x) dS_x - \int_U f(x) \text{div} v(x) dx.
\]  

(1.48)

Here, \( \nu(x) \) is the outward unit normal to the boundary \( \partial U \) at the point \( x \in \partial U \). Then, integrating by parts we get for \( J_\varepsilon \) (keep in mind that \( \Delta_x (x-y) = \Delta_y f(x-y) \)), since \( \Delta f = \text{div}(\nabla f) \):

\[
J_\varepsilon(x) = \int_{|y| \geq \varepsilon} \Phi(y) \Delta_x f(x-y) dy = \int_{|y| \geq \varepsilon} \Phi(y) \Delta_y f(x-y) dy = \int_{|y| \geq \varepsilon} \Phi(y) \text{div}_y (\nabla_y f(x-y)) dy
\]

\[
= -\int_{|y| \geq \varepsilon} \nabla_y \Phi(y) \cdot \nabla_y f(x-y) dy + \int_{|y| = \varepsilon} \Phi(y) (\nabla_y f(x-y) \cdot \nu(y)) dS_y = K_\varepsilon + L_\varepsilon.
\]

The second term above is small in the limit \( \varepsilon \to 0 \): let \( C'_f = \sup_{z \in \mathbb{R}^n} |\nabla f(z)| \), then

\[
|L_\varepsilon| \leq C'_f \int_{|y| = \varepsilon} |\Phi(y)| dS_y.
\]
so that when $n = 2$ we have

$$|L_{\varepsilon}| \leq C\varepsilon|\log \varepsilon|,$$

and when $n \geq 3$ we have

$$|L_{\varepsilon}| \leq C\varepsilon^{n-1}\varepsilon^{n-2} = C\varepsilon.$$

In both cases we have

$$\lim_{\varepsilon \downarrow 0} L_{\varepsilon} = 0.$$  (1.49)

Thus, the main contribution to $\phi$ must come from $K_{\varepsilon}$:

$$\Delta \phi(x) = \lim_{\varepsilon \downarrow 0} K_{\varepsilon}(x).$$  (1.50)

Let us look at that term: integrating by parts using Green’s formula once again gives

$$K_{\varepsilon}(x) = -\int_{|y| \geq \varepsilon} \nabla_y \Phi(y) \cdot \nabla_y f(x - y) dy = \int_{|y| \geq \varepsilon} \Delta \Phi(y) f(x - y) dy - \int_{|y| = \varepsilon} \frac{\partial \Phi(y)}{\partial \nu} f(x - y) dS_y.$$

As $\Delta \Phi(y) = 0$ for $y \neq 0$, the above is

$$K_{\varepsilon}(x) = -\int_{|y| = \varepsilon} \frac{\partial \Phi(y)}{\partial \nu} f(x - y) dS_y.$$  (1.51)

Consider only the case $n = 3$ – the case $n = 2$ is very similar. Then the normal derivative inside the integrand above is

$$\frac{\partial \Phi(y)}{\partial \nu} \bigg|_{|y| = \varepsilon} = \frac{1}{n\alpha(n)\varepsilon^{n-1}} \bigg|_{|y| = \varepsilon} = \frac{1}{n\alpha(n)\varepsilon^{n-1}},$$

and does not depend on the point $y$. The sign above comes from the fact that the outer normal to $\{|y| \geq \varepsilon\}$ points toward the origin $y = 0$. Using this in (1.51) gives

$$K_{\varepsilon}(x) = -\frac{1}{n\alpha(n)\varepsilon^{n-1}} \int_{|y| = \varepsilon} f(x - y) dS_y.$$  (1.52)

It remains only to observe that $n\alpha(n)\varepsilon^{n-1}$ is the surface area of the sphere of radius $\varepsilon$ in $\mathbb{R}^n$; in general, the volume $\alpha(n)$ of the unit sphere in $\mathbb{R}^n$ is related to its area $s_n$ by $\alpha(n) = s_n/n$ – this is easy to see from calculus. It follows, as $f(x)$ is continuous that

$$\lim_{\varepsilon \downarrow 0} K_{\varepsilon}(x) = -f(x).$$

Going back to (1.50) we conclude that

$$-\Delta \phi(x) = f(x),$$

as claimed. The end of the proof in dimension $n = 2$ is very similar to what we did after (1.51), so we do not present it here. □
A probabilistic interlude

Consider now the following problem: fix two radii $r$ and $R$, and let the Brownian motion start a point $x$ inside the annulus $D = \{ r < |x| < R \}$ in $\mathbb{R}^n$. The Brownian motion will spend some time inside $D$ but will eventually exit $D$ at some random point $\bar{x}$ such that either $|\bar{x}| = r$ or $|\bar{x}| = R$. We ask the following question: what is the probability that the Brownian motion will exit the annulus at the sphere $\{|x| = r\}$ and not on the sphere $\{|x| = R\}$. Let us call this probability $p(x)$. It is clear that if the starting point $x$ is such that $|x| = r$ then $p(x) = 1$ while if $|x| = R$ then $p(x) = 0$. One can show, as we did before, that if we replace the Brownian motion by a discrete random walk then the discretized $p(x)$ satisfies the discrete Laplace equation:

$$p(x) = \frac{1}{2n} \sum_{i=1}^{n} [p(x + e_i) + p(x - e_i)].$$

Here, $e_i$ is the unit vector in the direction of $x_i$ and $2n$ is the total number of the neighbours of the point $x$ on the lattice ($n$ is the spatial dimension). In the case of the Brownian motion which is a continuous limit of the random walks, the function $p(x)$ satisfies the Laplace equation

$$\Delta p(x) = 0 \text{ for } r < |x| < R,$$

supplemented by the boundary conditions

$$p(x) = 1 \text{ if } |x| = r \text{ and } p(x) = 0 \text{ if } |x| = R.$$

In dimension $n = 1$ the solution is given by

$$p(x) = Ax + B,$$

with the constants $A$ and $B$ determined by

$$Ar + B = 1, \quad AR + B = 0,$$

so that

$$A = -\frac{1}{R-r}, \quad B = \frac{R}{R-r}.$$ 

Note that if $R \to +\infty$, that is, if the right point moves to infinity, we have

$$A \to 0, \quad B \to 1 \quad \text{as } R \to +\infty.$$ 

This means that $p(x) \to 1$ as $R \to +\infty$ for any fixed point $x$. This is the reflection of the fact that the Brownian motion is recurrent in one dimension: no matter where it starts, it is certain to reach the point $x = r$. On the other hand, in dimensions $n \geq 3$ the function $p(x)$ is given by

$$p(x) = \frac{A}{|x|^{n-2}} + B,$$

with the constants $A$ and $B$ determined by

$$\frac{A}{r^{n-2}} + B = 1, \quad \frac{A}{R^{n-2}} + B = 0.$$
This gives
\[ A = \frac{r^{n-2}R^{n-2}}{R^{n-2} - r^{n-2}}, \quad B = -\frac{r^{n-2}}{R^{n-2} - r^{n-2}}, \]
We see that in dimension \( n \geq 3 \), as \( R \to +\infty \) we have
\[ A \to r^{n-2}, \quad B \to 0, \]
so that \( p(x) \) has a limit that depends on \( x \) as is always less than one:
\[ p(x) \to \frac{r^{n-2}}{|x|^{n-2}} \text{ as } R \to +\infty. \]
This reflect the fact that the Brownian motion is transitive in dimensions \( n \geq 3 \): no matter how close to the ball \( \{|x| \leq r\} \) it starts, there is a positive probability that it will never enter this ball.

1.3.2 Qualitative properties of harmonic functions
Our next task is to show, from several points of view that harmonic functions are beautifully well-behaved.

**The mean value property**
We begin with the mean value property that shows that locally \( u(x) \) is close to its average. Intuitively this implies that \( u(x) \) can not behave very irregularly and should have limited room to oscillate. And, indeed, the mean value property leads to an amazing number of qualitative conclusions of this kind. A word on notation: for a set \( S \) we denote by \( |S| \) its volume (or area), and by \( \partial S \) we denote its boundary (as we did before).

Let us first recall that in one dimension the mean value property is trivial: any harmonic function in one dimension is linear \( u(x) = ax + b \), and then, of course, for any \( x \in \mathbb{R} \) and any \( l > 0 \) we have
\[ u(x) = \frac{1}{2} (u(x + l) + u(x - l)) = \frac{1}{2l} \int_{x-l}^{x+l} u(y)dy. \]
Here is the generalization to harmonic functions in higher dimensions.

**Theorem 1.3** Let \( U \subset \mathbb{R}^n \) be an open set and let \( B(x, r) \) be a ball centered at \( x \in \mathbb{R}^n \) of radius \( r > 0 \) contained in \( U \). Assume that the function \( u(x) \) satisfies
\[ \Delta u = 0 \text{ for all } x \in U, \quad (1.53) \]
and that \( u \in C^2(U) \). Then we have
\[ u(x) = \frac{1}{|B(x, r)|} \int_{B(x, r)} u dy = \frac{1}{|\partial B(x, r)|} \int_{\partial B(x, r)} u dS. \quad (1.54) \]
The intuitive reason for the mean value property can be seen from the discrete version of the Laplace equation we have encountered when we discussed the probabilistic interpretation:

\[
\frac{1}{2n} \sum_{j=1}^{n} (u(x + he_j) + u(x - he_j)) = u(x).
\]

Here \( h \) is the mesh size, and \( e_j \) is the unit vector in the direction of the coordinate axis for \( x_j \). This discrete equation says exactly that the value \( u(x) \) is the average of the values of \( u \) at the neighbors of the point \( x \) on the lattice with mesh size \( h \), which is similar to the statement of Theorem 1.3 – though there is no meaning to “nearest” neighbor in the continuous case, and the average can be taken over an arbitrary large sphere or ball.

**Proof.** Let us fix the point \( x \in U \) and define

\[
\phi(r) = \frac{1}{|\partial B(x,r)|} \int_{\partial B(x,r)} u(z) dS(z).
\]  
(1.55)

It is easy to see that, since \( u(x) \) is continuous, we have

\[
\lim_{r \downarrow 0} \phi(r) = u(x).
\]  
(1.56)

Therefore, we would be done if we knew that \( \phi'(r) = 0 \) for all \( r > 0 \) such that the ball \( B(x,r) \) is contained in \( U \). To this end, using the polar coordinates \( z = x + ry \), with \( y \in \partial B(0,1) \), we may rewrite (1.55) as

\[
\phi(r) = \frac{1}{|\partial B(0,1)|} \int_{\partial B(0,1)} u(x + ry) dS(y).
\]

Then differentiating in \( r \) gives

\[
\phi'(r) = \frac{1}{|\partial B(0,1)|} \int_{\partial B(0,1)} y \cdot \nabla u(x + ry) dS(y).
\]

Going back to the \( z \)-variables gives

\[
\phi'(r) = \frac{1}{|\partial B(x,r)|} \int_{\partial B(x,r)} \frac{1}{r} (z - x) \cdot \nabla u(z) dS(z) = \frac{1}{|\partial B(x,r)|} \int_{\partial B(x,r)} \frac{\partial u}{\partial \nu} dS(z).
\]

Here, we used the fact that the outward normal to the ball \( B(x,r) \) at a point \( z \in \partial B(x,r) \) is \( \nu = (z - x)/r \). Using the Green’s formula

\[
\int_{U} f \Delta g dy = \int_{\partial U} f \frac{\partial g}{\partial \nu} dS - \int_{U} \nabla f \cdot \nabla g dy,
\]

with \( f = 1 \) and \( g = u \) gives now

\[
\phi'(r) = \frac{1}{|\partial B(x,r)|} \int_{B(x,r)} \Delta u(y) dy = 0,
\]

17
since \( u \) is harmonic – it satisfies (1.53). It follows that \( \phi(r) \) is a constant and then (1.56) implies that
\[
u(x) = \frac{1}{|\partial B(x, r)|} \int_{\partial B(x, r)} u \, dS,
\]
which is the second identity in (1.54).

In order to prove the first equality in (1.54) we use the polar coordinates once again:
\[
\frac{1}{|B(x, r)|} \int_{B(x, r)} u \, dy = \frac{1}{|B(x, r)|} \int_0^r \left( \int_{\partial B(x, s)} u \, dS \right) ds = \frac{1}{|B(x, r)|} \int_0^r u(x) n_\alpha(n) s^{n-1} ds
\]
\[
= u(x) \frac{n_\alpha(n) r^n}{\alpha(n) r^n} = u(x).
\]

In the second equality above we used two facts: first, the already proved identity (1.57) about averages on spherical shells, and, second, that the area of an \((n - 1)\)-dimensional unit sphere is \(n_\alpha(n)\). Now, the proof of (1.54) is complete. \(\Box\)

The maximum principle

The first consequence of the mean value property is the maximum principle that says that a harmonic function attains its maximum over any domain on the boundary and not inside the domain. Once again, in one dimension this is obvious: a linear function does not have any local extremal points.

**Theorem 1.4** (The maximum principle) Let \( u(x) \) be a harmonic function in a connected domain \( U \) and assume that \( u \in C^2(U) \cap C(\bar{U}) \). Then
\[
\max_{x \in U} u(x) = \max_{y \in \partial U} u(y).
\]
Moreover, if \( u(x) \) achieves its maximum at a point \( x_0 \) in the interior of \( U \) then \( u(x) \) is identically equal to a constant in \( U \).

**Proof.** Let us suppose that \( u(x) \) attains its maximum at an interior point \( x_0 \in U \), and set \( M = u(x_0) \). Then for any \( r > 0 \) sufficiently small (so that the ball \( B(x_0, r) \) is contained in \( U \)) we have
\[
M = u(x) = \frac{1}{|B(x_0, r)|} \int_{B(x_0, r)} u \, dy \leq M,
\]
with the equality above holding only if \( u(y) = M \) for all \( y \) in the ball \( B(x_0, r) \). Therefore, the set \( S \) of points where \( u(x) = M \) is open. Since \( u(x) \) is continuous, this set is also closed. Since \( S \) us both open and closed in \( U \), and \( U \) is connected, it follows that \( S = U \), hence \( u(x) = M \) at all points \( x \in U \). \(\Box\)

Of course, if we replace \( u \) by \((-u)\) (which is equally harmonic), we get the minimum principle for \( u \).

**Corollary 1.5** (Strict positivity) Assume that \( U \) is a connected domain, and \( u \) solves
\[
\Delta u = 0 \quad \text{in} \quad U
\]
\[
u = g \quad \text{on} \quad \partial U.
\]
Assume, in addition, that \( g \geq 0 \), \( g \) is continuous on \( \partial U \), and \( g(x) \neq 0 \). Then \( u(x) > 0 \) at all \( x \in U \).

**Proof.** This is an immediate consequence of the minimum principle: \( \min_{x \in \partial U} u(x) \geq 0 \), and \( u \) can not attain its minimum inside \( U \), thus \( u(x) > 0 \) for all \( x \in U \). \( \square \)

**Corollary 1.6 (Uniqueness)** Let \( g \) be continuous on \( \partial U \) and \( f \) be continuous in \( U \). Then there exists at most one solution \( u \in C^2(U) \cap C(\bar{U}) \) to the boundary value problem

\[
\begin{align*}
\Delta u &= f \quad \text{in } U \\
u &= g \quad \text{on } \partial U.
\end{align*}
\]

**Proof.** Let \( u_1 \) and \( u_2 \) be two such solutions to (1.60). Then the difference \( v = u_1 - u_2 \) satisfies the homogeneous problem

\[
\begin{align*}
\Delta v &= 0 \quad \text{in } U \\
v &= 0 \quad \text{on } \partial U.
\end{align*}
\]

The maximum principle implies that \( v \leq 0 \) in \( U \), and the minimum principle implies that \( v \geq 0 \) in \( U \), whence \( v \equiv 0 \), and we are done. \( \square \)

### Regularity of harmonic functions

Now, we prove that if \( u(x) \) is a twice continuously differentiable harmonic function then it is infinitely differentiable – this is quite an amazing result if you think about it, and is a very special property of elliptic equations. Another context such a result appears as in the study of holomorphic functions – a function that has a complex derivative is automatically infinitely differentiable. The “reason” is, of course, that the real and imaginary parts of a holomorphic function are harmonic! For the Laplace equation it can be deduced directly from the mean-value property, and in an arbitrary dimension.

**Theorem 1.7 (Regularity)** Let \( u \in C^2(U) \) be a harmonic function in a domain \( U \). Then \( u \) is infinitely differentiable in \( U \).

**Proof.** The proof is via a miracle: we first define a ”smoothed” version of \( u \), and then verify that the ”smoothed” version coincides with the original, hence original is also infinitely smooth. This is as close to a free lunch as it gets.

Consider a radial non-negative function \( \eta(x) \geq 0 \) that depends only on \( |x| \) such that (i) \( \eta(x) = 0 \) for \( |x| \geq 1 \), (ii) \( \eta(x) \) is infinitely differentiable, and (iii) \( \int_{\mathbb{R}^n} \eta(x)dx = 1 \). Also, for each \( \varepsilon \in (0, 1) \) define its rescaled version

\[ \eta_\varepsilon(x) = \frac{1}{\varepsilon^n} \eta \left( \frac{x}{\varepsilon} \right). \]

It is straightforward to verify that \( \eta_\varepsilon \) satisfies the same properties (i)-(iii) above. Moreover, the function

\[ u_\varepsilon(x) = \int_{\mathbb{R}^n} \eta_\varepsilon(x-y)u(y)dy \quad (1.62) \]
is infinitely differentiable in the slightly smaller domain $U_\varepsilon = \{ x \in \text{ dist}(x, \partial U) > \varepsilon \}$. The reason is that we can differentiate infinitely many times under the integral sign in (1.62) – this follows from the standard multivariable calculus theorem on differentiation of integrals depending on a parameter (the variable $x$ plays the role of a parameter here). Our main claim is that, because of the mean value property, we have

$$u_\varepsilon(x) = u(x) \text{ for all } x \in U_\varepsilon. \tag{1.63}$$

This will immediately imply that $u(x)$ is infinitely differentiable in the domain $U_\varepsilon$. And, as any point $x$ from $U$ lies in $U_\varepsilon$ if $\varepsilon < \text{ dist}(x, \partial U)$, it follows that $u(x)$ is infinitely differentiable at all points $x \in U$.

Let us now verify (1.63):

$$u_\varepsilon(x) = \int_{\mathbb{R}^n} \eta_\varepsilon(x-y)u(y)dy = \frac{1}{\varepsilon^n} \int_U \eta\left(\frac{|x-y|}{\varepsilon}\right)u(y)dy = \frac{1}{\varepsilon^n} \int_{B(x,\varepsilon)} \eta\left(\frac{|x-y|}{\varepsilon}\right)u(y)dy.$$ 

The last equality holds because $\eta(z) = 0$ if $|z| \geq 1$, whence $\eta_\varepsilon(z) = 0$ if $|z| \geq \varepsilon$. Changing variables $y = x + \varepsilon z$ gives

$$u_\varepsilon(x) = \int_{B(0,1)} \eta(z)u(x+\varepsilon z)dz.$$ 

Going to the polar coordinates leads to

$$u_\varepsilon(x) = \int_0^1 \eta(r) \left[ \int_{\partial B(0,1)} u(x + \varepsilon r \omega)dS(\omega) \right] r^{n-1}dr. \tag{1.64}$$

The mean value property implies that

$$\int_{\partial B(0,1)} u(x + \varepsilon r \omega)dS(\omega) = u(x)|\partial B(0,1)|.$$ 

Using this in (1.64), we obtain

$$u_\varepsilon(x) = u(x) \int_0^1 \eta(r)|\partial B(0,1)|r^{n-1}dr = u(x) \int_{B(0,1)} \eta(y)dy = u(x), \tag{1.65}$$

which is (1.63). We used the fact that $\eta$ has integral equal to one in the last step. □

This regularity property is quite fundamental and appears in one way or other for the class of elliptic equations (and not just for the Laplace equation) we will discuss later. One of their main qualitative properties is that solutions are more regular than the data prescribed, and they behave much nicer than, say, solutions of wave equations and other hyperbolic problems.

Let us now give a more quantitative estimate on how large the derivatives of the harmonic functions can be.

**Theorem 1.8** Let $u(x)$ be a harmonic function in a domain $U$ and let $B(y_0, r)$ be a ball contained in $U$ centered at a point $y_0 \in U$. Then there exist universal constants $C_n$ and $D_n$ that depends only on the dimension $n$ so that we have

$$|u(y_0)| \leq \frac{C_n}{r^n} \int_{B(y_0, r)} |u(y)|dy. \tag{1.66}$$
and
\[ |\nabla u(y_0)| \leq \frac{D_n}{r^{n+1}} \int_{B(y_0, r)} |u(y)| dy. \] (1.67)

The remarkable fact about the estimate (1.67) is that we are able to estimate the size of the derivatives of a harmonic function in terms of its values – this means that a harmonic function can not oscillate (oscillation means, essentially, that the function is much smaller than its derivative). It is a good time to pause and think about why taking a harmonic function \( u(x) \) and setting \( u_\varepsilon(x) = u(x/\varepsilon) \) with \( y_0 = 0 \) does not provide a counterexample for (1.67). We certainly would have \( |\nabla u_\varepsilon(0)| = |\nabla u(0)|/\varepsilon \). But what has to happen to the right side of (1.67) when we replace \( u \) by \( u_\varepsilon \) there?

**Proof.** First, the estimate (1.66) follows immediately from the first equality in the mean value formula (1.54). In order to obtain the derivative bound (1.67) note that if \( u(x) \) is harmonic then so are the partial derivatives \( \frac{\partial u}{\partial x_j} \), whence
\[ \left| \frac{\partial u(y_0)}{\partial x_j} \right| \leq \frac{1}{|B(y_0, r/2)|} \int_{B(y_0, r/2)} \left| \frac{\partial u(y)}{\partial x_j} \right| dy \leq \frac{1}{|B(y_0, r/2)|} \int_{\partial B(y_0, r/2)} u(y) \nu_j(y) dy, \] (1.68)
where \( \nu_j(y_0) \) is the \( j \)-th component of the outward normal. Continuing this estimate we see that (we use the fact that the area of the unit sphere is \( n\alpha(n) \))
\[ \left| \frac{\partial u(y_0)}{\partial x_j} \right| \leq \frac{2^n}{\alpha(n)r^n} \frac{n\alpha(n)r^{n-1}}{2^{n-1}} \sup_{z \in B(y_0, r/2)} |u(z)| = \frac{2n}{r} \sup_{z \in B(y_0, r/2)} |u(z)|. \] (1.69)
Now, we can use the estimate (1.66) applied at any point \( z \in B(y_0, r/2) \):
\[ |u(z)| \leq \frac{C_n}{(r/2)^n} \int_{B(z, r/2)} |u(z')| dz'. \] (1.70)
However, since \( |y_0 - z| \leq r/2 \) (this is why we took a smaller ball in (1.68)!), any such ball \( B(z, r/2) \) is contained inside the ball \( B(y_0, r) \), thus (1.70) implies that
\[ |u(z)| \leq \frac{C_n}{(r/2)^n} \int_{B(y_0, r)} |u(z')| dz'. \]
Now, it follows from (1.69) that
\[ \left| \frac{\partial u(y_0)}{\partial x_j} \right| \leq \frac{2n}{r} \frac{C_n}{(r/2)^n} \int_{B(y_0, r)} |u(z')| dz' = \frac{D_n}{r^{n+1}} \int_{B(y_0, r)} |u(z)| dz, \] (1.71)
which is (1.67). \( \Box \)

Theorem 1.8 is another expression of the fact that harmonic functions do not oscillate – the first estimate says that the value of the function at a point is bounded by its averages (but we have seen that already in the mean value property), while the second bound says in a quantitative way that derivative at a point can not be large without the function being large around the point. This rules out oscillatory behavior.
The Liouville theorem

The Liouville theorem says that a function which is harmonic in all of $\mathbb{R}^n$ is either unbounded or is identically equal to a constant.

**Theorem 1.9** Let $u(x)$ be a harmonic bounded function in $\mathbb{R}^n$. Then $u(x)$ is equal identically to a constant.

**Proof.** Let us assume that $|u(x)| \leq M$ for all $x \in \mathbb{R}^n$. We fix $x_0 \in \mathbb{R}^n$ and use Theorem 1.8:

$$|\nabla u(x_0)| \leq \frac{C}{r^{n+1}} \int_{B(x_0,r)} |u(y)| dy \leq \frac{C\alpha(n)r^n}{r^{n+1}} M \leq \frac{C\alpha(n)M}{r}.$$ 

As this is true for any $r > 0$ we may let $r \to \infty$ and conclude that $\nabla u(x_0) = 0$, thus $u(x)$ is equal identically to a constant. \(\square\)

This theorem is, of course, a direct generalization to higher dimensions of the familiar Liouville theorem in complex analysis that says that a bounded entire (analytic in all of $\mathbb{C}$) function has to be equal identically to a constant.

Harnack’s inequality

Here is another way to express lack of oscillations of nonnegative harmonic functions – their maximum cannot be much larger than their minimum. To trivialize, consider the one-dimensional situation. Let $u(x)$ be a non-negative harmonic function on the interval $(0, 1)$, that is, $u(x) = ax + b$ with some constants $a, b \in \mathbb{R}$. We claim that if $u(x) \geq 0$ for all $x \in [0, 1]$ then

$$\frac{1}{3} \leq \frac{u(x)}{u(y)} \leq 3,$$  \hspace{1cm} (1.72)

for all $x, y$ in the smaller interval $(1/4, 3/4)$. The constants $1/3$ and $3$ in (1.72) depend on the choice of the "smaller" interval – they would change of we would replace $(1/4, 3/4)$ by another subinterval of $[0, 1]$. But once we fix the subinterval, they do not depend on the choice of the harmonic function. Let us now show that (1.72) holds for all $x, y \in (1/4, 3/4)$. Without loss of generality we may assume that $x > y$. First, consider the case $a > 0$. Then, since $u(x)$ is increasing (because $a > 0$), we have

$$1 \leq \frac{u(x)}{u(y)} \leq \frac{u(3/4)}{u(1/4)} = \frac{3a + 4b}{a + 4b}.$$  \hspace{1cm} (1.73)

As $u(x) > 0$ on $[0, 1]$ we know that $b > 0$ (and $a > 0$ by assumption), using this in (1.73) gives, with $c = a/b$:

$$1 \leq \frac{u(x)}{u(y)} \leq \frac{3c + 4}{c + 4} = 3 - \frac{8}{c + 4} \leq 3.$$  \hspace{1cm} (1.74)

On the other hand, if $a < 0$ then the function $u$ is decreasing, and

$$1 \geq \frac{u(x)}{u(y)} \geq \frac{u(3/4)}{u(1/4)} = \frac{c + 4}{3c + 4} = \frac{1}{3} + \frac{8}{3(3c + 4)}.$$  \hspace{1cm} (1.75)
As \( u(1) > 0 \) we know that \( a + b > 0 \), and we still have \( b > 0 \) since \( u(0) > 0 \). Thus, \( c > -1 \), and therefore,

\[
1 \geq \frac{u(x)}{u(y)} \geq \frac{1}{3} + \frac{8}{3(3c + 4)} \geq \frac{1}{3}.
\]

We conclude that (1.72), indeed, holds. Geometrically, (1.72) expresses a very simple fact: if \( u(3/4) \gg u(1/4) \) then the slope of the straight line connecting the points \((1/4, u(1/4))\) and \((3/4, u(3/4))\) is too large so that it would go below the \(x\)-axis at \( x = 0 \) contradicting the assumption that the linear function is positive on the interval \((0, 1)\). On the other hand, if \( u(1/4) \gg u(3/4) \) then this line would go below that \(x\)-axis at \( x = 1 \). Therefore, the condition that \( u(x) > 0 \) on the larger interval \([0, 1]\) is very important here.

Now, we turn to the general case of dimension larger than one. We say that a set \( V \) is strictly contained in \( U \) if \( V \subset U \) and there exists \( \varepsilon_0 > 0 \) so that for any \( x \in V \) we have \( \text{dist}(x, \partial U) \geq \varepsilon_0 \).

**Theorem 1.10 (Harnack’s inequality)** Let \( U \) be an open set and let \( V \) be strictly contained in \( U \). Then there exists a constant \( C \) that depends on \( U \) and \( \varepsilon_0 \) but nothing else so that for any nonnegative harmonic function \( u \) in \( U \) we have

\[
\sup_{x \in V} u(x) \leq C \inf_{x \in V} u(x). \tag{1.74}
\]

**Proof.** Let \( r = (1/4)\text{dist}(V, \partial U) \) and choose two points \( x, y \in V \) such that \( |x - y| \leq r \). Then the ball \( B(x, 2r) \) is contained in \( U \) so \( u \) is harmonic in this ball, and the mean-value principle implies that

\[
u(x) = \frac{1}{|B(x, 2r)|} \int_{B(x, 2r)} u(z)dz. \tag{1.75}
\]

Note also that since \( |x - y| \leq r \), the ball \( B(y, r) \) is contained inside \( B(x, 2r) \), and \( u(z) \geq 0 \) everywhere. Hence, (1.75) implies that

\[
u(x) \geq \frac{1}{\alpha(n)2^{n}r^{n}} \int_{B(y, r)} u(z)dz. \tag{1.76}
\]

It follows, on the other hand, from the mean-value principle that

\[
u(y) = \frac{1}{\alpha(n)r^{n}} \int_{B(y, r)} u(z)dz. \tag{1.77}
\]

Putting (1.76) and (1.77) together gives

\[
u(x) \geq \frac{1}{2^n}u(y). \tag{1.78}
\]

Reversing the argument we can similarly conclude that

\[
u(y) \geq \frac{1}{2^n}u(x), \tag{1.79}
\]

hence

\[
\frac{1}{2^n}u(x) \leq u(y) \leq 2^n u(x), \quad \text{for all } x, y \in V \text{ such that } |x - y| \leq (1/4)\text{dist}(V, \partial U). \tag{1.80}
\]
In general, if $|x - y| \geq r$, there exists a number $N$ so that we may cover the compact set $\bar{V}$ by $N$ balls of radius $r/2$. Then given any two points $x, y \in V$ we can connect them by a piece-wise straight line curve with no more than $N$ segments, each segment at most $r$ long. It follows that for any $x, y \in V$ we have

$$\frac{1}{2^n} u(x) \leq u(y) \leq 2^N u(x), \text{ for all } x, y \in V.$$  \hspace{1cm} (1.81)

This, of course, implies (1.74) with $C = 2^{nN}$. \hfill \Box

### 1.3.3 Green’s function for the Poisson equation

We will now show a systematic way to construct solutions of the boundary value problem for the Poisson equation

$$-\Delta u = f \text{ in } U, \quad u = g \text{ on } \partial U. \hspace{1cm} (1.82)$$

When the domain $U$ is sufficiently simple (a ball, halfspace, etc.) then we will construct a more or less explicit formula for the solution. When $U$ is complicated we cannot get an explicit formula but we will reduce solving (1.82) with arbitrary functions $f$ and $g$ to the special case $f = 0$, and one particular function $g$. Having a solution to this one special case allows to construct solutions for general $f$ and $g$ immediately. This is useful when one needs to solve the Poisson equation in the same domain for various $f$ and $g$. It also helps to understand various qualitative properties of the solutions of the boundary value problem for the Poisson equation.

#### Definition of the Green’s function

Let us recall the fundamental solution of the Laplace equation $\Phi(x)$ we have constructed in Section 1.3 (compare to (1.39)-(1.40)):

$$\Phi(x) = -\frac{1}{2\pi} \log |x|, \quad n = 2, \hspace{1cm} (1.83)$$

and

$$\Phi(x) = \frac{1}{n(n-2)\alpha(n)} \frac{1}{|x|^{n-2}}, \quad n \geq 3. \hspace{1cm} (1.84)$$

Theorem 1.1 asserts that the function

$$u(x) = \int_{\mathbb{R}^n} \Phi(x - y)f(y)dy \hspace{1cm} (1.85)$$

is a solution of the Poisson equation

$$-\Delta u = f$$

posed in all of $\mathbb{R}^n$. What we would like to do is to adapt the representation (1.85) to the boundary value problem (1.82) posed in a bounded domain, and also taking into account the
correct boundary conditions. That is, we are hoping to get an integral representation of the solution of the boundary value problem (1.82) as

\[ u(x) = \int_U G(x, y)f(y)dy + \int_{\partial U} G_1(x, z)g(z)dz, \]

with some functions \( G(x, y) \) and \( G_1(x, z) \) that are to be determined (but they should not depend on the functions \( f \) and \( g \) – they should only depend on the domain \( U \) where the problem is posed).

To this end, we take a point \( x \in U \), and a small ball \( B(x, \varepsilon) \) around it. Consider the domain \( V_\varepsilon = U \setminus B(x, \varepsilon) \) (that is, \( U \) without the ball \( B(x, \varepsilon) \)) and use the Green’s formula:

\[ \int_{V_\varepsilon} [u(y)\Delta \Phi(y - x) - \Phi(y - x)\Delta u(y)] dy = \int_{\partial V_\varepsilon} \left[ u(z) \frac{\partial \Phi(z - x)}{\partial \nu} - \Phi(z - x) \frac{\partial u(z)}{\partial \nu} \right] dS(z). \]  

(1.87)

The reason we had to cut out the small ball around the point \( x \) is that now when \( y \in V_\varepsilon \) the argument \( (y - x) \) of the fundamental solution \( \Phi(x - y) \) can not vanish, and \( \Phi(z) \) is regular when \( z \neq 0 \). Otherwise, we would not be able to apply Green’s formula since \( \Phi(z) \) is singular at \( z = 0 \). As \( \Delta \Phi(y - x) = 0 \) when \( y \neq x \), the above is

\[ -\int_{V_\varepsilon} \Phi(y - x)\Delta u(y)dy = \int_{\partial V_\varepsilon} \left[ u(z) \frac{\partial \Phi(z - x)}{\partial \nu} - \Phi(z - x) \frac{\partial u(z)}{\partial \nu} \right] dS(z). \]  

(1.88)

This identity holds for all \( \varepsilon > 0 \) and we will now pass to the limit \( \varepsilon \downarrow 0 \) in (1.88), taking the size of the cut-out region to zero. The boundary \( \partial V_\varepsilon \) of the domain \( V_\varepsilon \) is the union of \( \partial U \) and the sphere \( S_\varepsilon = \{|z - x| = \varepsilon\} \). The integral over \( S_\varepsilon \) is computed as in the proof of Theorem 1.1 (we again do only the computations for \( n \geq 3 \), the case \( n = 2 \) is similar): first, we may use the explicit formula

\[ \Phi(x) = \frac{1}{n(n - 2)\alpha(n)|x|^{n-2}}. \]

Note that for \( z \) on the sphere \( S_\varepsilon \) we have \(|z - x| = \varepsilon\), meaning that

\[ \frac{\partial \Phi}{\partial \nu}(z - x) = \frac{1}{n\alpha(n)\varepsilon^{n-1}}, \]

so that

\[ \int_{\partial S_\varepsilon} u(z) \frac{\partial \Phi(z - x)}{\partial \nu} dS(z) = \frac{1}{n\alpha(n)\varepsilon^{n-1}} \int_{\partial S_\varepsilon} u(z)dS(z) = \frac{1}{|\partial S_\varepsilon|} \int_{\partial S_\varepsilon} u(z)dS(z). \]  

(1.89)

We used here the formula \(|\partial S_\varepsilon| = n\alpha(n)\varepsilon^{n-1}\) for a sphere of radius \( \varepsilon \) in \( \mathbb{R}^n \). As \( u \) is continuous at the point \( x \), letting \( \varepsilon \downarrow 0 \) we obtain

\[ \int_{\partial S_\varepsilon} u(z) \frac{\partial \Phi(z - x)}{\partial \nu} dS(z) = \frac{1}{|\partial S_\varepsilon|} \int_{\partial S_\varepsilon} u(z)dS(z) \rightarrow u(x) \quad \text{as} \quad \varepsilon \downarrow 0. \]  

(1.90)
The other term in the right side of (1.88) vanishes as $\varepsilon \downarrow 0$:

$$
\left| \int_{\partial S_\varepsilon} \Phi(z - x) \frac{\partial u(z)}{\partial \nu} dS(z) \right| = \frac{1}{n(n-2)\alpha(n)\varepsilon^{n-2}} \left| \int_{\partial S_\varepsilon} \frac{\partial u(z)}{\partial \nu} dS(z) \right| \leq \frac{M\varepsilon^{n-1}\alpha(n)}{n\alpha(n)(n-2)\varepsilon^{n-2}} \to 0, \quad (1.91)
$$

as $\varepsilon \downarrow 0$, where $M = \sup_{y \in U} |\nabla u|$. We used again the formula $|\partial S_\varepsilon| = n\alpha(n)\varepsilon^{n-1}$ above. Therefore, passing to the limit $\varepsilon \downarrow 0$ in (1.88) leads to

$$
u(x) = \int_{\partial U} \left[ \Phi(z - x) \frac{\partial u(z)}{\partial \nu} - u(z) \frac{\partial \Phi(z - x)}{\partial \nu} \right] dS(z) - \int_{U} \Phi(y - x) \Delta u(y) dy. \quad (1.92)
$$

Hence, in order to compute $u(x)$ we should know $\Delta u$ inside $U$ (which we do for the solution of the Poisson equation (1.82), it is $f$), as well as $u(z)$ on the boundary $\partial U$ (which we do know for the solution of the boundary value problem (1.82), it is $g$), but also the normal derivative $\partial u/\partial \nu$ at the boundary of $U$, and that we do not know a priori – this normal derivative can only be found after we solve (1.82). Therefore, (1.92) is not yet the answer we seek – it involves an unknown function $\partial u/\partial \nu$.

Note that this would not have been an issue if we would have $\Phi(x - y) = 0$ on $\partial U$ – then the corresponding term in (1.92) would have vanished. The idea is, then, to amend $\Phi(x - y)$ in such a way as to make this term disappear. This is done as follows. Fix a point $x \in U$ and let $\phi(y; x)$ be the solution of the boundary value problem

$$
-\Delta_y \phi = 0 \quad \text{for } y \in U,
$$

$$
\phi(y; x) = \Phi(y - x) \quad \text{for } y \in \partial U. \quad (1.93)
$$

Observe that, as $x$ lies inside the domain $U$, the function $\Phi(x - y)$ is regular when $y$ lies on the boundary $\partial U$ - the distance between $x$ and $y$ is uniformly positive. Therefore, (1.93) is simply a Laplace equation in $y$ with regular prescribed boundary data $\Phi(x - y)$ ($x$ here serves as a parameter). Hence, the function $\phi(y; x)$ is regular and has no singularity.

Using the Green’s formula as before (but without the need to throw out a small ball around the point $x$ since the function $\phi(y; x)$ is regular at $y = x$) gives

$$
- \int_{U} \phi(y; x) \Delta u(y) \, dy = \int_{\partial U} \left[ u(z) \frac{\partial \phi(z; x)}{\partial \nu} - \phi(z; x) \frac{\partial u(z)}{\partial \nu} \right] dS(z). \quad (1.94)
$$

The boundary condition for $\phi(y; z)$ shows that this is

$$
- \int_{U} \phi(y; x) \Delta u(y) \, dy = \int_{\partial U} \left[ u(z) \frac{\partial \phi(z; x)}{\partial \nu} - \phi(z; x) \frac{\partial u(z)}{\partial \nu} - \Phi(z - x) \frac{\partial u(z)}{\partial \nu} \right] dS(z). \quad (1.95)
$$

**Definition 1.11** The Green’s function for a domain $U$ is

$$
G(x; y) = \Phi(x - y) - \phi(y; x). \quad (1.96)
$$

Now, adding (1.92) and (1.95) up gives

$$
u(x) = - \int_{\partial U} u(z) \frac{\partial G(x; z)}{\partial \nu_z} dS(z) - \int_{U} G(x; y) \Delta u(y) dy. \quad (1.97)
$$
The advantage of (1.97) over (1.92) is that the normal derivative $\partial u/\partial \nu$ on $\partial U$ (which we do not know) no longer appears in the right side. Hence, solution of the Poisson boundary value problem (1.82) is given by

$$u(x) = -\int_{\partial U} g(z) \frac{\partial G(x; z)}{\partial \nu} dS(z) + \int_{U} G(x; y)f(y)dy.$$

(1.98)

This expression is particularly useful when $G(x; z)$ is known explicitly, and we will discuss below some examples when it can be computed analytically.

**How to approximate the Green’s function**

Let us now briefly discuss how one can approximate the Green’s function that we have just constructed. Let us think of $G_h(x, y)$ as an approximate Green’s function in a domain $U$ as $h \to 0$. We will keep the boundary condition $G(x, y) = 0$ for any $x \in U$ and $y \in \partial U$, so that

$$G_h(x, y) = 0 \text{ for } y \in \partial \Omega.$$  

Let now $u(x)$ be the solution of the boundary value problem

$$-\Delta u = f,$$

with the Dirichlet boundary condition $u = 0$ on $\partial \Omega$. Then, according to (1.98), we have

$$u(x) = \int_{U} G(x, y)f(y)dy.$$

As $G_h(x, y)$ should approximate $G(x, y)$, we should have

$$u(x) = \int_{U} G_h(x, y)f(y)dy + l.o.t.$$  

(1.99)

Let us then look for $f_h(x, y)$ such that if $G_h(x, y)$ satisfies the boundary value problem

$$-\Delta_y G_h = f_h(x, y), \text{ in } U$$

$$G_h(x, y) = 0 \text{ for } y \in \partial U,$$

then (1.99) holds. Since both $G_h$ and $u$ vanish on the boundary, Green’s formula says that

$$\int_{U} G_h(x, y)\Delta_y u(y) = \int_{U} u(y)\Delta G_h(y)dy.$$

We can write this as

$$\int_{U} G_h(x, y)f(y)dy = \int_{U} f_h(x, y)u(y)dy.$$

The question now becomes: what is a suitable choice $f_h(x, y)$ such that

$$\int_{U} f_h(x, y)u(y)dy - u(x) \to 0 \text{ as } h \to 0.$$
A suitable choice is to take a fixed $f(x) \geq 0$ such that $f(x) = 0$ for $|x| \geq 1$, and

$$
\int_{|x| \leq 1} f(y)dy = 1,
$$

and set (recall that $x \in \mathbb{R}^n$):

$$
f_h(x; y) = \frac{1}{h^n} f\left(\frac{x - y}{h}\right).
$$

Indeed, for a continuous function $u(x)$ we have for any $x$ fixed, and $h < \text{dist}(x, \partial U)$

$$
\int_U f_h(x, y)u(y)dy = \frac{1}{h^n} \int_U f\left(\frac{x - y}{h}\right)u(y)dy = \frac{1}{h^n} \int_{|x-y| \leq h} f\left(\frac{x - y}{h}\right)u(y)dy
$$

$$
= \int_{|z| \leq 1} f(z)u(x - hz)dz \rightarrow u(x) \int_{|z| \leq 1} f(z)dz = u(x).
$$

Thus, to find an approximation to $G_h(x,y)$ we need to solve (1.100).

Reciprocity of the Green’s function

Physically, the meaning of $G(x, y)$ is as follows: it is the value of the electric potential at the point $x$ when a localized charge is put at the point $y$. The physical principle of reciprocity implies that $G(x; y) = G(y; x)$. Let us show that this is, indeed, the case.

**Theorem 1.12** We have for all $x, y \in U, x \neq y$:

$$
G(x; y) = G(y; x).
$$

**Proof.** Once again, we will use Green’s formula. Let $x \neq y$ be two distinct points in $U$, and set $v(z) = G(x;z)$ and $w(z) = G(y;z)$. Let us cut out two small balls $B(x, \varepsilon)$ and $B(y, \varepsilon)$ with $\varepsilon > 0$ so small that the balls are not overlapping and are contained in $U$. Let

$$
V_\varepsilon = U \setminus (B(x, \varepsilon) \cup B(y, \varepsilon))
$$

be the domain $U$ with the two balls deleted. Then $\Delta_z w = \Delta_z v = 0$ in $V_\varepsilon$ as this set contains neither the point $x$ nor the point $y$. The Green’s formula then becomes

$$
\int_{\partial V_\varepsilon} \left[ w(z) \frac{\partial v(z)}{\partial \nu} - v(z) \frac{\partial w(z)}{\partial \nu} \right] dS(z) = 0.
$$

The boundary of $V_\varepsilon$ consists of three pieces: the outer boundary $\partial U$ where both $w$ and $v$ vanish, and the two spheres $\partial B(x, \varepsilon)$ and $\partial B(y, \varepsilon)$. Therefore, we have

$$
\int_{\partial B(x, \varepsilon)} \left[ w(z) \frac{\partial v(z)}{\partial \nu} - v(z) \frac{\partial w(z)}{\partial \nu} \right] dS(z) + \int_{\partial B(y, \varepsilon)} \left[ w(z) \frac{\partial v(z)}{\partial \nu} - v(z) \frac{\partial w(z)}{\partial \nu} \right] dS(z) = 0.
$$

(1.104)

Here $\nu$ is the normal pointing inside the spheres (it is the outside normal to $V_\varepsilon$ which is outside the two small balls). The functions $w$ and $v$ look as follows near $x$ and $y$: $v$ is regular in $B(y, \varepsilon)$, $w$ is regular in $B(x, \varepsilon)$. On the other hand, $v(z) = \Phi(x - z) - \phi(z; x)$, and $\phi(z; x)$ is
regular in $z$, including in $B(x; z)$, while $w(z) = \Phi(y - z) - \phi(z; y)$, and $\phi(z; y)$ is regular in $z$, including in $B(y; z)$. Hence, as in the discussion in the previous section leading up to (1.97), the main terms in (1.104) are

$$
\int_{\partial B(x, \varepsilon)} w(z) \frac{\partial \Phi(x - z)}{\partial \nu} dS(z) - \int_{\partial B(y, \varepsilon)} v(z) \frac{\partial \Phi(y - z)}{\partial \nu} dS(z) + l.o.t. = 0,
$$

where l.o.t. denotes terms that tend to zero as $\varepsilon \to 0$. Passing to the limit $\varepsilon \downarrow 0$ exactly as in (1.90), since $w(z)$ is continuous at $x$, and $v(z)$ is continuous in $y$, gives $w(x) - v(y) = 0$, which is exactly (1.103). □

**Green’s function for a half space**

We will now construct explicitly the Green’s function $G(x; y)$ for a half space

$$
\mathbb{R}^n_+ = \{ (x_1, \ldots, x_n) : x_n > 0 \}.
$$

This is done by the method of reflections. We need to find a function $\phi(x; y)$ that solves the following boundary value problem:

$$
\begin{align*}
-\Delta_y \phi(x; y) &= 0 \text{ for } y \in \mathbb{R}^n_+, \\
\phi(x; y) &= \Phi(x - y) \text{ for } y \text{ such that } y_n = 0,
\end{align*}
$$

with a given point $x \in \mathbb{R}^n_+$. Consider the reflected point $\bar{x} = (x_1, \ldots, x_{n-1}, -x_n)$, that lies in the lower half-space, and the corresponding fundamental solution $\Phi(\bar{x} - y)$. Then (i) as $\bar{x}$ lies outside of $\mathbb{R}^n_+$, this function is harmonic in $\mathbb{R}^n_+$, and (ii) since for any $y$ with $y_n = 0$ we have $|x - y| = |\bar{x} - y|$ and since the function $\Phi(x)$ is radial, we have $\Phi(x - y) = \Phi(\bar{x} - y)$ for all points $y$ on the hyper-plane $\{y_n = 0\}$. Therefore, $\phi(x; y) = \Phi(\bar{x} - y)$, and

$$
G(x; y) = \Phi(x - y) - \Phi(\bar{x} - y)
$$

is the Green’s function for the half space $\mathbb{R}^n_+$. We also need to compute its normal derivative at the hyper-plane $\{y_n = 0\}$:

$$
\frac{\partial G(x; y)}{\partial y_n} = \frac{\partial \Phi(y - x)}{\partial y_n} - \frac{\partial \Phi(y - \bar{x})}{\partial y_n} = -\frac{1}{n\alpha(n)} \left[ \frac{y_n - x_n}{|y - x|^n} - \frac{y_n + x_n}{|y - \bar{x}|^n} \right].
$$

Hence, if $y_n = 0$ then the outward normal to $\partial \mathbb{R}^n_+$ at the point $y$ is:

$$
\frac{\partial G(x; y)}{\partial \nu} = -\frac{2x_n}{n\alpha(n)|y - x|^n}.
$$

Let us define the Poisson kernel for the half-space

$$
K(x, y) = \frac{2x_n}{n\alpha(n)|y - x|^n}, \quad x \in \mathbb{R}^n_+, \quad y \in \partial \mathbb{R}^n_+.
$$

Then, (1.98) tells us that solution of the boundary value problem

$$
\begin{align*}
-\Delta u &= 0 \text{ for } y \in \mathbb{R}^n_+, \\
u(x) &= g(x) \text{ for } x \text{ such that } x_n = 0,
\end{align*}
$$

29
with a prescribed function \( g(x) \) is

\[
u(x) = \frac{2}{n\alpha(n)} \int_{\mathbb{R}^{n-1}} \frac{g(y)}{|x-y|^n} dy_1 \ldots dy_{n-1}, \tag{1.110}\]

for \( x \in \mathbb{R}^n \). Note that the denominator never vanishes since \( x_n > 0 \) and the distance \( |x-y| \) is computed in \( \mathbb{R}^n \) so that

\[
|x-y| = \left( (x_1 - y_1)^2 + \ldots + (x_{n-1} - y_{n-1})^2 + x_n^2 \right)^{1/2}.
\]

Of course, one should be more careful here: what (1.98) actually says is that the only possible smooth solution of (1.109) is given by the convolution with the Poisson kernel, but we do not know yet that the function \( u(x) \) defined by (1.110) is, indeed, a solution of (1.109). It is quite straightforward to verify that \( u(x) \) is harmonic in the upper half-space: all we need to check for that is that

\[
\Delta_x K(x, y) = 0,
\]

and that is true because (i) \( y \) is not in the interior of \( \mathbb{R}^n_+ \) – hence, \( K(x, y) \) is regular at all \( x \in \mathbb{R}^n_+ \), and (ii) \( K(x, y) = \partial G(x, y)/\partial y_n \), and \( G(x, y) \) is harmonic (in both \( x \) and \( y \) for \( x \neq y \)).

It is much more delicate to verify that the boundary condition for \( u(x) \) holds, that is, that \( u(x) \) is continuous up to the boundary \( \{ x_n = 0 \} \), and that

\[
\lim_{x_n \to 0} u(x_1, \ldots, x_n) = g(x_1, \ldots, x_{n-1}). \tag{1.111}
\]

The reason why the boundary condition holds is as follows. First, one can verify easily that

\[
\int_{\mathbb{R}^{n-1}} K(x, y) dy = 1, \tag{1.112}
\]

for each fixed \( x \in \mathbb{R}^n_+ \). Second, we can write the Poisson kernel as

\[
K(x_1, \ldots, x_{n-1}, x_n, y) = \frac{x_n}{n\alpha(n)} \left( (x_1 - y_1)^2 + \ldots + (x_{n-1} - y_{n-1})^2 + x_n^2 \right)^{n/2},
\]

\[
= \frac{2}{n\alpha(n)x_n^{n-1}} \left[ \left( \frac{x_1 - y_1}{x_n} \right)^2 + \ldots + \left( \frac{x_{n-1} - y_{n-1}}{x_n} \right)^2 + 1 \right]^{-n/2}.
\]

Denoting \( x' = (x_1, \ldots, x_{n-1}) \) we obtain then

\[
u(x', x_n) = \int_{\mathbb{R}^{n-1}} K(x, y) g(y) dy = \frac{2}{n\alpha(n)x_n^{n-1}} \int_{\mathbb{R}^{n-1}} \left[ 1 + \left| \frac{x' - y}{x_n} \right|^2 \right]^{-n/2} g(y) dy.
\]

Changing variables \( y' = (x' - y)/x_n \) gives

\[
u(x', x_n) = \frac{2}{n\alpha(n)} \int_{\mathbb{R}^{n-1}} \left[ 1 + |y'|^2 \right]^{-n/2} g(x' - x_n y') dy'.
\]
Letting $x_n \to 0$ gives

$$
\lim_{x_n \downarrow 0} u(x', x_n) = \frac{2}{n \alpha(n)} \int_{\mathbb{R}^{n-1}} \left[1 + |y'|^2\right]^{-n/2} g(x') dy' = g(x').
$$

(1.113)

In the last identity we used (1.112) with $x = (0, \ldots, 0, 1)$. The passage to the limit $x_n \to 0$ in (1.113) follows from continuity and boundedness of the function $g(x)$. Therefore, $u(x)$ satisfies the boundary condition in (1.109). Note that this computation is essentially the same as what we have done in the approximation of the Green’s function in (1.101)-(1.102).

1.3.4 Energy and variational methods

Uniqueness by the energy method

We now get a glimpse of what is known as energy methods which are a very useful tool in dealing with many PDE problems. We have previously shown that solution of the boundary value problem

$$
\begin{align*}
-\Delta u &= f \text{ in } U, \\
u &= g \text{ on } \partial U
\end{align*}
$$

(1.114)

in a bounded domain $U$ in $\mathbb{R}^n$ with a smooth boundary, is unique, using the maximum principle. Here is a way to show uniqueness without using the maximum principle. Let $u_1$ and $u_2$ be two solutions of (1.114). The function $v = u_1 - u_2$ satisfies

$$
\begin{align*}
-\Delta v &= 0 \text{ in } U, \\
v &= 0 \text{ on } \partial U.
\end{align*}
$$

(1.115)

Let us multiply the Laplace equation (1.115) by $v$ and integrate over the domain $U$:

$$
\int_U v \Delta v = 0.
$$

The Green’s formula then implies

$$
0 = \int_{\partial U} v \frac{\partial v}{\partial n} dS - \int_U |\nabla v|^2 dx.
$$

As $v = 0$ on $\partial U$, we conclude that

$$
\int_U |\nabla v|^2 dx = 0,
$$

whence $v = 0$ and $u_1 = u_2$ proving the uniqueness claim.

The variational formulation

We now discuss how PDEs can be reformulated in terms of variational problems. This is a very powerful tool (when it is available, and it is possible not for all PDEs) both to prove existence and uniqueness of the solution, and, maybe more importantly, to find it numerically.
In order to understand the idea, let us recall some linear algebra. Consider the equation

\[ Ax = b, \quad (1.116) \]

where \( A \) is a real-valued \( n \times n \) matrix, \( b \) is a given vector in \( \mathbb{R}^n \), and \( x \in \mathbb{R}^n \) is the unknown. Assume that the matrix \( A \) is symmetric positive-definite, meaning that there exists a positive number \( a > 0 \) so that for any \( y \in \mathbb{R}^n \) we have

\[ (Ay \cdot y) \geq a|y|^2. \quad (1.117) \]

Consider the function

\[ G(x) = \frac{1}{2}(Ax \cdot x) - (b \cdot x), \quad (1.118) \]

defined for all \( x \in \mathbb{R}^n \). Note that

\[ G(x) \geq a|x|^2 - |b|x, \]

and thus \( G(x) \to +\infty \) as \( |x| \to +\infty \). It follows that \( G(x) \) is bounded from below and attains its minimum at some point \( \bar{x} = (\bar{x}_1, \ldots, \bar{x}_n) \in \mathbb{R}^n \). Using basic calculus we find that this point must satisfy the equations

\[ \sum_{j=1}^{n} A_{ij} \bar{x}_j = b_i, \quad i = 1, \ldots, n, \quad (1.119) \]

which, of course, is nothing but (1.116). Therefore, solving equation (1.116) is exactly equivalent to finding the minimal point of the function \( G(x) \). The latter might be a much easier problem in many circumstances, especially so since the function \( G(x) \) is convex, thus it has a unique minimum.

This idea can be generalized to many PDEs, in particular, to the Poisson equation. We define the energy functional

\[ I[w] = \int_U \left( \frac{1}{2} |\nabla w|^2 - wf \right) dx, \quad (1.120) \]

and the class of admissible functions

\[ \mathcal{A} = \{ w \in C^2(\bar{U}) : w = g \text{ on } \partial U \}. \]

**Theorem 1.13** A function \( u \in C^2(\bar{U}) \) solves the boundary value problem

\[ -\Delta u = f \text{ in } U \]
\[ u = g \text{ on } \partial U \quad (1.121) \]

if and only if \( u \in \mathcal{A} \) and

\[ I[u] = \min_{w \in \mathcal{A}} I[w]. \quad (1.122) \]
Proof. Let us assume that $u$ solves (1.121), take $w \in \mathcal{A}$, multiply (1.121) by $(u - w)$ and integrate:
\[
\int_U (-\Delta u - f)(u - w) dx = 0.
\]
Integrating by parts using the Green’s formula gives
\[
\int_U (\nabla u \cdot \nabla (u - w) - f(u - w)) dx = 0.
\]
We used here the fact that $u - w = 0$ on $\partial U$ to kill the boundary terms in Green’s formula. It follows that
\[
\int_U (|\nabla u|^2 - fu) dx = \int_U (\nabla u \cdot \nabla w - fw) dx.
\]
(1.123)
Now comes the crucial trick: note that
\[
|\nabla u \cdot \nabla w| \leq \frac{1}{2} |\nabla u|^2 + \frac{1}{2} |\nabla w|^2.
\]
Using this in (1.123) leads to
\[
\int_U (|\nabla u|^2 - fu) \leq \int_U (\frac{1}{2} |\nabla u|^2 + \frac{1}{2} |\nabla w|^2 - fw) dx,
\]
(1.124)
hence,
\[
\int_U (\frac{1}{2} |\nabla u|^2 - fu) \leq \int_U (\frac{1}{2} |\nabla w|^2 - fw) dx,
\]
(1.125)
which is nothing but $I[u] \leq I[w]$. Therefore, if $u$ solves the boundary value problem (1.121) then it minimizes the functional $I[w]$ over $w \in \mathcal{A}$.

To show the other direction, let $u$ be a minimizer of $I[w]$ over $\mathcal{A}$. Take a function $v$ that is smooth in $U$ and vanishes on the boundary $\partial U$. Consider the increment of $I[w]$ in the direction $v$:
\[
r(s) = I[u + sv].
\]
Then, the function $u + sv$ is in $\mathcal{A}$, and, as $u$ minimizes $I[w]$ over $\mathcal{A}$, we should have $r(s) \geq r(0)$ for all $s \in \mathbb{R}$. The function $r(s)$ is a quadratic function of $s$:
\[
r(s) = \int_U \left( \frac{1}{2} |\nabla u + s\nabla v|^2 - (u + sv)f \right) dx
\]
\[= \int_U \left( \frac{1}{2} |\nabla u|^2 - uf \right) dx + s \int_U (\nabla u \cdot \nabla v - vf) dx + \frac{s^2}{2} \int_U |\nabla v|^2 dx.
\]
As $r(s)$ attains its minimum at $s = 0$, we have
\[
\int_U (\nabla u \cdot \nabla v - vf) dx = 0.
\]
(1.126)
Integrating by parts, as $v = 0$ on $\partial U$, gives
\[
\int_U (-\Delta u - f)v dx = 0.
\]
Since this identity holds for all smooth functions $v$ that vanish at the boundary $\partial U$, it follows that $u$ satisfies

$$-\Delta u = f,$$

and this finishes the proof – the boundary condition $u = g$ on $\partial U$ is satisfied automatically since $u \in A$. □

1.3.5 Image denoising as a variational problem

The image denoising problem is to find a function $u$ that is close to a recorded image $f$ but has "less noise". One approach to this problem is to consider the functional

$$J(w) = \lambda \int_U |\nabla w|^2 dx + \int_U (w - f)^2 dx. \tag{1.127}$$

Here, we think of $f$ as a noisy measured image, $U$ is the domain of the recording sensor, $\lambda$ is a small parameter, and we are looking for a function $w$ that is close to $f$ but is "reasonably smooth" – this is why the gradient term appears in (1.133). It is natural to assume that the normal derivative of $w$ vanishes at the image edges, so that our goal is to find a minimizer of $J(w)$ over the set

$$A = \{ u \in C^2(\Omega) \cap C(\bar{\Omega}) : \frac{\partial u}{\partial \nu} = 0 \text{ on } \partial \Omega \}.$$ 

An alternative is to minimize $J(w)$ over the set

$$A_g = \{ u \in C^2(\Omega) \cap C(\bar{\Omega}) : u = g \text{ on } \partial \Omega \},$$

where, $g$ is a smoothed version of $f$ on the boundary – the question how we smooth $f$ on the boundary is separate, and we do not touch upon it here. Let us assume that $w$ is a minimizer of $J(w)$ over $A$ and variate $J$ over $A$: take a smooth function $\eta(x)$ such that $\partial \eta/\partial \nu = 0$ on $\partial \Omega$ and compute

$$J(w + s\eta) = \lambda \int_\Omega |\nabla w + s\nabla \eta|^2 dx + \int_\Omega (w + s\eta - f)^2 dx = \lambda \int_\Omega |\nabla w|^2 dx + \int_\Omega (w - f)^2 dx + 2s \lambda \int_\Omega (\nabla \eta \cdot \nabla w) dx + s^2 \lambda \int_\Omega |\nabla \eta|^2 dx + 2s \int_\Omega (w - f) \eta dx + s^2 \int_\Omega \eta^2 dx. \tag{1.128}$$

We see that $J(w + s\eta)$ attains its minimum at $s = 0$ if

$$\lambda \int_\Omega (\nabla \eta \cdot \nabla w) dx + \int_\Omega (w - f) \eta dx = 0. \tag{1.129}$$

As $\partial w/\partial \nu = 0$ on the boundary, this can be re-written as

$$-\lambda \int_\Omega (\Delta w) \eta dx + \int_\Omega (w - f) \eta dx = 0. \tag{1.130}$$

Now, for $w$ to be a minimum of $J(w)$ over the admissible set $A$, the function

$$f(s) = J(w + s\eta),$$
should attain its minimum at \( s = 0 \) for all smooth test functions \( \eta \) that vanish on \( \partial \Omega \), and (1.130) should hold for all such \( \eta(x) \), hence \( w \) should be the solution of the boundary value problem

\[
-\lambda \Delta w + w = f \text{ in } \Omega, \tag{1.131}
\]
\[
\frac{\partial w}{\partial \nu} = 0 \text{ on } \partial \Omega.
\]

**Exercise 1.14** Show that if we were minimizing \( J(w) \) over the set \( A_g \) we would have arrived at the problem

\[
-\lambda \Delta w + w = f \text{ in } \Omega, \tag{1.132}
\]
\[
w = g \text{ on } \partial \Omega.
\]

The truth is that this denoising method oversmoothes and is terrible at preserving edges in an image even for small values of \( \lambda \), but it is a legitimate first attempt at denoising. A much better functional to use is

\[
\tilde{J}(w) = \lambda \int_U |\nabla w| dx + \int_U (w - f)^2 dx. \tag{1.133}
\]

However, minimizing \( \tilde{J}(w) \) leads to a rather complicated nonlinear PDE, so we will not consider this problem here.

### 1.4 The heat equation

#### 1.4.1 Another probabilistic interlude

We will now consider the heat (or diffusion) equation

\[
\frac{\partial u}{\partial t} - \Delta u = 0. \tag{1.134}
\]

Usually it is obtained from a balance of heat or concentration that assumes that the flux of heat is \( F = -\nabla u \), where \( u \) is the temperature – heat flows from hot to cold. Here, we derive it informally starting with a probabilistic model.

Consider a lattice on the real line of mesh size \( h \): \( x_n = nh \). Let \( X(t) \) be a random walk on this lattice that starts at some point \( x \), and after a delay \( \tau \) jumps to the left or right with probability \( 1/2 \): \( P(X(\tau) = x + h) = P(X(\tau) = x - h) = 1/2 \). Then it waits again for time \( \tau \), and again jumps to the left or right with probability \( 1/2 \), and so on. Let \( S \) be a subset of the real line and define \( u(t, x) = P(X(t) \in S | X(0) = x) \) – this is the probability that at a time \( t > 0 \) the particle is inside the set \( S \) given that it started at the point \( x \) at time \( t = 0 \).

Let us derive an equation for \( u(t, x) \). Since the process ”starts anew” after every jump we have the relation

\[
P(X(t) \in S | X(0) = x) = \frac{1}{2} P(X(t - \tau) \in S | X(0) = x + h) + \frac{1}{2} P(X(t - \tau) \in S | X(0) = x - h),
\]

which is

\[
u(t, x) = \frac{1}{2} u(t - \tau, x + h) + \frac{1}{2} u(t - \tau, x - h). \tag{1.135}
\]
Let us assume that $\tau$ and $h$ are small and use Taylor’s formula in the right side above. Then (1.135) becomes:

$$u(t,x) = \frac{1}{2} \left[ u(t,x) - \tau \frac{\partial u(t,x)}{\partial t} + h \frac{\partial u(t,x)}{\partial x} + \frac{h^2 \partial^2 u(t,x)}{2 \partial x^2} + \frac{\tau^2 \frac{\partial^2 u(t,x)}{\partial t^2}}{2} - \tau h \frac{\partial^2 u(t,x)}{\partial x \partial t} \right]$$

$$+ \frac{1}{2} \left[ u(t,x) - \tau \frac{\partial u(t,x)}{\partial t} - h \frac{\partial u(t,x)}{\partial x} + \frac{h^2 \partial^2 u(t,x)}{2 \partial x^2} + \frac{\tau^2 \frac{\partial^2 u(t,x)}{\partial t^2}}{2} + \tau h \frac{\partial^2 u(t,x)}{\partial x \partial t} \right] + \ldots,$$

which is

$$\tau \frac{\partial u}{\partial t} = \frac{h^2 \partial^2 u}{2 \partial x^2} + \frac{\tau^2 \frac{\partial^2 u(t,x)}{\partial t^2}}{2} + \ldots$$

In order to get a non-trivial balance we set $\tau = h^2$. Then the term involving $u_{tt}$ in the right side is smaller than the rest and in the leading order we obtain

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

which is the diffusion equation (we could get rid of the factor of 1/2 if we took $\tau = h^2/2$ but probabilists do not like that). It is supplemented by the initial condition

$$u(0,x) = \begin{cases} 1, & \text{if } x \in S \\ 0, & \text{if } x \notin S. \end{cases}$$

More generally, we can take a bounded function $f(x)$ defined on the real line and set

$$v(t,x) = \mathbb{E}\{f(X(t)) | X(0) = x\}.$$

Essentially an identical argument shows that if $\tau = h^2$ then in the limit $h \to 0$ we get the following Cauchy problem for $v(t,x)$:

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

$$v(0,x) = f(x).$$

What should we expect for the solutions of the Cauchy problem given this informal probabilistic representation? First, it should preserve positivity: if $f(x) \geq 0$ for all $x \in \mathbb{R}$, we should have $u(t,x) \geq 0$ for all $t > 0$ and $x \in \mathbb{R}$. Second, the maximum principle should hold: if $f(x) \leq M$ for all $x \in \mathbb{R}$, then we should have $u(t,x) \leq M$ for all $t > 0$ and $x \in \mathbb{R}$ because the expected value of a quantity can not exceed its maximum. We should also expect that $\max_{x \in \mathbb{R}} v(t,x)$ decays in time, at least if $f(x)$ is compactly supported – this is because the random walk will tend to spread around and at large times the probability to find it on the set where $f(x)$ does not vanish, is small.

### 1.4.2 The heat kernel and the Cauchy problem

The fundamental solution $\Phi_n(x)$ of the Laplace equation has the property that the convolution

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

36
gives a solution of the Poisson equation
\[-\Delta u = f(x) \text{ in } \mathbb{R}^n.\]

Let us now try to find the “moral equivalent” of the fundamental solution for the heat equation. More precisely, we will look for a function \( G(t, x) \) so that the convolution
\[ u(t, x) = \int G(t, x - y) f(y) dy \]  
(1.138)
gives a solution of the Cauchy problem (the initial value problem):
\[ \begin{align*}
  u_t &= \Delta u, \ t > 0, \ x \in \mathbb{R}^n, \\
  u(0, x) &= f(x).
\end{align*} \]  
(1.139)

Let us first look for some symmetries that the function \( G(t, x) \) has to satisfy. The key observation is that if \( u(t, x) \) is a solution to (1.139), then for all \( \lambda > 0 \) the rescaled function
\[ u_\lambda(t, x) = u(\lambda^2 t, \lambda x) \]
also solves the heat equation, but with the rescaled initial data
\[ u_\lambda(0, x) = u(0, \lambda x) = f(\lambda x). \]

It follows that for any initial data \( f(x) \) and any \( \lambda > 0 \) we should have
\[ u(\lambda^2 t, \lambda x) = u_\lambda(t, x) = \int G(t, x - y) f(\lambda y) dy. \]

Using expression (1.138) for \( u(\lambda^2 t, \lambda t) \) leads to
\[ \int G(\lambda^2 t, \lambda x - y) f(y) dy = \int G(t, x - y) f(\lambda y) dy = \frac{1}{\lambda^n} \int G(t, x - \frac{y}{\lambda}) f(y) dy. \]  
(1.140)

This identity holds for any continuous function \( f(y) \) with compact support, hence \( G(t, x) \) has to satisfy
\[ G(\lambda^2 t, \lambda x - y) = \frac{1}{\lambda^n} G(t, x - \frac{y}{\lambda}), \]  
(1.141)
for all \( \lambda > 0, t > 0 \) and \( x, y \in \mathbb{R}^n \). Denoting \( z = x - y/\lambda \), we get an equivalent form of (1.141): for any \( \lambda > 0, t > 0 \) and \( z \in \mathbb{R}^n \) we have
\[ G(\lambda^2 t, \lambda z) = \frac{1}{\lambda^n} G(t, z). \]  
(1.142)

Let us choose \( \lambda = 1/\sqrt{t} \) then (1.142) implies
\[ G(1, \frac{z}{\sqrt{t}}) = t^{n/2} G(t, z). \]  
(1.143)
In other words, the function \( G(t, z) \) has to be of the form
\[
G(t, z) = \frac{1}{t^{n/2}} v\left(\frac{z}{\sqrt{t}}\right). 
\]  (1.144)

Here we have denoted \( v(z) = G(1, z) \). This means that \( G(t, z) \) is self-similar: its shape at different times can be obtained by simple rescaling of the "shape" of the function \( v(y), y \in \mathbb{R}^n \).

Let us make a more general self-similar ansatz:
\[
u(t, x) = \frac{1}{t^m} v\left(\frac{x}{t^{\beta}}\right),
\]  (1.145)
and see for which \( m \) and \( \beta \) we can find a self-similar solution of the heat equation
\[
u_t = \Delta \nu. 
\]  (1.146)

We insert the ansatz (1.145) into (1.146) and compute, with \( y = x/t^{\beta} \):
\[
-m v(y) - \frac{\beta}{t^{m+\beta+1}} x \cdot \nabla v(y) - \frac{1}{t^{m+2\beta}} \Delta v(y) = 0,
\]  or
\[
-m v(y) - \frac{\beta}{t^{m+1}} y \cdot \nabla v(y) - \frac{1}{t^{m+2\beta}} \Delta v(y) = 0,
\]  (1.147)

Now, for (1.147) to be true for all \( t > 0 \) and \( y \in \mathbb{R}^n \), we need this equation to involve only the variable \( y \) – this forces us to take \( \beta = 1/2 \). With this choice of \( \beta \), equation (1.147) becomes
\[
m v(y) + \frac{y}{2} \cdot \nabla v(y) + \Delta v(y) = 0
\]  (1.148)

In order to simplify further we assume that \( v \) is radial. Actually, a good exercise is to convince yourself that the radial symmetry of the heat equation implies that the heat kernel, if it exists, has to be radial. In other words, this assumption should be automatically satisfied, and we must have \( v(y) = w(|y|) \), with some function \( w(r), r > 0 \), turning (1.148) into
\[
w''(r) + \frac{n-1}{r} w' + \frac{r}{2} w' + mw = 0. 
\]  (1.149)

Let us add the boundary condition \( w(0) = 1, w'(0) = 0 \) to (1.149), and look for \( w(r) > 0 \) that decays rapidly as \( r \to +\infty \) and is positive for all \( r > 0 \). Multiplying (1.149) by \( r^{n-1} \) gives
\[
(r^{n-1} w'(r))' + \left(\frac{r^n}{2} w\right)' + r^{n-1} \left(m - \frac{n}{2}\right) w = 0.
\]  (1.150)

The boundary conditions at \( r = 0 \) and rapid decay of \( w(r) \) as \( r \to +\infty \) imply that
\[
\int_0^\infty (r^n w'(r))' dr = \int_0^\infty \left(\frac{r^{n+1}}{2} w\right)' = 0.
\]

It follows that
\[
\left(m - \frac{n}{2}\right) \int_0^\infty r^n w(r) dr = 0.
\]
As we are looking for $w(r) > 0$, it forces the value $m = n/2$. With this choice of $m$, (1.150) can be solved giving a solution with the properties we have required: integrating this equation once gives
\[ r^{n-1}w'(r) + \frac{r^n}{2}w = 0, \tag{1.151} \]
so
\[ w'(r) = -\frac{r}{2}w, \]
and
\[ w(r) = be^{-r^2/4}, \tag{1.152} \]
with an arbitrary constant $b > 0$ gives a solution. Therefore, we obtain the following family of positive solutions to the heat equation:
\[ u(t, x) = \frac{b}{t^{n/2}}e^{-|x|^2/(4t)}, \quad b > 0. \tag{1.153} \]

This motivates the following

**Definition 1.15**  The function
\[ G(t, x) = \frac{1}{(4\pi t)^{n/2}}e^{-|x|^2/(4t)}, \quad t > 0, \quad x \in \mathbb{R}^n, \tag{1.154} \]
is called the heat kernel.

Note that
\[ \int_{\mathbb{R}^n} G(t, x)dx = 1 \text{ for all } t > 0. \tag{1.155} \]
This is because
\[ \int_{\mathbb{R}^n} G(t, x)dx = \frac{1}{(4\pi t)^{n/2}} \int_{\mathbb{R}^n} e^{-|x|^2/(4t)}dx = \frac{1}{(4\pi)^{n/2}} \int_{\mathbb{R}^n} e^{-z^2/4}dz = \frac{1}{(4\pi)^{n/2}} \int_{\mathbb{R}^n} e^{-z^2/4}dz = \left( \frac{1}{\sqrt{4\pi}} \int_{-\infty}^{\infty} e^{-s^2/4}ds \right)^n \]
while
\[ \left( \frac{1}{\sqrt{4\pi}} \int_{-\infty}^{\infty} e^{-s^2/4}ds \right)^2 = \frac{1}{4\pi} \int_{\mathbb{R}^n} e^{-(x^2+y^2)^2/4}dxdy = \frac{1}{4\pi} \int_{0}^{2\pi} \int_{0}^{\infty} e^{-r^2/4}rdrd\phi = \frac{1}{2} \int_{0}^{\infty} re^{-r^2/4}dr = \int_{0}^{\infty} e^{-s}ds = 1. \]

Actually, identity (1.155) is the reason why we chose the prefactor to be $(4\pi)^{-n/2}$ in the definition of the heat kernel.

Notice that if $u(t, x)$ is a solution of the heat equation, then so is the shifted function $u(t, x - y)$ for any $y \in \mathbb{R}^n$ fixed. It follows that the function
\[ v(t, x) = \int_{\mathbb{R}^n} G(t, x - y)f(y)dy, \quad t > 0, \quad x \in \mathbb{R}^n, \tag{1.156} \]
is also a solution of the heat equation, provided that the function \( f(y) \) is such that we can differentiate under the integral sign above. This is true, in particular, if \( f(y) \) is bounded and continuous. Moreover, under these assumptions we can differentiate \( v(t, x) \) as many times as we wish — that is, \( v(t, x) \) is infinitely differentiable even if \( f(y) \) is just continuous and bounded.

Let us now see what happens to \( v(t, x) \) as \( t \downarrow 0 \) — note that (1.156) defines it only for \( t > 0 \). To this end, we first note that
\[
\frac{1}{(4\pi)^{n/2}} \int_{\mathbb{R}^n} e^{-z^2/4} dz = 1, \tag{1.157}
\]
which can be seen by taking \( t = 1 \) in (1.155). Now, let \( t > 0 \) be small and let us make a change of variables \( z = (x - y)/\sqrt{t} \) in (1.156):
\[
v(t, x) = \frac{1}{(4\pi t)^{n/2}} \int_{\mathbb{R}^n} e^{-|x - y|^2/(4t)} f(y) dy = \frac{1}{(4\pi)^{n/2}} \int_{\mathbb{R}^n} e^{-z^2/4} f(x - z\sqrt{t}) dz. \tag{1.158}
\]
As \( f \) is continuous, we have
\[
f(x - z\sqrt{t}) \to f(x),
\]
for each \( x \in \mathbb{R}^n \) and \( z \in \mathbb{R}^n \) fixed. Since \( f \) is also globally bounded, \( |f(x)| \leq M \) for all \( x \in \mathbb{R}^n \), we can use Lebesgue dominated convergence theorem to conclude from (1.158) that
\[
v(t, x) = \frac{1}{(4\pi)^{n/2}} \int_{\mathbb{R}^n} e^{-z^2/4} f(x - z\sqrt{t}) dz \to f(x) \frac{1}{(4\pi)^{n/2}} \int_{\mathbb{R}^n} e^{-z^2/4} dz = f(x). \tag{1.159}
\]
We used identity (1.157) in the last step above. We summarize this discussion as follows.

**Theorem 1.16** Let \( f(x) \) be a continuous function in \( \mathbb{R}^n \) and assume that there exists a constant \( M > 0 \) so that \( |f(x)| \leq M \) for all \( x \in \mathbb{R}^n \). Then the function \( v(t, x) \) defined by (1.156) is infinitely differentiable for all \( t > 0 \) and \( x \in \mathbb{R}^n \) and, in addition, satisfies the heat equation
\[
\frac{\partial v}{\partial t} - \Delta v = 0, \quad t > 0, \ x \in \mathbb{R}^n,
\]
as well as the initial condition \( v(0, x) = f(x) \).

### 1.4.3 Duhamel’s principle for the heat equation

Let us now consider the inhomogeneous heat equation with the zero initial condition
\[
\begin{align*}
\frac{\partial v}{\partial t} - \Delta v &= f(t, x), \quad t > 0, \ x \in \mathbb{R}^n, \\
v(0, x) &= 0, \quad x \in \mathbb{R}^n.
\end{align*} \tag{1.160}
\]
Duhamel’s principle, as discussed at the end of Section 1.2 says that solution of (1.160) can be found as follows: for every fixed \( s \in [0, t] \) solve the following Cauchy problem for the function \( u(t, x; s) \), defined for \( t \geq s \):
\[
\begin{align*}
\frac{\partial u(t, x; s)}{\partial t} - \Delta u(t, x; s) &= 0, \quad t \geq s, \ x \in \mathbb{R}^n, \\
u(t = s, x; s) &= f(s, x), \quad x \in \mathbb{R}^n.
\end{align*} \tag{1.161}
\]
A solution of this problem is
\[ u(t, x; s) = \int_{\mathbb{R}^n} G(t - s, x - y) f(s, y) dy. \]

The Duhamel’s principle says that solution of (1.160) is given by
\[ v(t, x) = \int_0^t u(t, x; s) ds = \int_0^t \int_{\mathbb{R}^n} G(t - s, x - y) f(s, y) dy ds. \] (1.162)

Let us verify that this is, indeed, the case. If we could differentiate under the integral sign that would be simple:
\[ \frac{\partial v(t, x)}{\partial t} - \Delta v(t, x) = u(t, x; t) + \int_0^t \left( \frac{\partial u(t, x; s)}{\partial t} - \Delta u(t, x; s) \right) ds = f(t, x). \] (1.163)

The subtlety here is that \( G(t - s, x - y) \) is singular at \( s = t \) and we need to justify this formal procedure. This is done very similarly to what we did in the proof of Theorem 1.1 so we will just briefly describe how this can be done. Let us rewrite \( v(t, x) \) given by (1.162) as
\[ v(t, x) = \int_0^t \int_{\mathbb{R}^n} G(s, y) f(t - s, x - y) dy ds. \] (1.164)

Then, if \( f \) is sufficiently smooth, we can argue as in the proof of the aforementioned theorem to get
\[ \frac{\partial v(t, x)}{\partial t} - \Delta v(t, x) = \int_0^t \int_{\mathbb{R}^n} G(s, y) \left[ \frac{\partial f(t - s, x - y)}{\partial t} - \Delta_x f(t - s, x - y) \right] dy ds \\
+ \int_{\mathbb{R}^n} G(t, y) f(0, x - y) dy. \] (1.165)

The potential trouble point in the first integral in the right side of (1.165) is \( s = 0 \) where \( G(s, y) \) is singular. Hence, we take a small \( \varepsilon > 0 \) and, using also the fact that
\[ \Delta_x f(t - s, x - y) = \Delta_y f(t - s, x - y), \quad \frac{\partial f(t - s, x - y)}{\partial t} = -\frac{\partial f(t - s, x - y)}{\partial s}, \]
we split the integral in (1.165) as
\[ \frac{\partial v(t, x)}{\partial t} - \Delta v(t, x) = \int_\varepsilon^t \int_{\mathbb{R}^n} G(s, y) \left[ -\frac{\partial f(t - s, x - y)}{\partial s} - \Delta_y f(t - s, x - y) \right] dy ds \\
+ \int_0^\varepsilon \int_{\mathbb{R}^n} G(s, y) \left[ -\frac{\partial f(t - s, x - y)}{\partial s} - \Delta_y f(t - s, x - y) \right] dy ds + \int_{\mathbb{R}^n} G(t, y) f(0, x - y) dy \\
= I_\varepsilon + J_\varepsilon + K. \] (1.166)

The second term in the right side above is small when \( \varepsilon \to 0 \): let
\[ M_f = \max_{t \in \mathbb{R}, x \in \mathbb{R}^n} \left| \frac{\partial f(t, x)}{\partial x} \right| + \max_{t \in \mathbb{R}, x \in \mathbb{R}^n} |\Delta f(t, x)|, \]
then
\[ |J_\varepsilon| \leq M_f \int_0^\varepsilon \int_{\mathbb{R}^n} G(s, y) ds dy = \varepsilon M_f, \]
because of (1.155). In order to evaluate \( I_\varepsilon \) we integrate by parts in \( s \) and \( y \):

\[
I_\varepsilon = \int_\varepsilon^t \int_{\mathbb{R}^n} G(s, y) \left[ -\frac{\partial f(t-s, x-y)}{\partial s} - \Delta_y f(t-s, x-y) \right] dy ds
\]
\[
= \int_\varepsilon^t \int_{\mathbb{R}^n} \left[ \frac{\partial G(s, y)}{\partial s} - \Delta_y G(s, y) \right] f(t-s, x-y) dy ds
\]
\[
- \int_{\mathbb{R}^n} [G(t, y)f(0, x-y) - G(\varepsilon, y)f(t-\varepsilon, x-y)] dy.
\]
The term in the middle line above vanishes since
\[
\frac{\partial G(s, y)}{\partial s} - \Delta_y G(s, y) = 0.
\]
Moreover, the first term in the last line is simply \((-K)\) (with \( K \) as in (1.166)), hence
\[
I_\varepsilon + K = \int_{\mathbb{R}^n} G(\varepsilon, y)f(t-\varepsilon, x-y) dy = \frac{1}{(4\pi\varepsilon)^{n/2}} \int_{\mathbb{R}^n} e^{-|y|^2/(4\varepsilon)} f(x-y) dy.
\]
With the change of variables \( z = y/\sqrt{\varepsilon} \) this becomes
\[
I_\varepsilon + K = \frac{1}{(4\pi)^{n/2}} \int_{\mathbb{R}^n} e^{-|z|^2/4} f(t-\varepsilon, x-\sqrt{\varepsilon}z) dz.
\]
Therefore, as \( \varepsilon \downarrow 0 \) this converges to
\[
\frac{1}{(4\pi)^{n/2}} \int_{\mathbb{R}^n} e^{-|z|^2/4} f(t,x) dz = f(t,x).
\]
We conclude that Duhamel’s formula, indeed, gives the solution of the inhomogeneous problem (1.160).

1.4.4 The maximum principle for the heat equation

Let us now prove the maximum principle for the heat equation, first in a bounded domain. Let \( U \subset \mathbb{R}^n \) be a bounded domain and \( T > 0 \). We define
\[
U_T = \{(t, x) : x \in U, \ 0 \leq t \leq T\},
\]
and
\[
\Gamma_T = \{(t, x) : t = 0 \text{ and } x \in U, \text{ or } 0 \leq t \leq T \text{ and } x \in \partial U\}.
\]
Note that the definition of the "parabolic boundary" \( \Gamma_T \) does not include the final time \( t = T \) – it looks in time-space as a cylinder without the top.
Theorem 1.17 Let the function \( u(t,x) \) satisfy the heat equation
\[
\frac{\partial u}{\partial t} - \Delta u = 0
\]
in \( U_T \). Then \( u \) achieves its maximum and minimum over \( U_T \) on the parabolic boundary \( \Gamma_T \).

In other words, \( u \) attains its maximum either at some point \( x \in U \) at the initial time \( t = 0 \), or, if it attains its maximum at some point \( (t_0,x_0) \) with \( t_0 > 0 \) then \( x_0 \) has to belong to the boundary \( \partial U \).

Proof. Take \( \varepsilon > 0 \) and consider the function
\[
v(t,x) = u(t,x) - \varepsilon t.
\]
The function \( v(t,x) \) satisfies
\[
\frac{\partial v}{\partial t} - \Delta v = -\varepsilon.
\]
(1.167)

Consider the domain
\[
\bar{U}_T = U_T \bigcup \{(T,x) : x \in U\},
\]
which is the union of \( U_T \) and the "top" of the parabolic cylinder. The function \( v(t,x) \) must attain its maximum over the set \( \bar{U}_T \) at some point \( (t_0,x_0) \in \bar{U}_T \). We claim that this point has to lie on the parabolic boundary \( \Gamma_T \). Indeed, if \( 0 < t_0 < T \) and \( x_0 \) is not on the boundary \( \partial U \), then the point \( (t_0,x_0) \) is an interior maximum of \( v(t,x) \) and as such should satisfy
\[
\frac{\partial v(t_0,x_0)}{\partial t} = 0, \quad \Delta v(t_0,x_0) \leq 0,
\]
which is impossible because of (1.167). On the other hand, if \( t_0 = T \) and \( x_0 \) is an interior point of \( U \), and \( u \) attains its maximum over \( \bar{U} \) at this point, then we should have
\[
\frac{\partial v(t_0,x_0)}{\partial t} \geq 0, \quad \Delta v(t_0,x_0) \leq 0,
\]
which, once again contradicts (1.167). Hence, the function \( v \) attains its maximum over \( \bar{U}_T \) at a point \( (t_0,x_0) \) that belongs to \( \Gamma_T \). It means that
\[
\max_{(t,x) \in \bar{U}_T} v(t,x) = \max_{(t,x) \in \Gamma_T} v(t,x) = \max_{(t,x) \in \Gamma_T} u(t,x).
\]
However, we also have
\[
\max_{(t,x) \in \bar{U}_T} u(t,x) \leq \varepsilon T + \max_{(t,x) \in \bar{U}_T} v(t,x).
\]
Putting the last two inequalities together gives
\[
\max_{(t,x) \in \bar{U}_T} u(t,x) \leq \varepsilon T + \max_{(t,x) \in \Gamma_T} u(t,x).
\]
As \( \varepsilon > 0 \) is arbitrary, it follows that
\[
\max_{(t,x) \in \bar{U}_T} u(t,x) \leq \max_{(t,x) \in \Gamma_T} u(t,x),
\]
and the proof is complete. □

The situation with the maximum principle in the whole space is slightly more delicate. Let us just state the result.
Theorem 1.18 Let \( u(t, x) \) be a smooth solution of the Cauchy problem

\[
\begin{align*}
\frac{\partial u}{\partial t} - \Delta u &= 0, \quad t > 0, \quad x \in \mathbb{R}^n, \\
u(0, x) &= f(x).
\end{align*}
\] (1.168)

Assume that \( u \) satisfies the estimate

\[
u(t, x) \leq A e^{a|x|^2}, \quad \text{for all } x \in \mathbb{R}^n \text{ and } 0 \leq t \leq T,
\] (1.169)

with some constants \( a \) and \( A \). Then we have

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

The proof of this theorem is more technical than in a bounded domain but the idea is similar in spirit so we not present it here. We just mention that there are solutions of the homogeneous Cauchy problem

\[
\begin{align*}
\frac{\partial u}{\partial t} - \Delta u &= 0, \quad t > 0, \quad x \in \mathbb{R}^n, \\
u(0, x) &= 0
\end{align*}
\] (1.170)

that are not identically equal to zero and grow very rapidly at infinity. The role of assumption (1.169) is to preclude those.

1.4.5 Uniqueness

Let us now go back to the setting of Theorem 1.17.

Theorem 1.19 Let \( g \) be a continuous function on the parabolic boundary \( \Gamma_T \) and let \( f \) be continuous in \( U_T \). Then there exists at most one smooth solution to the initial boundary value problem

\[
\begin{align*}
\frac{\partial u}{\partial t} - \Delta u &= f, \quad \text{in } U_T \\
u &= g \text{ on } \Gamma_T.
\end{align*}
\] (1.171)

Note that the condition \( u = g \) on \( \Gamma_T \) prescribes both the initial data for \( u \) at the time \( t = 0 \) and the boundary data along \( \partial U \) at times \( t > 0 \).

**Proof.** This follows immediately from the maximum principle for bounded domains (Theorem 1.17). Indeed, if \( u_1 \) and \( u_2 \) solve (1.171) then the difference \( v = u_1 - u_2 \) satisfies

\[
\begin{align*}
\frac{\partial v}{\partial t} - \Delta v &= 0, \quad \text{in } U_T \\
v &= 0 \text{ on } \Gamma_T.
\end{align*}
\] (1.172)

The maximum principle now implies that \( v = 0 \). \( \Box \)
The same uniqueness result holds for solutions in the whole space: there exists at most one solution of the initial value problem
\[
\frac{\partial u}{\partial t} - \Delta u = 0, \quad t > 0, \quad x \in \mathbb{R}^n, \tag{1.173}
\]
\[
u(0, x) = f(x)
\]
that in addition satisfies the estimate
\[
u(t, x) \leq Ae^{a|x|^2}, \quad \text{for all } x \in \mathbb{R}^n \text{ and } 0 \leq t \leq T, \tag{1.174}
\]
with some constants \(a\) and \(A\). This follows from Theorem 1.18.

**Uniqueness by the energy method**

Let us now prove Theorem 1.19 without using the maximum principle, by what is known as the energy method. We need to show that the only solution of
\[
\frac{\partial v}{\partial t} - \Delta v = 0, \quad \text{in } U_T \tag{1.175}
\]
\[
v = 0 \text{ on } \Gamma_T,
\]
is \(v \equiv 0\). Let us define the energy
\[
E(t) = \int_U v^2(t, x)dx.
\]
We will show that any solution of the heat equation, regardless of the initial condition, satisfies
\[
\frac{dE}{dt} \leq 0, \tag{1.176}
\]
hence \(E(t) \leq E(0)\) for all \(t \geq 0\). However, if \(v(0, x) = 0\) then \(E(0) = 0\), so that \(E(t) \leq 0\) for all \(t \geq 0\). As \(E(t)\) is nonnegative by definition, it follows that \(E(t) = 0\) for all \(t \geq 0\), which means that \(v(t, x) \equiv 0\).

Let us now show that energy inequality (1.176) holds. We will do this for a general solution of the Cauchy problem with zero boundary conditions on \(\partial U\):
\[
\frac{\partial u}{\partial t} - \Delta u = 0, \quad \text{for } 0 \leq t \leq T, \quad x \in U, \tag{1.177}
\]
\[
u(t, x) = 0 \quad \text{for } 0 \leq t \leq T, \quad x \in \partial U,
\]
\[
u(0, x) = g(x) \quad \text{for } x \in U.
\]
The corresponding energy is
\[
E(t) = \int_U u^2(t, x)dx.
\]
We compute:
\[
\frac{dE}{dt} = 2 \int_U u(t, x) \frac{\partial u(t, x)}{\partial t} dx = 2 \int_U u(t, x) \Delta u(t, x) dx.
\]
Integration by parts gives

\[ \frac{dE}{dt} = 2 \int_{\partial U} u(t, x) \frac{\partial u(t, x)}{\partial \nu} \, dx - 2 \int_U \nabla u \cdot \nabla u \, dx. \]

The boundary condition \( u = 0 \) on \( \partial U \) implies that the above is

\[ \frac{dE}{dt} = -2 \int_U |\nabla u(t, x)|^2 \, dx \leq 0, \]

hence (1.176) holds.

1.4.6 Regularity of solutions of the heat equation

Let us now discuss how regular solutions of the heat equation are. For simplicity, we will consider only solutions in the whole space (generalization to bounded domains is not very painful), and we will always assume that they obey the bounds (1.174) in order to ensure uniqueness. Let \( u(t, x) \) satisfy the Cauchy problem

\[ \begin{align*}
\frac{\partial u}{\partial t} - \Delta u &= 0, \quad t > 0, \quad x \in \mathbb{R}^n, \\
 u(0, x) &= f(x).
\end{align*} \tag{1.178} \]

It is given then explicitly as

\[ u(t, x) = \int_{\mathbb{R}^n} G(t, x-y) f(y) \, dy = \frac{1}{(4\pi t)^{n/2}} \int_{\mathbb{R}^n} e^{-|x-y|^2/(4t)} f(y) \, dy. \tag{1.179} \]

The first fact to observe is that if, say, \( f(x) \) is continuous and of compact support, we can differentiate under the integral sign with respect to both \( t \) and \( x \) arbitrarily many times. This is because such differentiations will lead to expressions of the sort

\[ \int_{\mathbb{R}^n} p(t, x, y) e^{-|x-y|^2/(4t)} f(y) \, dy, \]

where \( p(t, x, y) \) is some polynomial in \( x \) and \( y \) with coefficients that are rational functions of \( t \) that is regular for \( t > 0 \). Such integrals converge, which justifies the differentiation under the integral sign (modulo some very minor details). Therefore, if \( f(y) \) is just continuous and bounded, the function \( u(t, x) \) is infinitely differentiable for all \( t > 0 \).

We also get directly from (1.179) that

\[ |u(t, x)| \leq \frac{1}{(4\pi t)^{n/2}} \int_{\mathbb{R}^n} |f(y)| \, dy. \tag{1.180} \]

Therefore, \( M_0(t) = \max_{x \in \mathbb{R}^n} |u(t, x)| \) decays as \( t \to +\infty \) as

\[ M_0(t) \leq \frac{1}{(4\pi t)^{n/2}} \int_{\mathbb{R}^n} |f(y)| \, dy. \tag{1.181} \]
It is remarkable that only the initial mass
\[ m_0 = \int_{\mathbb{R}^n} |f(y)| dy \]
enters into the upper bound (1.181), and not the maximum of the initial data. This means that solutions whose initial data consists of a narrow peak drops its maximum very quickly (a narrow peak has a small mass despite having a large maximum).

Let us now differentiate (1.179). We get for, say, the partial derivative with respect to \( x_1 \):
\[ \frac{\partial u(t, x)}{\partial x_1} = \frac{1}{(4\pi t)^{n/2}} \int_{\mathbb{R}^n} \frac{(x_1 - y_1)}{2t} e^{-|x-y|^2/(4t)} f(y) dy. \]

Note that the function \( p(z) = z_1 e^{-z^2/4} \) is globally bounded: it attains its maximum at the point \( \bar{z} = (\sqrt{2}, 0, 0, \ldots, 0) \) where it takes the value \( p(\bar{z}) = \sqrt{2}/e \). It follows that
\[ \left| \frac{\partial u(t, x)}{\partial x_1} \right| \leq \frac{1}{(4\pi t)^{n/2}} \sqrt{t} \int_{\mathbb{R}^n} \left| \frac{(x_1 - y_1)}{\sqrt{t}} e^{-|x-y|^2/(4t)} \right| f(y) dy \]
\[ \leq \frac{1}{(4\pi t)^{n/2}} \sqrt{2e} \int_{\mathbb{R}^n} |f(y)| dy = \frac{C_1}{t^{(n+1)/2}} \int_{\mathbb{R}^n} |f(y)| dy, \]
with \( C_1 = (4\pi)^{n/2} \sqrt{2e} \). Generally, we have
\[ |\nabla u(t, x)| \leq \frac{C_1}{t^{(n+1)/2}} \int_{\mathbb{R}^n} |f(y)| dy. \] (1.182)

We see from (1.182) that \( |\nabla u(t, x)| \) decays even faster as \( t \to +\infty \) than \( |u(t, x)| \), and, moreover, it is controlled by the initial mass of the solution, rather than by its derivatives. This is a very important property of the heat equation: as \( t \to +\infty \) both \( u(t, x) \) and its derivatives decay to zero at a rate controlled by the initial mass. Moreover, the higher the order of the derivative, the faster it decays in time – solution starts to look very smooth.

1.4.7 The self-similar behavior

Let us now analyze what happens to solutions of the heat equation
\[ \frac{\partial u}{\partial t} - \Delta u = 0, \] (1.183)

at large times in more detail. As we expect it to decay at the rate \( t^{-n/2} \) (from (1.181), for instance), we define \( w(t, x) = (1+t)^{n/2} u(t, x) \) (we take \( t+1 \) rather than \( t \) because as \( t \to +\infty \) they do not differ by much, and now \( w(0, x) = u(0, x) \) and we do not have a singularity at \( t = 0 \)). This function satisfies
\[ \frac{\partial w}{\partial t} - \frac{\partial^2 w}{\partial x^2} - \frac{n}{2(t+1)} w = 0, \quad w(0, x) = u(0, x). \] (1.184)
Let us now make a change of variables: look for \( w(t, x) = v(t, x/\sqrt{4(t + 1)}) \), then we calculate, with \( y = x/\sqrt{4(t + 1)} \):

\[
\frac{\partial w}{\partial t} = \frac{\partial v}{\partial t} - \sum_{j=1}^{n} \frac{x_j}{4(t + 1)^{3/2}} \frac{\partial v}{\partial y_j} = \frac{\partial v}{\partial t} - \frac{1}{2(t + 1)} \sum_{j=1}^{n} y_j \frac{\partial v}{\partial y_j} = \frac{\partial v}{\partial t} - \frac{1}{2(t + 1)} y \cdot \nabla_y v,
\]

and

\[
\frac{\partial w}{\partial x_j} = \frac{1}{\sqrt{4(t + 1)}} \frac{\partial v}{\partial y_j}.
\]

Differentiating once again gives

\[
\frac{\partial^2 w}{\partial x_j^2} = \frac{1}{4(t + 1)^2} \frac{\partial^2 v}{\partial y_j^2}.
\]

Inserting this into (1.184) gives

\[
\frac{\partial v}{\partial t} - \frac{1}{2(t + 1)} y \cdot \nabla_y v - \frac{1}{(t + 1)} \Delta_y v - \frac{n}{2(t + 1)} v = 0, \quad v(0, y) = u(0, 2y). \tag{1.185}
\]

The final change of variables is to set \( s = (1/4) \ln(t + 1) \), that is, \( v(t, y) = \tilde{v}((\ln(t + 1))/4, y) \), then

\[
\frac{\partial v}{\partial t} = \frac{1}{4(t + 1)} \frac{\partial \tilde{v}}{\partial s}.
\]

Now, (1.185) becomes

\[
\frac{\partial \tilde{v}}{\partial s} - 2y \cdot \nabla_y \tilde{v} - \Delta_y \tilde{v} - 2n \tilde{v} = 0, \quad \tilde{v}(0, y) = u(0, 2y). \tag{1.186}
\]

Equation (1.186) has a special solution \( \tilde{v}(y) = e^{-y^2} \) that can be checked directly. The main point of our analysis is to show that any solution of (1.186) converges as \( t \to +\infty \) to a multiple of \( \tilde{v}(y) \):

\[
\tilde{v}(s, y) \to m\tilde{v}(y) \text{ as } s \to +\infty. \tag{1.187}
\]

In order to find the constant \( m \) let us integrate (1.186) in \( y \), keeping in mind the following identities: first,

\[
\int_{\mathbb{R}^n} \frac{\partial \tilde{v}(s, y)}{\partial s} dy = \frac{d}{ds} \int_{\mathbb{R}^n} \tilde{v}(s, y) dy.
\]

Next, using Green’s formula we get

\[
\int_{\mathbb{R}^n} y \cdot \nabla \tilde{v} dy = -\int_{\mathbb{R}^n} \tilde{v}(s, y) \text{div}(y) dy = -n \int_{\mathbb{R}^n} \tilde{v}(s, y) dy,
\]

because \( \text{div}y = n \). We also have

\[
\int_{\mathbb{R}^n} \Delta \tilde{v}(s, y) dy = 0.
\]

Putting these expressions together gives

\[
\frac{d}{ds} \int_{\mathbb{R}^n} \tilde{v}(s, y) dy + 2n \int_{\mathbb{R}^n} \tilde{v}(s, y) dy - 2n \int_{\mathbb{R}^n} \tilde{v}(s, y) dy = 0,
\]

48
that is,

\[ \frac{d}{ds} \int_{\mathbb{R}^n} \tilde{v}(s, y) dy = 0. \]

This means that

\[ \int_{\mathbb{R}^n} \tilde{v}(s, y) dy = \int_{\mathbb{R}^n} u(0, 2y) dy \]

is a \textit{conserved quantity}. The only possible value for constant \( m \) is then determined from the condition

\[ m \int_{\mathbb{R}^n} \tilde{v}(y) dy = \int_{\mathbb{R}^n} u(0, 2y) dy, \quad (1.188) \]

whence

\[ m = \frac{1}{\pi^{n/2}} \int_{\mathbb{R}^n} u(0, 2y) dy = \frac{1}{(4\pi)^{n/2}} \int_{\mathbb{R}^n} u(0, y) dy. \quad (1.189) \]

How do we actually establish convergence (1.187) with the constant \( m \) determined by (1.188)? The ratio \( z(s, y) = \tilde{v}(s, y)/\tilde{v}(y) \) satisfies

\[ \frac{\partial z}{\partial s} + 2y \cdot \nabla z - \Delta z = 0, \quad z(0, y) = u(0, y) e^{y^2}. \quad (1.190) \]

Let us assume for the sake of simplicity of notation that the dimension \( n = 1 \) and set

\[ \phi(s, y) = \frac{\partial z(s, y)}{\partial y}. \]

The function \( \phi \) satisfies

\[ \frac{\partial \phi}{\partial s} + 2y \frac{\partial \phi}{\partial y} + 2\phi - \Delta \phi = 0, \quad \phi(0, y) = \frac{d}{dy} \left[ u(0, y) e^{y^2} \right]. \quad (1.191) \]

Let us multiply this equation by \( \phi \) and integrate in \( y \): first we have

\[ \int_{\mathbb{R}} \phi(s, y) \frac{\partial \phi(s, y)}{\partial s} dy = \frac{1}{2} \frac{d}{ds} \int_{\mathbb{R}} \phi^2(s, y) dy. \]

Next, we get, using integration by parts:

\[ 2 \int_{\mathbb{R}} y \phi(s, y) \frac{\partial \phi(s, y)}{\partial y} dy = - \int_{\mathbb{R}} \phi^2(s, y) dy, \]

and, finally:

\[ \int_{\mathbb{R}} \phi(s, y) \frac{\partial^2 \phi(s, y)}{\partial y^2} dy = - \int_{\mathbb{R}} \left| \frac{\partial \phi(s, y)}{\partial y} \right|^2 dy. \]

Together, these identities and (1.191) imply that

\[ \frac{1}{2} \frac{d}{ds} \int_{\mathbb{R}} \phi^2(s, y) dy + \int_{\mathbb{R}} \phi^2(s, y) dy + \int_{\mathbb{R}} \left| \frac{\partial \phi(s, y)}{\partial y} \right|^2 dy = 0. \]
As a consequence, we have
\[
\int_{\mathbb{R}} \phi^2(s, y) dy \leq \left( \int_{\mathbb{R}} \phi^2(0, y) dy \right) e^{-2s} \to 0 \text{ as } s \to +\infty.
\] (1.192)

Recalling that \( \phi \) is actually the derivative of \( z(s, y) \) with respect to \( y \), we conclude that \( z(s, y) \) tends to a constant as \( s \to +\infty \).

Let us now re-interpret the convergence result (1.187) in terms of the original variables \( t \) and \( x \), and the function \( u(t, x) \) that actually solves the heat equation (1.183). Recall that
\[
u(t, x) = \frac{1}{(1 + t)^{n/2}} w(t, x) = \frac{1}{(1 + t)^{n/2}} v \left( t, \frac{x}{\sqrt{4(t + 1)}} \right) = \frac{1}{(1 + t)^{n/2}} v \left( \frac{\ln(t + 1)}{4}, \frac{x}{\sqrt{4(t + 1)}} \right).
\]
Now, (1.187) tells us that
\[
u(t, x) \approx \frac{m}{(1 + t)^{n/2}} e^{-x^2/(4(t+1))} = \frac{1}{(4\pi(1 + t)^{n/2})} e^{-x^2/(4(t+1))} \int_{\mathbb{R}^n} u(0, y) dy.
\] (1.193)

Therefore, after long times solution starts to have the profile of a Gaussian weighted appropriately to have the correct mass. But the only information that remains from the initial condition is its total mass – the rest is completely lost. All solutions "look the same" – this is self-similarity.

2 The Fourier transform

2.1 The Fourier transform on the circle

2.1.1 Pointwise convergence on the circle

Given a periodic function \( f \) defined on the interval \([0, 1]\), we define the Fourier coefficients, for integers \( k \in \mathbb{Z} \) as
\[
\hat{f}(k) = \int_{0}^{1} f(x) e^{-2\pi i k x} dx.
\]
Trivially, we have \(|\hat{f}(k)| \leq \|f\|_{L^1} \) for all \( k \in \mathbb{Z} \), where
\[
\|f\|_{L^1} = \int_{0}^{1} |f(x)| dx.
\]
The Riemann-Lebesgue lemma shows that a continuous signal can not have too much high-frequency content and \( \hat{f}(k) \) have to decay for large \( k \) (this is actually true for a much more general class of functions). We will denote by \( C_{per}[0, 1] \) the class of functions that are continuous on \([0, 1]\) and \( f(0) = f(1) \).

**Lemma 2.1** (The Riemann-Lebesgue lemma) If \( f \in C_{per}[0, 1] \) then \( \hat{f}(k) \to 0 \) as \( k \to +\infty \).
**Proof.** Note that
\[
\hat{f}(k) = \int_0^1 f(x)e^{-2\pi ikx}dx = -\int_0^1 f(x)e^{-2\pi i(k+1/2)x}dx = -\int_0^1 f(x - \frac{1}{2k})e^{-2\pi ikx}dx,
\]
and thus
\[
\hat{f}(k) = \frac{1}{2} \int_0^1 \left[f(x) - f(x - \frac{1}{2k})\right] e^{-2\pi ikx}dx.
\]
As a consequence, since \(f(x)\) is continuous on \([0, 1]\), we have
\[
|\hat{f}(k)| \leq \frac{1}{2} \int_0^1 \left|f(x) - f(x - \frac{1}{2k})\right| dx,
\]
hence \(\hat{f}(k) \to 0\) as \(k \to +\infty\).

A simple implication of the Riemann-Lebesgue lemma is that
\[
\int_0^1 f(x) \sin(mx)dx \to 0
\]
as \(m \to \infty\) for any \(f \in C_{per}([0,1])\).

In order to investigate convergence of the Fourier series
\[
\sum_{k=-\infty}^{\infty} \hat{f}(k)e^{2\pi ikx}
\]
let us introduce the partial sums
\[
S_N f(x) = \sum_{k=-N}^{N} \hat{f}(k)e^{2\pi ikx}.
\]

A convenient way to represent \(S_N f\) is by writing it as a convolution:
\[
S_N f(x) = \int_0^1 f(t) \sum_{k=-N}^{N} e^{2\pi ik(x-t)} dt = \int_0^1 f(x-t)D_N(t) dt.
\]

Here the Dini kernel is
\[
D_N(t) = \sum_{k=-N}^{N} e^{2\pi ikt} = e^{-2\pi iNt}(1 + e^{2\pi it} + e^{4\pi it} + \ldots + e^{4\pi iNt}) = e^{-2\pi iNt} \frac{e^{2\pi i(2N+1)t} - 1}{e^{2\pi it} - 1}
\]
\[
= \frac{e^{2\pi i(N+1/2)t} - e^{-2\pi i(N+1/2)t}}{e^{\pi it} - e^{-\pi it}} = \frac{\sin((2N+1)\pi t)}{\sin(\pi t)}.
\]
The definition of the Dini kernel as a sum of exponentials implies immediately that
\[
\int_0^1 D_N(t) dt = 1 \tag{2.1}
\]
for all \( N \), while the expression in terms of sines shows that
\[
|D_N(t)| \leq \frac{1}{\sin(\pi \delta)}, \quad \delta \leq |t| \leq 1/2.
\]
The "problem" with the Dini kernel is that its \( L^1 \)-norm is not uniformly bounded in \( N \). Indeed, consider
\[
L_N = \int_{-1/2}^{1/2} |D_N(t)| dt. \tag{2.2}
\]
Let us show that
\[
\lim_{N \to +\infty} L_N = +\infty. \tag{2.3}
\]
We compute:
\[
L_N = 2 \int_0^{1/2} \frac{|\sin((2N + 1)\pi t)|}{|\sin \pi t|} dt \geq 2 \int_0^{1/2} \frac{|\sin((2N + 1)\pi t)|}{|\pi t|} dt
- 2 \int_0^{1/2} |\sin((2N + 1)\pi t)| \left| \frac{1}{\sin \pi t} - \frac{1}{\pi t} \right| dt = 2 \int_0^{N+1/2} \frac{|\sin(\pi t)|}{\pi t} dt + O(1)
\geq \frac{2}{\pi} \sum_{k=0}^{N-1} \int_0^1 \left| \frac{\sin \pi t}{t + k} \right| dt + O(1) \geq C \log N + O(1),
\]
which implies (2.3). This means that (2.1) holds because of cancellation of many oscillatory terms. These oscillations may cause difficulties in the convergence of the Fourier series.

Convergence of the Fourier series for regular functions

Nevertheless, for "reasonably regular" functions the Fourier series converges and Dini's criterion for the convergence of the Fourier series is as follows.

**Theorem 2.2** (Dini's criterion) Let \( f \in C[0, 1] \) satisfy the following condition at the point \( x \): there exists \( \delta > 0 \) so that
\[
\int_{|t|<\delta} \left| \frac{f(x+t) - f(x)}{t} \right| dt < +\infty, \tag{2.4}
\]
then \( \lim_{N \to +\infty} S_N f(x) = f(x) \). In particular, this is true if \( f \) is differentiable at the point \( x \).

Another criterion for the convergence of the Fourier series was given by Jordan:

**Theorem 2.3** (Jordan's criterion) If \( f \) is continuous and monotonic on some interval \((x - \delta, x + \delta)\) around the point \( x \), except possibly at the point \( x \) itself, then
\[
\lim_{N \to +\infty} S_N f(x) = \frac{1}{2} [f(x^+) + f(x^-)]. \tag{2.5}
\]
The du Bois-Raymond example

In 1873, surprisingly, du Bois-Raymond proved that the Fourier series of a continuous function may diverge at a point.

**Theorem 2.4** There exists a continuous function \( f \) so that its Fourier series diverges at \( x = 0 \).

Kolmogorov showed in 1926 that an \( L^1 \)-function may have a Fourier series that diverges at every point. Then Carelson in 1965 proved that the Fourier series of an \( L^2 \)-function converges almost everywhere and then Hunt improved this result to an arbitrary \( L^p \) for \( p > 1 \).

### 2.2 Approximation by trigonometric polynomials

**The Cesaro sums**

In order to 'improve' the convergence of the Fourier series consider the corresponding Cesaro sums

\[
\sigma_N f(x) = \frac{1}{N+1} \sum_{k=0}^{N} S_k f(x) = \int_{0}^{1} f(t) F_N(x-t) dt,
\]

where \( F_N \) is the Fejér kernel

\[
F_N(t) = \frac{1}{N+1} \sum_{k=0}^{N} D_k(t) = \frac{1}{(N+1) \sin^2(\pi t)} \sum_{k=0}^{N} \sin(\pi(2k+1)t) \sin(\pi t)
\]

\[
= \frac{1}{2(N+1) \sin^2(\pi t)} \sum_{k=0}^{N} [\cos(2\pi kt) - \cos(2\pi (k+1)t)]
\]

\[
= \frac{1}{2(N+1) \sin^2(\pi t)} [1 - \cos(2\pi (N+1)t)] = \frac{1}{N+1} \frac{\sin^2(\pi (N+1)t)}{\sin^2(\pi t)}.
\]

The definition and explicit form of \( F_N \) show that, unlike the Dini kernel, \( F_N \) is non-negative and has \( L^1 \)-norm

\[
\int_{0}^{1} |F_N(t)| dt = 1. \quad (2.6)
\]

Moreover, its mass outside of any finite interval around zero vanishes as \( N \to +\infty \):

\[
\lim_{N \to \infty} \int_{\delta < |t| < 1/2} F_N(t) dt = 0 \quad \text{for any } \delta > 0. \quad (2.7)
\]

This improvement is reflected in the following approximation theorem.

**Theorem 2.5** Let \( f \in L^2[0,1] \), \( 1 \leq p < \infty \), then

\[
\lim_{N \to \infty} \|\sigma_N f - f\|_{L^2} = 0. \quad (2.8)
\]

Moreover, if \( f \in C_{\text{per}}[0,1] \), then

\[
\lim_{N \to \infty} \|\sigma_N f - f\|_{C_{\text{per}}[0,1]} = 0. \quad (2.9)
\]
Corollary 2.6  The Parceval identity holds for any $f \in L^2[0, 1]$:

$$\sum_{k \in \mathbb{Z}} |\hat{f}(k)|^2 = \int_0^1 |f(x)|^2 dx. \tag{2.10}$$

2.3  An application to the heat equation: the periodic case

Let us consider the heat equation on the real line

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}, \tag{2.11}$$

with the initial data $u(0, x) = f(x)$ that is 1-periodic: $f(x + 1) = f(x)$. We claim that the solution $u(t, x)$ is 1-periodic in $x$ for any fixed $t \geq 0$, that is, $u(t, x + 1) = u(t, x)$ for all $t \geq 0$ and $x \in \mathbb{R}$. Indeed, the function $v(t, x) = u(t, x + 1) - u(t, x)$ satisfies

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

and $v(0, x) = f(x + 1) - f(x) = 0$, by the assumption that $f(x)$ is 1-periodic. It follows that $v(t, x) = 0$ for all $t > 0$, $x \in \mathbb{R}$, which means that $u(t, x + 1) = u(t, x)$. Consider the Fourier transform of both sides of (2.11):

$$\int_0^1 \frac{\partial u(t, x)}{\partial t} e^{-2\pi i nx} dx = \frac{d}{dt} \int_0^1 u(t, x) e^{-2\pi i nx} dx = \frac{du_n}{dt},$$

where

$$u_n = \int_0^1 u(t, x) e^{-2\pi i nx} dx$$

are the Fourier coefficients of the function $u$. On the other hand, since $u(t, x)$ is 1-periodic, we have

$$\int_0^1 \frac{\partial^2 u(t, x)}{\partial x^2} e^{-2\pi i nx} dx = 2\pi i n \int_0^1 \frac{\partial u(t, x)}{\partial x} e^{-2\pi i nx} dx = (2\pi i n)^2 u_n.$$

The boundary terms at $x = 0$ and $x = 1$ above disappear because $u(t, x)$ is periodic. Therefore, we get an ODE for $u_n$:

$$\frac{du_n}{dt} = -4\pi^2 n^2 u_n, \quad u_n(0) = f_n. \tag{2.12}$$

Here

$$f_n = \int_0^1 f(x) e^{-2\pi i nx} dx,$$

are the Fourier coefficients of the initial data $f(x)$. This gives

$$u_n(t) = e^{-4\pi^2 n^2 t} f_n,$$

and an explicit formula for the solution of the heat equation

$$u(t, x) = \sum_{n \in \mathbb{Z}} e^{-4\pi^2 n^2 t} f_n. \tag{2.13}$$
with \( f_n \) given by (2.12). Expression (2.13) conveys very effectively the regularizing properties of the heat equation: the high frequencies are attenuated very quickly due to the exponentially decaying factor. This is another way to express the fact that the heat equation kills oscillations. Another simple consequence of (2.13) is that

\[
u(t, x) \to f_0 = \int_0^1 f(y)dy, \text{ as } t \to +\infty,
\]

uniformly in \( x \) – that is, in the long time limit solution of the heat equation with periodic initial data converges to a constant equal to its spatial average.

### 2.4 The Fourier transform in \( \mathbb{R}^n \)

Given an \( L^1(\mathbb{R}^n) \)-function \( f \) its Fourier transform is

\[
\hat{f}(\xi) = \int f(x) e^{-2\pi i x \cdot \xi} dx.
\]

Then, obviously:

\[
|\hat{f}(\xi)| \leq \int_{\mathbb{R}^n} |f(x)| dx.
\]

Moreover, the function \( \hat{f}(\xi) \) is continuous, and the Riemann-Lebesgue lemma is easily generalized to the Fourier transform on \( \mathbb{R}^n \), and

\[
\lim_{\xi \to \infty} \hat{f}(\xi) = 0.
\]

For a smooth compactly supported function \( f \in C_0^\infty(\mathbb{R}^n) \) we have the following remarkable algebraic relations between taking derivatives and multiplying by polynomials:

\[
\frac{\partial f}{\partial x_i}(\xi) = 2\pi i \xi \hat{f}(\xi), \tag{2.14}
\]

and

\[
(-2\pi i) (x_j f)(\xi) = \frac{\partial \hat{f}}{\partial \xi_j}(\xi). \tag{2.15}
\]

This motivates the following definition.

**Definition 2.7** The Schwartz class \( S(\mathbb{R}^n) \) consists of functions \( f \) such that for any pair of multi-indices \( \alpha \) and \( \beta \)

\[
p_{\alpha\beta}(f) := \sup_x |x^\alpha D^\beta f(x)| < +\infty.
\]

As \( C_0^\infty(\mathbb{R}^n) \) lies inside the Schwartz class, the Schwartz functions are dense in \( L^1(\mathbb{R}^n) \).

The main reason to introduce the Schwartz class is the following theorem.
Theorem 2.8 (i) The Fourier transform of a Schwartz class function $f(x)$ is a Schwartz class function $\hat{f}(\xi)$.

(ii) For any $f, g \in \mathcal{S}(\mathbb{R}^n)$ we have

$$\int_{\mathbb{R}^n} f(x) \hat{g}(x) dx = \int_{\mathbb{R}^n} \hat{f}(x) g(x) dx. \quad (2.16)$$

(iii) The following inversion formula holds:

$$f(x) = \int \hat{f}(\xi) e^{2\pi i x \cdot \xi} d\xi \quad (2.17)$$

for all $f \in \mathcal{S}(\mathbb{R}^n)$.

Proof. We begin with a lemma that is one of the cornerstones of probability theory.

Lemma 2.9 Let $f(x) = e^{-\pi |x|^2}$, then $\hat{f}(x) = f(x)$.

Proof. First, as

$$f(x) = e^{-\pi |x_1|^2} e^{-\pi |x_2|^2} \ldots e^{-\pi |x_n|^2},$$

so that both $f$ and $\hat{f}$ factor into a product of functions of one variable, it suffices to consider the case $n = 1$. The proof is a glimpse of how useful the Fourier transform is for differential equations and vice versa: the function $f(x)$ satisfies an ordinary differential equation

$$u' + 2xu = 0, \quad (2.18)$$

with the boundary condition $u(0) = 1$. However, relations (2.14) and (2.15) together with (2.18) imply that $\hat{f}$ satisfies the same differential equation (2.18), with the same boundary condition $\hat{f}(0) = 0$. It follows that $f(x) = \hat{f}(x)$ for all $x \in \mathbb{R}$. $\Box$

We continue with the proof of Theorem 2.8. Relations (2.14) and (2.15) imply that the Fourier transform of a Schwartz class function is of the Schwartz class.

The Parceval identity can be verified directly as follows:

$$\int_{\mathbb{R}^n} f(x) \hat{g}(x) dx = \int_{\mathbb{R}^n} f(x) g(\xi) e^{-2\pi i \xi \cdot x} dx d\xi = \int_{\mathbb{R}^n} \hat{f}(\xi) g(\xi) d\xi.$$

Finally, we prove the inversion formula using a rescaling argument. Let $f, g \in \mathcal{S}(\mathbb{R}^n)$ then for any $\lambda > 0$ we have

$$\int_{\mathbb{R}^n} f(x) \hat{g}(\lambda x) dx = \int_{\mathbb{R}^{2n}} f(x) g(\xi) e^{-2\pi i \lambda \xi \cdot x} dx d\xi = \frac{1}{\lambda^n} \int_{\mathbb{R}^n} \hat{f}(\lambda \xi) g\left(\frac{\xi}{\lambda}\right) d\xi.$$

Multiplying by $\lambda^n$ and changing variables on the left side we obtain

$$\int_{\mathbb{R}^n} f\left(\frac{x}{\lambda}\right) \hat{g}(x) dx = \int_{\mathbb{R}^n} \hat{f}(\xi) g\left(\frac{\xi}{\lambda}\right) d\xi.$$

Letting now $\lambda \to \infty$ gives

$$f(0) \int_{\mathbb{R}^n} \hat{g}(x) dx = g(0) \int_{\mathbb{R}^n} \hat{f}(\xi) d\xi, \quad (2.19)$$
for all functions \( f \) and \( g \) in \( \mathcal{S}(\mathbb{R}^n) \). Taking \( g(x) = e^{-\pi |x|^2} \) in (2.19) and using Lemma 2.9 leads to

\[
 f(0) = \int_{\mathbb{R}^n} \hat{f}(\xi) d\xi. \tag{2.20}
\]

The inversion formula (2.17) now follows if we apply (2.20) to a shifted function \( f_y(x) = f(x + y) \), because

\[
 \hat{f}_y(\xi) = \int_{\mathbb{R}^n} f(x + y) e^{-2\pi i \xi \cdot x} dx = e^{2\pi i \xi \cdot y} \hat{f}(\xi),
\]

so that

\[
 f(y) = f_y(0) = \int_{\mathbb{R}^n} \hat{f}_y(\xi) d\xi = \int_{\mathbb{R}^n} e^{2\pi i \xi \cdot y} \hat{f}(\xi) d\xi,
\]

which is (2.17). □

The Schwartz distributions

**Definition 2.10** The space \( \mathcal{S}'(\mathbb{R}^n) \) of Schwartz distributions is the space of linear functionals \( T \) on \( \mathcal{S}(\mathbb{R}^n) \) such that \( T(\phi_k) \to 0 \) for all sequences \( \phi_k \to 0 \) in \( \mathcal{S}(\mathbb{R}^n) \).

Theorem 2.8 allows us to extend the Fourier transform to distributions in \( \mathcal{S}'(\mathbb{R}^n) \) by setting

\[
 \hat{T}(f) = T(\hat{f}) \quad \text{for} \quad T \in \mathcal{S}'(\mathbb{R}^n) \quad \text{and} \quad f \in \mathcal{S}(\mathbb{R}^n).
\]

The fact that \( \hat{T}(f_k) \to 0 \) for all sequences \( f_k \to 0 \) in \( \mathcal{S}(\mathbb{R}^n) \) follows from the continuity of the Fourier transform as a map \( \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n) \), hence \( \hat{T} \) is a Schwartz distribution for all \( T \in \mathcal{S}'(\mathbb{R}^n) \). For example, if \( \delta_0 \) is the Schwartz distribution such that \( \delta_0(f) = f(0) \), \( f \in \mathcal{S}(\mathbb{R}^n) \), then

\[
 \hat{\delta}_0(f) = \hat{f}(0) = \int_{\mathbb{R}^n} f(x) dx,
\]

so that \( \hat{\delta}(\xi) \equiv 1 \) for all \( \xi \in \mathbb{R}^n \).

Similarly, since differentiation is a continuous map \( \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n) \), we may define the distributional derivative as

\[
 \frac{\partial T}{\partial x_j}(f) = -T \left( \frac{\partial f}{\partial x_j} \right),
\]

for all \( T \in \mathcal{S}'(\mathbb{R}^n) \) and \( f \in \mathcal{S}(\mathbb{R}^n) \) – the minus sign here comes from the integration by parts formula, for if \( T \) happens to have the form

\[
 T_g(f) = \int_{\mathbb{R}^n} f(x) g(x) dx,
\]

with a given \( g \in \mathcal{S}(\mathbb{R}^n) \), then

\[
 T \left( \frac{\partial f}{\partial x_j} \right) = \int_{\mathbb{R}^n} \frac{\partial f}{\partial x_j}(x) g(x) dx = -\int_{\mathbb{R}^n} f(x) \frac{\partial g}{\partial x_j}(x) dx.
\]

For instance, in one dimension \( \delta_0(x) = 1/2(\text{sgn}(x))' \) in the distributional sense because for any function \( f \in \mathcal{S}(\mathbb{R}) \) we have

\[
 \langle (\text{sgn})', f \rangle = -\langle \text{sgn}, f' \rangle = -\int_{-\infty}^{\infty} \text{sgn}(x) f'(x) dx = \int_{-\infty}^{0} f'(x) dx - \int_{0}^{\infty} f'(x) dx = 2f(0) = 2\langle \delta_0, f \rangle.
\]
2.5 An application to the heat equation in the whole space

Consider the Cauchy problem for the heat equation

\[
\frac{\partial u}{\partial t} - \Delta u = 0, \quad t > 0, \quad x \in \mathbb{R}^n, \quad u(0, x) = f(x), \quad x \in \mathbb{R}^n. \tag{2.21}
\]

Taking the Fourier transform of the heat equation in the whole space gives, very similarly to the periodic case considered in Section 2.3:

\[
\frac{\partial \hat{u}(t, \xi)}{\partial t} + 4\pi^2 |\xi|^2 \hat{u}(t, \xi) = 0, \tag{2.22}
\]

with the initial data

\[
\hat{u}(0, \xi) = \hat{f}(\xi),
\]

where

\[
\hat{u}(t, \xi) = \int_{\mathbb{R}^n} u(t, x)e^{-2\pi i \xi \cdot x} dx, \quad \hat{f}(\xi) = \int_{\mathbb{R}^n} f(x)e^{-2\pi i \xi \cdot x} dx.
\]

The ODE (2.22) can be solved explicitly:

\[
\hat{u}(t, \xi) = \hat{f}(\xi)e^{-4\pi^2 |\xi|^2 t}. \tag{2.23}
\]

As in the periodic case, we see that high frequency components of the solutions of the heat equation decay rapidly in time, leading to the regularizing effects we have discussed previously.

The inverse Fourier transform then gives

\[
u(t, x) = \int_{\mathbb{R}^n} \hat{u}(t, \xi)e^{2\pi i \xi \cdot x} d\xi = \int_{\mathbb{R}^n} \hat{f}(\xi)e^{-4\pi^2 |\xi|^2 t} e^{2\pi i \xi \cdot x} d\xi. \tag{2.24}
\]

Let us make a couple of “regularizing observations” from (2.24). For example, we have

\[
|u(t, x)| \leq \int_{\mathbb{R}^n} |\hat{f}(\xi)|e^{-4\pi^2 |\xi|^2 t} d\xi \leq \int_{\mathbb{R}^n} e^{-4\pi^2 |\xi|^2 t} d\xi \left[ \max_{\xi \in \mathbb{R}^n} |\hat{f}(\xi)| \right].
\]

Recall that

\[
|\hat{f}(\xi)| \leq \int_{\mathbb{R}^n} |f(x)| dx,
\]

for all $\xi \in \mathbb{R}^n$. It follows that

\[
|u(t, x)| \leq \int_{\mathbb{R}^n} |f(x)| dx \int_{\mathbb{R}^n} e^{-4\pi^2 |\xi|^2 t} d\xi = \frac{1}{(4\pi t)^{n/2}} \int_{\mathbb{R}^n} |f(x)| dx,
\]

an estimate we have seen before but now obtained without the use of the Green’s function.

We may also recover self-similarity of the solution of the heat equation in the long time limit: consider $t \gg 1$ and let us look at $x \sim \sqrt{t}$; mathematically, this means that we fix $y \in \mathbb{R}$ and consider $u(t, y\sqrt{t})$ with the idea of then letting $t \to +\infty$ with $y$ fixed. We have

\[
u(t, y\sqrt{t}) = \int_{\mathbb{R}^n} \hat{f}(\xi)e^{-4\pi^2 |\xi|^2 t} e^{2\pi i \xi \cdot y} d\xi.
\]
Let us make a change of variable $k = 2\sqrt{\pi t}\xi$:

$$u(t, y\sqrt{t}) = \frac{1}{(2\sqrt{\pi t})^n} \int_{\mathbb{R}^n} \hat{f} \left( \frac{k}{2\sqrt{\pi t}} \right) e^{-\pi|k|^2} e^{i\sqrt{\pi}ky} dk \approx \frac{\hat{f}(0)}{(2\sqrt{\pi t})^n} \int_{\mathbb{R}^n} e^{-\pi|k|^2} e^{2\pi i k \cdot (y/2\sqrt{\pi})} dk.$$  

Lemma 2.9 implies that

$$u(t, y\sqrt{t}) \approx \frac{\hat{f}(0)}{4\pi t} e^{-\pi|y|^2/(4\pi)} = e^{-|y|^2/(4\pi)} \int_{\mathbb{R}^n} f(z) dz,$$

which is the self-similarity formula (1.193) we have obtained before by a rather more complicated approach.

Finally, in order to obtain the formula for the solution of the heat equation in terms of the heat kernel we start with (2.24) and write

$$u(t, x) = \int_{\mathbb{R}^n} \hat{f}(\xi) e^{-4\pi^2|\xi|^2 t} e^{-2\pi^2|\xi|^2 t} d\xi = \int_{\mathbb{R}^n} f(y) e^{-2\pi i\xi \cdot y} e^{-4\pi^2|\xi|^2 t} e^{2\pi i\xi \cdot x} d\xi dy,$$

and use Lemma 2.9 in the last step below:

$$u(t, x) = \int_{\mathbb{R}^n} f(y) \left( \int_{\mathbb{R}^n} e^{-2\pi i\xi \cdot (y-x)} e^{-4\pi^2|\xi|^2 t} d\xi \right) dy$$

$$= \frac{1}{(4\pi t)^{n/2}} \int_{\mathbb{R}^n} f(y) \left( \int_{\mathbb{R}^n} e^{-2\pi i\xi \cdot (y-x)/\sqrt{4\pi t}} e^{-\pi|\xi|^2} d\xi \right) dy$$

$$= \frac{1}{(4\pi t)^{n/2}} \int_{\mathbb{R}^n} f(y) e^{-(x-y)^2/(4t)} dy,$$

which is, of course, the expression for $u(t, x)$ in terms of the heat kernel.

### 3 The wave equation

#### 3.1 The wave equation in one dimension

The wave equation describes propagation of a disturbance $\phi(t, x)$ that moves with a local speed $c(x)$. It has the form

$$\frac{1}{c^2(x)} \frac{\partial^2 \phi}{\partial t^2} - \Delta \phi = 0. \tag{3.1}$$

Here

$$\Delta = \sum_{j=1}^{n} \frac{\partial^2}{\partial x_j^2}$$

is the standard Laplacian in $n$ dimensions, $t$ is the time variable, and $x$ is the spatial coordinate. Unless specified otherwise we will always assume $t \geq 0$ and $x \in \mathbb{R}^n$. When $x \in \mathbb{R}$ and the sound speed $c$ is constant then a general solution of (3.1) is given by the famous d’Alembert formula

$$\phi(t, x) = f(x - ct) + g(x + ct), \tag{3.2}$$

with arbitrary functions $f$ and $g$. It is straightforward to verify that any function of the form (3.2) solves the one-dimensional wave equation

$$\frac{1}{c^2} \frac{\partial^2 \phi}{\partial t^2} - \frac{\partial^2 \phi}{\partial x^2} = 0. \tag{3.3}$$
The functions \( f \) and \( g \) represent the right and left going waves, respectively.

Let us now consider the Cauchy problem

\[
\frac{1}{c^2} \frac{\partial^2 \phi}{\partial t^2} - \frac{\partial^2 \phi}{\partial x^2} = 0, \quad (3.4)
\]

\[
\phi(0, x) = p(x), \quad u_t(0, x) = q(x).
\]

Here we use the notation \( u_t = \partial u / \partial t \). Note that, since the wave equation is of the second order in time, we need to prescribe both the initial value of \( \phi \) and the value of its derivative at the time \( t = 0 \). We would like to express the solution of (3.4) in the form (3.3), that is, as a sum of left and right going waves. for this decomposition to hold we need

\[
p(x) = f(x) + g(x), \quad q(x) = -cf'(x) + cg'(x).
\]

it follows that

\[
g'(x) = \frac{1}{2c} (cp'(x) + q(x)), \quad f'(x) = \frac{1}{2c} (cp'(x) - q(x)),
\]

and thus we may take

\[
g(x) = \frac{1}{2} p(x) + \frac{1}{2c} \int_0^x q(y) dy, \quad f(x) = \frac{1}{2} p(x) - \frac{1}{2c} \int_0^x q(y) dy. \quad (3.5)
\]

Therefore, solution of the Cauchy problem (3.4) is given by

\[
u(t, x) = \frac{1}{2} p(x - ct) - \frac{1}{2c} \int_0^{x-ct} q(y) dy + \frac{1}{2} p(x + ct) + \frac{1}{2c} \int_0^{x+ct} q(y) dy, \quad (3.6)\]

or, equivalently,

\[
u(t, x) = \frac{1}{2} p(x - ct) + \frac{1}{2} p(x + ct) + \frac{1}{2c} \int_{x-ct}^{x+ct} q(y) dy. \quad (3.7)
\]

Expression (3.7) is known as d’Alembert’s formula. One corollary is that, unlike for the heat equation, solution at times \( t > 0 \) is no more regular than at \( t = 0 \): if, say, \( p(x) \) has only five derivatives, the function \( u(t, x) \) will not be any more regular than that. If \( p(x) \) is oscillatory, then \( u(t, x) \) will also be oscillatory and so on.

The idea that solution may be decomposed into such primitive building blocks goes much further than this trivial example. Another useful observation that is obvious in the one-dimensional case is that solutions propagate with a finite speed: if \( \phi(0, x) = 0 \) in an interval \((a - r, a + r)\) of length \(2r\) centered at \( x = a \) then \( \phi(t, a) = 0 \) for all \( t \leq T_0 = r/c \).

### 3.2 D’Alembert’s formula by the Fourier transform

Let us now use the Fourier transform to solve the Cauchy problem for the wave equation. Consider the Cauchy problem

\[
\frac{1}{c^2} \frac{\partial^2 \phi}{\partial t^2} - \Delta \phi = 0, \quad t \geq 0, \quad x \in \mathbb{R}, \quad (3.8)
\]

\[
\phi(0, x) = p(x), \quad \phi_t(0, x) = q(x).
\]
Let
\[ \hat{\phi}(t, \xi) = \int_{\mathbb{R}} \phi(t, x) e^{-2\pi i \xi x} dx \]
be the Fourier transform of \( \phi(t, x) \). It satisfies the ODE
\[ \frac{\partial^2 \hat{\phi}}{\partial t^2} + 4\pi^2 c^2 |\xi|^2 \hat{\phi} = 0, \tag{3.9} \]
with the initial data
\[ \hat{\phi}(0, \xi) = \hat{p}(\xi), \quad \hat{\phi}_t(0, \xi) = \hat{q}(\xi). \tag{3.10} \]
It follows from (3.9) that
\[ \hat{\phi}(t, \xi) = A(\xi) e^{-2\pi i \xi c t} + B(\xi) e^{2\pi i \xi c t}. \]
The coefficients \( A(\xi) \) and \( B(\xi) \) are determined by the initial condition (3.10):
\[ A(\xi) + B(\xi) = \hat{p}(\xi), \quad -2\pi ic\xi A(\xi) + 2\pi ic\xi B(\xi) = \hat{q}(\xi), \]
hence
\[ A(\xi) = \frac{1}{2} \hat{p}(\xi) - \frac{\hat{q}(\xi)}{4\pi ic\xi}, \quad B(\xi) = \frac{1}{2} \hat{p}(\xi) + \frac{\hat{q}(\xi)}{4\pi ic\xi}, \]
leading to
\[ \hat{\phi}(t, \xi) = \left( \frac{1}{2} \hat{p}(\xi) - \frac{\hat{q}(\xi)}{4\pi ic\xi} \right) e^{-2\pi i \xi c t} + \left( \frac{1}{2} \hat{p}(\xi) + \frac{\hat{q}(\xi)}{4\pi ic\xi} \right) e^{2\pi i \xi c t}. \tag{3.11} \]
Taking the inverse Fourier transform gives
\[ \phi(t, x) = \int_{\mathbb{R}} \left( \frac{1}{2} \hat{p}(\xi) - \frac{\hat{q}(\xi)}{4\pi ic\xi} \right) e^{-2\pi i \xi c t + 2\pi i \xi x} d\xi + \int_{\mathbb{R}} \left( \frac{1}{2} \hat{p}(\xi) + \frac{\hat{q}(\xi)}{4\pi ic\xi} \right) e^{2\pi i \xi c t + 2\pi i \xi x} d\xi \]
\[ = \frac{1}{2} (p(x - ct) + p(x + ct)) + \frac{1}{2c} \int_{x-ct}^{x+ct} q(y) dy, \tag{3.12} \]
which is nothing but d’Alembert’s formula (3.7).

### 3.3 Conservation of energy

The energy of solutions of (3.1) is defined as
\[ \mathcal{E}(t) = \frac{1}{2} \int_{\mathbb{R}^n} \left[ \frac{1}{c^2(x)} |\phi_t(t, x)|^2 + |\nabla \phi(t, x)|^2 \right] dx. \tag{3.13} \]
It is straightforward to verify that energy of (at least sufficiently smooth) solutions of (3.1) is preserved: \( \mathcal{E}(t) = \mathcal{E}(0) \). Indeed, we have
\[ \frac{d\mathcal{E}(t)}{dt} = \int_{\mathbb{R}^n} \left[ \frac{1}{c^2(x)} \phi_t(t, x) \phi_{tt}(t, x) + \nabla \phi(t, x) \cdot \nabla \phi_t(t, x) \right] dx. \]
Using the wave equation (3.13) we can re-write this as
\[
\frac{d\mathcal{E}(t)}{dt} = \int_{\mathbb{R}^n} \left[ \phi_t(t,x) \Delta \phi(t,x) + \nabla \phi(t,x) \cdot \nabla \phi_t(t,x) \right] dx.
\]
Integrating by parts in the first term (the boundary terms at infinity vanish) gives
\[
\frac{d\mathcal{E}(t)}{dt} = \int_{\mathbb{R}^n} \left[ -\nabla \phi_t(t,x) \cdot \nabla \phi(t,x) + \nabla \phi(t,x) \cdot \nabla \phi_t(t,x) \right] dx = 0,
\]
hence \(\mathcal{E}(t) \equiv \mathcal{E}(0)\), as claimed.

**Uniqueness for the Cauchy problem.** An immediate corollary of energy conservation is uniqueness of smooth solutions to the Cauchy problem for wave equation. More precisely, consider the Cauchy problem (3.4) in an arbitrary dimension \((x \in \mathbb{R}^n)\)
\[
\frac{1}{c^2(x)} \frac{\partial^2 \phi}{\partial t^2} - \Delta \phi = 0, \quad (3.14)
\]
\[
\phi(0,x) = p(x), \quad u_t(0,x) = q(x).
\]
Let us assume that that (3.14) has two solutions \(\phi_1(t,x)\) and \(\phi_2(t,x)\) and set \(v(t,x) = \phi_1(t,x) - \phi_2(t,x)\). The function \(v(t,x)\) satisfies
\[
\frac{1}{c^2(x)} \frac{\partial^2 v}{\partial t^2} - \Delta v = 0, \quad (3.15)
\]
\[
v(0,x) = 0, \quad v_t(0,x) = 0.
\]
As we have discussed, the energy
\[
\mathcal{E}(t) = \frac{1}{2} \int_{\mathbb{R}^n} \left[ \frac{1}{c^2(x)} |v_t(t,x)|^2 + |\nabla v(t,x)|^2 \right] dx
\]
is preserved by evolution \(\mathcal{E}(t) = \mathcal{E}(0)\). But at \(t = 0\) we have \(v(0,x) = v_t(0,x) = 0\), hence \(\mathcal{E}(0) = 0\). It follows that \(\mathcal{E}(t) = 0\) for all \(t > 0\), which, in turn, means that \(v(t,x) \equiv \text{const}\). Since \(v(0,x) = 0\), we conclude that \(v(t,x) \equiv 0\), and thus \(\phi_1(t,x) = \phi_2(t,x)\) for all \(t \geq 0\), \(x \in \mathbb{R}^n\).

### 3.4 Finite speed of propagation

Another beautiful corollary of the energy conservation is the fact that solutions of the wave equation propagate with a finite speed. Consider the wave equation in \(\mathbb{R}^n\) with a constant speed \(c\):
\[
\frac{1}{c^2} \frac{\partial^2 u}{\partial t^2} - \Delta u = 0, \quad (3.16)
\]
\[
u(t,x) = p(x), \quad x \in \mathbb{R}^n, \quad u_t(0,x) = q(x), \quad x \in \mathbb{R}^n.
\]
Fix a point \( x_0 \in \mathbb{R}^n \) and a time \( t_0 > 0 \). We will show that \( \phi(t_0, x_0) \) depends only on the values of the functions \( p(x) \) and \( q(x) \) in the ball \( B(x_0, ct_0) \) that is centered at \( x_0 \) and has the radius \( r_0 = ct_0 \). More precisely, if \( \phi(t, x) \) solves the Cauchy problem (3.17), and \( \psi(x) \) solves (3.17) with the initial data

\[
\psi(0, x) = \tilde{p}(x), \quad \psi_t(0, x) = \tilde{q}(x),
\]

and \( \tilde{p}(x) = p(x), \) \( \tilde{q}(x) = q(x) \) for all \( x \) such that \( |x - x_0| \leq ct_0 \), then \( \phi(t_0, x_0) = \psi(t_0, x_0) \). In order to show this, consider \( \phi(t, x) = u(t, x) - \psi(t, x) \). This function satisfies

\[
\frac{1}{c^2} \frac{\partial^2 \phi}{\partial t^2} - \Delta \phi = 0, \tag{3.17}
\]

\[
\phi(0, x) = p(x) - \tilde{p}(x), \quad x \in \mathbb{R}^n, \\
\phi_t(0, x) = q(x) - \tilde{q}(x), \quad x \in \mathbb{R}^n.
\]

and

\[
p(x) - \tilde{p}(x) = q(x) - \tilde{q}(x) = 0 \text{ for all } x \text{ such that } |x - x_0| \leq ct_0. \tag{3.18}
\]

Let us define

\[
e(t) = \frac{1}{2} \int_{|x-x_0|\leq ct_0-t} \left[ \frac{1}{c^2} \phi_t^2(t, x) + |\nabla \phi(t, x)|^2 \right] dx,
\]

that is, the portion of the total energy contained in the ball \( |x - x_0| \leq c(t_0 - t) \) at a time \( 0 \leq t \leq t_0 \). Note that we have \( e(0) = 0 \) because of the initial conditions. Let us re-write \( e(t) \) in the polar coordinates centered around the point \( x_0 \):

\[
e(t) = \frac{1}{2} \int_0^{ct_0-t} \int_{S^{n-1}} \left[ \frac{1}{c^2} \phi_t^2(t, x_0 + r \omega) + |\nabla \phi(t, x_0 + r \omega)|^2 \right] r^{n-1} dS(\omega) dr,
\]

where \( S^{n-1} \) is the \((n-1)\)-dimensional sphere of all possible directions. Let us now differentiate \( e(t) \):

\[
\frac{de(t)}{dt} = \int_0^{ct_0-t} \int_{S^{n-1}} \left[ \frac{1}{c^2} \phi_t(t, x_0 + r \omega) \phi_t(t, x_0 + r \omega) + \nabla \phi(t, x_0 + r \omega) \cdot \nabla \phi(t, x_0 + r \omega) \right] \\
\quad \times r^{n-1} dS(\omega) dr \\
- \frac{c}{2} \int_{S^{n-1}} \left[ \frac{1}{c^2} \phi_t^2(t, x_0 + c(t_0 - t) \omega) + |\nabla \phi(t, x_0 + c(t_0 - t) \omega)|^2 \right] c^{n-1} (t_0 - t)^{n-1} dS(\omega) dr.
\]

Going back to the original variables this is

\[
\frac{de(t)}{dt} = \int_{|x-x_0|\leq ct_0-t} \left[ \frac{1}{c^2} \phi_t(t, x) \phi_t(t, x) + \nabla \phi(t, x) \cdot \nabla \phi(t, x) \right] dx \tag{3.20}
\]

\[
- \frac{c}{2} \int_{|y-x_0|\leq ct_0-t} \left[ \frac{1}{c^2} \phi_t^2(t, y) + |\nabla \phi(t, y)|^2 \right] dS(y).
\]
Using the fact that \( \phi \) solves the wave equation in the first line above we get

\[
\int_{|x-x_0|\leq c(t_0-t)} \left[ \frac{1}{c^2} \phi_t(t,x)\phi_{tt}(t,x) + \nabla \phi(t,x) \cdot \nabla \phi_t(t,x) \right] dx
\]

Using Green's formula in the first term on the second line gives

\[
\int_{|x-x_0|\leq c(t_0-t)} [\phi_t(t,x) \Delta \phi(t,x) + \nabla \phi(t,x) \cdot \nabla \phi_t(t,x)] dx.
\]

Going back to (3.20), we obtain

\[
\frac{de(t)}{dt} = \int_{|y-x_0|=c(t_0-t)} \left[ \phi_t(t,y) \frac{\partial \phi(t,y)}{\partial \nu} - \frac{1}{2c} \phi_t^2(t,y) - \frac{c}{2} |\nabla \phi(t,y)|^2 \right] dS(y). \tag{3.22}
\]

As \( |\partial \phi / \partial \nu| \leq |\nabla \phi| \), we get

\[
\frac{de(t)}{dt} = \int_{|y-x_0|=c(t_0-t)} \left[ |\phi_t(t,y)| \nabla \phi(t,y) - \frac{1}{2c} \phi_t^2(t,y) - \frac{c}{2} |\nabla \phi(t,y)|^2 \right] dS(y) \tag{3.23}
\]

\[
= -\frac{1}{2} \int_{|y-x_0|=c(t_0-t)} \left( \frac{1}{c} |\phi_t(t,y)| \sqrt{c} |\nabla \phi(t,y)| \right)^2 dS(y) \leq 0.
\]

We conclude that \( e(t) \leq e(0) \) for all \( 0 \leq t \leq t_0 \). Recall that \( e(0) = 0 \) (from the initial data) and \( e(t) \geq 0 \) by its very definition. Therefore, we should have \( e(t) = 0 \) for all \( 0 \leq t \leq t_0 \), which means that

\[
\phi(t,x) = 0 \text{ for all } |x-x_0| \leq ct_0,
\]

as we have claimed.

### 3.5 Plane wave solutions and the Huygens principle

A plane wave is a solution of the wave equation whose level sets are parallel planes, that is, \( u \) has the form \( u = f(k \cdot x - ct) \), where \( k \in \mathbb{R}^n \) is a fixed wave vector. Such function solves the wave equation

\[
u_{tt} - c^2 \Delta u = 0 \tag{3.24}
\]

provided that \( |k|^2 = 1 \):

\[
u_{tt} - c^2 \Delta u = c^2 f'' - c^2 \sum_{j=1}^{n} k_j^2 f'' = 0.
\]
We now discuss how a general solution of the wave equation may be decomposed into plane waves and deduce the Huygens principle. The simplest example is given by the d’Alembert formula

\[ u(t, x) = h(x - ct) + p(x + ct) \]  

(3.25)

for a general solution for the Cauchy problem for the one-dimensional wave equation that we have already discussed in Section 3.1. The d’Alembert formula (3.25) decomposes an arbitrary solution of the wave equation into a sum of a left and right going waves. In dimensions higher than one we have to decompose over waves going in all directions. Hence, we seek a general solution of the wave equation

\[ u_{tt} - c^2 \Delta u = 0 \]  

(3.26)

in the form of a sum over plane waves

\[ u(t, x) = \int_{S^{n-1}} h(k \cdot x - ct, k)dk, \]  

(3.27)

where \( S^{n-1} \) is the \((n - 1)\)-dimensional sphere of directions. As each function \( h(k \cdot x - ct, k) \) is a solution of the wave equation, so is the integral over \( k \), thus we know that \( u(t, x) \) defined by (3.27) is a solution of the wave equation. Therefore, in order to ensure that \( u(t, x) \) given by (3.27) solves the Cauchy problem for the wave equation, we have to choose the function \( h(s, k) \) so as to match the initial data

\[ u(0, x) = f(x) = \int_{S^{n-1}} h(k \cdot x, k)dk \]  

(3.28)

\[ u_t(0, x) = g(x) = -c \int_{S^{n-1}} h_s(k \cdot x, k)dk. \]  

(3.29)

Here is how that is done. If we decompose \( h(s, k) \) into its even and odd parts, that is,

\[ h(s, k) = l(s, k) + q(s, k), \quad l(s, k) = \frac{h(s, k) + h(-s, -k)}{2}, \quad q(s, k) = \frac{h(s, k) - h(-s, -k)}{2}, \]  

(3.30)

we see that (3.28)-(3.29) become

\[ f(x) = \int_{S^{n-1}} l(k \cdot x, k)dk, \quad g(x) = -c \int_{S^{n-1}} q_s(k \cdot x, k)dk. \]  

(3.31)

We have used here the fact that the integral of an odd function over a sphere vanishes. Hence, we are reduced to the following problem: given a function \( f(x) \), \( x \in \mathbb{R}^n \) find an even function \( l(s, k) \), \( s \in \mathbb{R}, k \in S^{n-1} \), such that

\[ f(x) = \int_{S^{n-1}} l(k \cdot x, k)dk, \quad l(s, k) = l(-s, -k). \]  

(3.32)

The function \( l = Rf \) is called the Radon transform of the function \( f(x) \). Let us recall how it is constructed: we start with the Fourier transform of \( f \) and re-write it in the polar coordinates:

\[ f(x) = \int e^{i \xi \cdot x} \hat{f}(\xi) \frac{d\xi}{(2\pi)^n} = \int_{S^{n-1}} \int_0^\infty e^{i \rho k \cdot x} \hat{f}(\rho k) \rho^{n-1} \frac{d\rho dk}{(2\pi)^n}. \]
Now, if the space dimension $n$ is odd (so that $\rho^{n-1} = (-\rho)^{n-1}$), we may re-write the above integral as
\[
\hat{f}(x) = \frac{1}{2} \int_{S^{n-1}} \int_{0}^{\infty} e^{i\rho k \cdot x} \hat{f}(\rho k) \rho^{n-1} \frac{d\rho dk}{(2\pi)^n} + \frac{1}{2} \int_{S^{n-1}} \int_{-\infty}^{0} e^{-i\rho k \cdot x} \hat{f}(-\rho k) \rho^{n-1} \frac{d\rho dk}{(2\pi)^n}
\]
\[
= \frac{1}{2} \int_{S^{n-1}} \int_{-\infty}^{\infty} e^{i\rho k \cdot x} \hat{f}(\rho k) \rho^{n-1} \frac{d\rho dk}{(2\pi)^n}.
\]
We have constructed the Radon transform of the function $f$ explicitly:
\[
l(s, k) = \frac{1}{2} \int_{-\infty}^{\infty} e^{i\rho s} \hat{f}(\rho k) \rho^{n-1} \frac{d\rho}{(2\pi)^n}.
\] (3.33)

As a consequence, the Fourier transform of $l(s, k)$ in $s$ is
\[
\hat{l}(\rho, k) = \frac{1}{2} \hat{f}(\rho k) \rho^{n-1} (2\pi)^{n-1},
\]
where $\hat{l}(\rho, k) = \mathcal{F}_{s \rightarrow \rho} l(s, k)$. It follows that
\[
\int_{\mathbb{R}} s^j l(s, k) ds = s^j l(0, k) = 0 \text{ for } j = 0, 1, \ldots, n - 2.
\] (3.34)

Hence, we may define anti-derivatives of $l$ in the $s$-variable up to the order $n - 1$ that decay at infinity – they are compactly supported in $s$ if $l(s, k)$ is.

Expression (3.33) is, however, not as useful as another representation for the Radon transform that we will obtain now. We start with (3.32) and take another rapidly decaying function $g(x)$ with the Radon transform $m = Rg$. Multiply (3.32) by $\bar{g}(x)$ and integrate, interchanging the order of integration and integrating over planes $x \cdot k = s$:
\[
(f, g) = \int_{S^{n-1} \times \mathbb{R}^n} l(x \cdot k, k) \bar{g}(x) dk dx = \int_{S^{n-1}} \left( \int_{\mathbb{R}} l(s, k) \left( \int_{x \cdot k = s} \bar{g}(x) d\Sigma \right) ds \right) dk
\]
\[
= \int_{S^{n-1}} \int_{\mathbb{R}} l(s, k) \tilde{M}(s, k) dk ds.
\] (3.35)

Here we have defined
\[
\tilde{M}(s, k) = \int_{x \cdot k = s} g(x) d\Sigma
\]
as the integral of the function $g$ over the hyperplane $x \cdot k = s$. On the other hand we may also write, using the Plancherel identity
\[
(f, g) = \frac{1}{2} \int_{S^{n-1}} \int_{\mathbb{R}} \hat{f}(\rho k) \bar{g}(\rho k) \rho^{n-1} \frac{d\rho dk}{(2\pi)^n} = \int_{S^{n-1}} \int_{\mathbb{R}} \hat{l}(\rho, k) \bar{m}(\rho, k) \rho^{n-1} \frac{2(2\pi)^{2n-2} d\rho dk}{(2\pi)^n}
\]
\[
= \int_{S^{n-1}} \int_{\mathbb{R}} l(s, k) \tilde{h}(s, k) ds dk,
\] (3.36)

with the function $h(s, k)$ having the following Fourier transform in $s$:
\[
\tilde{h}(\rho, k) = (2\pi)^{2n-2} \bar{m}(\rho, k) \rho^{n-1}.
\]
Comparing (3.35) and (3.36) we see that

\[ \hat{M}(\rho, k) = 2(2\pi)^{2n-2} \frac{\tilde{m}(\rho, k)}{\rho^{n-1}}. \]

We summarize this computation as

\[ (Rg)(s, k) = C_n \frac{\partial^{n-1} M(s, k)}{\partial s^{n-1}}, \quad M(s, k) = \int_{x-k=s} g(x)d\Sigma, \quad (3.37) \]

with the constant \( C_n \) that depends only on the dimension \( n \). As a consequence of (3.37) we obtain the following theorem.

**Theorem 3.1** Let the dimension \( n \) of the space be odd. Then if the function \( g(x) \) vanishes outside a ball \( \{ |x| \leq r \} \) then its Radon transform \( m(s, k) = (Rg)(s, k) \) vanishes for \( |s| > r \).

In order to formulate the Huygens principle we need the notion of the domain of influence. We say that a space-time point \( (t, z) \) belongs to the domain of influence of the origin \( \mathcal{S} \) if any two solutions whose Cauchy data coincide outside \( \mathcal{S} \) then solution of the wave equation vanishes in \( \mathcal{S} \). This is equivalent to saying that if the Cauchy data vanishes in \( \mathcal{S} \) then the Huygens principle states the following.

**Theorem 3.2** The domain of influence of the origin in an odd number of dimensions \( n \) is the double cone \( |x| = ct \).

**Proof.** We have to show that the points \( (t, x) \) with \( |x| \neq ct \) lie outside the domain of influence of the origin. That is, we have to find a ball \( U_x = \{ y \in \mathbb{R}^n : |y| < \varepsilon \} \) and a space-time neighborhood \( V_{t,x} \) of \( (t, x) \) so that any two solutions whose Cauchy data coincide outside \( U_x \) then solution of the wave equation vanishes in \( V_{t,x} \). Note first that if \( |x| > ct \), then we can choose \( \varepsilon = (|x| - ct)/2 \). Indeed, if initially both \( u(0, x) \) and \( u(t, 0, x) \) vanish outside of the ball \( \{ |x| < \varepsilon \} \) then they vanish in the ball of radius \( ct \) around the point \( x \), hence the finite speed of propagation that we have already proved implies that \( u(t, x) = 0 \). Hence, we only need to consider the case \( |x| < ct \). Let \( u(0, x) = f(x) \), \( u(t, 0, x) = g(x) \) be such initial data. If we now choose \( h(s, k) \) to be

\[ h(s, k) = Rf - \frac{1}{c} \left( \frac{\partial}{\partial s} \right)^{-1} [Rg] \]

then \( u(t, x) \) is given by the plane wave decomposition (3.27), as can be seen from (3.30) and (3.31). The anti-derivative \( \partial_s^{-1} \) is well defined – see the remark after (3.34). Moreover, as both \( f(x) \) and \( g(x) \) vanish outside the ball \( U_x \), the function \( h(s, k) \) vanishes for \( |s| > \varepsilon \) as follows from Theorem 3.1 and the same remark. Therefore, expression (3.27) implies that \( u(t, x) = 0 \) if \( |x \cdot k - ct| > \varepsilon \) for all \( k \in S^{n-1} \). This condition is satisfied if we choose \( \varepsilon < |ct - x| \). Therefore, the points \( (t, x) \) with \( |x| < ct \) also lie outside the domain of influence of the origin.

\( \square \)

Theorem 3.2 is false in even dimensions – while it is true that the points in \( \{ |x| > ct \} \) lie outside the domain of influence of the origin (this follows from the finite speed of propagation results), the points inside \( \{ |x| < ct \} \) are influenced by the origin.
3.6 Very basic geometric optics

One of the main properties of the wave equations is propagation of oscillations that does not occur in the equation: compare the form of solutions of the wave equation

\[ w_{tt} - \Delta w = 0 \]

and of the heat equation

\[ h_t - \Delta h = 0 \]

with oscillatory initial data \( h(0, x) = w(0, x) = e^{ikx} \) (for the wave equation we in addition prescribe \( w_t(0, x) = 0 \)). Solution of the wave equation is

\[
w(t, x) = \frac{1}{2} e^{ik(x - ct)} + \frac{1}{2} e^{ik(x + ct)}
\]

which propagates without any decay. On the other hand, solution of the heat equation is

\[ h(t, x) = e^{ikx - |k|^2t} \]

which decays rapidly for large wave numbers \( k \) – oscillations dissipate.

The geometric optics studies oscillatory solutions and purports to estimate the error in terms of a small parameter \( \varepsilon \ll 1 \) which is the inverse of the non-dimensional wave length. We consider oscillatory solutions of the wave equation with constant coefficients:

\[ Lu = u_{tt} - c^2 \Delta u^\varepsilon = 0 \quad (3.38) \]

with oscillatory initial data \( u^\varepsilon(0, x) = e^{i\phi(x)/\varepsilon} a^\varepsilon(x), \ u^\varepsilon_t(0, x) = 0 \). When the phase \( \phi(x) = k \cdot x \) is a linear function we are back to the plane waves. Now, however, we are interested in more general phases and, rather more importantly, in the regime \( \varepsilon \ll 1 \). We look for solutions in the same form as the initial data:

\[
u^\varepsilon(t, x) = e^{i\phi(t, x)/\varepsilon} a^\varepsilon(t, x), \quad a^\varepsilon(t, x) = a_0(t, x) + \varepsilon a_1(t, x) + \ldots \quad (3.39)
\]

We insert this ansatz into (3.38) and compute

\[
L \left( e^{i\phi(t, x)/\varepsilon} a^\varepsilon(t, x) \right) = e^{i\phi(t, x)/\varepsilon} \left[ -\frac{1}{\varepsilon^2} \tilde{L}(\phi_t, \nabla \phi) a^\varepsilon + \frac{2i}{\varepsilon} V_{\phi} a^\varepsilon + \frac{i}{\varepsilon} (L\phi) a^\varepsilon + La^\varepsilon \right], \quad (3.40)
\]

where we have defined the symbol \( \tilde{L}(\omega, k) = \omega^2 - c^2 |k|^2 \) and the linear operator

\[
V_{\phi} a^\varepsilon = \frac{\partial \phi}{\partial t} \frac{\partial a^\varepsilon}{\partial t} - c^2 \nabla \phi \cdot \nabla a^\varepsilon.
\]

In order for \( u \) of the form (3.39) to be an approximate solution of (3.38) the leading order in \( \varepsilon \) in (3.40) (that is, \( O(\varepsilon^{-2}) \)) should vanish. This leads to the eikonal equation

\[
\tilde{L}(\phi_t, \nabla \phi) = \left( \frac{\partial \phi}{\partial t} \right)^2 - c^2 |\nabla \phi|^2 = 0. \quad (3.41)
\]
Solutions of the eikonal equation may be constructed in the same way as for the progressing waves, when initially $\nabla \phi_0 \neq 0$ everywhere. Then we pick up one branch of $\phi_t = \pm |\nabla \phi|$ and construct solutions by means of the Hamilton-Jacobi theory – we will return to this issue a little later.

Assuming that solutions of the eikonal equation exist, at least locally in time, we may proceed to find an equation for the amplitudes $a_j$. The term of order $\varepsilon^{-1}$ in (3.40) has the form

$$2V\phi a_0 + (L\phi)a_0 = 0. \quad (3.42)$$

This is a first order linear (once the phase $\phi$ is determined from the eikonal equation) equation for $a_0$, also known as the transport equation. We say that the characteristics of the linear operator $V\phi$ are called rays. Then (3.42) is an ODE along the rays. In order for the transport equation to have a solution for an initial value problem we require that $\phi_t \neq 0$ initially, which means that $|\nabla \phi| \neq 0$ as well. Then the eikonal equation becomes $\phi_t = \pm |\nabla \phi|$, depending on the sign of $\phi_t$ at $t = 0$, and the equation for $a_0$ admits a smooth solution.

Equations for the higher order coefficients are obtained from the higher order terms in $\varepsilon$ in (3.40): this leads to

$$2V\phi a_n + (L\phi)a_n + \frac{1}{i}La_{n-1} = 0. \quad (3.43)$$

This is a system of first order linear (again, once the phase $\phi$ is determined from the eikonal equation) equations for $a_n$ – each one has the same family of rays as its characteristics. They can be solved to provide an expansion for $a_\varepsilon$ up to any order.

The accuracy of the approximation of $u_\varepsilon(t, x)$ by a partial sum

$$u_\varepsilon^k(t, x) = e^{i\phi(t,x)/\varepsilon} \sum_{l=0}^{k} \varepsilon^l a_l(t, x)$$

depends on the chosen norm. This is related to the fact that as $u_\varepsilon$ is oscillatory, so its derivatives are large.

**Theorem 3.3** Let the initial data $\phi(0, x)$ and $a(0, x)$ be in $L^2(\mathbb{R}^n)$, and let $\bar{T}$ be such that a smooth solution of the eikonal equation (3.41) exists for all $0 \leq t < \bar{T}$ and let $0 \leq T < \bar{T}$. There exists a constant $C_T$ so that we have, uniformly in $0 \leq t \leq T$, the following estimate:

$$\|u_\varepsilon(t, \cdot) - u_\varepsilon^k(t, \cdot)\|_{L^2(\mathbb{R}^n)} \leq C_T\varepsilon^k. \quad (3.44)$$

The constant $C_T$ depends on the time $T$ and the Cauchy data for $u_\varepsilon$ at $t = 0$.

## 4 Method of characteristics

**The linear equations**

Consider a linear first order equation

$$\sum_{j=1}^{n} a_j(x) \frac{\partial u}{\partial x_j} = c(x)u + f(x), \quad x \in \mathbb{R}^n, \quad (4.1)$$

69
where \(a_j(x), c(x)\) and \(f(x)\) are all prescribed functions. We also prescribe the solution \(u(x)\) along some \((n-1)\)-dimensional surface \(\Gamma: u(y) = g(y)\) for all \(y \in \Gamma\), with a prescribed function \(g\). The question we are interested in, is whether we can extend \(u(x)\) outside of \(\Gamma\). The idea of the method of characteristics is as follows: we look for a family of curves \(X(s; y)\), that start at a point \(y\) on \(\Gamma\) \((X(0; y) = y)\), and \(s\) is a parameter along the curve, such that \(U(s) = u(X(s))\) can be computed "easily" given that we know \(U(0) = g(y)\). This will extend the function \(u\) to all points where characteristics can reach in a unique way.

In order to understand what characteristic curves we should choose, let us differentiate \(U(s)\) with respect to \(s\):

\[
\frac{dU}{ds} = \sum_{j=1}^{n} \frac{\partial u(X(s))}{\partial x_j} \frac{dX_j(s)}{ds}.
\] (4.2)

In order to relate (4.2) to the partial differential equation (4.1) it is convenient to choose \(X(s)\) as a solution of a system of ordinary differential equations

\[
\frac{dX_j}{ds} = a_j(X(s)), \quad j = 1, \ldots, n,
\] (4.3)

with a prescribed initial condition \(X(0) = y\), where \(y \in \Gamma\) is a given point on \(\Gamma\). With this choice of the curve \(X(s)\), equation (4.2) becomes

\[
\frac{dU}{ds} = \sum_{j=1}^{n} \frac{\partial u(X(s))}{\partial x_j} a_j(s) = c(X(s))U(s) + f(s).
\] (4.4)

Now, if we can solve the characteristics equations (4.3), we can insert the solution \(X(s)\) into the scalar ODE

\[
\frac{dU}{ds} = c(X(s))U(s) + f(X(s)),
\] (4.5)

and solve it. This will provide the solution to the PDE with the boundary condition, at the points that can be reached by the characteristics.

Consider an example

\[
u_{x_1} + u_{x_2} + u = 1,
\] (4.6)

subject to the initial condition

\[
u(y_1, y_2) = \sin y_1 \text{ on the curve } y_2 = y_1 + y_1^2, \quad y_1 > 0.
\] (4.7)

The characteristic curves are

\[
\frac{dX_1}{ds} = 1, \quad \frac{dX_2}{ds} = 1, \quad X_1(0) = y_1, \quad X_2(0) = y_2,
\]

so that

\[
X_1(s) = y_1 + s, \quad X_2(s) = y_2 + s.
\]

The ODE for \(U(s)\) is

\[
\frac{dU}{ds} = 1 - U(s), \quad U(0) = \sin y_1,
\]
whose solution is

\[ U(s) = 1 - (1 - \sin y_1) e^{-s}. \]

How do we find a solution at a given point \((x_1, x_2)\)? First, we look for a characteristic that passes through \((x_1, x_2)\):

\[ x_1 = y_1 + s, \quad x_2 = y_2 + s, \quad y_2 = y_1 + y_1^2, \quad y_1 > 0. \]

Let us now solve for \(y_1\): \( s = x_1 - y_1 \), hence \( y_2 = x_2 - (x_1 - y_1) \), thus

\[ x_2 - (x_1 - y_1) = y_1 + y_1^2, \]

hence

\[ y_1 = \sqrt{x_2 - x_1}. \]

We see immediately several issues: first, solution exists only if \(x_2 > x_1\) – otherwise, a characteristic curve that passes through a point \((x_1, x_2)\) does not cross the curve \(\Gamma\) at all, meaning that solution at \((x_1, x_2)\) is not defined! If \(x_2 > x_1\), then \(s\) is

\[ s = x_1 - y_1 = x_1 - \sqrt{x_2 - x_1}, \quad (4.8) \]

and

\[ u(x_1, x_2) = 1 - (1 - \sin(\sqrt{x_2 - x_1})) e^{-x_1 + \sqrt{x_2 - x_1}}. \]

Note that this solution is not differentiable at the points where \(x_1 = x_2\) – this is because this characteristic curve is tangent to the curve \(\Gamma\) at \((0, 0)\). In general, for the boundary value problem to have a unique solution characteristics should not be tangent to \(\Gamma\) at any point.

If we consider the same example as before but now prescribe the initial data along the whole curve \(y_2 = y_1 + y_1^2\), without the restriction \(y_1 > 0\) we will see that at each point \((x_1, x_2)\) with \(x_2 > x_1\) the characteristic curve will cross \(\Gamma\) at two points, with

\[ y_1 = \pm \sqrt{x_2 - x_1}. \]

This is problematic – which \(s\) should we choose then in the formula for \(u(x_1, x_2)\)? This problem comes, once again, from the fact that the characteristic is tangent to \(\Gamma\) at the point \((0, 0)\).

Generally, solutions exist only in a neighborhood of \(\Gamma\) and only if the characteristics coming out of \(\Gamma\) are never tangent to it.

Let us consider another example:

\[ -x_2 \frac{\partial u}{\partial x_1} + x_1 \frac{\partial u}{\partial x_2} = 0, \quad (4.9) \]

with the initial condition \(u(x_1, 0) = \psi(x_1)\). The characteristics are

\[ \frac{dX_1}{ds} = -X_2(s), \quad \frac{dX_2}{ds} = X_1(s), \quad X_1(0) = y_1, \quad X_2(0) = 0. \]

The solution is

\[ X_2(s) = y_1 \sin s, \quad X_1(s) = y_1 \cos s. \]

71
The equation for $U(s)$ is
\[ \frac{dU}{ds} = 0, \quad U(0) = \psi(y_1), \]
hence
\[ U(s) = \psi(y_1). \]
Therefore, characteristics are circles, and given a point $(x_1, x_2)$, we have
\[ y_1^2 = x_1^2 + x_2^2. \]
It follows that we can not define a solution in any open region that includes all of the real line – but if we restrict $\Gamma$ to be $\Gamma' = \{(x_1, 0) : x_1 > 0\}$, we can solve the problem with the solution being
\[ u(x_1, x_2) = \psi(\sqrt{x_1^2 + x_2^2}). \]
Once again, this solution is not differentiable at the origin, where the characteristics have a singular point.

**Nonlinear equations**

Now, we consider a nonlinear first-order PDE
\[ F(\nabla u, u, x) = 0, \quad (4.10) \]
with a boundary condition
\[ u = g \text{ on } \Gamma, \quad (4.11) \]
where $\Gamma$ is a given curve. Once again, we consider a curve $X(s)$ and try to choose it so that the evolution of $U(s) = u(X(s))$ would be tractable. Let us also introduce a vector-value function $P(s) = \nabla u(X(s))$ and consider
\[ \frac{dP_j}{ds} = n \sum_{i=1}^{n} \frac{\partial^2 u(X(s))}{\partial x_i \partial x_j} \frac{dX_i(s)}{ds}. \quad (4.12) \]
Let us also differentiate (4.10) with respect to $x_i$, with the variable $p = \nabla u$ in the argument of $F$:
\[ \frac{\partial F}{\partial x_i} + \frac{\partial F}{\partial u} \frac{\partial u}{\partial x_i} + \sum_{j=1}^{n} \frac{\partial F}{\partial p_j} \frac{\partial^2 u}{\partial x_i \partial x_j} = 0. \quad (4.13) \]
Let us then set the characteristic equation for $X(s)$ as:
\[ \frac{dX_i}{dt} = \frac{\partial F(P(s), U(s), X(s))}{\partial p_i}. \quad (4.14) \]
Then (4.12) becomes:
\[ \frac{dP_j}{ds} = \sum_{i=1}^{n} \frac{\partial^2 u(X(s))}{\partial x_i \partial x_j} \frac{dX_i(s)}{ds} = \sum_{i=1}^{n} \frac{\partial^2 u(X(s))}{\partial x_i \partial x_j} \frac{\partial F(P(s), U(s), X(s))}{\partial p_i}. \quad (4.15) \]
Using the differentiated equation (4.13) this becomes

\[
\frac{dP_j}{ds} = -\frac{\partial F(P(s), U(s), X(s))}{\partial x_j} - \frac{\partial F(P(s), U(s), X(s))}{\partial u} P_j(s).
\]  

(4.16)

Finally, the equation for \(U(s)\) is

\[
\frac{dU}{ds} = \sum_{j=1}^{n} P_j(s) \frac{\partial F(P(s), U(s), X(s))}{\partial p_j}.
\]

(4.17)

To summarize: we need to consider the full system of the characteristic equations

\[
\begin{align*}
\frac{dX_j}{ds} &= \frac{\partial F(P(s), U(s), X(s))}{\partial p_j} \\
\frac{dU}{ds} &= \sum_{j=1}^{n} P_j(s) \frac{\partial F(P(s), U(s), X(s))}{\partial p_j} \\
\frac{dP_j}{ds} &= -\frac{\partial F(P(s), U(s), X(s))}{\partial x_j} - \frac{\partial F(P(s), U(s), X(s))}{\partial u} P_j(s).
\end{align*}
\]

(4.18)

Let us consider an example

\[
\frac{\partial u}{\partial x_1} \frac{\partial u}{\partial x_2} = u,
\]

(4.19)

in the right half-plane \(\{x_1 > 0\}\) with the boundary condition \(u(0, x_2) = x_2^2\). Now, \(F(p, u, x) = p_1 p_2 - u\), so characteristics become

\[
\begin{align*}
\frac{dX_1}{ds} &= P_2(s), & \frac{dX_2}{ds} &= P_1(s) \\
\frac{dU}{ds} &= 2P_1(s)P_2(s), \\
\frac{dP_1}{ds} &= P_1(s), & \frac{dP_2}{ds} &= P_1(s).
\end{align*}
\]

(4.20)

Integrating these equations gives

\[
P_1(s) = P_1(0)e^s, \quad P_2(s) = P_2(0)e^s, \quad U(s) = U(0) + P_1(0)P_2(0) (e^{2s} - 1),
\]

and

\[
X_1(s) = X_1(0) + P_2(0)(e^s - 1), \quad X_2(s) = X_2(0) + P_1(0)(e^s - 1).
\]

(4.21)

At the boundary line \(\{x_1 = 0\}\) we have \(X_1(0) = 0, U(0) = (X_2(0))^2\) and \(P_2(0) = 2X_2(0)\). We may also find \(P_1(0)\) from the PDE

\[
\frac{\partial u}{\partial x_1} \frac{\partial u}{\partial x_2} = u,
\]

which at the point \((0, X_2(0))\) becomes

\[
P_1(0)P_2(0) = U(0),
\]

73
which gives $P_1(0) = X_2(0)/2$. Now, given a point $(x_1, x_2)$ with $x_1 > 0$ let us find the point $(0, X_2(0))$ such that the characteristic coming out of $(0, X_2(0))$ passes through $(x_1, x_2)$: using this in (4.20) leads to

$$x_1 = 2X_2(0)(e^s - 1), \quad x_2 = X_2(0) + \frac{1}{2}X_2(0)(e^s - 1),$$

so that

$$X_2(0) = x_2 - \frac{1}{4}x_1, \quad e^s = \frac{x_1 + 4x_2}{4x_2 - x_1}.$$  

It follows that

$$u(x_1, x_2) = U(0) + P_1(0)P_2(0) (e^{2s} - 1) = U(0)e^{2s} = (X_2(0))^2 e^{2s} = \frac{(x_1 + 4x_2)^2}{16}. \quad (4.22)$$

### 5 Basic introduction to conservation laws

#### 5.1 Rankine-Hugoniot conditions and integral solutions

We will consider here the initial value problem for the scalar conservation laws in one dimension:

$$u_t + (F(u))_x = 0, \quad t \geq 0, \quad x \in \mathbb{R}, \quad (5.1)$$

with a prescribed initial data $u(0, x) = g(x)$. The most canonical example is $F(u) = u^2/2$, which is known as the Burgers’ equation:

$$u_t + uu_x = 0. \quad (5.2)$$

As we will see, solutions of (5.1) do not necessarily stay continuous for all times $t > 0$ even if the initial data $g(x)$ is infinitely differentiable. Hence, we need first to devise a way of formulating the initial value problem for the smooth solutions of the PDE in a way that does not involve derivatives. That formulation would apply equally well then to non-smooth solutions. The main difficulty will be to choose this ”derivative-free” formulation in a physically meaningful way. To begin, we let $u(t, x)$ be a (smooth) solution of (5.1) and multiply this equation by a smooth ”test function” $v(t, x)$ of compact support, and integrate by parts (the terms at $x = \pm \infty$ vanish because $v(t, x)$ has compact support):

$$0 = \int_0^\infty \int_{-\infty}^{\infty} v(t, x)(u_t + (F(u))_x)dxdt$$

$$= -\int_0^\infty \int_{-\infty}^{\infty} [v_t u + F(u)v_x]dxdt - \int_{-\infty}^{\infty} v(0, x)u(0, x)dx. \quad (5.3)$$

Taking into account the initial condition gives

$$0 = \int_0^\infty \int_{-\infty}^{\infty} [v_t u + F(u)v_x]dxdt + \int_{-\infty}^{\infty} v(0, x)g(x)dx. \quad (5.4)$$

This identity should hold for any smooth compactly supported function $v(t, x)$. The advantage of (5.4) is that it makes sense for any bounded function $u(t, x)$ – it does not involve any
derivatives of $u(t, x)$. We say that $u(t, x)$ is an integral solution of (5.1) if the integral identity (5.3) holds for any smooth function $v(t, x)$ of compact support.

In order to understand the implications of this definition, consider a particularly simple piece-wise constant solution $u(t, x)$ such that $u(t, x) = u_l$ for $x < x(t)$, and $u(t, x) = u_r$ for $x > x(t)$. We would like to understand how the point $x(t)$ should evolve for $u(t, x)$ to be an integral solution of the conservation law (5.1). Take a smooth test function $v(t, x)$ that vanishes at $t = 0$, then (5.4) is

$$0 = \int_0^\infty \int_{-\infty}^\infty [v_t u + F(u) v_x] dx dt$$  \hspace{1cm} (5.5)$$

First, we note that

$$\int_0^\infty \int_{-\infty}^{x(t)} F(u_l) v_x dx dt + \int_0^\infty \int_{x(t)}^\infty F(u_r) v_x dx dt = \int_0^\infty [F(u_l)v(t, x(t)) - F(u_r)v(t, x(t))] dt.$$  \hspace{1cm} (5.6)

In order to deal with the time integral we note that

$$\int_0^\infty \int_{-\infty}^{x(t)} u_l v_t dt dx = \int_0^\infty \frac{d}{dt} \left( \int_{-\infty}^{x(t)} u_l v(t, x) dx \right) dt - \int_0^\infty u_l v(t, x(t)) \dot{x}(t) dt. \hspace{1cm} (5.7)$$

As $v(0, x) = 0$ for all $x$, and $v(t, x) = 0$ for all $t \geq T$ for some $T > 0$, the first term in the right side vanishes, giving

$$\int_0^\infty \int_{-\infty}^{x(t)} u_l v_t dt dx = -\int_0^\infty u_l v(t, x(t)) \dot{x}(t) dt.$$  \hspace{1cm} (5.8)$$

And similarly we have

$$\int_0^\infty \int_{x(t)}^\infty u_r v_t dt dx = \int_0^\infty \frac{d}{dt} \left( \int_{x(t)}^\infty u_r v(t, x) dx \right) dt + \int_0^\infty u_r v(t, x(t)) \dot{x}(t) dt$$

$$= \int_0^\infty u_r v(t, x(t)) \dot{x}(t) dt.$$ \hspace{1cm} (5.9)$$

Using expressions (5.6)-(5.9) in (5.5) gives

$$0 = -\int_0^\infty u_l v(t, x(t)) \dot{x}(t) dt + \int_0^\infty u_r v(t, x(t)) \dot{x}(t) dt + \int_0^\infty [F(u_l)v(t, x(t)) - F(u_r)v(t, x(t))] dt.$$ \hspace{1cm} (5.10)$$

or

$$0 = \int_0^\infty v(t, x(t))[-u_l \dot{x}(t) + u_r \dot{x}(t) + F(u_l) - F(u_r)] dt. \hspace{1cm} (5.11)$$

As this equality holds for all functions $v(t, x)$, we conclude that

$$\dot{x}(t) = \frac{F(u_l) - F(u_r)}{u_l - u_r}. \hspace{1cm} (5.12)$$
This is known as the Rankine-Hugoniot condition.

The Rankine-Hugoniot condition can be in a very straightforward manner generalized to solutions that are not piece-wise constant but piecewise smooth: let \( u(t, x) \) be an integral solution of (5.1) such that \( u(t, x) \) is smooth at \( x < x(t) \), and at \( x > x(t) \) but has a jump discontinuity at \( x = x(t) \) such that the limits

\[
u_l(t) = \lim_{x \to x(t)^-} u(t, x), \quad u_r(t) = \lim_{x \to x(t)^+} u(t, x),\]

exist. Then a nearly identical computation as above shows that \( u(t, x) \) should satisfy

\[
u_t + (F(u))_x = 0, \quad (5.13)\]

both in the region \( x < x(t) \) and in \( x > x(t) \), while the Rankine-Hugoniot condition should still hold:

\[
\dot{x}(t) = \frac{F(u_l(t)) - F(u_r(t))}{u_l(t) - u_r(t)}. \quad (5.14)
\]

The basic example of such solution is a shock of the Burgers' equation:

\[
u_t + \nu\nu_x = 0, \quad (5.15)
\]

with the initial condition

\[
u(0, x) = \begin{cases} 1, & x < 0 \\ 0, & x > 0. \end{cases} \quad (5.16)
\]

Then the solution is piece-wise constant:

\[
u(t, x) = \begin{cases} 1, & x < x(t) \\ 0, & x > x(t). \end{cases} \quad (5.17)
\]

The discontinuity point \( x(t) \) moves according to the Rankine-Hugoniot condition corresponding to the flux \( F(u) = u^2/2, u_l = 1, u_r = 0 \):

\[
\dot{x} = \frac{1/2 - 0}{1 - 0} = \frac{1}{2},
\]

and the shock position is \( x(t) = t/2 \) – the shock moves with the constant speed \( v = 1/2 \):

\[
u(t, x) = \begin{cases} 1, & x < t/2 \\ 0, & x > t/2. \end{cases} \quad (5.18)
\]

Let us now consider a flipped initial data:

\[
u(0, x) = \begin{cases} 0, & x < 0 \\ 1, & x > 0. \end{cases} \quad (5.19)
\]

Then the Rankine-Hugoniot condition would give

\[
\dot{x}(t) = \frac{0 - 1/2}{0 - 1} = \frac{1}{2},
\]
and the solution is

\[ u(t, x) = \begin{cases} 
0, & x < t/2 \\
1, & x > t/2.
\end{cases} \quad (5.20) \]

However, there is another integral solution with the same initial data:

\[ u(t, x) = \begin{cases} 
0, & x < 0, \\
x/t, & 0 < x < t, \\
1, & x > t.
\end{cases} \quad (5.21) \]

This solution is called a rarefaction wave. It is easy to check that \( u(t, x) \) is an integral solution because it is continuous, piece-wise differentiable, and on each interval where it is differentiable, it solves (5.1). This example shows that an integral solution is not, generally, unique, and an important question is to distinguish between various integral solutions.

### 5.2 Entropy condition

Let us look at the Burgers’ equation

\[ u_t + uu_x = 0, \quad (5.22) \]

and try to solve it by the method of characteristics. The full characteristics curves

\[(X(s), T(s), U(s), P_x(s), P_t(s))\]

satisfy

\[\begin{align*}
\frac{dT}{ds} &= 1, \\
\frac{dX}{ds} &= U, \\
\frac{dU}{ds} &= P_x U + P_t, \\
\frac{dP_x}{ds} &= -P_x^2, \\
\frac{dP_t}{ds} &= -P_x P_t,
\end{align*}\]

starting at

\[ T(0) = 0, \; X(0) = y, \; U(0) = g(y), \; P_x(0) = g'(y), \; P_t(0) = -g(y)g'(y). \]

The solution is \( T(s) = s \),

\[ P_x(s) = \frac{P_x(0)}{1 + sP_x(0)} = \frac{g'(y)}{g(y) + s}, \; P_t(s) = -g(y)P_x(s) = -\frac{g(y)g'(y)}{1 + sg'(y)}, \]

and, finally, \( U(s) = U(0) = g(y) \) – the function \( u(t, x) \) is constant along the characteristics. The projection of the characteristics on the physical space are straight lines

\[ X(t) = y + g(y)t. \quad (5.23) \]

We can not prevent the characteristics from crossing each other in the physical space – this is what leads to singularities and shocks. However, we can impose a condition that would ensure that if we start a characteristic at a point \((t, x)\) and run it backwards in time, then it will not hit any other characteristic before it reaches \( t = 0 \). To guarantee this, we require that if \( x = x(t) \) is the discontinuity curve for the solution, then no characteristic comes out of this curve. This is known as the entropy condition. In the case of a shock of the Burgers’ equation this is ensured if \( u_l > u_r \) – this follows from the simple geometric consideration on the the
plane \((x, t)\), the explicit formula (5.23) for the characteristic, and the fact that \(U(s) = g(y)\) along the characteristic.

Going back to the example of the initial data (5.19) for the Burgers’ equation, we see that the discontinuous solution does not satisfy the entropy condition, while the rarefaction wave does.

How does this change for a general conservation law? Consider a general equation of the form

\[
    u_t + F'(u)u_x = 0. \tag{5.24}
\]

The characteristics

\[
    (X(s), T(s), U(s), P_x(s), P_t(s))
\]

now satisfy

\[
    \frac{dT}{ds} = 1, \quad \frac{dX}{ds} = F'(U), \quad \frac{dU}{ds} = P_x F'(U) + P_t, \quad \frac{dP_x}{ds} = -F''(U) P_x^2, \quad \frac{dP_t}{ds} = -F''(U) P_x P_t,
\]

starting at

\[
    T(0) = 0, \quad X(0) = y, \quad U(0) = g(y), \quad P_x(0) = g'(y), \quad P_t(0) = -g(y)g'(y).
\]

The solution is \(T(s) = s\),

\[
    U(s) = U(0) = g(y), \quad P_x(s) = \frac{P_x(0)}{1 + sF''(g(y))P_x(0)} = \frac{g'(y)}{g'(y) + F''(g(y)) s},
\]

\[
    P_t(s) = -g(y)P_x(s) = -\frac{g(y)g'(y)}{1 + sF''(g(y))g'(y)}.
\]

The function \(u(t, x)\) is again constant along the characteristics. The projection of the characteristics on the physical space are now the straight lines

\[
    X(t) = y + F'(g(y)) t. \tag{5.25}
\]

The entropy condition for a jump discontinuity is now \(F'(u_l) > F'(u_r)\). If the flux is convex, this is equivalent to the (simpler!) condition

\[
    u_l > u_r.
\]

Let us now consider a step function initial data

\[
    u(0, x) = \begin{cases} 
        0, & x < 0 \\
        1, & 0 < x < 1, \\
        0, & x > 1.
    \end{cases} \tag{5.26}
\]

For \(0 \leq t \leq 2\) we have a rarefaction wave that follows a shock:

\[
    u(0, x) = \begin{cases} 
        0, & x < 0 \\
        x/t, & 0 < x < t, \\
        1, & t < x < 1 + t/2, \\
        0, & x > 1 + t/2.
    \end{cases} \tag{5.27}
\]
At the time $t = 2$ the rarefaction wave (whose front moves with the speed $v = 1$) catches up with the shock that moves with the speed $v = 1/2$. After this time solution is

$$u(0, x) = \begin{cases} 
0, & x < 0 \\
x/t, & 0 < x < x(t), \\
0, & x > x(t).
\end{cases} \tag{5.28}$$

The position $x(t)$ is determined by the Rankine-Hugoniot condition: $u_l = x(t)/t$, $u_r = 0$, so

$$\dot{x}(t) = \frac{u_l}{2} = \frac{x}{2t}, \quad x(2) = 2,$$

hence $x(t) = \sqrt{2t}$. Therefore, solution is

$$u(0, x) = \begin{cases} 
0, & x < 0 \\
x/t, & 0 < x < \sqrt{2t}, \\
0, & x > \sqrt{2t}.
\end{cases} \tag{5.29}$$

The maximum of the solution behaves as

$$u_{\text{max}}(t) = u(t, \sqrt{2t}) = \frac{\sqrt{2t}}{t} = \sqrt{\frac{2}{t}}, \tag{5.30}$$

that is, it is decaying in time. This brings a question: where does this dissipation come from?

### 5.3 The Lax-Oleinik formula for the Burgers’ equation

We will now obtain the Lax-Oleinik formula for the solution of the initial value problem

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} = 0, \tag{5.31}$$

with the initial data $u(0, x) = g(x)$ that is uniformly bounded: $|g(x)| \leq M$. We will assume for simplicity that $0 \leq g(x) \leq M$ and that $g(x)$ is piece-wise continuous. The Lax-Oleinik formula is obtained as follows: consider the function

$$s(t, x; y) = \frac{t}{2} \left( \frac{x - y}{t} \right)^2 + h(y), \tag{5.32}$$

with

$$h(y) = \int_0^y g(z)dz.$$ 

Let us see where the function $s(t, x; y)$ is minimized over $y$: $s_y(t, x, y) = 0$ if

$$x = y + tg(y), \quad (5.33)$$

because $g(y) = h'(y)$. As $g(y)$ is a bounded function, the function $h(y)$ satisfies $|h(y)| \leq M|y|$, and so $s(t, x; y)$ tends to $+\infty$ as $y \to \pm\infty$ for all $t$ and $x$ fixed. Moreover, $s(t, x; y)$ is continuous in $y$ for all $t$ and $x$ fixed. It follows that it attains its global minimum at some points $y$. We let $y(t, x)$ be the smallest such point.
Let us set
\[ w(t, x) = \frac{t}{2} \left( \frac{x - y(t, x)}{t} \right)^2 + h(y(t, x)) = \inf_{y \in \mathbb{R}} \left[ \frac{t}{2} \left( \frac{x - y}{t} \right)^2 + h(y) \right]. \] (5.34)

The function \( w(t, x) \) is Lipschitz continuous in \( x \) because for any \( x \) and \( x' \) we have (denote \( y = y(t, x) \))
\[ w(t, x') - w(t, x) = \inf_{z \in \mathbb{R}} \left[ \frac{t}{2} \left( \frac{x' - z}{t} \right)^2 + h(z) \right] - \frac{t}{2} \left( \frac{x - y}{t} \right)^2 - h(y) \leq h(x' - x + y) - h(y) \leq M|x' - x|. \] (5.35)

We took \( z = x' + y - x \) above as a "test point". Note also that as \( t \to 0 \) we have \( y(t, x) \to x - \) the first term in the definition of \( s(t, x, y) \) is blowing up as \( t \to 0 \) for \( x \neq y \), hence the minimizer has to converge to \( x \). As a consequence, we have
\[ w(t, x) \to h(x) \text{ as } t \to 0. \] (5.36)

Let us find an equation for the function \( w(t, x) \). The function \( y(t, x) \) where it is differentiable, satisfies
\[ y_x + tg'(y)y_x = 1, \]
and
\[ y_t + g(y) + tg'(y)y_t = 0. \]
Both of these relations follow from differentiating (5.33). It follows that
\[ y_t = -g(y)y_x. \] (5.37)

Let us now compute:
\[ w_t = -\frac{1}{2t^2} (x - y)^2 + \frac{y - x}{t} y_t + h'(y)y_t, \] (5.38)
and
\[ w_x = \frac{1}{t} (x - y)(1 - y_x) + h'(y)y_x. \] (5.39)

Using (5.33) and since \( h'(y) = g(y) \), we obtain
\[ w_t = -\frac{1}{2} g^2(y) - g(y)y_t + g(y)y_t = \frac{g^2(y)}{2}, \] (5.40)
and
\[ w_x = g(y)(1 - y_x) + g(y)y_x = g(y). \] (5.41)

We conclude that \( w(t, x) \) is Lipschitz continuous almost everywhere, and wherever it is differentiable, it solves the initial value problem
\[ w_t + \frac{w_x^2}{2} = 0, \quad w(0, x) = h(x). \] (5.42)
Therefore, we may define (almost everywhere)
\[
    u(t, x) = \frac{\partial w(t, x)}{\partial x} = \frac{\partial}{\partial x} \left[ \frac{t}{2} \left( \frac{x - y(t, x)}{t} \right)^2 - h(y(t, x)) \right].
\]  
(5.43)

Note that (5.41) and (5.33) imply that
\[
    u(t, x) = g(y(t, x)) = \frac{x - y(t, x)}{t}.
\]  
(5.44)

Let us show that \( u(t, x) \) is an integral solution of the initial value problem for the Burgers’ equation
\[
    u_t + uu_x = 0, \quad u(0, x) = g(x).
\]  
(5.45)

The fact that wherever \( u(t, x) \) is differentiable, it solves the Burgers’ equation follows immediately from (5.42) after taking the \( x \)-derivative. Let us see why it is an integral solution. Let \( v(t, x) \) be a test function and multiply (5.42) by \( v_x \):
\[
    \int_0^\infty \int_{-\infty}^\infty v_x \left( w_t + \frac{w_x^2}{2} \right) \, dxdt = 0. 
\]  
(5.46)

Now, since \( w \) is sufficiently regular in time, we may integrate by parts in time, leading to
\[
    \int_0^\infty \int_{-\infty}^\infty v_x w_t \, dxdt = - \int_0^\infty \int_{-\infty}^\infty v_{xt} w \, dxdt - \int_{-\infty}^\infty v_x w(0, x) \, dx.
\]

Next, we may integrate by parts in \( x \) since \( w \) is Lipschitz continuous in \( x \):
\[
    \int_0^\infty \int_{-\infty}^\infty v_x w_t \, dxdt = - \int_0^\infty \int_{-\infty}^\infty v_{xt} w \, dxdt - \int_{-\infty}^\infty v_x w(0, x) \, dx
    = - \int_0^\infty \int_{-\infty}^\infty v_t w_x \, dxdt + \int_{-\infty}^\infty v w_x(0, x) \, dx = - \int_0^\infty \int_{-\infty}^\infty v_t u \, dxdt + \int_{-\infty}^\infty v(0, x) g(x) \, dx.
\]

Inserting this into (5.46) gives
\[
    \int_0^\infty \int_{-\infty}^\infty \left( v_t u + v_x \frac{u_x^2}{2} \right) \, dxdt + \int_{-\infty}^\infty v(0, x) g(x) \, dx = 0, 
\]  
(5.47)

which means exactly that \( u(t, x) \) is an integral solution of (5.45).

Expression (5.43) is known as the Lax-Oleinik formula for the solution of the Burgers’ equation.

### 5.4 Entropy solutions and the Lax-Oleinik formula

Let us first show that the function \( u(t, x) \) given by the Lax-Oleinik formula satisfies a one-sided estimate
\[
    u(t, x + z) - u(t, x) \leq \frac{z}{t}, \quad \text{for all } t > 0, \ x \in \mathbb{R} \text{ and } z > 0. 
\]  
(5.48)
This estimate holds simply under the assumption that \( g \) is a bounded function. In order to establish (5.48) we first show that \( y(t, x) \) is non-decreasing in \( x \). Indeed, let \( x_1 < x_2 \) and \( t > 0 \), and \( y_1 = y(t, x_1), y_2 = y(t, x_2) \). We claim that \( s(t, x_2, y) > s(t, x_2, y_1) \) for all \( y < y_1 \), which means that \( y_2 \geq y_1 \). To see this, we note that

\[
(x_2 - y_1)^2 + (x_1 - y)^2 < (x_1 - y_1)^2 + (x_2 - y)^2.
\]

Observe also that, as \( y_1 \) is the smallest global minimizer of \( s(t, x_1, y) \), and \( y < y_1 \), we have

\[
t \left( \frac{x_1 - y_1}{t} \right)^2 + h(y_1) < t \left( \frac{x_1 - y}{t} \right)^2 + h(y).
\]

Combing this with (5.49) gives

\[
\frac{(x_2 - y_1)^2}{t} + \frac{(x_1 - y)^2}{t} + h(y_1) < \left( \frac{x_1 - y_1}{t} \right)^2 + \frac{(x_2 - y)^2}{t} + h(y_1)
\]

\[
< \left( \frac{x_1 - y}{t} \right)^2 + \left( \frac{x_2 - y}{t} \right)^2 + h(y),
\]

which is nothing but

\[
\frac{(x_2 - y_1)^2}{t} + h(y_1) < \frac{(x_2 - y)^2}{t} + h(y),
\]

and thus \( s(t, x_2, y) > s(t, x_2, y_1) \) for all \( y < y_1 \), which implies that \( y_2 > y_1 \).

We now use (5.44), and the fact that \( y(t, x) \) is non-decreasing in \( x \):

\[
u(t, x) = \frac{x - y(t, x)}{t} \geq \frac{x - y(t, x + z)}{t} = \frac{x + z - y(t, x + z)}{t} - \frac{z}{t} = u(t, x + z) - \frac{z}{t}, \tag{5.50}
\]

which is (5.50).

We will say that \( u(t, x) \) is an entropy solution of the Burgers’ equation if it satisfies the entropy condition (5.48). Note that if \( u(t, x) \) is discontinuous at some point and has left and right limits \( u_l \) and \( u_r \), then (5.48) implies that \( u_l > u_r \), which is our old entropy condition that we have introduced for piecewise continuous solutions. However, (5.48) is a more general form of the entropy condition as it does not require \( u(t, x) \) to be piecewise continuous. The key result is the following theorem.

**Theorem 5.1** There exists a unique entropy solution of the initial value problem

\[
u_t + uu_x = 0, \quad u(0, x) = g(x), \tag{5.51}
\]

with \( g \in L^\infty(\mathbb{R}) \), and it is given by the Lax-Oleinik formula.

We will not prove this result here.
5.5 Long time behavior

Let us now address the long-time decay of the solutions of the Burgers’ equation that we have mentioned previously. Consider the entropy solution of (5.51) and let us assume that

\[ \int_{-\infty}^{\infty} |g(x)| \, dx \leq M < +\infty. \]  

(5.52)

It follows that

\[ |h(y)| = \left| \int_{0}^{y} g(s) \, ds \right| \leq M, \]

as well. Then, we have

\[ s(t, x; y) = \frac{(x - y)^2}{2t} + h(y) \geq \frac{(x - y)^2}{2t} - M. \]

On the other hand, we also have

\[ s(t, x; x) = h(x) \leq M, \]

hence at the minimizing point \( y(t, x) \) we have

\[ \frac{(x - y(t, x))^2}{2t} + h(y(t, x)) \leq M, \]

whence

\[ \frac{(x - y(t, x))^2}{2t} \leq 2M, \]  

(5.53)

for all \( t > 0 \). Now, we simply use expression (5.44) for \( u(t, x) \):

\[ u(t, x) = \frac{x - y(t, x)}{t} \leq \frac{\sqrt{4Mt}}{t} = \frac{\sqrt{4M}}{\sqrt{t}}. \]  

(5.54)

Hence, in the long time regime the entropy solutions of the Burgers’ equation decay as \( O(t^{-1/2}) \). Therefore, the definition of the entropy solution includes an implicit form some “hidden” dissipation.

Finally, we establish ”decay to an \( N \)-wave” result for the solutions of the Burgers’ equation. Let \( g(x) \) have compact support and set

\[ p = -2 \min_{y \in \mathbb{R}} \int_{-\infty}^{y} g(s) \, ds, \quad q = 2 \max_{y \in \mathbb{R}} \int_{y}^{\infty} g(s) \, ds. \]

Define the \( N \)-wave as

\[ N(t, x) = \begin{cases} \frac{x}{t}, & \text{if } -\sqrt{pt} < x < \sqrt{qt}, \\ 0, & \text{otherwise.} \end{cases} \]

Theorem 5.2 Assume that \( p, q > 0 \). Then there exists \( C > 0 \) so that

\[ \int_{-\infty}^{\infty} |u(t, x) - N(t, x)| \, dx \leq \frac{C}{\sqrt{t}}. \]  

(5.55)
Proof. First, note that if \( g(x) = 0 \) for \( |x| \geq R \), then \( h(x) = h_- \) for all \( x < -R \), and \( h(x) = h_+ \) for all \( x > R \), with

\[
h_- = -\int_{-R}^{0} g(s) ds, \quad h_+ = \int_{0}^{R} g(s) ds.
\]

One can easily check that

\[
\min_y h(y) = -\frac{p}{2} + h_- = -\frac{q}{2} + h_+.
\]

(5.56)

We now claim that

\[
u(t, x) = 0 \quad \text{for all} \quad x < -R - (pt)^{1/2}, \quad \text{for all} \quad t > 0.
\]

(5.57)

In order to establish this, we will show that for \( x \) as in (5.57), we have

\[
s(t, x; x) = h(x) = h_-,
\]

while for \( y < -R \) we have

\[
s(t, x; y) = \frac{(x - y)^2}{2t} + h(y) = \frac{(x - y)^2}{2t} + h_- > h_-,
\]

unless \( y = x \). On the other hand, for \( y > -R \) we have

\[
s(t, x) = \frac{(x - y)^2}{2t} + h(y) \geq \frac{(x + R)^2}{2t} - \frac{p}{2} + h_- \geq \frac{pt}{2t} - \frac{p}{2} + h_- = h_-.
\]

Hence, \( y(t, x) = x \) is the smallest global minimizer of \( s(t, x; y) \). Then (5.44) implies that \( u(t, x) = 0 \). Similarly, we can show that

\[
u(t, x) = 0 \quad \text{for all} \quad x > R + (qt)^{1/2}, \quad \text{for all} \quad t > 0.
\]

(5.58)

Next, we claim that (for \( t \) sufficiently large so that this interval is not empty)

\[-R \leq y(t, x) \leq R, \quad \text{if} \quad R - \sqrt{pt} + 1 < x < -R + \sqrt{qt} - 1.
\]

(5.59)

Since \( y(t, x) \) is non-decreasing in \( x \) it suffices to show that

\[
y(t, R - \sqrt{pt} + 1) \geq -R,
\]

(5.60)

and

\[
y(t, -R + \sqrt{qt} - 1) \leq R.
\]

(5.61)

To see that (5.60) holds, take \( x_0 = R - \sqrt{pt} + 1 \). Note that if \( z < -R \), then \( h(z) = h_- \), and

\[
s(t, x_0; z) = \frac{(x_0 - z)^2}{2t} + h(z) \geq h_-.
\]

Choose now \( z_0 \) so that \( h(z) = \min_{y \in \mathbb{R}} h(y) = -p/2 + h_- \), and \(|z| \leq R\). Then we have

\[
s(t, x_0, z_0) = \frac{(x_0 - z_0)^2}{2t} + h(z_0) < \frac{pt}{2t} - \frac{p}{2} + h_- = h_-.
\]
It follows that the minimizer $y(t, x_0)$ of $s(t, x_0, y)$ satisfies (5.60). The proof of (5.61) is very similar. As

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

we conclude that

$$\left| u(t, x) - \frac{x}{t} \right| \leq \frac{C}{t} \quad \text{for} \quad R - \sqrt{pt} < x < -R + \sqrt{qt}.$$ (5.62)

So we have shown that $u(t, x) = 0$ for $x < -R - \sqrt{pt}$ and $x > R + \sqrt{qt}$, and, in addition, (5.62) holds for $R - \sqrt{pt} < x < -R + \sqrt{qt}$. As a consequence, we have

$$|u(t, x) - N(t, x)| \leq \frac{C}{t}, \quad \text{for} \quad x \notin (R - \sqrt{pt}, R - \sqrt{pt}) \text{ and } x \notin (R - \sqrt{qt}, R + \sqrt{qt}).$$

However, for $x \in (-R - \sqrt{pt}, R - \sqrt{pt})$ we have $u(t, x) \leq C/\sqrt{t}$, and $N(t, x) = x/t \leq C/\sqrt{t}$, and similarly for $x \in (R - \sqrt{qt}, R + \sqrt{qt})$. Putting all these pieces together implies that

$$\int_R \left| u(t, x) - N(t, x) \right| \leq \frac{C}{\sqrt{t}},$$

as we have claimed. □