SISSA - Università di Trieste
Corso di Laurea Magistrale in Matematica
A. A. 2009/2010
Istituzioni di Fisica Matematica, Modulo A

Boris DUBROVIN, with additions by M. Bertola

July 2, 2019

Contents

Chapter 1

Linear differential operators

1.1 Definitions and main examples

Let $\Omega \subset \mathbb{R}^d$ be an open subset. Denote $\mathcal{C}^\infty(\Omega)$ the set of all infinitely differentiable complex valued smooth functions on Ω. The Euclidean coordinates on $\mathbb{R}^d$ will be denoted $x_1, \ldots, x_d$. We will use short notations for the derivatives

$$\partial_k = \frac{\partial}{\partial x_k}$$

and we also introduce operators

$$D_k = -i\,\partial_k, \quad k = 1, \ldots, d. \tag{1.1.1}$$

For a multiindex

$$\mathbf{p} = (p_1, \ldots, p_d)$$

denote

$$\begin{aligned}
|\mathbf{p}| &= p_1 + \ldots + p_d \\
\mathbf{p}! &= p_1! \ldots p_d! \\
\mathbf{x}^{\mathbf{p}} &= x_1^{p_1} \ldots x_d^{p_d} \\
\partial^{\mathbf{p}} &= \partial_1^{p_1} \ldots \partial_d^{p_d}, \quad D^{\mathbf{p}} = D_1^{p_1} \ldots D_d^{p_d}.
\end{aligned}$$

The derivatives, as well as the higher order operators $D^{\mathbf{p}}$ define linear operators

$$D^{\mathbf{p}} : \mathcal{C}^\infty(\Omega) \to \mathcal{C}^\infty(\Omega), \quad f \mapsto D^{\mathbf{p}} f = (-i)^{|\mathbf{p}|} \frac{\partial^{|\mathbf{p}|} f}{\partial x_1^{p_1} \ldots \partial x_d^{p_d}}.$$

More generally, we will consider *linear differential operators* of the form

$$\begin{aligned}
A &= \sum_{|\mathbf{p}| \leqslant m} a_{\mathbf{p}}(x) D^{\mathbf{p}} \\
a_{\mathbf{p}}(x) &\in \mathcal{C}^\infty(\Omega) \\
A &: \mathcal{C}^\infty(\Omega) \to \mathcal{C}^\infty(\Omega).
\end{aligned} \tag{1.1.2}$$

We will define the *order* of the linear differential operator by

$$\operatorname{ord} A = \max|\mathbf{p}| \quad \text{such that} \quad a_{\mathbf{p}}(x) \neq 0. \tag{1.1.3}$$

Main examples are

1. Laplace operator
$$\Delta = \partial_1^2 + \ldots + \partial_d^2 = -(D_1^2 + \ldots D_d^2) \tag{1.1.4}$$

2. Heat operator
$$\frac{\partial}{\partial t} - \Delta \tag{1.1.5}$$

acting on functions on the $(d+1)$-dimensional space with the coordinates $(t, x_1, \ldots, x_d)$.

3. Wave operator
$$\frac{\partial^2}{\partial t^2} - \Delta. \tag{1.1.6}$$

4. Schrödinger operator
$$i\,\frac{\partial}{\partial t} + \Delta. \tag{1.1.7}$$

1.2 Principal symbol of a linear differential operator

Symbol of a linear differential operator (1.1.2) is a function
$$a(x,\xi) = \sum_{|\mathbf{p}|\leqslant m} a_{\mathbf{p}}(x)\xi^{\mathbf{p}}, \quad x \in \Omega \subset \mathbb{R}^d, \quad \xi \in \mathbb{R}^d. \tag{1.2.1}$$

If the order of the operator is equal to m then the *principal symbol* is defined by
$$a_m(x,\xi) = \sum_{|\mathbf{p}|=m} a_{\mathbf{p}}(x)\xi^{\mathbf{p}}. \tag{1.2.2}$$

The symbols (1.2.1), (1.2.2) are polynomials in d variables $\xi_1, \ldots, \xi_d$ with coefficients being smooth functions on Ω.

For the above examples we have the following symbols

1. For the Laplace operator Δ the symbol and principal symbol coincide
$$a = a_2 = -(\xi_1^2 + \ldots + \xi_d^2) \equiv -\xi^2.$$

2. For the heat equation the full symbol is
$$a = i\,\tau + \xi^2$$

while the principal symbol is ξ^2.

3. For the wave operator again the symbol and principal symbols coincide
$$a = a_2 = -\tau^2 + \xi^2.$$

4. The symbol of the Schrödinger operator is
$$-(\tau + \xi^2)$$

while the principal symbol is ξ^2.

Exercise 1.1 *Prove the following formula for the symbol of a linear differential operator*

$$a(x, i\xi) = e^{-i\,x\cdot\xi} A\left(e^{i\,x\cdot\xi}\right). \tag{1.2.3}$$

Here we use the notation

$$x \cdot \xi = x_1\xi_1 + \ldots + x_d \cdot \xi_d$$

for the natural pairing $\mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$.

Exercise 1.2 *Given a linear differential operator* A *with* constant *coefficients denote* $a(\xi)$ *its symbol (it does not depend on* x *for linear differential operators with constant coefficients). Prove that the exponential function*

$$u(x) = e^{i\,x\cdot\xi}$$

is a solution to the linear differential equation

$$A\,u = 0$$

iff *the vector* ξ *satisfies*

$$a(\xi) = 0.$$

Exercise 1.3 *Prove that for a pair of smooth functions* $u(x)$, $S(x)$ *and a linear differential operator* A *of order* m *the expression of the form*

$$e^{-i\,\lambda\,S(x)} A\left(u(x)e^{i\,\lambda\,S(x)}\right)$$

is a polynomial in λ *of degree* m. *Derive the following expression for the leading coefficient of this polynomial*

$$e^{-i\,\lambda\,S(x)} A\left(u(x)e^{i\,\lambda\,S(x)}\right) = i^m u(x) a_m(x, S_x(x))\lambda^m + O(\lambda^{m-1}). \tag{1.2.4}$$

Here

$$S_x = \left(\frac{\partial S}{\partial x_1}, \ldots, \frac{\partial S}{\partial x_d}\right)$$

is the gradient of the function $S(x)$.

Exercise 1.4 *Let* A *and* B *be two linear differential operators of orders* k *and* l *with the principal symbols* $a_k(x, \xi)$ *and* $b_l(x, \xi)$ *respectively. Prove that the superposition* $C = A \circ B$ *is a linear differential operator of order* $\leqslant k + l$. *Prove that the principal symbol of* C *is equal to*

$$c_{k+l}(x, \xi) = a_k(x, \xi)\, b_l(x, \xi) \tag{1.2.5}$$

in the case $\operatorname{ord} C = \operatorname{ord} A + \operatorname{ord} B$. *In the case of strict inequality* $\operatorname{ord} C < \operatorname{ord} A + \operatorname{ord} B$ *prove that the product (1.2.5) of principal symbols is identically equal to zero.*

The formula for computing the full symbol of the product of two linear differential operators is more complicated. We will give here the formula for the particular case of one spatial variable x.

Exercise 1.5 *Let* $a(x, \xi)$ *and* $b(x, \xi)$ *be the symbols of two linear differential operators* A *and* B *with one spatial variable. Prove that the symbol of the superposition* $A \circ B$ *is equal to*

$$a \star b = \sum_{k \geqslant 0} \frac{(-i)^k}{k!} \partial_\xi^k a\, \partial_x^k b. \tag{1.2.6}$$

1.3 Change of independent variables

Let us now analyze the transformation rules of the principal symbol $a(x, \xi)$ of an operator A under smooth invertible changes of variables

$$y_i = y_i(x), \quad i = 1, \dots, n. \tag{1.3.1}$$

Recall that the first derivatives transform according to the chain rule

$$\frac{\partial}{\partial x_i} = \sum_{k=1}^{d} \frac{\partial y_k}{\partial x_i} \frac{\partial}{\partial y_k}. \tag{1.3.2}$$

The transformation law of higher order derivatives is more complicated. For example

$$\frac{\partial^2}{\partial x_i \partial x_j} = \sum_{k,l=1}^{d} \frac{\partial y_k}{\partial x_i} \frac{\partial y_l}{\partial x_j} \frac{\partial^2}{\partial y_k \partial y^l} + \sum_{k=1}^{d} \frac{\partial^2 y_k}{\partial x_i \partial x_j} \frac{\partial}{\partial y_k}$$

etc. However it is clear that after the transformation one obtains again a linear differential operator of the same order m. More precisely define the operator

$$\tilde{A} = \sum (-i)^{|\mathbf{p}|} a_{\mathbf{p}}(y) \frac{\partial^{|\mathbf{p}|}}{\partial y_1^{p_1} \dots \partial y_d^{p_d}}$$

by the equation

$$A\, f(y(x)) = \left(\tilde{A}\, f(y) \right)_{y=y(x)}.$$

The transformation law of the principal symbol is of particular simplicity as it follows from the following

Proposition 1.6 *Let $a_m(x, \xi)$ be the principal symbol of a linear differential operator A. Denote $\tilde{a}_m(y, \tilde{\xi})$ the principal symbol of the same operator written in the coordinates y, i.e., the principal symbol of the operator $\tilde{A}$. Then*

$$a_m(y(x), \tilde{\xi}) = a_m(x, \xi) \quad \text{provided} \quad \xi_i = \sum_{k=1}^{d} \frac{\partial y_k}{\partial x_i} \tilde{\xi}_k. \tag{1.3.3}$$

Proof: Applying the formula (1.2.4) one easily derives the equality

$$\begin{aligned}
&a_m(x, S_x) = \tilde{a}_m(y, S_y) \\
&y = y(x) \\
&S_x = \left(\frac{\partial S}{\partial x_1}, \dots, \frac{\partial S}{\partial x_d} \right), \quad S_y = \left(\frac{\partial S}{\partial y_1}, \dots, \frac{\partial S}{\partial y_d} \right).
\end{aligned}$$

Applying the chain rule

$$\frac{\partial S}{\partial x_i} = \sum_{k=1}^{d} \frac{\partial y_k}{\partial x_i} \frac{\partial S}{\partial y_k}$$

we arrive at the transformation rule (1.3.3) for the particular case

$$\xi_i = \frac{\partial S}{\partial x_i}, \quad \tilde{\xi}_k = \frac{\partial S}{\partial y_k}.$$

This proves the proposition since the gradients can take arbitrary values. ∎

1.4 Canonical form of linear differential operators of order $\leqslant 2$ with constant coefficients

Consider a first order linear differential operator

$$A = a_1 \frac{\partial}{\partial x_1} + \ldots + a_d \frac{\partial}{\partial x_d} \tag{1.4.1}$$

with constant coefficients $a_1, \ldots, a_d$. One can find a linear transformation of the coordinates

$$\xi_i = \sum_{k=1}^{d} c_{ki} \tilde{\xi}_k, \quad i = 1, \ldots, d \tag{1.4.2}$$

that maps the vector $a = (a_1, \ldots, a_d)$ to the unit coordinate vector of the axis y_d. After such a transformation the operator A becomes the partial derivative operator

$$A = \frac{\partial}{\partial y_d}.$$

Therefore the general solution of the first order linear differential equation

$$A\,\varphi = 0$$

can be written in the form

$$\varphi(y_1, \ldots, y_d) = \varphi_0(y_1, \ldots, y_{d-1}). \tag{1.4.3}$$

Here φ_0 is an arbitrary smooth function of $(d-1)$ variables.

Exercise 1.7 *Prove that the general solution to the equation*

$$A\,\varphi + b\,\varphi = 0 \tag{1.4.4}$$

with A of the form (1.4.1) and a constant b reads

$$\varphi(y_1, \ldots, y_d) = \varphi_0(y_1, \ldots, y_{d-1}) e^{-b\, y_d}$$

for an arbitrary $\mathcal{C}^1$ function $\varphi_0(y_1, \ldots, y_{d-1})$.

Consider now a second order linear differential operator of the form

$$A = \sum_{i,j=1}^{d} a_{ij} \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{d} b_i \frac{\partial}{\partial x_i} + c \tag{1.4.5}$$

with constant coefficients. Without loss of generality one can assume the coefficient matrix a_{ij} to be symmetric. Denote

$$Q(\xi) = -a_2(x,\xi) = \sum_{i,j=1}^{d} a_{ij}\xi_i\xi_j \tag{1.4.6}$$

the quadratic form coinciding with the principal symbol, up to a common sign. Recall the following theorem from linear algebra.

Theorem 1.8 *There exists a linear invertible change of variables of the form (1.4.2) reducing the quadratic form (1.4.6) to the form*

$$Q = \tilde\xi_1^2 + \ldots + \tilde\xi_p^2 - \tilde\xi_{p+1}^2 - \ldots - \tilde\xi_{p+q}^2. \tag{1.4.7}$$

The numbers $p \geqslant 0$, $q \geqslant 0$, $p+q \leqslant d$ do not depend on the choice of the reducing transformation.

Note that, according to the Proposition 1.6 the transformation (1.4.2) corresponds to the linear invertible change of independent variables $x \to y$ of the form

$$y_k = \sum_{i=1}^{d} c_{ki}x_i, \quad k = 1,\ldots,d. \tag{1.4.8}$$

Invertibility means that the coefficient matrix of the transformation does not degenerate:

$$\det\left(c_{ki}\right)_{1\leqslant k,i\leqslant d} \neq 0.$$

We arrive at

Corollary 1.9 *A second order linear differential operator with constant coefficients can be reduced to the form*

$$A = \frac{\partial^2}{\partial y_1^2} + \ldots + \frac{\partial^2}{\partial y_p^2} - \frac{\partial^2}{\partial y_{p+1}^2} - \ldots - \frac{\partial^2}{\partial y_{p+q}^2} + \sum_{k=1}^{d}\tilde b_k y_k + c \tag{1.4.9}$$

by a linear transformation of the form (1.4.8). The numbers p and q do not depend on the choice of the reducing transformation.

1.5 Elliptic and hyperbolic operators. Characteristics

Let $a_m(x,\xi)$ be the principal symbol of a linear differential operator A.

Definition 1.10 *It is said that the operator $A : \mathcal{C}^\infty(\Omega) \to \mathcal{C}^\infty(\Omega)$ is* elliptic *if*

$$a_m(x,\xi) \neq 0 \quad \text{for any} \quad \xi \neq 0, \quad x \in \Omega. \tag{1.5.1}$$

For example the Laplace operator

$$\Delta = \frac{\partial^2}{\partial x_1^2} + \ldots + \frac{\partial^2}{\partial x_n^2}$$

is elliptic on $\Omega = \mathbb{R}^d$. The *Tricomi operator*

$$A = \frac{\partial^2}{\partial x^2} + x\,\frac{\partial^2}{\partial y^2} \tag{1.5.2}$$

is elliptic on the right half plane $x > 0$.

Definition 1.11 *Given a point $x_0 \in \Omega$, the hypersurface in the ξ-space defined by the equation*

$$a_m(x_0, \xi) = 0 \tag{1.5.3}$$

is called characteristic cone *of the operator A at x_0. The vectors ξ satisfying (1.5.3) are called* characteristic vectors *at the point x_0.*

Observe that the hypersurface (1.5.3) is invariant with respect to rescalings

$$\xi \mapsto \lambda\xi \quad \forall\ \lambda \in \mathbb{R} \tag{1.5.4}$$

since the polynomial $a_m(x_0, \xi)$ is homogeneous of degree m:

$$a_m(x, \lambda\,\xi) = \lambda^m a_m(x, \xi).$$

The characteristic cone of an elliptic operator is one point $\xi = 0$. For the example of wave operator

$$A = \frac{\partial^2}{\partial t^2} - \Delta, \quad \Delta = \frac{\partial^2}{\partial x_1^2} + \ldots + \frac{\partial^2}{\partial x_d^2} \tag{1.5.5}$$

the characteristic cone is given by the equation

$$\tau^2 - \xi_1^2 - \ldots - \xi_d^2 = 0. \tag{1.5.6}$$

Thus it coincides with the standard cone in the Euclidean $(d+1)$-dimensional space. The characteristic cone of the heat operator

$$\frac{\partial}{\partial t} - \Delta \tag{1.5.7}$$

is the τ-line

$$\xi_1 = \ldots = \xi_d = 0. \tag{1.5.8}$$

Definition 1.12 *The hypersurface in $\mathbb{R}^d$ is called* characteristic surface *or simply* characteristics *for the operator A if at every point x of the surface the normal vector ξ is a characteristic vector:*

$$a_m(x, \xi) = 0.$$

If the hypesurface is defined by a local equation

$$S(x) = 0 \tag{1.5.9}$$

then $S(x)$ satisfies the equation

$$a_m\left(x, S_x(x)\right) = 0 \tag{1.5.10}$$

at every point of the hypersurface (1.5.9).

As it follows from the Proposition 1.6 the characteristics do not depend on the choice of a system of coordinates.

Example. For a first order linear differential operator

$$A = a_1(x)\frac{\partial}{\partial x_1} + \ldots + a_d(x)\frac{\partial}{\partial x_d} \tag{1.5.11}$$

the function $S(x)$ defining a characteristic hypersurface must satisfy the equation

$$A\, S(x) = 0. \tag{1.5.12}$$

It is therefore a first integral of the following system of ODEs

$$\begin{aligned} \dot x_1 &= a_1(x_1, \ldots, x_d) \\ &\ldots \\ \dot x_d &= a_d(x_1, \ldots, x_d) \end{aligned} \tag{1.5.13}$$

Indeed, the equation (1.5.12) says that the function $S(x)$ is constant along the integral curves of the system (1.5.13). It is known from the theory of ordinary differential equations that locally, near a point x^0 such that $\big(a_1(x^0), \ldots, a_d(x^0)\big) \neq 0$ there exists a smooth invertible change of coordinates

$$(x_1, \ldots, x_d) \mapsto (y_1, \ldots, y_d), \quad y_k = y_k(x_1, \ldots, x_d)$$

such that, in the new coordinates the system reduces to the form

$$\begin{aligned} \dot y_1 &= 0 \\ &\ldots \\ \dot y_{d-1} &= 0 \\ \dot y_d &= 1 \end{aligned} \tag{1.5.14}$$

(the so-called rectification of a vector field). For the particular case of constant coefficients the needed transformation is linear (see above). In these coordinates the general solution to the equation (1.5.12) reads

$$S(y_1, \ldots, y_d) = S_0(y_1, \ldots, y_{d-1}). \tag{1.5.15}$$

Hyperbolic operators. Let us consider a linear differential operator A acting on smooth functions on a domain Ω in the $(d+1)$-dimensional space with Euclidean coordinates $(t, x_1, \ldots, x_d)$. Denote $a_m(t, x, \tau, \xi)$ the principal symbol of this operator. Here

$$\tau \in \mathbb{R}, \quad \xi = (\xi_1, \ldots, \xi_d) \in \mathbb{R}^d.$$

Recall that the principal symbol of an operator of order m is a polynomial of degree m in τ, ξ_1, $\ldots$, ξ_d.

Definition 1.13 *The linear differential operator A is called* hyperbolic *with respect to the time variable t if for any fixed $\xi \neq 0$ and any $(t, x) \in \Omega$ the equation for τ*

$$a_m(t, x, \tau, \xi) = 0 \tag{1.5.16}$$

has m pairwise distinct real roots

$$\tau_1(t, x, \xi), \ldots, \tau_m(t, x, \xi).$$

For brevity we will often say that a linear differential operator is hyperbolic if all its characteristics are real and pairwise distinct. For elliptic operators the characteristics are purely imaginary.

The wave operator (1.5.5) gives a simple example of a hyperbolic operator. Indeed, the equation

$$\tau^2 = \xi_1^2 + \ldots + \xi_d^2$$

has two distinct roots

$$\tau = \pm\sqrt{\xi_1^2 + \ldots + \xi_d^2}$$

for any $\xi \neq 0$. The heat operator (1.5.7) is neither hyperbolic nor elliptic.

Finding the j-th characteristic of a hyperbolic operator requires knowledge of solutions to the following Hamilton–Jacobi equation for the functions $S = S(x, t)$

$$\frac{\partial S}{\partial t} = \tau_j\left(t, x, \frac{\partial S}{\partial x}\right). \tag{1.5.17}$$

From the course of analytical mechanics it is known that the latter problem is reduced to integrating the Hamilton equations

$$\left.\begin{array}{rcl} \dot{x}_i & = & \frac{\partial H(t,x,p)}{\partial p_i} \\ \dot{p}_i & = & -\frac{\partial H(t,x,p)}{\partial x_i} \end{array}\right\} \tag{1.5.18}$$

with the time-dependent Hamiltonian $H(t, x, p) = \tau_j(t, x, p)$. In the next section we will consider the particular case $d = 1$ and apply it to the problem of canonical forms of the second order linear differential operators in a two-dimensional space.

1.6 Reduction to a canonical form of second order linear differential operators in a two-dimensional space

Consider a linear differential operator

$$A = a(x, y)\frac{\partial^2}{\partial x^2} + 2b(x, y)\frac{\partial^2}{\partial x \partial y} + c(x, y)\frac{\partial^2}{\partial y^2}, \quad (x, y) \in \Omega \subset \mathbb{R}^2. \tag{1.6.1}$$

The characteristics of these operator are curves

$$x = x(t), \quad y = y(t).$$

Here t is some parameter on the characteristic. Let (dx, dy) be the tangent vector to the curve. Then the normal vector $(-dy, dx)$ must satisfy the equation

$$a(x, y)dy^2 - 2b(x, y)dx\,dy + c(x, y)dx^2 = 0. \tag{1.6.2}$$

Assuming $a(x, y) \neq 0$ one obtains a quadratic equation for the vector dy/dx

$$a(x, y)\left(\frac{dy}{dx}\right)^2 - 2b(x, y)\frac{dy}{dx} + c(x, y) = 0. \tag{1.6.3}$$

The operator (1.6.1) is hyperbolic *iff* the discriminant of this equation is positive:

$$b^2 - a\,c > 0. \tag{1.6.4}$$

For elliptic operators the discriminant is strictly negative.

For a hyperbolic operator one has two families of characteristics to be found from the ODEs

$$\frac{dy}{dx} = \frac{b(x,y) + \sqrt{b^2(x,y) - a(x,y)\,c(x,y)}}{a(x,y)} \tag{1.6.5}$$

$$\frac{dy}{dx} = \frac{b(x,y) - \sqrt{b^2(x,y) - a(x,y)\,c(x,y)}}{a(x,y)}. \tag{1.6.6}$$

Let

$$\phi(x,y) = c_1, \quad \psi(x,y) = c_2 \tag{1.6.7}$$

be the equations of the characteristics[1]. Here c_1 and c_2 are two integration constants. Such curves pass through any point $(x,y) \in \Omega$. Moreover they are not tangent at every point. Let us introduce new local coordinates u, v by

$$u = \phi(x,y), \quad v = \psi(x,y). \tag{1.6.8}$$

Lemma 1.14 *The change of coordinates*

$$(x,y) \mapsto (u,v)$$

is locally invertible. Moreover the inverse functions

$$x = x(u,v), \quad y = y(u,v)$$

are smooth.

Proof: We have to check non-vanishing of the Jacobian

$$\det \begin{pmatrix} \partial u/\partial x & \partial u/\partial y \\ \partial v/\partial x & \partial v/\partial y \end{pmatrix} = \det \begin{pmatrix} \phi_x & \phi_y \\ \psi_x & \psi_y \end{pmatrix} \neq 0. \tag{1.6.9}$$

By definition the first derivatives of the functions ϕ and ψ correspond to two different roots of the same quadratic equation

$$a(x,y)\phi_x^2 + 2b(x,y)\phi_x\phi_y + c(x,y)\phi_y^2 = 0, \quad a(x,y)\psi_x^2 + 2b(x,y)\psi_x\psi_y + c(x,y)\psi_y^2 = 0.$$

The determinant (1.6.9) vanishes *iff* the gradients of ϕ and ψ are proportional:

$$(\phi_x, \phi_y) \sim (\psi_x, \psi_y).$$

This contradicts the requirement to have the roots distinct. ■

Let us rewrite the linear differential operator A in the new coordinates:

$$A = \tilde a(u,v)\frac{\partial^2}{\partial u^2} + 2\tilde b(u,v)\frac{\partial^2}{\partial u \partial v} + \tilde c(u,v)\frac{\partial^2}{\partial v^2} + \dots \tag{1.6.10}$$

where the dots stand for the terms with the low order derivatives.

[1]The function $\phi(x,y)$, resp. $\psi(x,y)$, is a first integral for the ODE (1.6.5), resp. (1.6.6), that is, it takes constant values along the integral curves of this differential equation.

Theorem 1.15 *In the new coordinates the linear differential operator reads*

$$A = 2\tilde{b}(u,v)\frac{\partial^2}{\partial u\,\partial v} + \dots$$

Proof: In the new coordinates the characteristic have the form

$$u = c_1, \quad v = c_2$$

for arbitrary constants c_1 and c_2. Therefore their tangent vectors $(1,0)$ and $(0,1)$ must satisfy the equation for characteristics

$$\tilde{a}(u,v)dv^2 - 2\tilde{b}(u,v)du\,dv + \tilde{c}(u,v)dv^2 = 0.$$

This implies $\tilde{a}(u,v) = \tilde{c}(u,v) = 0$. ■

For the case of elliptic operator (1.6.1) the analogue of the differential equations (1.6.5), (1.6.6) are complex conjugated equations

$$\frac{dy}{dx} = \frac{b \pm i\sqrt{a\,c - b^2}}{a}, \quad a = a(x,y), \quad b = b(x,y), \quad c = c(x,y). \tag{1.6.11}$$

Assuming analyticity of the functions $a(x,y)$, $b(x,y)$, $c(x,y)$ one can prove existence of a complex valued first integral

$$S(x,y) = \phi(x,y) + i\,\psi(x,y) \tag{1.6.12}$$

satisfying

$$a\,S_x + \left(b - i\sqrt{a\,c - b^2}\right)S_y = 0. \tag{1.6.13}$$

Let us introduce new system of coordinates by

$$u = \phi(x,y), \quad v = \psi(x,y). \tag{1.6.14}$$

Exercise 1.16 *Prove that the transformation*

$$(x,y) \mapsto (u,v)$$

is locally smoothly invertible. Prove that the operator A in the new coordinates takes the form

$$A = \tilde{a}(u,v)\left(\frac{\partial^2}{\partial u^2} + \frac{\partial^2}{\partial v^2}\right) + \dots \tag{1.6.15}$$

with some nonzero smooth function $\tilde{a}(u,v)$. Like above the dots stand for the terms with lower order derivatives.

Let us now consider the case of linear differential operators of the form (1.6.1) with identically vanishing discriminant

$$b^2(x,y) - a(x,y)\,c(x,y) \equiv 0. \tag{1.6.16}$$

Operators of this class are called *parabolic*. In this case we have only one characteristic to be found from the equation

$$\frac{dy}{dx} = \frac{b(x,y)}{a(x,y)}. \tag{1.6.17}$$

Let $\phi(x,y)$ be a first integral of this equation

$$a\,\phi_x + b\,\phi_y = 0, \quad \phi_x^2 + \phi_y^2 \neq 0. \tag{1.6.18}$$

Choose an arbitrary smooth function $\psi(x,y)$ such that

$$\det \begin{pmatrix} \phi_x & \phi_y \\ \psi_x & \psi_y \end{pmatrix} \neq 0.$$

In the coordinates

$$u = \phi(x,y), \quad v = \psi(x,y)$$

the coefficient $\tilde{a}(u,v)$ vanishes, since the line $\phi(x,y) = \text{const}$ is a characteristic. But then the coefficient $\tilde{b}(u,v)$ must vanish either because of vanishing of the discriminant

$$\tilde{b}^2 - \tilde{a}\,\tilde{c} = 0.$$

Thus the canonical form of a parabolic operator is

$$A = \tilde{c}(u,v)\frac{\partial^2}{\partial v^2} + \dots \tag{1.6.19}$$

where the dots stand for the terms of lower order.

1.7 General solution of a second order hyperbolic equation with constant coefficients in the two-dimensional space

Consider a hyperbolic operator

$$A = a\frac{\partial^2}{\partial x^2} + 2b\frac{\partial^2}{\partial x\,\partial y} + c\frac{\partial^2}{\partial y^2} \tag{1.7.1}$$

with constant coefficients a, b, c satisfying the hyperbolicity condition

$$b^2 - a\,c > 0.$$

The equations for characteristics (1.6.5), (1.6.6) can be easily integrated. This gives two linear first integrals

$$\begin{gathered} u = y - \lambda_1 x, \quad v = y - \lambda_2 x \\ \lambda_{1,2} = \frac{b \pm \sqrt{b^2 - a\,c}}{a}. \end{gathered} \tag{1.7.2}$$

In the new coordinates the hyperbolic equation $A\,\varphi = 0$ reduces to

$$\frac{\partial^2\varphi}{\partial u\,\partial v} = 0. \tag{1.7.3}$$

The general solution to this equation can be written in the form

$$\varphi = f(y - \lambda_1 x) + g(y - \lambda_2 x) \tag{1.7.4}$$

where f and g are two arbitrary smooth[2] functions of one variable.

For example consider the wave equation

$$\varphi_{tt} = a^2\varphi_{xx} \tag{1.7.5}$$

where a is a positive constant. The general solution reads

$$\varphi(x,t) = f(x - a\,t) + g(x + a\,t). \tag{1.7.6}$$

Observe that $f(x - a\,t)$ is a right-moving wave propagating with constant speed a. In a similar way $g(x + a\,t)$ is a left-moving wave. Therefore the general solution to the wave equation (1.7.5) is a superposition of two such waves.

[2]It suffices to take the functions of the $\mathcal{C}^2$ class.

1.8 Exercises for Chapter 1

Exercise 1.17 *Reduce to the canonical form the following equations*

$$u_{xx} + 2u_{xy} - 2u_{xz} + 2u_{yy} + 6u_{zz} = 0 \tag{1.8.1}$$

$$u_{xy} - u_{xz} + u_x + u_y - u_z = 0. \tag{1.8.2}$$

Exercise 1.18 *Reduce to the canonical form the following equations*

$$x^2 u_{xx} + 2x\,y\,u_{xy} - 3y^2 u_{yy} - 2x\,u_x + 4y\,u_y + 16x^4 u = 0 \tag{1.8.3}$$

$$y^2 u_{xx} + 2x\,y\,u_{xy} + 2x^2 u_{yy} + y\,u_y = 0 \tag{1.8.4}$$

$$u_{xx} - 2u_{xy} + u_{yy} + u_x + u_y = 0 \tag{1.8.5}$$

Exercise 1.19 *Find general solution to the following equations*

$$x^2 u_{xx} - y^2 u_{yy} - 2y\,u_y = 0 \tag{1.8.6}$$

$$x^2 u_{xx} - 2x\,y\,u_{xy} + y^2 u_{yy} + x\,u_x + y\,u_y = 0. \tag{1.8.7}$$

Chapter 2

Wave equation

2.1 Vibrating string

We consider small oscillations of an elastic string on the (x, u)-plane. Let the x-axis be the equilibrium state of the string. Denote $u(x,t)$ the displacement of the point x at a time t. It will be assumed to be orthogonal to the x-axis. Thus the shape of the string at the time t is given by the graph of the function $u(x,t)$. The velocity of the string at the point x is equal to $u_t(x,t)$. We will also assume that the only force to be taken into consideration is the tension directed along the string. In particular the string will be assumed to be totally elastic.

Consider a small interval of the string from x to $x + \Delta x$. We will write the equation of motion for this interval. Denote $T = T(x)$ the tension of the string at the point x. The horizontal and vertical components at the points x and $x + \Delta x$ are equal to

$$T_{\rm hor}(x) = T_1 \cos\alpha, \quad T_{\rm vert}(x) = T_1 \sin\alpha$$
$$T_{\rm hor}(x+\Delta x) = T_2 \cos\beta, \quad T_{\rm vert}(x+\Delta x) = T_2 \sin\beta$$

where $T_1 = T(x)$, $T_2 = T(x + \Delta x)$ (see Fig. 1).

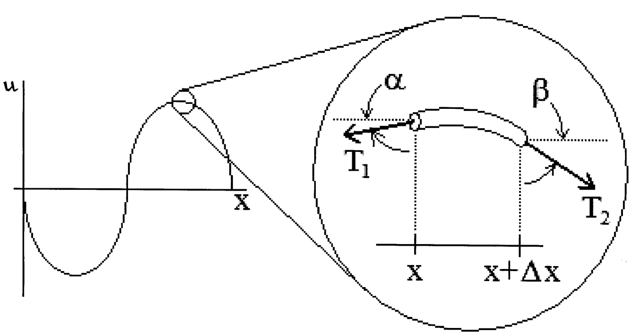

Fig. 1.

The angle α between the string and the x-axis at the point x is given by

$$\cos\alpha = \frac{1}{\sqrt{1+u_x^2}}, \quad \sin\alpha = \frac{u_x}{\sqrt{1+u_x^2}}.$$

The oscillations are assumed to be *small.* More precisely this means that the term u_x is small. So at the leading approximation we can neglect the square of it to arrive at

$$\cos\alpha \simeq 1, \quad \sin\alpha \simeq u_x(x)$$
$$\cos\beta \simeq 1, \quad \sin\beta \simeq u_x(x+\Delta x)$$

So the horizontal and vertical components at the points x and $x + \Delta x$ are equal to

$$T_{\text{hor}}(x) \simeq T_1, \quad T_{\text{vert}}(x) \simeq T_1 u_x(x)$$
$$T_{\text{hor}}(x+\Delta x) \simeq T_2, \quad T_{\text{vert}}(x+\Delta x) = T_2 u_x(x+\Delta(x),$$

Since the string moves in the u-direction, the horizontal components at the points x and $x + \Delta x$ must coincide:

$$T_1 = T(x) = T(x+\Delta x) = T_2.$$

Therefore $T(x) \equiv T = \text{const}$.

Let us now consider the vertical components. The resulting force acting on the piece of the string is equal to

$$f = T_2 \sin\beta - T_1 \sin\alpha = T\, u_x(x+\Delta x) - T\, u_x(x) \simeq T\, u_{xx}(x)\,\Delta x.$$

On another side the vertical component of the total momentum of the piece of the string is equal to

$$p = \int_x^{x+\Delta x} \rho(x) u_t(x,t)\, ds(x) \simeq \rho(x) u_t(x,t)\,\Delta x$$

where $\rho(x)$ is the linear mass density of the string and

$$ds(x) = \frac{dx}{\sqrt{1+u_x^2(x)}} \simeq dx$$

is the element of the length[1]. The second Newton law

$$p_t = f$$

in the limit $\Delta x \to 0$ yields

$$\rho(x) u_{tt} = T\, u_{xx}.$$

In particular in the case of constant mass density one arrives at the equation

$$u_{tt} = a^2 u_{xx} \tag{2.1.1}$$

where the constant a is defined by

$$a^2 = \frac{T}{\rho}. \tag{2.1.2}$$

[1] This means that the length s of the segment of the string between $x = x_1$ and $x = x_2$ is equal to

$$s = \int_{x_1}^{x_2} ds(x),$$

and the total mass m of the same segment is equal to

$$m = \int_{x_1}^{x_2} \rho(x)\, ds(x).$$

Exercise 2.1 *Prove that the* plane wave

$$u(x,t) = A\, e^{i(k\,x+\omega\,t)} \tag{2.1.3}$$

satisfies the wave equation (2.1.1) if and only if the real parameters ω *and* k *satisfy the following* dispersion relation

$$\omega = \pm a\, k. \tag{2.1.4}$$

The parameters ω and k are called resp. the *frequency*[2] and *wave number* of the plane wave. The arbitrary parameter A is called the *amplitude* of the wave. It is clear that the plane wave is periodic in x with the period

$$L = \frac{2\pi}{k} \tag{2.1.5}$$

since the exponential function is periodic with the period $2\pi\, i$. The plane wave is also periodic in t with the period

$$T = \frac{2\pi}{\omega}. \tag{2.1.6}$$

Due to linearity of the wave equation the real and imaginary parts of the solution (2.1.3) solve the same equation (2.1.1). Assuming A to be real we thus obtain the real valued solutions

$$\operatorname{Re} u = A\, \cos(k\,x + \omega\,t), \quad \operatorname{Im} u = A\, \sin(k\,x + \omega\,t). \tag{2.1.7}$$

2.2 D'Alembert formula

Let us start with considering oscillations of an *infinite string.* That is, the spatial variable x varies from $-\infty$ to ∞. The Cauchy problem for the equation (2.1.1) is formulated in the following way: find a solution $u(x,t)$ defined for $t \geqslant 0$ such that at $t = 0$ the initial conditions

$$u(x,0) = \phi(x), \quad u_t(x,0) = \psi(x) \tag{2.2.1}$$

hold true. The solution is given by the following *D'Alembert formula*:

Theorem 2.2 (D'Alembert formula) *For arbitrary initial data* $\phi(x) \in \mathcal{C}^2(\mathbb{R})$, $\psi(x) \in \mathcal{C}^1(\mathbb{R})$ *the solution to the Cauchy problem (2.1.1), (2.2.1) exists and is unique. Moreover it is given by the formula*

$$u(x,t) = \frac{\phi(x-a\,t) + \phi(x+a\,t)}{2} + \frac{1}{2a}\int_{x-a\,t}^{x+a\,t} \psi(s)\, ds. \tag{2.2.2}$$

Proof: As we have proved in Section 1.7 the general solution to the equation (2.1.1) can be represented in the form

$$u(x,t) = f(x-a\,t) + g(x+a\,t). \tag{2.2.3}$$

[2] In physics literature the number $-\omega$ is called frequency.

Let us choose the functions f and g in order to meet the initial conditions (2.2.1). We obtain the following system:

$$\begin{aligned} f(x) + g(x) &= \phi(x) \\ a\left[g'(x) - f'(x)\right] &= \psi(x). \end{aligned} \tag{2.2.4}$$

Integrating the second equation yields

$$g(x) - f(x) = \frac{1}{a}\int_{x_0}^{x} \psi(s)\, ds + C$$

where C is an integration constant. So

$$f(x) = \frac{1}{2}\phi(x) - \frac{1}{2a}\int_{x_0}^{x} \psi(s)\, ds - \frac{1}{2}C$$

$$g(x) = \frac{1}{2}\phi(x) + \frac{1}{2a}\int_{x_0}^{x} \psi(s)\, ds + \frac{1}{2}C.$$

Thus

$$u(x,t) = \frac{1}{2}\phi(x - a\,t) - \frac{1}{2a}\int_{x_0}^{x-a\,t} \psi(s)\, ds + \frac{1}{2}\phi(x + a\,t) + \frac{1}{2a}\int_{x_0}^{x+a\,t} \psi(s)\, ds.$$

This gives (2.2.2). It remains to check that, given a pair of functions $\phi(x) \in \mathcal{C}^2$, $\psi(x) \in \mathcal{C}^1$ the D'Alembert formula yields a solution to (2.1.1). Indeed, the function (2.2.2) is twice differentiable in x and t. It remains to substitute this function into the wave equation and check that the equation is satisfied. We leave it as an exercise for the reader. It is also straightforward to verify validity of the initial data (2.2.1). ■

Example. For the constant initial data

$$u(x,0) = u_0, \quad u_t(x,0) = v_0$$

the solution has the form

$$u(x,t) = u_0 + v_0 t.$$

This solution corresponds to the free motion of the string with the constant speed v_0.

Moreover the solution to the wave equation is stable with respect to small variations of the initial data. Namely,

Exercise 2.3 *For any $\epsilon > 0$ and any $T > 0$ there exists $\delta > 0$ such that the solutions $u(x,t)$ and $\tilde{u}(x,t)$ of the two Cauchy problems with initial conditions (2.2.1) and*

$$\tilde{u}(x,0) = \tilde{\phi}(x), \quad \tilde{u}_t(x,0) = \tilde{\psi}(x) \tag{2.2.5}$$

satisfy

$$\sup_{x\in\mathbb{R},\ t\in[0,T]} |\tilde{u}(x,t) - u(x,t)| < \epsilon \tag{2.2.6}$$

provided the initial conditions satisfy

$$\sup_{x\in\mathbb{R}} |\tilde{\phi}(x) - \phi(x)| < \delta, \quad \sup_{x\in\mathbb{R}} |\tilde{\psi}(x) - \psi(x)| < \delta. \tag{2.2.7}$$

Remark 2.4 *The property formulated in the above exercise is usually referred to as* well posedness *of the Cauchy problem (2.1.1), (2.2.1). We will return later to the discussion of this important property.*

2.3 Some consequences of the D'Alembert formula

Let (x_0, t_0) be a point of the (x, t)-plane, $t_0 > 0$. As it follows from the D'Alembert formula the value of the solution at the point (x_0, t_0) depends only on the values of $\phi(x)$ at $x = x_0 \pm a\,t_0$ and value of $\psi(x)$ on the interval $[x_0 - a\,t_0, x_0 + a\,t_0]$. The triangle with the vertices (x_0, t_0) and $(x_0 \pm a\,t_0, 0)$ is called *the dependence domain* of the segment $[x_0 - a\,t_0, x_0 + a\,t_0]$. The values of the solution *inside* this triangle are completely determined by the values of the initial data on the segment.

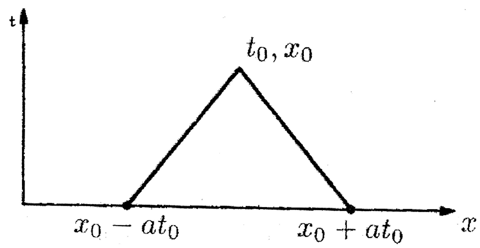

Fig. 2. The dependence domain of the segment $[x_0 - a\,t_0, x_0 + a\,t_0]$.

Another important definition is the *influence domain* for a given segment $[x_1, x_2]$ consider the domain defined by inequalities

$$x + a\,t \geqslant x_1, \quad x - a\,t \leqslant x_2, \quad t \geqslant 0. \tag{2.3.1}$$

Changing the initial data on the segment $[x_1, x_2]$ will not change the solution $u(x, t)$ outside the influence domain.

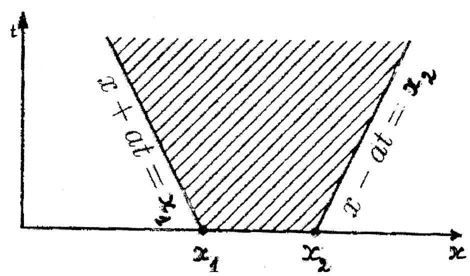

Fig. 3. The influence domain of the segment $[x_1, x_2]$.

Remark 2.5 *It will be convenient to slightly extend the class of initial data admitting piecewise smooth functions $\phi(x)$, $\psi(x)$ (all singularities of the latter must be integrable). If x_j are the singularities of these functions, $j = 1, 2, \ldots$, then the solution $u(x, t)$ given by the D'Alembert formula will satisfy the wave equation outside the lines*

$$x = \pm a\,t + x_j, \quad t \geqslant 0, \quad j = 1, 2, \ldots$$

The above formula says that the singularities of the solution propagate along the characteristics.

Example. Let us draw the profile of the string for the triangular initial data $\phi(x)$ shown on Fig. 4 and $\psi(x) \equiv 0$.

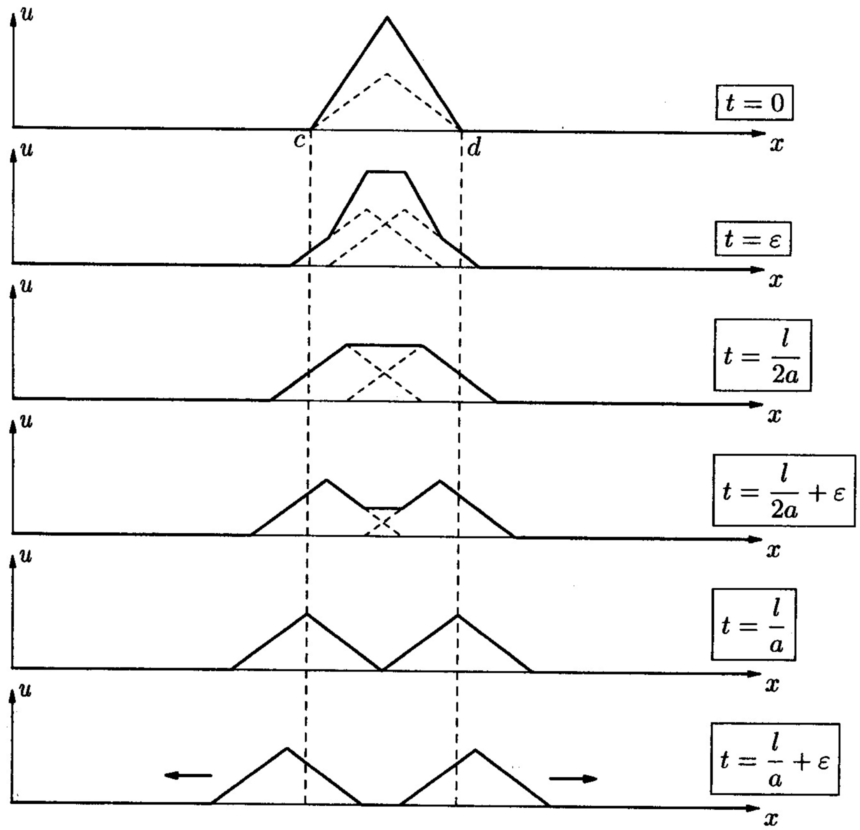

Fig. 4. The solution of the Cauchy problem for wave equation on the real line with a triangular initial profile at different instants of time.

2.4 Semi-infinite vibrating string

Let us begin with the following simple observation.

Lemma 2.6 *Let $u(x, t)$ be a solution to the wave equation. Then so are the functions*

$$\pm u(\pm x, \pm t)$$

with arbitrary choices of all three signs.

Proof: This follows from linearity of the wave equation and from its invariance with respect to the spatial reflection

$$x \mapsto -x$$

and time inversion

$$t \mapsto -t.$$

■

Let us consider oscillations of a string with a fixed point. Without loss of generality we can assume that the fixed point is at $x = 0$. We arrive at the following Cauchy problem for (2.1.1) on the half-line $x > 0$:

$$u(x,0) = \phi(x), \quad u_t(x,0) = \psi(x), \quad x > 0. \tag{2.4.1}$$

The solution must also satisfy the boundary condition

$$u(0,t) = 0, \quad t \geqslant 0. \tag{2.4.2}$$

The problem (2.1.1), (2.4.1), (2.4.2) is often called *mixed problem* since we have both initial conditions and boundary conditions.

The solution to the mixed problem on the half-line can be reduced to the problem on the infinite line by means of the following trick.

Lemma 2.7 *Let the initial data $\phi(x)$, $\psi(x)$ for the Cauchy problem (2.1.1), (2.2.1) be odd functions of x. Then the solution $u(x,t)$ is an odd function for all t.*

Proof: Denote

$$\tilde{u}(x,t) := -u(-x,t).$$

According to Lemma 2.6 the function $\tilde{u}(x,t)$ satisfies the same equation. At $t = 0$ we have

$$\tilde{u}(x,0) = -u(-x,0) = -\phi(-x) = \phi(x), \quad \tilde{u}_t(x,0) = -u_t(-x,0) = -\psi(-x) = \psi(x)$$

since ϕ and ψ are odd functions. Therefore $\tilde{u}(x,t)$ is a solution to the same Cauchy problem (2.1.1), (2.2.1). Due to uniqueness $\tilde{u}(x,t) = u(x,t)$, i.e. $-u(-x,t) = u(x,t)$ for all x and t. ■

We are now ready to present a recipe for solving the mixed problem for the wave equation on the half-line. Let us extend the initial data onto entire real line as odd functions. We arrive at the following Cauchy problem for the wave equation:

$$u(x,0) = \begin{cases} \phi(x), & x > 0 \\ -\phi(-x), & x < 0 \end{cases}, \quad u_t(x,0) = \begin{cases} \psi(x), & x > 0 \\ -\psi(-x), & x < 0 \end{cases} \tag{2.4.3}$$

According to Lemma 2.7 the solution $u(x,t)$ to the Cauchy problem (2.1.1), (2.4.3) given by the D'Alembert formula will be an odd function for all t. Therefore

$$u(0,t) = -u(0,t) = 0 \quad \text{for all} \quad t.$$

Example. Consider the evolution of a triangular initial profile on the half-line. The graph of the initial function $\phi(x)$ is non-zero on the interval $[l, 3l]$; the initial velocity $\psi(x) = 0$. The evolution is shown on Fig. 5 for few instants of time. Observe the reflected profile (the dotted line) on the negative half-line.

In a similar way one can treat the mixed problem on the half-line with a free boundary. In this case the vertical component $T\,u_x$ of the tension at the left edge must vanish at all times. Thus the boundary condition (2.4.2) has to be replaced with

$$u_x(0,t) = 0 \quad \text{for all} \quad t \geqslant 0. \tag{2.4.4}$$

One can solve the mixed problem (2.1.1), (2.4.1), (2.4.4) by using *even extension* of the initial data onto the negative half-line. We leave the details of the construction as an exercise for the reader.

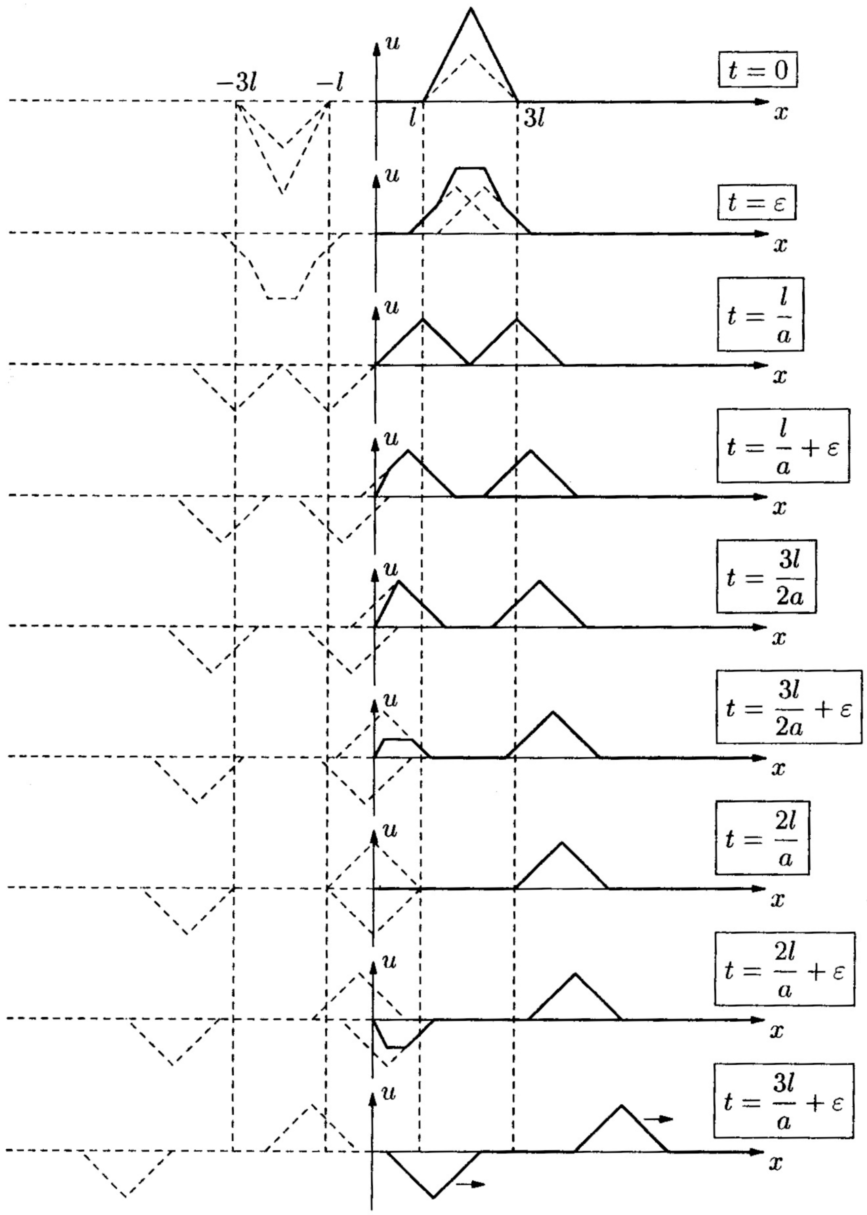

Fig. 5. The solution of the Cauchy problem for wave equation on the half-line with a triangular initial profile.

2.5 Periodic problem for wave equation. Introduction to Fourier series

Let us look for solutions to the wave equation (2.1.1) periodic in x with a given period $L > 0$. Thus we are looking for a solution $u(x,t)$ satisfying

$$u(x+L,t) = u(x,t) \quad \text{for any} \quad t \geqslant 0. \tag{2.5.1}$$

The initial data of the Cauchy problem

$$u(x,0) = \phi(x), \quad u_t(x,0) = \psi(x) \tag{2.5.2}$$

must also be L-periodic functions.

Theorem 2.8 *Given L-periodic initial data $\phi(x) \in \mathcal{C}^2(\mathbb{R})$, $\psi(x) \in \mathcal{C}^1(\mathbb{R})$ the periodic Cauchy problem (2.5.1), (2.5.2) for the wave equation (2.1.1) has a unique solution.*

Proof: According to the results of Section 2.2 the solution $u(x,t)$ to the Cauchy problem (2.1.1), (2.5.2) on $-\infty < x < \infty$ exists and is unique and is given by the D'Alembert formula. Denote

$$\tilde{u}(x,t) := u(x+L,t).$$

Since the coefficients of the wave equation do not depend on x the function $\tilde{u}(x,t)$ satisfies the same equation. The initial data for this function have the form

$$\tilde{u}(x,0) = \phi(x+L) = \phi(x), \quad \tilde{u}_t(x,t) = \psi(x+L) = \psi(x)$$

because of periodicity of the functions $\phi(x)$ and $\psi(x)$. So the initial data of the solutions $u(x,t)$ and $\tilde{u}(x,t)$ coincide. From the uniqueness of the solution we conclude that $\tilde{u}(x,t) = u(x,t)$ for all x and t, i.e. the function $u(x,t)$ is periodic in x with the same period L. ■

Exercise 2.9 *Prove that the complex exponential function e^{ikx} is L-periodic* iff *the wave number k has the form*

$$k = \frac{2\pi n}{L}, \quad n \in \mathbb{Z}. \tag{2.5.3}$$

In the following two exercises we will consider the particular case $L = 2\pi$. In this case the complex exponential

$$e^{\frac{2\pi i n x}{L}}$$

obtained in the previous exercise reduces to e^{inx}.

Exercise 2.10 *Prove that the solution of the periodic Cauchy problem with the Cauchy data*

$$u(x,0) = e^{inx}, \quad u_t(x,0) = 0 \tag{2.5.4}$$

is given by the formula

$$u(x,t) = e^{inx} \cos nat. \tag{2.5.5}$$

Exercise 2.11 *Prove that the solution of the periodic Cauchy problem with the Cauchy data*

$$u(x,0) = 0, \quad u_t(x,0) = e^{inx} \tag{2.5.6}$$

is given by the formula

$$u(x,t) = \begin{cases} e^{inx}\frac{\sin nat}{na}, & n \neq 0 \\ t, & n = 0. \end{cases} \tag{2.5.7}$$

Using the theory of *Fourier series* we can represent any solution to the periodic problem to the wave equation as a superposition of the solutions (2.5.5), (2.5.7). Let us first recall some basics of the theory of Fourier series.

Let $f(x)$ be a 2π-periodic continuously differentiable complex valued function on $\mathbb{R}$. The *Fourier series* of this function is defined by the formula

$$\sum_{n\in\mathbb{Z}} c_n e^{inx} \tag{2.5.8}$$

$$c_n = \frac{1}{2\pi}\int_0^{2\pi} f(x) e^{-inx} dx. \tag{2.5.9}$$

The following theorem is a fundamental result of the theory of Fourier series.

Theorem 2.12 *For any function $f(x)$ satisfying the above conditions the Fourier series is uniformly convergent to the function $f(x)$.*

In particular we conclude that any $\mathcal{C}^1$-smooth 2π-periodic function $f(x)$ can be represented as a sum of uniformly convergent Fourier series

$$f(x) = \sum_{n\in\mathbb{Z}} c_n e^{inx}, \quad c_n = \frac{1}{2\pi}\int_0^{2\pi} f(x) e^{-inx} dx. \tag{2.5.10}$$

For completeness we remind the proof of this Theorem.

Let us introduce *Hermitean inner product* in the space of complex valued 2π-periodic continuous functions:

$$(f,g) = \frac{1}{2\pi}\int_0^{2\pi} \bar f(x) g(x)\, dx. \tag{2.5.11}$$

Here the bar stands for complex conjugation. This inner product satisfies the following properties:

$$(g,f) = \overline{(f,g)} \tag{2.5.12}$$

$$\left.\begin{aligned} (\lambda f_1 + \mu f_2, g) &= \bar\lambda (f_1, g) + \bar\mu (f_2, g) \\ (f, \lambda g_1 + \mu g_2) &= \lambda (f, g_1) + \mu (f, g_2) \end{aligned}\right. \quad \text{for any} \quad \lambda,\ \mu \in \mathbb{C} \tag{2.5.13}$$

$$(f,f) > 0 \quad \text{for any nonzero continuous function} \quad f(x). \tag{2.5.14}$$

The real nonnegative number (f,f) will be used for defining the L_2-norm of the function:

$$\|f\| := \sqrt{(f,f)}. \tag{2.5.15}$$

Exercise 2.13 *Prove that the L_2-norm satisfies the triangle inequality:*

$$\|f+g\| \leqslant \|f\| + \|g\|. \tag{2.5.16}$$

Observe that the complex exponentials e^{inx} form an orthonormal system with respect to the inner product (2.5.11):

$$\left(e^{imx}, e^{inx}\right) = \delta_{mn} = \begin{cases} 1, & m = n \\ 0 & m \neq n \end{cases}. \tag{2.5.17}$$

(check it!).

Let $f(x)$ be a continuous function; denote c_n its Fourier coefficients. The following formula

$$c_n = (e^{inx}, f), \quad n \in \mathbb{Z} \tag{2.5.18}$$

gives a simple interpretation of the Fourier coefficients as the coefficients of decomposition of the function f with respect to the orthonormal system made from exponentials. Moreover, the partial sum of the Fourier series

$$S_N(x) = \sum_{n=-N}^{N} c_n e^{inx} \tag{2.5.19}$$

can be interpreted as the orthogonal projection of the vector f onto the $(2N+1)$-dimensional linear subspace

$$V_N = \text{span}\left(1, e^{\pm ix}, e^{\pm 2ix}, \dots, e^{\pm iNx}\right) \tag{2.5.20}$$

consisting of all *trigonometric polynomials*

$$P_N(x) = \sum_{n=-N}^{N} p_n e^{inx} \tag{2.5.21}$$

of degree N. Here p_0, $p_{\pm 1}$, $\dots p_{\pm N}$ are arbitrary complex numbers.

Lemma 2.14 *The following inequality holds true:*

$$\sum_{n=-N}^{N} |c_n|^2 \leqslant \|f\|^2. \tag{2.5.22}$$

The statement of this lemma is called *Bessel inequality*.

Proof: We have

$$\begin{aligned} 0 \leqslant \|f(x) - \sum_{n=-N}^{N} c_n e^{inx}\|^2 &= \left(f(x) - \sum_{n=-N}^{N} c_n e^{inx}, f(x) - \sum_{n=-N}^{N} c_n e^{inx}\right) \\ &= (f,f) - \sum_{n=-N}^{N} \left[c_n\left(f, e^{inx}\right) + \bar{c}_n\left(e^{inx}, f\right)\right] + \sum_{m,n=-N}^{N} \bar{c}_m c_n \left(e^{imx}, e^{inx}\right). \end{aligned}$$

Using (2.5.18) and orthonormality (2.5.17) we recast the right hand side of the last equation in the form

$$(f,f) - \sum_{n=-N}^{N} |c_n|^2.$$

This proves Bessel inequality. ■

Geometrically the Bessel inequality says that the square length of the orthogonal projection of a vector onto the linear subspace V_N cannot be longer than the square length of the vector itself.

Corollary 2.15 *For any continuous function $f(x)$ the series of squares of absolute values of Fourier coefficients converges:*

$$\sum_{n\in\mathbb{Z}} |c_n|^2 < \infty. \tag{2.5.23}$$

The following *extremal property* says that the N-th partial sum of the Fourier series gives the best L_2-approximation of the function $f(x)$ among all trigonometric polynomials of degree N.

Lemma 2.16 *For any trigonometric polynomial $P_N(x)$ of degree N the following inequality holds true*

$$\|f(x) - S_N(x)\| \leqslant \|f(x) - P_N(x)\|. \tag{2.5.24}$$

Here $S_N(x)$ is the N-th partial sum (2.5.19) of the Fourier series of the function f. The equality in (2.5.24) takes place iff *the trigonometric polynomial $P_N(x)$ coincides with $S_N(x)$, i.e.,*

$$p_n = \frac{1}{2\pi}\int_0^{2\pi} f(x)e^{-inx}dx, \quad n = 0, \pm 1, \pm 2, \ldots, \pm N,$$

Proof: From (2.5.18) we derive that

$$(f(x) - S_N(x), P_N(x)) = 0 \quad \text{for any} \quad P_N(x) \in V_N.$$

Hence

$$\begin{aligned}
&\|f(x) - P_N(x)\|^2 = \|(f - S_N) + (S_N - P_N\|^2 = \\
&= (f - S_N, f - S_N) + (f - S_N, Q_N) + (Q_N, f - S_N) + (Q_N, Q_N) \\
&= (f - S_N, f - S_N) + (Q_N, Q_N) \geqslant (f - S_N, f - S_N) = \|f - S_N\|^2.
\end{aligned}$$

Here we denote

$$Q_N = S_N(x) - P_N(x) \in V_N.$$

Clearly the equality takes place *iff* $Q_N = 0$, i.e. $P_N = S_N$. ■

Lemma 2.17 *For any continuous 2π-periodic function the following* Parseval equality *holds true:*

$$\sum_{n\in\mathbb{Z}} |c_n|^2 = \|f\|^2. \tag{2.5.25}$$

The Parseval equality can be considered as an infinite-dimensional analogue of the Pythagoras theorem: sum of the squares of orthogonal projections of a vector on the coordinate axes is equal to the square length of the vector.

Proof: According to Stone – Weierstrass theorem[3] any continuous 2π-periodic function can be uniformly approximated by Fourier polynomials

$$P_N(x) = \sum_{n=-N}^{N} p_n e^{inx}. \tag{2.5.26}$$

That means that for a given function $f(x)$ and any $\epsilon > 0$ there exists a trigonometric polynomial $P_N(x)$ of some degree N such that

$$\sup_{x\in[0,2\pi]} |f(x) - P_N(x)| < \epsilon.$$

Then

$$\|f - P_N\|^2 = \frac{1}{2\pi}\int_0^{2\pi} |f(x) - P_N(x)|^2 dx < \epsilon^2.$$

Therefore, due to the extremal property (see Lemma 2.16 above), we obtain the following inequality

$$\|f - S_N\|^2 < \epsilon^2.$$

Repeating the computation used in the proof of Bessel inequality

$$\|f - S_N\|^2 = \|f\|^2 - \sum_{n=-N}^{N} |c_n|^2 < \epsilon^2$$

we arrive at the proof of Lemma. ■

[3]The Stone – Weierstrass theorem is a very general result about uniform approximation of continuous functions on a compact K in a metric space. Let us recall this important theorem. Let $A \subset \mathcal{C}(K)$ be a subset of functions in the space of continuous real- or complex-valued functions on a compact K. The following requirements must hold true.

1. A must be a *subalgebra* in $\mathcal{C}(K)$, i.e. for $f,\ g \in A$, $\alpha,\ \beta \in \mathbb{R}$ (or $\alpha,\ \beta \in \mathbb{C}$) the linear combination and the product belong to A:

$$\alpha\, f + \beta\, g \in A, \quad f \cdot g \in A.$$

2. The functions in A must *separate points* in K, i.e., $\forall x,\ y \in K$, $x \neq y$ there exists $f \in A$ such that

$$f(x) \neq f(y).$$

3. The subalgebra is *non-degenerate*, i.e., $\forall x \in K$ there exists $f \in A$ such that $f(x) \neq 0$.
   The last condition has to be imposed in the complex situation.
4. The subalgebra A is said to be *self-adjoint* if for any function $f \in A$ the complex conjugate function $\bar{f}$ also belongs to A.

Theorem 2.18 *Given an algebra of functions $A \subset \mathcal{C}(K)$ that separates points, is non-degenerate and, for complex-valued functions, is self-adjoint then A is an everywhere dense subset in $\mathcal{C}(K)$.*

Recall that density means that for any continuous function $F \in \mathcal{C}(K)$ and an arbitrary $\epsilon > 0$ there exists $f \in A$ such that

$$\sup_{x\in K}|F(x) - f(x)| < \epsilon.$$

In the particular case of algebra of polynomials one obtains the classical Weierstrass theorem about polynomial approximations of continuous functions on a finite interval. For the needs of the theory of Fourier series one has to apply the Stone – Weierstrass theorem to the subalgebra of Fourier polynomials in the space of continuous 2π-periodic functions. We leave as an exercise to verify applicability of the Stone – Weierstrass theorem in this case.

The Parseval equality is also referred to as *completeness* of the trigonometric system of functions

$$1, e^{\pm ix}, e^{\pm 2ix}, \dots .$$

For the case of infinite-dimensional spaces equipped with a Hermitean (or Euclidean) inner product the property of completeness is the right analogue of the notion of an orthonormal *basis* of the space.

Corollary 2.19 *Two continuous 2π-periodic functions $f(x)$, $g(x)$ with all equal Fourier coefficients identically coincide.*

Proof: Indeed, the difference $h(x) = f(x) - g(x)$ is continuous function with zero Fourier coefficients. The Parseval equality implies $\|h\|^2 = 0$. So $h(x) \equiv 0$. ∎

We can now prove that uniform convergence of the Fourier series of a $\mathcal{C}^1$-function. Denote c_n' the Fourier coefficients of the derivative $f'(x)$. Integrating by parts we derive the following formula:

$$c_n = \frac{1}{2\pi}\int_0^{2\pi} f(x)e^{-inx}\,dx = -\frac{1}{2\pi in}f(x)e^{-inx}\,|_0^{2\pi} + \frac{1}{2\pi in}\int_0^{2\pi} f'(x)e^{-inx}\,dx = -\frac{i}{n}c_n'.$$

This implies convergence of the series

$$\sum_{n\in\mathbb{Z}} |c_n|.$$

Indeed,

$$|c_n| = \frac{|c_n'|}{n} \leqslant \frac{1}{2}\left(|c_n'|^2 + \frac{1}{n^2}\right).$$

The series $\sum |c_n'|^2$ converges according to the Corollary 2.15; convergence of the series $\sum \frac{1}{n^2}$ is well known. Using Weierstrass theorem we conclude that the Fourier series converges absolutely and uniformly

$$\sum_{n\in\mathbb{Z}} \left|c_n e^{inx}\right| = \sum_{n\in\mathbb{Z}} |c_n| < \infty.$$

Denote $g(x)$ the sum of this series. It is a continuos function. The Fourier coefficients of g coincide with those of f:

$$\left(e^{inx}, g\right) = c_n.$$

Hence $f(x) \equiv g(x)$. ∎

For the specific case of real valued function the Fourier coefficients satisfy the following property.

Lemma 2.20 *The function $f(x)$ is real valued* iff *its Fourier coefficients satisfy*

$$\bar{c}_n = c_{-n} \quad \text{for all} \quad n \in \mathbb{Z}. \tag{2.5.27}$$

Proof: Reality of the function can be written in the form

$$\bar{f}(x) = f(x).$$

Since

$$\overline{e^{inx}} = e^{-inx}$$

we have

$$\bar{c}_n = \frac{1}{2\pi}\int_0^{2\pi} \bar{f}(x)e^{inx}dx = c_{-n}.$$

■

Note that the coefficient

$$c_0 = \frac{1}{2\pi}\int_0^{2\pi} f(x)\, dx$$

is always real if $f(x)$ is a real valued function.

Let us establish the correspondence of the complex form (2.5.10) of the Fourier series of a real valued function with the real form.

Lemma 2.21 *Let $f(x)$ be a real valued 2π-periodic smooth function. Denote c_n its Fourier coefficients (2.5.9). Introduce coefficients*

$$a_n = c_n + c_{-n} \quad = \frac{1}{\pi}\int_0^{2\pi} f(x)\cos nx\, dx, \quad n = 0, 1, 2, \ldots \tag{2.5.28}$$

$$b_n = i(c_n - c_{-n}) = \frac{1}{\pi}\int_0^{2\pi} f(x)\sin nx\, dx, \quad n = 1, 2, \ldots \tag{2.5.29}$$

Then the function $f(x)$ is represented as a sum of uniformly convergent Fourier series of the form

$$f(x) = \frac{a_0}{2} + \sum_{n\geqslant 1}\left(a_n \cos nx + b_n \sin nx\right). \tag{2.5.30}$$

We leave the proof of this Lemma as an exercise for the reader.

Exercise 2.22 *For any real valued continuous function $f(x)$ prove the following version*[4] *of Bessel inequality (2.5.22):*

$$\frac{a_0^2}{2} + \sum_{n=1}^{N}(a_n^2 + b_n^2) \leqslant \frac{1}{\pi}\int_0^{2\pi} f^2(x)\, dx \tag{2.5.31}$$

and Parseval equality (2.5.25)

$$\frac{a_0^2}{2} + \sum_{n=1}^{\infty}(a_n^2 + b_n^2) = \frac{1}{\pi}\int_0^{2\pi} f^2(x)\, dx. \tag{2.5.32}$$

The following statement can be used in working with functions with an arbitrary period.

Exercise 2.23 *Given an arbitrary constant $c \in \mathbb{R}$ and a solution $u(x,t)$ to the wave equation (2.1.1) then*

$$\tilde{u}(x,t) = u\left(c\, x, c\, t\right) \tag{2.5.33}$$

also satisfies (2.1.1).

[4]Notice a change in the normalization of the L_2 norm.

Note that for $c \neq 0$ the function $\tilde{u}(x,t)$ is periodic in x with the period $L = \frac{2\pi}{c}$ if $u(x,t)$ was 2π-periodic.

For non-smooth functions the problem of convergence of Fourier series is more delicate. Let us consider an example giving some idea about the convergence of Fourier series for piecewise smooth functions. Consider the function

$$\operatorname{sign} x = \begin{cases} 1, & x > 0 \\ 0, & x = 0 \\ -1, & x < 0 \end{cases}. \tag{2.5.34}$$

This function will be considered on the interval $[-\pi, \pi]$ and then continued 2π-periodically onto entire real line. The Fourier coefficients of this function can be easily computed:

$$a_n = 0, \quad b_n = \frac{2}{\pi} \frac{(1 - (-1)^n)}{n}.$$

So the Fourier series of this functions reads

$$\frac{4}{\pi} \sum_{k \geqslant 1} \frac{\sin(2k-1)x}{2k-1}. \tag{2.5.35}$$

One can prove that this series converges to the sign function at every point of the interval $(-\pi, \pi)$. Moreover this convergence is uniform on every closed subinterval non containing 0 or $\pm\pi$. However the character of convergence near the discontinuity points $x = 0$ and $x = \pm\pi$ is more complicated as one can see from the following graph of a partial sum of the series (2.5.35).

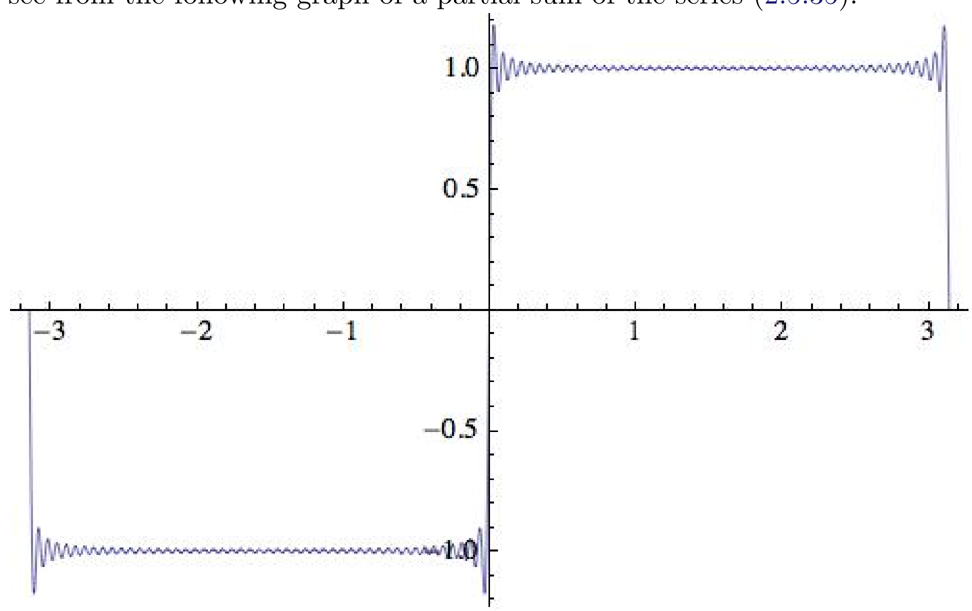

Fig. 6. Graph of the partial sum $S_n(x) = \frac{4}{\pi} \sum_{k=1}^{n} \frac{\sin(2k-1)x}{2k-1}$ for $n = 50$.

In general for piecewise smooth functions $f(x)$ with some number of discontinuity points one can prove that the Fourier series converges to the mean value $\frac{1}{2}(f(x_0+0) + f(x_0-0))$ at every first kind discontinuity point x_0. The non vanishing oscillatory behavior of partial sums near discontinuity points is known as *Gibbs phenomenon* (see Exercise 2.51 below).

Let us return to the wave equation. Using the theory of Fourier series we can represent any periodic solution to the Cauchy problem (2.5.2) as a superposition of solutions of the form (2.5.5), (2.5.7). Namely, let us expand the initial data in Fourier series:

$$\phi(x) = \sum_{n \in \mathbb{Z}} \phi_n e^{inx}, \quad \psi(x) = \sum_{n \in \mathbb{Z}} \psi_n e^{inx}. \tag{2.5.36}$$

Then the solution to the periodic Cauchy problem reads

$$u(x,t) = \sum_{n \in \mathbb{Z}} \phi_n e^{inx} \cos ant + \psi_0 t + \frac{1}{a} \sum_{n \in \mathbb{Z} \setminus 0} \psi_n e^{inx} \frac{\sin ant}{n}. \tag{2.5.37}$$

Remark 2.24 *The formula (2.5.37) says that the solutions*

$$\begin{aligned} u_n^{(1)}(x,t) &= e^{inx} \cos ant \\ u_n^{(2)}(x,t) &= \begin{cases} t, & n = 0 \\ e^{inx} \frac{\sin ant}{n}, & n \neq 0 \end{cases} \end{aligned} \tag{2.5.38}$$

for $n \in \mathbb{Z}$ form a basis in the space of 2π-periodic solutions to the wave equation. Observe that all these solutions can be written in the so-called separated form

$$u(x,t) = X(x)T(t) \tag{2.5.39}$$

for some smooth functions $X(x)$ and $T(t)$. A rather general method of separation of variables *for solving boundary value problems for linear PDEs has this observation as a starting point. This method will be explained later on.*

2.6 Finite vibrating string. Standing waves

Let us proceed to considering a finite string of the length l. We begin with considering the oscillations of the string with fixed endpoints. So we have to solve the following mixed problem for the wave equation (2.1.1)

$$u(x,0) = \phi(x), \quad u_t(x,0) = \psi(x), \quad x \in [0,l] \tag{2.6.1}$$

$$u(0,t) = 0, \quad u(l,t) = 0 \quad \text{for all} \quad t > 0. \tag{2.6.2}$$

The idea of solution is, again, in a suitable extension of the problem onto entire line.

Lemma 2.25 *Let the initial data $\phi(x)$, $\psi(x)$ of the Cauchy problem (2.2.1) for the wave equation on $\mathbb{R}$ be odd $2l$-periodic functions. Then the solution $u(x,t)$ will also be an odd $2l$-periodic function for all t satisfying the boundary conditions (2.6.2).*

Proof: As we already know from Lemma 2.7 the solution is an odd function for all t. So

$$u(0,t) = 0 \quad \text{for all} \quad t > 0.$$

Next, the solution will be $2l$-periodic for all t according to Theorem 2.8 above. So

$$u(l-x,t) = -u(x-l,t) = -u(x+l,t).$$

Substituting $x=0$ we get

$$u(l,t) = -u(l,t), \quad i.e. \quad u(l,t) = 0.$$

■

The above Lemma gives an algorithm for solving the mixed problem (2.6.1), (2.6.2) for the wave equation. Namely, we extend the initial data $\phi(x)$, $\psi(x)$ from the interval $[0,x]$ onto the real axis as odd $2l$-periodic functions. After this we apply D'Alembert formula to the extended initial data. The resulting solution will satisfy the initial conditions (2.6.1) on the interval $[0,l]$ as well as the boundary conditions (2.6.2) at the end points of the interval.

We will apply now the technique of Fourier series to the mixed problem (2.6.1), (2.6.2).

Lemma 2.26 *Let a 2π-periodic functions $f(x)$ be represented as the sum of its Fourier series*

$$f(x) = \sum_{n\in\mathbb{Z}} c_n e^{inx}, \quad c_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-inx}dx.$$

The function $f(x)$ is even/odd iff *the Fourier coefficients satisfy*

$$c_{-n} = \pm c_n$$

respectively.

Proof: For an even function one must have

$$\sum_{n\in\mathbb{Z}} c_n e^{inx} = f(x) = f(-x) = \sum_{n\in\mathbb{Z}} c_n e^{-inx} = \sum_{n\in\mathbb{Z}} c_{-n} e^{inx}.$$

This proves $c_{-n} = c_n$. A similar argument gives $c_{-n} = -c_n$ for the case of an odd function. ■

Corollary 2.27 *Any even/odd smooth 2π-periodic function can be expanded in Fourier series in cosines/sines:*

$$f(x) = \frac{a_0}{2} + \sum_{n\geqslant 1} a_n \cos nx, \quad a_n = \frac{2}{\pi}\int_0^{\pi} f(x)\cos nx\, dx, \quad f(x) \quad \text{is even} \tag{2.6.3}$$

$$f(x) = \sum_{n\geqslant 1} b_n \sin nx, \quad b_n = \frac{2}{\pi}\int_0^{\pi} f(x)\sin nx\, dx, \quad f(x) \quad \text{is odd.} \tag{2.6.4}$$

Proof: Let us consider the case of an odd function. In this case we have $c_{-n} = -c_n$, and, in particular, $c_0 = 0$, so we rewrite the Fourier series in the following form

$$\begin{aligned} f(x) &= \sum_{n\geqslant 1} c_n e^{inx} + \sum_{n\leqslant -1} c_n e^{inx} \\ &= \sum_{n\geqslant 1} c_n \left(e^{inx} - e^{-inx}\right) = 2i \sum_{n\geqslant 1} c_n \sin nx. \end{aligned}$$

Denote

$$b_n = 2ic_n, \quad n \geqslant 1.$$

For this coefficient we obtain

$$b_n = \frac{2i}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-inx}dx = \frac{i}{\pi}\int_0^{\pi} f(x)e^{-inx}dx + \frac{i}{\pi}\int_{-\pi}^{0} f(x)e^{-inx}dx.$$

In the second integral we change the integration variable $x \mapsto -x$ and use that $f(-x) = -f(x)$ to arrive at

$$b_n = \frac{i}{\pi}\int_0^{\pi} f(x)e^{-inx}dx + \frac{i}{\pi}\int_{\pi}^{0} f(x)e^{inx}dx = \frac{i}{\pi}\int_0^{\pi} f(x)\left[e^{-inx} - e^{inx}\right] dx = \frac{2}{\pi}\int_0^{\pi} f(x)\sin nx\, dx.$$

■

Let us return to the solution to the wave equation on the interval $[0, l]$ with fixed endpoints boundary condition. Summarizing the previous considerations we arrive at the following

Theorem 2.28 *Let $\phi(x) \in \mathcal{C}^2([0, l])$, $\psi(x) \in \mathcal{C}^1([0, l])$ be two arbitrary smooth functions. Then the solutions to the mixed problem (2.6.1), (2.6.2) for the wave equation is written in the form*

$$u(x, t) = \sum_{n \geqslant 1} \sin\frac{\pi n x}{l}\left(b_n \cos\frac{\pi a n t}{l} + \dot{b}_n \sin\frac{\pi a n t}{l}\right) \tag{2.6.5}$$

$$b_n = \frac{2}{l}\int_0^l \phi(x)\sin\frac{\pi n x}{l}\, dx, \quad \dot{b}_n = \frac{2}{\pi a n}\int_0^l \psi(x)\sin\frac{\pi n x}{l}\, dx.$$

Particular solutions to the wave equation giving a basis in the space of all solutions satisfying the boundary conditions (2.6.1) have the form

$$u_n^{(1)}(x, t) = \sin\frac{\pi n x}{l}\cos\frac{\pi a n t}{l}, \quad u_n^{(2)}(x, t) = \sin\frac{\pi n x}{l}\sin\frac{\pi a n t}{l}, \quad n = 1, 2, \dots \tag{2.6.6}$$

are called *standing waves*. Observe that these solutions have the separated form (2.5.39). The shape of these waves essentially does not change in time, only the size does change. In particular the location of the *nodes*

$$x_k = k\,\frac{l}{n}, \quad k = 0, 1, \dots, n \tag{2.6.7}$$

of the n-th solution $u_n^{(1)}(x, t)$ or $u_n^{(2)}(x, t)$ does not depend on time. The n-th standing waves (2.6.6) has $(n + 1)$ nodes on the string. The solution takes zero values at the nodes at all times.

2.7 Energy of vibrating string

Let us consider the vibrating string with fixed points $x = 0$ and $x = l$. It is clear that the kinetic energy of the string at the moment t is equal to

$$K = \frac{1}{2}\int_0^l \rho\, u_t^2(x, t)\, dx. \tag{2.7.1}$$

Let us now compute the potential energy U of the string. By definition U is equal to the work done by the elastic force moving the string from the equilibrium $u \equiv 0$ to the actual position given by the graph $u(x)$. The motion can be described by the one-parameter family of curves

$$v(x; s) = s\, u(x) \tag{2.7.2}$$

where the parameter s changes from $s = 0$ (the equilibrium) to $s = 1$ (the position of the string). As we already know the vertical component of the force acting on the interval of the string (2.7.2) between x and $x + \Delta x$ is equal to

$$F = T\, \left(v_x(x + \Delta x; s) - v_x(x; s)\right) \simeq s\, T\, u_{xx}(x)\, \Delta x.$$

The work A to move the string from the position $v(x; s)$ to $v(x; s + \Delta s)$ is therefore equal to

$$A = -F \cdot [v(x; s + \Delta s) - v(x; s)] \simeq -s\, T\, u(x) \Delta x\, \Delta s$$

(the negative sign since the direction of the force is opposite to the direction of the displacement). The total work of the elastic forces for moving the string of length l from the equilibrium $s = 0$ to the given configuration at $s = 1$ is obtained by integration:

$$U = -\int_0^1 ds \int_0^l s\, T\, u_{xx}(x) u(x)\, dx = -\frac{1}{2} \int_0^l T\, u_{xx}(x) u(x)\, dx.$$

By definition this work is equal to the potential energy of the string. Integrating by parts and using the boundary conditions

$$u(0) = u(l) = 0$$

we finally arrive at the following expression for the potential energy:

$$U = \frac{1}{2} \int_0^l T\, u_x^2(x)\, dx. \tag{2.7.3}$$

Summarizing (2.7.1) and (2.7.3) gives the formula for the total energy $E = E(t)$ of the vibrating string at the moment t

$$E = K + U = \int_0^l \left(\frac{1}{2}\, \rho\, u_t^2(x, t) + \frac{1}{2}\, T\, u_x^2(x, t) \right) dx. \tag{2.7.4}$$

Exercise 2.29 *Prove that the same expression (2.7.3) holds true for the total work of elastic forces moving the string from the equilibrium to the given position $u(x)$ along an* arbitrary path

$$v(x; s), \quad v(x; 0) \equiv 0, \quad v(x; s) = u(x)$$

in the space of configurations.

It is understood that $v(x; t)$ is a smooth function on $[0, l] \times [0, 1]$.

We will now prove that the total energy E of vibrating string with fixed end points does not depend on time.

Lemma 2.30 *Let the function $u(x, t)$ satisfy the wave equation. Then the following identity holds true*

$$\frac{\partial}{\partial t} \left(\frac{1}{2}\, \rho\, u_t^2(x, t) + \frac{1}{2}\, T\, u_x^2(x, t) \right) = \frac{\partial}{\partial x} \left(T\, u_x u_t \right). \tag{2.7.5}$$

Proof: A straightforward differentiation using $u_{tt} = a^2 u_{xx}$ yields

$$\frac{\partial}{\partial t}\left(\frac{1}{2}\rho\, u_t^2(x,t) + \frac{1}{2} T\, u_x^2(x,t)\right) = \rho\, a^2 u_t u_{xx} + T\, u_x u_{xt}.$$

Since

$$a^2 = \frac{T}{\rho}$$

(see above) we rewrite the last equation in the form

$$= T\left(u_t u_{xx} + u_{tx} u_x\right) = T\,\left(u_t u_x\right)_x.$$

■

Corollary 2.31 *Denote $E_{[a,b]}(t)$ the energy of a segment of vibrating string*

$$E_{[a,b]}(t) = \int_a^b \left(\frac{1}{2}\rho\, u_t^2(x,t) + \frac{1}{2} T\, u_x^2(x,t)\right) dx. \tag{2.7.6}$$

The following formula describes the dependence of this energy on time:

$$\frac{d}{dt} E_{[a,b]}(t) = T\, u_t u_x|_{x=b} - T\, u_t u_x|_{x=a}. \tag{2.7.7}$$

Remark 2.32 *In physics literature the quantity*

$$\frac{1}{2}\rho\, u_t^2(x,t) + \frac{1}{2} T\, u_x^2(x,t) \tag{2.7.8}$$

is called energy density*. It is equal to the energy of a small piece of the string from x to $x + dx$ at the moment t. The total energy of a piece of a string is obtained by integration of this density in x. Another important notion is the* flux density

$$-T\, u_t u_x. \tag{2.7.9}$$

The formula (2.7.7) says that the change of the energy of a given piece of the string for the time dt is given by the total flux though the boundary of the piece.

Finally we arrive at the *conservation law* of the total energy of a vibrating string with fixed end points.

Theorem 2.33 *The total energy (2.7.4) of the vibrating string with fixed end points does not depend on t:*

$$\frac{d}{dt} E = 0.$$

Proof: The formula (2.7.7) for the particular case $a = 0$, $b = l$ gives

$$\frac{d}{dt}E = T\,\left(u_t(l,t)u_x(l,t) - u_t(0,t)u_x(0,t)\right) = 0$$

since

$$u_t(0,t) = \partial_t u(0,t) = 0, \quad u_t(l,t) = \partial_t u(l,t) = 0$$

due to the boundary conditions $u(0,t) = u(l,t) = 0$. ■

The conservation law of total energy makes it evident that the vibrating string is a *conservative system.*

Exercise 2.34 *Derive the formula for the total energy and prove the conservation law for a vibrating string of finite length with free boundary conditions* $u_x(0,t) = u_x(l,t) = 0$.

Exercise 2.35 *Prove that the energy of the vibrating string represented as sum (2.6.5) of standing waves (2.6.6) is equal to the sum of energies of standing waves.*

The conservation of total energy can be used for proving uniqeness of solution for the wave equation. Indeed, if $u^{(1)}(x,t)$ and $u^{(2)}(x,t)$ are two solutions vanishing at $x = 0$ and $x = l$ with the same initial data. The difference

$$u(x,t) = u^{(2)}(x,t) - u^{(1)}(x,t)$$

solves wave equation, satisfies the same boundary conditions and has zero initial data $u(x,0) = \phi(x) = 0$, $u_t(x,0) = \psi(x) = 0$. The conservation of energy for this solution gives

$$E(t) = \int_0^l \left(\frac{1}{2}\,\rho\,u_t^2(x,t) + \frac{1}{2}\,T\,u_x^2(x,t)\right)\,dx = E(0) = \int_0^l \left(\frac{1}{2}\,\rho\,\psi^2(x) + \frac{1}{2}\,T\,\phi_x^2(x)\right)\,dx = 0.$$

Hence $u_x(x,t) = u_t(x,t) = 0$ for all x, t. Using the boundary conditions one concludes that $u(x,t) \equiv 0$,

2.8 Inhomogeneous wave equation: Duhamel principle

To give a heuristic motivation of the method we start by reminding that for solving linear first order ODEs

$$\dot{u}(t) + Lu(t) = g(t), \tag{2.8.1}$$

with L a constant (in t) we can use *variation of parameters* which gives the particular solution $u_p(t)$

$$u_p(t) = \mathrm{e}^{-Lt}\int_0^t \mathrm{e}^{Ls}g(s)\mathrm{d}s = \int_0^t \mathrm{e}^{-L(t-s)}g(s)\mathrm{d}s; \quad u_p(0) = 0. \tag{2.8.2}$$

Denoting the integrand of this latter equation by $f(t;s) = \mathrm{e}^{L(t-s)}g(s)$ we note that it is also a solution of the **homogeneous** ODE

$$\partial_t f(t;s) + Lf(t;s) = 0, \qquad f(t;s)|_{t=s} = g(s). \tag{2.8.3}$$

This shows that the particular solution (2.8.2) of the non-homogeneous equation (2.8.1) can be written as a superposition (integral) of *homogeneous* solutions with $g(s)$ is the initial value at $t = s$.

Similarly for second order ODEs:

$$\ddot{u}(t) + Lu(t) = g(t) \ . \tag{2.8.4}$$

A particular solution given by the variation of parameters formula appears in the form

$$u_p(t) = \int_0^t \frac{\sin(\sqrt{L}(t-s))}{\sqrt{L}} g(s)\mathrm{d}s \ , \quad u_p(0) = 0, \quad \dot{u}_p(0) = 0. \tag{2.8.5}$$

Once more we observe that the integral of the above formula $f(t;s) = \frac{\sin(\sqrt{L}(t-s))}{\sqrt{L}} g(s)$ is the solution of the Cauchy problem

$$\partial_t^2 f(t;s) + Lf(t;s) = 0 \ ; \qquad f(t;s)|_{t=s} = 0, \qquad \partial_t f(t;s)|_{t=s} = g(s). \tag{2.8.6}$$

With appropriate interpretation, the same formulæ would hold if $u(t)$ is a function taking values in an arbitrary vector space (even infinite dimensional, formally) as long as L is a linear operator *independent of* t. Since ∂_x^2 could be construed as such, this motivates the following theorem

Theorem 2.36 (Duhamel formula (principle)) *Consider the inhomogeneous equation of the string with external forcing* $g(x,t) \in \mathcal{C}^0(\mathbb{R}^2)$*:*

$$u_{tt}(x,t) - a^2 u_{xx}(x,t) = g(x,t), \quad u(x,0) = 0 = u_t(x,0). \tag{2.8.7}$$

Then the solution is given by the formula

$$u(x,t) = \int_0^t F(x,t;s)\mathrm{d}s \tag{2.8.8}$$

where $F(x,t;s)$ *is the solution of the homogeneous wave equation with initial conditions at* $t = s$*;*

$$F_{tt} - a^2 F_{xx} = 0 \tag{2.8.9}$$
$$F(x,t;s)|_{t=s} = 0 \tag{2.8.10}$$
$$F_t(x,t;s)|_{t=s} = g(x,s) \tag{2.8.11}$$

Proof. We verify that the formula gives the solution; first of all we observe that from the conditions we deduce that (using the chain rule)

$$(F_t + F_s)|_{t=s} \equiv 0, \quad \forall x. \tag{2.8.12}$$

Now we can compute the derivatives of u as follows

$$u_{tt} = \partial_t \left(F(x;t,t) + \int_0^t F_t(x,t;s)\mathrm{d}s \right) \overset{(2.8.12)}{=} F_t(x,t;s)|_{s=t} + \int_0^t F_{tt}(x,t;s)\mathrm{d}s =$$
$$= g(x,t) + a^2 \int_0^t F_{xx}(x,t;s)\mathrm{d}x = g(x,t) + a^2 u_{xx}. \tag{2.8.13}$$

We need to verify the initial conditions: now, clearly $u(x,0) = 0$ because of the integral. Secondly we have

$$u_t(x,0) = F(x,t;s)|_{t=s=0} + \int_0^0 F_t(x,t;s)\mathrm{d}s = 0. \tag{2.8.14}$$

This conclues the proof. ■

If we need to solve the nonhomogeneous wave equation with different initial conditions, we simply write the solution as the sum of the particular solution provided for by Duhamel's principle plus the solution of the homogeneous problem with the given initial conditions. See Problem 2.44.

Solution using D'Alembert's formula Combining Duhamel's principle (Thm. 2.36) with D'Alembert's formula (Thm. 2.2) we obtain

$$u(x,t) = \frac{1}{2a}\int_0^t \int_{x-a(t-s)}^{x+a(t-s)} g(\xi,s)\mathrm{d}\xi\mathrm{d}s. \tag{2.8.15}$$

Remark 2.37 *The integral in (2.8.15) has the following nice interpretation: the value of u at (x,t) in the spacetime plane, is the area integral of $g(x',t')$ over the whole characteristic cone at (x,t) up to $t = 0$. (Picture on board!)*

2.9 The weak solutions of the wave equation

In some applications (and some exercises) it is convenient to extend the meaning of the wave equation to a larger class. As one can plainly see, the D'Alembert equation (Thm. (2.2)) is rather "agnostic" regarding the regularity class of the functions ϕ, ψ, as long as the integration makes sense. However it is not immediately clear what meaning to attribute to the differential equation itself if -say- ϕ is a piecewise continuous function.

For this reason we introduce the notion *weak solutions*, while we refer to the $\mathcal{C}^2$ solutions as *classical solutions*.

Definition 1 (Weak solutions of the wave equation) *A function $u(x,t)$ is called a* **weak solution** *of the wave equation $u_{tt} - a^2 u_{xx} = 0$ on $(x,t) \in \mathbb{R}\times\mathbb{R}$ if, for every $\varphi \in \mathcal{C}_0^\infty(\mathbb{R}^2)$ the following holds:*

$$\int_{\mathbb{R}}\int_{\mathbb{R}} u(x,t)(\varphi_{tt}(x,t) - a^2\varphi_{xx}(x,t))\mathrm{d}x\mathrm{d}t = 0 \tag{2.9.1}$$

This is accompanied with the definition of weak solution subject to IC and also external forge

Definition 2 *A function $u(x,t)$ is called a* **weak solution** *of the wave equation $u_{tt} - a^2 u_{xx} = g(x,t)$ on $(x,t) \in \mathbb{R} \times [0,\infty)$ subject to the initial conditions (IC)*

$$u(x,0) = \phi(x); \qquad u_t(x,0) = \psi(x) \tag{2.9.2}$$

if [5] *for every $\varphi \in \mathcal{C}_0^\infty(\mathbb{R} \times [0,\infty))$ the following holds*

$$\int_0^\infty \int_{\mathbb{R}} u(\varphi_{tt} - a^2\varphi_{xx})\mathrm{d}x\mathrm{d}t + \int_{\mathbb{R}} \Big(\varphi_t(x,0)\phi(x) - \varphi(x,0)\psi(x)\Big) = \int_0^\infty\int_{\mathbb{R}} g\varphi \mathrm{d}x\mathrm{d}t \tag{2.9.3}$$

The motivation of these definitions relies on the notion of "distribution" that the reader may have already encountered. It is motivated by the following

Proposition 2.38 *If $u(x,t)$ is a classical solution of the forced DE + IC, then it is also a weak solution in the sense of Def 2.*

Proof. The proof consists of the following chain of identities. For an arbitrary $\varphi \in \mathcal{C}_0^\infty(\mathbb{R} \times [0,\infty))$ let $R > 0$ be sufficiently large so that $\mathrm{supp}\varphi \subset [-R,R] \times [0,R]$. The value R is understood to be such that the support of φ does not intersect the left, right and top sides of the boundary of the rectangle $[-R,R] \times [0,R]$ but, of course, it may intersect the segment $(x,t) \in (-R,R) \times \{0\}$ (picture on board!)

$$\iint\limits_{\mathbb{R}_+\times\mathbb{R}} u\varphi = \iint\limits_{\mathbb{R}_+\times\mathbb{R}} \Big(u_{tt} - a^2 u_{xx}\Big)\varphi = \int_{\mathbb{R}} \mathrm{d}x \int_0^\infty \mathrm{d}t\, u_{tt}\varphi - a^2 \int_0^\infty \mathrm{d}t \int_{-R}^R \mathrm{d}x\mathrm{d}t\, u_{xx}\varphi \tag{2.9.4}$$

The inner integral in the second term can be integrated by parts twice without contribution from the boundary $x = -R, R$; the inner integral in the first term, on the other hand should be handled with some care:

$$\begin{aligned}\int_{\mathbb{R}} \mathrm{d}x \int_0^\infty \mathrm{d}t u_{tt}\varphi &= \int_{\mathbb{R}}\mathrm{d}x\int_0^R \mathrm{d}t\; u_{tt}\varphi = \int \mathrm{d}x(u_t\varphi)\Big|_{t=0}^{t=R} - \int_{\mathbb{R}}\mathrm{d}x\int_0^R \mathrm{d}t\; u_t\varphi_t = \\ &= -\int \mathrm{d}x\, \psi\varphi|_{t=0} - \int \mathrm{d}x(u\varphi_t)\Big|_{t=0}^{t=R} + \int_{\mathbb{R}}\mathrm{d}x\int_0^R \mathrm{d}t\; u\varphi_{tt} = \\ &= -\int \mathrm{d}x\, \psi\varphi|_{t=0} + \int \mathrm{d}x(\phi\varphi_t)|_{t=0} + \int_{\mathbb{R}}\mathrm{d}x\int_0^\infty \mathrm{d}t\; u\varphi_{tt}.\end{aligned} \tag{2.9.5}$$

Recombining the terms yields

$$\iint\limits_{\mathbb{R}_+\times\mathbb{R}} u\varphi = -\int \mathrm{d}x\, \psi\varphi|_{t=0} + \int \mathrm{d}x(\phi\varphi_t)|_{t=0} + \int_{\mathbb{R}}\mathrm{d}x\int_0^\infty \mathrm{d}t\; u(\varphi_{tt} - a^2\varphi_{xx}) \tag{2.9.6}$$

This proves the statement. ■

2.10 Exercises for Chapter 2

Exercise 2.39 *We know that a solution (weak or classical) of* $u_{tt} - u_{xx} = 0$ *is the sum of a left and right traveling waves:* $u(x,t) = f(x-t) + g(x+t)$*. Suppose now that* f, g *are only* $\mathcal{C}_0^1(\mathbb{R})$ *so that* u *is a weak(er) solution.*

1. *Show that for any* $t \in \mathbb{R}$ *fixed, the function* $u(x,t)$ *is compactly supported with respect to* x*.*
2. *Show that the energy*
$$E = \frac{1}{2}\int_{\mathbb{R}} (u_t^2 + u_x^2)\mathrm{d}x \tag{2.10.1}$$
is well defined (i.e. not infinite).
3. *Show that the energy is still conserved. Show also that the energy is the sum of the energy of the left and right traveling waves. Note that* f, g *are not assumed to be twice differentiable and hence you cannot use this for showing the conservation of energy.*

Exercise 2.40 *Prove a similar statement as Prop. 2.38 for the first definition of weak solution, Def. 1*

Exercise 2.41 *Give an appropriate definition of the notion of weak solution for the following DE+IC+BC for the finite string* $x \in [0, \ell]$

$$\begin{aligned} &(DE) \;\; u_{tt} - u_{xx} = g, \\ &(IC) \;\; u(x,0) = \phi(x), \;\; u_t(x,0) = \psi(x), \\ &(BC) \; u(x,0) = 0 = u(\ell, 0) \end{aligned} \tag{2.10.2}$$

Exercise 2.42 *Let* $f(x)$ *be a piecewise continuous function on* $\mathbb{R}$*. Show that* $u(x,t) = f(x-t)$ *is a weak solution of* $u_{tt} - u_{xx} = 0$*.*

Exercise 2.43 *For few instants of time* $t \geqslant 0$ *make a graph of the solution* $u(x,t)$ *to the wave equation with the initial data*

$$u(x,0) = 0, \quad u_t(x,0) = \left\{ \begin{array}{ll} 1, & x \in [x_0, x_1] \\ 0 & \text{otherwise} \end{array} \right., \quad -\infty < x < \infty.$$

Exercise 2.44 *Solve the following DE + IC on the whole line* $x \in \mathbb{R}$*:*

$$u_{tt} - u_{xx} = x - t \tag{2.10.3}$$
$$u(x,0) = x^4 \tag{2.10.4}$$
$$u_t(x,0) = \sin(x) \tag{2.10.5}$$

Exercise 2.45 *For few instants of time* $t \geqslant 0$ *make a graph of the solution* $u(x,t)$ *to the wave equation on the half line* $x \geqslant 0$ *with the free boundary condition*

$$u_x(0,t) = 0$$

and with the initial data

$$u(x,0) = \phi(x), \quad u_t(x,0) = 0, \quad x > 0$$

where the graph of the function $\phi(x)$ *is an isosceles triangle of height* 1 *and the base* $[l, 3l]$*.*

Exercise 2.46 *For few instants of time $t \geqslant 0$ make a graph of the solution $u(x,t)$ to the wave equation on the half line $x \geqslant 0$ with the fixed point boundary condition*

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

and with the initial data

$$u(x,0) = 0, \quad u_t(x,0) = \begin{cases} 1, & x \in [l, 3l] \\ 0, & \text{otherwise} \end{cases}, \quad x > 0.$$

Exercise 2.47 *Prove that*

$$\sum_{n=1}^{\infty} \frac{\sin nx}{n} = \frac{\pi - x}{2} \quad \text{for} \quad 0 < x < 2\pi.$$

Compute the sum of the Fourier series for all other values of $x \in \mathbb{R}$.

Exercise 2.48 *Compute the sums of the following Fourier series:*

$$\sum_{n=1}^{\infty} \frac{\sin 2nx}{2n}, \quad 0 < x < \pi;$$
$$\sum_{n=1}^{\infty} \frac{(-1)^n}{n} \sin nx, \quad |x| < \pi.$$

Exercise 2.49 *Prove that*

$$x^2 = \frac{\pi^2}{3} + 4 \sum_{n=1}^{\infty} \frac{(-1)^n}{n^2} \cos nx, \quad |x| < \pi.$$

Exercise 2.50 *Compute the sums of the following Fourier series:*

$$\sum_{n=1}^{\infty} \frac{\cos(2n-1)x}{(2n-1)^2}$$
$$\sum_{n=1}^{\infty} \frac{\cos nx}{n^2}.$$

Exercise 2.51 *Denote*

$$S_n(x) = \frac{4}{\pi} \sum_{k=1}^{n} \frac{\sin(2k-1)x}{2k-1}$$

the n-th partial sum of the Fourier series (2.5.35). Prove that

1) for any $x \in (-\pi, \pi)$

$$\lim_{n\to\infty} S_n(x) = \operatorname{sign} x.$$

2) Verify that the n-th partial sum has a maximum at

$$x_n = \frac{\pi}{2n}.$$

Hint: derive the following expression for the derivative

$$S_n'(x) = \frac{2}{\pi}\frac{\sin 2nx}{\sin x}.$$

3) Prove that

$$S_n(x_n) = \frac{2}{\pi}\sum_{k=1}^{n}\frac{\pi}{n}\cdot\frac{\sin\frac{(2k-1)\pi}{2n}}{\frac{(2k-1)\pi}{2n}} \to \frac{2}{\pi}\int_0^\pi \frac{\sin x}{x}\,dx \simeq 1.17898$$

for $n \to \infty$.

Thus for the trigonometric series (2.5.35)

$$\limsup_{n\to\infty} S_n(x) > 1 \quad \text{for} \quad x > 0.$$

In a similar way one can prove that

$$\liminf_{n\to\infty} S_n(x) < -1 \quad \text{for} \quad x < 0.$$

Exercise 2.52 *Consider the DE $u_{tt} - u_{xx} = 0$ on the semi-infinite axis $x \in [0, \infty)$ with Neumann boundary conditions and the following IC:*

$$u(x,0) = \phi(x); \quad u_t(x,0) = \phi'(x) \tag{2.10.6}$$

where ϕ is the smooth compactly supported function

$$\phi(x) = \begin{cases} (x-1)^3(2-x)^3 & x =\in [1,2] \\ 0 & x \notin [1,2]. \end{cases} \tag{2.10.7}$$

Give a sketch of ϕ and describe the evolution of the string in the following three intervals of time:

$$t \in [0,1], \quad t \in [1,2], \quad t \geqslant 2. \tag{2.10.8}$$

Also answer the same question where the Neumann condition is replaced with a Dirichlet condition.

Chapter 3

Laplace equation

3.1 Ill-posedness of the Cauchy problem for the Laplace equation

In the study of various classes of solutions to the Cauchy problem for the wave equation we were able to establish

- *existence* of the solution in a suitable class of functions;
- *uniqueness* of the solution;
- *continuous dependence* of the solution on the initial data (see Exercise 2.3 above) with respect to a suitable topology.

One may ask whether these properties remain valid for all evolutionary PDEs satisfying conditions of the Cauchy – Kovalevskaya theorem?

Let us consider a counterexample found by J.Hadamard (1922). Changing the sign in the wave equation one arrives at an equation of *elliptic* type

$$u_{tt} + a^2 u_{xx} = 0. \tag{3.1.1}$$

(The equation (3.1.1) is usually called *Laplace equation*.) Does the change of the type of equation affect seriously the properties of solutions?

To be more specific we will deal with the periodic Cauchy problem

$$u(x,0) = \phi(x), \quad u_t(x,0) = \psi(x) \tag{3.1.2}$$

with two 2π-periodic smooth initial functions $\phi(x)$, $\psi(x)$. For simplicity let us choose $a = 1$. We will see that the solution to this Cauchy problem *does not* depend continuously on the initial data. To do this let us consider the following sequence of initial data: for any integer $k > 0$ denote $u_k(x,t)$ solution to the Cauchy problem

$$u(x,0) = 0, \quad u_t(x,0) = \frac{\sin k\, x}{k}. \tag{3.1.3}$$

The 2π-periodic solution can be expanded in Fourier series

$$u_k(x,t) = \frac{a_0(t)}{2} + \sum_{n=1}^{\infty} \left[a_n(t)\cos nx + b_n(t)\sin nx\right]$$

with some coefficients $a_n(t)$, $b_n(t)$. Substituting the series into equation

$$u_{tt} + u_{xx} = 0$$

we obtain an infinite system of ODEs

$$\begin{aligned} \ddot{a}_n &= n^2 a_n \\ \ddot{b}_n &= n^2 b_n, \end{aligned}$$

$n = 0, 1, 2, \ldots$. The initial data for this infinite system of ODEs follow from the Cauchy problem (3.1.2):

$$\begin{aligned} a_n(0) = 0, &\quad \dot{a}_n = 0 \quad \forall n, \\ b_n(0) = 0, &\quad \dot{b}_n(0) = \begin{cases} 1/k, & n = k \\ 0, & n \neq k. \end{cases} \end{aligned}$$

The solution has the form

$$\begin{aligned} a_n(t) &= 0 \quad \forall n, \quad b_n(t) = 0 \quad \forall n \neq k \\ b_k(t) &= \frac{1}{k^2} \sinh kt. \end{aligned}$$

So the solution to the Cauchy problem (3.1.2) reads

$$u_k(x,t) = \frac{1}{k^2} \sin kx \, \sinh kt. \tag{3.1.4}$$

Using this explicit solution we can prove the following

Theorem 3.1 *For any positive ϵ, M, t_0 there exists an integer K such that for any $k \geqslant K$ the initial data (3.1.3) satisfy*

$$\sup_{x \in [0, 2\pi]} \left(|u_k(x,0)| + |\partial_t u_k(x,0)| \right) < \epsilon \tag{3.1.5}$$

but the solution $u_k(x,t)$ at the moment $t = t_0 > 0$ satisfies

$$\sup_{x \in [0, 2\pi]} \left(|u_k(x,t_0)| + |\partial_t u_k(x,t_0)| \right) \geqslant M. \tag{3.1.6}$$

Proof: Choosing an integer K_1 satisfying

$$K_1 > \frac{1}{\epsilon}$$

we will have the inequality (3.1.5) for any $k \geqslant K_1$. In order to obtain a lower estimate of the form (3.1.6) let us first observe that

$$\sup_{x \in [0, 2\pi]} \left(|u_k(x,t)| + |\partial_t u_k(x,t)| \right) = \frac{1}{k^2} \sinh kt + \frac{1}{k} \cosh kt > \frac{e^{kt}}{k^2}$$

where we have used an obvious inequality

$$\frac{1}{k} > \frac{1}{k^2} \quad \text{for} \quad k > 1.$$

The function

$$y = \frac{e^x}{x^2}$$

is monotone increasing for $x > 2$ and

$$\lim_{x \to +\infty} \frac{e^x}{x^2} = +\infty.$$

Hence for any $t_0 > 0$ there exists x_0 such that

$$\frac{e^x}{x^2} > \frac{M}{t_0^2} \quad \text{for} \quad x > x_0.$$

Let K_2 be a positive integer satisfying

$$K_2 > \frac{x_0}{t_0}.$$

Then for any $k > K_2$

$$\frac{e^{k\,t_0}}{k^2} = t_0^2 \frac{e^{k\,t_0}}{(k\,t_0)^2} > t_0^2 \frac{e^{x_0}}{x_0^2} > M.$$

Choosing

$$K = \max(K_1, K_2)$$

we complete the proof of the Theorem. ■

The statement of the Theorem is usually referred to as *ill-posedness* of the Cauchy problem (3.1.1), (3.1.2).

A natural question arises: what kind of initial or boundary conditions can be chosen in order to uniquely specify solutions to Laplace equation without violating the continuous dependence of the solutions on the boundary/initial conditions?

3.2 Dirichlet and Neumann problems for Laplace equation on the plane

The *Laplace operator* in the d-dimensional Euclidean space is defined by

$$\Delta = \frac{\partial^2}{\partial x_1^2} + \ldots + \frac{\partial^2}{\partial x_d^2}. \tag{3.2.1}$$

The symbol (coinciding with the principal symbol) of this operator is equal to

$$-\left(\xi_1^2 + \ldots + \xi_d^2\right) < 0 \quad \text{for all} \quad \xi \neq 0.$$

So Laplace operator is an example of an *elliptic* operator.

In this section we will formulate the two main boundary value problems (b.v.p.'s) for the *Laplace equation*

$$\Delta u = 0, \quad u = u(x), \quad x \in \Omega \subset \mathbb{R}^d. \tag{3.2.2}$$

The solutions to the Laplace equation are called *harmonic functions* in the domain Ω.

We will assume that the boundary $\partial\Omega$ of the domain Ω is a smooth hypersurface. Moreover we assume that the domain Ω does not go to infinity, i.e., Ω belongs to some ball in $\mathbb{R}^d$. Denote $n = n(x)$ the unit external normal vector at every point $x \in \partial\Omega$ of the boundary.

Problem 1 (*Dirichlet problem*). Given a function $f(x)$ defined at the points of the boundary find a function $u = u(x)$ satisfying the Laplace equation on the internal part of the domain Ω and the boundary condition

$$u(x)|_{x\in\partial\Omega} = f(x) \tag{3.2.3}$$

on the boundary of the domain.

Problem 2 (*Neumann problem*). Given a function $g(x)$ defined at the points of the boundary find a function $u = u(x)$ satisfying the Laplace equation on the internal part of the domain Ω and the boundary condition

$$\left(\frac{\partial u(x)}{\partial n}\right)_{x\in\partial\Omega} = g(x) \tag{3.2.4}$$

on the boundary of the domain.

Example 1. For $d = 1$ the Laplace operator is just the second derivative

$$\Delta = \frac{d^2}{dx^2}.$$

The Dirichlet b.v.p. in the domain $\Omega = (a, b)$

$$u''(x) = 0, \quad u(a) = f_a, \quad u(b) = f_b$$

has an obvious unique solution

$$u(x) = \frac{f_b - f_a}{b - a}(x - a) + f_a.$$

The Neumann b.v.p. in the same domain

$$u''(x) = 0, \quad -u'(a) = g_a, \quad u'(b) = g_b$$

has solution only if

$$g_a + g_b = 0. \tag{3.2.5}$$

Example 2. In two dimensions the Laplace operator reads.

$$\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}. \tag{3.2.6}$$

Exercise 3.2 *Prove that in the polar coordinates*

$$\left.\begin{aligned} x =& \; r\cos\phi \\ y =& \; r\sin\phi \end{aligned}\right\} \tag{3.2.7}$$

the Laplace operator takes the form

$$\Delta = \frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial \phi^2}. \tag{3.2.8}$$

In the particular case

$$\Omega = \{(x, y) \,|\, x^2 + y^2 < \rho^2\} \tag{3.2.9}$$

(a circle of radius ρ) the Dirichlet b.v.p. is formulated as follows: find a solution to the Laplace equation

$$\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0, \quad u = u(x, y), \quad \text{for} \quad x^2 + y^2 < \rho^2 \tag{3.2.10}$$

satisfying the boundary condition

$$u\,|_{r=\rho} = f(\phi). \tag{3.2.11}$$

Here we represent the boundary condition defined on the boundary of the circle as a function depending only on the polar angle ϕ. Similarly, the Neumann problem consists of finding a solution to the Laplace equation satisfying

$$\left(\rho\,\frac{\partial u}{\partial r}\right)_{r=\rho} = g(\phi) \tag{3.2.12}$$

for a given function $g(\phi)$. The factor ρ in the left side is only a convenient normalization of the boundary data.

Let us return to the general d-dimensional case. The following identity will be useful in the study of harmonic functions.

Theorem 3.3 (Green's formula) . *For arbitrary smooth functions u, v on the* **closed and bounded** *domain $\bar{\Omega}$ with a piecewise smooth boundary $\partial\Omega$ the following identity holds true*

$$\int_\Omega \nabla u \cdot \nabla v \,\mathrm{d}V + \int_\Omega u\,\Delta v \,\mathrm{d}V = \int_{\partial\Omega} u\,\frac{\partial v}{\partial n}\,\mathrm{d}S. \tag{3.2.13}$$

where $\partial v/\partial n$ denotes the directional derivative of v along the outer normal vector $\mathbf{n}$

Here

$$\nabla u \cdot \nabla v = \sum_{i=1}^{d} \frac{\partial u}{\partial x_i}\frac{\partial v}{\partial x_i}$$

is the inner product of the gradients of the functions,

$$dV = dx_1 \dots dx_d$$

is the Euclidean volume element, n the external normal and dS is the area element on the hypersurface $\partial\Omega$.

This identity is a consequence of another identity known as the **Divergence Theorem**:

Theorem 3.4 (Divergence theorem) *Let $\Omega \subset \mathbb{R}^d$ be a bounded domain (open and connected set) with piecewise smooth boundary $\partial\Omega$. Let $\vec{F} : \Omega \to \mathbb{R}^d$ be a vector field of class $\mathcal{C}^1(\Omega)$ and $\mathcal{C}^0(\overline{\Omega})$. Then*

$$\int_\Omega \operatorname{div} \vec{F}\,\mathrm{d}V = \int_{\partial\Omega} \vec{F}\cdot\mathbf{n}\,\mathrm{d}S \tag{3.2.14}$$

Example 1. For $d = 1$ and $\Omega = (a, b)$ the Green's formula reads

$$\int_a^b u_x\, v_x\, dx + \int_a^b u\, v_{xx}\, dx = u\, v_x|_a^b$$

since the oriented boundary of the interval consists of two points $\partial[a, b] = b - a$. This is an easy consequence of integration by parts.

Example 2. For $d = 2$ and a rectangle $\Omega = (a, b) \times (c, d)$ the Green's formula becomes

$$\int_\Omega (u_x v_x + u_y v_y)\, dx\, dy + \int_\Omega u\,(v_{xx} + v_{yy})\, dx\, dy = \int_a^b (u\, v_y)_c^d\, dx + \int_c^d (u\, v_x)_a^b\, dy$$

(the sum of integrals over four pieces of the boundary $\partial\Omega$ stands in the right hand side of the formula).

Let us return to the general discussion of Laplace equation. The following corollary follows immediately from the Green's formula.

Corollary 3.5 *For a function u harmonic in a* bounded *domain Ω with a piecewise smooth boundary the following identity holds true*

$$\int_\Omega (\nabla u)^2 = \int_{\partial\Omega} \frac{1}{2} \partial_n u^2\, dS. \tag{3.2.15}$$

Proof: This is obtained from (3.2.13) by choosing $u = v$. ■

Using this identity we can easily derive uniqueness of solution to the Dirichlet problem.

Theorem 3.6 *1) Let u_1, u_2 be two functions harmonic in the bounded domain Ω and smooth in the closed domain $\bar{\Omega}$ coinciding on the boundary $\partial\Omega$. Then $u_1 \equiv u_2$.*

2) Under the same assumptions about the functions u_1, u_2, if the normal derivatives on the boundary coincide

$$\frac{\partial u_1}{\partial n} = \frac{\partial u_2}{\partial n}$$

then the functions differ by a constant.

Proof: Applying to the difference $u = u_2 - u_1$ the identity (3.2.15) one obtains

$$\int_\Omega (\nabla u)^2 dV = 0$$

since the right hand side vanishes. Hence $\nabla u = 0$, and thus the function u is equal to a constant. The value of this constant on the boundary is zero. Therefore $u \equiv 0$. The second statement has a similar proof. ■

The following counterexample shows that the uniqueness does not hold true for *infinite* domains. Let Ω be the upper half plane:

$$\Omega = \{(x, y) \in \mathbb{R}^2 \,|\, y > 0.\}.$$

The linear function $u(x, y) = y$ is harmonic in Ω and vanishes on the boundary. Clearly $u \neq 0$ on Ω.

Our goal is to solve the Dirichlet and Neumann boundary value problems. The first result in this direction is the following

Theorem 3.7 (Solution of the Laplace equation on a disk: Dirichlet problem) *For an arbitrary $\mathcal{C}^1$-smooth 2π-periodic function $f(\phi)$ the solution to the Dirichlet b.v.p. (3.2.10), (3.2.11) exists and is unique. Moreover it is given by the following formula*

$$u(r, \phi) = \frac{1}{2\pi} \int_0^{2\pi} \frac{\rho^2 - r^2}{\rho^2 - 2\rho\, r\, \cos(\phi - \psi) + r^2} f(\psi)\, d\psi. \tag{3.2.16}$$

The expression (3.2.16) for the solution to the Dirichlet b.v.p. in the circle is called *Poisson formula.*

Proof: We will first use the method of separation of variables in order to construct particular solutions to the Laplace equation. At the second step we will represent solutions to the Dirichlet b.v.p. as a linear combination of the particular solutions.

The method of *separation of variables* starts from looking for solutions to the Laplace equation in the form

$$u = R(r)\Phi(\phi). \tag{3.2.17}$$

Here r, ϕ are the polar coordinates on the plane (see Exercise 3.2 above). Using the form (3.2.8) we reduce the Laplace equation $\Delta u = 0$ to

$$R''(r)\Phi(\phi) + \frac{1}{r} R'(r)\Phi(\phi) + \frac{1}{r^2} R(r)\Phi''(\phi) = 0.$$

After division by $\frac{1}{r^2} R(r)\Phi(\phi)$ we can rewrite the last equation in the form

$$\frac{R''(r) + \frac{1}{r} R'(r)}{\frac{1}{r^2} R(r)} = -\frac{\Phi''(\phi)}{\Phi(\phi)}.$$

The left hand side of this equation depends on r while the right hand side depends on ϕ. The equality is possible only if both sides are equal to some constant λ. In this way we arrive at two ODEs for the functions $R = R(r)$ and $\Phi = \Phi(\phi)$

$$R'' + \frac{1}{r} R' - \frac{\lambda}{r^2} R = 0 \tag{3.2.18}$$

$$\Phi'' + \lambda \Phi = 0. \tag{3.2.19}$$

We have now to determine the admissible values of the parameter λ. To this end let us begin from the second equation (3.2.19). Its solutions have the form

$$\Phi(\phi) = \begin{cases} A\, e^{\sqrt{-\lambda}\,\phi} + B\, e^{-\sqrt{-\lambda}\,\phi}, & \lambda < 0 \\ A + B\,\phi, & \lambda = 0 \\ A\,\cos\sqrt{\lambda}\,\phi + B\,\sin\sqrt{\lambda}\,\phi, & \lambda > 0 \end{cases}.$$

Since the pairs of polar coordinates (r,ϕ) and $(r,\phi+2\pi)$ correspond to the same point on the Euclidean plane the solution $\Phi(\phi)$ must be a 2π-periodic function. Hence we must discard the negative values of λ. Moreover λ must have the form

$$\lambda = n^2, \quad n = 0,\, 1,\, 2, \dots. \tag{3.2.20}$$

This gives

$$\Phi(\phi) = A\,\cos n\phi + B\,\sin n\phi. \tag{3.2.21}$$

The first ODE (3.2.18) for $\lambda = n^2$ becomes

$$R'' + \frac{1}{r}R' - \frac{n^2}{r^2}R = 0.$$

This is a particular case of Euler equation. One can look for solutions in the form

$$R(r) = r^k.$$

The exponent k has to be determined from the characteristic equation

$$k(k-1) + k - n^2 = 0$$

obtained by the direct substitution of $R = r^k$ into the equation. The roots of the characteristic equation are $k = \pm n$. For $n > 0$ this gives the general solution of the equation (3.2.18) in the form

$$R = a\,r^n + \frac{b}{r^n}$$

with two integration constants a and b. For $n = 0$ the general solution is

$$R = a + b\,\log r.$$

As the solution must be smooth at $r = 0$ one must always choose $b = 0$ for all n. In this way we arrive at the following family of particular solutions to the Laplace equation

$$u_n = r^n\left(a_n \cos n\phi + b_n \sin n\phi\right), \quad n = 0,\, 1,\, 2, \dots \tag{3.2.22}$$

We want now to represent any solution to the Dirichlet b.v.p. in the circle of radius ρ as a linear combination of these solutions:

$$\begin{aligned} u &= \frac{A_0}{2} + \sum_{n\geqslant 1} r^n\left(A_n \cos n\phi + B_n \sin n\phi\right) \\ u|_{r=\rho} &= f(\phi). \end{aligned} \tag{3.2.23}$$

The boundary data function $f(\phi)$ must be a 2π-periodic function. Assuming this function to be $\mathcal{C}^1$-smooth let us expand it in Fourier series

$$f(\phi) = \frac{a_0}{2} + \sum_{n\geqslant 1} (a_n \cos n\phi + b_n \sin n\phi)$$

$$a_n = \frac{1}{\pi}\int_0^{2\pi} f(\phi)\cos n\phi\, d\phi, \quad b_n = \frac{1}{\pi}\int_0^{2\pi} f(\phi)\sin n\phi\, d\phi. \tag{3.2.24}$$

Comparison of (3.2.23) with (3.2.24) yields

$$A_n = \frac{a_n}{\rho^n}, \quad B_n = \frac{b_n}{\rho^n},$$

or, equivalently

$$u = \frac{a_0}{2} + \sum_{n\geqslant 1}\left(\frac{r}{\rho}\right)^n (a_n \cos n\phi + b_n \sin n\phi)\,. \tag{3.2.25}$$

Recall that this formula holds true on the circle of radius ρ, i.e., for

$$r \leqslant \rho.$$

The last formula can be rewritten as follows:

$$\begin{aligned} u &= \frac{1}{\pi}\int_0^{2\pi}\left[\frac{1}{2} + \sum_{n\geqslant 1}\left(\frac{r}{\rho}\right)^n (\cos n\phi\cos n\psi + \sin n\phi \sin n\psi)\right] f(\psi)\, d\psi \\ &= \frac{1}{\pi}\int_0^{2\pi}\left[\frac{1}{2} + \sum_{n\geqslant 1}\left(\frac{r}{\rho}\right)^n \cos n(\phi-\psi)\right] f(\psi)\, d\psi. \end{aligned}$$

To compute the sum in the square bracket we represent it as a geometric series converging for $r < \rho$:

$$\begin{aligned} &\frac{1}{2} + \sum_{n\geqslant 1}\left(\frac{r}{\rho}\right)^n \cos n(\phi-\psi) = \frac{1}{2} + \mathrm{Re}\sum_{n\geqslant 1}\left(\frac{r}{\rho}\right)^n e^{in(\phi-\psi)} \\ &= \frac{1}{2} + \mathrm{Re}\,\frac{r\, e^{i(\phi-\psi)}}{\rho - r\, e^{i\,(\phi-\psi)}} = \frac{1}{2} + \frac{1}{2}\left(\frac{r\, e^{i(\phi-\psi)}}{\rho - r\, e^{i\,(\phi-\psi)}} + \frac{r\, e^{-i(\phi-\psi)}}{\rho - r\, e^{-i\,(\phi-\psi)}}\right) \\ &= \frac{1}{2}\frac{\rho^2 - r^2}{\rho^2 - 2\rho\, r\,\cos(\phi-\psi) + r^2}. \end{aligned}$$

■

In a similar way one can treat the Neumann boundary problem. However in this case one has to impose an additional constraint for the boundary value of the normal derivative (cf. (3.2.5) above in dimension 1).

Lemma 3.8 *Let v be a smooth function on the closed domain $\bar\Omega$ harmonic inside the domain. Then the integral of the normal derivative of v over the boundary $\partial\Omega$ vanishes:*

$$\int_{\partial\Omega}\frac{\partial v}{\partial n}\, dS = 0. \tag{3.2.26}$$

Proof: Applying the Green formula to the pair of functions $u \equiv 1$ and v one obtains

$$\int_\Omega \Delta v \, dV = \int_{\partial\Omega} \frac{\partial v}{\partial n} \, dS.$$

The left hand side of the equation vanishes since $\Delta v = 0$ in Ω. ■

Corollary 3.9 *The Neumann problem (3.2.4) can have a solution only if the boundary function g satisfies*

$$\int_{\partial\Omega} g \, dS = 0. \tag{3.2.27}$$

We will now prove, for the particular case of a circle domain in the dimension $d = 2$ that this necessary condition of solvability is also a sufficient one.

Theorem 3.10 (Solution of the Laplace equation on a disk: Neumann problem.) *For an arbitrary $\mathcal{C}^1$-smooth 2π-periodic function $g(\phi)$ satisfying*

$$\int_0^{2\pi} g(\phi) \, d\phi = 0 \tag{3.2.28}$$

the Neumann b.v.p. (3.2.10), (3.2.12) has a solution unique up to an additive constant. This solution can be represented by the following integral formula

$$u(r, \phi) = \frac{1}{2\pi} \int_0^{2\pi} \log \frac{\rho^2}{\rho^2 - 2\rho\, r\, \cos(\phi - \psi) + r^2} \, g(\psi) \, d\psi. \tag{3.2.29}$$

Proof: Repeating the above arguments one arrives at the following expression for the solution $u = u(r, \phi)$:

$$\begin{aligned} u &= \frac{A_0}{2} + \sum_{n \geqslant 1} r^n \left(A_n \cos n\phi + B_n \sin n\phi \right) \\ \left(\rho \frac{\partial u}{\partial r} \right)_{r=\rho} &= g(\phi). \end{aligned} \tag{3.2.30}$$

Let us consider the Fourier series of the function $g(\phi)$

$$g(\phi) = \frac{a_0}{2} + \sum_{n \geqslant 1} \left(a_n \cos n\phi + b_n \sin n\phi \right).$$

Due to the constraint (3.2.28) the constant term vanishes:

$$a_0 = 0.$$

Comparing this series with the boundary condition (3.2.30) we find that

$$u(r,\phi) = \frac{A_0}{2} + \sum_{n\geqslant 1} \frac{1}{n}\left(\frac{r}{\rho}\right)^n (a_n \cos n\phi + b_n \sin n\phi)$$

$$a_n = \frac{1}{\pi}\int_0^{2\pi} \cos n\psi\, g(\psi)\, d\psi, \quad b_n = \frac{1}{\pi}\int_0^{2\pi} \sin n\psi\, g(\psi)\, d\psi.$$

Here A_0 is an arbitrary constant. Combining the two last equations we arrive at the following expression:

$$u(r,\phi) = \frac{1}{\pi}\int_0^{2\pi} \sum_{n\geqslant 1} \frac{1}{n}\left(\frac{r}{\rho}\right)^n \cos n(\phi-\psi) g(\psi)\, d\psi. \tag{3.2.31}$$

It remains to compute the sum of the trigonometric series in the last formula.

Lemma 3.11 *Let R and θ be two real numbers, $R < 1$. Then*

$$\sum_{n=1}^{\infty} \frac{1}{n} R^n \cos n\theta = \frac{1}{2}\log\frac{1}{1-2R\cos\theta + R^2}. \tag{3.2.32}$$

Proof: The series under consideration can be represented as the real part of a complex series

$$\sum_{n=1}^{\infty} \frac{1}{n} R^n \cos n\theta = \mathrm{Re} \sum_{n=1}^{\infty} \frac{1}{n} R^n e^{in\theta}.$$

The latter can be written as follows:

$$\sum_{n=1}^{\infty} \frac{1}{n} R^n e^{in\theta} = \int_0^R \sum_{n=1}^{\infty} \frac{1}{R} R^n e^{in\theta}\, dR.$$

We can easily compute the sum of the geometric series with the denominator $Re^{i\theta}$. Integrating we obtain

$$\sum_{n=1}^{\infty} \frac{1}{n} R^n e^{in\theta} = \int_0^R \frac{e^{i\theta}}{1-Re^{i\theta}}\, dR = -\log\left(1-Re^{i\theta}\right).$$

Hence

$$\sum_{n=1}^{\infty} \frac{1}{n} R^n \cos n\theta = \frac{1}{2}\left[\log\frac{1}{1-Re^{i\theta}} + \log\frac{1}{1-Re^{-i\theta}}\right] = \frac{1}{2}\log\frac{1}{1-2R\cos\theta+R^2}.$$

■

Applying the formula of the Lemma to the series (3.2.31) we complete the proof of the Theorem.
■

3.3 Properties of harmonic functions: mean value theorem, the maximum principle

In this section we will establish, for the specific case of dimension $d = 2$, the two fundamental properties of harmonic functions.

Let $\Omega \subset \mathbb{R}^d$ be a domain. Recall that a point $x_0 \in \Omega$ is called *internal* if there exists a ball of some radius $R > 0$ with the centre at x_0 entirely belonging to Ω. For an internal point $x_0 \in \Omega$ denote

$$S^{d-1}(x_0, R) = \{x \in \mathbb{R}^d \,|\, |x - x_0| = R\}$$

a sphere of radius $R > 0$ with the center at x_0.

Remark 3.12 *The area a_{d-1} of the unit sphere in $\mathbb{R}^d$ can be computed with the following "trick": we start from the d–dimensional Gaussian integral*

$$\int_{\mathbb{R}^d} \mathrm{e}^{-\|x\|^2} \mathrm{d}V = \pi^{\frac{d}{2}}. \tag{3.3.1}$$

Rewriting it in "spherical" coordinates it reads

$$\int_0^\infty r^{d-1} \mathrm{e}^{-r^2} \mathrm{d}r \int_{S^{d-1}} \mathrm{d}S = a_{d-1} \frac{1}{2} \Gamma\left(\frac{d}{2}\right) \tag{3.3.2}$$

Comparing the two formulas we obtain

$$a_{d-1} = \frac{2\pi^{\frac{d}{2}}}{\Gamma\left(\frac{d}{2}\right)}. \tag{3.3.3}$$

The radius is chosen small enough to guarantee that the sphere belongs to the domain Ω. Denote a_{d-1} the area of the unit sphere in $\mathbb{R}^d$. For any continuous function $f(x)$ on the sphere the *mean value* is defined by the formula

$$\bar{f} = \frac{1}{a_{d-1} R^{d-1}} \int_{S^{d-1}(x_0, R)} f(x) \, dS. \tag{3.3.4}$$

In the particular case of a constant function the mean value coincides with the value of the function.

For example, in dimension $d = 1$ the "sphere" consists of two points $x_0 \pm R$. The formula (3.3.3) for the area of the zero-dimensional sphere gives

$$a_0 = \frac{\pi^{1/2}}{\Gamma\left(\frac{3}{2}\right)} = 2.$$

So the mean value of a function is just the arithmetic mean value of the two numbers $f(x_0 \pm R)$:

$$\bar{f} = \frac{f(x_0 + R) + f(x_0 - R)}{2}.$$

In the next case $d = 2$ the sphere is just a circle of radius R with the centre at x_0. The area (i.e., the length) element is $dS = R \, d\phi$. The restriction of f to the circle is a 2π-periodic function $f(\phi)$. So the mean value on this circle is given by

$$\bar{f} = \frac{1}{2\pi} \int_0^{2\pi} f(\phi) \, d\phi.$$

Theorem 3.13 *Let $u = u(x)$ be a function harmonic in a domain Ω. Then the mean value of u over a small sphere centered at a point $x_0 \in \Omega$ is equal to the value of the function at this point:*

$$u(x_0) = \frac{1}{a_{d-1}R^{d-1}} \int_{S^{d-1}(x_0,R)} u(x)\, dS. \tag{3.3.5}$$

Moreover we also have

$$u(x_0) = \frac{1}{V_d(R)} \int_{B_R(x_0)} u \,\mathrm{d}V \tag{3.3.6}$$

where $V_d(R) = a_{d-1}\frac{R^d}{d}$ is the volume of the ball of radius R.

Proof. We start with $d = 2$. Denote $f(\phi)$ the restriction of the harmonic function u onto the small circle $|x - x_0| = R$. By definition the function $u(x)$ satisfies the Dirichlet b.v.p. inside the circle:

$$\begin{aligned} &\Delta u(x) = 0, \quad |x - x_0| < R \\ &u(x)|_{|x-x_0|=R} = f(\phi). \end{aligned}$$

As we already know from the proof of Theorem 3.7 the solution to this b.v.p. can be represented by the Fourier series

$$u(r, \phi) = \frac{a_0}{2} + \sum_{n\geqslant 1} \left(\frac{r}{R}\right)^n (a_n \cos n\phi + b_n \sin n\phi) \tag{3.3.7}$$

for $r := |x - x_0| < R$ (cf. (3.2.25) above). In this formula a_n and b_n are the Fourier coefficients of the boundary function

$$f(\phi) = u(x)|_{|x-x_0|=R}.$$

In particular

$$\frac{a_0}{2} = \frac{1}{2\pi} \int_0^{2\pi} f(\phi)\, d\phi$$

is the mean value of the function u on the circle. On the other side the value of the function u at the center of the circle can be evaluated substituting $r = 0$ in the formula (3.3.7):

$$u(x_0) = \frac{a_0}{2}.$$

Comparing the last two equations we arrive at (3.3.5).

For general dimension we can proceed as follows: Let $B_r(x_0)$ be the ball of radius r centered at $x_0 \in \Omega \subset \mathbb{R}^d$. Then

$$\begin{aligned} 0 = \int_{B_r(x_0)} \triangle u \overset{\text{Div. Thm.}}{=} \int_{\partial B_r(x_0)} \nabla_{\mathbf{n}} u \,\mathrm{d}S = r^{d-1} \int_{S^{d-1}} \frac{\partial}{\partial r} u(x_0 + ry)\mathrm{d}S(y) = \\ = r^{d-1} \frac{\partial}{\partial r} \int_{S^{d-1}} u(x_0 + ry)\mathrm{d}S(y) \end{aligned} \tag{3.3.8}$$

Now divide by the volume of the sphere $V_d = a_{d-1}\frac{r^d}{d}$ so that (denoting by $⨍$ the average)

$$0 = ⨍_{B_r(x_0)} \triangle u = \frac{d}{r}\frac{\partial}{\partial r} ⨍_{S^{d-1}} u(x_0 + ry)\mathrm{d}S(y) \tag{3.3.9}$$

The integral under differentiation is the average of u over the surface of the ball $B_r(x_0)$. Thus we conclude that

$$⨍_{\partial B_r(x_0)} u\mathrm{d}S = C(x_0) \tag{3.3.10}$$

is a constant independent of the radius of the ball (within the domain Ω). Since $u \in \mathcal{C}^2(\Omega)$ we know that it takes a maximum and minimum on $\partial B_d(x_0)$ (which is compact), and a simple continuity argument shows that, as $r \to 0$ the average must converge to $u(x_0)$. Thus $⨍_{\partial B_r(x_0)} u\mathrm{d}S = u(x_0)$.

The second formula is proven by integration of the first:

$$\int_{B_R(x_0)} u\mathrm{d}V = \int_0^R \left(\int_{S^{d-1}} u(x_0 + ry)\mathrm{d}S(y) \right) r^{d-1}\mathrm{d}r =$$
$$= a_{d-1}u(x_0) \int_0^R r^{d-1}\mathrm{d}r = V_d(R)u(x_0). \tag{3.3.11}$$

Dividing by the volume $V_d(R)$ concludes the proof. ■

Using the mean value theorem we will now prove another important property of harmonic functions, namely the *maximum principle*. Recall that a function $u(x)$ defined on a domain $\Omega \subset \mathbb{R}^d$ is said to have a *local maximum* at the point x_0 if the inequality

$$u(x) \leqslant u(x_0) \tag{3.3.12}$$

holds true for any $x \in \Omega$ sufficiently close to x_0. A *local minimum* is defined in a similar way.

Theorem 3.14 *Let a function $u(x)$ be harmonic in a bounded connected domain Ω and continuous in a closed domain $\bar{\Omega}$. Denote*

$$M = \sup_{x \in \partial\Omega} u(x), \quad m = \inf_{x \in \partial\Omega} u(x).$$

Then

1) $m \leqslant u(x) \leqslant M$ for all $x \in \Omega$;

2) if $u(x) = M$ or $u(x) = m$ for some internal point $x \in \Omega$ then the function u is constant.

Proof: It is based on the following Main Lemma.

Lemma 3.15 *Let the harmonic function $u(x)$ have a local maximum/minimum at an internal point $x_0 \in \Omega$. Then $u(x) \equiv u(x_0)$ on some neighborhood of the point x_0.*

Proof: Let us consider the case of a local maximum. Choosing a sufficiently small sphere with the centre at x_0 we obtain, according to the mean value theorem, that

$$u(x_0) = \frac{1}{a_{d-1}R^{d-1}} \int_{|x-x_0|=R} u(x)\, dS.$$

We can assume the inequality (3.3.12) holds true for all x on the sphere. So

$$u(x_0) = \frac{1}{a_{d-1}R^{d-1}} \int_{|x-x_0|=R} u(x)\, dS \leqslant \frac{1}{a_{d-1}R^{d-1}} \int_{|x-x_0|=R} u(x_0)\, dS = u(x_0). \tag{3.3.13}$$

If there exists a point x sufficiently close to x_0 such that $u(x) < u(x_0)$ then also the inequality (3.3.13) is strict. Such a contradiction shows that the function $u(x)$ takes constant values on some ball with the centre at x_0. The case of a local minimum can be treated in a similar way. ■

Let us return to the proof of the Theorem. Denote

$$M' = \sup_{x\in\bar\Omega} u(x)$$

the maximum of the function u continuous on the compact $\bar\Omega$. We want to prove that $M' \leqslant M$. Indeed, if $M' > M$ then there exists an internal point $x_0 \in \Omega$ such that $u(x_0) = M'$. Denote $\Omega' \subset \Omega$ the set of points x of the domain where the function u takes the same value M'. According to the Main Lemma this subset is open. Clearly it is also closed and nonempty. Hence $\Omega' = \Omega$ since the domain is connect. In other words the function is constant everywhere in Ω. Because of continuity it takes the same value M' at the points of the boundary $\partial\Omega$. Hence $M' \leqslant M$. The contradiction we arrived at shows that the value of a harmonic function at an internal point of the domain cannot be bigger than the value of this function on the boundary of the domain. Moreover if the harmonic function takes the value M at an internal point then it is constant. In a similar way we prove that a non-constant harmonic function cannot have a minimum outside the boundary of the domain. ■

Corollary 3.16 *Given two functions $u_1(x)$, $u_2(x)$ harmonic in a bounded domain Ω and continuous in the closed domain $\bar\Omega$. If*

$$|u_1(x) - u_2(x)| \leqslant \epsilon \quad \text{for} \quad x \in \partial\Omega$$

then

$$|u_1(x) - u_2(x)| \leqslant \epsilon \quad \text{for any} \quad x \in \Omega$$

Proof: Denote

$$u(x) = u_1(x) - u_2(x).$$

The function u is harmonic in Ω and continuous in $\bar\Omega$. By assumption we have $-\epsilon \leqslant u(x) \leqslant \epsilon$ for any $x \in \partial\Omega$. So

$$-\epsilon \leqslant \inf_{x\in\partial\Omega} u(x), \quad \sup_{x\in\partial\Omega} u(x) \leqslant \epsilon.$$

According to the maximum principle it must be also

$$-\epsilon \leqslant \inf_{x\in\Omega} u(x), \quad \sup_{x\in\Omega} u(x) \leqslant \epsilon.$$

■

The Corollary implies that the the solution to the Dirichlet boundary value problem, if exists, *depends continuously* on the boundary data.

3.3.1 Boundary problem on annuli

The Poisson kernels for the Dirichlet and Neumann boundary conditions on circles does not work for other domains. We consider here an annulus $\Omega := \{\rho^2 < x^2 + y^2 < R^2\}$.

Since the domain does not contain the origin, the same considerations already used allow us to say that any harmonic function on Ω must take the form

$$u(r,\theta) = \frac{A_0}{2} + \frac{C_0}{2}\ln r + \sum_{n\geqslant 1} r^n\Big(A_n\cos(n\theta) + B_n\sin(n\theta)\Big) + \frac{C_n\cos(n\theta) + D_n\sin(n\theta)}{r^n} \quad (3.3.14)$$

There are clearly four types of boundary conditions: D-D, D-N, N-D, N-N, where D stands for Dirichlet and N for Neumann. Here we consider only D-D.

Suppose we want to find the kernel for D-D BCs

$$\triangle u = 0, \quad (x,y)\in\Omega \quad (3.3.15)$$

$$u|_{r=\rho} = f(\theta); \quad u|_{r=R} = g(\theta). \quad (3.3.16)$$

Let the Fourier expansion of f, g be

$$f = \frac{\alpha_0}{2} + \sum_{n\geqslant 1}\alpha_n\cos(n\theta) + \beta_n\sin(n\theta); \quad (3.3.17)$$

$$g = \frac{\gamma_0}{2} + \sum_{n\geqslant 1}\gamma_n\cos(n\theta) + \delta_n\sin(n\theta). \quad (3.3.18)$$

The coefficients A_n, B_n, C_n, D_n must solve the system

$$\begin{cases} A_0 + C_0\ln(\rho) = \alpha_0 \\ A_0 + C_0\ln(R) = \gamma_0 \\ A_n\rho^n + \frac{C_n}{\rho^n} = \alpha_n \\ B_n\rho^n + \frac{D_n}{\rho^n} = \beta_n \\ A_nR^n + \frac{C_n}{R^n} = \beta_n \\ B_nR^n + \frac{D_n}{R^n} = \delta_n \end{cases} \quad (3.3.19)$$

It is more practical, in concrete problems, to solve directly the system rather than writing a kernel.

3.3.2 Laplace equation on rectangles

Consider the equation

$$D.E.: \qquad \triangle u = 0, \quad (x,y)\in[0,L]\times[0,M] \quad (3.3.20)$$

$$B.C.: \qquad \begin{cases} u(x,0) = f(x) \\ u(0,y) = h(y) \end{cases} \qquad \begin{cases} u(x,M) = g(x) \\ u(L,y) = k(y) \end{cases} \quad (3.3.21)$$

The B.C. are assumed to be continuous; so, for example $f(0) = h(0)$ and so on. We consider here the simpler case where $f(0) = f(M) = h(0) = h(L) = g(0) = g(L) = k(0) = k(M) = 0$ so that each of the functions f, h, g, k admits periodic odd extensions to continuous functions of periods $2L$ or $2M$.

Namely we assume that all of them have a sin Fourier series representation:

$$f(x) = \sum_{n\geqslant 1} a_n \sin\left(n\pi \frac{x}{L}\right) \qquad g(x) = \sum_{n\geqslant 1} b_n \sin\left(n\pi \frac{x}{L}\right) \tag{3.3.22}$$

$$h(x) = \sum_{n\geqslant 1} c_n \sin\left(n\pi \frac{y}{M}\right) \qquad k(x) = \sum_{n\geqslant 1} d_n \sin\left(n\pi \frac{y}{M}\right) \tag{3.3.23}$$

Consider first the problem where $g = h = k \equiv 0$; if we solve this BVP, then we can analogously solve the others and the complete solution will simply be the sum of the various solutions.

First we look for factorized solutions $u(x,y) = X(x)Y(y)$; plugging into the equation yields separation of variables

$$X''Y + XY'' = 0 \qquad X'' = -\lambda X; \qquad Y'' = \lambda Y. \tag{3.3.24}$$

Depending on the sign of λ we have various possibilities. Since we must have $u(0,y) = u(L,y) = 0$ we quickly conclude that $-\lambda = \frac{n^2\pi^2}{L}$ with $n \in \mathbb{N}$, and we arrive at possible solutions

$$u_n(x,y) = \sin\left(\frac{n\pi x}{L}\right)\left(A_n \mathrm{e}^{\frac{n\pi y}{L}} + B_n \mathrm{e}^{-\frac{n\pi y}{L}}\right) \tag{3.3.25}$$

Imposing also that $u_n(x,M) = 0$ gives

$$u_n(x,y) = \sin\left(\frac{n\pi x}{L}\right)\sinh\left(\frac{n\pi(M-y)}{L}\right) \tag{3.3.26}$$

so that

$$u(x,y) = \sum_{n\geqslant 1} A_n \sin\left(\frac{n\pi x}{L}\right)\sinh\left(\frac{n\pi(M-y)}{L}\right) \tag{3.3.27}$$

Finally, imposing $u(x,0) = f(x)$ yields:

$$u(x,y) = \sum_{n\geqslant 1} \frac{a_n}{\sinh\left(\frac{n\pi M}{L}\right)} \sin\left(\frac{n\pi x}{L}\right)\sinh\left(\frac{n\pi(M-y)}{L}\right) \tag{3.3.28}$$

Solution of the full problem. Therefore we have the solution of the full problem as follows:

$$u(x,y) = \sum_{n\geqslant 1} \frac{a_n}{\sinh\left(\frac{n\pi M}{L}\right)} \sin\left(\frac{n\pi x}{L}\right)\sinh\left(\frac{n\pi(M-y)}{L}\right) + \tag{3.3.29}$$

$$+ \sum_{n\geqslant 1} \frac{b_n}{\sinh\left(\frac{n\pi M}{L}\right)} \sin\left(\frac{n\pi x}{L}\right)\sinh\left(\frac{n\pi y}{L}\right) + \tag{3.3.30}$$

$$+ \sum_{n\geqslant 1} \frac{c_n}{\sinh\left(\frac{n\pi L}{M}\right)} \sin\left(\frac{n\pi y}{M}\right)\sinh\left(\frac{n\pi(L-x)}{M}\right) + \tag{3.3.31}$$

$$+ \sum_{n\geqslant 1} \frac{d_n}{\sinh\left(\frac{n\pi L}{M}\right)} \sin\left(\frac{n\pi y}{M}\right)\sinh\left(\frac{n\pi x}{M}\right) \tag{3.3.32}$$

3.3.3 Poisson equation

The Poisson equation is the non-homogeneous version of the Laplace equation

$$\triangle u(x) = g(x) \tag{3.3.33}$$

possibly subject to some boundary conditions.

Note that if $\Omega = \mathbb{R}^d$ typically one requires g to be either compactly supported or decaying at infinity. Uniqueness of a solution is then based on the following Lemma left as exercise

Lemma 3.17 *Let $u \in \mathcal{C}^2(\mathbb{R}^d)$ be harmonic. If $\lim_{|x|\to\infty} u(\vec{x}) = 0$ then u vanishes identically.*

The Lemma 3.17 allows to replace the Dirichlet conditions on a finite domain with an "asymptotic" Dirichlet condition.

The solution can be found according to the general philosophy of finding a particular solution of the non-homogeneous equation and then adding a suitable solution of the homogeneous equation that also takes care of the boundary conditions.

We start with the Lemma

Lemma 3.18 *The functions*

$$G_1(x) = \frac{1}{2}|x|, \qquad x \in \mathbb{R}^1 \tag{3.3.34}$$

$$G_2(\vec{x}) = \frac{1}{2\pi}\ln(|\vec{x}|), \qquad \vec{x} \in \mathbb{R}^2 \backslash \{\vec{0}\} \tag{3.3.35}$$

$$G_d(\vec{x}) = -\frac{\Gamma(\frac{d}{2}-1)}{4\pi^{\frac{d}{2}}|\vec{x}|^{d-2}}, \qquad \vec{x} \in \mathbb{R}^d \backslash \{\vec{0}\},\ d \geqslant 3 \tag{3.3.36}$$

are all harmonic in $\mathbb{R}^d \backslash \{\vec{0}\}$. Here the multiplicative constants are chosen for later convenience.

Observe that all G_d's are functions only of the distance from the origin; furthermore the formula for G_d gives the same result for $d = 1$. For $d = 2$ the function G_2 is the limit

$$G_2(r) = \lim_{d=2}\left(G_d(r) + \frac{1}{2\pi(d-2)} + \frac{\gamma + \ln(\pi)}{4\pi}\right) \tag{3.3.37}$$

where $\gamma \simeq 0.5772...$ here is the Euler–Mascheroni constant (this is an example of *renormalization*).

Exercise 3.19 *Prove Lemma 3.18.*

Definition 3 *The functions G_d are called "Green functions" for the Laplace operator in d–dimensions.*

Remark 3.20 *For the readers who know what the Dirac delta distribution is, we can say that the Green's functions of the Laplace operator are function that satisfy the following equation in the distributional sense (which is precisely what we prove below):*

$$\triangle_y G_d(y-x) = \delta_x^d(y) \tag{3.3.38}$$

where $\delta_x^d(y)$ denotes the Dirac distribution in the variable y in d–dimensions supported at $y = x$.

Remark 3.21 (Connection with Maxwell's equations of electromagnetism) *Maxwell's equations are a set of PDEs for two 3–dimensional vector-fields $\mathbb{E}(\vec{x},t), \mathbb{B}(\vec{x},t)$. (electric/magnetic fields). They read:*

$$\operatorname{div} \mathbb{E} = \frac{\rho(\vec{x},t)}{\epsilon_0} \tag{3.3.39}$$

$$\operatorname{div} \mathbb{B} = 0 \tag{3.3.40}$$

$$\mathrm{curl}\mathbb{E} = -\frac{\partial \mathbb{B}}{\partial t} \tag{3.3.41}$$

$$\mathrm{curl}\mathbb{B} = \mu_0 \left(\epsilon_0 \frac{\partial \mathbb{E}}{\partial t} + \mathbb{J}(x,t)\right) \tag{3.3.42}$$

where ρ is the density of charge per unit volume, $\mathbb{J}$ is the electric current, ϵ_0 is the permittivity of space (dielectric constant) and μ_0 the permeability of space (magnetic constant).

If the sources $\rho, \mathbb{J}$ are independent of time and we seek for static solutions (independent of time) we see that $\mathrm{curl}\, \mathbb{E} = 0$ and hence (in $\mathbb{R}^3$) we can write $\mathbb{E} = -\nabla V$ (the sign is conventional), where V is the electrostatic potential.

Thus the potential solves the Poisson equation $\triangle V = -\frac{\rho(\vec{x})}{\epsilon_0}$.

You may also notice that $G_3(\vec{x})$ is (up to a suitable constant) the Coulomb potential for an isolated point-like charge placed at the origin.

For these reasons, the study of the Laplace/Poisson equation is usually part of the branch of mathematics called potential theory.

Proposition 3.22 *Let $g(\vec{x})$ be $\mathcal{C}_0^1(\mathbb{R}^d)$. Then*

$$u(\vec{x}) := \int_{\mathbb{R}^d} G_d(\vec{y}-\vec{x})g(\vec{y})\mathrm{d}V(\vec{y}) \tag{3.3.43}$$

is a solution of the Poisson equation $\triangle u(\vec{x}) = g(\vec{x})$.

Proof. We sketch the proof (we discount some analytical details for simplicity).

First of all we observe that the integral is well defined; this is seen by passing to polar coordinates centered at x and writing $\vec{y} = \vec{x} + \rho\mathbf{n}$, $\mathrm{d}V(y) = \rho^{d-1}\mathrm{d}\rho\mathrm{d}S(\mathbf{n})$. One can also see that it is possible to differentiate G_d with respect to $\vec{x}$ **once** and still have a convergent integral.

With that in mind we can integrate by parts

$$\nabla_x u(x) = \nabla_x \int_{\mathbb{R}^d} G_d(\vec{y}-\vec{x})g(\vec{y})\mathrm{d}V(\vec{y}) = -\int_{\mathbb{R}^d} \nabla_y G_d(\vec{y}-\vec{x})g(\vec{y})\mathrm{d}V(\vec{y}) = \\ = \int_{\mathbb{R}^d} G_d(\vec{y}-\vec{x})\nabla_y g(\vec{y})\mathrm{d}V(\vec{y}) \tag{3.3.44}$$

Now we can compute the divergence:

$$\triangle u = \int_{\mathbb{R}^d} \nabla_x G_d(\vec{y}-\vec{x})\nabla_y g(\vec{y})\mathrm{d}V(\vec{y}) = -\int_{\mathbb{R}^d} \nabla_y G_d(\vec{y}-\vec{x})\nabla_y g(\vec{y})\mathrm{d}V(\vec{y}) \tag{3.3.45}$$

Now the integral can be split into $\mathbb{R}^d \backslash B_\epsilon(\vec{x})$ and $B_\epsilon(\vec{x})$; since the value is independent of ϵ, we are allowed to take the limit as $\epsilon \to 0^+$:

$$-\int_{\mathbb{R}^d} \nabla_y G_d(\vec{y}-\vec{x})\nabla_y g(\vec{y})\mathrm{d}V(\vec{y}) = \\ = \lim_{\epsilon\to 0^+}\left(\int_{\mathbb{R}^d\backslash B_\epsilon(x)} + \int_{B_\epsilon(\vec{x})}\right) \nabla_y G_d(\vec{y}-\vec{x})\nabla_y g(\vec{y})\mathrm{d}V(\vec{y}) \tag{3.3.46}$$

Since the integrand is integrable, the second limit tends to zero and we reach the conclusion that

$$\triangle u(\vec{x}) = -\lim_{\epsilon\to 0^+}\int_{\mathbb{R}^d\backslash B_\epsilon(x)} \nabla_y G_d(\vec{y}-\vec{x})\nabla_y g(\vec{y})\mathrm{d}V(\vec{y}) \tag{3.3.47}$$

Applying Thm. 3.3 again to the first integral and keeping in mind that $\triangle_y G_d(y-x) = 0$ for $y \in \mathbb{R}^d \backslash B_\epsilon(x)$, we get

$$\triangle u(x) = -\lim_{\epsilon\to 0^+}\int_{\partial B_\epsilon(x)} \nabla_{\mathbf{n}} G_d(\vec{y}-\vec{x})\mathfrak{g}(\vec{y})\mathrm{d}S(\vec{y}) \tag{3.3.48}$$

The normal $\mathbf{n}$ is the normal pointing *towards* x (the outer normal of $\mathbb{R}^d\backslash B_\epsilon(x)$) and the gradient is with respect to y

$$-\nabla_{\mathbf{n}} G(\vec{y}-\vec{x})|_{|\vec{y}-\vec{x}|=\epsilon|} = \begin{cases} \partial_r G_d|_{r=\epsilon} = \dfrac{(d-2)\Gamma(\frac{d}{2}-1)}{4\pi^{\frac{d}{2}}\epsilon^{d-1}} = \dfrac{1}{\alpha_{d-1}\epsilon^{d-1}} & d \geqslant 3 \\ \partial_r G_d|_{r=\epsilon} = \dfrac{1}{2\pi\epsilon} & d = 2 \\ \partial_r G_d|_{r=\epsilon} = \dfrac{1}{2} & d = 1 \end{cases} \tag{3.3.49}$$

In all cases $d = 1, 2, \ldots$ we have

$$\triangle u(x) = \lim_{\epsilon\to 0^+}\frac{1}{\alpha_{d-1}\epsilon^{d-1}}\int_{\partial B_\epsilon(x)} g(y)\mathrm{d}S(y) \tag{3.3.50}$$

Since $g(y)$ is continuous at $y = x$, its average on the surface of the ϵ–sphere at x tends to $g(x)$ as ϵ tends to zero. ■

Green's functions for domains

The Green functions presented in Lemma 3.18 allow to solve the Poisson equation on $\mathbb{R}^d$ (i.e. unbounded domains). For domains Ω with boundary the corresponding Green's function is, using the distributional notation,

$$\triangle_y G_\Omega(y,x) = \delta_x(y) \qquad G_\Omega(y,x)\Big|_{y\in\partial\Omega} = 0 \tag{3.3.51}$$

for the Dirichlet problem (analogous formulation for the Neumann problem). In other words, they allow to solve the Poisson equation

$$\triangle u = g; \qquad u\Big|_{\partial\Omega} = 0, \tag{3.3.52}$$

for $g \in \mathcal{C}_0^\infty(\Omega)$. In general these Green functions are not invariant under translations.

While the general theory is beyond the scope of the present course, we present here a simple example of the Green functions of special domains in $\mathbb{R}^2$.

It is convenient to identify $\mathbb{R}^2 \simeq \mathbb{C}$ and write a point in complex notation (see also next section)

$$z = x + iy. \tag{3.3.53}$$

Definition 4 *Given a domain $\Omega \subset \mathbb{C} \simeq \mathbb{R}^2$ the* **Green's function** *$G_\Omega(z;w)$ is a function defined for $w \in \Omega$, $z \in \Omega\backslash\{w\}$ satisfying the following properties:*

1. *$G_\Omega(z;w)$ is harmonic with respect to z in $\Omega\backslash\{w\}$;*
2. *$G_\Omega(z;w) - \frac{1}{2\pi}\ln|z-w|$ extends to a harmonic function with respect to z in the whole Ω.*
3. *$G_\Omega(z;w)$ extends to a continuous function for $z \in \overline{\Omega}\backslash\{w\}$ and $G_\Omega(z;w) = 0$ for $z \in \partial\Omega$.*

The complex conjugation geometrically represents the reflection around the real axis. Let $\mathbb{H} = \{z;\ \Im(z) > 0\}$ and define

$$G_{\mathbb{H}}(z,w) = G(z-w) - G(z-w^\star) = \frac{1}{2\pi}\ln\frac{|z-w|}{|z-w^\star|}. \tag{3.3.54}$$

Note that if $z \in \mathbb{R}$ then $G_{\mathbb{H}}(z,w) = 0$ for any $w \in \mathbb{H}$.

Proposition 3.23 *The function $G_{\mathbb{H}}$ is the Green's function of the upper half plane with Dirichlet boundary conditions; namely, for any $g \in \mathcal{C}_0^\infty(\mathbb{H})$, the solution of the Poisson-Dirichlet problem*

$$\triangle u = g \quad u(x,0) = 0 \tag{3.3.55}$$

is given by

$$u(z) = \int_{\mathbb{H}} \frac{g(w)}{2\pi} \ln \frac{|z-w|}{|z-\overline{w}|} \mathrm{d}^2 w \tag{3.3.56}$$

where $\mathrm{d}^2 w$ *denotes the Lebesgue area measure in* $\mathbb{C} \simeq \mathbb{R}^2$.

Similarly one can write the Green's function for the unit disk (or any other disk by simple arguments)

Proposition 3.24 *Let* $\mathbb{D} = \{|z| < 1\} \subset \mathbb{C} \simeq \mathbb{R}^2$ *and define*

$$G_{\mathbb{D}}(z,w) := \frac{1}{2\pi} \ln \frac{|z-w|}{|w| \left|z - \frac{1}{\overline{w}}\right|}, \qquad z, w \in \mathbb{D}. \tag{3.3.57}$$

Then $G_{\mathbb{D}}$ *is the Green's function of* $\mathbb{D}$ *with Dirichlet boundary conditions.*

3.4 Harmonic functions on the plane and complex analysis

In solving the wave equation $u_{xx} - u_{yy}$ we have factorized the wave operator into two derivatives along the characteristic directions:

$$\partial_x^2 - c^2 \partial_y^2 = (\partial_x - c\partial_y)(\partial_x + c\partial_y) \tag{3.4.1}$$

so that one easily concludes that the solutions are sums of functions of $x + cy$ and $x - cy$.

On a formalistic level we may assume that $c = i = \sqrt{-1}$ and proceed in the same way:

$$\partial_x^2 + \partial_y^2 = (\partial_x - i\partial_y)(\partial_x + i\partial_y) \tag{3.4.2}$$

This (heuristic) observation ushers the methods of complex analysis into the study of the Laplace equation in two-dimensions. In a certain sense, as we shall see momentarily, this is also correct.

First of all we identify $\mathbb{R}^2 \simeq \mathbb{C}^2$ via the obvious map $(x,y) \mapsto z = x + iy$. Then we recall that a differentiable complex valued function $f(x,y) = u(x,y) + iv(x,y)$ on a domain in $\mathbb{R}^2$ is called *holomorphic* if it satisfies the following system of *Cauchy – Riemann equations*

$$\left.\begin{aligned} \tfrac{\partial u}{\partial x} - \tfrac{\partial v}{\partial y} &= 0 \\ \tfrac{\partial v}{\partial x} + \tfrac{\partial u}{\partial y} &= 0 \end{aligned}\right\} \tag{3.4.3}$$

or, in complex form

$$\frac{\partial f}{\partial x} + i \frac{\partial f}{\partial y} = 0. \tag{3.4.4}$$

Introducing complex combinations of the Euclidean coordinates

$$z = x + iy \qquad \bar{z} = x - iy$$

we can also introduce the following two vector fields.

$$\frac{\partial}{\partial z} = \frac{1}{2}\left(\frac{\partial}{\partial x} - i\,\frac{\partial}{\partial y}\right) \qquad \frac{\partial}{\partial \overline{z}} = \frac{1}{2}\left(\frac{\partial}{\partial x} + i\,\frac{\partial}{\partial y}\right). \tag{3.4.5}$$

Note that, by construction

$$\frac{\partial}{\partial z}z = 1 = \frac{\partial}{\partial \overline{z}}\overline{z}, \qquad \frac{\partial}{\partial z}\overline{z} = 0 = \frac{\partial}{\partial z}\overline{z}. \tag{3.4.6}$$

so that, in a certain sense, we can view z and $\overline{z}$ as independent coordinates.

With the aid of the vectors (3.4.5) the Cauchy – Riemann equations can be rewritten in the form

$$\frac{\partial f}{\partial \bar{z}} = 0. \tag{3.4.7}$$

Example 3.25 Let $f(x, y)$ be a polynomial

$$f(x, y) = \sum_{k,l} a_{kl} x^k y^l.$$

It is a holomorphic function *iff*, after the substitution

$$x = \frac{z + \bar{z}}{2}$$
$$y = \frac{z - \bar{z}}{2i}$$

there will be no dependence on $\bar{z}$:

$$\sum_{k,l} a_{kl} \left(\frac{z + \bar{z}}{2}\right)^k \left(\frac{z - \bar{z}}{2i}\right)^l = \sum_m c_m z^m.$$

In that case the result will be a polynomial in z. For example a quadratic polynomial

$$f(x, y) = ax^2 + 2bxy + cy^2$$

is holomorphic *iff* $a + c = 0$ and $b = \frac{i}{2}(a - c)$.

More generally holomorphic functions are denoted $f = f(z)$. The partial derivative $\partial/\partial z$ of a holomorphic function is denoted df/dz or $f'(z)$. One can also define *antiholomorphic* functions $f = f(\bar{z})$ satisfying equation

$$\frac{\partial f}{\partial z} = 0. \tag{3.4.8}$$

Notice that the complex conjugate $\overline{f(z)}$ to a holomorphic function is an antiholomorphic function.

From complex analysis it is known that any function f holomorphic on a neighborhood of a point z_0 is also a *complex analytic* function, i.e., it can be represented as a sum of a power series

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n \tag{3.4.9}$$

convergent uniformly and absolutely for sufficiently small $|z - z_0|$. In particular it is continuously differentiable any number of times. Its real and imaginary parts $u(x, y)$ and $v(x, y)$ are infinitely smooth functions of x and y.

Theorem 3.26 *The real and imaginary parts of a function holomorphic in a domain Ω are harmonic functions on the same domain.*

Proof: Differentiating the first equation in (3.4.3) in x and the second one in y and adding we obtain

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0.$$

Similarly, differentiating the second equation in x and subtracting the first one differentiated in y gives

$$\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} = 0.$$

■

Corollary 3.27 *For any integer $n \geqslant 1$ the functions*

$$\operatorname{Re} z^n \quad \text{and} \quad \operatorname{Im} z^n \tag{3.4.10}$$

are polynomial solutions to the Laplace equation.

Polynomial solutions to the Laplace equation are called *harmonic polynomials*. We obtain a sequence of harmonic polynomials

$$x, \; y, \; x^2 - y^2, \; xy, \; x^3 - 3xy^2, \; 3x^2 y - y^3, \dots .$$

Observe that the harmonic polynomials of degree n can be represented in the polar coordinates r, ϕ as

$$\operatorname{Re} z^n = r^n \cos n\phi, \quad \operatorname{Im} z^n = r^n \sin n\phi.$$

These are exactly the same functions we used to solve the main boundary value problems for the circle.

Exercise 3.28 *Prove that the Laplace operator*

$$\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$$

in the coordinates z, $\bar z$ becomes

$$\Delta = 4 \frac{\partial^2}{\partial z \partial \bar z}. \tag{3.4.11}$$

To a certain extent the converse of Theorem 3.26 holds as well

Theorem 3.29 *Let $u(x,y)$ be a harmonic function in a* **simply connected** *domain $\Omega \subset \mathbb{R}^2 \simeq \mathbb{C}$. Then there is a holomorphic function $f(z)$ such that $u = \Re f$. The function $v = \Im f$ is called the* **harmonic conjugate function** *to u.*

Proof. Consider the total differential of u,

$$\mathrm{d}u = u_x \mathrm{d}x + u_y \mathrm{d}y \tag{3.4.12}$$

This is clearly an exact form; now consider the "Hodge dual"

$$\star \mathrm{d}u := -u_y \mathrm{d}x + u_x \mathrm{d}y. \tag{3.4.13}$$

Due to the fact that u is harmonic, this form is also closed: $u_{yy} = -u_{xx}$. Now, since Ω is simply connected, we can define

$$v(x,y) := \int_{(x_0,y_0)}^{(x,y)} \big(-u_y \mathrm{d}x + u_x \mathrm{d}y\big) \quad \Rightarrow \quad \mathrm{d}v = \star \mathrm{d}u. \tag{3.4.14}$$

and the integration is independent of the path (here (x_0, y_0) is some choice of point in Ω).

By construction we have $v_x = -u_y$ and $v_y = u_x$; namely the function $f(x,y) = u(x,y) + iv(x,y)$ satisfies the Cauchy–Riemann equations in the domain Ω and hence it is holomorphic. ■

Remark 3.30 If we lift the assumption that Ω is simply connected, then we can only assert the local existence of v, but in general f will not be single valued. The prototypical example is $u(x,y) = \ln(\sqrt{x^2+y^2})$ on $\mathbb{C}\backslash\{0\}$. In this case $f(z)$ is the complex logarithm, which is not single–valued.

Using the representation (3.4.11) of the two-dimensional Laplace operator one can describe all complex valued solutions to the Laplace equation.

Theorem 3.31 *Any complex valued solution u to the Laplace equation $\Delta u = 0$ on the plane can be represented as a sum of a holomorphic and an antiholomorphic function:*

$$u(x,y) = f(z) + g(\bar{z}). \tag{3.4.15}$$

Proof: Let the $\mathcal{C}^2$-smooth function $u(x,y)$ satisfy the Laplace equation

$$\frac{\partial^2 u}{\partial z \partial \bar{z}} = 0.$$

Denote

$$F = \frac{\partial u}{\partial z}.$$

The Laplace equation implies that this function is holomorphic, $F = F(z)$. From complex analysis it is known that any holomorphic function admits a holomorphic primitive,

$$F(z) = f'(z).$$

Consider the difference $g := u - f$. It is an antiholomorphic function, $g = g(\bar{z})$. Indeed,

$$\frac{\partial g}{\partial z} = \frac{\partial u}{\partial z} - f' = 0.$$

So $u = f(z) + g(\bar{z})$. ■

Corollary 3.32 *Any harmonic function on the plane can be represented as the real part of a holomorphic function.*

Notice that the imaginary part of a holomorphic function $f(z)$ is equal to the real part of the function $-i\, f(z)$ that is holomorphic as well.

Corollary 3.33 *Any harmonic function on the plane is $\mathcal{C}^\infty$-smooth.*

Another important consequence of the complex representation (3.4.11) of the Laplace operator on the plane is invariance of the Laplace equation under conformal transformation. Recall that a smooth map

$$f : \Omega \to \Omega'$$

is called *conformal* if it preserves the angles between smooth curves. The dilatations

$$(x, y) \mapsto (k\, x, k\, y)$$

with $k \neq 0$, rotations by the angle ϕ

$$(x, y) \mapsto (x \cos\phi - y \sin\phi,\ x \sin\phi + y \sin\phi)$$

and reflections

$$(x, y) \mapsto (x, -y)$$

are examples of linear conformal transformations. These examples and their superpositions exhaust the class of linear conformal maps. The general description of conformal maps on the plane are given by

Lemma 3.34 *Let $f(z)$ be a function holomorphic in the domain Ω with never vanishing derivative:*

$$\frac{df(z)}{dz} \neq 0 \quad \forall z \in \Omega.$$

Then the map

$$z \mapsto f(z)$$

of the domain Ω to $\Omega' = f(\Omega$ is conformal. Same for antiholomorphic functions. Conversely, if the smooth map $(x, y) \mapsto (u(x, y), v(x, y))$ is conformal then the function $f = u + iv$ is holomorphic or antiholomorphic with nonvanishing derivative.

Proof: Let us consider the differential of the map $(x, y) \mapsto (u(x, y), v(x, y))$ given by the real $u = \operatorname{Re} f$ and imaginary $v = \operatorname{Im} f$ parts of the holomorphic function f. It is a linear map defined by the Jacobi matrix

$$\begin{pmatrix} \partial u/\partial x & \partial u/\partial y \\ \\ \partial v/\partial x & \partial v/\partial y \end{pmatrix} = \begin{pmatrix} \partial u/\partial x & -\partial v/\partial x \\ \\ \partial v/\partial x & \partial u/\partial x \end{pmatrix}$$

(we have used the Cauchy – Riemann equations). Since

$$0 \neq |f'(z)|^2 = \left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial v}{\partial x}\right)^2,$$

we can introduce the numbers $r > 0$ and ϕ by

$$r = |f'(z)|, \quad \cos\phi = \frac{\partial u/\partial x}{\sqrt{\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial v}{\partial x}\right)^2}}, \quad \sin\phi = \frac{\partial v/\partial x}{\sqrt{\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial v}{\partial x}\right)^2}}.$$

The Jacobi matrix then becomes a composition of the rotation by the angle ϕ and a dilatation with the coefficient r:

$$\begin{pmatrix} \partial u/\partial x & \partial u/\partial y \\ \\ \partial v/\partial x & \partial v/\partial y \end{pmatrix} = r \begin{pmatrix} \cos\phi & -\sin\phi \\ \\ \sin\phi & \cos\phi \end{pmatrix}.$$

This is a linear conformal transformation preserving the angles. A similar computation works for an antiholomorphic map with nonvanishing derivatives $f'(\bar{z}) \neq 0$.

Conversely, the Jacobi matrix of a conformal transformation must have the form

$$r \begin{pmatrix} \cos\phi & -\sin\phi \\ \\ \sin\phi & \cos\phi \end{pmatrix}$$

or

$$r \begin{pmatrix} \cos\phi & \sin\phi \\ \\ \sin\phi & -\cos\phi \end{pmatrix}.$$

In the first case one obtains the differential of a holomorphic map while the second matrix corresponds to the antiholomorphic map. ■

We are ready to prove

Theorem 3.35 *Let*

$$f : \Omega \to \Omega'$$

be a conformal map. Then the pull-back (composition with f) of any function harmonic in Ω' will be harmonic in Ω.

Proof: According to the Lemma the conformal map is given by a holomorphic or an antiholomorphic function. Let us consider the holomorphic case,

$$z \mapsto w = f(z).$$

The transformation law of the Laplace operator under such a map is clear from the following formula:

$$\frac{\partial^2}{\partial z \partial \bar{z}} = |f'(z)|^2 \frac{\partial^2}{\partial w \partial \bar{w}}. \tag{3.4.16}$$

Thus any function U on Ω' satisfying

$$\frac{\partial^2 U}{\partial w \partial \bar{w}} = 0$$

will also satisfy

$$\frac{\partial^2 U}{\partial z \partial \bar{z}} = 0.$$

The case of an antiholomorphic map can be considered in a similar way. ■

A conformal map

$$f : \Omega \to \Omega'$$

is called *conformal transformation* if it is one-to-one. In that case the inverse map

$$f^{-1} : \Omega' \to \Omega$$

exists and is also conformal. The following fundamental *Riemann theorem* is the central result of the theory of conformal transformations on the plane.

Theorem 3.36 (Riemann uniformization theorem) *For any connected and simply connected domain Ω on the plane not coinciding with the plane itself there exists a conformal transformation of Ω to the unit circle $f : \Omega \to \mathbb{D}$.*

There is an interesting application of the Riemann Uniformization Theorem.

Theorem 3.37 *For an arbitrary simply connected domain Ω not coinciding with the plane, the Green's function for the Dirichlet problem is given by*

$$G_\Omega(z, w) = \frac{1}{2\pi} \ln \left| \frac{f(z) - f(w)}{1 - f(z)\overline{f(w)}} \right| \tag{3.4.17}$$

The Riemann theorem, together with conformal invariance of the Laplace equation gives a possibility to reduce the main boundary value problems for any connected simply connected domain to similar problems for the unit circle.

The proof of Riemann's theorem belongs to an advanced course in complex analysis and will not be reported here.

3.4.1 Conformal maps in fluid-dynamics

Suppose $\vec{V}(x,y,z,t)$ is a vector field representing the motion of a fluid.

The fluid is called **incompressible** if $\operatorname{div}\vec{V} \equiv 0$.

Suppose that $\vec{V}$ represents a stationary flow (i.e. independent of time) and bi-dimensional (i.e. independent of z and with zero z–component). Denote by $\mathbf{v}(x,y)$ the projection of $\vec{V}$ to the x,y components.

Proposition 3.38 *If* $\mathbf{v}$ *is incompressible and irrotational, then there are two functions* Φ, Ω *such that*

$$\mathbf{v} = \Phi_x \mathbf{i} + \Phi_y \mathbf{j} = \Omega_y \mathbf{i} - \Omega_x \mathbf{j}. \tag{3.4.18}$$

These two functions are called the **velocity potential** *and the* **stream function** *respectively.*

The naming stems from the observation that the level-sets of Ω are stream lines.

A simple consequence of the above proposition is that the function

$$F(z) = \Phi(x,y) + i\Omega(x,y) \tag{3.4.19}$$

is analytic. It is called the **complex velocity potential**. Since $F'(z) = \Phi_x + i\Omega_x = v_1 - iv_2$ we see that $|F'(x)|$ is the speed of the flow. I.e. the complex conjugate of $F'(z)$ represents the velocity field $\mathbf{v}$ interpreted as a complex number.

Suppose that $\mathscr{D}$ is a region in $\mathbb{C}$ and we imagine the stationary 2-dimensional motion of an incompressible and irrotational fluid around it. Assuming that the fluid goes around the region (an *obstacle*) then the flow-lines must tightly surround $\mathscr{D}$. Thus we can use the Riemann mapping theorem by mapping the *complement* of $\mathscr{D}$ to the upper half plane. If $F(z)$ is said map, then it provides the velocity potential and stream function. This is used for example for some early models of airfoils (follow for example this link: Joukowsky map).

3.4.2 Some examples of mappings of regions

We list here some conformal mappings of regions. We start by observing that the upper half-plane $\mathbb{H}$ and the unit disk $\mathbb{D}$ are conformally equivalent;

Lemma 3.39 *The map*

$$w = f(z) := \frac{z-i}{z+i} : \mathbb{H} \to \mathbb{D} \tag{3.4.20}$$

is a conformal map. The inverse is

$$z = g(w) := i\frac{w+1}{1-w} : \mathbb{D} \to \mathbb{H} \tag{3.4.21}$$

In view of the above lemma, many conformal maps bring a domain Ω to $\mathbb{H}$ instead of $\mathbb{D}$.

Example 3.40 *The semiinfinite strip*

$$\Omega := \left\{ z : \quad \Re z > 0, \quad \Im z \in [0, \pi] \right\} \tag{3.4.22}$$

is mapped to $\mathbb{H}$ *by*

$$f(z) = \cosh(z) : \Omega \to \mathbb{H} \tag{3.4.23}$$

The source of many examples is the *Schwarz–Christoffel* formula

Theorem 3.41 *Let* $\alpha_j \in (0, 2\pi]$, $j = 1, \ldots, n$ *such that* $\sum_{j=1}^{n} \alpha j = (n-2)\pi$ *and* $a_1 < a_2 < \ldots < a_{n-1}$. *Let* $f(\zeta)$ *be the function*

$$f(\zeta) = \int^{\zeta} \frac{\mathrm{d}w}{\prod_{j=1}^{n} (w - a_j)^{1 - \frac{\alpha_j}{\pi}}} \tag{3.4.24}$$

Then this function maps conformally the upper half–plane $\mathbb{H}$ *to the interior of a polygon with interior angles* α_j.

The proof (whose detail we skip) consists in verifying that the the argument of $f'(\zeta)$ as ζ traverses the real axis (i.e. the tangent vector to the boundary of the region) is constant and with jumps as $\zeta = a_j$ given precisely by the angles α_j. The condition on the sum allows to conclude, by the argument principle, that the function $f(\zeta)$ is univalent (injective) in the upper half plane.

For example in the case of the semi-infinte strip one takes $a_1 = -1, a_2 = 1$ and $\alpha_1 = \alpha_2 = \frac{\pi}{2}$. Then

$$\int \frac{\mathrm{d}\zeta}{(\zeta - 1)^{\frac{1}{2}} (\zeta + 1)^{\frac{1}{2}}} \propto \operatorname{arccosh}(\zeta) \tag{3.4.25}$$

(the proportionality constant and the additive constant allow to translate/rotate/dilate the resulting polygon).

3.5 Exercises for Chapter 3

Exercise 3.42 *Prove Lemma 3.17.*

Exercise 3.43 *Let Ω be a bounded open set in $\mathbb{R}^d$ with p.w. smooth boundary. Let $u_1, u_2, u_3 \in \mathcal{C}^2(\Omega) \cap \mathcal{C}^0(\overline{\Omega})$ be harmonic functions. Show that if $u_1 \leqslant u_2 \leqslant u_3$ when restricted to the boundary of Ω then the same inequality holds throughout Ω.*

Exercise 3.44 *Let Ω be the semi-infinite strip $(0, \pi) \times (0, \infty)$. Consider the Laplace problem on Ω with B.C.*

$$u(x, 0) = 0, \qquad u(0, y) = 0 \quad x \in [0, \pi], \quad y \in [0, \infty) \tag{3.5.1}$$

Find more than one harmonic solution to this problem. Explain how that does not contradict the uniqueness theorem for harmonic functions.

Solution. One solution is the zero solution; another solution is $u(x, y) = xy$. There at least two reasons the solution is not unique. First we did not specify boundary values on the whole boundary. Second the domain is unbounded.

Even if the domain were the quarter plane $[0, \infty) \times [0, \infty)$ with zero B.C., we would still have $u(x, y) = xy$ as harmonic function (and the trivial solution). ■

Exercise 3.45 *Prove Proposition 3.24.*

Exercise 3.46 *Prove Proposition 3.23.*

Exercise 3.47 *Prove that any harmonic polynomial is a linear combination of the polynomials (3.4.10). [Hint: if p is a harmonic polynomial, then solve the Laplace equation on the disk with BC $u|_{|z|=1} = p|_{|z|=1}$.]*

Exercise 3.48 *Find a Green function for the upper half plane $\mathbb{H}$ but with Neumann conditions on $\mathbb{R}$.*

Exercise 3.49 *Suppose that we have the Poisson equation $\triangle u = g$ where g is only an L^1 function. Show that*

$$u(x) = \int_{\mathbb{R}^d} G_d(y - x) g(y) \mathrm{d}V(y) \tag{3.5.2}$$

is a solution in the weak sense, namely show that for all $\varphi \in \mathcal{C}_0^\infty(\mathbb{R}^d)$ we have

$$\int_{\mathbb{R}^d} u \triangle \varphi \, \mathrm{d}V(x) = \int_{\mathbb{R}^d} g \varphi \, \mathrm{d}V(x). \tag{3.5.3}$$

Exercise 3.50 *Let $\Omega \subset \mathbb{C} \simeq \mathbb{R}^2$ be an open domain and $a \in \Omega$. Recalling Def. 4 prove that*

1. *If a bounded open set Ω admits a Green's function then $G_\Omega(z; w)$ is unique.*

2. With Ω as above plus connected, prove that $G_\Omega(z;w)$ is negative on $\Omega\backslash\{a\}$.

Exercise 3.51 *Find a function $u(x,y)$ satisfying*

$$\Delta u = x^2 - y^2$$

for $r < a$ and the boundary condition $u|_{r=a} = 0$.

Solution. We see that $x^2 - y^2 = \Re(z^2) = \frac{z^2+\overline{z}^2}{2}$. Since $\partial_z \overline{z} = 0$ we can "integrate" independently to find a particular solution; recall $4\partial_z\partial_{\overline{z}} = \triangle$ so that we compute

$$\partial_z 4u_p(z,\overline{z}) = \int \mathrm{d}\overline{z}\frac{z^2+\overline{z}^2}{2} = \frac{z^2\overline{z}}{2} + \frac{\overline{z}^3}{6} \quad \Rightarrow \quad 4u_p(z,\overline{z}) = \frac{z^3\overline{z} + \overline{z}^3 z}{6} = \frac{|z|^2}{6}(z^2+\overline{z}^2) \tag{3.5.4}$$

On the boundary of the disk a we have $u_p = \frac{a^2}{24}(z+\overline{z})\Big|_{|z|=a}$ which is by inspection the restriction of a harmonic polynomial. In conclusion we find:

$$u = \frac{|z|^2 - a^2}{24}(z^2+\overline{z}^2). \tag{3.5.5}$$

■

Exercise 3.52 *Let $\chi_{\mathbb{D}}(z)$ be the characteristic function of the unit disk $\mathbb{D} := \{z \in \mathbb{C} : \quad |z| < 1\}$. Compute*

$$U(z) := \int_{\mathbb{C}} \chi_{\mathbb{D}}(w)\frac{\ln|z-w|}{2\pi}\mathrm{d}^2 w. \tag{3.5.6}$$

Solution The function $U(z)$ is clearly harmonic outside of the unit disk. At infinity it behaves like $\frac{\ln|z|}{2\pi}\mathcal{A}(\mathbb{D}) = \frac{1}{2}\ln|z|$.

Inside the disk it satisfies $\triangle U = 1$ and hence it must be of the form $U(z) = \frac{|z|^2}{4}+$ harmonic.

The function is also easily seen to be continuous across the boundary; thus we start by subtracting the solution of a Dirichlet problem $\triangle V = 0, \quad V|_{|\partial\mathbb{D}} = 1$. This is easily seen to be the constant 1.

$$U(z) = \frac{|z|^2-1}{4}\chi_{\mathbb{D}} + \frac{1}{2}\ln|z|\,\chi_{\mathbb{C}\backslash\mathbb{D}}. \tag{3.5.7}$$

■

Exercise 3.53 *Let $G_\Omega(z;w)$ as in Theorem 3.37. Let $\varphi(x,y) \in \mathcal{C}_0^\infty(\Omega)$. Show, directly, that $\int_\Omega G_\Omega(z;w)\triangle\varphi(w)\mathrm{d}^2 w = \varphi(z)$ using the fact that G_Ω is the pullback of the Green's function of the unit disk via the uniformizing map $f : \Omega \to \mathbb{D}$.pp*

Solution. Let $\zeta = F(z) : \Omega \to \mathbb{D}$ be the Riemann uniformization map to the unit disk

$$4\partial_z\partial_{\overline{z}}\int_\Omega G_\Omega(z;w)\triangle\varphi(w)\mathrm{d}^2w = 4\partial_z\partial_{\overline{z}}\int_\Omega \frac{1}{2\pi}\ln\left|\frac{F(z)-F(w)}{1-F(z)\overline{F(w)}}\right|\triangle\varphi(w)\mathrm{d}^2w \tag{3.5.8}$$

Denote $\zeta = F(z)$ and $\xi = F(w)$; then the above reads

$$4\left|\frac{\mathrm{d}\zeta}{\mathrm{d}z}\right|^2\partial_\zeta\partial_{\overline{\zeta}}\int_\Omega \frac{1}{2\pi}\ln\left|\frac{\zeta-\xi}{1-\zeta\overline{\xi}}\right|\triangle\varphi(w)\frac{\mathrm{d}^2\xi}{|\mathrm{d}\xi/\mathrm{d}w|^2} \overset{\text{why?}}{=} 4\left|\frac{\mathrm{d}\zeta}{\mathrm{d}z}\right|^2\triangle\varphi(z)\frac{1}{|\mathrm{d}\zeta/\mathrm{d}z|^2} = \varphi(z) \tag{3.5.9}$$

■

Exercise 3.54 *Find a harmonic function on the annular domain*

$$a < r < b$$

with the boundary conditions

$$u|_{r=a} = 1, \quad \left(\frac{\partial u}{\partial r}\right)_{r=b} = \cos^2\theta.$$

Solution We have

$$u = \frac{A_0}{2} + \frac{B_0}{2}\ln|z| + \sum_{n\geqslant 1} A_n\Re z^n + B_n\Im z^n + C_n\Re z^{-n} + D_n\Im z^{-n} \tag{3.5.10}$$

Note that $\cos^2(\theta) = \frac{1+\cos(2\theta)}{2}$. Imposing the BC gives the system:

$$\frac{A_0}{2} = 1 \quad\Rightarrow\quad A_0 = 2. \tag{3.5.11}$$
$$A_n = -C_n \tag{3.5.12}$$
$$B_n = -D_n \tag{3.5.13}$$

■

Exercise 3.55 *Find a harmonic function $u(x, y)$ solving the the Dirichlet b.v.p. in the rectangle*

$$0 \leqslant x \leqslant a, \quad 0 \leqslant y \leqslant b$$

satisfying the boundary conditions

$$u(0, y) = A\,y(b-y), \quad u(a, y) = 0$$
$$u(x, 0) = B\,\sin\frac{\pi x}{a}, \quad u(x, b) = 0.$$

Hint: use separation of variables in Euclidean coordinates.

Chapter 4

Heat equation

4.1 Derivation of the heat equation

The heat equation for the function $u = u(x,t)$, $x \in \mathbb{R}^d$, $t \in \mathbb{R}_{>0}$ reads

$$\frac{\partial u}{\partial t} = a^2 \Delta u. \tag{4.1.1}$$

Here Δ is the Laplace operator in $\mathbb{R}^d$. We will consider only the case of constant coefficients $a =$ const. For $d = 3$ this equation describes the distribution of temperature $u(x,t)$ in a homogeneous and isotropic medium at the time t. It is also used to describe the diffusion of the concentration of some quantity (e.g. a solute in a solution).

The derivation of heat equation is based on the following assumptions.

1. The heat Q necessary for changing from u_1 to u_2 the temperature of a portion of mass m is proportional to the mass and to the difference of temperatures:

$$Q = c_p\, m(u_2 - u_1).$$

The coefficient c_p is called *specific heat capacity.*

2. The *Fourier law* describing the quantity of heat speading in the time Δt through a surface S during the time interval Δt. It says that this quantity ΔQ is proportional to the area $A(S)$ of the surface, to the time Δt and to the derivative of the temperature u along the normal n to the surface:

$$\Delta Q = -k\, A(S) \frac{\partial u}{\partial n}\, \Delta t.$$

Here the coefficient $k > 0$ is called *thermal conductivity.* The negative sign means that the heat is spreading from hot to cold regions.

In order to derive the heat equation let us consider the heat balance within a domain $\Omega \subset \mathbb{R}^d$ with a smooth boundary $\partial\Omega$.

We remind that the divergence of $u(x,t)$ can be interpreted as the infinitesimal net *flux* of the gradient of u across the six sides (in 3 dimensions) of an infinitesimal box of sides $\mathrm{d}x, \mathrm{d}y, \mathrm{d}z$ centered at x.

Then the integral form of our first assumption says that the total exchange of heat between the domain Ω and the exterior during the time interval Δt is

$$\Delta Q = \int_{\Omega} c_p \rho(x) \left[u(t + \Delta t, x) - u(t, x) \right] \mathrm{d}V \simeq \int_{\Omega} c_p \rho(x)\, u_t(x, t)\, \mathrm{d}V \Delta t$$

where $\rho(x)$ is the mass density, such that the mass of the media contained in the volume is equal to

$$m = \int_{\Omega} \rho(x)\, \mathrm{d}V.$$

On the other hand, the Fourier law gives another expression for ΔQ in terms of the flux

$$\Delta Q = k \int_{\partial\Omega} \nabla_{\mathbf{n}} u(x, t) \mathrm{d}S \Delta t, \tag{4.1.2}$$

where the normal here is the outer one and the sign is because we are measuring the heat *acquired* by Ω (but the normal is opposite).

Equating these two expressions yields

$$\int_{\Omega} c_p \rho(x)\, u_t(x, t)\, \mathrm{d}V = k \int_{\partial\Omega} \nabla_{\mathbf{n}} u(x, t) \mathrm{d}S \tag{4.1.3}$$

The divergence theorem now can be used on the right side of (4.1.3) to yield the *balance equation*

$$0 = \int_{\Omega} \left(c_p \rho(x)\, u_t(x, t) + k \triangle u(x, t)\, \right) \mathrm{d}V \tag{4.1.4}$$

Since this identity must hold for arbitrary domains Ω we must have the equation

$$\frac{\partial u(x, t)}{\partial t} = -\frac{k}{c_p \rho(x)} \triangle u(x, t). \tag{4.1.5}$$

In fact, this equation could be derived also if the thermal conductivity k and or the specific heat capacity c_p depends on space (inhomogeneous medium).

In the case of a homogeneous media the mass density ρ, the specific heat c_p and the thermal conductivity k are constant and we reach the form (4.1.1) with a constant a

$$a^2 = \frac{k}{c_p \rho}$$

which is called *thermal conductivity* or *thermal diffusivity.*

4.2 Main boundary value problems for heat equation

The simplest is the Cauchy problem of finding a function $u(x, t)$ satisfying

$$\begin{aligned} &\frac{\partial u}{\partial t} = a^2 \Delta u \\ &u(x, 0) = \phi(x), \quad x \in \mathbb{R}^d. \end{aligned} \tag{4.2.1}$$

The physical meaning of this problem is clear: given the initial temperature distribution in the space to determine the temperature at any time $t > 0$ at any point x of the space.

Often we are interested in the temperature distribution only within the bounded domain $\Omega \subset \mathbb{R}^d$. In this case one has to add to the Cauchy data within Ω also the information about the temperature on the boundary $\partial\Omega$ or about the heat flux through the boundary. In this way we arrive at two main mixed problems in a bounded domain:

The *first mixed problem*: find a function $u(x,t)$ satisfying

$$\begin{aligned} &\frac{\partial u}{\partial t} = a^2 \Delta u, \quad t > 0, \quad x \in \Omega \\ &u(0,x) = \phi(x), \quad x \in \Omega \\ &u(x,t) = f(x,t), \quad t > 0, \quad x \in \partial\Omega. \end{aligned} \tag{4.2.2}$$

The *second mixed problem* is obtained from (4.2.2) by replacing the last condition by

$$\left(\frac{\partial u}{\partial n}\right)_{x \in \partial\Omega} = g(x,t), \quad t > 0, \quad x \in \partial\Omega. \tag{4.2.3}$$

In this equation n is the unit external normal to the boundary.

In the particular case of the boundary data independent of time

$$f = f(x) \quad \text{or} \quad g = g(x)$$

one can look for a *stationary solution* u satisfying

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

In this case the first and the second mixed problem for the heat equation reduce respectively to the Dirichlet and Neumann boundary value problem for the Laplace equation in $\mathbb{R}^d$.

4.3 Fourier transform

Our next goal is to solve the one-dimensional Cauchy problem for heat equation on the line. To this end we will develop a continuous analogue of Fourier series.

Let $f(x)$ be an absolutely integrable complex valued function on the real line, i.e.,

$$\int_{-\infty}^{\infty} |f(x)|\, dx < \infty. \tag{4.3.1}$$

Definition 4.1 *The function*

$$\hat{f}(p) := \frac{1}{2\pi} \int_{-\infty}^{\infty} f(x) e^{-ipx} dx \tag{4.3.2}$$

of the real variable p is called the Fourier transform *of $f(x)$.*

Due to the condition (4.3.1) the integral converges absolutely and uniformly with respect to $p \in \mathbb{R}$. Thus the function $\hat{f}(p)$ is continuous in p.

Example. Let us compute the Fourier transform of the Gaussian function

$$f(x) = e^{-\frac{x^2}{2}}.$$

We have

$$\int_{-\infty}^{\infty} e^{-\frac{x^2}{2}-ipx}dx = \int_{-\infty}^{\infty} e^{-\frac{1}{2}(x+ip)^2-\frac{p^2}{2}}dx$$

We want to perform a change of variables

$$s = x + ip.$$

To do this one can consider the integral

$$\oint_C e^{-\frac{z^2}{2}-\frac{1}{2}p^2}dz, \quad z = x + iy \tag{4.3.3}$$

over the boundary C of the rectangle on the complex z-plane

$$-R \leqslant x \leqslant R, \quad 0 \leqslant y \leqslant p.$$

It is easy to see that the integrals over the vertical segments $x = \pm R$, $0 \leqslant y \leqslant p$ in (4.3.3) tend to zero when $R \to \infty$. The total integral is equal to zero since the integrand is holomorphic on the entire complex plane. Hence

$$\int_{-R}^{R} e^{-\frac{1}{2}x^2-\frac{1}{2}p^2}dx + \int_{R}^{-R} e^{-\frac{1}{2}(x+ip)^2-\frac{p^2}{2}}dx \to 0 \quad \text{as} \quad R \to \infty,$$

so

$$\int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2-\frac{1}{2}p^2}dx = \int_{-\infty}^{\infty} e^{-\frac{1}{2}(x+ip)^2-\frac{p^2}{2}}dx.$$

Using the Euler integral

$$\int_{-\infty}^{\infty} e^{-\frac{x^2}{2}}dx = \sqrt{2\pi} \tag{4.3.4}$$

we finally obtain the Fourier transform of the Gaussian function

$$\hat{f}(p) = \frac{1}{\sqrt{2\pi}}e^{-\frac{p^2}{2}}. \tag{4.3.5}$$

Definition 5 (Convolution) *Let $f, g \in L^1(\mathbb{R}^n, \mathrm{d}^n h)$ and define their* **convolution**

$$(f \star g)(x) := \int_{\mathbb{R}^n} f(h)g(x-h)\mathrm{d}^n h. \tag{4.3.6}$$

We observe that the result of the convolution of two L^1 functions is still in L^1; to see this, let us denote $F(x) = (f \star g)(x)$ and observe that

$$|F(x)| \leqslant \int_{\mathbb{R}^n} |f(h)g(x-h)|\mathrm{d}^n h. \tag{4.3.7}$$

Now we integrate

$$\int_{\mathbb{R}^n} |F(x)|\mathrm{d}^n x \leqslant \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} |f(h)g(x-h)|\mathrm{d}^n h\mathrm{d}^n x \overset{\text{Fubini}}{=} \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} |f(h)g(x-h)|\mathrm{d}^n x\mathrm{d}^n h =$$
$$= \int_{\mathbb{R}^n} |f(h)|\mathrm{d}^n h \int_{\mathbb{R}^n} |g(x-h)|\mathrm{d}^n x = \int_{\mathbb{R}^n} |f(h)|\mathrm{d}^n h \int_{\mathbb{R}^n} |g(x)|\mathrm{d}^n x < \infty \quad (4.3.8)$$

Reading the chain of inequalities fro right to left shows that $F(x)$ is well defined in L^1.

Exercise 4.2 *Show that the convolution is Abelian ($f \star g = g \star f$) and bilinear. Show that L^1 with this operation becomes an Abelian Banach algebra (i.e. $\|f \star g\|_1 \leqslant \|f\|_1 \|g\|_1$).*

Proposition 4.3 (Elementary properties of the Fourier transform) *The following properties hold:*

1. *If $f \in L^1(\mathbb{R})$ then $\widehat{f}(0) = \frac{\int_{\mathbb{R}} f \mathrm{d}x}{2\pi}$ and $|\widehat{f}(p)| \leqslant \frac{\|f\|_1}{2\pi}$.*
2. *If $f \in L^1(\mathbb{R})$ and $g(x) := \mathrm{e}^{iax} f(x)$ for $a \in \mathbb{R}$ then $\widehat{g}(p) = \widehat{f}(p-a)$.*
3. *If $g(x) = f(x-a)$ for $a \in \mathbb{R}$ then $\widehat{g}(p) = \mathrm{e}^{-ipa}\widehat{f}(p)$.*
4. *If $g(x) = f(ax)$ for $a \in \mathbb{R}\backslash\{0\}$ then $\widehat{g}(p) = \frac{1}{a}\widehat{f}\left(\frac{p}{a}\right)$.*
5. *If $g(x) = f(-x)$ then $\widehat{g}(p) = \widehat{f}(t)^*$ (complex conjugate).*
6. *The Fourier tranform of a convolution is the product of the Fourier transforms:*

$$\widehat{f \star g}(p) = 2i\pi \widehat{f}(p)\widehat{g}(p). \quad (4.3.9)$$

Proof. We only prove the last point and leave the rest for exercise. Consider the chain of equalities (we do it in $L^1(\mathbb{R})$ but the proof extends to $\mathbb{R}^n$ without obstacle)

$$\begin{aligned}\widehat{f \star g}(p) &= \frac{1}{2i\pi}\int_{\mathbb{R}} \mathrm{d}x \mathrm{e}^{-ipx}\int_{\mathbb{R}} f(h)g(x-h) = \frac{1}{2i\pi}\int_{\mathbb{R}} \mathrm{d}x \mathrm{e}^{-ip(x-h)}\mathrm{e}^{-iph}\int_{\mathbb{R}} \mathrm{d}h f(h)g(x-h) = \\ &= \frac{1}{2i\pi}\int_{\mathbb{R}} \mathrm{d}h f(h)\mathrm{e}^{-iph}\int_{\mathbb{R}} \mathrm{d}x \mathrm{e}^{-ip(x-h)}g(x-h) = 2i\pi\widehat{f}(p)\widehat{g}(p) \end{aligned} \quad (4.3.10)$$

where the exchange of integrals is justified by Fubini's theorem. ■

Other structural properties of the Fourier transform are as follows:

Lemma 4.4 *If $f \in L^1(\mathbb{R})$ then $\widehat{f}$ is continuous and bounded.*

Proof.

$$\left|\widehat{f}(p+\epsilon) - \widehat{f}(p)\right| \leqslant \frac{1}{2\pi}\int_{\mathbb{R}} \mathrm{d}x |f(x)| \left|\mathrm{e}^{-i(p+\epsilon)x} - \mathrm{e}^{-ipx}\right| \quad (4.3.11)$$

The last integral is less than $2\|f\|_1$ and hence we can use dominated convergence to say that we can exchange the limit $\epsilon \to 0$ with the integral. Thus $\lim_{\epsilon\to 0} \widehat{f}(p+\epsilon) = \widehat{f}(p)$.

Next we clearly have $|\widehat{f}(p)| \leqslant \frac{\|f\|_1}{2\pi}$. ■

Remark 4.5 *If $f \in L^1(\mathbb{R})$ then in $\widehat{f}$ is bounded, continuous and tend to zero, as we have seen. However in general it does not belong to L^1.*

Lemma 4.6 (Further properties of the Fourier transform) **[1]** *Suppose that $f \in L^1(\mathbb{R})$ and also $xf(x) \in L^1(\mathbb{R})$. Then $\widehat{f} \in \mathcal{C}^1(\mathbb{R})$ and*

$$\frac{\mathrm{d}}{\mathrm{d}p}\widehat{f}(p) = (\widehat{-ixf(x)})(p). \tag{4.3.12}$$

[2] *Suppose that $f \in L^1$ and also $f \in \mathcal{C}^1(\mathbb{R})$, with $f' \in L^1(\mathbb{R})$. Then*

$$\widehat{f'(x)}(p) = \int_{-\infty}^{\infty} f'(x)e^{-ipx}\frac{\mathrm{d}x}{2\pi} = ip\,\hat{f}(p). \tag{4.3.13}$$

Proof. **[1]** We need to see that we can differentiate under the integral sign. This is allowed because of the assumption $xf \in L^1$ together with Fubini's theorem.

[2] From the integrability of $f'(x)$ it follows that both limits below exist

$$f(\pm\infty) := \lim_{x\to\pm\infty} f(x) = f(0) + \lim_{x\to\pm\infty}\int_0^x f'(y)\,dy.$$

Because of absolute integrability of f the limiting values $f(\pm\infty)$ must be equal to zero. Integrating by parts

$$\int_{-\infty}^{\infty} f'(x)e^{-ipx}\frac{\mathrm{d}x}{2\pi} = \left(e^{-ipx}\frac{f(x)}{2\pi}\right)_{-\infty}^{\infty} + ip\int_{-\infty}^{\infty} f'(x)e^{-ipx}dx = ip\hat{f}(p)$$

we arrive at the needed formula. ■

Example 4.7 **[1]** *Let $f(x) = \frac{a}{\pi}\frac{1}{x^2+a^2}$ with $a > 0$ (Laurentzian function). Then $\widehat{f} = \frac{1}{2\pi}\mathrm{e}^{-a|p|}$. (***Exercise***).*

Exercise 4.8 *Show that the convolution of two Gaussians or two Laurentzians (with the same centers) are still Gaussians/Laurentzians, respectively. Use property* **[6]** *in Prop. 4.3.*

4.3.1 Invertibility of the Fourier transform.

We will now establish, under certain additional assumptions, validity of the *inversion formula* for the Fourier transform:

$$\int_{-\infty}^{\infty} \hat{f}(p)e^{ipx}\mathrm{d}p = f(x). \tag{4.3.14}$$

Given a function $g(p)$ in $L^1(\mathbb{R})$ We shall use the notation $\check{}$ for this "inverse" Fourier transform

$$\check{g}(x) := \int_{\mathbb{R}} g(p)\mathrm{e}^{ipx}\mathrm{d}p. \tag{4.3.15}$$

Note that this is almost the same formula as the Fourier transform, up to the factor 2π and the sign in the exponent.

Exercise 4.9 *Show that the inverse Fourier transform of* $\frac{\mathrm{e}^{-a|p|}}{2\pi}$ *is* $\frac{a}{\pi}\frac{1}{x^2+a^2}$. *Namely (using Ex. 4.7) that the inverse Fourier transform is indeed an identity in this case.*

Proof. We have

$$\begin{aligned}\int_{\mathbb{R}} \frac{\mathrm{e}^{ipx-a|p|}}{2\pi}\mathrm{d}p &= \left(\int_0^{\infty} + \int_{-\infty}^{0}\right)\frac{\mathrm{e}^{ipx-a|p|}}{2\pi}\mathrm{d}p = \int_0^{\infty}\mathrm{e}^{(ix-a)p}\frac{\mathrm{d}p}{2\pi} + \int_{-\infty}^{0}\mathrm{e}^{(ix+a)p}\frac{\mathrm{d}p}{2\pi} = \\ &= \frac{1}{2\pi}\left(\frac{1}{a-ix} + \frac{1}{a+ix}\right) = \frac{a}{\pi}\frac{1}{x^2+a^2}. \end{aligned} \tag{4.3.16}$$

■

Remark 4.10 *In general the statement is false. Consider* $f(x) := \mathrm{e}^{-x}\chi_{[0,\infty)}(x)$. *Then its Fourier transform is* $\widehat{f}(p) = \frac{1}{2\pi}\frac{1}{ip+1}$ *and this is not in* L^1. *The inverse Fourier transform is not defined.*

We recall now (without proof) the fact that

Lemma 4.11 *The set* $\mathcal{C}_0^0(\mathbb{R})$ *is dense in* $L^1(\mathbb{R})$. *(In fact in any* L^p, $p \geqslant 1$*).*

Based on the above density statement we state and prove the

Lemma 4.12 (Moving average lemma) *Let* $\rho \in L^1$ *be a non-negative function of total mass 1:* $\rho(x) \geqslant 0$, $\int \rho \mathrm{d}x = 1$. *Define* $\rho_n(x) := n\rho(nx)$ *so that* $\int \rho_n \mathrm{d}x = 1$ *for all* $n \in N$.

For an arbitrary $f \in L^1$ *define* $f_n(x) := (f \star \rho_n)(x) = \int_{\mathbb{R}} \rho_n(x-h)f(h)\mathrm{d}h$ *("moving average"). Then* $\|f_n - f\|_1 \to 0$.

Proof. We start with $f \in \mathcal{C}_0^0(\mathbb{R})$ and compute

$$f_n(x) - f(x) = \int_{\mathbb{R}} \rho_n(h)f(x-h)\mathrm{d}h - f(x) \overset{\int\rho_n=1}{=} \int_{\mathbb{R}} \rho_n(h)\Big(f(x-h) - f(x)\Big)\mathrm{d}h \tag{4.3.17}$$

Consequently we have

$$\begin{aligned}\Big\|f_n(x) - f(x)\Big\|_1 &\leqslant \int_{\mathbb{R}}\mathrm{d}x\int_{\mathbb{R}}\rho_n(h)\Big|f(x-h) - f(x)\Big|\mathrm{d}h = \int_{\mathbb{R}}\mathrm{d}x\int_{\mathbb{R}}\rho(s)\left|f\left(x-\frac{s}{n}\right) - f(x)\right|\mathrm{d}s = \\ &\overset{\text{Fubini}}{=} \int_{\mathbb{R}}\mathrm{d}s\rho(s)\int_{\mathbb{R}}\mathrm{d}x\left|f\left(x-\frac{s}{n}\right) - f(x)\right| \end{aligned} \tag{4.3.18}$$

Since $f \in \mathcal{C}_0^0$ then $\Delta f := f\left(x - \frac{s}{n}\right) - f(x)$ is also in $\mathcal{C}_0^0$ and for any $s \in \mathbb{R}$ it converges uniformly to 0 as $n \to \infty$. Using dominated convergence we establish thus that

$$J_n(s) := \int_{\mathbb{R}}\mathrm{d}x\left|f\left(x-\frac{s}{n}\right) - f(x)\right| \tag{4.3.19}$$

is bounded, positive and tends to zero. Thus the whole integral (4.3.18) tends to zero.

If $f \in L^1$ is now arbitrary, then a simple density argument using the triangle inequality shows that $f_n \to f$ in L^1. ■

Theorem 4.13 (Injectivity of the Fourier transform and invertibility) *Let $f \in L^1$ be such that $\widehat{f}$ belongs to L^1. Then $\check{\widehat{f}}(x) = f(x)$ almost everywhere (a.e.).*

Proof. We use the moving average lemma 4.12 with

$$\rho(x) = \frac{1}{\pi}\frac{1}{x^2+1} = \rho_1(x). \tag{4.3.20}$$

We know that $\check{\widehat{\rho_n}} = \rho_n$ because of exercise 4.9:

$$\rho_m(h) = \int_{\mathbb{R}} \mathrm{d}h\, \widehat{\rho_m}(p)\mathrm{e}^{ihp}. \tag{4.3.21}$$

Thus

$$(f \star \rho_m)(x) = \int_{\mathbb{R}} \mathrm{d}h\, f(x-h) \int_{\mathbb{R}} \mathrm{d}p\, \widehat{\rho_m}(p)\mathrm{e}^{ihp}. \tag{4.3.22}$$

Since $\int_{\mathbb{R}} \mathrm{d}h\, |f(x-h)| \int_{\mathbb{R}} \mathrm{d}p\, |\widehat{\rho_m}(p)| < \infty$ we can apply Fubini's theorem and exchange the order of integrations in (4.3.22):

$$\begin{aligned}(f \star \rho_m)(x) &= \int_{\mathbb{R}} \mathrm{d}p \int_{\mathbb{R}} \mathrm{d}h\, \widehat{\rho_m}(p)\mathrm{e}^{ihp}\, f(x-h) = \\ &= \int_{\mathbb{R}} \mathrm{d}p\, \widehat{\rho_m}(p)\mathrm{e}^{ipx} \int_{\mathbb{R}} \mathrm{d}h\, f(x-h)\mathrm{e}^{-ip(x-h)} = \\ &= 2\pi \int_{\mathbb{R}} \mathrm{d}p\, \widehat{f}(p)\mathrm{e}^{ipx}\widehat{\rho}_m(p) \overset{\text{Ex. 4.7}}{=} \int_{\mathbb{R}} \mathrm{d}p\, \widehat{f}(p)\mathrm{e}^{ipx}\mathrm{e}^{-\frac{|p|}{m}}\end{aligned} \tag{4.3.23}$$

In this last expression we can pass to the limit $m \to \infty$ inside the integral sign because we have assumed that $\widehat{f} \in L^1$; the limit gives $\check{\widehat{f}}(x)$. On the left side, using the Moving Average Lemma 4.12 we have also that the limit converges in L^1 to $f(x)$ (i.e. almost everywhere). ∎

A more elementary approach

Given the importance of the subject we provide a more direct proof of invertibility (but for stronger assumptions on f). For this we need the

Lemma 4.14 (Riemann–Lebesgue lemma) *Let a $f(x)$ be absolutely integrable on $\mathbb{R}$. Then*

$$\lim_{\lambda\to\infty} \int_{-\infty}^{\infty} f(x)e^{i\lambda x}dx = 0.$$

Proof:

We start from a simple exercise: the characteristic function, $\chi_{[a,b]}$ of an interval $[a,b]$ has the Fourier transform

$$\int_{-\infty}^{\infty} \mathrm{e}^{i\lambda x}\chi_{[a,b]}(x)\mathrm{d}x = \int_a^b \mathrm{e}^{i\lambda x}\mathrm{d}x = \frac{1}{i\lambda}(\mathrm{e}^{ib\lambda} - \mathrm{e}^{ia\lambda}) = \mathrm{e}^{i\lambda\frac{a+b}{2}}\frac{2\sin\left(\frac{\lambda(b-a)}{2}\right)}{\lambda} \tag{4.3.24}$$

which clearly tends to zero as $|\lambda| \to \infty$.

Next: a function $f \in L^1(\mathbb{R})$ can be uniformly approximated in the L^1 norm by a **finite** linear combination of characteristic functions of bounded intervals (i.e. by a *simple function*):

$$\forall f \in L^1(\mathbb{R}) \ \ \forall \epsilon > 0 \quad \exists g(x) \text{ simple} : \int_{\mathbb{R}} |f - g| \mathrm{d}x \leqslant \epsilon. \tag{4.3.25}$$

This statement should be familiar in the context of Lebesgue integration and it is a consequence of monotone convergence.

For any simple function $g(x)$ it follows from (4.3.24) and from the finiteness of the sum appearing in the simple function, that its Fourier transform tends to zero as $|\lambda| \to \infty$:

$$\forall g \in L^1(\mathbb{R}), \ \ \text{simple}, \forall \epsilon > 0 \ \ \exists R > 0 : \ \ |\lambda| > R \ \Rightarrow \ \ |\widehat{g}(\lambda)| < \epsilon. \tag{4.3.26}$$

We combine these two properties with the triangle inequality yields the result as follows. Let $f \in L^1$ and g simple within $\pi\epsilon$ distance (in L^1) from f. Then let R as in (4.3.26). We conclude from the linearity of the Fourier transform that for all $|\lambda| \geqslant R$ we have

$$\begin{aligned} |\widehat{f}(\lambda)| \leqslant \left|\widehat{f}(\lambda) - \widehat{g}(\lambda)\right| + |\widehat{g}(\lambda)| < \int_{\mathbb{R}} \left|(f(x) - g(x))\mathrm{e}^{i\lambda x}\right| \frac{\mathrm{d}x}{2\pi} + \frac{\epsilon}{2} \leqslant \\ \leqslant \int_{\mathbb{R}} |(f(x) - g(x))| \frac{\mathrm{d}x}{2\pi} + \frac{\epsilon}{2} < \epsilon. \end{aligned} \tag{4.3.27}$$

This is the definition of $\lim_{|\lambda| \to \infty} |\widehat{f}(\lambda)| = 0$.

The above proof extends to higher dimensions by using the same steps and characteristic functions of bounded multi-intervals. ■

We also need

Exercise 4.15 (Dirichlet integral) *Prove that*

$$\int_0^\infty \frac{\sin x}{x} \, dx = \frac{\pi}{2}. \tag{4.3.28}$$

Proof. Consider the function

$$G(t) := \int_0^\infty \mathrm{e}^{-tx} \frac{\sin(x)}{x} \mathrm{d}x, \qquad t > 0. \tag{4.3.29}$$

We can compute its derivative

$$G'(t) = \int_0^\infty \mathrm{e}^{-tx} \sin(x) \mathrm{d}x = -\frac{1}{1 + t^2} \tag{4.3.30}$$

from which it follows that

$$G(t) = -\arctan(t) + C. \tag{4.3.31}$$

Since clearly $G(t) \to 0$ as $t \to \infty$, we conclude that

$$G(t) = \frac{\pi}{2} - \arctan(t). \tag{4.3.32}$$

Thus $\lim_{t \to 0^+} G(t) = \frac{\pi}{2}$. ■

Theorem 4.16 *Let the absolutely integrable function $f(x)$ be differentiable at any point $x \in \mathbb{R}$. Then*

$$\lim_{R\to\infty} \int_{-R}^{R} \hat{f}(p) e^{ipx} dp = f(x). \tag{4.3.33}$$

Remark 4.17 *In this theorem the integral is the standard Riemann integral.*

Proof: Denote $I_R(x)$ the integral in the left hand side of (4.3.33). Using continuity and uniform convergence of the Fourier integral (4.3.2) we can apply Fubini theorem to this integral and thus rewrite it as follows:

$$\begin{aligned} I_R(x) &= \int_{-R}^{R} \hat{f}(p) e^{ipx} dp = \int_{-R}^{R} \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} f(y) e^{-ipy} dy \right) e^{ipx} dp \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} f(y) \left(\int_{-R}^{R} e^{ip(x-y)} dp \right) dy = \frac{1}{\pi} \int_{-\infty}^{\infty} f(y) \frac{\sin R(x-y)}{x-y} dy \\ &= \frac{1}{\pi} \int_{-\infty}^{\infty} f(x+s) \frac{\sin R\, s}{s} ds = \frac{1}{\pi} \int_0^{\infty} [f(x+s) + f(x-s)] \frac{\sin R\, s}{s} ds. \end{aligned}$$

Using the Dirichlet integral (4.3.28) we can rewrite the difference $I_R(x) - f(x)$ in the form

$$I_R(x) - f(x) = \frac{1}{\pi} \int_0^{\infty} \frac{f(x+s) - 2f(x) + f(x-s)}{s} \sin R\, s\, ds.$$

Because of differentiablity

$$\lim_{s\to 0} \frac{f(x+s) - 2f(x) + f(x-s)}{s} = \lim_{s\to 0} \frac{f(x+s) - f(x)}{s} + \lim_{s\to 0} \frac{f(x-s) - f(x)}{s} = f'(x) - f'(x) = 0$$

the integrand

$$F(s;x) = \begin{cases} \frac{f(x+s)-2f(x)+f(x-s)}{s}, & s \neq 0 \\ 0, & s = 0 \end{cases}$$

is a continuous functions in s depending on the parameter x. In order to complete the proof of the Theorem let us represent the last integral in the form

$$\begin{aligned} \int_0^{\infty} \frac{f(x+s) - 2f(x) + f(x-s)}{s} \sin R\, s\, ds &= \int_0^1 F(s;x) \sin R\, s\, ds \\ + \int_1^{\infty} \frac{f(x+s) + f(x-s)}{s} \sin R\, s\, ds &- 2f(x) \int_1^{\infty} \frac{\sin R\, s}{s} ds. \end{aligned}$$

The first integral in the r.h.s. vanishes according to the Riemann–Lebesgue lemma. The same is true for the second and third integrals. Finally the last integral by a change of integration variable $x = R\, s$ reduces to

$$\int_1^{\infty} \frac{\sin R\, s}{s} ds = \int_R^{\infty} \frac{\sin x}{x} dx \to 0 \quad \text{for} \quad R \to \infty.$$

■

Exercise 4.18 *Let $f(x)$ be an absolutely integrable piecewise continuous function of $x \in \mathbb{R}$ differentiable on every interval of continuity. Let us also assume that at every discontinuity point x_0 the left and right limits $f_-(x_0)$ and $f_+(x_0)$ exists and, moreover, the left and right derivatives*

$$\lim_{s\to 0-} \frac{f(x_0+s)-f_-(x_0)}{s} \quad \text{and} \quad \lim_{s\to 0+} \frac{f(x_0+s)-f_+(x_0)}{s}$$

exist as well. Prove the following modification of the inversion formula for the Fourier transform

$$\lim_{R\to\infty} \int_{-R}^{R} \hat{f}(p) e^{ipx} dp = \begin{cases} f(x), & x \text{ is a continuity point} \\ \frac{f_-(x)+f_+(x)}{2}, & x \text{ is a discontinuity point} \end{cases} \tag{4.3.34}$$

Hint: *re-trace the proof of Theorem 4.16. Replace $f(x)$ in the definition of $F(s;x)$ by $\frac{f_+(x)+f_-(x)}{2}$.*

The main property of Fourier transform used for solving linear PDEs is given by the following formula:

Denote by $\mathcal{F}_{x\to p}$ the map of the space of functions in x variable to the space of functions in the variable p given by the Fourier transform:

$$\mathcal{F}_{x\to p}(f) = \hat{f}(p). \tag{4.3.35}$$

The inverse Fourier transform will now be denoted $\mathcal{F}_{p\to x}$. The property formulated in the Lemma 4.6 says that the operator of x-derivative transforms to the operator of multiplication by the independent variable, up to a factor i:

$$\mathcal{F}_{x\to p}\left(\frac{d}{dx} f\right) = ip\, \mathcal{F}_{x\to p}(f). \tag{4.3.36}$$

and

$$\mathcal{F}_{x\to p}(x\, f) = i \frac{d}{dp} \mathcal{F}_{x\to p}(f) \tag{4.3.37}$$

valid for functions $f = f(x)$ absolutely integrable together with $xf(x)$. We leave the proof of this formula as an exercise for the reader.

4.3.2 Extension to $\mathbb{R}^n$

For completeness we give the definition of Fourier transform in n–dimensions; we denote by $\mathbf{x} = (x_1, \ldots, x_n)$ and $\mathbf{p} = (p_1, \ldots, p_n)$. Then

$$\mathcal{F}[f](\mathbf{p}) := \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \mathrm{e}^{i\mathbf{p}\cdot\mathbf{x}} f(\mathbf{x}) \mathrm{d}^n\mathbf{x}. \tag{4.3.38}$$

$$\mathcal{F}^{-1}[g](\mathbf{x}) = \int_{\mathbb{R}^n} \mathrm{e}^{-i\mathbf{p}\cdot\mathbf{x}} g(\mathbf{p}) \mathrm{d}^n\mathbf{p} \tag{4.3.39}$$

The properties that we have stated for $n = 1$ extend as follows

$$(f \star g)(\mathbf{x}) = \int_{\mathbb{R}^n} \mathrm{d}^n\mathbf{h} f(\mathbf{x}-\mathbf{h}) g(\mathbf{h}), \tag{4.3.40}$$

and

$$\mathcal{F}(f \star g)(\mathbf{p}) = (2\pi)^n \widehat{f}(\mathbf{p}) \widehat{g}(\mathbf{p}). \tag{4.3.41}$$

The formulation of the interplay of derivatives and Fourier transform is left as **exercise**.

4.4 Solution to the Cauchy problem for heat equation on the line

Let us consider the one-dimensional Cauchy problem for the heat equation

$$\begin{aligned} &\frac{\partial u}{\partial t} = a^2 \frac{\partial^2 u}{\partial x^2}, \quad t > 0 \\ &u(x,0) = \phi(x), \quad x \in \mathbb{R}. \end{aligned} \tag{4.4.1}$$

Theorem 4.19 *Let the initial data $\phi(x)$ be absolutely integrable function on $\mathbb{R}$. Then the Cauchy problem (4.4.1) has a unique solution $u(x,t)$ absolutely integrable in $x \in \mathbb{R}$ for all $t > 0$ represented by the formula*

$$u(x,t) = \int_{-\infty}^{\infty} G(x-y;t)\, \phi(y)\, dy. \tag{4.4.2}$$

where G is the Gaussian distribution

$$G(x;t) = \frac{1}{2a\sqrt{\pi t}} e^{-\frac{x^2}{4a^2 t}}. \tag{4.4.3}$$

The integral representation (4.4.2) of solutions to the Cauchy problem is called *Poisson integral.*

Remark 4.20 *We have the following remarks:*

1. *The formula above is rewritten also using the convolution operator as*

$$u(x,t) = (G(\bullet;t) \star \phi)(x) \tag{4.4.4}$$

2. *Even if $\phi \in L^1$ is not continuous/differentiable, one observes that $u(x,t)$ is $\mathcal{C}^\infty$ for any $t > 0$; thus the heat equation "regularizes" (or has a smoothing effect) on the initial datum. Note that the formula does not make sense for negative times (for general ϕ).*

Proof: Denote

$$\hat{u}(p,t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} u(x,t) e^{-ipx} dx$$

the Fourier-image of the unknown solution. According to Lemma 4.6 the function $\hat{u}(p,t)$ satisfies equation

$$\frac{\partial \hat{u}(p,t)}{\partial t} = -a^2 p^2 \hat{u}(p,t).$$

This equation can be easily solved

$$\hat{u}(p,t) = \hat{u}(p,0) e^{-a^2 p^2 t}.$$

Due to the initial condition we obtain

$$\hat{u}(p,0) = \hat{\phi}(p) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \phi(x) e^{-ipx} dx.$$

Thus

$$\hat{u}(p,t) = \hat{\phi}(p)e^{-a^2p^2t}. \tag{4.4.5}$$

It remains to apply the inverse Fourier transform to this formula:

$$u(x,t) = \int_{-\infty}^{\infty} e^{ixp}\hat{\phi}(p)e^{-a^2p^2t}dp = \frac{1}{2\pi}\int_{-\infty}^{\infty}\left(e^{ixp-a^2p^2t}\int_{-\infty}^{\infty} e^{-ipy}\phi(y)\,dy\right)dp$$
$$= \frac{1}{2\pi}\int_{-\infty}^{\infty}\phi(y)\left(\int_{-\infty}^{\infty} e^{ip(x-y)-a^2p^2t}dp\right)dy$$

The integral in p is nothing but the (inverse) Fourier transform of the Gaussian function. A calculation similar to the above one gives the value for this integral

$$\int_{-\infty}^{\infty} e^{ip(x-y)-a^2p^2t}dp = \frac{\sqrt{\pi}}{a\sqrt{t}}e^{-\frac{(x-y)^2}{4a^2t}}.$$

This completes the proof of the Theorem. ■

Remark 4.21 *The formula (4.4.2) can work also for not necessarily absolutely integrable functions. For example for the constant initial data $\phi(x) \equiv \phi_0$ we obtain $u(x,t) \equiv \phi_0$ due to the following integral*

$$\frac{1}{2a\sqrt{\pi t}}\int_{-\infty}^{\infty} e^{-\frac{(x-y)^2}{4a^2t}}dy \equiv 1. \tag{4.4.6}$$

4.4.1 Heat equation in $\mathbb{R}^d$

We can similarly pose and solve the Cauchy problem in$\mathbb{R}^d$ as follows:

$$\frac{\partial u}{\partial t} = a^2\triangle u, \quad t > 0 \tag{4.4.7}$$
$$u(\mathbf{x};0) = \phi(\mathbf{x}). \tag{4.4.8}$$

The idea is to take the Fourier transform of both sides in $\mathbf{x}$; we proceed a bit formally (without great regards to inversion of order of derivatives and integrals) and we obtain

$$\frac{\partial\widehat{u}}{\partial t}(\mathbf{p},t) = -a^2\|\mathbf{p}\|^2\widehat{u}(\mathbf{p},t) \tag{4.4.9}$$

where $\|\mathbf{p}\| = \sum_{j=1}^d p_j^2$ is just the square of the Euclidean norm of $\mathbf{p}$. This ODE is easily solved:

$$\widehat{u}(\mathbf{p},t) = \mathrm{e}^{-a^2\|\mathbf{p}\|^2t}\widehat{u}(\mathbf{p},0) \tag{4.4.10}$$

This now appears to be the product of two Fourier transforms, where

$$\mathrm{e}^{-a^2\|\mathbf{p}\|^2t} = \frac{1}{(a^2\pi t)^{\frac{d}{2}}}\int_{\mathbb{R}^d}\mathrm{d}^d\mathbf{x}\;\mathrm{e}^{-\frac{\|\mathbf{x}\|^2}{4a^2t}+i\mathbf{p}\cdot\mathbf{x}} \tag{4.4.11}$$

Therefore the Poisson kernel (or "heat kernel") is defined as

$$G(\mathbf{x};t) := \frac{1}{(4a^2\pi t)^{\frac{d}{2}}}\mathrm{e}^{-\frac{\|\mathbf{x}\|^2}{4a^2t}}. \tag{4.4.12}$$

This allows us to solve the Cauchy problem (4.4.7) as a convolution exactly as before;

$$u(\mathbf{x};t) = (G(\cdot;t) \star \phi)(\mathbf{x}) = \int_{\mathbb{R}^d} \mathrm{d}^d\mathbf{y} \frac{1}{(4a^2\pi t)^{\frac{d}{2}}} \mathrm{e}^{-\frac{\|\mathbf{x}-\mathbf{y}\|^2}{4a^2 t}} \phi(\mathbf{y}). \tag{4.4.13}$$

We will now use the Poisson integral (4.4.2) (or (4.4.12) in order to prove an analogue of the maximum principle for solutions to the heat equation.

Theorem 4.22 *The solution to the Cauchy problem represented by the Poisson integral (4.4.2) for all $t > 0$ satisfies*

$$\inf_{x\in\mathbb{R}} \phi(x) \leqslant u(x,t) \leqslant \sup_{x\in\mathbb{R}} \phi(x). \tag{4.4.14}$$

Moreover, if some of the inequalities becomes equality for some $t > 0$ and $x \in \mathbb{R}$ then $u(x,t) \equiv \mathrm{const}$.

Proof: The inequalities (4.4.14) easily follow from positivity of the Gaussian function and from the integral (4.4.6). Due to the same positivity the equality can have place only if $\phi(x) = \mathrm{const}$. But then also $u(x,t) = \mathrm{const}$. ■

Corollary 4.23 *The solution to the Cauchy problem (4.4.1) for the heat equation depends continuously on the initial data in the* sup *norm.*

Proof: Let $u_1(x,t)$, $u_2(x,t)$ be two solutions to the heat equation with the initial data $\phi_1(x)$ and $\phi_2(x)$ respectively. If the initial data differ by ϵ, i.e.

$$|\phi_1(x) - \phi_2(x)| \leqslant \epsilon \quad \forall x \in \mathbb{R}$$

then from the maximum principle applied to the solution $u(x,t) = u_1(x,t) - u_2(x,t)$ it follows that

$$|u_1(x,t) - u_2(x,t)| \leqslant \epsilon.$$

■

4.5 Mixed boundary value problems for the heat equation

Let us begin with the periodic problem

$$\begin{aligned} &\frac{\partial u}{\partial t} = a^2 \frac{\partial^2 u}{\partial x^2}, \quad t > 0 \\ &u(x + 2\pi, t) = u(x,t), \quad t > 0 \\ &u(x,0) = \phi(x) \end{aligned} \tag{4.5.1}$$

where $\phi(x)$ is a smooth 2π-periodic function.

Theorem 4.24 *There exists a unique solution to the problem (4.5.1). It can be represented in the form*

$$u(x,t) = \frac{1}{2\pi}\int_0^{2\pi} \Theta(x-y;t)\phi(y)\,dy, \quad t>0 \tag{4.5.2}$$

where

$$\Theta(x;t) = \sum_{n\in\mathbb{Z}} e^{-a^2n^2t+inx}. \tag{4.5.3}$$

Proof: Let us expand the unknown periodic function $u(x,t)$ in the Fourier series:

$$u(x,t) = \sum_{n\in\mathbb{Z}} \hat{u}_n(t)e^{inx}$$

$$\hat{u}_n(t) = \frac{1}{2\pi}\int_0^{2\pi} u(x,t)e^{-inx}dx.$$

The substitution to the heat equation yields

$$\frac{\partial \hat{u}_n(t)}{\partial t} = -a^2n^2\hat{u}_n(t),$$

so

$$\hat{u}_n(t) = \hat{u}_n(0)e^{-a^2n^2t}, \quad n\in\mathbb{Z}.$$

At $t=0$ one must meet the initial conditions, hence we arrive at the formula

$$\hat{u}_n(t) = \hat{\phi}_n e^{-a^2n^2t}$$

$$\hat{\phi}_n = \frac{1}{2\pi}\int_0^{2\pi} \phi(y)e^{-iny}dy.$$

For the function $u(x,t)$ we obtain

$$u(x,t) = \frac{1}{2\pi}\sum_{n\in\mathbb{Z}}\int_0^{2\pi} e^{-a^2n^2t+in(x-y)}\phi(y)dy.$$

In order to complete the proof of the Theorem it suffices to show that the series (4.5.3) converges absolutely and uniformly for all $x\in\mathbb{R}$ and all $t>0$. This easily follows from convergence of the integral

$$\int_0^\infty e^{-a^2x^2t}dx < \infty \quad \text{for} \quad t>0.$$

In a similar way one can prove that the series (4.5.3) can be differentiated any number of times. The theorem is proved. ■

The functioned defined by the series (4.5.3) is called *theta-function*. It is expressed via the *Jacobi theta-function*

$$\theta_3(\phi\,|\,\tau) = \sum_{n\in\mathbb{Z}} e^{\pi in^2\tau+2\pi in\phi} \tag{4.5.4}$$

by a change of variables

$$\Theta(x;t) = \theta(\phi\,|\,\tau), \quad \phi = \frac{1}{2\pi}x, \quad \tau = i\frac{a^2 t}{\pi}. \tag{4.5.5}$$

The convergence of the series (4.5.4) for Jacobi theta function takes place for all complex values of τ provided

$$\operatorname{Im}\tau > 0. \tag{4.5.6}$$

The function $\Theta(x;t)$ is periodic in x with the period 2π while the Jacobi theta-function is periodic in ϕ with the period 1. It satisfies many remarkable properties. Some them will be now formulated as a series of exercises.

Exercise 4.25 *Prove that*

$$\int_0^{2\pi} \Theta(x;t)\,dx = 2\pi. \tag{4.5.7}$$

Exercise 4.26 *Prove that the series*

$$\sum_{n\in\mathbb{Z}} e^{-a^2 n^2 t + inz} \tag{4.5.8}$$

converges for any complex number $z = x+iy$ uniformly on the strips $|\operatorname{Im} z| \leqslant M$ for any positive M. Derive that the theta-function (4.5.3) can be analytically continued to a function $\Theta(z;t)$ holomorphic on the entire complex z-plane.

Exercise 4.27 *Prove that the function $\Theta(z;t)$ satisfies the identity*

$$\Theta\left(z + 2ia^2 t; t\right) = e^{a^2 t - iz}\Theta(x;t). \tag{4.5.9}$$

The complex number $2ia^2 t$ is called *quasi-period* of the theta-function.

Exercise 4.28 *Prove that the theta-function has zeroes at the points*

$$x_{k\,\ell} = \pi(2k+1) + ia^2 t\,(2\ell+1), \quad k,\ \ell \in \mathbb{Z}. \tag{4.5.10}$$

Exercise 4.29 *Prove that the theta-function has no other zeroes on the complex plane. Derive that, in particular*

$$\Theta(x;t) > 0 \quad \text{for} \quad x \in \mathbb{R}. \tag{4.5.11}$$

Hint for the last two exercises: Compute the integrals

$$\frac{1}{2\pi i}\oint_C \frac{d\Theta(z;t)}{\Theta(z,t)} \qquad \frac{1}{2\pi i}\oint_C z\frac{d\Theta(z;t)}{\Theta(z,t)}$$

over the oriented boundary of the rectangle

$$C = \{0 \leqslant x \leqslant 2\pi,\ 0 \leqslant y \leqslant 2a^2 t\}$$

on the complex z-plane, $z = x + iy$. The first counts the number of zeroes inside, the second computes the sum of their positions. □

Another proof of positivity of the theta-function follows from the following *Poisson summation formula* that is of course of interest on its own.

Lemma 4.30 (Poisson summation formula) *Let $f(x)$ be a continuously differentiable absolutely integrable function satisfying the inequalities*

$$|f(x)| < C(1+|x|)^{-1-\epsilon}, \quad |\hat{f}(p)| < C(1+|p|)^{-1-\epsilon}$$

for some positive ϵ. Here $\hat{f}(p)$ is the Fourier transform of $f(x)$. Then

$$\sum_{n\in\mathbb{Z}} f(2\pi n) = \sum_{m\in\mathbb{Z}} \hat{f}(m). \tag{4.5.12}$$

Proof: We will actually prove a somewhat more general formula

$$\sum_{n\in\mathbb{Z}} f(x+2\pi n) = \sum_{m\in\mathbb{Z}} \hat{f}(m)e^{imx}. \tag{4.5.13}$$

Since the function in the left hand side is 2π-periodic in x, it suffices to check that the Fourier coefficients c_m of this function coincide with $\hat{f}(m)$. Indeed, the m-th Fourier coefficient of the left hand side is equal to

$$c_m = \frac{1}{2\pi}\int_0^{2\pi}\left(\sum_{n\in\mathbb{Z}} f(x+2\pi n)\right)e^{-imx}dx.$$

Due to absolute and uniform (in x) convergence of the series

$$\sum_{n\in\mathbb{Z}} f(x+2\pi n)$$

one interchange the order of summation and integration to arrive at

$$c_m = \frac{1}{2\pi}\sum_{n\in\mathbb{Z}}\int_0^{2\pi} f(x+2\pi n)e^{-imx}dx.$$

Doing a shift in the n-the integral

$$y = x + 2\pi n$$

one rewrites the sum as follows:

$$c_m = \frac{1}{2\pi}\sum_{n\in\mathbb{Z}}\int_{2\pi n}^{2\pi(n+1)} f(y)e^{-imy-2\pi imn}dy = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(y)e^{-imy}dy = \hat{f}(m)$$

since $e^{-2\pi imn} = 1$. ■

Using the Poisson summation formula we can prove the following remarkable identity for the theta-function.

Proposition 4.31 *The theta-function (4.5.1) satisfies the following identity*

$$\Theta(x;t) = \frac{1}{a}\sqrt{\frac{\pi}{t}}\sum_{n\in\mathbb{Z}} e^{-\frac{(x+2\pi n)^2}{4a^2 t}}. \tag{4.5.14}$$

Proof: It can be obtained by applying the Poisson summation formula to the function

$$f(x) = \frac{1}{a}\sqrt{\frac{\pi}{t}}\, e^{-\frac{x^2}{4a^2 t}}, \quad \hat{f}(p) = e^{-a^2 p^2 t}.$$

■

Remark 4.32 *The formula (4.5.14) is the clue to derivation of the transformation law for the Jacobi theta-function under* modular transformations

$$\tau \mapsto \frac{a\tau + b}{c\tau + d}, \quad a,\, b,\, c,\, d \in \mathbb{Z}, \quad ad - bc = 1.$$

Let us now consider the first mixed problem for heat equation on the interval $[0, l]$ with zero boundary conditions:

$$\begin{aligned} &\frac{\partial u}{\partial t} = a^2 \frac{\partial^2 u}{\partial x^2}, \quad 0 \leqslant x \leqslant l, \quad t > 0 \\ &u(0,t) = u(l,t) = 0 \\ &u(x,0) = \phi(x), \quad 0 \leqslant x \leqslant l. \end{aligned} \tag{4.5.15}$$

Like in Section 2.6 above, let us extend the initial data $\phi(x)$ to the real line as an odd $2l$-periodic function. We leave as an exercise for the reader to check that the solution to this periodic Cauchy problem will remain an odd periodic function for all times and, hence, it will vanish at the points $x = 0$ and $x = l$. In this way one arrives at the following

Theorem 4.33 *The mixed b.v.p. (4.5.15) has a unique solution for an arbitrary smooth function $\phi(x)$. It can be represented by the following integral*

$$u(x,t) = \frac{1}{l}\int_0^l \tilde{\Theta}(x,y;t)\phi(y)\, dy \tag{4.5.16}$$

where

$$\Theta(x,y;t) = 2\sum_{n=1}^{\infty} e^{-a^2 n^2 t} \sin\frac{\pi n x}{l} \sin\frac{\pi n y}{l}. \tag{4.5.17}$$

4.6 More general boundary conditions for the heat equation. Solution to the inhomogeneous heat equation

In the previous section the simplest b.v.p. for the heat equation has been considered. We will now address the more general problem

$$\begin{aligned} &\frac{\partial u}{\partial t} = a^2 \frac{\partial^2 u}{\partial x^2}, \quad t > 0, \quad 0 < x < l \\ &u(0,t) = f_0(t), \quad u(l,t) = f_1(t), \quad t > 0 \\ &u(x,0) = \phi(x), \quad 0 < x < l. \end{aligned} \tag{4.6.1}$$

The following simple procedure reduces the above problem to the b.v.p. with zero boundary condition for the *inhomogeneous* heat equation

$$\begin{aligned} &\frac{\partial v}{\partial t} = a^2 \frac{\partial^2 v}{\partial x^2} + F(x,t), \quad t>0, \quad 0<x<l \\ &v(0,t) = v(l,t) = 0, \quad t>0 \\ &v(x,0) = \Phi(x), \quad 0<x<l \end{aligned} \tag{4.6.2}$$

where the functions $F(x,t)$, $\Phi(x)$ are given by

$$\begin{aligned} F(x,t) &= -\left[\frac{df_0(t)}{dt} + \frac{x}{l}\left(\frac{df_1(t)}{dt} - \frac{df_0(t)}{dt}\right)\right] \\ \Phi(x) &= \phi(x) - \left[f_0(0) + \frac{x}{l}(f_1(0) - f_0(0))\right]. \end{aligned} \tag{4.6.3}$$

Indeed, it suffices to do the following substitution

$$u(x,t) = v(x,t) + \left[f_0(t) + \frac{x}{l}(f_1(t) - f_0(t))\right] \tag{4.6.4}$$

observing that the expression in the square brackets is annihilated by the operator $\partial^2/\partial x^2$. Moreover, the function in the square brackets takes the needed values $f_0(t)$ and $f_1(t)$ at the endpoints of the interval.

In the more general case of multidimensional heat equation with non-vanishing boundary conditions

$$\begin{aligned} &\frac{\partial u}{\partial t} = a^2 \Delta u, \quad t>0, \quad x \in \Omega \subset \mathbb{R}^d \\ &u(x,t)|_{x\in\partial\Omega} = f(x,t), \quad t>0 \\ &u(x,0) = \phi(x), \quad x \in \Omega \end{aligned} \tag{4.6.5}$$

the procedure is similar to the above one. Namely, denote $u_0(x,t)$ the solution to the Dirichlet boundary value problem for the Laplace equation in x depending on t as on the parameter:

$$\begin{aligned} &\Delta u_0 = 0, \quad x \in \Omega \subset \mathbb{R}^d \\ &u_0(x,t)|_{x\in\partial\Omega} = f(x,t). \end{aligned} \tag{4.6.6}$$

We already know that the solution to the Dirichlet boundary value problem is unique and depends continuously on the boundary conditions. Therefore the solution $u_0(x,t)$ is a continuous function on $\Omega \times \mathbb{R}_{>0}$. One can also prove that this functions is smooth, if the boundary data $f(x,t)$ are so. Then the substitution

$$u(x,t) = v(x,t) + u_0(x,t) \tag{4.6.7}$$

reduces the mixed b.v.p. (4.6.6) to the one with zero boundary conditions

$$v(x,t)|_{x\in\partial\Omega} = 0, \quad t>0$$

with the modified initial data

$$v(x,0) = \phi(x) - u_0(x,0), \quad x \in \Omega$$

but the heat equation becomes inhomogeneous one:

$$\frac{\partial v}{\partial t} = a^2 \Delta v + F(x,t), \quad F(x,t) = -\frac{\partial u_0(x,t)}{\partial t}, \quad x \in \Omega.$$

We will now explain a simple method for solving the inhomogeneous heat equation. For the sake of simplicity let us consider in details the case of one spatial variable. Moreover we will concentrate on the infinite line case. So the problem under consideration is in finding a function $u(x,t)$ on $\mathbb{R} \times \mathbb{R}_{>0}$ satisfying

$$\begin{aligned} &\frac{\partial u}{\partial t} = a^2 \frac{\partial^2 u}{\partial x^2} + f(x,t), \quad x \in \mathbb{R}, \quad t > 0 \\ &u(x,0) = \phi(x). \end{aligned} \tag{4.6.8}$$

The solution is found using the same Duhamel principle explained in Sect. 2.8.

Theorem 4.34 *The solution to the inhomogeneous problem (4.6.8) has the form*

$$u(x,t) = \int_0^t d\tau \int_{-\infty}^{\infty} G(x-y; t-\tau)\, f(y,\tau) dy + \int_{-\infty}^{\infty} G(x-y;t)\, \phi(y)\, dy \tag{4.6.9}$$

where the function $G(x;t)$ was defined in (4.4.3).

Proof: As we already know from Theorem 4.19 the second term

$$u_2(x,t) = \int_{-\infty}^{\infty} G(x-y;t)\, \phi(y)\, dy$$

in (4.6.9) solves the homogeneous heat equation and satisfies initial condition

$$u_2(x,0) = \phi(x).$$

The first term

$$u_1(x,t) = \int_0^t d\tau \int_{-\infty}^{\infty} G(x-y; t-\tau)\, f(y,\tau) dy$$

clearly vanishes at $t = 0$. Let us prove that it satisfies the inhomogeneous heat equation

$$\frac{\partial u_1}{\partial t} = a^2 \frac{\partial^2 u_1}{\partial x^2} + f(x,t).$$

Denote

$$v(x,t;\tau) = \int_{-\infty}^{\infty} G(x-y; t-\tau)\, f(y,\tau) dy.$$

Like in the Theorem 4.19 we derive that this is a solution to the homogeneous heat equation in x, t depending on the parameter τ. This solution is defined for $t \geqslant \tau$; for $t = \tau$ it satisfies the initial condition

$$v(x,\tau;\tau) = f(x,\tau).$$

Applying the heat operator to the function

$$u_1(x,t) = \int_0^t v(x,t;\tau)\, d\tau$$

one obtains

$$\left(\frac{\partial}{\partial t} - a^2 \frac{\partial^2}{\partial x^2}\right) u_1(x,t) = v(x,t;t) + \int_0^t \left(\frac{\partial}{\partial t} - a^2 \frac{\partial^2}{\partial x^2}\right) v(x,t;\tau)\, d\tau = v(x,t;t) = f(x,t).$$

■

4.6.1 A caveat

Let us reconsider the infinite rod case with zero boundary condition:

$$u_t = u_{xx}, \qquad u(x,0) = 0. \tag{4.6.10}$$

If we do not *assume* that u is in L^1 for $t > 0$ then the uniqueness fails. The following (counter)example is provided by Tychonov.

Let $f(t) = \mathrm{e}^{-\frac{1}{t^2}}$ (and extended to $f(0) = 0$). A calculus exercise shows that f and all derivatives of $f(t)$ at $t = 0$ exist and are zero.

Define

$$u(x,t) := \sum_{n=0}^{\infty} f^{(n)}(t) \frac{x^{2n}}{(2n)!}. \tag{4.6.11}$$

The series is absolutely convergent for all $t > 0, \quad x \in \mathbb{R}$ and we can differentiate under the summation symbol. A direct inspection then shows that $u(x,t)$ solves the DE with zero initial conditions. The key to reconcile the apparent contradiction is that $u(x,t)$ is unbounded in x for any $t > 0$.

Indeed, in deriving the solution we tacitly assumed that $u(x,t)$ is L^1 in x for $t > 0$ (since we took the Fourier transform). Thus our derivation would not apply if we allow $u(x,t)$ to be unbounded in the x–direction.

In general the uniqueness holds if we add the following "boundary conditions"

$$\lim_{x \to \pm\infty} \sup_{t \in (0,T]} |u(x,t)| = 0. \tag{4.6.12}$$

4.7 Exercises for Chapter 4

Exercise 4.35 *Let the function $f(x)$ belong to the class $\mathcal{C}^k(\mathbb{R})$ and, moreover, all the functions $f(x)$, $f'(x)$, ..., $f^{(k)}(x)$ be absolutely integrable on $\mathbb{R}$. Prove that then*

$$\hat{f}(p) = \mathcal{O}\left(\frac{1}{p^k}\right) \quad \text{for} \quad |p| \to \infty. \tag{4.7.1}$$

Exercise 4.36 *Let $\hat{f}(p)$ be the Fourier transform of the function $f(x)$. Prove that $e^{iap}\hat{f}(p)$ is the Fourier transform of the shifted function $f(x+a)$.*

Exercise 4.37 *Find Fourier transforms of the following functions.*

$$f(x) = \Pi_A(x) = \begin{cases} \frac{1}{2A}, & |x| < A \\ 0, & \text{otherwise} \end{cases} \tag{4.7.2}$$

$$f(x) = \Pi_A(x)\cos\omega x \tag{4.7.3}$$

$$f(x) = \begin{cases} \frac{1}{A}\left(1 - \frac{|x|}{A}\right), & |x| < A \\ 0, & \text{otherwise} \end{cases} \tag{4.7.4}$$

$$f(x) = \cos ax^2 \quad \text{and} \quad f(x) = \sin ax^2 \quad (a > 0) \tag{4.7.5}$$

$$f(x) = |x|^{-\frac{1}{2}} \quad \text{and} \quad f(x) = |x|^{-\frac{1}{2}}e^{-a|x|} \quad (a > 0). \tag{4.7.6}$$

Exercise 4.38 *Find the function $f(x)$ if its Fourier transform is given by*

$$\hat{f}(p) = e^{-k|p|}, \quad k > 0. \tag{4.7.7}$$

Exercise 4.39 *Let $u = u(x,y)$ be a solution to the Laplace equation on the half-plane $y \geqslant 0$ satisfying the conditions*

$$\begin{aligned} &\Delta u(x,y) = 0, \quad y > 0 \\ &u(x,0) = \phi(x) \\ &u(x,y) \to 0 \quad \text{as} \quad y \to +\infty \quad \text{for every} \quad x \in \mathbb{R} \end{aligned} \tag{4.7.8}$$

1) Prove that the Fourier transform of u in the variable x

$$\hat{u}(p,y) = \frac{1}{2\pi}\int_{-\infty}^{\infty} u(x,y)e^{-ipx}dx$$

has the form

$$\hat{u}(p,y) = \hat{\phi}(p)e^{-y\,|p|}.$$

Here $\hat{\phi}(p)$ is the Fourier transform of the boundary function $\phi(x)$.

2) Derive the following formula for the solution to the b.v.p. (4.7.8)

$$u(x,y) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{y}{(x-s)^2 + y^2}\phi(s)\,ds. \tag{4.7.9}$$

Exercise 4.40 *Show, by a (formal) use of the Fourier transform and convolution theorem, that*

$$y''(x) - y(x) = f(x) \tag{4.7.10}$$

has a particular solution of the form

$$y_p(x) = -\frac{1}{2}\int_{-\infty}^{\infty} \mathrm{e}^{-|x-s|} f(s)\mathrm{d}s. \tag{4.7.11}$$

Then show directly by differentiation that this expression is a valid solution if $f(x)$ is continuous and absolutely integrable.

Exercise 4.41 *Using Poisson's summation formula, compute:*

$$(\mathbf{a}) \quad \sum_{k=1}^{\infty} \frac{\sin(ak)}{k}; \qquad (\mathbf{b}) \quad \sum_{k=1}^{\infty} \left(\frac{\sin(ak)}{k}\right)^2; \qquad (\mathbf{c}) \quad \sum_{k=0}^{\infty} \frac{1}{a^2+k^2} \quad (a \neq 0). \tag{4.7.12}$$

(d) *Using the same formula, show that the function*

$$\omega(t) := \frac{1}{4}\mathrm{e}^{t} \sum_{n\in\mathbb{Z}} \mathrm{e}^{-\pi n^2 \mathrm{e}^{4t}} \tag{4.7.13}$$

is even: $\omega(t) = \omega(-t)$.
Remark ($\star$) *Consider the cosine transform of ω*

$$\phi(p) := \int_0^{\infty} \omega(t)\cos(tp)\mathrm{d}t.$$

Show that it has only real zeros. If you manage let me know.

Exercise 4.42 *Consider the Heat kernel on $\mathbb{R}$*

$$G(x;t) = \frac{1}{\sqrt{4\pi a^2 t}}\mathrm{e}^{-\frac{x^2}{4a^2 t}} \tag{4.7.14}$$

Why is it obvious that $G(\cdot;t) \star G(\cdot;s) = G(x;s+t)$? (i.e. without computing).

Exercise 4.43 *This exercise gives an alternative proof of Riemann–Lebesgue lemma. Suppose that*

$$\lim_{a\to 0}\int_{\mathbb{R}} \Big| f(x) - f(x+a)\Big| \mathrm{d}x = 0 \tag{4.7.15}$$

(for example: $f \in L^1$ and p.wise continuous). We can show that $\widehat{f}(p) \to 0$ as $|p| \to \infty$ as follows: first show that

$$\widehat{f}(p) = -\int_{\mathbb{R}} f(x)^{-ip(x-\frac{\pi}{p})}\frac{\mathrm{d}x}{2\pi} = -\int_{\mathbb{R}} f\left(x+\frac{\pi}{p}\right)^{-ipx}\frac{\mathrm{d}x}{2\pi}. \tag{4.7.16}$$

Then note that $2\widehat{f}(p) = \int \left(f(x) - f\left(x+\frac{\pi}{p}\right)\right)\mathrm{e}^{-ipx}\frac{\mathrm{d}x}{2\pi}$.

Exercise 4.44 *Solve the problem*

$$(DE): \qquad u_t = a^2 u_{xx} + \frac{1}{\sqrt{4a^2 t}} \mathrm{e}^{-\frac{x^2}{4a^2 t}} \tag{4.7.17}$$

$$(IC): \qquad u(x,0) = 0. \tag{4.7.18}$$

Exercise 4.45 *Let $f(x)$ be p.wise continuous, bounded.*

- *Show that $u(x,t) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} \mathrm{e}^{-s^2} f(x + 2s\sqrt{a^2 t}) \mathrm{d}s$ is the solution of the heat equation $u_t = a^2 u_{xx}$ with IC $u(x,0) = f(x)$.*
- *using Dominated Convergence show that if f is continuous at x_0 then $u(x_0,t) \to f(x_0)$ for $t \to 0_+$. What happens if f is discontinuous at x_0 but both one-sided limits exist?*

Chapter 5

Non homogeneous Heat equation and Sturm–Liouville theory

5.1 Generalization of the homogeneous rod

If we reconsider the derivation of the heat equation in Sect. 4.1 when the thermal conductivity k, the specific heat capacity c and the (linear) density ρ depend on the position, we get a more general PDE

$$c(x)\rho(x)u_t(x,t) = \left(k(x)u_x(x,t)\right)_x . \tag{5.1.1}$$

On a finite (or infinite) rod, the method of separation of variables is still viable; seeking a solution of the form $u(x,t) = T(x)X(x)$ gives promptly

$$\frac{T_t(t)}{T(t)} = \frac{(kX_x)_x}{c\rho X} \tag{5.1.2}$$

and hence one is lead to studying the second order ODE

$$(kX_x)_x = \lambda c\rho X. \tag{5.1.3}$$

This equation fall in the general framework of Sturm–Liouville theory which we set out to analyze.

Definition 6 *The Sturm–Liouville problem consists in finding solutions of the ODE*

$$(DE) \qquad (P(x)f_x(x))_x + Q(x)f(x) = -\lambda R(x)f(x), \quad x \in I \tag{5.1.4}$$

where I is an interval and $P \in \mathcal{C}^1(I)$, $Q, R \in \mathcal{C}^0(I)$ are given functions; the function $P(x)$ is a positive function on I, $P(x) > 0$. If the interval I is bounded $I = [a,b]$ then the problem is supplemented by general boundary value conditions (BVCs)

$$(BC) \qquad \alpha_1 f(a) + \alpha_2 P(a)f(a) = 0 \qquad \beta_1 f(b) + \beta_2 P(b)f(b) = 0. \tag{5.1.5}$$

where the constants α_j, β_j (not identically zero) are part of the data of the problem.

The values of the constant λ for which there exist solutions of (5.1.4) + (5.1.5) are called the **eigenvalues** *of the Sturm-Liouville problem.*
If P, R are strictly positive on I then the problem is called **regular**.
If $P > 0$ on (a, b) and $R \geqslant 0$ and $P(a) = P(b) = 0$ then the problem is called **singular**.
If we are in the finite-interval case $I = [a, b]$ and impose boundary conditions $f(a) = f'(b)$ and $f'(a) = f'(b)$, then we speak of the **periodic** *problem.*

The conditions (BC) include both Dirichlet and Neumann like conditions, depending on the choice of the values of the constants α_j, β_j, $j = 1, 2$.

If P, Q, R are constant functions, we are reduced to the problem we have already studied in the context of the wave equation.

Definition 7 *The form of the Sturm–Liouville equation (5.1.4) is called self–adjoint form.*

Any second order ODE of the form

$$\mathcal{H}[\varphi] = P_2(x)\varphi''(x) + P_1(x)\varphi'(x) + P_0(x)\varphi(x) = 0 \tag{5.1.6}$$

can be recast in the form (5.1.22) by a simple change of dependent variable; indeed, setting $\varphi(x) = \mu(x)f(x)$ for a function $\mu(x)$ to be determined, the equation (5.1.6) becomes

$$\begin{aligned} &P_2\Big(\mu f'' + 2\mu' f' + \mu'' f\Big) + P_1\Big(\mu f' + \mu' f\Big) + P_0\mu f = \\ &= \mu P_2 f'' + \big(2P_2\mu' + P_1\mu\big)\, f' + (P_2\mu'' + P_1\mu' + P_0\mu)f = 0 \end{aligned} \tag{5.1.7}$$

To identify with the equation $\mathcal{L}[f] = 0$ and $\mathcal{L}$ given by (5.1.22) we need to impose $P = P_2\mu$ and

$$(P_2\mu)' = 2P_2\mu' + P_1\mu \quad \Rightarrow \quad P_2\mu' + (P_1 - P_2')\mu = 0 \quad \Rightarrow \quad \mu = P_2 \exp\left(-\int \frac{P_1}{P_2}\mathrm{d}x\right). \tag{5.1.8}$$

Example 5.1 *The (parametric)* **Bessel equation**

$$x^2y'' + xy' + (\lambda^2x^2 - m^2)y = 0 \tag{5.1.9}$$

is recast in the self-adjoint form (5.1.4) by dividing by x:

$$xy'' + y' + \left(\lambda^2 x - \frac{m^2}{x}\right) y = 0 \tag{5.1.10}$$

so that $P = x, \quad Q = -\frac{m^2}{x}, \quad R = x$.

Example 5.2 *The* **Legendre differential equation**

$$y'' - \frac{2x}{1 - x^2}y' + \frac{\mu}{1 - x^2}y = 0 \tag{5.1.11}$$

is equivalently written

$$\underbrace{(1-x^2)y'' - 2xy'}_{\left((1-x^2)y'\right)'} + \mu y = 0 \tag{5.1.12}$$

and hence it is of the form (5.1.4) with $P = (1-x^2),\ Q = 0, \quad R = 1.$

Example 5.3 *The* **Chebyshev equation**

$$(1-x^2)y'' - xy' + n^2 y = 0 \tag{5.1.13}$$

can be recast in the form (5.1.4) by dividing by $\sqrt{1-x^2}$ *so that*

$$\underbrace{\sqrt{1-x^2}y'' - \frac{x}{\sqrt{1-x^2}}y'}_{\left(\sqrt{1-x^2}y'\right)'} + \frac{n^2}{\sqrt{1-x^2}}y = 0 \tag{5.1.14}$$

so that now $P = \sqrt{1-x^2}, \quad Q = 0, \quad R = \frac{1}{\sqrt{1-x^2}}.$

Example 5.4 Consider the SL problem

$$(DE) \quad f'' + \lambda f = 0, \quad x \in [0, L], \quad L < \frac{\pi}{2} \tag{5.1.15}$$

$$(BC) \quad f(0) - f'(0) = 0; \quad f(L) + f'(L) = 0. \tag{5.1.16}$$

Let us find the spectrum (eigenvalues) of this problem.

For $\lambda \leqslant 0$ it is easy (**exercise**) to see that there is only the trivial solution. Let $\lambda > 0$: then the solution of the DE is

$$f(x) = A\cos(x\sqrt{\lambda}) + B\sin(x\sqrt{\lambda}). \tag{5.1.17}$$

Imposing the BCs gives the system for A, B

$$\begin{cases} A - \sqrt{\lambda}B = 0 \\ A\cos(L\sqrt{\lambda}) + B\sin(L\sqrt{\lambda}) + \sqrt{\lambda}\Big(-A\sin(L\sqrt{\lambda}) + B\cos(L\sqrt{\lambda})\Big) = 0 \end{cases} \tag{5.1.18}$$

The problem has nontrivial solution if and only if the matrix is degenerate:

$$\det\begin{bmatrix} 1 & -\sqrt{\lambda} \\ \cos(q) - \sqrt{\lambda}\sin(q) & \sin(q) + \sqrt{\lambda}\cos(q) \end{bmatrix} = 0, \quad q := L\sqrt{\lambda}. \tag{5.1.19}$$

Thus we have

$$(1-\lambda)\sin(L\sqrt{\lambda}) + 2\sqrt{\lambda}\cos(L\sqrt{\lambda}) = 0 \tag{5.1.20}$$

Since $L < \frac{\pi}{2}$ the value $\lambda = 1$ is not a solution; the eigenvalues are given by the implicit equation

$$\tan(L\sqrt{\lambda}) = \frac{2\sqrt{\lambda}}{\lambda - 1}. \tag{5.1.21}$$

At this point one is usually reduced to a graphical analysis; considering that $L < \frac{\pi}{2}$ the typical spectrum consists of the intersection of the graphs of the two functions as in the Figure 5.1

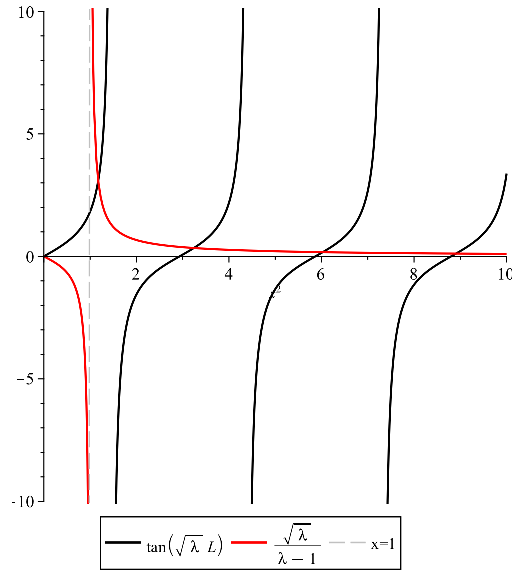

Figure 5.1: The spectrum of the Sturm-Liouville problem in Example 5.4. The horizontal axis is the $\sqrt{\lambda}$ axis. There is an eigenvalue in each interval $\sqrt{\lambda} \in \left(k\frac{\pi}{2L}, (k+2)\frac{\pi}{2L}\right)$, $k = 1, 2, \ldots$.

5.1.1 Spectral properties

Denote by $\mathcal{L}$ the Sturm–Liouville linear differential operator

$$\mathcal{L}[f] = \frac{\mathrm{d}}{\mathrm{d}x}\left(P(x)\frac{\mathrm{d}f}{\mathrm{d}x}\right) + Q(x)f(x). \tag{5.1.22}$$

Definition 8 *Given an arbitrary linear differential operator $\mathcal{H}$ of the form*

$$\mathcal{H} = \sum_{j=0}^{n} P_j(x)\frac{\mathrm{d}^j}{\mathrm{d}x^j} \tag{5.1.23}$$

the **formal adjoint** *operator is defined to be*

$$\mathcal{H}^\star = \sum_{j=0}^{n}\left(-\frac{\mathrm{d}}{\mathrm{d}x}\right)^j P_j(x). \tag{5.1.24}$$

For example, the adjoint of $\mathcal{H}$ (5.1.6) is

$$\mathcal{H}^\star(\psi) = \frac{\mathrm{d}^2}{\mathrm{d}x^2}(P_2\psi) - \frac{\mathrm{d}}{\mathrm{d}x}(P_1\psi) + P_0\psi = P_2\psi'' + (2P_2' - P_1')\psi' + (P_0 + P_2'' - P_1')\psi. \tag{5.1.25}$$

The relevance of this notion to our discussion is the following:

Proposition 5.5 *The Sturm–Liouville operators $\mathcal{L}$ are formally self-adjoint: $\mathcal{L} = \mathcal{L}^\star$.*

Proof. We have

$$\mathcal{L}[f] = Pf'' + P'f' + Qf. \tag{5.1.26}$$

Thus the formal adjoint is

$$\begin{aligned}\mathcal{L}^\star[g] = (Pg)'' - (P'g)' + Qg = Pg'' + \cancel{2}P'g' + \cancel{P''g} - \cancel{P''g} - \cancel{P'g'} + Qg = \\ = Pg'' + P'g' + Qg = \mathcal{L}[g].\end{aligned} \tag{5.1.27}$$

This completes the proof. ■

The notion is intimately related to "integration by parts". To explain this remark we prove

Proposition 5.6 *Let $I = [a, b]$ be a finite interval and consider the space*

$$\mathcal{C}^2_{BC} := \Big\{ f \in \mathcal{C}^2(I) : \quad f \textit{ satisfies the BCs } (5.1.5) \Big\} \tag{5.1.28}$$

Then, for every $f, g \in \mathcal{C}^2_{BC}$ we have

$$\int_I f\mathcal{L}[g]\mathrm{d}x = \int_I g\mathcal{L}[f]\mathrm{d}x. \tag{5.1.29}$$

Namely, $\mathcal{L}$ is symmetric on the domain $\mathcal{C}^2_{BC}$.

Proof. This is a direct computation:

$$\begin{aligned}\int_a^b g\Big((Pf')' + Qf\Big)\mathrm{d}x = gPf'\Big|_a^b + \int_a^b \Big(- g'Pf' + Qgf\Big)\mathrm{d}x = \\ = gPf'\Big|_a^b - g'Pf\Big|_a^b + \int_a^b \Big((g'P)'f + Qgf\Big)\mathrm{d}x = R + \int_I f\mathcal{L}[g]\mathrm{d}x\end{aligned} \tag{5.1.30}$$

where R is the contribution of the boundary terms. Because of the conditions (5.1.5) we have $g'(b)P(b) = -\frac{\beta_1}{\beta_2}g(b)$ and similarly $f'(b)P(b) = -\frac{\beta_1}{\beta_2}f(b)$ (we assume $\beta_2 \neq 0$ and leave the case $\beta_2 = 0$ as exercise) and similarly for the values at $x = a$. Thus

$$R = \big(gPf' - g'Pf\big)\Big|_a^b = \cancel{-\frac{\beta_2}{\beta_1}g(b)f(b)} + \cancel{\frac{\beta_2}{\beta_1}g(b)f(b)} + \cancel{\frac{\alpha_2}{\alpha_1}g(b)f(b)} - \cancel{\frac{\alpha_2}{\alpha_1}g(b)f(b)} = 0 \tag{5.1.31}$$

Thus we have proved the statement. ■

The Proposition 5.6 has the following simple but important consequence.

Corollary 5.7 [**1**] *The spectrum of a regular Sturm-Liouville problem*

$$\mathcal{L}[f] = -\lambda R(x) f(x) \tag{5.1.32}$$

with any BC (5.1.5) is real.
[**2**] *Eigenfunctions with distinct eigenvalues are orthogonal with respect to the standard inner product in $L^2(I, R(x)\mathrm{d}x)$.*

Proof. [**1**] Recall that $R(x) > 0$ and that we are looking for a nontrivial solution of the equation (5.1.32) within $\mathcal{C}^2_{BC}$. Suppose that λ is an eigenvalue. Since P, Q, R are real–valued functions, we can take the complex conjugate of the equation (5.1.32) and obtain that $\overline{f}$ satisfies

$$\mathcal{L}[\overline{f}] = -\lambda R\overline{f}. \tag{5.1.33}$$

Now we multiply by f and integrate over I;

$$\int_a^b f\mathcal{L}[\overline{f}]\mathrm{d}x = -\int_a^b \overline{\lambda} R|f|^2 \mathrm{d}x \tag{5.1.34}$$

Using that $\mathcal{L}$ is symmetric on the domain we obtain:

$$\int_a^b \overline{f}\mathcal{L}[f]\mathrm{d}x = -\int_a^b \overline{\lambda} R|f|^2 \mathrm{d}x \tag{5.1.35}$$

Repeating the argument with the role of f and $\overline{f}$ interchanged and subtracting we obtain

$$0 = \int_a^b (f\mathcal{L}[\overline{f}] - \overline{f}\mathcal{L}[f])\mathrm{d}x = \int_a^b (\overline{\lambda} - \lambda) R(x)|f|^2 \mathrm{d}x \tag{5.1.36}$$

Since $R(x) > 0$ (strictly) on I, we conclude that $\lambda = \overline{\lambda}$ is real.
[**2**] Let $\lambda \neq \mu$ be two eigenvalues and f, g the respective eigenfunctions. Then

$$\lambda \int f(x)\overline{g}(x)R(x)\mathrm{d}x = -\int_I \mathcal{L}[f]\overline{g}\mathrm{d}x \overset{\text{symmetry}}{=} -\int_I f\mathcal{L}[\overline{g}]\mathrm{d}x \overset{\mu=\overline{\mu}}{=} \int_I \mu f\overline{g}R\mathrm{d}x \tag{5.1.37}$$

Thus, taking the difference of the two sides of the equation we get:

$$(\lambda - \mu)\int_I \mu f\overline{g}R\mathrm{d}x = 0. \tag{5.1.38}$$

Since $\lambda \neq \mu$ by assumption, the two functions f, g must be orthogonal. ■

Theorem 5.8 (Simplicity of the spectrum) *Consider a regular Surm-Liouville problem with BC as in (5.1.5). Then the spectrum is simple, namely, if f_1, f_2 are eigenfunctions with the same eigenvalue, then they are proportional to each other.*

Proof. Consider the function

$$W(x) := f_2'(a)f_1(x) - f_1'(a)f_2(x). \tag{5.1.39}$$

Since it is a linear combination of f_1, f_2, it is also an eigenfunction with the same eigenvalue. Assume that $(f_1'(a))^2 + (f_2'(a))^2 \neq 0$. Then W solves

$$\mathcal{L}[W] = -\lambda R W \ , \qquad W(a) = W'(a) = 0. \tag{5.1.40}$$

By the theorem of existence and uniqueness for linear ODEs, we conclude $W(x) \equiv 0$. If $(f_1'(a))^2 + (f_2'(a))^2 \neq 0$ then it means that both f_j have zero derivative at $x = a$ and hence they are linearly dependent. ■

5.1.2 Definite Sturm–Liouville operators

Theorem 5.9 *Consider a regular Sturm–Liouville problem*

$$(Pf')' + Qf = -\lambda R f \tag{5.1.41}$$

with $Q \leqslant 0$ and with BC as in (5.1.5); we assume now that the signs of the coefficients α_j, β_j satisfy

$$\alpha_1\alpha_2 \leqslant 0, \quad \beta_1\beta_2 \geqslant 0. \tag{5.1.42}$$

Then the eigenvalues are non-negative. If $\lambda = 0$ is an eigenvalue, then necessarily $Q(x) = 0$ and $\alpha_1 = \beta_1 = 0$.

Proof. Let λ be an eigenvalue and f the corresponding eigenfunction, which we assume real without loss of generality (the ODE is real and the BC are real, so both real and imaginary parts of a solutions are solutions in their own regard). Then (we write $Q = -|Q|$ to emphasize the assumption on its sign)

$$\begin{aligned}-\lambda \int_I R|f|^2 \mathrm{d}x = \int_I \mathcal{L}[f] f \mathrm{d}x = \int_I \Big((Pf')' - |Q| f \Big) f \mathrm{d}x = \\ = Pf'f \Big|_a^b - \int_a^b \Big(P(f')^2 + |Q| f^2 \Big) \mathrm{d}x\end{aligned} \tag{5.1.43}$$

The boundary terms are non-positive because $P(a) > 0 < P(b)$ (the operator is regular) and $\alpha_1 f(a) = -\alpha_2 f'(a)$ so that $f(a)f'(a) \geqslant 0$. Similarly $f(b)f'(b) \leqslant 0$ and thus

$$Pf'f \Big|_a^b \leqslant 0. \tag{5.1.44}$$

Therefore each term on the right side of (5.1.43) is non-positive and thus the eigenvalue λ is $\geqslant 0$. The only possibility to have $\lambda = 0$ is that $Q = 0$, $f' = 0$ and (hence), $\alpha_1 = \beta_1 = 0$. ■

5.1.3 Sturm comparison theorem

Consider the following motivating example

Example 5.10 Suppose that $\lambda_2 > \lambda_1 > 0$ and y_1, y_2 are solutions of $y_j'' = -\lambda_j y_j$ respectively. We now see that between any two consecutive zeros of y_1 there is a zero of y_2. Indeed $y_1 = A\sin(\sqrt{\lambda_1}x + \phi)$ and $y_2 = B\sin(\sqrt{\lambda_2}x + \delta)$. The period of y_1 is longer than the period of y_2 and thus between two zeroes of y_1 there is (at least) one zero of y_1.

The above example is the motivation for the following theorem

Theorem 5.11 (Sturm comparison theorem) *Let $\mathcal{L}[f] = (Pf')' + Qf$ with $P > 0$ and $P', Q \in \mathcal{C}^0([a,b])$. Suppose that f_1, f_2 are two nontrivial solutions of the problems*

$$(Pf_1')' + (\lambda_1 R_1(x) + Q(x))f_1(x) = 0\ , \qquad (Pf_2')' + (\lambda_2 R_2(x) + Q(x))f_2(x) = 0. \tag{5.1.45}$$

with $R_j \in \mathcal{C}^0(I)$. Suppose that $\lambda_2 R_2(x) \geqslant \lambda_1 R_1(x)$ for $x \in I$. Let $a \leqslant z < w \leqslant b$ be two zeros of f_1 and assume that the inequality $\lambda_2 R_2 \geqslant \lambda_1 R_1$ is strict somewhere in between. Then between two zeros $a \leqslant z < w \leqslant b$ of f_1 there is at least one zero of f_2.

Remark 5.12 *The theorem is stated in this form for convenience later on (only $\lambda_j R_j$ enters). Note also that in typical examples $R_j(x)$ are analytic functions and hence the zeroes are isolated and the inequality strict almost everywhere.*

Proof. Let z, w be the two consequtive zeroes of y_1 and consider the following chain of equalities

$$\int_z^w \left(\lambda_2 R_2 - \lambda_1 R_1\right) f_1 f_2 \mathrm{d}x = \int_z^w (\mathcal{L}[f_1]f_2 - \mathcal{L}[f_2]f_1)\mathrm{d}x \overset{\text{b.p.}}{=} P(f_1'f_2 - f_1 f_2')\Big|_z^w . \tag{5.1.46}$$

Since f_1 vanishes at the endpoints we are left with $Pf_1'f_2\Big|_z^w$ Since f_1 has the same sign within the interval $[z,w]$ (and is not identically zero) it follows that $f_1'(z)f_1'(w) < 0$ (the two derivatives have opposite signs). Note that $f_1'(z) \neq 0 \neq f_1'(w)$ (**why?**, exercise).

For clarity and without loss of generality we can assume that $f_1 > 0$ in (z,w); then $f_1'(z) > 0 > f_1'(w)$. Thus the right side of (5.1.46) is $f_2(w)P(w)f_1'(w) - f_2(z)f_1'(w)P(w) = -f_2(w)P(w)|f_1'(w)| - f_2(z)f_1'(w)P(w)$. Suppose that f_2 does not have any zero within (z,w) (w.l.o.g $f_2 > 0$): then $f_2(w) \geqslant 0 \leqslant f_2(z)$ and we have thus found the right side of (5.1.46) to be non-positive. But this is in contradiction with the fact that the left side is *strictly* positive (it can't be zero because of the assumptions on $\lambda_j R_j$). ■

5.1.4 Existence of eigenvalues

For a general regular Sturm–Liouville problem there are infinitely many (countable) eigenvalues. In order not to obfuscate the idea of the proof we consider a special case

Theorem 5.13 *Let $\mathcal{L}[f] = f'' + Q(x)f(x)$ and $R(x) > 0$ and continuous on $[a, b]$. Consider the problem*

$$f(x)'' + Q(x)f(x) = -\lambda R(x)f(x) \tag{5.1.47}$$

$$f(a) = 0 = f(b) \tag{5.1.48}$$

Let $\Phi(x; \lambda)$ be the solution of

$$\Phi'' + \Big(Q + \lambda R\Big)\Phi = 0 \tag{5.1.49}$$

$$\Phi(a; \lambda) = 0, \quad \Phi'(a; \lambda) = 1. \tag{5.1.50}$$

Then $\Phi(x; \lambda)$ is jointly continuous and the problem (5.1.48) has infinitely many eigenvalues, bounded from below $\lambda_1 < \lambda_2 < \dots$, with $\lim_{n\to\infty} \lambda_n = +\infty$. Furthermore the eigenfunctions f_n have exactly n nodes; the between two nodes of f_n there is exactly one node of f_{n+1} (interlacing).

Proof. The proof exploits the Sturm comparison Theorem 5.11. Consider the function $\Phi(b; \lambda)$ as a function of λ (*Evans' function*). Then we claim (**why?**, exercise) that its zeros are precisely the eigenvalues.

First off we claim that $\Phi(x; \lambda)$ cannot have infinitely many zeros in $[a, b]$, see Exercise 5.23.

We also know that $\Phi(x; \lambda)$ is continuous in λ in the sup–norm (hence, jointly continuous) by the general theory of (global) existence and uniqueness of solutions to linear ODEs.

Next, if $x_0 \in [a, b]$ is a zero of Φ, then $\Phi'(x_0; \lambda) \neq 0$ (**why?**) and hence necessarily $\Phi(x; \lambda)$ changes sign at x_0. Thus if λ_0 is a value such that $\Phi(b; \lambda_0) \neq 0$ then the number of zeroes in (a, b) is constant in $(\lambda_0 - \epsilon, \lambda_0 + \epsilon)$ for sufficiently small ϵ. This reasoning shows that the number of zeroes jumps by one exactly for λ_j such that $\Phi(b, \lambda_j) = 0$.

We know from the Comparison Theorem that if $\Phi(x; \lambda)$ has a root at x_0 then for $\widetilde{\lambda} > \lambda$ there must be a root at $\widetilde{x}_0 < x_0$. In fact, as an application of the theorem of continuity with respect to parameters, one could show that any root of $\Phi(x; \lambda)$ is a continuous decreasing function of λ; the fact that it is decreasing is precisely a direct consequence of the Comparison Theorem.

Combining this with the previous observations, we see that all the (finitely many) roots of $\Phi(x; \lambda)$ in (a, b) move to the left (but never cross $x = a$ because $\Phi'(a, \lambda) = 1$).

It remains to show that zeroes do occur; to this end, let us compare $\Phi(x; \lambda)$ with the solution of

$$F'' + k^2 F = 0, \quad F(a; k) = 0, \quad F'(a; k) = 1, \quad \Rightarrow \qquad F(x; k) = \frac{1}{k}\sin(k(x - a)) \tag{5.1.51}$$

Let $k > 0$ be fixed and choose λ sufficiently large so that

$$Q(x) + \lambda R(x) > k^2 \quad \forall x \in [a, b]. \tag{5.1.52}$$

This is possible because $\inf R = \min R > 0$ since R is continuous on $[a, b]$. Then F plays the rôle of f_1 in the comparison theorem 5.11 and Φ of f_2; between two consecutive zeroes of F there is at least one of Φ. Clearly, for k large than $\frac{\pi}{b-a}$ there is at least one zero of F in (a, b) and thus also one of Φ. As k grows, the number of zeroes of F grows as well, and thus the number of zeroes of $\Phi(x;\lambda)$ in (a,b)

$$Z(\lambda) := \sharp\{x \in (a,b) : \quad \Phi(x,\lambda) = 0\} \tag{5.1.53}$$

is a monotonic, integer–valued function that grows without bounds. By the discussion above, the points of discontinuity of $Z(\lambda)$ are the eigenvalues of the SL problem. Since we have shown that $Z(\lambda)$ is unbounded, there must be infinitely many eigenvalues. ∎

Completeness of eigenfunctions

The important question we address now is the completeness of the eigenfunction set for a regular Sturm–Liouville problem on an interval I (possibly infinite).

$$(Pf')' + Qf = -\lambda Rf; \quad P, P', Q, R \in \mathcal{C}^0(I), \ P > 0, \ R > 0. \tag{5.1.54}$$

The boundary conditions can be set in many different ways without changing the properties of completeness. We will stick with (5.1.5) in the case of finite intervals (although one can impose periodic boundary conditions as well). If I is (semi)-infinite we require that the solutions of the problem be square-summable (L^2).

Let

$$\mathcal{H} := \overline{\mathcal{C}^2_{BC}(I)} \tag{5.1.55}$$

where the bar on top denotes the closure with respect to the norm of $L^2(I, R(x)\mathrm{d}x)$. We have seen that there are infinitely many eigenvalues bounded from below which we label $\lambda_0 < \lambda_1 < \ldots$ and correspondingly eigenfunctions $f_0, f_1, \ldots,$.

Without loss of generality we assume that $f_n(x)$ are real–valued and that $\int_I |f_n|^2 R\mathrm{d}x = 1$ (normalized).

We now want to give a (sketch of) proof of the following theorem

Theorem 5.14 (Completeness of eigenfunctions) *The eigenfunctions $\{f_n\}_{n\in\mathbb{N}}$ form a complete orthonormal set; namely, for every $\phi \in \mathcal{H}$ we have*

$$\lim_{N\to\infty} \left\| \phi - \sum_{j=0}^{N-1} \langle \phi, f_n \rangle f_n \right\| = 0. \tag{5.1.56}$$

where the inner product and norm are in $L^2(I, R\mathrm{d}x)$.

Proof. Let us define (**note** the different use of the symbol $\mathcal{L}$ in this proof, as opposed to earlier in the chapter)

$$\mathcal{L}[f] := \frac{-1}{R(x)} \left[(Pf')' + Qf\right] \tag{5.1.57}$$

In particular now the eigenfunctions satisfy simply $\mathcal{L}[f_n] = \lambda_n f_n$ with $\lambda_n \to +\infty$. For every $\phi \in \mathcal{C}^2_{BC}$ be arbitrarily chosen we use the **Rayleigh quotient**

$$\mathcal{R}[\phi] := \frac{\langle \phi, \mathcal{L}[\phi] \rangle}{\langle \phi, \phi \rangle} \tag{5.1.58}$$

Define

$$W_N := \mathrm{Span}\{f_0, \ldots, f_{N-1}\}. \tag{5.1.59}$$

The eigenvalues can be expressed by the following variational problems (Courant–Fisher)

$$\lambda_0 = \min_{\phi \in \mathcal{H} \setminus \{0\}} \mathcal{R}[\phi]. \tag{5.1.60}$$

$$\lambda_j = \min_{\phi \in W_j^{\perp} \setminus \{0\}} \mathcal{R}[\phi]. \tag{5.1.61}$$

Now fix $\phi \in \mathcal{C}^2_{BC}$ and define

$$\phi_N^{\perp} := \phi - \sum_{j=0}^{N-1} \langle \phi, f_j \rangle f_j \in W_N^{\perp}. \tag{5.1.62}$$

We assume that $\phi_N^{\perp}$ is nonzero for every $N \in \mathbb{N}$ (otherwise there is nothing to prove). Because of the variational property of the eigenvalues we have

$$\mathcal{R}[\phi_N^{\perp}] \geqslant \lambda_N \tag{5.1.63}$$

Using the orthonormality of the eigenfunctions one finds

$$\langle \phi_N^{\perp}, \mathcal{L}[\phi_N^{\perp}] \rangle = \langle \phi, \mathcal{L}[\phi] \rangle - \sum_{j=0}^{N-1} |c_j|^2 \lambda_j. \tag{5.1.64}$$

Therefore we have

$$\frac{1}{\lambda_N} \left(\langle \phi, \mathcal{L}[\phi] \rangle - \sum_{j=0}^{N-1} |c_j|^2 \lambda_j \right) \geqslant \|\phi_n^{\perp}\|^2 \tag{5.1.65}$$

where c_j are the Fourier coefficients in the orthonormal system: $c_j = \langle \phi, f_n \rangle$.

Recall that $\lambda_N \to +\infty$ and hence the eigenvalues are positive from some N_0 onwards. This allows us to write

$$\frac{1}{\lambda_N} \left(\langle \phi, \mathcal{L}[\phi] \rangle - \sum_{j=0}^{N_0} |c_j|^2 \lambda_j \right) \geqslant \frac{1}{\lambda_N} \left(\langle \phi, \mathcal{L}[\phi] \rangle - \sum_{j=0}^{N-1} |c_j|^2 \lambda_j \right) \geqslant \|\phi_n^{\perp}\|^2 \tag{5.1.66}$$

The numerator in the left side of (5.1.66) is independent of N and thus the left side tends to zero since $\lambda_N \to \infty$. This way we have proved by the squeeze theorem that

$$\lim_{N \to \infty} \|\phi_N^{\perp}\|^2 = 0 \tag{5.1.67}$$

which proves the completeness. ■

5.2 Examples

The above theorems allow generalizations of the ordinary Fourier (series) analysis. To see this consider the simple SL problem

$$f'' + \lambda f = 0, \quad f(0) = 0 = f(L). \tag{5.2.1}$$

We know that the eigenfunctions are $\sin(n\pi/Lx)$ with eigenvalues $\lambda_n = (n\pi/L)^2$. The completeness that we proved directly in the case of Fourier series follows from the general Sturm-Liouville theory.

There are other cases of class of functions of great importance in applications and we report on some of them.

For each SL problem one obtains a different "Fourier series" decomposition theorem where the role of $\sin(nx)$ is played by the eigenfunctions of the SL problem.

5.2.1 Separation of variables in polar and spherical coordinates

In solving the heat or wave or Schrödinger equations:

$$\begin{aligned}
(W): &\qquad u_{tt} = c^2 \triangle u \\
(H): &\qquad u_t = k^2 \triangle u \\
(S): &\qquad i\hbar \Psi_t = \left(-\frac{\hbar^2}{2m}\triangle + V(\vec{x})\right)\Psi
\end{aligned} \tag{5.2.2}$$

it is convenient to rewrite them in different coordinate systems and seek factorized solutions, especially depending on the geometry of the boundary conditions (and the group of invariance of $V(\vec{x})$, called the "potential" of the Schrödinger equation).

The separation of variables between the time t and the space variables immediately brings about the Helmholtz equation $(\triangle + \lambda^2)u = 0$. In turns this equation needs to be solved according to the supplementary boundary conditions. In the case of the Schrödinger equation we get the "time independent" Schrödinger equation

$$\left(-\frac{\hbar^2}{2m}\triangle + V(\vec{x})\right)\psi(\vec{x};\lambda) = \lambda^2 \psi(\vec{x};\lambda). \tag{5.2.3}$$

5.2.2 Spherical coordinates

Exercise 5.15 *Show that in $\mathbb{R}^3$ the Laplace operator in the spherical coordinates $(r, \phi, \theta) \in \mathbb{R}_+ \times [0, \pi] \times [0, 2\pi)$ reads as follows*

$$\partial_x^2 + \partial_y^2 + \partial_z^2 = \frac{1}{r^2}\partial_r\left(r^2 \partial_r\right) + \frac{1}{r^2}\overbrace{\left(\frac{1}{\sin\phi}\partial_\phi\left(\sin\phi\,\partial_\phi\right) + \frac{1}{\sin^2\phi}\partial_\theta^2\right)}^{\triangle_{S^2}} \tag{5.2.4}$$

where

$$x = r\sin\phi\cos\theta\ , \qquad y = r\sin\phi\sin\theta\ , \qquad z = r\cos\phi. \tag{5.2.5}$$

and $\triangle_{S^2}$ indicates the Laplace operator on the sphere with the induced Riemannian metric.

Thus the Helmholtz equation can be solved by separation of variables $u(r, \phi, \theta) == R(r)Y(\phi, \theta) = R(r)\Phi(\phi)\Theta(\theta)$. In the case of the Schrödinger equaton with a potential V that depends only on the distance r we the equation can also be effectively separated. In either cases we get something of the form

$$\frac{(r^2 R')'}{r^2 R} + \frac{1}{r^2}\frac{\triangle_{S^2} Y}{Y} = \frac{\left(r^2 R'\right)'}{r^2 R(r)} + \frac{1}{r^2}\frac{(\sin\phi\,\Phi')'}{\Phi\sin\phi} + \frac{1}{r^2\sin^2\phi}\frac{\Theta''}{\Theta} = -\lambda^2 - V(r) \tag{5.2.6}$$

Thus we need to find first off the **eigenfunctions/eigenvalues** of the Laplace operator on the sphere S^2, called **spherical harmonics**:

$$\triangle Y_{\ell m}(\phi, \theta) = \frac{\partial_\phi\Big(\sin\phi(\partial_\phi Y_{\ell,m})\Big)}{\sin\phi} + \frac{1}{\sin^2\phi}\partial_\theta^2 Y_{\ell,m} = -\ell(\ell+1)Y_{\ell,m}(\phi,\theta) \tag{5.2.7}$$

where we anticipate that the spectrum of consists of the numbers $-\ell(\ell+1)$ for $\ell = 0, 1, 2, \dots$ and $m = -\ell, \dots, \ell$ labels the eigenfunctions with the same eigenvalue.

This equation can be further factorized $Y_{\ell,m}(\phi,\theta) = \Phi_{\ell,m}(\phi)\Theta_m(\theta)$ and we are lead to three equations (where the constants of separation of variables are written in this form for later convenience)

$$\frac{\mathrm{d}^2 R}{\mathrm{d}r^2} + \frac{2}{r}\frac{\mathrm{d}R}{\mathrm{d}r} + \left(V(r) - \frac{\ell(\ell+1)}{r^2} - \lambda^2\right)R = 0 \tag{5.2.8}$$

$$\frac{\mathrm{d}}{\mathrm{d}x}\left((1-x^2)\frac{\mathrm{d}P_{\ell m}}{\mathrm{d}x}\right) + \left(\ell(\ell+1) - \frac{m^2}{1-x^2}\right)P_{\ell m} = 0 \qquad (x = \cos\phi \in [-1,1]) \tag{5.2.9}$$

$$\Theta'' + m^2\Theta = 0 \tag{5.2.10}$$

The continuity and boundedness of the solution implies that Θ must be periodic and $P_{\ell,m}(x)$ should be bounded at $x = \pm 1$. This implies immediately that $m = 0, 1, 2, \dots$ and $\Theta = A_m\cos(m\theta) + B_m\sin(m\theta)$.

The values of the parameter $\ell(\ell+1)$ are determined by the boundary conditions (the equation (5.2.9) is an *irregular* Sturm-Liouville problem because the function K of the general form (5.1.4) vanishes at the endpoints). We anticipate that $\ell = 0, 1, \dots$ turns out to be a nonnegative integer.

Remark 5.16 *There is a simple way of describing spherical harmonics in $\mathbb{R}^n$. In general it can be shown that $\triangle = r^{1-n}\partial_r(r^{n-1}\partial_r) + \frac{1}{r^2}\triangle_{S^{n-1}}$.*

Take a **homogeneous** *polynomial $P(\vec{x})$ of degree ℓ that satisfies $\triangle P = 0$; since $r\partial_r P(\vec{x}) = \ell P(\vec{x})$ we see that*

$$0 = \triangle P = \ell r^{1-n}\partial_r\left(r^{n-2}P\right) + \frac{1}{r^2}\triangle_{S^{n-1}}P = \frac{\ell(\ell+n-2)}{r^2}P + \frac{1}{r^2}\triangle_{S^{n-1}}P \tag{5.2.11}$$

so that we conclude that any harmonic polynomial, restricted to the sphere $r = 1$ gives an eigenfunction of the Laplace operator of the sphere with eigenvalue $-\ell(\ell+n-2)$. It can be shown that all spherical harmonics come in this form.

Furthermore (Miles-Williams, "A basic set of Homogeneous Harmonic Polynomials in k variables", Proc. AMS '55) one can actually construct such harmonic polynomials.

Radial equation in spherical coordinates

The radial equation (5.2.8) depends on $V(r)$; we focus now on that equation for $V(r) = 0$.

Proposition 5.17 *The solutions of the radial equation on $r > 0$ (*radial Bessel functions*)*

$$\frac{\mathrm{d}^2 R}{\mathrm{d}r^2} + \frac{2}{r}\frac{\mathrm{d}R}{\mathrm{d}r} + \left((\pm)\lambda^2 - \frac{\ell(\ell+1)}{r^2}\right) R = 0 \tag{5.2.12}$$

are of the form

$$r^\ell \left(\frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}r}\right)^\ell R_0^\pm(\lambda r) \tag{5.2.13}$$

where $R_0^\pm$ are the solutions of the equation (5.2.12) for $\ell = 0$

$$R_0^{(+)} = \frac{a_0\cos(r) + b_0\sin(r)}{r} \qquad R_0^{(-)} = \frac{a_0\cosh(r) + b_0\sinh(r)}{r} \tag{5.2.14}$$

Proof. The proof for $\ell = 0$ is a direct inspection (the change of variables $R(r) = f(r)/r$ turns the equation into a constant–coefficient equation for f). For the proof of the first statement the interested reader is referred to, e.g. [”Basic Partial Differential Equations” (Bleecker-Csordas ’92), section 9.4.]

A direct proof can be obtained by playing with commutators. Let

$$\mathcal{L} := \partial_r^2 + \frac{2}{r}\partial_r, \qquad \mathcal{Q}_\ell := r^\ell\left(\frac{1}{r}\partial_r\right)^\ell. \tag{5.2.15}$$

One verifies by induction that

$$[\mathcal{L}, \mathcal{Q}_\ell] = \frac{\ell(\ell+1)}{r^2}\mathcal{Q}_\ell. \tag{5.2.16}$$

Suppose we have proved (5.2.16) and R_0 is a solution of $\mathcal{L}R_0 = cR_0$ for some c. Then we define $R_\ell := \mathcal{Q}_\ell R_0$ and

$$\begin{aligned}\mathcal{L}R_\ell &= \mathcal{L}(Q_\ell R_0) = \mathcal{Q}_\ell\mathcal{L}R_0 + \frac{\ell(\ell+1)}{r^2}\mathcal{Q}_\ell R_0 = \left(c + \frac{\ell(\ell+1)}{r^2}\right)\mathcal{Q}_\ell R_0 = \\ &= \left(c + \frac{\ell(\ell+1)}{r^2}\right) R_\ell.\end{aligned} \tag{5.2.17}$$

which proves the theorem.

To verify the formula (5.2.16) we observe the following commutation relations

$$[\mathcal{L}, \partial_r] = \frac{2}{r^2}\partial_r, \qquad \left[\mathcal{L}, \frac{1}{r}\right] = -\frac{2}{r^2}\partial,$$

$$\mathcal{Q}_{\ell+1} = r^{\ell+1}\left(\frac{1}{r}\partial_r\right)^{\ell+1} = r^\ell\partial_r\left(\frac{1}{r}\partial_r\right)^\ell = \partial_r\mathcal{Q}_\ell - \frac{\ell}{r}\mathcal{Q}_\ell \tag{5.2.18}$$

Using (5.2.18) one verifies the induction step directly without obstacles. ■

Angular equation and spherical Harmonics: Legendre functions

The functions

$$Y_{\ell,m}(\phi,\theta) = P_{\ell m}(\cos(\phi))\mathrm{e}^{im\theta}, \quad m \in \mathbb{Z} \tag{5.2.19}$$

are called *spherical harmonics* because they are eigenfunctions of the Laplace operator on the 2–sphere

The functions $P_{\ell,m}(x)$ solve (5.2.9) which we now analyze.

Consider first the ODE

$$\Big((1-x^2)u'\Big)' + \lambda u = 0 \tag{5.2.20}$$

with the boundary condition that $u(x)$ is bounded at $x = \pm 1$.

Proposition 5.18 *The solutions of (5.2.20) which are bounded at ± 1 occur only for $\lambda = \ell(\ell+1)$ and $\ell \in \mathbb{N}$. They are given by the* **Legendre polynomials**

$$P_0 = 1, \qquad P_\ell = \frac{1}{2^\ell \ell!}\frac{\mathrm{d}^n((x^2-1)^n)}{\mathrm{d}x^n}, \qquad \ell = 1,2,\dots \tag{5.2.21}$$

which are alternatively written as

$$P_\ell(x) = \frac{1}{2^\ell \ell!}\sum_{s=0}^{\ell}(-1)^{\ell-s}\binom{\ell}{s}\frac{(2s)!}{(2s-\ell)!}x^{2s-\ell}. \tag{5.2.22}$$

Proof. The fact that P_ℓ solves the ODE is derived by differentiating $(\ell+1)$ times the equation

$$(x^2-1)\frac{\mathrm{d}}{\mathrm{d}x}\left(x^2-1\right)^\ell = 2\ell x(x^2-1)^\ell \tag{5.2.23}$$

(**Exercise**)

The equation (5.2.20) in normal form is

$$u'' - \frac{2x}{1-x^2}u' + \frac{\lambda}{1-x^2}u = 0 \tag{5.2.24}$$

which shows that the solutions are defined in the interval $(-1,1)$ but in general have a singularity at one or both endpoints.

If $u = f(x)$ is a nontrivial solution then so is $f(-x)$; thus $f(x)+f(-x)$ and $f(x)-f(-x)$ are solutions as well and at least one of them is nonzero.

We now show that if u is an even or odd solution to (5.2.20) and bounded at ± 1 then it is a the Legendre polynomial.

We seek power-series solution $u(x) = \sum_{n=0}^{\infty} a_n x^n$ and plugging into the equation yields the recurrence relation

$$a_n = \frac{(n-1)(n-2)-\lambda}{n(n-1)}a_{n-2}. \tag{5.2.25}$$

We claim that the solution is a polynomial if and only if $\lambda = \ell(\ell+1)$ for $\ell = 1, 2, ...$ The sufficiency is clear; if $\lambda = \ell(\ell-1)$ we see that $a_\ell = 0$ even if $a_{\ell-2} \neq 0$. Subsequent terms in the series are also all zero and u has the same parity as ℓ.

For $\ell \notin \mathbb{N}$ the series is infinite and standard criteria show that the radius of convergence is 1. It is known also that the only singularities of a solution to the ODE (5.2.20) can be poles at the singularities of the coefficients (this we don't show but it is a standard result in the study of Fuchsian singularities of ODEs in the complex plane). Thus $x = 1$ or $x = -1$ (or both) are a pole and the solution cannot be bounded. (One can prove directly that the solution is unbounded at $x = 1$ because the coefficients have all the same sign for n sufficiently large and the series can be estimated by the harmonic series from below. See Courant–Hilbert's texbook, pag. 326). ■

The general equation for $P_{\ell,m}$ is obtained as follows.

Definition 9 *The Legendre functions of m-th order are defined by*

$$P_{\ell,m}(x) := (1-x^2)^{\frac{m}{2}} \frac{\mathrm{d}^m}{\mathrm{d}x^m} P_\ell(x) \tag{5.2.26}$$

Proposition 5.19 *The Legendre functions $P_{\ell,m}$ solve the ODE (5.2.9)*

$$\frac{\mathrm{d}}{\mathrm{d}x}\left((1-x^2)\frac{\mathrm{d}P_{\ell m}}{\mathrm{d}x}\right) + \left(\ell(\ell+1) - \frac{m^2}{1-x^2}\right) P_{\ell m} = 0 \qquad m = 0, 1, \ldots, \ell. \tag{5.2.27}$$

and are only solutions which are bounded at $x = \pm 1$. For $m > \ell$ there are no bounded solutions.

Proof. We give a sketch of the proof but we won't prove that they are the only bounded solutions. Taking the derivative of (5.2.20) with $\lambda = \ell(\ell+1)$, and defining $u_{\ell,m} := \sqrt{1-x^2}^m \partial_x^m P_\ell(x)$ one sees that it solves (5.2.27) by direct computation. Note also that if $m > \ell$ the proposed expression yields the trivial solution (because P_ℓ is a polynomial of degree ℓ). ■

5.2.3 Separation in polar coordinates: Bessel equation.

Exercise 5.20 *Show that in $\mathbb{R}^2$ the Laplace operator in polar coordinates $(r, \theta) \in \mathbb{R}_+ \times [0, 2\pi)$ reads as follows*

$$\partial_x^2 + \partial_y^2 = \frac{1}{r}\partial_r (r\partial_r) + \frac{1}{r^2}\partial_\theta^2 = \partial_r^2 + \frac{1}{r}\partial_r + \frac{1}{r^2}\partial_\theta^2. \tag{5.2.28}$$

where

$$x = r\cos\theta \ , \qquad y = r\sin\theta. \tag{5.2.29}$$

In this case the Helmholtz (Schrödinger) equation leads to

$$\frac{\mathrm{d}^2R}{\mathrm{d}r^2} + \frac{1}{r}\frac{\mathrm{d}R}{\mathrm{d}r} + \left(V(r) - \frac{m^2}{r^2} + \lambda^2\right)R = 0 \tag{5.2.30}$$

$$\Theta'' + m^2\Theta = 0 \tag{5.2.31}$$

Note that (5.2.30) and (5.2.8) differ only in the coefficient in front of R'. The Bessel (or cylindrical) equation is the case with $V(r) = 0$:

$$R''(r) + \frac{1}{r}R'(r) + \left(\lambda^2 - \frac{m^2}{r^2}\right)R(r) = 0 \tag{5.2.32}$$

or, in self–adjoint form,

$$(rR')' - \frac{m^2}{r}R + \lambda^2 rR = 0 \tag{5.2.33}$$

The interval is typically $[0, \infty)$ or $[0, a]$. In either case this is an *irregular* problem because the function $K(x) = x$ in (5.1.4) vanishes at one of the endpoints.

In keeping with the introduction to the chapter we will consider the case of the wave equation (or heat) on a circular membrane;

$$r^2F'' + rF' + (r^2\lambda^2 - m^2)F = 0\ , \qquad F(R) = 0 \tag{5.2.34}$$

With the transformation $\rho = \lambda r$ the problem is recast into the ODE

$$\frac{\mathrm{d}^2}{\mathrm{d}\rho^2}F + \frac{1}{\rho}\frac{\mathrm{d}F}{\mathrm{d}\rho} + \left(1 - \frac{m^2}{\rho^2}\right)F = 0. \tag{5.2.35}$$

$$F(\ell) = 0 \tag{5.2.36}$$

where $\ell = R\lambda$ and we recall that $m = 0, 1, 2, \ldots$ is an integer because the only periodic solutions of (5.2.10) occur for $m \in \mathbb{Z}$ (as we have observed in the study of the Laplace equation in two–dimensions).

Proposition 5.21 *Consider the Bessel equation on $\mathbb{R}_+$*

$$x^2f'' + xf' + (x^2 - m^2)f = 0 \tag{5.2.37}$$

Then any solution has infinitely many zeroes on $\mathbb{R}_+$.

The proof is contained in the two Exercises 5.24 and 5.25.

Expressions for the Bessel functions are contained in Exercises 5.30 and 5.31

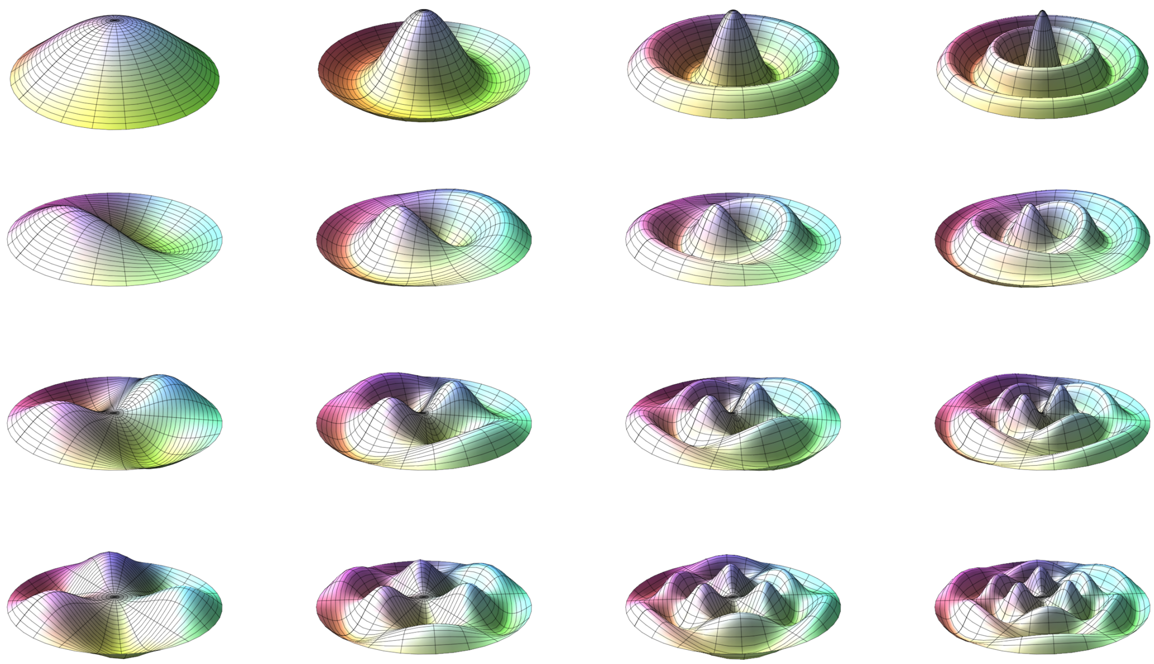

Figure 5.2: The first few modes of the membrane; the plot is of $J_m(\lambda_{m,n} r)\cos(m\theta)$ for $m = 0, 1, .., 4$ (rows) and $n = 0, 1, 2, 3$ (columns), $r \in [0, 1]$, $\theta \in [0, 2\pi]$.

The circular membrane

Suppose now that we want to solve the Helmholtz equation for the circular membrane (disk) i.e. with the need to solve the ODE with boundary condition

$$\frac{\mathrm{d}^2 R}{\mathrm{d}r^2} + \frac{1}{r}\frac{\mathrm{d}R}{\mathrm{d}r} + \left(\lambda^2 - \frac{m^2}{r^2}\right) R = 0 \ , \quad R(\rho) = 0 \tag{5.2.38}$$

and also $R(r)$ bounded at $r = 0$. Then we see that the Bessel function $J_m(\lambda r)$ is a solution of the ODE and that λ needs to be chosen so that $\lambda\rho$ is one of the zeroes of J_m. We know from the Proposition 5.21 that there are infinitely many zeros and so there are infinitely many eigenvalues. A plot of the first few Bessel functions is in Fig. 5.3

For each $m = 0, 1, 2, \ldots$ denote by $\lambda_{m,s}$, $s = 1, 2, \ldots$ the positive zeroes of $J_m(x)$; Then the solution of the Heat equation

$$u_t = k^2(u_{xx} + u_{yy}) \ , \tag{5.2.39}$$

$$u(x, y, t) = 0, \qquad x^2 + y^2 = \rho^2 \tag{5.2.40}$$

$$u|_{t=0} = \phi(x, y) \tag{5.2.41}$$

is of the form

$$u = \sum_{m\in\mathbb{Z}} \sum_{s=1}^{\infty} A_{m,s} \mathrm{e}^{-k^2 \lambda_{|m|,s} t} J_m\left(\lambda_{|m|,s}\frac{r}{\rho}\right) \mathrm{e}^{im\theta} \tag{5.2.42}$$

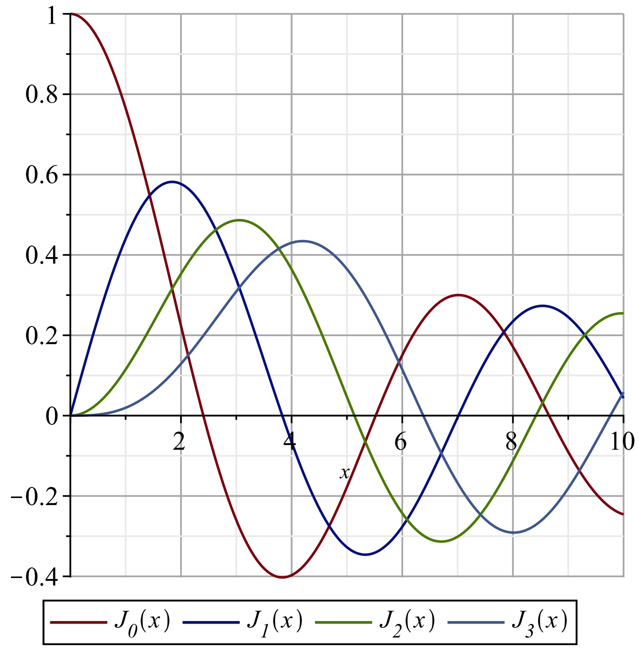

Figure 5.3: The plot of the first few Bessel functions.

where we have used the fact (see Ex. 5.30) that $J_{-m}(x) = (-1)^m J_m(x)$. The coefficients of $A_{m,s}$ are the Fourier coefficients of ϕ and are given by

$$A_{m,s} = \frac{1}{\Omega_{m,s}} \int_0^{2\pi} \mathrm{d}\theta \int_0^{\rho} \phi(r,\theta) \mathrm{e}^{-im\theta} J_m(\lambda_{|m|,s} r/\rho) r \mathrm{d}r \tag{5.2.43}$$

Remark 5.22 *If ϕ depends on r alone, then we have the "Fourier–Bessel series" of order 0 for the function ϕ.*

5.3 Exercises for Chapter 5

Exercise 5.23 *Show that for any fixed λ the function $\Phi(x;\lambda)$ solving (5.1.49) in Thm. 5.13 cannot have infinitely many zeroes in $[a,b]$.* **Hint:** *If there were infinitely many zeroes, they would accumulate somewhere. Show that the derivative at the point of accumulation is zero and deduce a contradiction.*

Exercise 5.24 *Show that if $f_m(x)$ is a solution of (5.2.37) and we pose $y_m = \sqrt{x}f_m(x)$ then the function y_m solves the ODE*

$$y'' + \left(1 + \frac{1/4 - m^2}{x^2}\right) y = 0, \quad x > 0 \tag{5.3.1}$$

Exercise 5.25 *Let $U_m(x) := \left(1 + \frac{1/4-m^2}{x^2}\right)$. Show that for every $\epsilon > 0$ we have the inequality*

$$U_m(x) \geqslant 1 - \epsilon, \quad x \in \left(\frac{m}{\sqrt{\epsilon}}, \infty\right). \tag{5.3.2}$$

Deduce that the solution y_m (and hence $f_m(x)$ has infinitely many zeroes on the positive half-line.

Exercise 5.26 *Find eigenvalues λ_n and eigenfunctions y_n of the Sturm–Liouville problem*

$$(DE): \qquad y'' + \lambda y = 0, \quad x \in [0,L], \tag{5.3.3}$$

subject to the BC's

$$(\mathbf{a}) \;\; y(0) = 0; \;\; y'(L) = 0; \qquad (\mathbf{b}) \;\; y'(0) = 0 = y'(L). \tag{5.3.4}$$

Solution to Problem (5.26) **(a)** For $\lambda < 0$ we have $y = \sinh(\sqrt{|\lambda|}x)$ to match the condition at 0. To match the condition at L is impossible because y' is never zero for $x \neq 0$. For $\lambda = 0$ we have a straight line with zero slope (because $y'(L) = 0$ and starting from zero; hence also trivial. So we are left with $\lambda > 0$.

We must have $y = \sin(\sqrt{\lambda}x)$; $y'(L) = \sqrt{\lambda}\cos(\sqrt{\lambda}L) = 0$ and hence

$$\lambda_k = \frac{1}{L^2}\left(\frac{\pi}{2} + k\pi\right)^2, \qquad k = 0,1,2,... \tag{5.3.5}$$

(the negative k's give the same set of eigenvalues).

(b) For $\lambda < 0$ we must have $y = \cosh(\sqrt{\lambda}x)$ ($y'(0) = 0$) and $y'(L) = \sqrt{\lambda}\sinh(\sqrt{\lambda}L) = 0$ which has only $\sqrt{\lambda} = 0$ solution. For $\lambda = 0$ we have the constant solution $y_0(x) = 1$. For $\lambda > 0$ we have

$$y = \cos(\sqrt{\lambda}x), \quad y'(L) = \sqrt{\lambda}\sin(\sqrt{\lambda}L) = 0 \qquad \Leftrightarrow \qquad \lambda = \left(\frac{n\pi}{L}\right)^2, \quad n = 1,2,\dots. \tag{5.3.6}$$

■

Exercise 5.27 **(a)** *Show that every real value λ is an eigenvalue of the problem*

$$\begin{aligned} (DE):&\quad y'' + \lambda y = 0, \qquad x \in [0,1] \\ (BC):&\quad y(0) - y(1) = 0, \qquad y'(0) + y'(1) = 0 \end{aligned} \tag{5.3.7}$$

(b) *Show that the problem*

$$\begin{aligned} (DE): &\quad y'' + \lambda y = 0, \quad x \in [0,\pi] \\ (BC): &\quad \pi y(0) + -y(\pi) = 0, \qquad \pi y'(0) + y'(\pi) = 0 \end{aligned} \tag{5.3.8}$$

has no real eigenvalues.

Are the above examples in contradiction with Theorem 5.13?

Exercise 5.28 **(a)** *Show that every linear second order ODE*

$$P_2 y'' + P_1 y' + P_0 y = 0 \tag{5.3.9}$$

(with $P_2 > 0$ and $P_2, P_2' P_1, P_0$ continuous) can be transformed in an equation in self–adjoint form by multiplication with $\exp\left(\int \frac{P_1 - P_2'}{P_2} \mathrm{d}x\right)$. **(b)** *Transform the following ODE's in self–adjoint form;*

$$\begin{array}{lll} \textbf{(i)} \quad x^2 y'' + x y' + (x^2 - m^2) y = 0, \quad x > 0 & \textit{(Bessel)} & \\ \textbf{(ii)} \quad (1 - x^2) y'' - 2 x y' + m(m+1) y = 0 \quad x \in [-1,1] & \textit{(Legendre)} & \\ \textbf{(iii)} \quad (1 - x^2) y'' - x y' + m^2 y = 0 \quad x \in [-1,1] & \textit{(Chebyshev)} & \\ \textbf{(iv)} \quad y'' - 2 x y' + 2 m y = 0 \quad x \in \mathbb{R} & \textit{(Hermite)} & \\ \textbf{(iv)} \quad x y'' + (1 - x) y' + m y = 0 \quad x \in \mathbb{R}_+ & \textit{(Laguerre)} & (5.3.10) \end{array}$$

Exercise 5.29 *Consider the Legendre polynomials*

$$P_\ell = \frac{1}{2^\ell \ell!} \frac{\mathrm{d}^n((x^2-1)^\ell)}{\mathrm{d}x^\ell} \tag{5.3.11}$$

(a) *Show that they are orthogonal to each other in $L^2([-1,1], \mathrm{d}x)$. Deduce that they are obtained by Gram–Schmidt orthogonalization process from the ordered basis $\{1, x, x^2, \ldots, x^n, \ldots\}$ of L^2.*
(b) *Consider their generating function*

$$G(x,t) := \sum_{\ell=0}^{\infty} P_\ell(x) t^\ell \tag{5.3.12}$$

Show that it equals

$$G(x,t) = \frac{1}{\sqrt{1 - 2tx + t^2}} \tag{5.3.13}$$

Hint: *express P_ℓ as a residue using Cauchy's formula. Swap sum and integral and find a geometric series. Evaluate the resulting integral.*
(c) *Deduce that*

$$\int_{-1}^{1} P_\ell^2 \mathrm{d}x = \frac{2}{2\ell + 1}. \tag{5.3.14}$$

Exercise 5.30 *Show that the following two series are solutions of the Bessel equation (5.2.32) for $m \notin \mathbb{N}$ and $\lambda = 1$ (for $\lambda \neq 1$ a simple rescaling can be used):*

$$J_m(x) = \left(\frac{x}{2}\right)^m \sum_{n=0}^{\infty} \frac{(-x^2/4)^n}{n!(m+n)!}$$
$$J_{-m}(x) = \left(\frac{x}{2}\right)^{-m} \sum_{n=0}^{\infty} \frac{(-x^2/4)^n}{n!(n-m)!} \tag{5.3.15}$$

Show also that for integer m $J_{-m}(x) = (-1)^m J_m(x)$. Using that the Wronskian of two independent solutions of (5.2.32) is $W = \exp(-\int \frac{1}{x}\mathrm{d}x) = \frac{C}{x}$ deduce that there can be at most one bounded solution at $x = 0$.

Exercise 5.31 **[1]** *Show that the following integrals*

$$J_m(z) = \oint_{|t|=1} \mathrm{e}^{\frac{z}{2}(t-t^{-1})} t^{-m-1} \frac{\mathrm{d}t}{2i\pi} = \frac{1}{\pi}\int_0^\pi \cos\Big(m\theta - z\sin(\theta)\Big)\mathrm{d}\theta, \qquad m \in \mathbb{Z} \tag{5.3.16}$$

give a solution of the equation (5.2.32);

$$J_m'' + \frac{1}{z}J_m' + \left(1 - \frac{m^2}{z^2}\right) J_m = 0 \tag{5.3.17}$$

[2] *Show that the above Bessel functions solve the recurrence relation*

$$J_{m+1}(z) + J_{m-1}(z) = \frac{2m}{z} J_m(z). \tag{5.3.18}$$

[3] *Show that*

$$\frac{\mathrm{d}}{\mathrm{d}z}\left(z^m J_m(z)\right) = z^m J_{m-1}(z). \tag{5.3.19}$$

[4] *Compute*

$$\int_0^\infty J_0(x)\mathrm{e}^{-px}\mathrm{d}x = \frac{1}{\sqrt{1+p^2}}, \qquad (p > 0). \tag{5.3.20}$$

Give a formula for

$$\int_0^\infty J_m(x)\mathrm{e}^{-px}\mathrm{d}x, \qquad m = 1, 2, \ldots, \quad p > 1. \tag{5.3.21}$$

Solution of part [1]. We use the expression with the cos:

$$J_n' = \int_0^\pi \sin(\theta)\sin(n\theta - z\sin\theta)\frac{\mathrm{d}\theta}{\pi} \tag{5.3.22}$$
$$J_n'' = -\int_0^\pi \sin^2(\theta)\cos(n\theta - z\sin\theta)\frac{\mathrm{d}\theta}{\pi} \tag{5.3.23}$$
$$\tag{5.3.24}$$

We compute J' by parts

$$
\begin{aligned}
J_n' = \int_0^\pi \sin(\theta)\sin(n\theta - z\sin\theta)\frac{\mathrm{d}\theta}{\pi} = \\
= \underbrace{-\cos(\theta)\sin(n\theta - z\sin(\theta))\Big|_0^\pi}_{=0} + \int_0^\pi \cos\theta\cos\left(n\theta - z\sin\theta\right)\left(n - z\cos\theta\right) = \\
= n\int \cos\theta\cos(n\theta - z\sin\theta) - zJ_n + z\int_0^\pi \sin^2(\theta)\cos(n\theta - z\sin\theta) = \\
= n\int \cos\theta\cos(n\theta - z\sin\theta) - zJ_n - zJ_n''
\end{aligned}
\tag{5.3.25}
$$

Rearranging the terms above we have shown

$$
J_n'' + \frac{1}{z}J_n' + J_n = \frac{n}{z}\int \cos\theta\cos(n\theta - z\sin\theta) \tag{5.3.26}
$$

Thus

$$
\begin{aligned}
J_n'' + \frac{1}{z}J_n' + \left(1 - \frac{n^2}{z^2}\right)J_n = \frac{n}{z}\int_0^\pi \cos\theta\cos(n\theta - z\sin\theta)\frac{\mathrm{d}\theta}{\pi} - \frac{n^2}{z^2}J_n \overset{\text{trig. id.}}{=} \\
= \frac{n}{z}\int_0^\pi \left(\cos(n\theta)\overbrace{\cos(z\sin\theta)\cos\theta}^{=\frac{1}{z}(\sin(z\sin\theta))'} + \sin(n\theta)\overbrace{\cos\theta\sin(z\sin\theta)}^{=-\frac{1}{z}(\cos(z\sin(\theta))'}\right)\frac{\mathrm{d}\theta}{\pi} - \frac{n^2}{z^2}J_n \overset{\text{b.p.}}{=} \\
= \frac{n}{z}\int_0^\pi \underbrace{\left(\frac{n}{z}\sin(n\theta)\sin(z\sin\theta) + \frac{n}{z}\cos(n\theta)\cos(z\sin\theta)\right)}_{=\cos(n\theta - z\sin\theta)}\frac{\mathrm{d}\theta}{\pi} - \frac{n^2}{z^2}J_n = 0
\end{aligned}
\tag{5.3.27}
$$

Chapter 6

Introduction to nonlinear PDEs

6.1 Method of characteristics for first order quasilinear equations

Let us recall (see Section 1.5 above) the procedure of construction of the general solution for the first order linear homogeneous equation

$$\frac{\partial u}{\partial t} = \sum_{i=1}^{d} a_i(x,t) \frac{\partial u}{\partial x_i}. \tag{6.1.1}$$

Here and below $x = (x_1, \dots, x_d)$. One has to consider the system of equations for the characteristics of (6.1.1)

$$\begin{aligned} \dot{x}_i &= a_i(x,t), \quad i = 1, \dots, d \\ \dot{t} &= -1. \end{aligned}$$

Using t as the parameter along the characteristics one can recast the above system into the form

$$\frac{dx_i}{dt} + a_i(x,t) = 0, \quad i = 1, \dots, d. \tag{6.1.2}$$

Any solution to the system (6.1.1) is a function $u = u(x,t)$ constant along the characteristics. Recall that such functions are called *first integrals* of the system of ODEs (6.1.2).

In order to construct the general solution to (6.1.1) one has to find d independent first integrals, i.e., d particular solutions $v_1(x,t)$, $\dots$, $v_d(x,t)$ to the PDE (6.1.1) satisfying the condition

$$\det \begin{pmatrix} \partial v_1/\partial x_1 & \dots & \partial v_1/\partial x_d \\ \dots & \dots & \dots \\ \partial v_d/\partial x_1 & \dots & \partial v_d/\partial x_d \end{pmatrix} \neq 0 \tag{6.1.3}$$

at a given point $(x_0, t_0) \in \mathbb{R}^d \times \mathbb{R}$. Then the general solution to the PDE (6.1.1) near this point can be written as follows

$$u(x,t) = U\left(v_1(x,t), \dots, v_d(x,t)\right) \tag{6.1.4}$$

where $U(u_1, \dots, u_d)$ is an arbitrary smooth function of d variables. Indeed, the following simple statement holds true.

Proposition 6.1 *Let $u(x,t)$ be a solution to the Cauchy problem for the equation (6.1.1) defined in a neighborhood of the point (x_0, t_0) and satisfying the initial condition*

$$u(x, t_0) = \phi(x), \quad |x - x_0| < \rho \tag{6.1.5}$$

with a smooth function $\phi(x)$ defined on the ball $|x - x_0| < \rho$ for some positive ρ. Then there exists a smooth function $U(v_1, \dots, v_d)$ on some neighbourhood of the point

$$\mathbf{u}^0 := (v_1^0, \dots, v_d^0) = (v_1(x_0, t_0), \dots, v_d(x_0, t_0)) \in \mathbb{R}^d$$

such that the solution $u(x,t)$ can be represented in the form (6.1.4) for $|x - x_0| < \rho_1$ for some positive $\rho_1 \leqslant \rho$.

Proof: Applying the theorem about the inverse mapping to the system

$$\begin{aligned} v_1 &= v_1(x, t_0) \\ \dots & \quad \dots \quad \dots \\ v_d &= v_d(x, t_0) \end{aligned}$$

one obtains smooth functions

$$\begin{aligned} x_1 &= x_1(v_1, \dots, v_d) \\ \dots & \quad \dots \quad \dots \quad \dots \\ x_d &= x_d(v_1, \dots, v_d) \end{aligned}$$

defined on some neighborhood of the point u^0 and uniquely determined by the conditions

$$x_i(\mathbf{u}^0) = x_i^0, \quad i = 1, \dots, d.$$

This can be done due to the assumption (6.1.3). We put

$$U(u_1, \dots, u_d) := \phi\left(x_1(u_1, \dots, u_d), \dots, x_d(u_1, \dots, u_d)\right).$$

Such a function gives the needed representation of the solution $u(x,t)$. ∎

Let us now consider a *quasilinear* equation, not necessarily homogeneous. By definition such an equation has the form

$$\frac{\partial u}{\partial t} = \sum_{i=1}^{d} a_i(u, \mathbf{x}, t) \frac{\partial u}{\partial x_i} + b(u, \mathbf{x}, t) \tag{6.1.6}$$

with the coefficients $a_1(u, \mathbf{x}, t)$, ..., $a_d(u, \mathbf{x}, t)$, $b(u, \mathbf{x}, t)$ being smooth functions on some neighborhood of a point $(u_0, x_0, t_0) \in \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}$. The following trick reduces the problem (6.1.6) to the previous one. Let us look for solutions to (6.1.6) written in the implicit form

$$f(u, \mathbf{x}, t) = 0 \tag{6.1.7}$$

where $f(u, \mathbf{x}, t)$ is a smooth function defined on some neighborhood of the point $(u_0, \mathbf{x}_0, t_0)$ satisfying the condition

$$f_u(u_0, \mathbf{x}_0, t_0) \neq 0. \tag{6.1.8}$$

According to the implicit function theorem, the assumption (6.1.8) implies existence and uniqueness of a smooth function $u(\mathbf{x}, t)$ defined on some neighborhood of the point $(\mathbf{x}_0, t_0) \in \mathbb{R}^d \times \mathbb{R}$ and satisfying $u(\mathbf{x}_0, t_0) = u_0$. Let us derive the condition for the function f that guarantees that $u(\mathbf{x}, t)$ satisfies (6.1.6). According to the implicit function theorem the partial derivatives of the function $u(\mathbf{x}, t)$ determined by (6.1.7) can be written in the form

$$\frac{\partial u}{\partial t} = -\frac{f_t(u, \mathbf{x}, t)}{f_u(u, \mathbf{x}, t)}, \quad \frac{\partial u}{\partial x_i} = -\frac{f_{x_i}(u, \mathbf{x}, t)}{f_u(u, \mathbf{x}, t)}, \quad i = 1, \ldots, d. \tag{6.1.9}$$

The substitution to (6.1.6) yields a linear homogeneous PDE for the function f of $d + 2$ variables

$$\frac{\partial f}{\partial t} = \sum_{i=1}^{d} a_i(u, \mathbf{x}, t) \frac{\partial f}{\partial x_i} - b(u, \mathbf{x}, t) \frac{\partial f}{\partial u}. \tag{6.1.10}$$

The solution $f(u, \mathbf{x}, t)$ to this PDE with the initial data chosen in the form

$$f(u, \mathbf{x}, t_0) = u - \phi(\mathbf{x}) \tag{6.1.11}$$

give a solution to the original PDE (6.1.6) specified by the initial data

$$u(\mathbf{x}, t_0) = \phi(\mathbf{x}), \quad |\mathbf{x} - \mathbf{x}_0| < \rho \tag{6.1.12}$$

for some positive ρ. Note that the function ϕ must satisfy $\phi(\mathbf{x}_0) = u_0$. The PDE (6.1.10) can be solved by the method of characteristics. The characteristics in the $(d + 2)$-dimensional space with the coordinates u, x_1, $\ldots$, x_d, t can be determined from the following system of ODEs

$$\begin{aligned} &\frac{\partial x_i}{\partial t} + a_i(u, \mathbf{x}, t) = 0, \quad i = 1, \ldots, d \\ &\frac{\partial u}{\partial t} = b(u, \mathbf{x}, t). \end{aligned} \tag{6.1.13}$$

Like above, one has to find $(d + 1)$ independent first integrals, i.e., $(d + 1)$ particular solutions $f_0(u, \mathbf{x}, t)$, $\ldots$, $f_d(u, \mathbf{x}, t)$ satisfying

$$\det \begin{pmatrix} \partial f_0/\partial u & \partial f_0/\partial x_1 & \ldots & \partial f_0/\partial x_d \\ \partial f_1/\partial u & \partial f_1/\partial x_1 & \ldots & \partial f_1/\partial x_d \\ \ldots & \ldots & \ldots & \ldots \\ \partial f_d/\partial u & \partial f_d/\partial x_1 & \ldots & \partial f_d/\partial x_d \end{pmatrix} \neq 0 \tag{6.1.14}$$

at the given point (u_0, x_0, t_0). The general solution to the PDE (6.1.10) can be represented in the form

$$f(u, \mathbf{x}, t) = F\left(f_0(u, \mathbf{x}, t), f_1(u, x, t), \ldots, f_d(u, \mathbf{x}, t)\right). \tag{6.1.15}$$

The smooth function F of $(d+1)$ variables has to be determined from the Cauchy data (6.1.11)

$$F\left(f_0(u,\mathbf{x},t_0),\ldots,f_d(u,\mathbf{x},t_0)\right)=u-\phi(\mathbf{x}). \tag{6.1.16}$$

Then the function $u=u(\mathbf{x},t)$ is the function defined implicitly by (6.1.16).

The local existence and uniqueness of such a solution is established as before.

Let us consider in more details the case of quasilinear homogeneous equations in one spatial dimension with coefficients independent from x and t

$$u_t=a(u)u_x. \tag{6.1.17}$$

The equations for the characteristics become very simple in this particular case:

$$\begin{aligned}&\frac{dx}{dt}+a(u)=0\\ &\frac{du}{dt}=0.\end{aligned} \tag{6.1.18}$$

The solutions are straight lines

$$u=\text{const},\quad x+a(u)t=\text{const}. \tag{6.1.19}$$

Thus the general solution can be written in the implicit form

$$x+a(u)(t-t_0)=f(u). \tag{6.1.20}$$

The function $f(u)$ has to be determined from the initial condition

$$u(x,t_0)=\phi(x).$$

This gives

$$x=f\left(\phi(x)\right).$$

The solution to the last equation exists if the initial function $\phi(x)$ is monotonous near the point $x=x_0$. Then the function f coincides with the inverse function ϕ^{-1}.

Example. In the proof of the Cauchy–Kovalevskaya theorem we arrived at the following Cauchy problem

$$\begin{aligned}v_t&=\frac{M\,n}{1-\frac{n}{r}v}v_x\\ v(x,0)&=\frac{M\,x}{\rho-x}.\end{aligned}$$

(see (7.2.26) above). The general solution to the PDE in the implicit form reads

$$x+\frac{M\,n}{1-\frac{n}{r}v}\,t=f(v)$$

for an arbitrary function $f(v)$ to be determined by the initial data. To do this one has to solve the equation

$$v=\frac{M\,x}{\rho-x}$$

for x. This gives

$$x = \frac{\rho\, v}{M+v} =: f(v).$$

Thus the solution to the above Cauchy problem has to be determined from the algebraic equation

$$x + \frac{M\, n}{1 - \frac{n}{r} v}\, t = \frac{\rho\, v}{M+v}. \tag{6.1.21}$$

This coincides with (7.2.27).

For the particular case $a(u) = c$=const the equation

$$u_t + a(u) u_x = 0 \tag{6.1.22}$$

describes propagation of waves with constant speed c. The characteristics in this case are just parallel lines

$$x = c\, t + x_0.$$

We will now concentrate our attention at the simplest example of a nonlinear PDE of the above form

$$v_t + v\, v_x = 0 \tag{6.1.23}$$

called *Hopf equation*. This equation can be used as the simplest example of equations describing motion of an ideal incompressible fluid. The fluid can be considered as a system of an infinite number of particles distributed with some density ρ that in the *incompressible* case will be assumed to be constant. The particles can be "labeled" in two different ways. In the *Lagrange* parameterization one can label the particles by their positions $\xi \in \mathbb{R}$ at a certain initial moment of time. The motion then will be described by a pair of functions

$$\begin{aligned} x &= x(\xi, t) \\ v &= v(\xi, t) \end{aligned} \tag{6.1.24}$$

where $x(\xi, t)$ and $v(\xi, t)$ are the coordinate and the velocity of the particle with the "number ξ" at the moment t. By definition we have

$$\frac{\partial x(\xi, t)}{\partial t} = v(\xi, t). \tag{6.1.25}$$

In the *Euler* parameterization we just follow the motion of the particle passing through the point x at the moment t. Any physical quantity f assigned to every particle (e.g., the temperature[1] of the particle) will be characterized by a function $f = f(x, t)$.

Proposition 6.2 *If the quantity f is conserved, i.e., it depends only on the initial position of the particles, $f = f(\xi)$, then the function $f(x, t)$ satisfies the equation*

$$\frac{\partial f(x,t)}{\partial t} + v(x,t) \frac{\partial f(x,t)}{\partial x} = 0. \tag{6.1.26}$$

[1] In the case f=temperature of the water in the river the function $f(x, t)$ is obtained by measuring the temperature sitting on the beach while $f(\xi, t)$ can be measured from the boat drifting freely along the stream of the river.

Proof: By using the chain rule along with (6.1.25) we obtain

$$0 = \frac{d}{dt} f(\xi) = \frac{\partial f}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial f}{\partial t} = \frac{\partial f}{\partial t} + v\,\frac{\partial f}{\partial x}.$$

■

Exercise 6.3 *In the three-dimensional case of a fluid moving with the velocity $v = (v_x, v_y, v_z)$ derive a similar equation for dependence of a conserved quantity $f = f(x, y, z; t)$:*

$$\frac{\partial f}{\partial t} + v_x \frac{\partial f}{\partial x} + v_y \frac{\partial f}{\partial y} + v_z \frac{\partial f}{\partial z} = 0. \tag{6.1.27}$$

Let us consider the free motion of an ideal compressible fluid; we recall that this means that the fluid does not oppose any ”resistance” to being compressed, as opposed to an incompressible fluid, for which the velocity field must satisfy div$\vec{v} = 0$ at all times; clearly in dimension 1 the incompressibility forces v to be independent of x and hence nothing interesting can happen.

Thus, in the case of a free ideally compressible fluid, no external forces act on the particles of the fluid. Because of this the momentum of every particle conserved. From the Proposition 6.2 one immediately obtains

Corollary 6.4 *For the free motion of an ideal incompressible fluid the velocity $v(x, t)$ satisfies equation (6.1.23).*

According to the general procedure the Cauchy problem for the equation (6.1.23) with the initial data

$$v(x, 0) = \phi(x) \tag{6.1.28}$$

for small time t can be written in the implicit form

$$\begin{aligned} x &= v\,t + f(v) \\ f\left(\phi(x)\right) &= x. \end{aligned} \tag{6.1.29}$$

on every interval of monotonicity of the initial data $\phi(x)$. Let us try to figure out what can happen when the time is not that small.

The solution $v = v(x, t)$ to the equation (6.1.29) exists provided the conditions of the implicit function theorem hold true:

$$t + f'(v) \neq 0. \tag{6.1.30}$$

At the moment where fails, let's say t_0, v_0, the function $v\,t + f(v)$ is not locally monotone any more, so the equation (6.1.29) cannot be solved for v. Let us assume for simplicity that the initial data is a globally monotone *decreasing* function. Then the inverse function $f(v)$ will be monotone decreasing as well. Denote t_0 the first moment of time for which the function $X(v, t) := v\,t + f(v)$ becomes not a monotone function at some point v_0. Since t_0 is the first such time, the function $X(v, t_0)$ must be decreasing on the left and the right of $v = v_0$, and hence v_0 must be an inflection point of the graph

$$x = v\,t + f(v)$$

i.e., at this point one must have

$$f''(v_0) = 0.$$

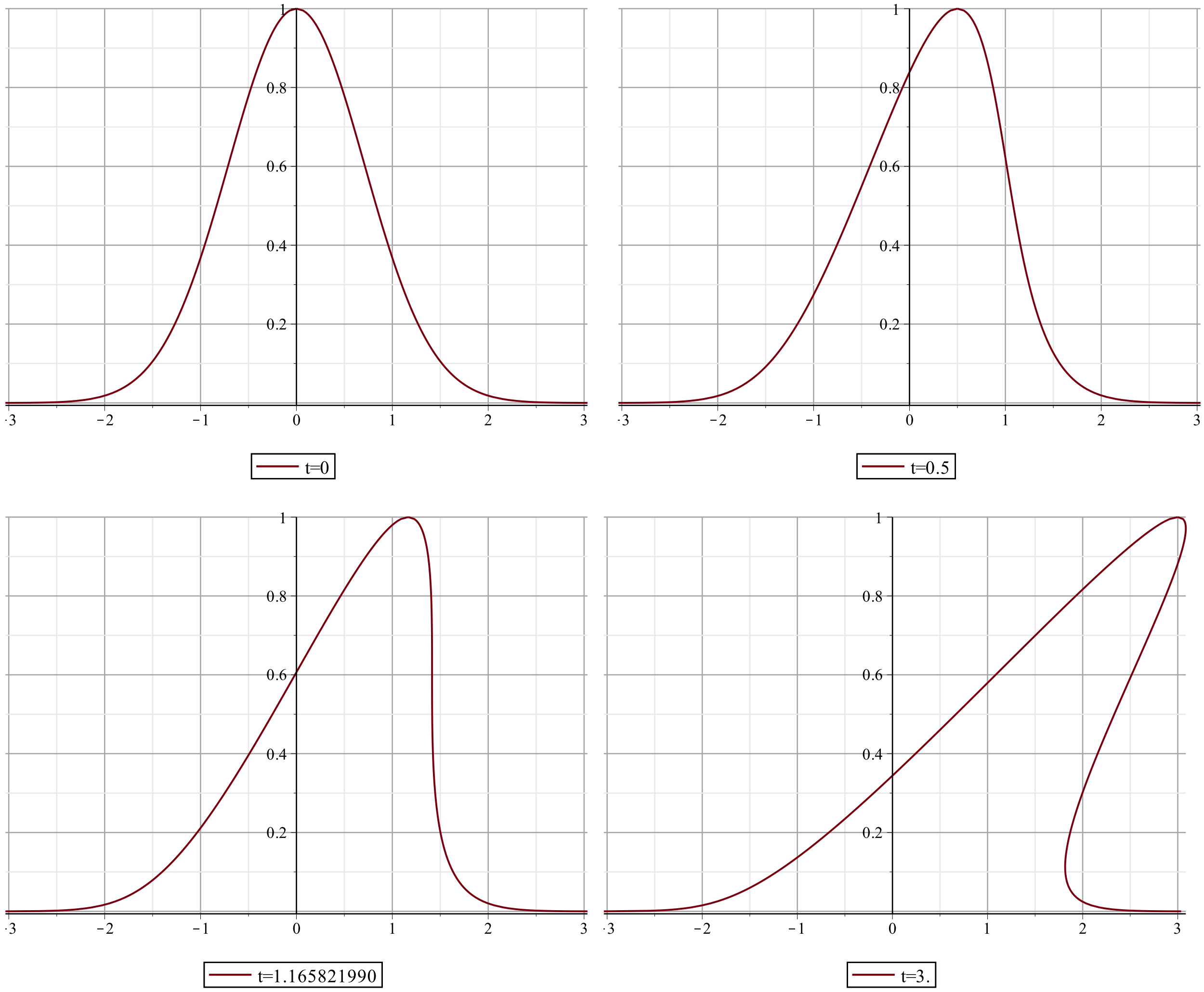

Figure 6.1: The Hopf-evolution. Each point (x, v) on the graph of $v = u(x, t)$ travels to the right with constant speed equal to v; after the time of catastrophe the solution given by the method of characteristics loses significance because it is not a function of x any longer.

In this way we arrive at 1the following "bad points" (x_0, t_0, v_0) where the implicit function theorem does not work any more. The coordinates of these points can be determined from the following system

$$\begin{aligned} &x_0 = v_0 t_0 + f(v_0) \\ &t_0 + f'(v_0) = 0 \\ &f''(v_0) = 0. \end{aligned} \tag{6.1.31}$$

Such a point (x_0, t_0) is called the point of *gradient catastrophe*. The solution to the Cauchy problem (6.1.28) exists for all $x \in \mathbb{R}$ only for $t < t_0$; the derivatives u_x and u_t become infinite at the point of gradient catastrophe.

Example 6.5 *Consider the Hopf equation with the initial datum (see Fig. 6.1)*

$$\phi(x) = \mathrm{e}^{-x^2} \tag{6.1.32}$$

We can find the inverse function in the two intervals $\mathbb{R}_\pm$*:*

$$f_\pm(v) = \pm\sqrt{-\ln v}, \qquad v \in (0,1] \tag{6.1.33}$$

Then the two branches evolve according to the implicit equations

$$x = vt + f_\pm(v) \tag{6.1.34}$$

The point $(x_0, v_0) = (0,1)$ *follows the characteristic line* $x_0(t) = v_0 t = t$ *and for each* t *fixed we can expect two branches of the solution in the intervals* $(-\infty, t)$ *and* (t, ∞)*.*

For $t \neq 0$ *we cannot explicitly invert the equation (6.1.34) but we can nonetheless plot the* **parametric curves** $x = X_\pm(v,t) = vt + f_\pm(v)$*; the plots for various times are in Fig. 6.1.*

Let us find the point of gradient catastrophe: we have

$$f''_\pm = \mp\frac{1}{4}\frac{2\ln v + 1}{(-\ln(v))^{\frac{3}{2}}v^2} = 0 \quad\Rightarrow\quad v_0 = \mathrm{e}^{-\frac{1}{2}} \quad\Rightarrow\quad t_0 = -f'(\mathrm{e}^{-\frac{1}{2}}) = \frac{\mathrm{e}^{\frac{1}{2}}}{\sqrt{2}} \simeq 1.1658 \tag{6.1.35}$$

and $x_0 = \sqrt{2}$*. For times* $t > t_0$ *the parametric curves cannot be functions of* x*.*

6.2 Higher order perturbations of the first order quasilinear equations. Solution of the Burgers equation

As we have seen in the previous section the life span of a typical solution to the equations of motion of an ideal incompressible fluid is finite: the solution does not exist beyond the point of gradient catastrophe. Such a phenomenon suggests that the real physical process can be only approximately described by the equation (6.1.23). Near the point of catastrophe higher corrections have to be taken into account.

We will consider two main classes of such perturbations of Hopf equation: *Burgers equation*

$$v_t + v\, v_x = \mu\, v_{xx} \tag{6.2.1}$$

and *Korteweg - de Vries* (KdV) equation

$$v_t + v\, v_x + \epsilon^2 v_{xxx} = 0. \tag{6.2.2}$$

The small parameters μ and ϵ will be assumed to be positive. The Burgers equation arises in the description of one-dimensional waves in the presence of small *dissipative* effects; the small parameter μ is called the *viscosity coefficient.* The Korteweg - de Vries (KdV) equation describes one-dimensional waves with no dissipation but in the presence of small *dispersion.* It turns out that in both cases the perturbation, whatever small it be, resolves the problem with non-existence of solutions to the Cauchy problem for large time. However we will see that the properties of solutions to the equations (6.2.1) and (6.2.2) are rather different.

Let us first explain in what sense the equation (6.2.1) has to be considered as a dissipative equation but there is no dissipation in (6.2.2). First observe that both equations have a family of constant solutions

$$v = c.$$

We will now apply the general *linearization* procedure in order to study small perturbations of constant solutions. The idea is to look for the perturbed solutions in the form

$$v(x,t) = c + \delta v(x,t). \tag{6.2.3}$$

The perturbation is assumed to be small, so we will neglect the terms quadratic in δv. In such a way we arrive at the linearized Burgers equation

$$\delta v_t + c\,\delta v_x = \mu\,\delta v_{xx} \tag{6.2.4}$$

and the linearized KdV equation

$$\delta v_t + c\,\delta v_x + \epsilon^2 \delta v_{xxx} = 0. \tag{6.2.5}$$

Let us look for the plane wave solutions to these equations:

$$\delta v = a\,e^{ikx - i\omega t}.$$

The substitution to (6.2.4) and (6.2.7) yields the *dispersion relation* between the wave number k and the frequency ω. Namely, we obtain that

$$\omega = c\,k - i\mu\,k^2 \tag{6.2.6}$$

for Burgers equation and

$$\omega = c\,k - \epsilon^2 k^3 \tag{6.2.7}$$

for the KdV equation. We conclude that the small perturbations of the constant solutions to the Burgers equation exponentially decay at $t \to +\infty$

$$\delta v = a\,e^{ik(x-ct) - \mu\,k^2 t}, \quad |\delta v| = |a| e^{-\mu k^2 t} \to 0$$

while for the KdV equation the magnitude of small perturbations remains unchanged:

$$\delta v = a\,e^{ik(x-ct) + i\epsilon^2 k^3 t}, \quad |\delta v| \equiv |a|.$$

We postpone the explanation of the dispersive nature of the KdV equation till Section 6.4. We will now concentrate our attention on the solutions to Burgers equation. We will first prove global solvability for (6.2.1) for a suitable class of initial data.

Theorem 6.6 *The solution to the Cauchy problem*

$$v(x,0) = \phi(x)$$

for the Burgers equation (6.2.1) exists and is unique for all $t > 0$. It can be represented in the following form

$$v(x,t) = -2\mu \frac{\partial}{\partial x} \log \left\{ \frac{1}{2\sqrt{\pi\,\mu\,t}} \int_{-\infty}^{\infty} \exp\left[-\frac{(x-y)^2}{4\mu\,t} - \frac{1}{2\mu} \int_0^y \phi(y')\,dy' \right] dy \right\} \tag{6.2.8}$$

Proof: The central step in the derivation of the formula (6.2.8) is in the following

Lemma 6.7 (Cole - Hopf transformation). *The substitution*

$$v = -2\mu \frac{\partial}{\partial x} \log u \tag{6.2.9}$$

transforms the Burgers equation (6.2.1) to the heat equation

$$u_t = \mu\, u_{xx}. \tag{6.2.10}$$

Proof: We have

$$v_t = 2\mu\, \frac{u_t u_x - u\, u_{xt}}{u^2}$$

$$v_x = 2\mu\, \frac{u_x^2 - u\, u_{xx}}{u^2}$$

$$v_{xx} = 2\mu\, \frac{3u\, u_x u_{xx} - u^2 u_{xxx} - 2u_x^3}{u^3}.$$

After substitution into the Burgers equation and division by (-2μ) we arrive at

$$0 = \frac{u\, (u_t - \mu\, u_{xx})_x - u_x\, (u_t - \mu\, u_{xx})}{u^2} = \frac{\partial}{\partial x} \frac{u_t - \mu\, u_{xx}}{u}.$$

So, if $u = u(x,t)$ satisfies heat equation (6.2.10) then the function v given by (6.2.9) satisfies Burgers equation. ■

We can now complete the proof of the Theorem. The solution to the heat equation (6.2.10) with the initial data $u(x,0) = \psi(x)$ can be represented by the Poisson integral

$$u(x,t) = \frac{1}{2\sqrt{\pi\, \mu\, t}} \int_{-\infty}^{\infty} e^{-\frac{(x-y)^2}{4\mu\, t}}\, \psi(y)\, dy.$$

According to (6.2.9) the initial data for the Burgers and heat equations must be related by

$$\phi(x) = -2\mu\, [\log \psi(x)]_x\, .$$

Integrating[2] one obtains

$$\psi(x) = e^{-\frac{1}{2\mu}\int_0^x \phi(x')\, dx'}.$$

Hence

$$u(x,t) = \frac{1}{2\sqrt{\pi\, \mu\, t}} \int_{-\infty}^{\infty} e^{-\frac{(x-y)^2}{4\mu\, t} - \frac{1}{2\mu}\int_0^y \phi(y')\, dy'}\, dy.$$

Applying the transformation (6.2.9) one arrives at the formula (6.2.8). ■

Example. Let us consider the solution to the Burgers equation (6.2.1) with the step-like initial data

$$\phi(x) = \left\{ \begin{array}{ll} 1, & x < 0 \\ -1, & x > 0 \end{array} \right. \tag{6.2.11}$$

[2] It is easy to see that another choice of the integration constant changes $u \mapsto c\, u$ with a nonzero constant c. Such a change leaves invariant the logarithmic derivative $\frac{\partial}{\partial x} \log u$.

Integrating one obtains the initial data for the heat equation The Poisson integral gives the solution in the form

$$\psi(x) = e^{\frac{|x|}{2\mu}}.$$

So

$$u(x,t) = \frac{1}{2\sqrt{\pi\,\mu\,t}}\left[\int_{-\infty}^{0} e^{-\frac{(x-y)^2}{4\mu\,t}-\frac{y}{2\mu}}dy + \int_0^{\infty} e^{-\frac{(x-y)^2}{4\mu\,t}+\frac{y}{2\mu}}dy\right]$$
$$= \frac{1}{2\sqrt{\pi\,\mu\,t}}\left[\int_{-\infty}^{0} e^{-\frac{(y-x+t)^2}{4\mu\,t}+\frac{t-2x}{4\mu}}dy + \int_0^{\infty} e^{-\frac{(y-x-t)^2}{4\mu\,t}+\frac{t+2x}{4\mu}}dy\right]$$
$$= \frac{1}{\sqrt{\pi}}\left[e^{\frac{t-2x}{4\mu}}\int_{\frac{x-t}{2\sqrt{\mu\,t}}}^{\infty} e^{-s^2}ds + e^{\frac{t+2x}{4\mu}}\int_{-\infty}^{\frac{x+t}{2\sqrt{\mu\,t}}} e^{-s^2}ds\right]$$
$$= \frac{1}{\sqrt{\pi}}\left\{e^{\frac{t-2x}{4\mu}}\left[\int_0^{\infty} e^{-s^2}ds - \int_0^{\frac{x-t}{2\sqrt{\mu\,t}}} e^{-s^2}ds\right] + e^{\frac{t+2x}{4\mu}}\left[\int_{-\infty}^{0} e^{-s^2}ds + \int_0^{\frac{x+t}{2\sqrt{\mu\,t}}} e^{-s^2}ds\right]\right\}$$
$$= \frac{1}{2}e^{\frac{t}{4\mu}}\left(e^{\frac{x}{2\mu}} + e^{-\frac{x}{2\mu}}\right) + \frac{1}{2}e^{\frac{t}{4\mu}}\left[e^{\frac{x}{2\mu}}\,\mathrm{Erf}\left(\frac{x+t}{2\sqrt{\mu\,t}}\right) - e^{-\frac{x}{2\mu}}\,\mathrm{Erf}\left(\frac{x-t}{2\sqrt{\mu\,t}}\right)\right]$$

where

$$\mathrm{Erf}(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{-s^2}ds$$

is the *error function.*

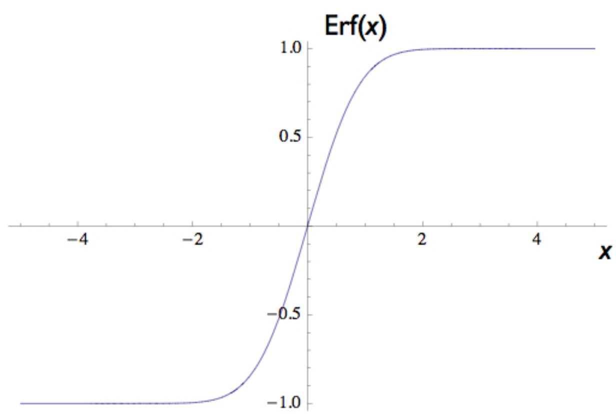

Fig. 7. Graph of the error function

Observe that the error function takes values very close to ± 1 away from the interval $(-2, 2)$; near the origin it is well approximated by the linear function with the slope

$$\mathrm{Erf}'(0) = \frac{2}{\sqrt{\pi}} \simeq 1.128.$$

Substitution to the formula (6.2.8) gives, after simple computations, the solution to the Burgers equation with the step-like initial data

$$v(x,t) = -\frac{\sinh\frac{x}{2\mu} + \frac{1}{2}\left[e^{\frac{x}{2\mu}}\,\mathrm{Erf}\,\frac{x+t}{2\sqrt{\mu\,t}} + e^{-\frac{x}{2\mu}}\,\mathrm{Erf}\,\frac{x-t}{2\sqrt{\mu\,t}}\right]}{\cosh\frac{x}{2\mu} + \frac{1}{2}\left[e^{\frac{x}{2\mu}}\,\mathrm{Erf}\,\frac{x+t}{2\sqrt{\mu\,t}} - e^{-\frac{x}{2\mu}}\,\mathrm{Erf}\,\frac{x-t}{2\sqrt{\mu\,t}}\right]} \tag{6.2.12}$$

When $t \to +0$ the arguments of the Erf functions tend to $\pm\infty$ for $x > 0$ or $x < 0$ respectively. So for positive x the numerator and the denominator both tend to $\sinh\frac{x}{2\mu} + \cosh\frac{x}{2\mu}$, thus the function $v(x,t)$ tends to -1. For negative x the numerator tends to $\sinh\frac{x}{2\mu} - \cosh\frac{x}{2\mu}$, and the denominator tends to $\cosh\frac{x}{2\mu} - \sinh\frac{x}{2\mu}$, thus $v(x,t) \to +1$.

It is also easy to describe the large time asymptotics of the solution (6.2.12). Indeed, for $t \to +\infty$ one has

$$\frac{x+t}{2\sqrt{\mu\,t}} \to +\infty, \quad \frac{x-t}{2\sqrt{\mu\,t}} \to -\infty.$$

Hence

$$\lim_{t\to+\infty} v(x,t) = -\tanh\frac{x}{2\mu}. \tag{6.2.13}$$

Observe that for small μ the limiting curve is very close to the original step-like profile.

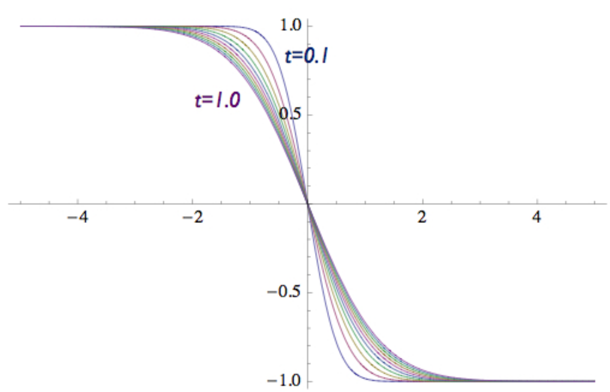

Fig. 8. Solution to the Burgers equation with $\mu = 1$ with the step-like initial data (6.2.11)

From Fig. 8 it is clear that, for small time the solution $v(x,t)$ departs rapidly from the initial data but then the deviation becomes more slow. The next picture suggests that the smaller is the viscosity μ the closer to the initial step-like data remains the solution.

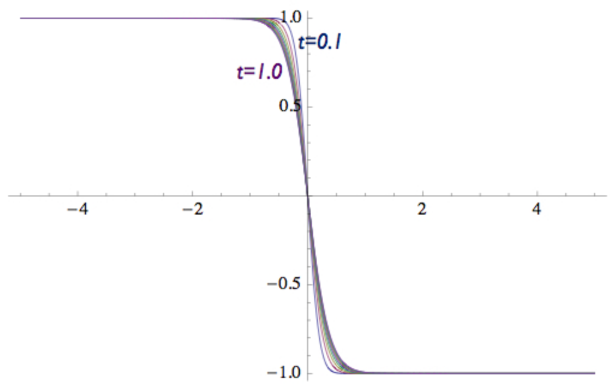

Fig. 9. Solution to the Burgers equation with $\mu = 0.2$ with the step-like initial data (6.2.11)

Exercise 6.8 *Prove that the solution (6.2.12) for any x, t (with $t > 0$) for $\mu \to 0$ tends to the step-like function (6.2.11).*

One can prove that, more generally speaking a generic solution to the Burgers equation in the limit of small viscosity $\mu \to 0$ tends to a discontinuous function within a certain region of the (x,t)-half-plane. In fluid dynamics such discontinuities can be interpreted as *shock waves*. The proof of this statement will not be given in the lectures. Our nearest goal is to study the behaviour of generic solutions to the Burgers equation for $\mu \to 0$. In the next section we will introduce a necessary analytic tool for such a study.

6.3 Asymptotics of Laplace integrals. Stationary phase asymptotic formula

If we rewrite the solution of Burgers' equation (6.2.8)

$$\begin{aligned} v(x,t) &= -2\mu\partial_x \ln\left\{\int_{\mathbb{R}} \exp\left[-\frac{1}{\mu}\left(\frac{(y-x)^2}{4t} + \frac{1}{2}\int^y \phi(s)\mathrm{d}s\right)\right]\mathrm{d}y\right\} = \\ &= -2\mu\partial_x \ln\left[\int_{\mathbb{R}} \mathrm{e}^{-\frac{1}{\mu}S(y;x,t)}\mathrm{d}y\right] = \frac{\displaystyle\int_{\mathbb{R}}(x-y)\mathrm{e}^{-\frac{1}{\mu}S(y;x,t)}\mathrm{d}y}{\displaystyle 2t\int_{\mathbb{R}}\mathrm{e}^{-\frac{1}{\mu}S(y;x,t)}\mathrm{d}y} \end{aligned} \tag{6.3.1}$$

where the **phase function** $S(y;x,t)$ is in our case

$$S(y;x,t) = \frac{(y-x)^2}{4t} + \frac{1}{2}\int^y \phi(s)\mathrm{d}s = \frac{(y-x)^2}{4t} + \frac{1}{2}Q(y) \tag{6.3.2}$$

(where we put $Q(y)$ an antiderivative of ϕ).

Problem: Find the behaviour of the solution $v(x,t)$ as $\mu \to 0_+$.

The type of integrals we are facing are the so–called *Laplace integrals* of the form

$$I(\mu) = \int_{\mathbb{R}} f(y) e^{-\frac{S(y)}{\mu}} \mathrm{d}y \tag{6.3.3}$$

with smooth functions $f(y)$, $S(y)$. The basic idea is that, for $\epsilon \to +0$ the main contribution to the integral comes from the minima of the phase function $S(x)$. More precise statements are contained in the following propositions.

Proposition 6.9 *Let $S : \mathbb{R} \to \mathbb{R}$ be $\mathcal{C}^2$ and bounded from below, and $f(y)$ continuous. Consider*

$$I(\mu) := \int_{\mathbb{R}} f(y) e^{-\frac{S(y)}{\mu}} \mathrm{d}y. \tag{6.3.4}$$

Then
[1] *If $I(\mu)$ converges for $\mu_0 > 0$ then it converges for every $0 < \mu < \mu_0$.*
[2] *Suppose now that $S(y)$ has a unique global minimum at $y = y_0$ with $S''(y_0) > 0$ (strict inequality) and that for $\delta > 0$ sufficiently small the set*

$$J_\delta = \left\{ y : \quad 0 \leqslant S(y) - S_{\min} < \delta \right\} \tag{6.3.5}$$

consists of a single open interval containing y_0. Then

$$I(\mu) = \sqrt{\frac{2\pi\mu}{S''(y_0)}} f(y_0) \mathrm{e}^{-\frac{S_{min}}{\mu}} \Big(1 + o(1) \Big) \tag{6.3.6}$$

Remark 6.10 *There are several generalizations to integrals on bounded intervals. If the functions S, f have higher regularity, the infinitesimal $o(1)$ can be shown to be an asymptotic series in μ.*

Proof. **[1]** Since S is bounded below, let $S_{\min} = \inf_{\mathbb{R}} S$. Suppose $I(\mu_0)$ is a convergent integral. Then for $\mu < \mu_0$ we have $-\frac{1}{\mu} < -\frac{1}{\mu_0}$ and

$$|I(\mu)| \leqslant \int_{\mathbb{R}} |f(y)| \mathrm{e}^{-\frac{S(y)}{\mu}} = \mathrm{e}^{-\frac{S_{\min}}{\mu}} \int_{\mathbb{R}} |f(y)| \mathrm{e}^{-\frac{S(y)-S_{\min}}{\mu}} < \mathrm{e}^{-\frac{S_{\min}}{\mu}} \int_{\mathbb{R}} |f(y)| \mathrm{e}^{-\frac{S(y)-S_{\min}}{\mu_0}} < \infty. \tag{6.3.7}$$

[2] Let J_δ as in (6.3.5) and the assumptions hold. Then we consider the expression

$$F(\mu) := \frac{\mathrm{e}^{\frac{S_{\min}}{\mu}}}{\sqrt{\mu}} I(\mu) = \frac{1}{\sqrt{\mu}} \left[\int_{J_\delta} f(y) \mathrm{e}^{-\frac{1}{\mu}(S(y)-S_{\min})} \mathrm{d}y + \int_{\mathbb{R}\setminus J_\delta} f(y) \mathrm{e}^{-\frac{1}{\mu}(S(y)-S_{\min})} \mathrm{d}y \right] \tag{6.3.8}$$

We show that the second integral above tends to zero faster than any positive power of μ. Indeed

$$\left| \int_{\mathbb{R}\setminus J_\delta} f(y) \mathrm{e}^{-\frac{1}{\mu}(S(y)-S_{\min})} \mathrm{d}y \right| \leqslant \mathrm{e}^{-\delta/\mu} \int_{\mathbb{R}\setminus J_\delta} |f(y)| \mathrm{e}^{-\frac{1}{\mu}(S(y)-S_{\min}-\delta)} \mathrm{d}y = (\star) \tag{6.3.9}$$

This latter integral is easily seen to be a increasing function of μ and bounded for $\mu = \mu_0$; hence

$$(\star) \leqslant K\mathrm{e}^{-\frac{\delta}{\mu}} \tag{6.3.10}$$

for $\mu < \mu_0$. Thus this expression tends to zero faster than any power of μ. We express this by the symbol $\mathcal{O}(\mu^\infty)$. Thus we have shown

$$F(\mu) = \frac{1}{\sqrt{\mu}} \int_{J_\delta} f(y)\mathrm{e}^{-\frac{1}{\mu}(S(y)-S_{\min})}\mathrm{d}y + \mathcal{O}(\mu^\infty). \tag{6.3.11}$$

Now define

$$\eta(y) := \sqrt{S(y) - S(y_0)}\mathrm{sign}\,(y - y_0). \tag{6.3.12}$$

We observe that $\eta(y)$ is a $\mathcal{C}^1$ function with $\eta'(y) > 0)$ in J_δ. In particular

$$\eta'(y_0) = \sqrt{\frac{S''(y_0)}{2}} \tag{6.3.13}$$

We can thus rewrite the integral over J_δ in the coordinate η

$$\frac{1}{\sqrt{\mu}} \int_{J_\delta} f(y)\mathrm{e}^{-\frac{1}{\mu}(S(y)-S_{\min})}\mathrm{d}y = \frac{1}{\sqrt{\mu}} \int_{-\sqrt{\delta}}^{\sqrt{\delta}} f(y(\eta))\frac{\mathrm{d}y}{\mathrm{d}\eta}\mathrm{e}^{-\frac{\eta^2}{\mu}}\mathrm{d}\eta \tag{6.3.14}$$

The latter integral is of the general form

$$\frac{1}{\sqrt{\mu}} \int_{-\sqrt{\delta}}^{\sqrt{\delta}} g(\eta)\mathrm{e}^{-\frac{\eta^2}{\mu}}\mathrm{d}\eta \tag{6.3.15}$$

with $g(\eta)$ continuous over the interval. Letting $\widehat{g}(\eta) := g(\eta)\chi_J(\eta)$ we can write the integral as

$$\frac{1}{\sqrt{\mu}} \int_{\mathbb{R}} \widehat{g}(\eta)\mathrm{e}^{-\frac{\eta^2}{\mu}}\mathrm{d}\eta = \int_{\mathbb{R}} g(\sqrt{\mu}s)\mathrm{e}^{-s^2}\mathrm{d}s. \tag{6.3.16}$$

By Lebesgue dominated convergence theorem, we can take the limit under the integral sign as $\mu \to 0_+$ and see that the limit gives $g(0)\sqrt{\pi}$. Thus,

$$\frac{1}{\sqrt{\mu}} \int_{-\sqrt{\delta}}^{\sqrt{\delta}} f(y(\eta))\frac{\mathrm{d}y}{\mathrm{d}\eta}\mathrm{e}^{-\frac{\eta^2}{\mu}}\mathrm{d}\eta = f(y_0)\sqrt{\pi}\frac{1}{\left.\frac{\mathrm{d}\eta}{\mathrm{d}y}\right|_{y=y_0}} + o(1) = f(y_0)\sqrt{\frac{2\pi}{S''(y_0)}} + o(1) \tag{6.3.17}$$

Thus, summarizing

$$\frac{\mathrm{e}^{\frac{S_{\min}}{\mu}}}{\sqrt{\mu}} I(\mu) = f(y_0)\sqrt{\frac{2\pi}{S''(y_0)}} + o(1). \tag{6.3.18}$$

whence the asymptotic formula

$$I(\mu) = f(y_0)\sqrt{\frac{2\pi\mu}{S''(y_0)}}\mathrm{e}^{-\frac{S_{\min}}{\mu}}(1 + o(1)). \tag{6.3.19}$$

This concludes the proof. We remark that if we have higher regularity for f and S one can prove that the term $o(1)$ is $o(\mu^k)$ with k the smoothness class of f and S''. In particular if S, f are analytic, then there is a full asymptotic expansion in place of $o(1)$ (in the Poincaré sense). ■

Let us apply the Laplace formula to the study of small viscosity solutions to the Burgers equation (6.2.1). According to the previous Section the solution is proportional to the logarithmic derivative of the function

$$u(x,t) = \frac{1}{2\sqrt{\pi\,\mu\,t}} \int_{-\infty}^{\infty} e^{-\frac{S(y;x,t)}{\mu}}\,dy$$

where the phase function $S(y; x, t)$ depending on the parameters x, t is given by

$$S(y;x,t) = \frac{(x-y)^2}{4t} + \frac{1}{2}\int_0^y \phi(y')\,dy'. \tag{6.3.20}$$

Here $\phi(x)$ is the initial data for the Burgers equation.

Theorem 6.11 *Let $\phi(x)$ be a monotone increasing smooth function. Then*
[1] *The solution of the Hopf equation*

$$w_t + ww_x = 0, \qquad w(x,0) = \phi(x) \tag{6.3.21}$$

exists globally.
[2] *The solution $v = v(x,t;\mu)$ to the Cauchy problem to the Burgers equation*

$$v_t + vv_x = \mu v_{xx}, \qquad v(x,0;\mu) = \phi(x) \tag{6.3.22}$$

with the initial data $\phi(x)$ satisfies

$$|v(x,t;\mu) - w(x,t)| \to 0 \quad \text{for} \quad \mu \to 0 \tag{6.3.23}$$

uniformly on compact subsets of the half-plane (x,t). The same asymptotics (6.3.23) holds true for monotone decreasing initial data for the times before the time t_0 of gradient catastrophe for the solution to the Hopf equation (6.3.21) provided that the derivative $\phi'(x)$ of the initial function is bounded on the real line.

Proof. **[1]** Since $\phi(x)$ is increasing the method of characteristics gives a global solution;

$$w = \phi(x) \quad x = \Psi(w) \tag{6.3.24}$$

and

$$tw + \Psi(w) = x \tag{6.3.25}$$

Note that $\Psi(w)$ is an increasing function and $t > 0$, so that the LHS is a monotone increasing function and hence defines a unique function $w(x,t)$ for all $t > 0$ and $x \in \mathbb{R}$.

[2] The stationary point $y = y(x,t)$ of the phase function is determined from the equation

$$S_y(y;x,t) = \frac{y-x}{2t} + \frac{1}{2}\phi(y) = 0$$

equivalent to

$$x = y + t\,\phi(y). \tag{6.3.26}$$

For $t = 0$ the solution is unique, $y(x,0) = x$. For a monotone increasing function ϕ the solution remains a unique one also for all $t > 0$ since the y-derivative of the equation (6.3.26) remains positive for all $y \in \mathbb{R}$. This stationary point is a nondegenerate minimum. Indeed, the second derivative at the stationary point is always positive

$$S_{yy}(y(x,t);x,t) = \frac{1 + t\,\phi'(y)}{2t} > 0.$$

Applying the Laplace formula one obtains

$$u(x,t) = \frac{1}{\sqrt{1 + t\,\phi'(y)}} e^{-\frac{S(y(x,t);x,t)}{\mu}} \left(1 + \mathcal{O}(\mu)\right).$$

Taking the logarithmic derivative yields

$$v(x,t) = -2\mu \frac{\partial u(x,t)}{\partial t} = 2S_x(y(x,t);x,t) + \mathcal{O}(\mu) = \phi(y(x,t)) + \mathcal{O}(\mu).$$

It remains to observe that the function $w = \phi(y(x,t))$ satisfies the implicit function equation

$$x = \Psi(w) + t\,w$$

where, as above, the function Ψ is the inverse function to ϕ. Thus $w = w(x,t)$ coincides with the solution to the Cauchy problem (6.3.21).

For the case of monotone decreasing initial function $\phi(x)$ with bounded derivative $\phi'(x)$ all above arguments remain valid for small times, $t < t_0$, where t_0 is the time of the gradient catastrophe for the solution to the Cauchy problem (6.3.21). ■

6.3.1 A worked out example; shock formation

What happens when ϕ is not increasing? We know that if ϕ has intervals of decrease, then eventually the solution of Hopf equation given by the method of characteristics will cease to exist because it gives ambiguous solutions.

By the way of example we consider in some detail the initial condition $\phi(x) = 5\mathrm{e}^{-x^2}$. We know that the point of gradient catastrophe occurs for

$$(x_0, t_0, w_0) = \left(\sqrt{2}, \frac{\sqrt{2}\mathrm{e}^{1/2}}{10}, 5\mathrm{e}^{-\frac{1}{2}}\right). \tag{6.3.27}$$

Now the limit of Burgers' equation $v(x,t;\mu)$ tends to

$$v(x,t;0^+) = \phi(y_0(x,t)) = 5\mathrm{e}^{-y_0^2(x,t)} \tag{6.3.28}$$

where $y_0(x,t)$ is the point of **absolute** minimum of $S(y;x,t)$:

$$S(y;x,t) = \frac{(y-x)^2}{4t} + \frac{1}{2}\int_0^y \phi(s)\mathrm{d}s = \frac{1}{4}\frac{(y-x)^2}{t} + \frac{5\sqrt{\pi}}{4}\mathrm{Erf}\,(y) \tag{6.3.29}$$

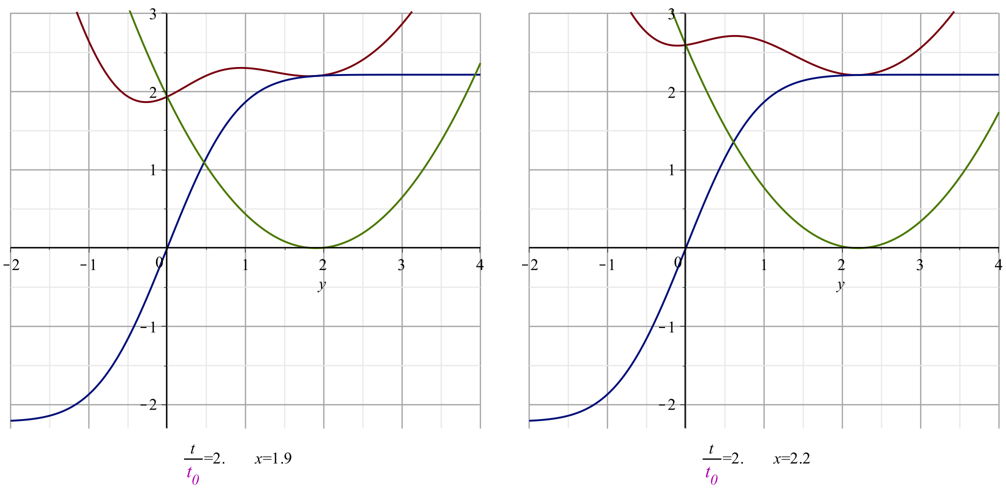

Figure 6.2: The plot of S (red) for $t/t_0 = 2$ and for two different values of x. Note that S has three critical points, of which two are local minima. The location of the absolute minimum has a discontinuity w.r.t. x somewhere in between the two indicated values. Note that for $x = 1.9$ the minimum is on the left and for $x = 2.2$ the minimum is the other critical point. In green is plotted $(y-x)^2/4t$ and blue is the term containing the Erf function.

Let us consider $t > t_0$; the plot of $S(y; x, t)$ is shown in Fig. 6.2. As one can see, there are as many as three critical points y_0, y_1, y_2 depending on x, t, of which two are local minima and the one in between is a local maximum. According to the strategy for Laplace integrals, only the **absolute** maximum enters in the leading order asymptotics.

It does not take long to convince oneself that there is a critical $x_\star(t)$ such that:

for $x < x_\star(t)$ the global minimum of S is at $y_0(x,t)$ and for $x > x_\star(t)$ it is at $y_2(x,t)$.

Therefore the solution $v(x, t; 0^+)$ exhibits a discontinuous behaviour:

$$v(x,t;0^+) = \begin{cases} 5\mathrm{e}^{-(y_0(x,t))^2} & x < x_\star(t) \\ 5\mathrm{e}^{-(y_2(x,t))^2} & x > x_\star(t) \end{cases} \tag{6.3.30}$$

Example 6.12 *The Euler Gamma function*

$$\Gamma(x) = \int_0^\infty \mathrm{e}^{-s} s^{-x} \frac{\mathrm{d}s}{s} \tag{6.3.31}$$

can be studied in the limit $x \to +\infty$ by the use of Laplace integrals (here the interval of integration is $[0, \infty)$ but it won't cause problems as we see not). With a change of variable $s = xy$ we get

$$\Gamma(x) = \int_0^\infty \mathrm{e}^{-s-x\ln s} \frac{\mathrm{d}s}{s} = \mathrm{e}^{x\ln x} \int_0^\infty \mathrm{e}^{-x(y+\ln y)} \frac{\mathrm{d}y}{y}. \tag{6.3.32}$$

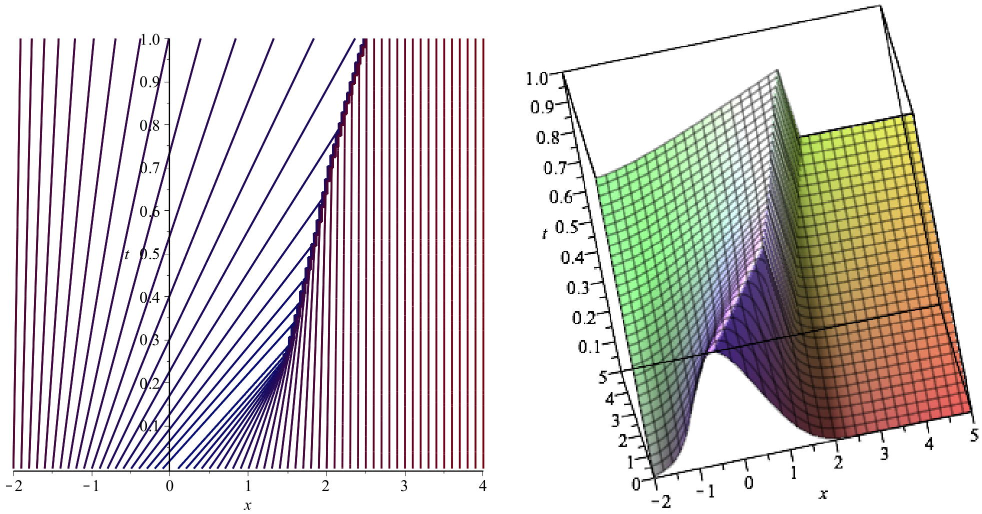

Figure 6.3: Plot of the level lines of $v(x, t; 0^+)$. Note that the level lines follow the characteristics and there is a visible line of shock (the jagged line; this jaggedness is just an issue with the numerics, the shock-line is smooth in reality). The same plot in 3d on the right.

The phase function is $S(y) = y + \ln y$ *which has an absolute minimum at* $y = 1$ *with* $S''(1) = 1$ *and* $S(1) = 1$*. Thus the formula of Laplace gives (here* $\mu = \frac{1}{x}$*)*

$$\Gamma(x) = \sqrt{\frac{2\pi}{x}} \mathrm{e}^{-x \ln x - x}(1 + \mathcal{O}(x^{-1})) \tag{6.3.33}$$

In this case the $\mathcal{O}(x^{-1})$ *can be expanded into a full asymptotic series in* x^{-1} *and the coefficients have been worked out.*

6.3.2 Oscillatory integrals

At the end of this Section let us give, without proof, the complex version of the Laplace formula. This is the so-called *stationary phase formula* for the asymptotics of the integrals with complex phase function

$$I(\epsilon) = \int_a^b f(x) e^{\frac{i\, S(x)}{\epsilon}} dx. \tag{6.3.34}$$

Like in the case of Laplace integrals the localization principle says that the main contribution to the asymptotics comes from the stationary points of the phase function $S(x)$ and from the boundary of the integration segment. However, differently from the Laplace method, the stationary phase asymptotics involve contributions from *all* stationary points of the phase function, not only from the minima. More precisely,

Proposition 6.13 *Let* $f(x)$*,* $S(x)$ *be* $\mathcal{C}^\infty$ *functions, such that* $f(x)$ *vanishes at the boundary of the segment* $[a, b]$ *with all derivatives, and* $S'(x) \neq 0$ $\forall x \in (a, b)$*. Then*

$$\int_a^b f(x) e^{\frac{i\, S(x)}{\epsilon}} dx = \mathcal{O}(\epsilon^n) \quad \forall n \in \mathbb{Z}_{>0}.$$

Proposition 6.14 *Let $f(x)$, $S(x)$ be $\mathcal{C}^\infty$ functions, such that $f(x)$ vanishes at the boundary of the segment $[a,b]$ with all derivatives, and $S(x)$ has a unique nondegenerate stationary point $x_0 \in (a,b)$. Then*

$$\int_a^b f(x) e^{\frac{i\,S(x)}{\epsilon}} dx = \sqrt{\frac{2\pi\,\epsilon}{|S''(x_0)|}} e^{\frac{i\,S(x_0)}{\epsilon} + \frac{i\,\pi}{4} \operatorname{sign} S''(x_0)} \left(f(x_0) + \mathcal{O}(\epsilon) \right). \tag{6.3.35}$$

The crucial step in the derivation of the stationary phase formula is in the computation of the following integral.

Exercise 6.15 *Prove that*

$$\int_{-1}^1 e^{\frac{ix^2}{\epsilon}} dx = \sqrt{\pi\,\epsilon}\, e^{\frac{i\pi}{4}} \left(1 + \mathcal{O}(\epsilon)\right). \tag{6.3.36}$$

6.4 Dispersive waves. Solitons for KdV

We are now in a position to explain the effect of dispersion in the theory of linear waves. Let us assume that a linear PDE admits plane wave solutions

$$v(x,t) = a\, e^{i(kx - \omega(k)t)} \tag{6.4.1}$$

for any real k. Moreover we assume that the *dispersion law*

$$\omega = \omega(k) \tag{6.4.2}$$

is a real valued function satisfying

$$\omega''(k) \neq 0 \quad \text{for} \quad k \neq 0. \tag{6.4.3}$$

These assumptions hold true, e.g., for the linearized KdV equation (6.2.7) where

$$\omega(k) = c\,k - \epsilon^2 k^3.$$

Another example is given by the Klein–Gordon equation

$$v_{tt} - v_{xx} + m^2 v = 0. \tag{6.4.4}$$

In this case the dispersion relation splits into two branches

$$\omega(k) = \pm\sqrt{k^2 + m^2}. \tag{6.4.5}$$

For the linear wave equation

$$v_{tt} = a^2 v_{xx}$$

the dispersion relation reads

$$\omega(k) = \pm a\, k.$$

The condition (6.4.3) does not hold.

More general solutions can be written as linear superpositions of the plane wave

$$v(x,t) = \int_K a(k) e^{i(kx-\omega(k)t)} dk \tag{6.4.6}$$

where the integration is taken over a domain in the space of wave numbers. Here $a(k)$ is the complex amplitude of the k-th wave. Let us describe the asymptotic behaviour of the solution (6.4.6) for large x and t. More precisely the question is: what will see the observer moving with a constant speed c for sufficiently large time? The answer is given by the following

Lemma 6.16 *Let us assume that the equation*

$$c = \omega'(k) \tag{6.4.7}$$

has a unique root $k = k_0$ belonging to K. Then for $t \to \infty$ the solution (6.4.6) restricted onto the line

$$x = ct + x_0$$

behaves as follows

$$v(x,t) = \sqrt{\frac{2\pi}{t\,|\omega''(k_0)|}}\, a(k_0) e^{it[ck_0-\omega(k_0)]-\frac{i\pi}{4}\operatorname{sign}\omega''(k_0)+ik_0x_0} \left(1 + \mathcal{O}\left(\frac{1}{t}\right)\right). \tag{6.4.8}$$

Proof: It follows immediately from the stationary phase formula (6.3.35). ■

Let us apply the result of the Lemma to the case of *wave-trains*, i.e., solutions of the form (6.4.6) obtained by integration over a small neighborhood of a point k_*. In this case the remote observer will be able to detect a nonzero value of the wave only if

$$\frac{x}{t} \simeq \omega'(k_*).$$

We conclude that., from the point of view of the remote observer the wave-train with the wave number k_* propagates with the velocity $\omega'(k_*)$. For this reason the number $\omega'(k_*)$ is called the *group velocity* of the wave-train.

In short we can say that the velocity of propagation of dispersive waves depends on the wave number.

The linearized KdV equation is an example of a dispersive system. Indeed, the group velocity is equal to

$$\omega'(k) = c - 3\epsilon^2 k^2.$$

That means that **the rapidly oscillating (i.e., $|k| >> 1$) small perturbations propagate from right to left**. At the same time, as we know from the analysis of Hopf equation, **the slow varying solutions with positive magnitude propagate from left to right**.

The full mathematical theory of solutions to the KdV equation is too complicated to present here. Here we will present only a small output of this theory describing an important class of particular solutions to the KdV equation. They are created by a balance between the nonlinear and dispersive effects. The idea is to look for solutions in the form of simple waves

$$v(x,t) = V\left(\frac{x-ct}{\epsilon}\right). \tag{6.4.9}$$

Substitution to (6.2.2) yields an ODE for the function $V = V(X)$

$$-c\,V' + V\,V' + V''' = 0.$$

Integrating one obtains a second order ODE

$$V'' + \frac{1}{2}V^2 - c\,V = a \tag{6.4.10}$$

where a is an integration constant. This equation can be interpreted as the Newton law for the motion of a particle in the field of a cubic potential

$$V'' = -\frac{\partial P(V)}{\partial V}, \quad P(V) = \frac{1}{6}V^3 - \frac{c}{2}V^2 - a\,V. \tag{6.4.11}$$

One should expect to apply the law of conservation of energy to integration of this equation. Indeed, after multiplication of (6.4.10) by V' one can integrate once more to arrive at a first order equation

$$\frac{1}{2}V'^2 + \frac{1}{6}V^3 - \frac{c}{2}V^2 - a\,V = b$$

where b is another integration constant (the total energy of the system (6.4.11)). The last equation can be integrated by quadratures

$$X - X_0 = \int \frac{dV}{\sqrt{2\left(-\frac{1}{6}V^3 + \frac{c}{2}V^2 + a\,V + b\right)}}. \tag{6.4.12}$$

For general values of the constants a, b, c the solution (6.4.12) can be expressed via elliptic functions. We will now determine the values of these parameters that allow a reduction to elementary functions. This can happen when the cubic polynomial under the square root has a multiple root. Moreover we will assume that this double root is at $V = 0$. To meet such a requirement one must have

$$a = b = 0.$$

We arrive at computation of the integral

$$X - X_0 = \int \frac{dV}{V\sqrt{c - \frac{1}{3}V}} = -\frac{2}{\sqrt{c}}\tanh^{-1}\frac{\sqrt{c - \frac{V}{3}}}{\sqrt{c}}$$

Inverting one obtains

$$V = 3c\left(1 - \tanh^2\frac{\sqrt{c}(X - X_0)}{2}\right) = \frac{3c}{\cosh^2\frac{\sqrt{c}(X-X_0)}{2}}$$

We arrive at the following family of solutions to the KdV equation

$$v(x,t) = \frac{3k^2}{\cosh^2\frac{k(x-x_0)-k^3t}{2\epsilon}} \tag{6.4.13}$$

where we put $k = \sqrt{c}$.

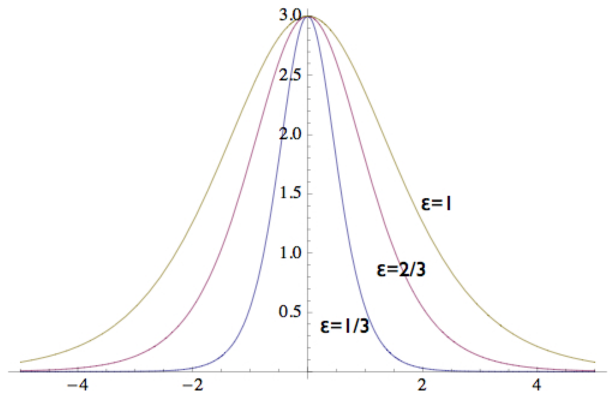

Fig. 10. Soliton solutions to the KdV equation with $t = 0$, $k = 1$ for various values of ϵ

6.5 Exercises for Chapter 6

Exercise 6.17 *Derive the following formula for the solution to the Cauchy problem*

$$\delta v(x,0) = \phi(x)$$

for the linearized Burgers equation (6.2.4):

$$\delta v(x,t) = \frac{1}{2\sqrt{\pi\,\mu\,t}}\int_{-\infty}^{\infty} e^{-\frac{(x-y-ct)^2}{4\mu\,t}}\phi(y)\,dy.$$

Exercise 6.18 *Obtain the following representation for solutions to the linearized KdV equation (6.2.7) with the initial data $\delta v(x,0) = \phi(x)$ rapidly decreasing at $|x| \to \infty$:*

$$\delta v(x,t) = \int_{-\infty}^{\infty} A(x-y-c\,t, \epsilon^2 t)\phi(y)\,dy$$

where

$$A(x,t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i(k\,x+k^3 t)}dk. \tag{6.5.1}$$

The integral (6.5.1) can be expressed via *Airy function*

$$A(x,t) = \frac{1}{(3t)^{1/3}} Ai\left(\frac{x}{(3t)^{1/3}}\right)$$

defined by the integral

$$Ai(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i\left(s\,x+\frac{s^3}{3}\right)}ds.$$

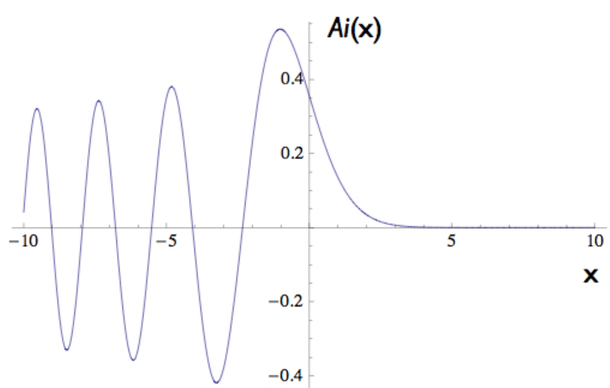

Fig. 11. Graph of Airy function

Chapter 7

Cauchy problem for systems of PDEs. Cauchy - Kovalevskaya theorem

7.1 Formulation of Cauchy problem

Consider a system of order $K \geqslant 1$ of n PDEs for the functions $u_1(x,t), \ldots, u_n(x,t)$ of the form

$$\frac{\partial^K u_i}{\partial t^K} = f_i\left(t, x, \mathbf{u}, \mathbf{u}_t, \ldots, \mathbf{u}_x, \mathbf{u}_{xx}, \ldots, \partial_t^\ell \partial_x^m \mathbf{u}\right), \quad i = 1, \ldots, n. \tag{7.1.1}$$

$$\ell \leqslant K-1, \quad \ell + m \leqslant K \tag{7.1.2}$$

where the functions f_j depend on the derivatives of $\mathbf{u}(t,x)$ up to $\ell = K-1$ in t. We say that the vector-valued function $\mathbf{u}(x,t) = (u_1(x,t), \ldots, u_n(x,t))$ defined for $x \in (x_0, x_1)$, $t \in (t_0, t_1)$ satisfies the system (7.1.1) if, after the substitution

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

$$\mathbf{u} = (u_1(x,t), \ldots, u_n(x,t)), \quad \mathbf{u}_x = \left(\frac{\partial u_1(x,t)}{\partial x}, \ldots, \frac{\partial u_n(x,t)}{\partial x}\right), \ldots$$

the system is identically satisfied. Without loss of generality we can assume that

$$t_0 < 0 < t_1.$$

The *Cauchy problem* is formulated as follows. Given n functions $\phi_1(x), \ldots, \phi_n(x)$ find a solution $u_1(x,t), \ldots, u_n(x,t)$ defined for $0 \leqslant t \leqslant t_1$ such that

$$u_1(x,0) = \phi_1(x), \ldots, u_n(x,0) = \phi_n(x). \tag{7.1.3}$$

In the next section we will prove that the Cauchy problem (7.1.1), (7.1.3) has a unique solution provided the right hand sides of the equations and the initial data are analytic functions.

With some preliminary manipulation similar to those used to reduce higher order ODEs to (systems) of first order ODEs one can always reduce the problem to the first order case $K = 1$ where the equation is thus

$$\mathbf{u}_t = \mathbf{f}(t, x, \mathbf{u}, \mathbf{u}_x). \tag{7.1.4}$$

Remark 7.1 *The theorem can be formulated also in the case where x is a vector; this generalization only adds complication in the notation but introduces no significant conceptual hurdles.*

Example 7.2 *Consider the case of a scalar PDE of second order*

$$u_{tt} = f(t, x, u, u_x, u_t, u_{tx}, u_{xx}) \tag{7.1.5}$$

If we introduce the vector $\mathbf{y} = [u, u_x, u_t] = [q_0, q_1, p]$ we note that now f depends on $\mathbf{y}, \mathbf{y_x}$ and the system takes the form

$$\begin{bmatrix} q_0 \\ q_1 \\ p \end{bmatrix}_t = \begin{bmatrix} u \\ u_x \\ u_t \end{bmatrix}_t = \begin{bmatrix} p \\ p_x \\ f(t, x, q_0, q_1, p, p_x, q_{1,x}) \end{bmatrix} \tag{7.1.6}$$

so that we are reduced to a system of first order PDEs. □

The idea of the proof is very simple: using the system (7.1.1) we can compute the time derivatives of any order of the solution at the point $t = 0$. For example for the first derivative we have

$$\dot{\phi}_i(x) := \frac{\partial u_i}{\partial t}|_{t=0} = f_i\left(0, x, \phi(x), \phi_x(x)\right),$$

$$\ddot{\phi}_i(x) := \frac{\partial^2 u_i}{\partial t^2}|_{t=0} = \frac{\partial f_i}{\partial t} + \sum_{j=1}^n \frac{\partial f_i}{\partial u_j} f_j + \sum_{j=1}^n \frac{\partial f_i}{\partial u_{j,x}} \partial_x f_j$$

etc. Here all the functions f_i, f_j and their derivatives have to be evaluated at the point $(0, x, \phi(x), \phi_x(x))$. The operator ∂_x is defined as follows

$$\partial_x f\left(t, x, \mathbf{u}, \mathbf{u}_x\right) = \frac{\partial}{\partial x} f\left(t, x, \mathbf{u}, \mathbf{u}_x\right) + \sum_{j=1}^n \left(u'_j \frac{\partial}{\partial u_j} + u''_j \frac{\partial}{\partial u'_j} \right) f\left(t, x, \mathbf{u}, \mathbf{u}_x\right). \tag{7.1.7}$$

(we use interchangeably the flyspeck $'$ notation or the subscript $_x$ for the differentiation w.r.t. x). In a similar way one can compute all the derivatives $\partial^k u_i / \partial t^k$ at $t = 0$. We obtain then the solution in the form of Taylor series

$$u_i(x, t) = \phi_i(x) + \dot{\phi}_i(x) \frac{t}{1!} + \ddot{\phi}_i(x) \frac{t^2}{2!} + \dots \tag{7.1.8}$$

In the next section we will prove convergence of this series.

7.2 Cauchy - Kovalevskaya theorem

Theorem 7.3 *Let the functions in the right hand sides of the system (7.1.4) be analytic in some neighborhood of the point*

$$t = 0, \quad x = 0, \quad \mathbf{u} = \phi(0), \quad \mathbf{u}_x = \phi'(0). \tag{7.2.1}$$

Moreover assume that the initial data (7.1.3) is analytic at $x = 0$. Then the Cauchy problem (7.1.1), (7.1.3) has a unique solution analytic in some neighborhood of the point $x = t = 0$.

Proof: First of all, without loss of generality, we assume $\phi(0) = \mathbf{0} = \phi'(0)$; simply re-define $\mathbf{u}(x,t) \mapsto \mathbf{u}(x,t) - \phi(0) - x\phi'(0)$.] At the first step we will reduce the Cauchy problem (7.1.1), (7.1.3) to another problem for a system of first order quasilinear equations. For simplicity let us consider the case $n = 1$, $m = 1$:

$$u_t = f(t, x, u, u_x), \quad u(x, 0) = \phi(x). \tag{7.2.2}$$

Introduce new functions

$$p = u_t, \quad q = u_x.$$

One obtains a first order quasilinear (i.e., linear in derivatives) system of three equations

$$\begin{aligned} u_t &= p \\ q_t &= p_x \\ p_t &= f_t(t, x, u, q) + f_u(t, x, u, q)p + f_q(t, x, u, q)p_x \end{aligned} \tag{7.2.3}$$

along with the initial data

$$u(x, 0) = \phi(x), \quad q(x, 0) = \phi'(x), \quad p(x, 0) = f(0, x, \phi(x), \phi'(x)). \tag{7.2.4}$$

Conversely, let us show that the Cauchy problem (7.2.3), (7.2.4) gives a solution to the original Cauchy problem (7.2.2). First, using the first and the last equations one obtains

$$u_{tt} = p_t = \frac{\partial}{\partial t} f(t, x, u, q).$$

Integrating in t we obtain

$$u_t = f(t, x, u, q) + g(x)$$

where $g(x)$ is the integration constant. At $t = 0$ we have

$$u_t(x, 0) = p(x, 0) = f(0, x, \phi(x), \phi'(x)).$$

Hence $g(x) \equiv 0$, that is

$$u_t = f(t, x, u, q).$$

Next, differentiating the first equation in (7.2.3) in x and using the second equation gives

$$u_{xt} = q_t.$$

Integrating in t we arrive at

$$u_x = q + h(x)$$

with a new integration constant $h(x)$. The initial data (7.2.4) imply

$$u_x(x, 0) = \phi'(x) = q(x, 0).$$

So $h(x) \equiv 0$ and thus $u_x = q$.

We have reduced the original problem to a Cauchy problem for a system of first order quasilinear equations

$$\mathbf{u}_t = \mathbf{A}(t, x, \mathbf{u})\mathbf{u_x} + \mathbf{b}(t, x, \mathbf{u}) \tag{7.2.5}$$

with a Cauchy data

$$\mathbf{u}(x, 0) = \phi(x). \tag{7.2.6}$$

Here $\mathbf{A}(t, x, \mathbf{u})$ and $\mathbf{b}(t, x, \mathbf{u})$ are some matrix-valued and vector-valued functions respectively. At the next step we eliminated the explicit dependence on x and t by means of the following trick. Introduce two new unknown functions τ, σ and consider the new Cauchy problem

$$\begin{aligned} \mathbf{u}_t &= \mathbf{A}(\tau, \sigma, \mathbf{u})\mathbf{u}_{\mathbf{x}} + \mathbf{b}(\tau, \sigma, \mathbf{u})\sigma_x \\ \tau_t &= \sigma_x \\ \sigma_t &= 0 \\ \\ \mathbf{u}(x, 0) &= \phi(x), \quad \tau(x, 0) = 0, \quad \sigma(x, 0) = x. \end{aligned} \tag{7.2.7}$$

It can easily be derived that the functions $\tau(x, t)$, $\sigma(x, t)$ satisfying (7.2.7) must be of the form

$$\tau(x, t) = t, \quad \sigma(x, t) = x.$$

So $\sigma_x \equiv 1$ and the problem (7.2.7) is equivalent to (7.2.5), (7.2.6).

We have arrived at a system of first order quasilinear PDEs with coefficients with no explicit dependence on the space-time variables x and t. Moreover, the right hand sides of the system are linear *homogeneous* functions of the derivatives. For these reasons it suffices to prove the Theorem for the Cauchy problem of the form

$$\begin{aligned} \mathbf{u}_t &= \mathbf{A}(\mathbf{u})\mathbf{u}_x \\ \\ \mathbf{u}(x, 0) &= \phi(x) \end{aligned} \tag{7.2.8}$$

with

$$\mathbf{u} = (u_1(x, t), \dots, u_n(x, t)), \quad \mathbf{A}(\mathbf{u}) = (A_{ij}(\mathbf{u}))_{1 \leqslant i,j \leqslant n}, \quad \phi(x) = (\phi_1(x), \dots, \phi_n(x)).$$

We will now apply the procedure of solving the system (7.2.8) in the form of power series explained in the previous section and prove convergence of this procedure.

Without loss of generality we may assume that

$$\phi(0) = 0.$$

Indeed, if this was not the case then one can shift the dependent function

$$\mathbf{u} \mapsto \mathbf{u} - \phi(0).$$

The analyticity assumption implies that the functions $\phi_i(x)$ and $A_{ij}(\mathbf{u})$ can be represented as sums of power series

$$\begin{aligned} \phi_i(x) &= \sum_{p=1}^{\infty} \phi_{i,p} x^p \\ \\ A_{ij}(\mathbf{u}) &= \sum_{p_1, \dots, p_n = 0}^{\infty} A_{ij,\mathbf{p}} u_1^{p_1} \dots u_n^{p_n} \end{aligned} \tag{7.2.9}$$

convergent for

$$|x| < \rho, \quad |u_i| < r, \quad i = 1, \dots, n. \tag{7.2.10}$$

We want to prove that the Cauchy problem (7.2.8) admits a solution in the form of a power series

$$u_i(x,t) = \sum_{p,q\geqslant 0} u_{i,pq} x^p t^q, \quad u_{i,0,0} = 0 \tag{7.2.11}$$

convergent for

$$|x| < \delta, \quad |t| < \tau \tag{7.2.12}$$

for some positive δ, τ. From the previous arguments it is clear that such a solution is unique.

Observe that the coefficients $u_{i,pq}$ can be expressed as polynomials in $\phi_{i,p}$, $A_{ij,\mathbf{p}}$

$$u_{i,pq} = P_{i,pq}\left(\phi_{j,r}, A_{jk,\mathbf{s}}\right) \tag{7.2.13}$$

with *universal* coefficients. Universality means that these coefficients do not depend on the particular choice of the system. For example,

$$u_{i,p\,0} = \phi_{i,p}.$$

In order to compute the coefficients $u_{i,p\,1}$ of the Taylor expansion of the function

$$\frac{\partial u_i(x,0)}{\partial t}$$

one has to use the equations (7.2.8) along with the initial data:

$$\sum_{p\geqslant 0} u_{i,p\,1} x^p = \sum_{j=1}^n \sum_{s_1,\dots,s_n} A_{ij,\mathbf{s}} \phi_1^{s_1}(x)\dots\phi_n^{s_n}(x)\phi_j'(x). \tag{7.2.14}$$

Expanding the right hand sides in Taylor series one obtains expressions for $u_{i,p\,1}$. For example,

$$u_{i,0\,1} = \sum_{j=1}^n A_{ij,0}\phi_{j,1}$$

etc. Observe that the assumption $\phi(0) = 0$ is crucial to arrive at polynomial expressions. It is clear that the coefficients of these polynomials are nonnegative integers.

We will consider also another Cauchy problem

$$\begin{aligned} \mathbf{v}_t &= \mathbf{B}(\mathbf{v})\mathbf{v}_x \\ \mathbf{v}(x,0) &= \psi(x) \end{aligned} \tag{7.2.15}$$

of the same size with analytic initial data and analytic coefficients

$$\begin{aligned} \psi_i(x) &= \sum_{p=1}^\infty \psi_{i,p} x^p \\ B_{ij}(\mathbf{v}) &= \sum_{p_1,\dots,p_n=0}^\infty B_{ij,\mathbf{p}} v_1^{p_1}\dots v_n^{p_n} \end{aligned} \tag{7.2.16}$$

that gives a *majorant* for the Cauchy problem (7.2.8), that is, all coefficients of the series (7.2.16) are nonnegative real numbers satisfying inequalities

$$\psi_{i,p} \geqslant |\phi_{i,p}|, \quad B_{ij,\mathbf{p}} \geqslant |A_{ij,\mathbf{p}}|. \tag{7.2.17}$$

Let

$$v_i(x,t) = \sum_{p\geqslant 1,\ q\geqslant 0} v_{i,pq} x^p t^q \tag{7.2.18}$$

be the solution to the Cauchy problem (7.2.15) in the class of formal power series. Like above one has

$$v_{i,pq} = P_{i,pq}\left(\psi_{j,r}, B_{jk,\mathbf{s}}\right) \tag{7.2.19}$$

with the *same* polynomials $P_{i,pq}$ with nonnegative integer coefficients. Hence the inequalities (7.2.17) imply

$$v_{i,pq} \geqslant |u_{i,pq}|. \tag{7.2.20}$$

Our goal is to find a majorant for the Cauchy problem (7.2.8) in such a way that the formal solution (7.2.18) to (7.2.15) converges for sufficiently small $|x|$ and $|t|$. Then the inequalities (7.2.20) will imply convergence of the series (7.2.11) for the same values of x and t.

In order to construct such a majorant let us recall the Cauchy inequalities for the coefficients of convergent power series:

$$\begin{aligned} |\phi_{i,p}| &\leqslant \frac{M}{\rho^p} \\ |A_{ij,\mathbf{p}}| &\leqslant \frac{M}{r^{p_1+\ldots+p_n}} \end{aligned} \tag{7.2.21}$$

for some positive constant M. The radii ρ and r are defined in (7.2.10). We choose

$$\begin{aligned} \psi_{i,p} &= \frac{M}{\rho^p} \\ B_{ij,\mathbf{p}} &= \frac{(p_1+\ldots+p_n)!}{p_1!\ldots p_n!}\frac{M}{r^{p_1+\ldots+p_n}}. \end{aligned} \tag{7.2.22}$$

Observe obvious inequality

$$B_{ij,\mathbf{p}} \geqslant \frac{M}{r^{p_1+\ldots+p_n}},$$

so

$$B_{ij,\mathbf{p}} \geqslant |A_{ij,\mathbf{p}}|.$$

We obtain the initial data for the majorant Cauchy problem

$$\psi_i(x) = M\sum_{p=1}^{\infty}\left(\frac{x}{\rho}\right)^p = \frac{M\,x}{\rho - x}, \quad |x|<\rho \tag{7.2.23}$$

and the coefficient matrix

$$B_{ij}(\mathbf{v}) = M\sum_{p_1,\ldots,p_n\geqslant 0}\frac{(p_1+\ldots+p_n)!}{p_1!\ldots p_n!}\left(\frac{v_1}{r}\right)^{p_1}\cdots\left(\frac{v_n}{r}\right)^{p_n} \tag{7.2.24}$$

We arrive at the following majorant Cauchy problem

$$\begin{aligned}\frac{\partial v_i}{\partial t} &= \frac{M}{1-\frac{v_1+\ldots+v_n}{r}}\sum_{j=1}^{n}\frac{\partial v_i}{\partial x}, \quad i=1,\ldots,n\\ v_i(x,0) &= \frac{M\,x}{\rho - x}.\end{aligned} \tag{7.2.25}$$

Let us look for a solution to the problem (7.2.25) in the form

$$v_i(x,t) = v(x,t), \quad i = 1,\ldots,n.$$

The function $v = v(x,t)$ must satisfy the following scalar Cauchy problem

$$\begin{aligned}v_t &= \frac{M\,n}{1-\frac{n}{r}v}v_x\\ v(x,0) &= \frac{M\,x}{\rho - x}.\end{aligned} \tag{7.2.26}$$

Lemma 7.4 *The solution of the Cauchy problem (7.2.26) is determined from the quadratic equation*

$$(v+M)\left[\left(1-\frac{n}{r}v\right)x + M\,n\,t\right] = \rho\,v\left(1-\frac{n}{r}v\right) \tag{7.2.27}$$

where one has to choose the root of the quadratic equation vanishing at $x=0$, $t=0$.

Proof: Let us apply the implicit function theorem to the equation (7.2.27). Differentiating the quadratic equation in x and t one finds

$$\begin{aligned}v_x &= \frac{(M+v)\left(1-\frac{n}{r}v\right)}{\rho\left(1-\frac{2n}{r}v\right) - Mnt - \left(1-\frac{Mn}{r}-\frac{2nv}{r}\right)x}\\ v_t &= \frac{M\,n\,(M+v)}{\rho\left(1-\frac{2n}{r}v\right) - Mnt - \left(1-\frac{Mn}{r}-\frac{2nv}{r}\right)x}.\end{aligned}$$

Applicability of the implicit function theorem is guaranteed by non-vanishing of the denominator at the point $x=t=0$:

$$\rho\left(1-\frac{2n}{r}v\right) = \rho \neq 0$$

(we have used the condition $v(0,0)=0$). Substituting the above formula into the PDE we obtain

$$v_t = \frac{M\,n}{1-\frac{n}{r}v}v_x.$$

At $t=0$ the quadratic equation simplifies to

$$(v+M)\left(1-\frac{n}{r}v\right)x = \rho\,v\left(1-\frac{n}{r}v\right)$$

that gives the needed solution

$$v = \frac{M\,x}{\rho - x}.$$

It remains to observe that at $x = t = 0$ the quadratic equation (7.2.27) reduces to

$$v\left(1 - \frac{n}{r} v\right) = 0.$$

The latter has two distinct roots

$$v_1 = 0 \quad \text{and} \quad v_2 = \frac{r}{n}.$$

Hence the roots of the quadratic equation remain distinct for sufficiently small $|x|$ and $|t|$. The lemma is proved. ∎

The root we are looking for can be written explicitely

$$v = \frac{1}{2} \frac{\frac{M}{\rho}(x - r\,t) + \frac{r}{n}\left(1 - \frac{x}{\rho}\right)}{1 - \frac{x}{\rho}} - \frac{1}{2} \frac{\sqrt{\left[\frac{M}{\rho}(x + r\,t) - \frac{r}{n}\left(1 - \frac{x}{\rho}\right)\right]^2 - 4M^2 \frac{r\,t}{\rho}}}{1 - \frac{x}{\rho}} \tag{7.2.28}$$

This function is analytic for

$$|x| < \rho \quad \text{and} \quad \left[\frac{M}{\rho}(x + r\,t) - \frac{r}{n}\left(1 - \frac{x}{\rho}\right)\right]^2 - 4M^2 \frac{r\,t}{\rho} > 0.$$

These inequalities hold true for sufficiently small $|x|$ and $|t|$. Hence the above arguments based on the technique of majorants prove convergence of the series for the solution of (7.2.8). ∎

Remark 7.5 *The theorem can be extended to the systems with complex coefficients replacing the real variable x to a complex one z. The assumption of analyticity remains crucial in the proof. Recall that a complex analytic function $f = f(z)$ can be considered as a function of two real variables x, y, where $z = x + iy$, satisfying the* Cauchy - Riemann equation

$$\begin{gathered}\frac{\partial f}{\partial \bar z} = 0 \\ \frac{\partial}{\partial \bar z} = \frac{1}{2}\left(\frac{\partial}{\partial x} + i\,\frac{\partial}{\partial y}\right).\end{gathered} \tag{7.2.29}$$

Remark 7.6 *The analyticity assumption is fundamental for validity of the theorem. Indeed, in 1956 Hans Lewy found the following counterexample. He considered the following equation*

$$u_x + i\,u_y - 2i(x + iy)u_t = g(x, y, t). \tag{7.2.30}$$

This equation has solutions analytic near the origin provided the right hand side is analytic. However Lewy proved existence of $\mathcal{C}^\infty$ functions g such that (7.2.30) has no solutions in any neighborhood of $x = y = t = 0$. Later (1962) S.Mizohata found another counterexample considering the equation

$$u_x + i\,x\,u_y = g(x, y). \tag{7.2.31}$$