

Advanced Mathematical Physics I. Notes

Mikhail Bershtein

14 October 2026

Abstract

Typeset by AI (Claude Opus 5.5) from my handwritten lecture notes, with corrections of misprints and occasional added details. Edited, but not fully proofread by me; please report errors.

Contents

1	Introduction to PDEs	1
2	First-order equations	5
3	Linear differential operators	9
4	Second-order operators in the plane. The wave equation	12
5	The wave equation with boundary conditions	16
6	Periodic problem for the wave equation. Fourier series	19
	References	24

Lecture 1. Introduction to PDEs

References: [Strauss, Ch. 1], [Evans, Ch. 1].

1.1 What is a PDE

PDE stands for partial differential equation(s).

Let u be a function of some variables $(x, y, \dots)$. Derivatives are written as subscripts:

$$u_x = \frac{\partial u}{\partial x}, \quad \text{etc.} \quad (1.1)$$

A PDE of the first order is

$$F(x, y, u, u_x, u_y) = 0, \quad (1.2)$$

a PDE of the second order is

$$F(x, y, u, u_x, u_y, u_{xx}, u_{xy}, u_{yy}) = 0, \quad \text{etc.} \quad (1.3)$$

A solution is a function u that satisfies the equation in some region.

- A system of PDEs: $u = (u^1, \dots, u^m)$ and

$$F^1 = F^2 = \dots = F^n = 0 \quad (\text{usually } n = m). \quad (1.4)$$

- There is not much general theory, rather examples.

1.2 Examples (constants are hidden)

- Transport equation, first order, linear homogeneous:

$$u_t + cu_x = 0. \quad (1.5)$$

- Burgers–Hopf equation, first order:

$$u_t + uu_x = 0. \quad (1.6)$$

- Laplace equation, second order, linear:

$$\Delta u = \sum_i u_{x_i x_i} = 0. \quad (1.7)$$

- Poisson equation, linear:

$$\Delta u = f, \quad (1.8)$$

where f is a density.

- Wave equation, second order, linear:

$$u_{tt} = \Delta u. \quad (1.9)$$

- Heat (diffusion) equation, second order, linear:

$$u_t = \Delta u. \quad (1.10)$$

- KdV equation:

$$u_t = uu_x + u_{xxx}. \quad (1.11)$$

- Maxwell equations:

$$\begin{cases} \operatorname{div} E = \rho, \\ \operatorname{div} B = 0, \\ \nabla \times E = -B_t, \\ \nabla \times B = E_t + J. \end{cases} \quad (1.12)$$

Here $E = (E^1, E^2, E^3)$ is the electric field, $B = (B^1, B^2, B^3)$ is the magnetic field, ρ is the charge density, J is the current. In coordinates,

$$\operatorname{div} E = \sum_i E_{x_i}^i, \quad \nabla \times E = (E_{x_2}^3 - E_{x_3}^2, E_{x_3}^1 - E_{x_1}^3, E_{x_1}^2 - E_{x_2}^1). \quad (1.13)$$

- Euler equations:

$$\begin{cases} u_t + (u \cdot \nabla)u = -\nabla p, \\ \operatorname{div} u = 0. \end{cases} \quad (1.14)$$

Here $u = (u^1, u^2, u^3)$ and p is the pressure. In coordinates,

$$(u^i)_t + \sum_j u^j u_{x_j}^i = -p_{x_i}. \quad (1.15)$$

1.3 Generalization of ODE

A PDE is a generalization of an ODE (we will need some basics, not the tricky parts).

The solution of an ODE of order m depends on m parameters. For a PDE, see the following examples.

Example 1.1. Let $u(x, y)$ satisfy

$$u_{xx} = 0. \quad (1.16)$$

Then u_x is a constant in x , that is, a function of y . Integrating again, we get

$$u = xf(y) + g(y), \quad (1.17)$$

where f, g are arbitrary functions. This is like an ODE, only the constants depend on y .

Example 1.2. Let $u(x, y)$ satisfy

$$u_{xx} + u = 0. \quad (1.18)$$

As for an ODE,

$$u = f(y) \cos x + g(y) \sin x. \quad (1.19)$$

Example 1.3. Let $u(x, y)$ satisfy

$$u_{xy} = 0. \quad (1.20)$$

Then $u_x = f(x)$ and

$$u = F(x) + G(y). \quad (1.21)$$

Moral. The solution of a PDE contains arbitrary functions.

1.4 Physics leads to PDEs

- Transport equation:

$$u_t + cu_x = 0. \quad (1.22)$$

A fluid in a tube; the rate of change is proportional to the gradient.

- Wave equation:

$$u_{tt} = \Delta u. \quad (1.23)$$

A vibrating string, more generally a drumhead.

- Heat (diffusion) equation:

$$u_t = \Delta u. \quad (1.24)$$

- Maxwell equations.

- Schrödinger equation:

$$-iu_t = \Delta u - V(x)u, \quad (1.25)$$

where $V(x)$ is a potential.

- Euler equations: an ideal fluid.

1.5 Initial and boundary conditions

If one of the coordinates is the time t , there are initial conditions

$$u(x, t_0) = \phi(x) \quad (\text{position}), \quad (1.26a)$$

$$u_t(x, t_0) = \psi(x) \quad (\text{velocity}). \quad (1.26b)$$

Their number depends on the order of the equation.

More generally, let D be the domain where the PDE is valid.

(D) Dirichlet conditions:

$$u|_{\partial D} = g. \quad (1.27)$$

(N) Neumann conditions:

$$\left. \frac{\partial u}{\partial n} \right|_{\partial D} = g. \quad (1.28)$$

Here $\frac{\partial u}{\partial n}$ is the normal derivative: for the normal $n = (n^1, n^2, n^3)$,

$$\frac{\partial u}{\partial n} = \sum_i n^i u_{x_i}. \quad (1.29)$$

(R) Robin condition:

$$\frac{\partial u}{\partial n} + au = g. \quad (1.30)$$

In one dimension, $D = (0, \ell)$:

(D) $u(0, t) = g(t)$, $u(\ell, t) = h(t)$;

(N) $u_x(0, t) = g(t)$, $u_x(\ell, t) = h(t)$.

Example 1.4 (String). If the string is fixed at both ends, then (D): $u(0, t) = u(\ell, t) = 0$. A free end means that the vertical component of the tension vanishes, then (N): $u_x(\ell, t) = 0$.

Conditions at infinity are more tricky. For example, for the Schrödinger equation: $\int |u|^2 dx < \infty$.

1.6 Well-posed problems

A problem consists of a PDE, a domain and initial/boundary conditions. It is well posed if the following hold.

(i) Existence.

(ii) Uniqueness.

(iii) Stability: the solution depends continuously on the data given in the problem.

Stability (iii) is important for physical applications, because we know the initial data only with some precision.

Not all problems are well posed.

Example 1.5 (Diffusion). Consider

$$u_t = u_{xx}. \quad (1.31)$$

For the initial data $u(x, 0) = f(x)$ we expect a unique solution. But backwards in time (antidiffusion, more chaotic) existence and stability fail.

Example 1.6 (Laplace). Consider the Laplace equation

$$u_{xx} + u_{yy} = 0, \quad D = \{-\infty < x < \infty, 0 < y < \infty\}. \quad (1.32)$$

It has the solutions

$$u_n = e^{-\sqrt{n}} \sin nx \sinh ny, \quad (1.33)$$

since $(u_n)_{xx} = -n^2 u_n$ and $(u_n)_{yy} = n^2 u_n$. The boundary data as $n \rightarrow \infty$:

$$u_n(x, 0) = 0, \quad (u_n)_y(x, 0) = ne^{-\sqrt{n}} \sin nx \rightarrow 0. \quad (1.34)$$

But for $y > 0$ the functions $u_n(x, y)$ of x do not tend to 0. And for the zero boundary data the solution is 0.

Example 1.7. Consider

$$u_{xx} + u = 0 \quad (1.35)$$

with the Dirichlet conditions (D): $u(0) = u(\ell) = 0$. The general solution of the equation is $u = \alpha \sin x + \beta \cos x$. If $\ell \in \pi\mathbb{Z}$, the solution of the problem is not unique.

1.7 What do we mean by a solution?

Is u analytic? Infinitely differentiable? At least k times continuously differentiable for a PDE of order k : this is a *classical solution*.

In general this is not the case, and we need *generalized* or *weak* solutions.

Example 1.8. The equation

$$u_{xy} = 0 \quad (1.36)$$

has the solution

$$u = F(x) + G(y) + C, \quad (1.37)$$

where

$$F(x) = u(x, 0) - C, \quad G(y) = u(0, y) - C, \quad C = u(0, 0). \quad (1.38)$$

If F, G are not smooth, then u is not smooth.

Moral. Today we discussed what a PDE is, the main examples of PDEs and the initial and boundary conditions imposed on their solutions. We also discussed what it means for a problem to be well posed (existence, uniqueness, stability) and what we mean by a solution (classical or weak).

Lecture 2. First-order equations

References: [Strauss, Sec. 1.2, 14.1], [Arnold, Ch. 1].

2.1 First-order linear equations

Consider the transport equation

$$u_t + cu_x = 0. \quad (2.1)$$

We can change the coordinates:

$$\tilde{t} = t + cx, \quad \tilde{x} = ct - x. \quad (2.2)$$

Then

$$(1 + c^2)\partial_{\tilde{t}}u = 0, \quad u = f(\tilde{x}) = f(ct - x). \quad (2.3)$$

More generally, consider a linear equation with constant coefficients

$$\sum_i a_i u_{x_i} = 0, \quad (2.4)$$

where $v = (a_1, \dots, a_n)$ is a vector field. There exists a change of variables $x_1, \dots, x_n \rightsquigarrow \tilde{x}_1, \dots, \tilde{x}_n$ such that

$$u_{\tilde{x}_1} = 0 \quad \Rightarrow \quad u = u(\tilde{x}_2, \dots, \tilde{x}_n). \quad (2.5)$$

The boundary data are

$$u|_S = f, \quad (2.6)$$

where S is a hypersurface. The necessary and sufficient condition (locally) is that v is transversal to S .

Inhomogeneous equation. Consider

$$\sum_i a_i u_{x_i} = \rho. \quad (2.7)$$

Again we change the variables:

$$u_{\tilde{x}_1} = \rho \quad \Rightarrow \quad u = \int_{\tilde{x}_1^{(0)}}^{\tilde{x}_1} \rho d\tilde{x}_1 + f(\tilde{x}_2, \dots, \tilde{x}_n). \quad (2.8)$$

Geometric interpretation. For the equation (2.1) consider the vector field (a tangent vector at every point) $v = (c, 1)$. The equation reads $u_v = 0$, that is, u is constant on the lines $ct - x = k$. So u is a function of k .

2.2 Linear equations with non-constant coefficients

Example 2.1. Consider equation

$$u_t + xu_x = 0. \quad (2.9)$$

Geometrically, we have the vector field $v = (x, 1)$ and $L_v u = 0$. The integral curves of v (curves $\gamma(s)$ with the tangent vector v at every point) are called *characteristic curves*. For $\gamma(s) = (x(s), t(s))$ we get

$$\frac{dx}{ds} = x, \quad \frac{dt}{ds} = 1. \quad (2.10)$$

Without loss of generality $t = s$. Then

$$\frac{dx}{dt} = x \quad \Rightarrow \quad x = Ce^t, \quad xe^{-t} = C. \quad (2.11)$$

The function u is constant on the characteristic curves:

$$\frac{d}{dt} u(Ce^t, t) = u_t + Ce^t u_x = 0. \quad (2.12)$$

The general solution is

$$u = f(xe^{-t}). \quad (2.13)$$

For example, consider the initial condition $u(x, 0) = x^2$. Then

$$f(xe^{-0}) = x^2 \quad \Rightarrow \quad u(x, t) = (xe^{-t})^2. \quad (2.14)$$

Example 2.2. Consider equation

$$u_t + 2tx^2u_x = 0. \quad (2.15)$$

Again we have the vector field $v = (2tx^2, 1)$ and $L_v u = 0$. Then

$$\frac{dx}{dt} = 2tx^2, \quad \frac{1}{x^2}dx = 2t dt, \quad -\frac{1}{x} = t^2 - c. \quad (2.16)$$

The characteristic curves are

$$x = \frac{1}{c - t^2}, \quad (2.17)$$

and the general solution has the form (for $x \neq 0$)

$$u = f\left(t^2 + \frac{1}{x}\right). \quad (2.18)$$

Example 2.3. Consider

$$u_t + u_x + u = e^{x+2t}, \quad u(x, 0) = 0. \quad (2.19)$$

Here $v = (1, 1)$, and the integral (characteristic) curves are $x = t + x_0$. On such a curve

$$\frac{du}{dt} + u = e^{3t+x_0}, \quad (2.20)$$

hence

$$u = \frac{e^{3t+x_0} - e^{x_0-t}}{4} \Rightarrow u = \frac{e^{x+2t} - e^{x-2t}}{4}. \quad (2.21)$$

Lesson. The equation

$$L_v u = F(u, x) \quad (2.22)$$

reduces to an ODE on the characteristic curves. For example,

$$L_v u + cu = 0 \rightsquigarrow u_{\tilde{x}_1} + cu = 0, \quad u = e^{-c\tilde{x}_1} f(\tilde{x}_2, \dots, \tilde{x}_n). \quad (2.23)$$

2.3 Quasi-linear equations

Consider

$$\sum_i A_i(u, x) u_{x_i} = F, \quad (2.24)$$

where A_i do not contain derivatives.

Geometrically, consider the $(n+1)$ -dimensional space $(x_1, \dots, x_n, y)$. The function u defines a surface Γ of codimension one, the graph of u . Let v be the vector field $(A_1, \dots, A_n, F)$.

Proposition 2.4. *The function u is a solution of the equation if and only if v is tangent to Γ .*

Proof. The surface Γ is given by $u(x) - y = 0$. Its tangent space is orthogonal to $(u_{x_1}, \dots, u_{x_n}, -1)$. The equation reads

$$(A_1, \dots, A_n, F) \cdot (u_{x_1}, \dots, u_{x_n}, -1) = 0. \quad (2.25)$$

□

Remark 2.5. *If A does not depend on u , then the projection does not lose information. If $F = 0$, then the vector field is horizontal.*

Remark 2.6. *Now the vector field depends on the layer.*

Example 2.7 (Hopf–Burgers equation). Consider

$$u_t + uu_x = 0, \quad v = (y, 1, 0) \quad (2.26)$$

in the coordinates (x, t, y) . A characteristic curve (an integral curve of v) satisfies

$$\dot{t} = 1, \quad \dot{x} = u, \quad \dot{y} = 0. \quad (2.27)$$

We take t as the parameter:

$$\frac{d}{dt}(u(x(t), t)) = u_t + u_x \frac{dx}{dt} = u_t + uu_x = 0. \quad (2.28)$$

On a characteristic curve u is constant. Now we know the characteristic curves: they are lines with slope u .

Remark 2.8. *Free particles*

$$\frac{d^2x}{dt^2} = 0 \quad (2.29)$$

are described in terms of the velocity field.

Consider the initial data

$$u(x, 0) = \phi(x). \quad (2.30)$$

For every x_0 we draw the line with slope $\phi(x_0)$. If the lines intersect, we are in trouble: the wave breaks (*gradient catastrophe*).

- In general, we want the map $x = x_0 + t\phi(x_0)$ to be injective. A sufficient condition is positivity of its derivative $1 + t\phi'(x_0)$; it holds for all $t > 0$ if ϕ is increasing. If $\inf \phi' < 0$, the derivative first vanishes at $t_* = -\frac{1}{\inf \phi'}$.
- We can allow weak solutions (not continuous).
- We can say that the solution exists near the initial line and then breaks. Hence the real physical process can be described by the equation only approximately.

In formulas,

$$\begin{cases} \frac{x - x_0}{t - 0} = \phi(x_0), \\ u(x, t) = \phi(x_0). \end{cases} \quad (2.31)$$

Example 2.9. Let $\phi(x) = x^2$. Then

$$x - x_0 = tx_0^2 \quad \Rightarrow \quad x_0 = \frac{-1 \pm \sqrt{1 + 4tx}}{2t}. \quad (2.32)$$

As $t \rightarrow 0$ we want $x_0 \rightarrow x$, so we take the upper sign:

$$x_0 = \frac{-1 + \sqrt{1 + 4tx}}{2t}, \quad t \neq 0. \quad (2.33)$$

Then

$$u = \left(\frac{-1 + \sqrt{1 + 4tx}}{2t} \right)^2 = \frac{1 + 2tx - \sqrt{1 + 4tx}}{2t^2}. \quad (2.34)$$

This is a solution in the region $1 + 4tx > 0$. Here $\inf \phi' = -\infty$, so there is no solution in any strip $0 \leq t < T$.

In a bit more general situation, consider

$$u_t + A(u)u_x = 0. \quad (2.35)$$

The characteristic curves are the integral curves of $v = (A(u), 1, 0)$. On a characteristic curve u is constant, hence A is constant, hence the curve is a line. We can solve the equation provided the characteristic lines do not intersect.

Moral. Today we solved first-order linear equations, with constant and with non-constant coefficients, by reducing them to ODEs on the characteristic curves. For quasi-linear equations the characteristic curves live in the space of the graph of the solution; for the Hopf–Burgers equation they are lines, and when they intersect the wave breaks.

Lecture 3. Linear differential operators

References: [Dubrovin, Ch. 2].

3.1 Linear differential operators

Consider $\mathbb{R}^d$ with coordinates $x_1, \dots, x_d$, and let $\Omega \subset \mathbb{R}^d$ be an open subset. Let

$$\partial_k = \frac{\partial}{\partial x_k}, \quad D_k = -i\partial_k, \quad k = 1, \dots, d, \quad i = \sqrt{-1}. \quad (3.1)$$

This is the notation of Dubrovin; the operator D_k is self-adjoint on $L^2(\mathbb{R}^d)$.

For a multi-index $\mathbf{p} = (p_1, \dots, p_d)$ we write

$$|\mathbf{p}| = p_1 + \dots + p_d, \quad \mathbf{p}! = p_1! \dots p_d!, \quad (3.2)$$

and

$$x^{\mathbf{p}} = x_1^{p_1} \dots x_d^{p_d}, \quad \partial^{\mathbf{p}} = \partial_1^{p_1} \dots \partial_d^{p_d}, \quad D^{\mathbf{p}} = D_1^{p_1} \dots D_d^{p_d}. \quad (3.3)$$

Thus

$$D^{\mathbf{p}}: C^\infty(\Omega) \rightarrow C^\infty(\Omega), \quad f \mapsto D^{\mathbf{p}}f = (-i)^{|\mathbf{p}|} \frac{\partial^{|\mathbf{p}|} f}{\partial x_1^{p_1} \dots \partial x_d^{p_d}}. \quad (3.4)$$

Definition 3.1. A *linear differential operator* is

$$A = \sum_{\mathbf{p}} a_{\mathbf{p}}(x) D^{\mathbf{p}}, \quad a_{\mathbf{p}}(x) \in C^\infty(\Omega), \quad A: C^\infty(\Omega) \rightarrow C^\infty(\Omega). \quad (3.5)$$

Its *order* is

$$\text{ord } A = \max_{\mathbf{p}: a_{\mathbf{p}} \neq 0} |\mathbf{p}|. \quad (3.6)$$

Example 3.2. 1. $A = \sum_i a_i \partial_{x_i}$ is a linear operator of order one.

2. The Laplace operator

$$\Delta = \sum_{i=1}^d \partial_i^2 = - \sum_{i=1}^d D_i^2. \quad (3.7)$$

3. The wave operator

$$\frac{\partial^2}{\partial t^2} - \Delta. \quad (3.8)$$

4. The diffusion (heat) operator

$$\partial_t - \Delta. \quad (3.9)$$

3.2 Symbol

Definition 3.3. The *symbol* of $A = \sum_{|\mathbf{p}| \leq m} a_{\mathbf{p}}(x) D^{\mathbf{p}}$ is

$$a(x, \xi) = \sum_{|\mathbf{p}| \leq m} a_{\mathbf{p}}(x) \xi^{\mathbf{p}}. \quad (3.10)$$

The *principal symbol* is

$$a_m(x, \xi) = \sum_{|\mathbf{p}|=m} a_{\mathbf{p}}(x) \xi^{\mathbf{p}}, \quad m = \text{ord } A. \quad (3.11)$$

The symbol is a function on $\Omega \times \mathbb{R}^d$; more geometrically, on the cotangent bundle $T^*\Omega$. Here x is a coordinate and ξ is a *covector* (dual to a vector). Both a and a_m are polynomials in ξ .

Example 3.4. The symbols of the operators of Example 3.2 are as follows; for the operators with t , the variable dual to t is τ .

1. $i \sum_i a_i \xi_i$.
2. $-\sum_i \xi_i^2$.
3. $-\tau^2 + \sum_i \xi_i^2$.
4. $i\tau + \sum_i \xi_i^2$.

Lemma 3.5. Let A have constant coefficients and $u = e^{ix \cdot \xi}$, where $x \cdot \xi = \sum_i x_i \xi_i$. Then

$$Au = 0 \quad \Leftrightarrow \quad a(\xi) = 0. \quad (3.12)$$

Proof. We have $D_i u = \xi_i u$, hence $D^{\mathbf{p}} u = \xi^{\mathbf{p}} u$ and $Au = a(\xi)u$. $\square$

Lemma 3.6. Let A and B be linear differential operators of orders k and l with principal symbols $a_k(x, \xi)$ and $b_l(x, \xi)$, respectively. Then $C = A \circ B$ has order $\leq k + l$; in the case of equality its principal symbol is $a_k \cdot b_l$.

Usually the equality holds; it fails only because of zero-divisor effects, such as non-intersecting supports of a_k and b_l .

Proof. We have

$$a_{\mathbf{p}} D^{\mathbf{p}} b_{\mathbf{q}} D^{\mathbf{q}} = a_{\mathbf{p}} b_{\mathbf{q}} D^{\mathbf{p}+\mathbf{q}} + \text{lower terms}. \quad (3.13)$$

$\square$

3.3 Change of variables

Consider a change of variables $y_i = y_i(x)$, $i = 1, \dots, d$. Then

$$\frac{\partial}{\partial x_j} = \sum_k \frac{\partial y_k}{\partial x_j} \frac{\partial}{\partial y_k}, \quad (3.14)$$

$$\frac{\partial^2}{\partial x_i \partial x_j} = \sum_{k,l} \frac{\partial y_k}{\partial x_i} \frac{\partial y_l}{\partial x_j} \frac{\partial^2}{\partial y_k \partial y_l} + \sum_k \frac{\partial^2 y_k}{\partial x_i \partial x_j} \frac{\partial}{\partial y_k}, \quad (3.15)$$

and so on. We get an operator of the same order with lower-order corrections:

$$A \rightsquigarrow \tilde{A} = \sum_{\mathbf{p}} \tilde{a}_{\mathbf{p}}(y) \frac{\partial^{|\mathbf{p}|}}{\partial y_1^{p_1} \dots \partial y_d^{p_d}}. \quad (3.16)$$

Proposition 3.7. *Let a_m be the principal symbol of A and $\tilde{a}_m$ the principal symbol of $\tilde{A}$. Then*

$$\tilde{a}_m(y, \tilde{\xi}) = a_m(x, \xi) \quad \text{if} \quad \xi_j = \sum_k \frac{\partial y_k}{\partial x_j} \tilde{\xi}_k. \quad (3.17)$$

Proof. For a first-order operator, the linear formula (3.14) is exactly the transformation from ξ to $\tilde{\xi}$. For D^p of higher order we use multiplicativity, and for an arbitrary operator additivity. $\square$

Example 3.8. Let

$$A = \frac{\partial^2}{\partial x_k \partial x_l}, \quad a_2 = -\xi_k \xi_l. \quad (3.18)$$

Then

$$\tilde{A} = \sum_{i,j} \frac{\partial y_i}{\partial x_k} \frac{\partial y_j}{\partial x_l} \frac{\partial^2}{\partial y_i \partial y_j} + \text{lower terms}, \quad \tilde{a}_2 = - \sum_{i,j} \frac{\partial y_i}{\partial x_k} \frac{\partial y_j}{\partial x_l} \tilde{\xi}_i \tilde{\xi}_j. \quad (3.19)$$

3.4 Operators with constant coefficients

First order. For a first-order operator

$$A = \sum_k a_k \frac{\partial}{\partial x_k} + b \quad (3.20)$$

a change of coordinates gives

$$\tilde{A} = \frac{\partial}{\partial y_1} + \tilde{b}. \quad (3.21)$$

Second order. Consider a second-order operator

$$A = \sum_{k,l} a_{kl} \frac{\partial^2}{\partial x_k \partial x_l} + \sum_k b_k \frac{\partial}{\partial x_k} + c. \quad (3.22)$$

Without loss of generality $a_{kl} = a_{lk}$. A linear change of variables transforms the coefficients as

$$a_{kl} \rightsquigarrow \tilde{a}_{ij} = \sum_{k,l} \frac{\partial y_i}{\partial x_k} \frac{\partial y_j}{\partial x_l} a_{kl}. \quad (3.23)$$

In matrix form,

$$\tilde{a} = C a C^T, \quad C = \left(\frac{\partial y_i}{\partial x_k} \right). \quad (3.24)$$

By linear algebra, a can be transformed to

$$\text{diag}(\underbrace{1, \dots, 1}_p, \underbrace{-1, \dots, -1}_q, 0, \dots, 0), \quad (3.25)$$

and p, q do not depend on the choice of the reducing transformation.

Corollary 3.9. *There exists a change of variables such that*

$$\tilde{A} = \sum_{i=1}^p \frac{\partial^2}{\partial y_i^2} - \sum_{i=p+1}^{p+q} \frac{\partial^2}{\partial y_i^2} + \sum_{k=1}^d \tilde{b}_k \frac{\partial}{\partial y_k} + c, \quad (3.26)$$

and p, q do not depend on the change of variables.

3.5 Characteristic covectors

Definition 3.10. A covector $\xi \in \mathbb{R}^d$ is called *characteristic* at a point x if

$$a_m(x, \xi) = 0. \quad (3.27)$$

Remark 3.11. (a) For constant coefficients, $u = e^{ix \cdot \xi}$ is a solution if $a(\xi) = 0$.

(b) We use a_m , not a , since a_m is invariant under changes of variables.

(c) The principal symbol a_m is homogeneous, so if ξ is a solution, then $\lambda\xi$ is a solution. The characteristic covectors form a union of lines, that is, a cone; 0 is characteristic.

Definition 3.12. A differential operator A is *elliptic* if

$$a_m(x, \xi) \neq 0 \quad \forall \xi \neq 0, \quad x \in \Omega. \quad (3.28)$$

Example 3.13. 1. The Laplace operator $\Delta = \sum_i \partial_{x_i}^2$.

2. The operator $\partial_x^2 + x \partial_y^2$ is elliptic on the half-plane $x > 0$.

Remark 3.14. In the case of constant coefficients (second order) this means $p = d$, $q = 0$ or $q = d$, $p = 0$: the form is positive definite or negative definite. Its level sets are ellipses.

Definition 3.15. An operator A is *hyperbolic with respect to the time t* if for all $\xi \neq 0$ and all $(t, x) \in \Omega$ the equation

$$a_m(t, x, \tau, \xi) = 0 \quad (3.29)$$

has m pairwise distinct real roots $\tau_1, \dots, \tau_m$.

Usually one omits the choice of time and just says *hyperbolic*.

Example 3.16. For the wave operator

$$\frac{\partial^2}{\partial t^2} - \sum_i \frac{\partial^2}{\partial x_i^2} \quad (3.30)$$

the characteristic covectors form the cone $\tau^2 = \sum_i \xi_i^2$. The level sets of the principal symbol (in two dimensions) are hyperbolas.

Example 3.17. For the linear differential operator $\sum_i a_i \partial_{x_i}$ the characteristic covectors are given by

$$\sum_i a_i \xi_i = 0, \quad (3.31)$$

a hyperplane of covectors normal to the vector field $v = \sum_i a_i \partial_{x_i}$. In two dimensions these are the normals to the characteristic curves of the last lecture.

Moral. Today we discussed linear differential operators, their symbols and principal symbols, and how the principal symbol transforms under a change of variables. For constant coefficients we reduced operators of the first and second order to a canonical form, and defined characteristic covectors, elliptic and hyperbolic operators.

Lecture 4. Second-order operators in the plane. The wave equation

References: [Dubrovin, Sec. 2.6, 2.7, 3.2, 3.3], [Strauss, Sec. 2.1, 2.2].

4.1 Second-order operators with arbitrary coefficients

Consider the operator

$$A = a(x, y)\partial_x^2 + 2b(x, y)\partial_x\partial_y + c(x, y)\partial_y^2. \quad (4.1)$$

Recall that for constant coefficients we can bring A to one of the forms

$$\partial_x^2 + \partial_y^2, \quad -\partial_x^2 - \partial_y^2 \quad (\text{elliptic}), \quad (4.2)$$

$$\partial_x^2 - \partial_y^2 \quad (\text{hyperbolic}), \quad (4.3)$$

$$\partial_y^2 + \text{lower terms} \quad (\text{parabolic}). \quad (4.4)$$

The three cases are distinguished by the determinant

$$\text{Det} = \det \begin{pmatrix} a & b \\ b & c \end{pmatrix} = ac - b^2, \quad (4.5)$$

or by the number of lines of vectors of zero length, that is, of characteristic covectors.

Case 1: hyperbolic, $\text{Det} < 0$. The characteristic equation

$$a(x, y)\xi_x^2 + 2b(x, y)\xi_x\xi_y + c(x, y)\xi_y^2 = 0 \quad (4.6)$$

has two lines of solutions (ξ_x^1, ξ_y^1) and (ξ_x^2, ξ_y^2) . They give two families of characteristic curves, given by the equations

$$\frac{dy}{dx} = -\frac{\xi_x^i}{\xi_y^i}, \quad i = 1, 2. \quad (4.7)$$

Locally the solutions $y = y^i(x)$ are given by the equations

$$u(x, y) = C_1, \quad v(x, y) = C_2, \quad (4.8)$$

where du is proportional to (ξ_x^1, ξ_y^1) and dv to (ξ_x^2, ξ_y^2) . Here we use the straightening theorem for vector fields from the theory of ODEs.

Lemma 4.1. *The functions u, v are local coordinates.*

Proof. We need to check that

$$\det \begin{pmatrix} u_x & u_y \\ v_x & v_y \end{pmatrix} \neq 0. \quad (4.9)$$

The rows (u_x, u_y) and (v_x, v_y) are proportional to the covectors ξ^1 and ξ^2 , which are not proportional. Hence the determinant does not vanish. $\square$

In the coordinates u, v

$$\tilde{A} = \tilde{a}\partial_u^2 + 2\tilde{b}\partial_u\partial_v + \tilde{c}\partial_v^2 + \dots, \quad (4.10)$$

where the dots denote terms of lower order. The characteristic covectors are $du \sim (1, 0)$ and $dv \sim (0, 1)$. Hence $(1, 0)$ and $(0, 1)$ are roots of

$$\tilde{a}\tilde{\xi}_u^2 + 2\tilde{b}\tilde{\xi}_u\tilde{\xi}_v + \tilde{c}\tilde{\xi}_v^2 = 0, \quad (4.11)$$

so $\tilde{a} = \tilde{c} = 0$ and $\tilde{A} = 2\tilde{b}\partial_u\partial_v + \dots$

Theorem 4.2. *If $\text{Det} < 0$, then in some coordinates*

$$\tilde{A} = 2\tilde{b}(u, v)\frac{\partial^2}{\partial u\partial v} + \dots \quad (4.12)$$

Case 2: elliptic, $\text{Det} > 0$. Now equation (4.6) has no nonzero real solutions: there are no real characteristic covectors. Assume that a, b, c are analytic. Then we have complex conjugate characteristics ξ^1, ξ^2 such that $\xi^2 = \overline{\xi^1}$. Similarly to the hyperbolic case there are functions

$$S(x, y) = u(x, y) + iv(x, y), \quad \bar{S}(x, y) = u(x, y) - iv(x, y), \quad (4.13)$$

such that $dS \sim \xi^1, d\bar{S} \sim \xi^2$.

Lemma 4.3. *The functions u, v are local coordinates.*

Proof. The differentials dS and $d\bar{S}$ are proportional to the covectors ξ^1 and ξ^2 , which are not proportional over $\mathbb{C}$. Hence du and dv are linearly independent. $\square$

In these coordinates $\tilde{A}$ has the form (4.10), and the characteristic covectors are

$$dS = (1, i), \quad d\bar{S} = (1, -i). \quad (4.14)$$

Hence $\tilde{b} = 0$ and $\tilde{a} = \tilde{c}$.

Theorem 4.4. *If $\text{Det} > 0$, then in some coordinates*

$$\tilde{A} = \tilde{a} (\partial_u^2 + \partial_v^2) + \dots \quad (4.15)$$

Case 3: parabolic, $\text{Det} = 0$. Assume that $(a, b, c) \neq (0, 0, 0)$; otherwise all covectors are characteristic. There is one line of characteristic covectors (ξ_x, ξ_y) . Let u be the corresponding integral, and let v be an arbitrary function such that u, v are local coordinates. Then for $\tilde{A}$ the only characteristic covector is $(1, 0)$, with multiplicity two. Hence $\tilde{a} = \tilde{b} = 0$.

Theorem 4.5. *If $\text{Det} = 0$ and $(a, b, c) \neq (0, 0, 0)$, then in some coordinates*

$$\tilde{A} = \tilde{c} \partial_v^2 + \dots \quad (4.16)$$

Remark 4.6. *It is important that $d = 2$. For example, in the elliptic case we used two functions u, v to impose two conditions $\tilde{b} = 0, \tilde{a} = \tilde{c}$. For d variables we have d unknown functions and need $d - 1 + \frac{d(d-1)}{2} = \frac{(d+2)(d-1)}{2}$ equations. For example, for $d = 3$ we have five equations on three functions.*

Remark 4.7. *We worked only with the principal symbol. So this is actually a classical problem of geometry, namely we were looking for the canonical form of metric on a surface. For example, in the elliptic case we have shown that a positive metric is conformally equivalent to $dx^2 + dy^2$.*

4.2 Hyperbolic case with constant coefficients

In some coordinates the operator is

$$\partial_t^2 - c^2 \partial_x^2, \quad c > 0. \quad (4.17)$$

The constant c can be excluded, but we keep it, since it has a physical meaning. Physically this is the vibrating string. The wave equation is

$$(\partial_t - c \partial_x)(\partial_t + c \partial_x)u = 0. \quad (4.18)$$

The change of variables

$$\tilde{x} = x + ct, \quad \tilde{y} = x - ct \quad (4.19)$$

gives

$$\partial_{\tilde{x}} \partial_{\tilde{y}} u = 0, \quad (4.20)$$

discussed in Example 1.3. The general solution is

$$u = f(x + ct) + g(x - ct). \quad (4.21)$$

Geometry. The characteristic covectors are the solutions of

$$\xi_t^2 - c^2 \xi_x^2 = 0, \quad (\xi_t, \xi_x) \sim (c, 1), (c, -1). \quad (4.22)$$

There are two families of characteristic lines

$$x \pm ct = \text{const.} \quad (4.23)$$

The function $f(x + ct)$ travels to the left with speed c , and $g(x - ct)$ travels to the right with speed c .

Initial value problem. Consider, for $x \in \mathbb{R}$,

$$u_{tt} = c^2 u_{xx}, \quad u(x, 0) = \phi(x), \quad u_t(x, 0) = \psi(x). \quad (4.24)$$

We use $u(x, t) = f(x + ct) + g(x - ct)$. Then

$$\begin{cases} f(x) + g(x) = \phi(x), \\ cf'(x) - cg'(x) = \psi(x), \end{cases} \quad \Rightarrow \quad \begin{cases} f' + g' = \phi', \\ f' - g' = \frac{1}{c}\psi. \end{cases} \quad (4.25)$$

Hence

$$f' = \frac{1}{2} \left(\phi' + \frac{1}{c}\psi \right), \quad g' = \frac{1}{2} \left(\phi' - \frac{1}{c}\psi \right), \quad (4.26)$$

and

$$f(x) = \frac{1}{2}\phi(x) + \frac{1}{2c} \int_0^x \psi(s) ds + A, \quad g(x) = \frac{1}{2}\phi(x) - \frac{1}{2c} \int_0^x \psi(s) ds + B, \quad A + B = 0. \quad (4.27)$$

Theorem 4.8 (d'Alembert formula).

$$u(x, t) = f(x + ct) + g(x - ct) = \frac{1}{2} \left(\phi(x + ct) + \phi(x - ct) + \frac{1}{c} \int_{x-ct}^{x+ct} \psi(s) ds \right). \quad (4.28)$$

Example 4.9. For $\phi(x) = u_0$, $\psi(x) = v_0$ we get

$$u(x, t) = u_0 + v_0 t, \quad (4.29)$$

a free string moving with speed v_0 .

Remark 4.10. Let $\tilde{u}$ be the solution with the initial data $\tilde{\phi}$, $\tilde{\psi}$. For every $\varepsilon > 0$ and $T > 0$ there is $\delta = \delta(\varepsilon, T)$ such that

$$\sup_{x \in \mathbb{R}, t \in [0, T]} |\tilde{u}(x, t) - u(x, t)| < \varepsilon \quad \text{if} \quad \sup |\tilde{\phi} - \phi| < \delta, \quad \sup |\tilde{\psi} - \psi| < \delta. \quad (4.30)$$

Hence the problem is well posed.

Example 4.11. The plucked string: see the visualization.

Remark 4.12. If ϕ , ψ have singularities, then they propagate for u along the characteristic lines $x \pm ct = \text{const.}$

Remark 4.13 (Causality). No part of the wave goes faster than the speed c . The domain of influence: the data on $[x_1, x_2]$ at $t = 0$ influence the solution at $t = t_0$ only on $[x_1 - ct_0, x_2 + ct_0]$. Dually, the domain of dependence: the solution on $[x_1, x_2]$ at $t = t_0$ depends only on the data on $[x_1 - ct_0, x_2 + ct_0]$ at $t = 0$.

Moral. Today we reduced second-order operators in the plane to a canonical form in the hyperbolic, elliptic and parabolic cases, and saw that this works only in two variables. For the wave equation we found the general solution, solved the initial value problem by d'Alembert's formula, and discussed well-posedness, propagation of singularities and causality.

Lecture 5. The wave equation with boundary conditions

References: [Strauss, Sec. 2.2, 3.2, 3.4], [Dubrovin, Sec. 3.4].

5.1 Energy

Assume that u decays sufficiently fast as $|x| \rightarrow \infty$, and let

$$\text{KE} = \frac{1}{2} \int u_t^2 dx. \quad (5.1)$$

Then

$$\frac{d\text{KE}}{dt} = \int u_t u_{tt} dx = c^2 \int u_t u_{xx} dx = -c^2 \int u_{tx} u_x dx = -\frac{c^2}{2} \frac{d}{dt} \int u_x^2 dx. \quad (5.2)$$

Hence the *energy*

$$E = \int_{\mathbb{R}} (u_t^2 + c^2 u_x^2) dx \quad (5.3)$$

is a conserved quantity.

5.2 Half-line: fixed end

Usually a string has a boundary. Consider the wave equation on the half-line:

$$u_{tt} = c^2 u_{xx}, \quad u(x, 0) = \phi(x), \quad u_t(x, 0) = \psi(x), \quad x > 0, \quad (5.4)$$

with the boundary condition

$$u|_{x=0} = 0. \quad (5.5)$$

This is a fixed end, or the *Dirichlet condition*.

There are two regions (Figure 5.1).

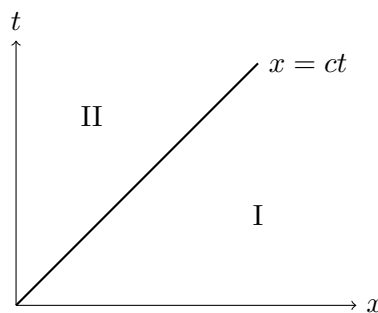


Figure 5.1: Half-line: the regions I ($x \geq ct$) and II ($x \leq ct$).

In region I, $x \geq ct$, d'Alembert's formula (4.28) from the previous lecture works. In region II, $x \leq ct$, we write

$$u(x, t) = f(x + ct) + g(x - ct). \quad (5.6)$$

For $x \geq 0$, as on the whole line,

$$f(x) = \frac{1}{2} \left(\phi(x) + \frac{1}{c} \int_0^x \psi(s) ds \right). \quad (5.7)$$

The boundary condition gives $f(x) + g(-x) = 0$, that is, $g(-x) = -f(x)$ for $x \geq 0$. Hence, for $0 \leq x \leq ct$,

$$u(x, t) = \frac{1}{2} \left(\phi(x + ct) - \phi(ct - x) + \frac{1}{c} \int_{ct-x}^{ct+x} \psi(s) ds \right). \quad (5.8)$$

Remark 5.1. *Reflection of the characteristic lines: in region II one of the two characteristic lines through (x, t) hits the boundary $x = 0$ and is reflected (Figure 5.2; see also the visualization).*

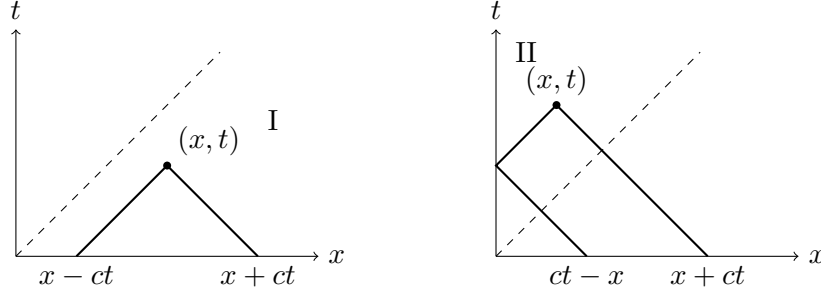


Figure 5.2: Characteristic lines through (x, t) . Left: region I, as on the whole line. Right: region II, one line is reflected at $x = 0$.

More elegant way. The equation has a symmetry: if $u(x, t)$ is a solution, then

$$-u(x, t), \quad u(-x, t), \quad -u(-x, t) \quad (5.9)$$

are solutions. Take the odd extensions of ϕ, ψ :

$$\phi_{\text{ext}}(x) = \begin{cases} \phi(x), & x \geq 0, \\ -\phi(-x), & x < 0, \end{cases} \quad \psi_{\text{ext}}(x) = \begin{cases} \psi(x), & x \geq 0, \\ -\psi(-x), & x < 0, \end{cases} \quad (5.10)$$

and let u_{ext} be the solution on the whole line with the initial data $\phi_{\text{ext}}, \psi_{\text{ext}}$.

Proposition 5.2. u_{ext} is odd for all t , and $u_{\text{ext}} = u$ for $x \geq 0$.

First proof. Use d'Alembert's formula and the oddness of $\phi_{\text{ext}}, \psi_{\text{ext}}$. □

Second proof. The function $\tilde{u}(x, t) = -u_{\text{ext}}(-x, t)$ satisfies the same equation with the same initial data. Hence $\tilde{u} = u_{\text{ext}}$. □

Example 5.3. Let ϕ be a pulse on the half-line and $\psi = 0$. The pulse splits into two pulses moving in opposite directions. In the odd extension the left pulse meets the inverted pulse coming from $x < 0$, and after the reflection it moves to the right upside down. See the visualization.

Remark 5.4. To have a C^2 solution we need $\phi \in C^2$, $\psi \in C^1$, $\phi(0) = 0$, $\phi''(0) = 0$, $\psi(0) = 0$. Otherwise there is a singularity on the line $x = ct$, the boundary between the regions I and II.

5.3 More general Dirichlet condition

Consider the boundary condition

$$u(x, t)|_{x=0} = h(ct). \quad (5.11)$$

The problem is linear, hence

$$u = u_{\phi, \psi} + u_h, \quad (5.12)$$

where $u_{\phi, \psi}$ is the solution with the initial data ϕ, ψ and $h = 0$, and u_h is the solution with zero initial data and the boundary value h . It remains to find u_h . In region I, $u_h = 0$. In region II, $u_h = f(x + ct) + g(x - ct)$ with $f = 0$ and $g(-ct) = h(ct)$. Hence

$$u_h(x, t) = h(ct - x), \quad 0 \leq x \leq ct. \quad (5.13)$$

5.4 Free end

Consider the boundary condition

$$u_x|_{x=0} = 0. \quad (5.14)$$

This is a free end, or the *Neumann condition*. Take the even extensions

$$\phi_{\text{ext}}(x) = \phi(|x|), \quad \psi_{\text{ext}}(x) = \psi(|x|). \quad (5.15)$$

Proposition 5.5. u_{ext} is even for all t and coincides with u for $x \geq 0$; in particular, $u_x(0, t) = 0$.

Example 5.6. For a pulse on the half-line, the left pulse is reflected at $x = 0$ and moves to the right with the same sign. See the visualization.

5.5 Segment

Consider the string $0 < x < l$ with the boundary conditions

$$u|_{x=0} = u|_{x=l} = 0. \quad (5.16)$$

Now there are two reflections:

$$\phi_{\text{ext}}(-x) = -\phi_{\text{ext}}(x), \quad \phi_{\text{ext}}(2l - x) = -\phi_{\text{ext}}(x). \quad (5.17)$$

Take

$$\phi_{\text{ext}}(x) = \begin{cases} \phi(x), & 0 < x < l, \\ -\phi(-x), & -l < x < 0, \end{cases} \quad (5.18)$$

extended to the whole line with period $2l$, and similarly ψ_{ext} . Then

$$u_{\text{ext}}(x, t) = \frac{1}{2} \left(\phi_{\text{ext}}(x + ct) + \phi_{\text{ext}}(x - ct) + \frac{1}{c} \int_{x-ct}^{x+ct} \psi_{\text{ext}}(s) ds \right). \quad (5.19)$$

Proposition 5.7. u_{ext} is odd with respect to both reflections, $u_{\text{ext}}|_{x=0} = u_{\text{ext}}|_{x=l} = 0$, and $u_{\text{ext}} = u$ for $0 \leq x \leq l$.

We have many regions. A region is determined by the number of reflections crossed by the characteristic lines. The regions are diamonds (triangles at the ends of the segment), and in each region there is a new formula (Figure 5.3; see also the visualization).

For example, in the case

$$3l \leq x + ct \leq 4l, \quad -2l \leq x - ct \leq -l, \quad (5.20)$$

we get

$$u(x, t) = \frac{1}{2} \left(-\phi(4l - x - ct) + \phi(2l + x - ct) + \frac{1}{c} \int_{x-ct+2l}^{4l-x-ct} \psi(s) ds \right). \quad (5.21)$$

More general boundary condition. Consider

$$u|_{x=0} = h(ct), \quad u|_{x=l} = 0. \quad (5.22)$$

As before, $u = u_{\phi, \psi} + u_h$. Now the regions are triangles bounded by the characteristic line from the corner $(0, 0)$, reflected at both ends (Figure 5.4). Going up from $t = 0$, in the successive regions

$$\begin{aligned} u_h &= 0, & u_h &= h(ct - x), & u_h &= h(ct - x) - h(ct + x - 2l), \\ u_h &= h(ct - x) - h(ct + x - 2l) + h(ct - x - 2l), \end{aligned} \quad (5.23)$$

and so on: each reflection adds a term. The second term is needed for $u_h|_{x=l} = 0$; the third one because at $x = 0$ the first two give $h(ct) - h(ct - 2l)$.

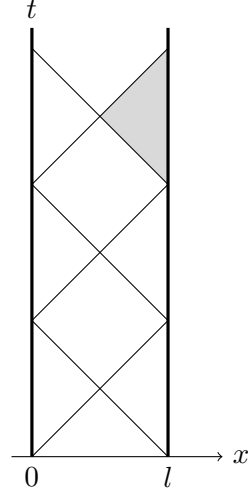


Figure 5.3: Segment: the regions are bounded by the characteristic lines through the ends. Shaded: the region $3l \leq x + ct \leq 4l$, $-2l \leq x - ct \leq -l$ of the example below.

Moral. The main topic of this lecture is the wave equation with boundary conditions. We used two basic methods: reflection or or explicit formula $f(x + ct) + g(x - ct)$ in each region.

Lecture 6. Periodic problem for the wave equation. Fourier series

References: [Dubrovin, Sec. 3.5], [Arnold, Sec. 4.6, Ch. 5], [Strauss, Sec. 5.5].

6.1 Periodic boundary condition

Consider the periodic boundary condition, or the *circle condition*,

$$u(x + l, t) = u(x, t) \quad \text{for all } x, t, \quad (6.1)$$

with l -periodic initial data ϕ, ψ . The ends of the segment $[0, l]$ are identified (Figure 6.1).

Proposition 6.1. *There exists a unique solution.*

First proof. d'Alembert's formula gives an l -periodic function. □

Second proof. Let u be the solution on $\mathbb{R}$. Then $u(x + l, t)$ is also a solution of the same problem. The solution on $\mathbb{R}$ is unique, hence $u(x + l, t) = u(x, t)$. A periodic solution is a solution on $\mathbb{R}$, hence it is unique. □

Below $l = 2\pi$ (otherwise change the scale of x , which changes c). Typical periodic functions are $\sin x, \cos x, e^{ikx}$ with $k \in \mathbb{Z}$.

Lemma 6.2. *If $\phi(x) = e^{inx}$ and $\psi(x) = 0$, then*

$$u(x, t) = e^{inx} \cos(nct). \quad (6.2)$$

If $\phi(x) = 0$ and $\psi(x) = e^{inx}$, then

$$u(x, t) = \begin{cases} e^{inx} \frac{1}{nc} \sin(nct), & n \neq 0, \\ t, & n = 0. \end{cases} \quad (6.3)$$

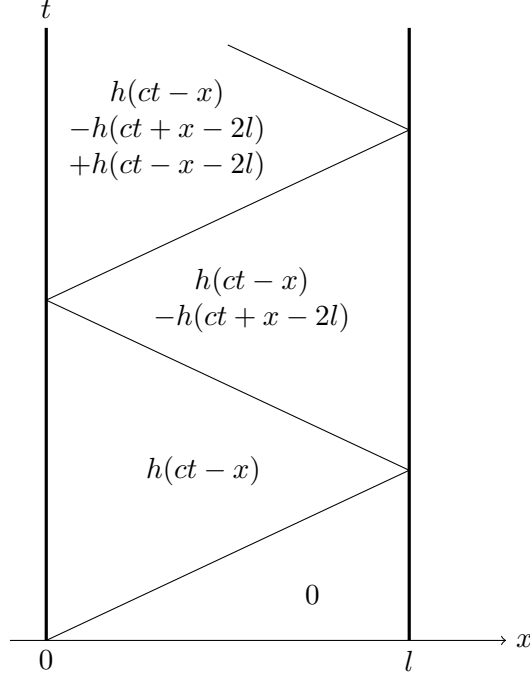


Figure 5.4: Segment with $u|_{x=0} = h(ct)$, $u|_{x=l} = 0$: the values of u_h in the regions bounded by the characteristic line from $(0,0)$ and its reflections.

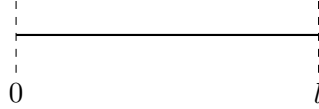


Figure 6.1: The circle condition: the ends of the segment are identified.

6.2 Another idea: eigenvectors

Recall the ODE

$$u_t = Au, \quad (6.4)$$

where u lies in a finite-dimensional vector space $\mathbb{C}^N$. If A has a basis of eigenvectors $u_1, \dots, u_N$ with eigenvalues $\lambda_1, \dots, \lambda_N$, then

$$u|_{t=0} = \sum_i c_i u_i \implies u(t) = \sum_i c_i u_i e^{\lambda_i t}. \quad (6.5)$$

That is, to solve the equation we need to

1. find the eigenvectors;
2. find the coefficients c_i .

Step 2 is simplified if A is self-adjoint, $(Au, \tilde{u}) = (u, A\tilde{u})$. Then there exists an orthogonal eigenbasis, and

$$c_i = \frac{(u, u_i)}{(u_i, u_i)}. \quad (6.6)$$

Idea. For $A = c^2 \partial_x^2$ consider a space of functions as an analogue of $\mathbb{C}^N$, and find orthogonal eigenvectors. The eigenfunctions of ∂_x are e^{ikx} ; periodicity gives $k \in \mathbb{Z}$.

Consider the Hermitian scalar product

$$(f, g) = \frac{1}{2\pi} \int_0^{2\pi} f(x) \overline{g(x)} dx. \quad (6.7)$$

It satisfies

$$(f, g) = \overline{(g, f)}, \quad (6.8)$$

$$(\lambda f_1 + \mu f_2, g) = \lambda(f_1, g) + \mu(f_2, g), \quad (f, \lambda g_1 + \mu g_2) = \bar{\lambda}(f, g_1) + \bar{\mu}(f, g_2), \quad (6.9)$$

$$(f, f) > 0 \quad \text{for } f \neq 0. \quad (6.10)$$

Integrating by parts, we get

$$\partial_x^* = -\partial_x, \quad (\partial_x^2)^* = \partial_x^2. \quad (6.11)$$

The exponentials form an orthonormal basis:

$$(e^{imx}, e^{inx}) = \delta_{m,n}. \quad (6.12)$$

Remark 6.3. Over $\mathbb{R}$ one can use the orthogonal basis

$$\cos nx = \frac{1}{2} (e^{inx} + e^{-inx}), \quad \sin nx = \frac{1}{2i} (e^{inx} - e^{-inx}). \quad (6.13)$$

6.3 Fourier series

Definition 6.4. Let f be a 2π -periodic function on $\mathbb{R}$. Its *Fourier series* is

$$\sum_{n \in \mathbb{Z}} c_n e^{inx}, \quad c_n = (f, e^{inx}) = \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-inx} dx. \quad (6.14)$$

We want the Fourier series to converge to f . This requires some smoothness of f . Consider the partial sums

$$S_N(x) = \sum_{n=-N}^N c_n e^{inx}. \quad (6.15)$$

This is the orthogonal projection of f onto the $(2N+1)$ -dimensional space

$$V_N = \langle e^{-iNx}, \dots, e^{-ix}, 1, e^{ix}, \dots, e^{iNx} \rangle, \quad (6.16)$$

since $S_N \in V_N$ and $(S_N, g) = (f, g)$ for all $g \in V_N$. In particular,

$$(S_N, S_N) = \sum_{n=-N}^N |c_n|^2. \quad (6.17)$$

Lemma 6.5 (Bessel inequality). For $f \in L^2$,

$$\sum_{n \in \mathbb{Z}} |c_n|^2 \leq (f, f). \quad (6.18)$$

Proof. We have

$$0 \leq (f - S_N, f - S_N) = (f, f) - (f, S_N) - (S_N, f) + (S_N, S_N) = (f, f) - (S_N, S_N). \quad (6.19)$$

□

Geometrically, the length of the projection is at most the length of the vector.

Corollary 6.6. $\sum_{n \in \mathbb{Z}} |c_n|^2 < \infty$; in particular, $c_n \rightarrow 0$ as $|n| \rightarrow \infty$.

Theorem 6.7. *If $f \in C^1$ is 2π -periodic, then $S_N(x) \rightarrow f(x)$ for every x .*

Proof. Step 1. Substituting the coefficients and then changing the variable of integration $y = x + \theta$ (the integral of a 2π -periodic function over any period is the same), we get

$$\begin{aligned} S_N(x) &= \sum_{n=-N}^N c_n e^{inx} = \frac{1}{2\pi} \int_{-\pi}^{\pi} dy f(y) \sum_{n=-N}^N e^{-iny} e^{inx} \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} d\theta f(x + \theta) \sum_{n=-N}^N e^{-in\theta} = \frac{1}{2\pi} \int_{-\pi}^{\pi} d\theta f(x + \theta) \frac{\sin(N + \frac{1}{2})\theta}{\sin \frac{1}{2}\theta}. \end{aligned} \quad (6.20)$$

Here we used

$$\left(e^{iN\theta} + \dots + e^{-iN\theta} \right) \left(e^{i\theta/2} - e^{-i\theta/2} \right) = e^{i(N+\frac{1}{2})\theta} - e^{-i(N+\frac{1}{2})\theta}. \quad (6.21)$$

Step 2. Note that

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{n=-N}^N e^{in\theta} d\theta = 1, \quad \text{hence} \quad f(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) \frac{\sin(N + \frac{1}{2})\theta}{\sin \frac{1}{2}\theta} d\theta. \quad (6.22)$$

Therefore

$$f(x) - S_N(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} d\theta (f(x) - f(x + \theta)) \frac{\sin(N + \frac{1}{2})\theta}{\sin \frac{1}{2}\theta} = \frac{1}{4\pi} \int_{-2\pi}^{2\pi} d\theta g(\theta) \sin(N + \frac{1}{2})\theta, \quad (6.23)$$

where

$$g(\theta) = \frac{f(x) - f(x + \theta)}{\sin \frac{1}{2}\theta} \quad (6.24)$$

is a 4π -periodic continuous function (at $\theta = 0, \pm 2\pi$ because $f \in C^1$).

Step 3. The functions $\sin(N + \frac{1}{2})\theta$ are orthogonal on $[-2\pi, 2\pi]$, hence the scalar products tend to zero. Indeed, they are bounded by $\frac{1}{2}(|c_{N+\frac{1}{2}}| + |c_{-N-\frac{1}{2}}|)$, where $c_{k/2}$, $k \in \mathbb{Z}$, are the Fourier coefficients of the 4π -periodic function g , and these tend to zero by the Bessel inequality. $\square$

Theorem 6.8. *If $f \in C^1$, then $S_N \rightarrow f$ uniformly.*

Proof. Let

$$f' = \sum_{n \in \mathbb{Z}} c'_n e^{inx}. \quad (6.25)$$

Integrating by parts, we get

$$c'_n = \frac{1}{2\pi} \int_0^{2\pi} f'(x) e^{-inx} dx = -\frac{1}{2\pi} \int_0^{2\pi} f(x) (-ine^{-inx}) dx = inc_n, \quad (6.26)$$

and $\sum |c'_n|^2 < \infty$ by the Bessel inequality. By Theorem 6.7,

$$f(x) - S_N(x) = \sum_{|n| > N} c_n e^{inx}, \quad (6.27)$$

which is bounded by $\sum_{|n| > N} |c_n|$. Finally,

$$|c_n| = \left| \frac{1}{n} c'_n \right| \leq \frac{1}{2} \left(|c'_n|^2 + \frac{1}{n^2} \right), \quad (6.28)$$

and both series converge. $\square$

Here we used the inequality $|ab| \leq \frac{1}{2}(|a|^2 + |b|^2)$. Equally, one can use the Cauchy–Bunyakovsky–Schwarz inequality

$$\sum_n |a_n b_n| \leq \left(\sum_n |a_n|^2 \right)^{1/2} \left(\sum_n |b_n|^2 \right)^{1/2}, \quad |(a, b)| \leq (a, a)^{1/2} (b, b)^{1/2}. \quad (6.29)$$

Remark 6.9. *The conditions of Theorems 6.7 and 6.8 can be weakened. In Theorem 6.7 it suffices that f is bounded and integrable and the difference quotient $(f(x + \theta) - f(x))/\theta$ is bounded for small θ , for example, f is differentiable at the point x : the proof uses only that g is square integrable. In Theorem 6.8 it suffices that f is continuous and piecewise C^1 : the proof uses only the integration by parts and the Bessel inequality for f' . Continuity alone is not sufficient: there are continuous functions whose Fourier series diverges at a point.*

Example 6.10. Let $f(x) = \text{sign } x$ for $-\pi < x < \pi$, extended 2π -periodically. Its Fourier series converges at every point, but not uniformly: the sum is discontinuous. See the visualization.

Corollary 6.11 (Parseval equality). *For $f \in C^1$,*

$$\sum_{n \in \mathbb{Z}} |c_n|^2 = (f, f). \quad (6.30)$$

This is an infinite-dimensional analogue of the Pythagorean theorem.

Proof. $S_N \rightarrow f$ uniformly, hence $\sum_{|n| \leq N} |c_n|^2 = (S_N, f) \rightarrow (f, f)$. □

Corollary 6.12. *If $f, g \in C^1$ and $c_n(f) = c_n(g)$ for all n , then $f = g$.*

Smoothness is needed here: if we modify f at one point, the coefficients c_n remain the same.

Remark 6.13. *More generally, if $f \in C^k$, then $|c_n n^k| \rightarrow 0$.*

6.4 Solution of the periodic problem

Expand the initial data:

$$\phi(x) = \sum_{n \in \mathbb{Z}} \phi_n e^{inx}, \quad \psi(x) = \sum_{n \in \mathbb{Z}} \psi_n e^{inx}. \quad (6.31)$$

By Lemma 6.2, the solution is

$$u(x, t) = \sum_{n \in \mathbb{Z}} \phi_n \cos(nct) e^{inx} + \frac{1}{c} \sum_{n \neq 0} \psi_n \frac{\sin(nct)}{n} e^{inx} + \psi_0 t. \quad (6.32)$$

For $\phi \in C^1$, $\psi \in C^0$ the series converges.

Remark 6.14 (Real solutions). *One can use the basis $\cos nx$, $\sin nx$ of Remark 6.3:*

$$f(x) = \frac{a_0}{2} + \sum_{n \geq 1} (a_n \cos nx + b_n \sin nx), \quad (6.33)$$

convergent for $f \in C^1$, where

$$a_n = c_n + c_{-n} = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos nx \, dx, \quad b_n = i(c_n - c_{-n}) = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin nx \, dx. \quad (6.34)$$

Moral. Today we solved the wave equation with the periodic boundary condition in two ways. The first is d'Alembert's formula; the second, as for a linear ODE, is the expansion of the solution in the eigenbasis e^{inx} of the self-adjoint operator ∂_x^2 . This method is very general, and we will need it again. In the answer (6.32) each term is a product of a function of t and a function of x : this is the separation of variables. To justify the method we proved the convergence of Fourier series.

References

- [Arnold] V. I. Arnold, *Lectures on Partial Differential Equations*, Springer, 2004.
- [Dubrovin] B. Dubrovin, course notes, https://people.sissa.it/~dubrovin/fm1_web.pdf.
- [Evans] L. C. Evans, *Partial Differential Equations*, AMS, 2010.
- [Strauss] W. A. Strauss, *Partial Differential Equations: An Introduction*, Wiley, 2008.