University of Luxembourg
The Faculty of Science, Technology and Medicine

Analyse 4b

Damjan Pištalo

Contents

1 Curves and Surfaces

Aim of the present section is to introduce curves and surfaces; their tangent spaces, and normal vectors. Solid understanding of those elementary concepts in differential geometry comes to use in the method of characteristic curves for solving first order linear PDE-s. We begin by recalling the notion of a C^1-function within the context of vector calculus:

1.1 C^1 functions

A *multivariable function* is a function $f : \Omega \to \mathbb{R}$, where $\Omega \subseteq \mathbb{R}^n$ is an open set. In physics, it is commonly referred to as a *scalar field*: indeed, it assigns a single number (ex. temperature, electric or gravitational potential, etc.) to each point of a region in space. A multivariable function f is called a C^1 *function* if all the partial derivatives

$$\partial_k f,\ \ 1 \le k \le n$$

exist and are continuous as functions from Ω to $\mathbb{R}$.

A *vector-valued multivariable function* is a function $\vec{f} : \Omega \to \mathbb{R}^m$. In the language of physics, we speak of *vector fields*: to each point in a region of space, they assign a vector. In coordinates, vector-valued functions are determined by m multivariable functions $f_1, \dots, f_m : \Omega \to \mathbb{R}$. A multivariable vector-valued function $\vec{f} = (f_1, \dots, f_m)$ is called a C^1 function if for every $1 \le l \le m$, the function $f_l : \Omega \to \mathbb{R}$ is C^1.

Next, we define curves and surfaces in two ways: as parametric objects, and as level sets.

1.2 Parametric curves and surfaces

Parametric curves are best understood through kinematics. In fact, using the language of physics, parametric curves are trajectories of a point particle. Indeed, they are defined as C^1 functions $\gamma(t) = (x(t), y(t), z(t)) : I \to \mathbb{R}^3$ from an interval (call it a time interval) to the 3-dimensional space. During the course, we also consider parametric plane curves which are valued in $\mathbb{R}^2$. *Tangent vector* to a parametric curve γ at the point $\gamma(t)$ is the velocity of the point particle, i.e. the vector

$$\vec{\gamma}'(t) = (\frac{dx}{dt}(t), \frac{dy}{dt}(t), \frac{dz}{dt}(t)).$$

Parametric curve is called *regular* if $\vec{\gamma}'(t) \neq 0$ for all $t \in I$.

Example 1.1. 1. For real parameters $x_0, y_0, z_0, a, b,$ and c such that at least one of the numbers a, b, c is non-zero,

$$x(t) = x_0 + at, \;\; y(t) = y_0 + bt, \;\; z(t) = z_0 + ct,$$

defines the straight line which passes through the point (x_0, y_0, z_0) in direction of the vector (a, b, c).

2. For $R > 0$,

$$x(t) = x_0 + R\cos(t), \;\; y(t) = y_0 + R\sin(t)$$

defines a circle in $\mathbb{R}^2$ around the point (x_0, y_0) of radius R.

Similarly, *parametric surfaces* are defined as C^1 functions $\phi : \Omega \to \mathbb{R}^3$, for $\Omega \subset \mathbb{R}^2$ open. A parametric surface is called *regular* if the Jacobian matrix of ϕ is of maximal rank at every point of the curve, that is, denoting $\phi(r,s) = (x(r,s), y(r,s), z(r,s))$, the matrix

$$\mathrm{Jac}(\phi)(r,s) = \begin{bmatrix} \frac{\partial x}{\partial r}(r,s) & \frac{\partial x}{\partial s}(r,s) \\ \frac{\partial y}{\partial r}(r,s) & \frac{\partial y}{\partial s}(r,s) \\ \frac{\partial z}{\partial r}(r,s) & \frac{\partial z}{\partial s}(r,s) \end{bmatrix}$$

has linearly independent columns for every $(r,s) \in \Omega$.

1.3 Curves and surfaces as a level-sets

Another way to define a curve in $\mathbb{R}^2$ is by specifying its equation. For example,

$$Ax + By + C = 0, \;\text{ for } A, B, C \in \mathbb{R}$$

specifies a straight line if at least one of the parameters A and B is non-zero. Similarly,

$$(x - x_0)^2 + (y - y_0)^2 - R^2 = 0, \;\text{ for } x_0, y_0, \in \mathbb{R}, \;\text{ and } R \in \mathbb{R}_{>0}$$

defines a circle around the point (x_0, y_0) of radius R. In general, a C^1 function $F : \Omega \to \mathbb{R}$, where $\Omega \subset \mathbb{R}^2$ is an open set, specifies a *level curve* as the set of all the points in Ω where F vanishes:

$$\gamma_F = \{(x,y) \in \Omega : F(x,y) = 0\}.$$

A level curve is *regular* if it is a non-empty set, and

$$\vec{\nabla} F(x,y) = (\partial_x F(x,y), \partial_y F(x,y)) \neq 0, \quad \forall (x,y) \in \gamma_F. \tag{1}$$

Importantly, the regularity condition (1) guarantees that a tangent line to a curve at (x, y) is well-defined. In fact, as shown in the following section, $\vec{\nabla} F(x, y)$ is perpendicular to the tangent line of γ_F at the point (x, y). Points on a curve at which regularity condition is not satisfied are called *singularities*. They include cusps, self-intersections, and isolated points.

Example 1.2 (cusp). Given

$$F : \mathbb{R}^2 \to \mathbb{R},\ \ F(x, y) = x^3 - y^2,$$

$\vec{\nabla} F(x, y) = (3x^2, -2y)$. Hence, $\vec{\nabla} F(x, y)$ vanishes at the origin. As origin is contained in γ_F, it is a singularity. This is an example of a cusp (see Figure 1).

Example 1.3 (self-intersection). Given

$$G : \mathbb{R}^2 \to \mathbb{R},\ \ G(x, y) = x^3 + x^2 - y^2,$$

$\vec{\nabla} G(x, y) = (3x^2 + 2x, -2y)$. Hence, $\vec{\nabla} G(x, y)$ vanishes at the origin. As origin is contained in γ_G, it is a singularity. This is an example of a self-intersection (see Figure 2).

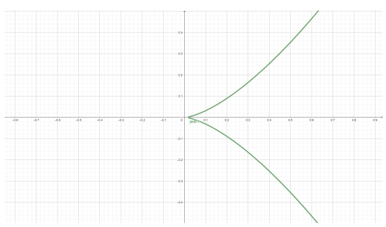


Figure 1: $x^3 - y^2 = 0$

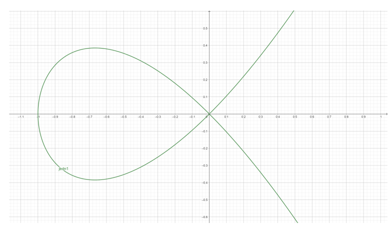


Figure 2: $x^3 + x^2 - y^2 = 0$

Similarly to curves in $\mathbb{R}^2$, surfaces in $\mathbb{R}^3$ can be defined by their equation. For example,

$$Ax + By + Cz + D = 0,\ \text{for } A, B, C, D \in \mathbb{R}$$

specifies a plane if at least one of the parameters A, B and C is non-zero. Similarly,

$$(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2 - R^2 = 0,\ \text{for } x_0, y_0, z_0 \in \mathbb{R},\ \text{and } R \in \mathbb{R}_{>0}$$

defines a sphere around the point (x_0, y_0) of radius R, and

$$(x - x_0)^2 + (y - y_0)^2 - R^2 = 0,\ \text{for } x_0, y_0 \in \mathbb{R},\ \text{and } R \in \mathbb{R}_{>0}$$

defines a vertical cylinder of radius R centered at the vertical line which intersects the xy plane at the point (x_0, y_0). In general, a C^1 function $F : \Omega \to \mathbb{R}$, where $\Omega \subset \mathbb{R}^3$ is an open set, specifies a *level surface*

$$S_F = \{(x, y, z) \in \Omega : F(x, y, z) = 0\}.$$

S_F is *regular* if it is a non-empty set, and

$$\vec{\nabla} F(x, y, z) = (\partial_x F(x, y, z), \partial_y F(x, y, z), \partial_z F(x, y, z)) \neq 0 \quad \forall (x, y, z) \in S_F.$$

Notably, any regular level surface can be locally parametrized. We will prove this using the implicit function theorem. In short, the theorem gives sufficient conditions for an equation $F(x_1, \ldots, x_n, y) = 0$ to have a local solution $y = f(x_1, \ldots, x_n)$. In other words, it provides conditions for a level surface (if $n = 2$) or a level curve (if $n = 1$) to be the graph of a function. For example, the circle $x^2 + y^2 = 1$ is the graph of a function locally around any point (x_0, y_0) of the circle other than $(\pm 1, 0)$: if $y_0 > 0$ the function is $y = \sqrt{1 - x^2}$, and if $y_0 < 0$ the function is $y = -\sqrt{1 - x^2}$.

Theorem 1.4. *(Implicit function theorem) Let $F : \Omega \to \mathbb{R}$, for $\Omega \subset \mathbb{R}^{n+1}$ open, be a C^1 function. Given a point $(\mathbf{x}, y) = (x_1, \ldots, x_n, y) \in \Omega$ such that $F(\mathbf{x}, y) = 0$ and $\partial_n F(\mathbf{x}, y) \neq 0$, there exists $U \subset \mathbb{R}^n$ open which contains $\mathbf{x}$, and a function $g : U \to \mathbb{R}$ such that $g(\mathbf{x}) = y$ and $F(\mathbf{x}, g(\mathbf{x})) = 0$.*

Let's come back to the proof that every surface can be locally parametrized! Assume the regularity condition to be satisfied at a point (x_0, y_0, z_0) of the level surface S_f. By the regularity condition, at least one partial derivative of F is non-zero. Permuting the variables if necessary, assume $\partial_z F(x_0, y_0, z_0) \neq 0$. For $g : U \to \mathbb{R}$ provided by the implicit function theorem, a parametrization of S_F in the neighborhood of (x_0, y_0, z_0) is given by

$$x(r, s) = r, \;\; y(r, s) = s, \;\; z(r, s) = g(r, s).$$

1.4 Tangent spaces and normal vector fields

Let S_F be a regular surface and $\mathbf{x} = (x, y, z) \in S_F$ a point. Let $I \subset \mathbb{R}$ be an interval, and let

$$\gamma : I \to \mathbb{R}^3, \;\; \gamma(0) = \mathbf{x}, \gamma(I) \subset S_F$$

be a parametric curve through $\mathbf{x}$, whose image is contained in the surface S_F. As $F \circ \Gamma \equiv 0$ is a constant function, by the chain rule for multi-variable functions,

$$0 = (F \circ \gamma)'(0) = \vec{\nabla} F(\gamma(0)) \cdot \vec{\gamma}'(0).$$

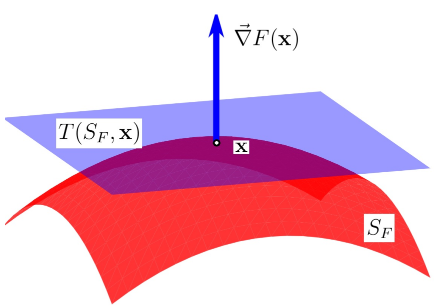
Hence, the tangent vector at $\mathbf{x}$ to an arbitrary parametrized curve in S_F which passes through $\mathbf{x}$ is perpendicular to the gradient $\vec{\nabla} F(\mathbf{x})$.

The set of tangent vectors at $\mathbf{x}$ of all such curves

$$T(S_F, \mathbf{x}) = \{\vec{\gamma}'(0) : \gamma : I \to \mathbb{R}^3 \text{ is a curve s.t. } \gamma(0) = \mathbf{x}, \gamma(I) \subset S_F\}$$

is called the *tangent space* to S_F at $\mathbf{x}$. A vector field which is at each point $\mathbf{x} \in S_F$ an element of the tangent space is called a *tangent vector field*. On the other side, a vector field which is at every point $\mathbf{x} \in S_F$ perpendicular to the tangent space $T(S_F, \mathbf{x})$ is called a *normal vector field*. In particular, $\vec{\nabla} F$ is a normal vector field. We end the section by proving that

$$T(S_F, \mathbf{x}) = \{\vec{v} \in \mathbb{R}^3 : \vec{v} \cdot \vec{\nabla} F(\mathbf{x}) = 0\}. \tag{2}$$



Recall that for two vector spaces V and W of the same dimension, $V \subseteq W$ implies $V = W$. As the right-hand-side of the equation (2) is a 2-dimensional vector space which contains $T(S_F, \mathbf{x})$, it suffices to find a two-dimensional vector space contained in $T(S_F, \mathbf{x})$. Let

$$g : U \to \mathbb{R}^3, \quad (r, s) \mapsto (x(r, s), y(r, s), z(r, s))$$

be a parametrization of S_F in a neighborhood of $\mathbf{x}$. Let $(r_0, s_0) \in U$ be such that $g(r_0, s_0) = \mathbf{x}$. Given a parametrized curve $\gamma : I \to U$ with $\gamma(0) = (r_0, s_0)$, $g \circ \gamma$ is a parametrized curve in S_F. By the chain rule for multi-variable functions,

$$(g \circ \gamma)'(0) = \mathrm{Jac}(g)(r_0, s_0) \cdot \gamma'(0) = \begin{bmatrix} \frac{\partial x}{\partial r}(r_0, s_0) & \frac{\partial x}{\partial s}(r_0, s_0) \\ \frac{\partial y}{\partial r}(r_0, s_0) & \frac{\partial y}{\partial s}(r_0, s_0) \\ \frac{\partial z}{\partial r}(r_0, s_0) & \frac{\partial z}{\partial s}(r_0, s_0) \end{bmatrix} \begin{bmatrix} r'(0) \\ s'(0) \end{bmatrix}$$

Thus, the set $T(S_F, \mathbf{x})$ contains the image of $\mathrm{Jac}(g)(r_0, s_0)$, which is by the regularity condition a two dimensional subspace of $\mathbb{R}^3$.

1.5 Curves and surfaces as graphs of functions

Important examples of level-sets are graphs of C^1 functions. Given a C^1 function $u : \mathbb{R} \to \mathbb{R}$, its graph is the set of all the points $(x, y) \in \mathbb{R}^2$ such that $y = u(x)$; equivalently, such that the function $F : \mathbb{R}^2 \to \mathbb{R}$ defined by $F(x, y) = y - u(x)$ vanishes. As $\vec{\nabla} F$ is nowhere-vanishing, graph of u is a regular curve.

Similarly, graphs of C^1 functions from $\mathbb{R}^2$ to $\mathbb{R}$ define regular surfaces. Indeed, let

$$u : \Omega \to \mathbb{R}, \;\; \Omega \subseteq \mathbb{R}^2$$

be a C^1 function. The graph of u is defined as

$$\Gamma_u = \{(x, y, z) \in \Omega \times \mathbb{R} : z = u(x, y)\}.$$

Proposition 1.5. *Γ_u is a regular surface and $\vec{n} = (\partial_x u, \partial_y u, -1)$ is its normal vector field.*

Proof. Define

$$F : \Omega \times \mathbb{R} \to \mathbb{R}, \;\; F(x, y, z) = u(x, y) - z.$$

Clearly, $\Gamma_u = S_F$. The gradient $\vec{\nabla} F = (\partial_x u, \partial_y u, -1)$ is nowhere-vanishing, and it is a normal vector field to S_F. □

2 First order partial differential equations: method of characteristic curves

2.1 Linear equations

Let $\Omega \subset \mathbb{R}^2$ be open, and let $a, b, c : \Omega \to \mathbb{R}$ be continuous functions. Equation

$$a(x, t)\partial_x u + b(x, t)\partial_t u = c(x, t) \tag{3}$$

is called a *first order linear PDE*. Its *local solution* is a C^1 function

$$u : U \to \mathbb{R}, \;\; U \subset \Omega \text{ open}$$

which satisfies the equation (3). It is a *global solution* if $U = \Omega$.

Given a local solution u, its graph Γ_u is called the *solution surface*. Observe that the scalar product

$$(a,b,c)\cdot(\partial_x u,\partial_t u,-1)=a\partial_x u+b\partial_t u-c$$

vanishes if and only if equation (3) is satisfied. Thus, the vector field (a,b,c) is tangent to the solution surface. To solve the equation (3) amounts to reconstruct the solution surface from its tangent vector field (a,b,c) and the appropriate initial condition.

As the first step, starting from a vector field $\vec{v}$ and a point $\mathbf{x}_0\in\mathbb{R}^3$, one would like to reconstruct a parametric curve $\gamma: I\to\mathbb{R}^3$ which passes through $\mathbf{x}_0$ and to which the vector fields $\vec{v}$ is tangent; meaning that $\vec{\gamma}'(s)=\vec{v}(\gamma(s))$ for every $s\in I$. Such curves are called *integral curves* of the vector field $\vec{v}$, and they always exist. In fact, under additional mild conditions on the vector field $\vec{v}$ they are even unique. Indeed:

Theorem 2.1. *For $V\subset\mathbb{R}^3$, given a continuous vector field $\vec{v}: V\to\mathbb{R}^3$ and a point $\mathbf{x}_0\in V$, there exists an open interval $I\ni 0$, and a curve $\gamma: I\to V$ such that $\gamma(0)=\mathbf{x}_0$ and $\vec{\gamma}'(t)=\vec{v}(\gamma(t))$.*

Proof. Denote $\vec{v}(\mathbf{x})=(a(\mathbf{x}),b(\mathbf{x}),c(\mathbf{x}))$, $\mathbf{x}_0=(x_0,t_0,z_0)$. A parametrized curve $\gamma(s)=(x(s),t(s),z(s))$ which satisfies the theorem is a local solution of the system of ordinary differential equations

$$\begin{array}{ccc} x'=a(x,t,z) & t'=b(x,t,z) & z'=c(x,t,z)\\ x(0)=x_0 & t(0)=t_0 & z(0)=z_0. \end{array}\tag{4}$$

By the fundamental theorem of ordinary differential equations, the system has a solution, which is unique if $\vec{v}$ is Lipschitz continuous (whatever it might mean). □

We are now ready to solve the initial value problem posed by the equation (3), together with the initial condition $u(x,0)=f(x)$, where $f: I\to\mathbb{R}$, with $I\times 0\subset\Omega$, is a given function. For every $x_0\in I$, there exists a (unique) integral curve

$$\gamma_{x_0}: I\to\Omega\times\mathbb{R},\ \ \gamma_{x_0}(s)=(x(x_0,s),t(x_0,s),z(x_0,s))$$

of the continuous vector field

$$(a,b,c):\Omega\times\mathbb{R}\to\mathbb{R},\ \ (x,t,z)\mapsto(a(x,t),b(x,t),c(x,t))$$

which passes through the point $(x_0,0,f(x_0))$. Those curves (for different x_0) are called *characteristic curves* of the equation. Their projections to Ω are integral curves or the vector field (a, b), and are called *projected characteristic curves*. The solution surface is the union of all characteristic curves. All that remains is to determine if the solution surface is a graph of a function, and, if so, to compute the function in question.

Assume the function

$$(x_0, s) \mapsto (x(x_0, s), t(x_0, s))$$

to be invertible, and denote its inverse by $(x_0(x, t), s(x, t))$. With this, the solution surface is indeed the graph of function $u(x, t) = z(x_0(x, t), s(x, t))$.

Example 2.2. (transport equation) Solve the initial value problem

$$c\partial_x u + \partial_t u = 0, \;\; a \in \mathbb{R}, \;\; u(x_0, 0) = f(x_0).$$

The unique integral curve $\gamma_{x_0}(s) = (x(x_0, s), t(x_0, s), z(x_0, s))$ of the constant vector field $(c, 1, 0)$ which passes through the point $(x_0, 0, f(x_0))$ is the solution to the system of ordinary differential equations (x_0 is a parameter whose derivations do not enter the system, compare with (4))

$$\begin{array}{lll} \partial_s x(x_0, s) = c & \partial_s t(x_0, s) = 1 & \partial_s z(x_0, s) = 0 \\ x(x_0, 0) = x_0 & t(x_0, 0) = 0 & z(x_0, 0) = f(x_0). \end{array}$$

which is clearly

$$x(x_0, s) = x_0 + cs \quad t(x_0, s) = s \quad z(x_0, s) = f(x_0).$$

Inverting the function $(x(x_0, s), t(x_0, s))$, we get $s = t$, $x_0 = x - ct$. Hence, $u(x, t) = z(x_0(x, t), s(x, t)) = f(x - ct)$. Conceptually, the graph of function f travels in the positive direction at speed c. Hence the name transport equation. □

Next, we determine, in a slightly greater generality, conditions under which the function $(x_0, s) \to (x, t)$ is invertible. So-far, the initial condition was determined by the values of u over the x-axis. Now, we specify the initial condition over any parametrized regular C^1 curve $\Gamma : I \to \Omega$, called the *non-characteristic curve*, by setting $u \circ \Gamma(r) = f(r)$. As before, for each $r \in I$, there exists a unique integral curve

$$\gamma_r : I_r \to \Omega \times \mathbb{R}, \;\; \gamma_r(s) = (x(r, s), t(r, s), z(r, s))$$

of the vector field (a, b, c) – a characteristic curve of the equation – such that $\gamma_r(0) = (\Gamma(r), f(r))$. As above, the union of those characteristic curves determines the solution surface if the function

$$(r, s) \mapsto (x(r, s), t(r, s))$$

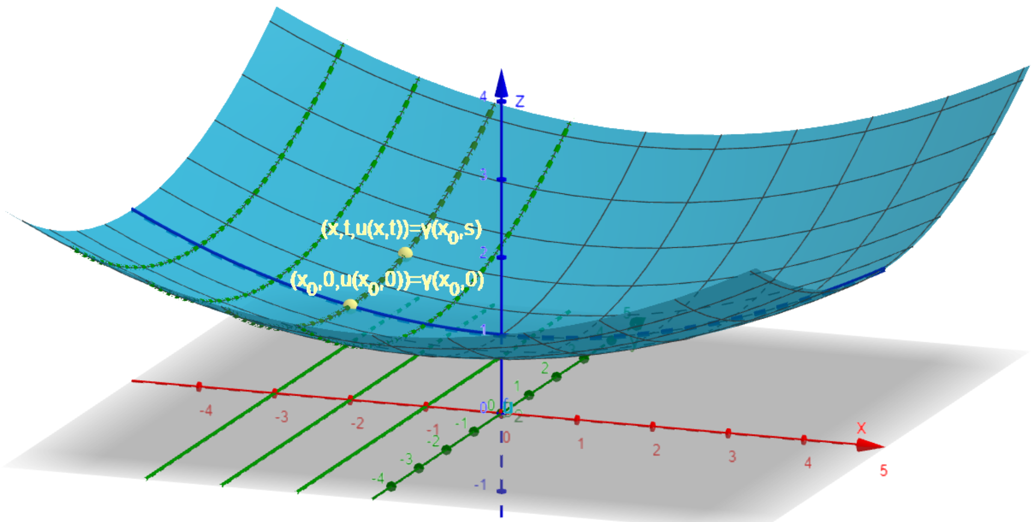


Figure 3: The blue line is the curve specified by the initial condition. The non-characteristic curve (its projection to the (x,t)-plane) is the red line (x-axis). In green are the characteristic curves and projected characteristic curves – their projections to the (x,t)-plane.

is invertible.

A sufficient condition for this is that the vector field (a, b) in the (x, t) plane is nowhere tangent to non-characteristic curve. Indeed, recall the inverse function theorem:

Theorem 2.3. *Let $\Omega \subset \mathbb{R}^n$ be an open subset, and $F : \Omega \to \mathbb{R}^n$ a C^1 function, whose Jacobian*

$$\mathbf{J}_F = \begin{bmatrix} \frac{\partial F_1}{\partial x_1} & \cdots & \frac{\partial F_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial F_n}{\partial x_1} & \cdots & \frac{\partial F_n}{\partial x_n} \end{bmatrix}$$

is regular at a point $\mathbf{x} \in \Omega$. Then there exists an open an open set $U \subset \Omega$ containing $\mathbf{x}$ and an open set $V \subset \mathbb{R}^n$ containing $F(\mathbf{x})$ such that $F|_U : U \to V$ is invertible and whose inverse is a C^1 function.

Applying to $F(r, s) = (x(r, s), t(r, s))$, we find that F is locally invertible in a neighborhood of a point $(r, 0)$ if vectors

$$(\partial_r x(r, 0), \partial_r t(r, 0)) = \vec{\Gamma}'(r), \text{ and } (\partial_s x(r, 0), \partial_s t(r, 0)) = (a(\Gamma(r)), b(\Gamma(r))$$

are linearly independent.

Example 2.4. Solve the differential equation

$$3\partial_x u - 2\partial_t u = x,$$

if $u(x, t)$ is zero when restricted to the line $x + t = 1$.

The non-characteristic curve in the example is determined by the equation $x + t = 1$. A parametrization is given by $\Gamma(r) = (r, 1 - r)$, and the initial condition is given by $u \circ \Gamma(r) = 0$. The unique integral curve $\gamma_r(s) = (x(r, s), t(r, s), z(r, s))$ of the vector field $(3, -2, x)$ which passes through the point $(\Gamma(r), 0)$ is the solution to the system of ordinary differential equations

$$\begin{array}{ccc} \partial_s x(r, s) = 3 & \partial_s t(r, s) = -2 & \partial_s z(r, s) = x(r, s) \\ x(r, 0) = r & t(r, 0) = 1 - r & z(r, 0) = 0, \end{array}$$

which is

$$x(r, s) = 3s + r \quad t(r, s) = -2s + 1 - r, z(r, s) = z(r, 0) + \textstyle\int_0^s x(r, s')ds' = \frac{3}{2}s^2 + rs.$$

Inverting the function $(x(r, s), t(r, s))$, we get $r = -2x - 3t + 3$, $s = x + t - 1$. Hence, $u(x, t) = z(r(x, t), s(x, t)) = \frac{3}{2}(x + t - 1)^2 + (-2x - 3t + 3)(x + t - 1)$. The projected characteristic curves are the parametric curves $s \mapsto (x(r,s), t(r,s))$. Their level-set equations are obtained by eliminating the variable s in the system of linear equations $x = 3s + r$, $t = -2s + 1 - r$, and read $t = -\frac{1}{3}x - \frac{5}{3}r + 1$. Clearly, projected characteristic curves are not tangent to the non-characteristic curve.

Another generalization of the method is to allow the function c to depend on the indeterminate. Concretely, assume the equation is given by

$$a(x,y)\partial_x u + b(x,y)\partial_t u = c(x,t,u).$$

As before, denote the integral curves of (a,b,c) by $\gamma_r(s) = (s(r,s), t(r,s), z(r,s))$. Since they lie on the solution surface, $z(r,s) = u(x(r,s), t(r,s))$. Consequently, the curves γ_r are solutions of the following system:

$$\partial_s x = a(x,t) \quad \partial_s t = b(x,t) \quad \partial_s z = c(x,t,z).$$

2.2 Shock waves

Shock waves appear when one allows the functions a and b to depend on the indeterminate. An example of such equation is a simple model of road traffic. Assume that the traffic flows in one line, with the density of $\rho(x,t)$ cars per meter. The number of cars within a short road segment Δx is given by $\rho(x,t)\Delta x$. Denote by $j(x,t)$ the flux – number of cars which pass during one second through the point x at time t. Total amount of cars passing through x during a short period of time Δt is $\rho(x,t)\Delta t$. Since the number of cars in the segment Δx can change only if cars enter or leave the segment through endpoints, assuming that the traffic flows in the positive direction, we get

$$\rho(x,t+\Delta t)\Delta x - \rho(x,t)\Delta x = j(x,t)\Delta t - j(x+\Delta x,t)\Delta t.$$

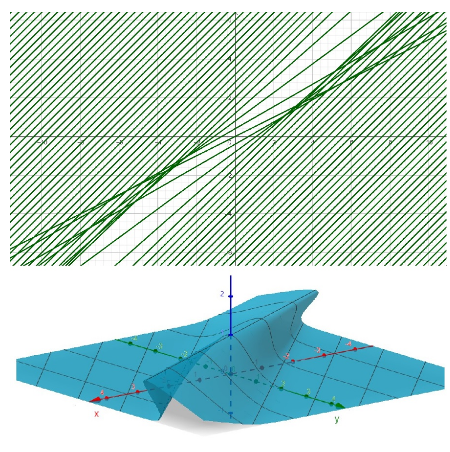


Figure 4: Projected characteristic curves and the solution surface for the shock wave equation with initial condition $f(x) = \frac{1}{4a}e^{-x^2}$.

Dividing both sides of the equation by $\Delta x \Delta t$, and taking the limit $\Delta x, \Delta t \to 0$, we get

$$\partial_t \rho(x,t) = -\partial_x j(x,t). \tag{5}$$

In physics, this equation is referred to as the *continuity equation*.

Observe that if the traffic density at a point x is m cars per meter, and if their velocity in n meters per second, then mn cars pass through the point x per the second. Denoting by $v(x,t)$ the velocity of the cars at the position x and time t, we have

$$j = v\rho. \tag{6}$$

Equations 5 and 6 yield

$$\partial_t \rho = -\partial_x(\rho v).$$

In a simple model, the velocity depends only on the traffic's density. When the density approaches zero, the velocity approaches a constant value c_0 (for example 120 km/h). On the other side, when the density becomes critical ($\rho = 1/a$ for an appropriately determined constant a), the traffic stops. Linear interpolation between the two extremes gives

$$v(\rho) = c_0(1 - a\rho).$$

Using the product rule and chain rule for derivations we get

$$\partial_t \rho = -\partial_x(\rho v) = -v\partial_x \rho - \rho(\partial_\rho v)(\partial_x \rho) = -c_0(1 - 2a\rho)\partial_x \rho.$$

Hence, differential equation which models the trafic floow is

$$c_0(1 - 2a\rho)\partial_x \rho + \partial_t \rho = 0.$$

Assume the initial condition is given as $\rho(x,0) = f(x)$. Characteristic curves are solutions of the system

$$\begin{array}{ccc} \partial_s x(x_0,s) = c_0(1-2az) & \partial_s t(x_0,s) = 1 & \partial_s z(x_0,s) = 0 \\ x(x_0,0) = x_0 & t(x_0,0) = 0 & z(x_0,0) = f(x_0) \end{array}$$

which is

$$t(x_0,s) = s; \quad z(x_0,s) = f(x_0); \text{ and } \quad x(x_0,s) = c_0 s(1 - 2af(x_0)) + x_0. \tag{7}$$

From here we get $x_0 = x - c_0 t(1 - 2az)$, and finally

$$\rho = f(x - c_0 t(1 - 2\rho)). \tag{8}$$

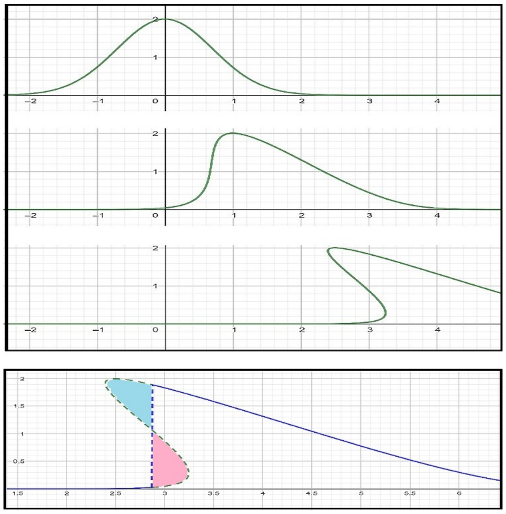


Figure 5: Propagation of the wave front and shock fitting.

By the implicit function theorem 1.4, there exists a unique locally defined function ρ which satisfies the equation as long as

$$\partial_\rho(f(x - c_0 t(1-2\rho)) - \rho) \neq 0.$$

As an illustration, for $f(x) = x$, $\rho = x - c_0 t(1-2a\rho)$, hence $\rho(x,t) = \frac{x-c_0 t}{1-2ac_0 t}$. Notice that the solution u is well-defined only for $t < 1/2ac_0$ – even though the formula itself makes sense for $t > 1/2ac_0$, solutions to initial value problems are considered as valid only before the singularity. The singularity itself is no coincidence. By (7), projected characteristic curves are the straight lines whose equations are

$$x = x_0 + c_0(1 - 2af(x_0))t.$$

Unless f is a constant function, they are bound to intersect. As the z-coordinate is constant along each characteristic curve and differs among different characteristic curves, the solution is not uniquely determined over the intersection. Hence, the parametric surface $(x_0, s) \mapsto (x(x_0, s), t(x_0 s), z(x_0, s))$ is no longer the graph of a function $\rho(x,t)$.

An instructive initial condition is the bump function $f(x) = \frac{1}{4a}e^{-x^2}$. Projected characteristic curves and parametric solution surface for this initial condition are visualized in Figures 4a and 4b. Conceptually, the equation that we are dealing with

$$\partial_t \rho + c_0(1-2a\rho)\partial_x \rho = 0$$

is a transport equation, in which the propagation speed of the wave-front depends on its height: denser the traffic, slower it moves. Consequently, the top of the bump $f(x) = \frac{1}{4a}e^{-x^2}$ "drags behind" and eventually gets "overpassed" by other parts of the wave-front (Figure 5 up). Nevertheless, the solution can be extanded beyond the breaking point with the help of a conservation law: the total amount of cars on the road (the integral $\int_{-\infty}^{\infty} \rho(x,t)dx$, or the area under the wave-front) is constant. This law can be proven if the function f is well behaved. In particular, a bounded C^1 function $|f(x)| < M_1$ which is zero outside of a sufficiently large segment $[-M_2, M_2]$ will do the job. Indeed, by the triangle inequality, at any time t,

$$|x - ct(1-2\rho)| \geq ||x| - |ct(1+2M_1)||.$$

Thus,

$$\rho(x,t) = f(x - ct(1-2\rho)) = 0 \text{ for } |x| > M_2 + |ct(1+2M_1)|.$$

Denote $M = M_2 + |ct(1+2M_1)|$. The continuity equation (5) implies

$$\partial_t \int_{-\infty}^{\infty} \rho(x,t)dx = \partial_t \int_{-M}^{M} \rho(x,t)dx = \int_{-M}^{M} \partial_t \rho(x,t)dx = -\int_{-M}^{M} \partial_x j(\rho(x,t))dx$$
$$= -j(\rho(M,t)) + j(\rho(-M,t)) = -j(0) + j(0) = 0$$

Beyond the braking point, we apply the procedure called *shock fitting*, which is visualized in the Figure (5 below): We remove a segment from the wave-front to obtain the graph of a single-valued function with one discontinuity, without changing the area below. Obtained solution is globally well-defined, satisfies the equation away from the discontinuity, and keeps the conservation law valid.

Physically, the discontinuity should not come as a surprise. In fact, traffic density on a highway is rarely continuous, especially during rush hours.

Infinitesimal version of the conservation law can be used to determine the speed at which the discontinuity (shock) propagates: Assume that at time t, the function $\rho(x,t)$ has a single discontinuity at $x = \xi(t)$. Left (resp. right) of the shock, u is a continuous function denoted by ρ_- (resp. ρ_+). Let Δt be a small time interval, and $\Delta x = \xi'(t)\Delta t$ the distance covered by the discontinuity in time Δt. The change in the number of cars within the interval $[x, \Delta x]$ from time t to $t + \Delta t$ is on one hand equal to $\Delta x(\rho_- - \rho_+)$ and on the other hand equal to the number of cars that passed through the end-points of the segment $\Delta t(j_- - j_+)$. Finally,

$$\frac{d\xi}{dt}(t) = \frac{\Delta x}{\Delta t} = \frac{j_- - j_+}{\rho_- - \rho_+}.$$

This equation is known as *Rankine-Hungoniot formula.*

We conclude the section by noting that all the conclusions made for the simple traffic model in fact apply to a general PDE of type

$$\partial_t u + c(u)\partial_x u = 0, \tag{9}$$

where $c(u)$ is a continuous function. Indeed, denoting by j the primitive of c, equation (9) becomes the continuity equation

$$\partial_t u = -c(u)\partial_x u = -j'(u)\partial_x u = -\partial_x j(u),$$

which is, together with the method of characteristic curves, all we ever used.

2.3 Exercises

Exercise 2.5. Solve the initial value problem

$$\partial_t u + c\partial_x u + au = 0, \ \ u(x,0) = g(x)$$

Why is this equation called transport with decay?

Exercise 2.6. Solve

$$x\partial_y u - y\partial_x u = 0, \ \ u(x,0) = f(x) \text{ for x>0.}$$

Draw the projected characteristic curves, and draw the curve on which the initial condition is specified.

Exercise 2.7. Solve

$$x\partial_y u + y\partial_x u = 0, \ \ u(x,0) = e^{-x^2}.$$

Draw the projected characteristic curves, and draw the curve on which the initial condition is specified. Determine the domain of the solution.

3 Wave equation

We derive the wave equation from the elastic string model.

Consider a flexible homogeneous elastic string of linear density ρ, taut between two walls, subject exclusively to the transversal vibrations. Assume that at a given moment t in time, the shape of the string is the graph of a function $x \mapsto u(x,t)$. Observe the short string segment whose endpoints are $(x, u(x,t))$ and $(x+\Delta x, u(x+\Delta x, t))$. Forces that act on it (see Figure **??**) are the elastic forces on its endpoint. As by the assumption, vibrations are exclusively transversal, the horizontal components of the elastic forces must cancel out, that is,

$$T_1 \cos\theta_1 = T_2 \cos\theta_2 =: T_0. \tag{10}$$

By the second Newton law, the difference of its vertical components is the product of segment's mass and vertical acceleration

$$T_2 \sin\theta_2 - T_1 \sin\theta_1 = \rho \Delta x \partial_t^2 u(x,t). \tag{11}$$

Equations (10) and (11) give

$$T_0 \operatorname{tg}\theta_2 - T_0 \operatorname{tg}\theta_1 = \rho \Delta x \partial_t^2 u(x,t). \tag{12}$$

As $\operatorname{tg}\theta = \partial_x u(x,t)$, dividing the Equation 12 by Δx, and taking the limit $\Delta t \to 0$, we arrive to the wave equation

$$T_0 \partial_x^2 u(x,t) = \rho \partial_t^2 u(x,t).$$

3.1 Wave equation on infinite string

Denoting $c^2 = T_0/\rho$, the wave equation on infinite strings becomes

$$\partial_t^2 u(x,t) - c^2 \partial_x^2 u(x,t) = 0, \ \ (x,t) \in \mathbb{R}^2. \tag{13}$$

As

$$\partial_t^2 u - c^2 \partial_x^2 u = (\partial_t - c\partial_x)(\partial_t + c\partial_x)u = 0,$$

any solution to the transport equation $(\partial_t + c\partial_x)u = 0$ is also a solution of the wave equation. Similarly,

$$\partial_t^2 u - c^2 \partial_x^2 u = (\partial_t + c\partial_x)(\partial_t - c\partial_x)u = 0$$

implies that any solution of $(\partial_t - c\partial_x)u = 0$ is a solution of the wave equation as well. Recalling the general solution of the transport equation (Example 2.2), we conclude that for any C^2 functions ϕ and ψ, $u_1(c,t) = \phi(x+ct)$ and $u_2(c,t) = \psi(x-ct)$ both solve the wave equation. Notice further that the set of solutions of the wave equation is a vector space:

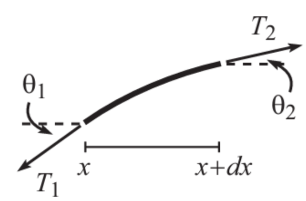


Figure 6:

$$\partial_t^2(a_1u_1 + a_2u_2) - c^2\partial_x^2 u(a_1u_1 + a_2u_2) = a_1(\partial_t^2 u_1 - c^2\partial_x^2 u_1) + a_2(\partial_t^2 u_2 - c^2\partial_x^2 u_2) = 0 + 0 = 0.$$

In conclusion, $u(x,t) = \phi(x+ct) + \psi(x-ct)$ is a solution of the wave equation. In fact, it can be shown that every solution of the wave equation is of the latter form:

Theorem 3.1. *General solution of the wave equation 13 is given by* $u(x,t) = \phi(x+ct) + \psi(x-ct)$*, where* ϕ *and* ψ *are arbitrary* C^2 *functions.*

3.2 Initial value problem

Since the wave equation is in fact the Newton's second law of motion applied to a string, it is reasonable to expect that the motion is uniquely determined by the initial position and velocity of the string. In other words, the initial value problem

$$\begin{aligned} &\partial_t^2 u - c^2 \partial_x^2 u = 0, \ \ (x,t) \in \mathbb{R}^2 \\ &u(x,0) = f(x), \ \ \partial_t u(x,0) = g(x) \end{aligned} \tag{14}$$

should have unique solution. There are multiple approaches to the above problem available in the literature. For example, one could start from the general solution $u(x,t) = \phi(x - ct) + \psi(x+ct)$ of the wave equation, and determine the unknown functions ϕ and ψ from the initial conditions. Alternatively, the method of characteristic curves can be used. Let us follow the latter approach.

Substituting $v(x,t) = (\partial_t - c\partial_x)u$, we get the initial value problem

$$\begin{aligned} &(\partial_t + c\partial_x)v = 0; \\ &v(x,0) = \partial_t u(x,0) - c\partial_x u(x,0) = g(x) - c\partial_x f(x). \end{aligned}$$

The solution is clearly $v(x,t) = g(x-ct) - c\partial_x f(x-ct)$. With this, $u(x,t)$ is the solution of the initial value problem

$$\begin{aligned} &(\partial_t - c\partial_x)u = v = g(x-ct) - c\partial_x f(x-ct) = g(x-ct) - cf'(x-ct) \\ &u(x,0) = f(x). \end{aligned}$$

Using the method of characteristic curves, we reduce to the system of ODE-s

$$\begin{array}{ccc} \partial_s x(x_0,s) = -c & \partial_s t(x_0,s) = 1 & \partial_s z(x_0,s) = g(x-ct) - cf'(x-ct) \\ x(x_0,0) = x_0 & t(x_0,0) = 0 & z(x_0,0) = f(x_0). \end{array}$$

Solution of the first two equations is $x(x_0,s) = x_0 - cs$, $t(x_0,s) = s$, and its inverse is $x_0(x,t) = x + ct$, $t(x_0,s) = s$. As for the third equation, we have

$$\begin{aligned} \partial_s z(x_0,s) &= g(x(x_0,s) - ct(x_0,s)) - cf'(x(x_0,s) - ct(x_0,s)) \\ &= g(x_0 - 2cs) - cf'(x_0 - 2cs). \end{aligned}$$

Integrating, we get

$$\begin{aligned} z(x_0,s) - z(x_0,0) &= \int_0^s g(x_0 - 2cs') - cf'(x_0 - 2cs') ds' \\ &= \frac{1}{2}(f(x_0 - 2cs') - f(x_0)) + \int_0^s g(x_0 - 2cs') ds' \end{aligned}$$

Finally,

$$u(x,t) = z(x_0(x,t), s(x,t)) = \frac{f(x-ct) + f(x+ct)}{2} + \int_0^t g(x_0 - 2cs') ds'.$$

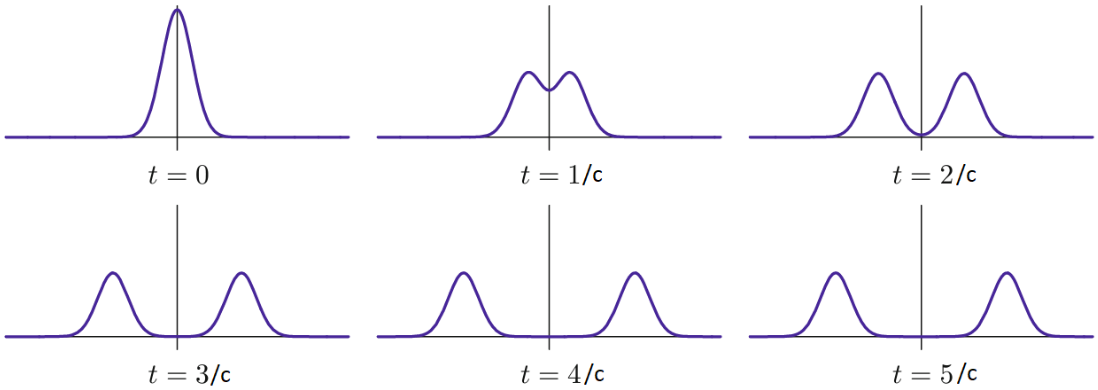


Figure 7: Solution to the wave equation for the initial displacement $f(x) = e^{-x^2}$ and no initial velocity is two wave front traveling in the opposite direction

Substituting $\xi = x_0 - 2cs'$, we arrive to d'Alambert's solution of the wave equation

$$u(x,t) = \frac{f(x-ct)+f(x+ct)}{2} + \frac{1}{2c}\int_{x-ct}^{x+ct} g(\xi)d\xi. \tag{15}$$

A *classical solution* of the wave equation is a C^2 function. For this, f has to be a C^2 function, and g a C^1 function. However, d'Alambert's solution makes sense as long as g is an integrable function. In this case, we speak of *weak solutions.*

3.3 Inhomogeneous wave equation

If $g(x) = 0$, d'Alambert's solution consists of two waves, each of the shape $f(x)/2$, travelling in opposite directions at the constant speed c (see Figure 7). If the initial displacement $f(x)$ is localized, wave front is localized as well, and each point of the string returns to the equilibrium position after the wave passes. In contrary, for $f(x) = 0$, even if the initial velocity is localized, it leaves the string permanently deformed (see Figure 8).

When the vibrating string is subject to external forcing, the wave equation acquires an additional inhomogeneous term:

$$\partial_t^2 u - c^2 \partial_x^2 u = F(x,t). \tag{16}$$

Using the method of characteristic curves, we now solve the inhomogeneous wave equation 16 for the initial conditions

$$u(x,0) = \partial_t u(x,0) = 0.$$

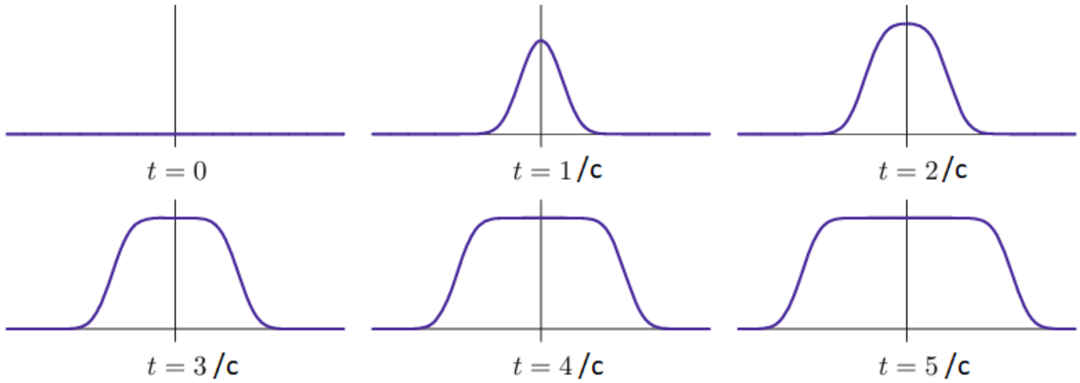


Figure 8: Solution of the wave equation for initial velocity $g(x) = e^{-x^2}$ without initial displacement leaves the string permanently deformed.

Substituting $v(x,t) = (\partial_t + c\partial_x)u$, we get the initial value problem

$$(\partial_t - c\partial_x)v = F(x,t),\ \ v(x,0) = 0$$

Using the method of integral curves, we reduce to the system of ODE-s

$$\begin{array}{ccc} \partial_s x(x_0,s) = -c & \partial_s t(x_0,s) = 1 & \partial_s z(x_0,s) = F(x,t) \\ x(x_0,0) = x_0 & t(x_0,0) = 0 & z(x_0,0) = 0. \end{array}$$

Solution of the first two equations is $x(x_0,s) = x_0 - cs$, $t(x_0,s) = s$, and its inverse is $x_0(x,t) = x + ct$, $t(x_0,s) = s$. As for the third equation, we get

$$\partial_s z(x_0,s) = F(x(x_0,s), t(x_0,s)) = F(x_0 - cs, s)$$

Integrating and inverting, we get

$$v(x,t) = z(x_0(x,t), s(x,t)) = z(x_0,0) + \int_0^s F(x_0 - cs', s')ds' = \int_0^t F(x + c(t-s'), s')ds'$$

With this, $u(x,t)$ is the solution of the initial value problem

$$(\partial_t + c\partial_x)u = \int_0^t F(x + c(t-s'), s')ds',\ \ u(x,0) = 0.$$

Using once again the method of integral curves, we get

$$\begin{array}{ccc} \partial_s x(x_0,s) = c & \partial_s t(x_0,s) = 1 & \partial_s z(x_0,s) = \int_0^t F(x + c(t-s'), s')ds' \\ x(x_0,0) = x_0 & t(x_0,0) = 0 & z(x_0,0) = 0. \end{array}$$

Solution of the first two equations is $x(x_0, s) = x_0 + cs$, $t(x_0, s) = s$, and its inverse is $x_0(x, t) = x - ct$, $t(x_0, s) = s$. Substituting in the third equation, we get

$$\partial_s z(x_0, s) = \int_0^s F(x_0 + c(2s'' - s'), s')ds'.$$

Integrating and inverting, we arrive at

$$\begin{aligned} u(x,t) = z(x_0(x,t), s(x,t)) &= z(x_0, 0) + \int_0^s \int_0^{s''} F(x_0 + c(2s'' - s'), s')ds'ds'' \\ &= \int_0^t \int_0^{s''} F(x + c(2s'' - s' - t), s')ds'ds''. \end{aligned}$$

Finally, we simplify the integral. First, one changes the order of integration. The area over which we integrate is determined by the system of inequalities

$$0 < s' < s'', \;\; 0 < s'' < t,$$

and it is the triangle whose vertices in the (s', s'') plane are $(0,0)$, $(t, 0)$ and (t, t). The same triangle is determined by the system

$$0 < s' < t, \;\; s' < s'' < t.$$

Thus,

$$u(x,t) = \int_0^t \int_{s''}^t F(x + c(2s'' - s' - t), s')ds''ds'.$$

Substituting $\tau(s', s'') = s'$, $\chi(s', s'') = x + c(2s'' - s' - t)$, the integral area becomes

$$0 < \tau < t, \;\; x - c(t - \tau) < \chi < x + c(t - \tau),$$

while the Jacobian determinant of the function $(s', s'') \mapsto (\tau, \chi)$ is

$$\mathrm{Det}(\mathbf{J}) = \det \begin{bmatrix} 1 & 0 \\ -c & 2c \end{bmatrix} = 2c.$$

Hence,

$$u(x,t) = \frac{1}{2c} \int_0^t \int_{x-c(t-\tau)}^{x+c(t-\tau)} F(\chi, \tau)d\chi d\tau.$$

From here it is easy to deduce the solution of the general initial value problem for inhomogeneous wave equation.

Theorem 3.2. *The solution to the initial value problem*

$$\begin{aligned} &\partial_t^2 u - c^2 \partial_x^2 u = F(x,y), \ \ (x,t) \in \mathbb{R}^2 \\ &u(x,0) = f(x), \ \ \partial_t u(x,0) = g(x) \end{aligned} \tag{17}$$

is given by

$$u(x,t) = \frac{f(x-ct)+f(x+ct)}{2} + \frac{1}{2c}\int_{x-ct}^{x+ct} g(\xi)d\xi + \frac{1}{2c}\int_0^t \int_{x-c(t-\tau)}^{x+c(t-\tau)} F(\chi,\tau)d\chi d\tau.$$

Proof. Denoting

$$u_1(x,t) = \frac{f(x-ct)+f(x+ct)}{2} + \frac{1}{2c}\int_{x-ct}^{x+ct} g(\xi)d\xi$$

and

$$u_2(x,t) = \frac{1}{2c}\int_0^t \int_{x-c(t-\tau)}^{x+c(t-\tau)} F(\chi,\tau)d\chi d\tau,$$

$u = u_1 + u_2$. Thus,

$$\begin{aligned} &\partial_t^2 u - c^2\partial_x^2 u = (\partial_t^2 u_1 - c^2\partial_x^2 u_1) + (\partial_t^2 u_2 - c^2\partial_x^2 u_2) = 0 + F(x,y), \\ &u(x,0) = u_1(x,0) + u_2(x,0) = f(x) + 0, \\ &\partial_t u(x,0) = \partial_t u_1(x,0) + \partial_t u_2(x,0) = g(x) + 0. \end{aligned}$$

□

3.4 Causality

Given a point (x,t) in space-time, the triangle whose vertices are $(x-ct,0)$, $(x+ct,0)$, and (x,t) is called the *domain of dependence* of the point (x,t). This is because for $t > 0$, the solution $u(x,t)$ of the initial value problem (17) depends only on the values of the initial data and forcing function at points within the triangle. Indeed, in the solution formula, the first term requires only the initial displacement at the vertices $(x-ct,0)$, $(x+ct,0)$; the second term requires only the initial velocity at the edge which joins those two vertices; while the final term requires the value of the force in the entire triangular region.

Similarly, initial position at the point $(x_0,0)$ influences the solution only along the rays $x - ct = x_0$ and $x + ct = x_0$ (for $t > 0$), whereas the initial velocity and forcing at the same point influence only the region between those two rays, called *the domain of influence.* In other words, the effects of initial displacement propagate at the speed c, while those of initial velocity and forcing persist, but none of the effects propagate faster than the speed c. Similar observations hold in two dimensions. However, in the three-dimensional wave equation, the effects of initial velocity and forcing do not persist, making the causality principle is sharper. This sharp form is called *Huygens's principle*.

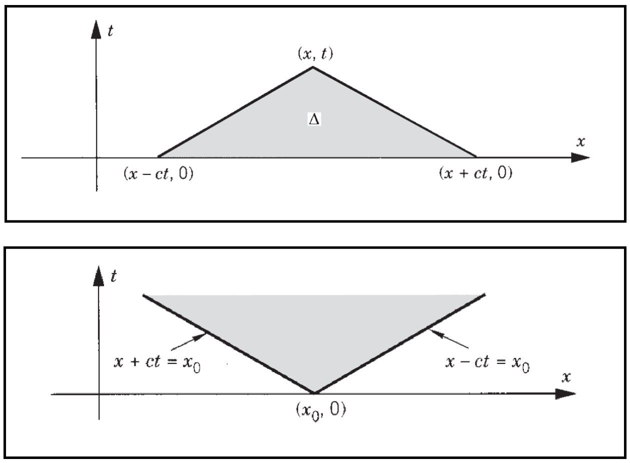


Figure 9: Domain of dependence (above), and domain of influence (below).

A legitimate question to ask is if in one dimension - a concrete example is the propagation of sound along a railway track - each noise that a train makes would automatically mix with the "echoes" of the noise created by trains which passed before it, and the answer is no. Namely, a localized impulse does keep the string permanently deformed, but it doesn't keep it permanently vibrating. After the time at which the initial localized impulse traveling at speed c surpasses the point of the observation x, d'Alambert's solution at x remains constant in time, equal to the integral of the entire impulse, scaled by $1/2c$ (see the Exercise 3.4).

3.5 Exercises

Exercise 3.3. Find the week solutions to the initial value problem

$$\partial_x^2 u - \partial_x^2 = 0, \;\; u(x,0) = \begin{cases} 2x+1 & \text{for } -\frac{1}{2} < x \leq 0 \\ -2x+1 & \text{for } 0 < x < \frac{1}{2} \\ 0 & \text{otherwise} \end{cases}, \;\; \partial_t u(x,0) = 0.$$

in successive times $t_0 = 0, \frac{1}{8}, \frac{1}{4}, \frac{1}{2}, 1$, and sketch the graphs $u(x, t_0)$.

Exercise 3.4. The point $x = 0$ of an infinite string of tension T and density ρ (recall that $c^2 = \frac{T}{\rho}$) is hit by a hammer whose head diameter is $2a$, so that initial conditions are

$$u(x,0) = 0, \;\; \partial_t u(x,0) = \begin{cases} 1 & \text{if } -a < x < a \\ 0 & \text{otherwise.} \end{cases}$$

A flea is sitting at a distance l from the point $x = 0$. (Assume that $a < l$; otherwise, poor flea!) How long does it take for the disturbance to reach the flea? At what time will the point where the flea is sitting be still again?

Exercise 3.5. Solve the initial value problem

$$\partial_t^2 u - 3\partial_x\partial_t u - 4\partial_x^2 = \sin(x+t),\ \ u(x,0) = \partial_t u(x,0) = 0.$$

Exercise 3.6. Solve the equation

$$\partial_t^2 u - c^2\partial_x^2 u = \sin(\omega t)\sin(kx),\ \ u(x,0) = \partial_t u(x,0) = 0$$

for both $kc \neq \omega$, and $kc = \omega$,

Exercise 3.7. Prove the following facts:

1. If initial conditions for the wave equation $u(x,0) = f(x)$, and $\partial_t u(x,0) = g(x)$ are both even functions ($f(x) = f(-x), g(x) = g(-x)$), at any time $t \in \mathbb{R}$, d'Alambert solution $u(x,t)$ is also an even function in the x-variable.

2. If initial conditions for the wave equation $u(x,0) = f(x)$, and $\partial_t u(x,0) = g(x)$ are both odd functions ($f(x) = -f(-x), g(x) = -g(-x)$), at any time $t \in \mathbb{R}$, d'Alambert solution $u(x,t)$ is also an odd function in the x-variable.

3. If initial conditions for the wave equation $u(x,0) = f(x)$, and $\partial_t u(x,0) = g(x)$ are both periodic functions of period l ($f(x) = f(x+l), g(x) = g(x+l)$), at any time $t \in \mathbb{R}$, d'Alambert solution $u(x,t)$ is also a periodic function of period l in the x-variable.

4 Diffusion or heat equation

Imagine a straight horizontal tube filled with motionless liquid. Our aim is to deduce a differential equation which governs diffusion (movement) of a chemical substance – for instance dye – through the liquid-filled tube. Transfer of heat through a motionless liquid is governed by the same differential equation, hence the name heat equation.

Denote by $u(x,t)$ the concentration of dye in terms of the amount of substance per meter, at the distance x from the first end of the tube, at the time t. By Flick's law, flux of the substance in a motionless fluid goes from regions of high concentration to regions of low concentration, with a magnitude that is proportional to the concentration gradient. In one dimension (inside the tube) we have

$$j(x,t) = -k\partial_x u(x,t),$$

where k is the diffusion coefficient. Flick's law, together with the continuity equation

$$\partial_x j(x,t) = -\partial_t u(x,t),$$

yields the diffusion equation

$$\partial_t u(x,t) - k\partial_x^2 u(x,t) = 0.$$

Initial value problem for the diffusion equation inside an infinite tube is given by

$$\begin{array}{l} \partial_t u(x,t) - k\partial_x^2 u(x,t) = 0, \quad -\infty < x < \infty, \quad t > 0, \\ \lim_{t\to 0} u(x,t) = f(x). \end{array} \tag{18}$$

Observe a fundamental difference between (18) and initial value problem for either wave equation, or first order PDE-s. In first order PDE-s, a local solution to the initial value problem is specified in an open neighborhood of the x-axis, which includes negative time. Initial value problem for wave equation has a globally defined solution, again in both positive and negative time. In contrast to this, we aim to solve the diffusion equation only in strictly positive time, assigning the initial condition at the limit $t \to 0$. This is not a mere computational convenience. Instead, it is rooted in the physical fact that diffusion is an irreversible process. Although for a certain class of initial conditions, it is possible from the mathematical point of view to determine solutions to the diffusion equation in negative time, such solutions are ill-behaved, in the sense that they are numerically unstable under small perturbations of the initial condition, and can have singularities even if the initial condition is a smooth function.

4.1 Fundamental solution of the diffusion equation

In this section, we study diffusion of the unit amount (1 mol) of the substance (dye) initially placed at one single point $x = 0$. The corresponding initial condition is given by the Dirac delta function: concentration of the dye is zero away from the origin and infinite at the origin, whereas the total amount of dye is one:

$$\delta(x) = \begin{cases} \infty & \text{for } x = 0, \text{ and} \\ 0 & \text{for } x \neq 0 \end{cases}; \quad \int_{-\infty}^{\infty} \delta(x)dx = 1.$$

The solution to the initial value problem in hands

$$\begin{aligned} &\partial_t S(x,t) - k\partial_x^2 S(x,t) = 0, \quad -\infty < x < \infty, \quad t > 0, \\ &\lim_{t\to 0} S(x,t) = \delta(x). \end{aligned} \tag{19}$$

is called *the fundamental solution* to the diffusion equation. As we will see in the next section, solution of the general initial value problem (18) is deduced from the fundamental solution.

On the face of it, the initial value problem (19) looks hopeless, as Dirac delta function is not a function at all (in fact, it is a distribution, see Appendix A). However, we can use to our advantage the fact that its "antiderivative" is a function:

$$H(x) = \int_{-\infty}^{x} \delta(x')dx' = \begin{cases} 0 & \text{for } x < 0, \text{ and} \\ 1 & \text{for } x > 0. \end{cases}$$

Observe that the partial derivative $\partial_x u(x,t)$ of a solution $u(x,t)$ to the diffusion equation is itself a solution:

$$\partial_t(\partial_x u)(x,t) - k\partial_x^2(\partial_x u)(x,t) = \partial_x(\partial_t u(x,t) - k\partial_x^2 u(x,t)) = 0.$$

Hence, if $Q(x,t)$ solves the initial value problem

$$\begin{aligned} &\partial_t Q(x,t) - k\partial_x^2 Q(x,t) = 0, \quad -\infty < x < \infty, \quad t > 0, \\ &\lim_{t\to 0} Q(x,t) = H(x). \end{aligned} \tag{20}$$

one expects the function

$$S(x,t) := \partial_x Q(x,t)$$

to solve the initial value problem (19). Again, the reasoning can be formalized in the language of distributional derivatives, but we will not go deep into this. Instead, the Corollary A1.3 of the Theorem 4.2 directly establishes the convergence $\lim_{t\to 0} S(x,t) = \delta(x)$ in the space of distributions.

Let's now solve (20)! The strategy is to eliminate one variable (and hence reduce to an ordinary differential equation) by the use of symmetry: a parametrized family of transformations $(x,t) \mapsto (\chi(x,t), \tau(x,t))$ such that $Q(\chi(x,t), \tau(x,t))$ solves a given initial value problem as long as $Q(x,t)$ does. As a physical process is uniquely determined by the initial condition, the (unique) *physical* solution must be invariant under symmetry:

$$Q(x,t) = Q(\chi(x,t), \tau(x,t)).$$

In the initial value problem (20), it is reasonable to search for a symmetry of type $(x,t) \mapsto (\lambda x, \mu t)$, where $\lambda, \mu > 0$ are real parameters. Indeed, let Q be a solution to 20, and set $\chi(x) = \lambda x$, $\tau(t) = \mu t$. By the chain rule

$$\begin{aligned}
&\partial_x Q(\lambda x, \mu t) = \partial_x \chi(x) \partial_\chi Q(\chi, \tau) = \lambda \partial_\chi Q(\chi, \tau),\\
&\partial_x^2 Q(\lambda x, \mu t) = \lambda \partial_x \partial_\chi Q(\chi, \tau) = \lambda \partial_x \chi(x) \partial_\chi^2 Q(\chi, \tau) = \lambda^2 \partial_\chi^2 Q(\chi, \tau), \text{ and}\\
&\partial_t Q(\lambda x, \mu t) = \partial_t \tau(t) \partial_\tau Q(\chi, \tau) = \mu \partial_\tau Q(\chi, \tau).
\end{aligned}$$

As Q is a solution to the diffusion equation,

$$0 = \partial_\tau Q(\chi, \tau) - k \partial_\chi^2 Q(\chi, \tau) = \frac{1}{\mu} \partial_t Q(\lambda x, \mu t) - (\frac{1}{\lambda})^2 k \partial_x^2 Q(\lambda x, \mu t).$$

For $\mu = \sqrt{\lambda}$,

$$\partial_t Q(\lambda x, \sqrt{\lambda} t) - k \partial_x^2 Q(\lambda x, \sqrt{\lambda} t) = \lambda^2 (\partial_\tau Q(\chi, \tau) - k \partial_\chi^2 Q(\chi, \tau)) = 0,$$

so $Q(\lambda x, \sqrt{\lambda} t)$ satisfies the diffusion equation. As

$$\lim_{t \to 0^+} Q(\lambda x, \sqrt{\lambda} t) = H(\lambda x) = H(x),$$

$Q(\lambda x, \sqrt{\lambda} t)$ is a solution of (20). Hence, symmetry of our problem is the λ-parametrized family of transformations $(x,t) \mapsto (\lambda x, \sqrt{\lambda} t)$.

Denote $g(\xi) = Q(\xi, 1)$. For any fixed $x \in \mathbb{R}$ and $t, \lambda > 0$, $Q(x,t) = Q(\lambda x, \sqrt{\lambda} t)$. In particular, for $\lambda = 1/\sqrt{t}$,

$$Q(x,t) = Q(\frac{x}{\sqrt{t}}, 1) = g(\frac{x}{\sqrt{t}}).$$

As the above equality holds for all $x \in \mathbb{R}$ and $t > 0$, it is an equality of functions. By the chain rule,

$$\partial_t Q(x,t) = -\frac{x}{2t\sqrt{t}} g'(\frac{x}{\sqrt{t}}), \quad \partial_x Q(x,t) = \frac{1}{\sqrt{t}} g'(\frac{x}{\sqrt{t}}), \quad \partial_x^2 Q(x,t) = \frac{1}{t} g'(\frac{x}{\sqrt{t}}).$$

The diffusion equation now reads

$$\frac{x}{2t\sqrt{t}} g'(\frac{x}{\sqrt{t}}) + k \frac{1}{t} g''(\frac{x}{\sqrt{t}}) = 0.$$

Multiplying by t, and substituting $\xi = x/\sqrt{t}$, $g' = h$, we get

$$\frac{1}{2} \xi h(\xi) + k h'(\xi) = 0.$$

By the separation of variables

$$\frac{1}{2}\int \xi d\xi = -k\int \frac{dh}{h},$$

and hence

$$h(\xi) = C_0 e^{-\xi^2/4k}.$$

Since $g' = h$,

$$g(\xi) = C_0\int_0^{\xi} e^{-r^2/4k}dr + D = [s = r/\sqrt{4k},\, ds = dr/\sqrt{4k}] = \sqrt{4k}C_0\int_0^{\frac{\xi}{\sqrt{4k}}} e^{-s^2}ds + D.$$

For $C = \sqrt{4k}C_0$,

$$Q(x,t) = g(\frac{x}{\sqrt{t}}) = C\int_0^{\frac{x}{\sqrt{4kt}}} e^{-s^2}ds + D.$$

Evaluating at the initial condition, for $x > 0$,

$$\lim_{t\to 0^+} Q(x,t) = C\int_0^{\infty} e^{-s^2}ds + D = \frac{C\sqrt{\pi}}{2} + D = 1,$$

and for $x < 0$,

$$\lim_{t\to 0^+} Q(x,t) = C\int_0^{-\infty} e^{-s^2}ds + D = -C\int_{-\infty}^{0} e^{-s^2}ds + D = -\frac{C\sqrt{\pi}}{2} + D = 0.$$

From here, $C = \frac{1}{\sqrt{\pi}}$, and $D = \frac{1}{2}$.

Finally, we calculate the fundamental solution $S(x,t) = \partial_x H(x,t)$. By the Leibniz integral rule[1],

$$S(x,t) = \partial_x Q(x,t) = \partial_x(\frac{1}{\sqrt{\pi}}\int_0^{\frac{x}{\sqrt{4kt}}} e^{-s^2}ds + \frac{1}{2}) = \frac{1}{\sqrt{\pi}}e^{-\frac{x^2}{4kt}}\partial_x\frac{x}{\sqrt{4kt}} = \frac{1}{\sqrt{4k\pi t}}e^{-\frac{x^2}{4kt}}.$$

There are a few observations to be made about the solution:

- Even though the initial condition is singular, the solution in positive time is a smooth (or C^∞) function, meaning that its partial derivatives of any order are well-defined. Physically, diffusion "smooths out" the initial distribution, making it more uniform as particles spread throughout the space.

[1] Given a continuous function $f(x,t)$ which is differentiable in x, and C^1 functions $a(x)$ and $b(x)$,

$$\partial_x\int_{a(x)}^{b(x)} f(x,t)dt = f(x,b(x))b'(x) - f(x,a(x))a'(x) + \int_{a(x)}^{b(x)} \partial_x f(x,t)dt$$

- Although the initial condition is non-zero only at the point $x = 0$, in all the positive times, it is non-zero everywhere. This means that in the model, speed at which substance diffuses is not bounded by the speed of light. We say that, in contrast with the wave equation, diffusion equation is non-relativistic.

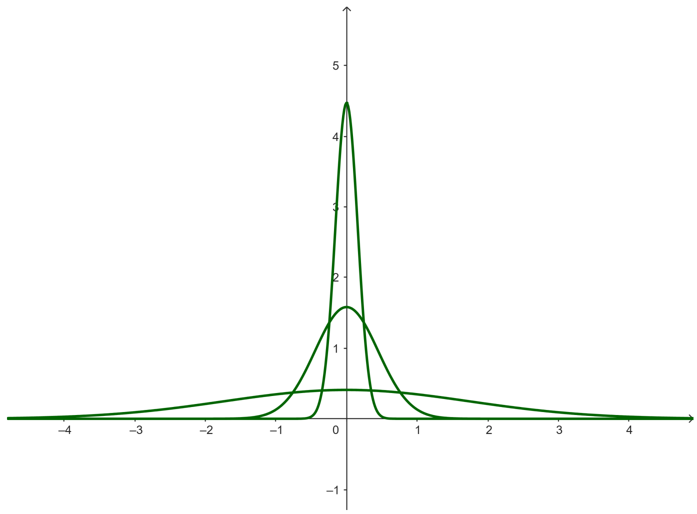


Figure 10: Solution $S(x,t)$ for different t.

4.2 Solution to the general initial value problem, and non homogeneous diffusion equation

To determine the solution of 18, we first approximate the initial continuous distribution $\lim_{t\to 0} u(x,t) = f(x)$ by a discrete one, where at the points x_n which are distributed along the x-axis at the distance Δx we place $f(x_n)\Delta x$ amount of substance. The initial value problem in hands is

$$\partial_t u(x,t) - k\partial_x^2 u(x,t) = 0, \quad -\infty < x < \infty, \quad t > 0,$$
$$\lim_{t\to 0} u(x,t) = \sum_n \delta(x - x_n) f(x_n)\Delta x.$$

Lemma 4.1 below implies that the solution is given by

$$u(x,t) = \sum_n S(x - x_n, t) f(x_n)\Delta x.$$

Physically, the substance initially placed at each point x_n spreads out in time and contributes to the distribution $u(x,t)$ by the term $S(x - x_n, t)f(x_n)\Delta x$. In the limit $\Delta x \to 0$, we get

$$u(x,t) = \int_{-\infty}^{\infty} S(x - y, t) f(y) dy. \tag{21}$$

Theorem 4.2 below provides a rigorous mathematical proof that the solution (21) which we derived using physical arguments is correct as long as the function f sastisfies certain mild conditions.

Lemma 4.1. 1. The set of solutions to the diffusion equation is a vector space. That is, if $u_n(x,t)$ are solutions, and a_n real numbers, then $u(x,t) = \sum_n u_n(x,t)a_n$ is a solution.

2. If $u(x,t)$ is a solution to the diffusion equation, then $u(x-y,t)$ is also a solution for any $y \in \mathbb{R}$.

Proof. Part 1 is a simple calculation:

$$\partial_t(\sum_n u_n(x,t)a_n) - k\partial_x^2(\sum_n u_n(x,t)a_n) = \sum_n(\partial_t u_n(x,t) - k\partial_x^2 u_n(x,t))a_n = 0.$$

For part 2, denote $\chi = x - y$. We have

$$\partial_x u(x-y,t) = \partial_x\chi\partial_\chi u(\chi,t) = \partial_\chi u(\chi,t), \text{ and}$$
$$\partial_x^2 u(x-y,t) = \partial_x\partial_\chi u(\chi,t) = \partial_\chi^2 u(\chi,t).$$

Since $u(x,t)$ satisfies the diffusion equation,

$$\partial_t u(x-y,t) - k\partial_x^2 u(x-y,t) = \partial_t u(\chi,t) - k\partial_\chi^2 u(\chi,t) = 0. \tag{22}$$

□

A similar strategy can be used to determine the solution of the initial-value problem for the non-homogeneous diffusion equation

$$\partial_t u(x,t) - k\partial_x^2 u(x,t) = F(x,t), \quad -\infty < x < \infty, \quad t > 0,$$
$$\lim_{t\to 0} u(x,t) = 0.$$

Physically, the non-homogeneous diffusion equation describes diffusion in the presence of a source. Again, we can approximate the continuous source with a discrete one. At points x_n (distributed along the x-axis at the distance Δx) at times t_m (distributed on the t-axis at the distance Δt), we place $F(x_n,t_m)\Delta x\Delta t$ amount of substance. The amount of substance placed at a point (x_n,t_m) diffuses in time $t > t_m$, and contributes to the distribution $u(x,t)$ by the term $S(x-x_n,t-t_m)F(x_n,t_m)\Delta x\Delta t$. Summing over all the contributions, and taking the limit $\Delta x, \Delta t \to 0$ we get

$$u(x,t) = \int_0^t \int_{-\infty}^{\infty} S(x-y,t-s)F(y,s)dyds.$$

A mathematical proof that the obtained solution is correct is rather simple, and it is the content of exercise 4.5.

Finally, just as with wave equation, the solution to the general intial-value problem

$$\begin{aligned} &\partial_t u(x,t) - k\partial_x^2 u(x,t) = F(x,t), \quad -\infty < x < \infty, \quad t > 0, \\ &\lim_{t\to 0} u(x,t) = f(x) \end{aligned}$$

is the sum of the solution of the non-homoegenous equation with the trivial initial condition, and the homogeneous equation with the given initial condition

$$u(x,t) = \int_{-\infty}^{\infty} S(x-y,t)f(y)dy + \int_0^t \int_{-\infty}^{\infty} S(x-y,t-s)F(y,s)dyds.$$

Theorem 4.2. *Given a function $f : \mathbb{R} \to \mathbb{R}$ which is bounded (there exists $M > 0$ s.t $|f(x)| < M$ for all $x \in \mathbb{R}$) and compactly supported (there exists $N > 0$ s.t. $f(x) = 0$ if $|x| > N$), formula (21) defines an infinitely derivable (smooth) function, which is a solution of the initial value problem (18).*

Remark 4.3. *For the theorem to hold it suffices for the function f to be bounded. However, the proof in this general case requires the use of certain tools from analysis such as uniform convergence, which we prefer to avoid, in order to keep the exposition as simple as possible. Be it mentioned, by the Weierstrass theorem from Analysis 1, continuous compactly supported functions are automatically bounded, hence the latter condition is obsolete in the (weaker) theorem that we intend prove.*

Remark 4.4. *Interestingly, the solution (21) is not the unique solution to the diffusion equation. However, it is the unique solution for which there exist constants $C, D > 0$ such that $|u(x,t)| < Ce^{Dx^2}$. This result, called Tychonoff's uniqueness theorem, is outside the scope Analysis 4b. Clearly, the non-unique highly divergent solutions are non-physical.*

Proof. We begin by proving that $u(x,y)$ is a well-defined smooth function. The integral (21) is convergent as $f(x)$ is compactly supported. Indeed,

$$u(x,t) = \int_{-\infty}^{\infty} S(x-y,t)f(y)dy = \int_{-N}^{N} S(x-y,t)f(y)dy$$

is well-defined for all $x \in \mathbb{R}$, $t > 0$. By the Leibniz integral rule, the same holds for partial derivatives of first order

$$\partial_x u(x,t) = \int_{-N}^{N} \partial_x S(x-y,t)f(y)dy; \quad \partial_t u(x,t) = \int_{-N}^{N} \partial_t S(x-y,t)f(y)dy.$$

By induction, higher partial derivatives

$$\partial_x^k \partial_t^l u(x,t) = \int_{-N}^{N} \partial_x^k \partial_t^l S(x-y,t) f(y) dy$$

are also well-defined, and $u(x,t)$ is a smooth function.

Next, we prove that $u(x,t)$ satisfies the diffusion equation. By Lemma 4.1, $S(x-y,t)$ satisfies the diffusion equation for every $y \in \mathbb{R}$, hence

$$k\partial_t u(x,t) - k\partial_x^2 u(x,t) = \int_{-N}^{N} (k\partial_t S(x-y,t) - k\partial_x^2 S(x-y,t)) f(y) dy = 0.$$

It remains to verify that $\lim_{t\to 0} u(x,t) = f(x)$. The first observation to be made is that $\int_{-\infty}^{\infty} S(x-y,t)dy = 1$. Indeed,

$$\begin{aligned}\int_{-\infty}^{\infty} S(x-y,t)dy &= \frac{1}{\sqrt{4k\pi t}} \int_{-\infty}^{\infty} e^{-\frac{(x-y)^2}{4kt}} dy = [\chi = \frac{y-x}{\sqrt{4kt}}, d\chi = \frac{dy}{\sqrt{4kt}}] \\ &= \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-\chi^2} d\chi = 1.\end{aligned}$$

Consequently,

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

By the triangle inequality for integrals,

$$|u(x,t) - f(x)| = |\int_{-\infty}^{\infty} S(x-y,t)(f(y) - f(x))dy| \le \int_{-\infty}^{\infty} S(x-y,t)|f(y) - f(x)|dy.$$

Proving that $\lim_{t\to 0} u(x,t) = f(x)$ amounts to showing that for any given $\epsilon > 0$, and small enough $t > 0$, we get $|u(x,t) - f(x)| < \epsilon$.

Fix $\epsilon > 0$. As f is continuous, there exists $\delta > 0$ such that $|f(y) - f(x)| < \epsilon/2$ for all $y \in (x-\delta, x+\delta)$. We have

$$\begin{aligned}|u(x,t) - f(x)| &\le \int_{-\infty}^{\infty} S(x-y,t)|f(y) - f(x)|dy \\ &= \int_{x-\delta}^{x+\delta} S(x-y,t)|f(y) - f(x)|dy + \int_{\mathbb{R}\setminus(x-\delta,x+\delta)} S(x-y,t)|f(y) - f(x)|dy \\ &< \frac{\epsilon}{2} \int_{x-\delta}^{x+\delta} S(x-y,t)dy + 2M \int_{\mathbb{R}\setminus(x-\delta,x+\delta)} S(x-y,t)dy.\end{aligned}$$

As

$$\frac{\epsilon}{2} \int_{x-\delta}^{x+\delta} S(x-y,t)dy < \frac{\epsilon}{2} \int_{-\infty}^{\infty} S(x-y,t)dy = \frac{\epsilon}{2},$$

it remains to prove that for small enough $t > 0$,

$$2M \int_{\mathbb{R}\setminus(x-\delta,x+\delta)} S(x-y,t)dy < \frac{\epsilon}{2}.$$

The latter condition is satisfied as long as

$$\int_{\mathbb{R}\setminus(x-\delta,x+\delta)} S(x-y,t)dy \xrightarrow{t\to 0} 0.$$

We have

$$\begin{aligned}\int_{\mathbb{R}\setminus(x-\delta,x+\delta)} S(x-y,t)dy &= \frac{1}{\sqrt{4k\pi t}} \int_{\mathbb{R}\setminus(x-\delta,x+\delta)} e^{-\frac{(x-y)^2}{4kt}} dy = [\chi = \frac{y-x}{\sqrt{4kt}}, d\chi = \frac{dy}{\sqrt{4kt}}] \\ &= \frac{1}{\sqrt{\pi}} \int_{\mathbb{R}\setminus(\frac{-\delta}{\sqrt{4kt}},\frac{\delta}{\sqrt{4kt}})} e^{-\chi^2} d\chi \xrightarrow{t\to 0} 0,\end{aligned}$$

as $\lim_{t\to 0}(\pm\delta/\sqrt{4kt}) = \pm\infty$, and the integral $\int_{-\infty}^{\infty} e^{-\chi^2} d\chi$ is convergent (the integral over the "tail" $|\xi| > A$ of the Gaussian distribution is arbitrarily small for sufficiently large A). □

Exercise 4.5. Let $F(x,t)$ be a bounded continuous function such that there exists $M > 0$ with $f(x,t) = 0$ if $|x| > M$. Using the Leibniz integral rule and Theorem 4.2, show that the formula

$$u(x,t) = \int_0^t \int_{-\infty}^{\infty} S(x-y,t-s)F(y,s)dyds$$

defines a $C^{\infty,1}$ function (smooth in the x variable, and C^1 in the t variable) which is a solution to the initial valuse problem

$$\begin{aligned}&\partial_t u(x,t) - k\partial_x^2 u(x,t) = F(x,t), \quad -\infty < x < \infty, \quad t > 0, \\ &\lim_{t\to 0} u(x,t) = 0.\end{aligned}$$

4.3 Maximum and minimum principle

Assume that a function $u(x,t)$ solves the homogeneous diffusion equation in the rectangle $0 \le x \le l$, $0 \le t \le T$. Physically, $u(x,t)$ describes diffusion in positive time within the segment $[0,l]$, and in absence of a source. As substance diffuses from the areas of higher concentration towards the areas of lower concentration, a local maximum of $u(x,t)$ can occur either at the edges of the segment (in the presence of a source outside the segment $[0,l]$, form which the substance diffuses through the edges), or initially at $t = 0$. This property of the diffusion equation is known as the maximum principle. Analogous minimum principle is also satisfied.

Theorem 4.6. *Given $u(x,t)$ as above, its maximal and minimal values are assumed either initially ($t=0$), or on the lateral sides ($x=0$ or $x=l$).*

Proof. Observe that it suffices to prove the maximum principle, as the minimum principle of $u(x,t)$ is equivalent to the maximum principle of $-u(x,t)$.

Let M be the maximal value of $u(x,t)$ on the three sides ($t=0$, $x=0$, and $x=l$) of the rectangle $[0,l]\times[0,T]$. We need to show that $u(x,t)\le M$ in the entire rectangle. For a fixed $\epsilon>0$, define

$$u_\epsilon(x,t)=u(x,t)+\epsilon x^2.$$

On the three sides of the rectangle,

$$u_\epsilon(x,t)\le M+\epsilon l^2.$$

If we can prove that this inequality holds on the entire rectangle, it would follow that on the entire rectangle

$$u(x,t)\le u_\epsilon(x,t)\le M+\epsilon l^2$$

for arbitrarily small $\epsilon>0$, which is possible only if $u(x,t)\le M$.

Assume the opposite, namely that there is a point (x_0,t_0) either in the interior of the rectangle $(x_0,t_0)\in(0,l)\times(0,T)$ or on the upper side of the rectangle away from the edges ($t_0=T$, $x_0\neq 0$, $x_0\neq l$), where u_ϵ attains a local maximum. For any $(x,t)\in[0,l]\times[0,T]$ of the rectangle,

$$\partial_t u_\epsilon(x,t)-k\partial_x^2 u_\epsilon(x,t)=\partial_t u(x,t)-k\partial_x^2 u(x,t)-2k\epsilon=-2k\epsilon<0.$$

If (x_0,t_0) is in the interior, t_0 must be a local maximum of the one-variable function $t\mapsto u_\epsilon(x_0,t)$, hence its first derivative at t_0, $\partial_t u_\epsilon(x_0,t_0)$ must vanish; and x_0 must be a local maximum of the one-variable function $x\mapsto u_\epsilon(x,t_0)$, hence its second derivative at x_0, $\partial_x^2 u_\epsilon(x_0,t_0)$ must be non-positive. From here,

$$\partial_t u_\epsilon(x_0,t_0)-k\partial_x^2 u_\epsilon(x_0,t_0)>0$$

which is a contradiction.

If $(x_0,t_0)=(x_0,T)$ is on the upper side of the rectangle away from the edges, the one-variable function $t\mapsto u_\epsilon(x_0,t)$ must be non-decreasing at T, hence $\partial_t u_\epsilon(x_0,T)\ge 0$, and the one-variable function $x\mapsto u_\epsilon(x,T)$, must have a local maximum at x_0, hence $\partial_x^2 u_\epsilon(x_0,T)\le 0$. We again conclude that

$$\partial_t u_\epsilon(x_0,T)-k\partial_x^2 u_\epsilon(x_0,T)>0,$$

which is a contradiction. □

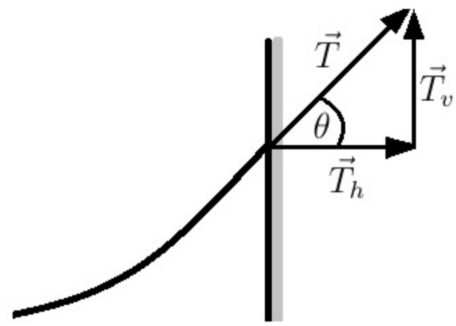


Figure 11

5 Boundary problems

In the final chapter of the course, we deal with heat and wave equations on a finite interval $0 \leq x \leq l$, subjected to constraints on the boundary ($x = 0$, $x = l$), called the *boundary conditions*. We begin by a review of the most common boundary conditions together with their physical interpretations. Let $u(x,t)$ be a solution to either wave equation or heat equation on the interval $0 \leq x \leq l$.

1. Solution $u(x,t)$ satisfies the *Dirichlet boundary condition* if at all times

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

   In the case of heat transfer on a rod of length l, it is realized when its end-points are kept at constant temperature. For transverse vibrations of a taut string (governed by the wave equation), Dirichlet boundary condition is realized when the end-points are fixed; example is a guitar string. For longitudinal vibrations of air pressure in a tube (also governed by the wave equation), it is realized when both ends of a tube are open, in which case the pressure at the ends is fixed to the atmospheric pressure; example is a flute.

2. Solution $u(x,t)$ satisfies the *Neumann boundary condition* if at all times

$$\partial_x u(0,t) = \partial_x u(l,t) = 0.$$

   In the case of heat transfer on a rod, it is realized when the endpoints are isolated. Indeed, by the Fourier's law, heat flux is proportional to the negative gradient of the temperature, that is to $-\partial_x u$ in one dimensional case. If the end points are isolated, there is no flux through them, hence $\partial_x u(0,t) = \partial_x u(l,t) = 0$.

For transverse vibrations of a taut string, $\partial_x u(l,t) = \operatorname{tg}\theta$, where θ is the angle which the right end ($x = l$) of the vibrating string closes with the horizontal axis (see Figure 11). Denoting by $\vec{T}$ the force by which the right wall acts on the string, and by T_h and T_v its vertical and horizontal components, we get $\operatorname{tg}\theta = \frac{T_v}{T_h}$. As vibrations are purely transverse, T_h is constant, hence $\partial_x u(l,t)$ is proportional to T_v. Neumann boundary condition is hence satisfied when there is no vertical force acting on the string at the endpoints. We speak of free endpoints, which can be physically realized by attaching both ends of the string to a massless ring which is free to slide on a frictionless poles situated at $x = 0$ and $x = l$.

For longitudinal vibrations of air pressure in a tube, Neumann boundary condition models the tube whose endpoints are closed. Namely, by the Bernoulli's principle flux of air is proportional to the negative gradient of the air pressure $-\partial_x u$. If the ends of a tube are closed, air flux through them is zero, hence $\partial_x u(0,t) = \partial_x u(l,t) = 0$.

3. Solution $u(x,t)$ satisfies the *Robin boundary condition* if at all times

$$\partial_x u(0,t) - a_0 u(0,t) = \partial_x u(l,t) + a_l u(l,t) = 0 \quad a_0, a_l \in \mathbb{R}.$$

Let us return to the above examples Dirichlet boundary condition. How can one assure that the ends of a rod are kept at a constant temperature, and to which extent is this even possible? The best that I can do (maybe you can do better as young physicists) is to keep them in contact with a liquid of constant temperature T. In this situation, we get transfer of heat due to convection, so that the flux from the end-points is proportional to the negative difference between the temperature of the liquid and of the rod. For $x = l$, we have $J(l,t) = -c(T - u(x,l))$ where c is the convection coefficient. Applying Fourier's law, we arrive to the Robin boundary condition

$$\partial_x u(l,t) = -J(l,t) = -c(u(x,l) - T_0).$$

For transverse vibrations of a taut string with fixed end-points, one can ask what happens if the end-points are not perfectly rigid. In this case, there is a harmonic vertical force acting on the endpoints of the vibrating string $T_v(l,t) = -ku(l,t)$. $T_v(l,t)$ being proportional to $\partial_x u(l,t)$, we again arrive to Robin boundary condition.

4. Solution $u(x,t)$ satisfies the *circular boundary condition* if at all times

$$u(0,t) = u(l,t), \text{ and } \partial_x u(0,t) = \partial_x u(l,t).$$

   As the name suggest, this boundary condition is used to model vibrations or heat transfer on a circular domain, such a ring of circumference l. Indeed, in this case, $x = 0$ and $x = l$ determine the same point.

5. When different boundary conditions are satisfied on the two ends of the interval $0 \le x \le l$, we speak *mixed boundary conditions*. Common combination is Neumann boundary condition on one end, and the Dirichlet boundary condition on the other. There is a number of wind instruments that obey this mixed boundary condition. For example, in the clarinet, mouthpiece is a closed end, and bell is an open end. In some traditional instruments such as Irish uilleann pipes and Slovak koncovka, the foot end of the instrument is kept open for some notes, and closed for other.

5.1 Separation of variables

Our next task is to find general solutions to the boundary value problems for both wave and heat equation. Here, we will focus on Dirichlet and Neumann boundary conditions. Other boundary conditions are dealt with in the exercises section.

Dirichlet boundary condition

We start with the wave equation

$$\begin{aligned} &\partial_t^2 u - c^2 \partial_x^2 u = 0, \ \ 0 \le x \le l \\ &u(0,t) = u(l,t) = 0. \end{aligned} \tag{23}$$

The strategy is to first determine those solutions which are products of a function which depends only on x and a function which depends only on t

$$u(x,t) = X(x)T(t).$$

Such solutions are called *separated solutions*. It turns out that every solution is a superposition of separated solutions.

Substituting $u(x,t) = X(x)T(t)$ into the wave equation, we get $XT'' - c^2 X'' T = 0$. Dividing by $c^2 XT$,

$$\frac{X''}{X} = \frac{T''}{c^2 T} = \lambda.$$

This defines a quantity λ which doesn't depend on t as the first expression doesn't, and doesn't depend on x as the second expression doesn't. Hence, it is a constant. As $X(0)T(t) = u(0,t) = 0$, and $X(l)T(t) = u(l,t) = 0$, either $X(0) = X(l) = 0$, or T is identically zero. In the latter case, $u(x,t)$ is also identically zero. A *nontrivial* separated solution is hence determined by the system of ODEs

$$X" = \lambda X, \;\; X(0) = X(l) = 0; \tag{24a}$$
$$T'' = c^2 \lambda T. \tag{24b}$$

For $\lambda = 0$, $X(x) = Ax + B$. Evaluating at the boundaries, we get $A = B = 0$, hence u is again a trivial solution.

For $\lambda \neq 0$,

$$X(x) = Ae^{\sqrt{\lambda}x} + Be^{-\sqrt{\lambda}x}.$$

Evaluating at the boundary $x = 0$, we get $A + B = 0$, and evaluating at $x = l$,

$$A(e^{\sqrt{\lambda}l} - e^{-\sqrt{\lambda}l}) = 0.$$

Multiplying by $e^{\sqrt{\lambda}l}/A$, we get

$$e^{2\sqrt{\lambda}l} = 1,$$

which is satisfied if and only if $2\sqrt{\lambda}l = 2n\pi i$ for an integer n, that is, if $\lambda = -(\frac{n\pi}{l})^2$. As for any $z \in \mathbb{C}$, $(e^{iz} - e^{-iz})/2i = \sin(z)$, renaming the constant ($A' = 2Ai$), we get for each n a solution to the Equation (24a)

$$X_n(x) = A' \sin(\frac{n\pi}{l}x).$$

For $\lambda = -(\frac{n\pi}{l})^2$, general solution of the Equation (24b) is

$$T_n(t) = A \sin(\frac{nc\pi}{l}t) + B \cos(\frac{nc\pi}{l}t).$$

Renaming the constants as $A_n = A'A$, and $B_n = A'B$,

$$u_n(x,t) := X_n(x)T_n(t) = (A_n \sin(\frac{nc\pi}{l}t) + B_n \cos(\frac{nc\pi}{l}t)) \sin(\frac{n\pi}{l}x).$$

Clearly, solutions of 26, and the same is true for the other boundary value problems listed above, form a vector space. In particular, superposition (linear combination) of finitely many separated solutions $u_n(x,t)$ is again a solution. However, this is not enough to deal with general initial value problems. Instead, we must consider superpositions of infinitely many separated solutions, that is the convergent function series

$$u(x,t) = \sum_{n=1}^{\infty} (A_n \sin(\frac{nc\pi}{l}t) + B_n \cos(\frac{nc\pi}{l}t)) \sin(\frac{n\pi}{l}x). \tag{25}$$

Next we solve the heat equation for Dirichlet boundary condition

$$\begin{aligned} &\partial_t u - k\partial_x^2 u = 0, \;\; 0 \le x \le l \\ &u(0,t) = u(l,t) = 0, \end{aligned} \tag{26}$$

also using separation of variables. Assume that

$$u(x,t) = X(x)T(t)$$

is a separated solution. Substituting in the heat equation, we get $XT' - kX''T = 0$. Dividing by kXT,

$$\frac{X''}{X} = \frac{T'}{kT} = \lambda.$$

A *nontrivial* separated solution is hence determined by the system of ODEs

$$X" = \lambda X, \;\; X(0) = X(l) = 0; \tag{27a}$$

$$T' = k\lambda T. \tag{27b}$$

Equation for X is the same as above, so we already know that it a has non-trivial solution only for $\lambda = -(\frac{n\pi}{l})^2$, in which case,

$$X_n(x) = A' \sin(\frac{n\pi}{l}x).$$

For the same λ, general solution to the Equation (27b) is

$$T_n(t) = Ae^{-(\frac{n\pi}{l})^2 kt}.$$

Renaming the constants,

$$u_n(x,t) := X_n(x)T_n(t) = A_n e^{-(\frac{n\pi}{l})^2 kt} \sin(\frac{n\pi}{l}x).$$

General solution is again the superposition of separated solutions

$$u(x,t) = \sum_{n=1}^{\infty} A_n e^{-(\frac{n\pi}{l})^2 kt} \sin(\frac{n\pi}{l}x). \tag{28}$$

Observe that in the time limit, $\lim_{t\to\infty} u(x,t) = 0$. Indeed, if both ends of a heated rod are kept at the zero temperature, the heat initially present in the rod will gradually leave the rod through its ends.

Neumann boundary condition

We now solve the wave equation for Neumann boundary condition

$$\begin{aligned} &\partial_t^2 u - c^2 \partial_x^2 u = 0,\ \ 0 \leq x \leq l \\ &\partial_x u(0,t) = \partial_x u(l,t) = 0, \end{aligned} \tag{29}$$

again using separation of variables. Let

$$u(x,t) = X(x)T(t)$$

be a separated solution. Substituting in the equation, we get

$$\frac{X''}{X} = \frac{T''}{c^2 T} = \lambda.$$

A *nontrivial* separated solution is hence determined by the system of ODEs

$$X'' = \lambda X,\ \ \partial_x X(0) = \partial_x X(l) = 0; \tag{30a}$$

$$T'' = c^2 \lambda T. \tag{30b}$$

For $\lambda = 0$, $X_0(x) = Ax + B$, $T_0(t) = A't + B'$. Evaluating at the boundaries, we get $A = 0$. Renaming the constants as $A_0 = 2BA'$, $B_0 = 2BB'$, we get

$$u_0(x,t) := X_0(x)T_0(t) = \frac{1}{2}(A_0 t + B_0).$$

For $\lambda \neq 0$,

$$X(x) = Ae^{\sqrt{\lambda}x} + Be^{-\sqrt{\lambda}x}.$$

$$X'(0) = \sqrt{\lambda}(A - B) = 0, \text{ and } X'(l) = \sqrt{\lambda}(Ae^{\sqrt{\lambda}l} - Be^{-\sqrt{\lambda}l}).$$

First equation implies $A = B$. Substituting into the second equation and multiplying by $e^{\sqrt{\lambda}l}/A\sqrt{\lambda}$, we get

$$e^{2\sqrt{\lambda}l} = 1,$$

which is satisfied if and only if $2\sqrt{\lambda}l = 2n\pi i$ for an integer n, that is, if $\lambda = -(\frac{n\pi}{l})^2$. As for any $z \in \mathbb{C}$, $(e^{iz} + e^{-iz})/2 = \cos(z)$, renaming the constant ($A' = 2A$), we get for each n a solution to the Equation (30a)

$$X_n(x) = A' \cos(\frac{n\pi}{l} x).$$

For $\lambda = -(\frac{n\pi}{l})^2$, general solution of the Equation (30b) is again

$$T_n(t) = A\sin(\frac{nc\pi}{l}t) + B\cos(\frac{nc\pi}{l}t).$$

Renaming the constants as $A_n = A'A$, and $B_n = A'B$,

$$u_n(x,t) := X_n(x)T_n(t) = (A_n\sin(\frac{nc\pi}{l}t) + B_n\cos(\frac{nc\pi}{l}t))\cos(\frac{n\pi}{l}x).$$

General solution is the superposition of separated solutions

$$u(x,t) = \frac{1}{2}(A_0 + B_0 t) + \sum_{n=1}^{\infty}(A_n\sin(\frac{nc\pi}{l}t) + B_n\cos(\frac{nc\pi}{l}t))\cos(\frac{n\pi}{l}x). \tag{31}$$

As for the heat equation with Dirichlet boundary condition

$$\begin{aligned} &\partial_t^2 u - k\partial_x^2 u = 0, \ \ 0 \le x \le l \\ &\partial_x u(0,t) = \partial_x u(l,t) = 0, \end{aligned} \tag{32}$$

the same method yields the solution

$$u(x,t) = \frac{1}{2}A_0 + \sum_{n=1}^{\infty} A_n e^{-(\frac{n\pi}{l})^2 kt}\cos(\frac{n\pi}{l}x). \tag{33}$$

In the following section, we determine constants A_n and B_n from the initial conditions.

5.2 Initial-boundary value problem and Fourier series

Evaluating the solution of Dirichlet boundary value problem for the heat equation (25) at initial the condition $u(x,0) = f(x)$, we get

$$f(x) = \sum_{n=1}^{\infty} A_n \sin\frac{n\pi}{l}x. \tag{34}$$

The series on the right-hand side is called *Fourier sine series.* Thus, to solve this initial-boundary value problem, one is required to calculate the Fourier sine expansion of the function $f(x)$.

In Dirichlet initial-boundary value problem for wave equation (25), the first initial condition $u(x,0) = f(x)$ gives the same series as above, and the second initial condition $\partial_t u(x,0) = g(x)$ gives

$$g(x) = \sum_{n=1}^{\infty} \frac{nc\pi}{l} B_n \sin\frac{n\pi}{l}x. \tag{35}$$

This is again a Fourier sine series, but with re-scaled coefficients.

Similarly, to solve Neumann initial-boundary value problem, we must calculate the *Fourier cosine series.* Indeed, evaluating the solution (33) of the heat equation at the initial condition $u(x,0) = f(x)$, we get

$$f(x) = \frac{1}{2}A_0 + \sum_{n=1}^{\infty} A_n \cos\frac{n\pi}{l}x. \tag{36}$$

Evaluating the solution (33) of wave equation at the first initial condition $u(x,0) = f(x)$ gives the same series as in the heat equation, whereas the second initial condition $\partial_t u(x,0) = g(x)$ gives

$$g(x) = \frac{1}{2}B_0 + \sum_{n=1}^{\infty} \frac{nc\pi}{l} B_n \cos\frac{n\pi}{l}x. \tag{37}$$

To derive the formula for coefficients in Fourier sine and cosine series, we first recall a similar computation that we encountered in linear algebra. Let V be a finite dimensional vector space equipped with a scalar product $\langle -|-\rangle$. Let $\{\vec{e}_1, \ldots, \vec{e}_n\}$ be an orthogonal basis for V (meaning that for every $i \neq j$, $\langle \vec{e}_i|\vec{e}_j\rangle = 0$). Given a vector $\vec{v} \in V$, we wish to determine its expansion in the basis $\{\vec{e}_i\}$ in terms of the scalar product. Assume

$$\vec{v} = \sum_{i=1}^{n} A_i \vec{e}_i.$$

For $1 \leq m \leq n$, applying the scalar product $\langle -|\vec{e}_m\rangle$ on both sides of the above equation, and using the fact that the basis is orthogonal, we get

$$\langle \vec{v}|\vec{e}_m\rangle = A_m \langle \vec{e}_m|\vec{e}_m\rangle.$$

Finally, the formula for the coefficient A_m yields

$$A_m = \frac{\langle \vec{v}|\vec{e}_m\rangle}{\langle \vec{e}_m|\vec{e}_m\rangle}.$$

With this in mind, we return to Fourier sine series. In this case, the vector space V consists of (sufficiently well behaved) functions $f : [0,l] \to \mathbb{R}$ which satisfy the Dirichlet boundary condition $f(0) = f(l) = 0$. Scalar product on V is given by the formula

$$\langle f|g\rangle = \int_0^l f(x)g(x)dx,$$

and the orthogonal "basis[2]" of V is the set of functions $\{\sin \frac{n\pi}{l}x : n \in \mathbb{N}_{>0}\}$. Concerning orthogonality, a direct verification shows that

$$\int_{x=0}^{l} \sin \frac{n\pi}{l} x \sin \frac{m\pi}{l} x dx = \begin{cases} l/2, & \text{if } m = n; \\ 0 & \text{otherwise.} \end{cases}$$

The question if the space V is spanned by those functions, i.e. whether a function $f : [0, l] \to \mathbb{R}$ which satisfies the Dirichlet boundary conditions is the sum of a Fourier sine series, is postponed. In fact, when it comes to function series, there are different types of convergence (pointwise, uniform, L^2 etc.) and for each of them to be guaranteed, there are conditions that the function f needs to satisfy.

Assuming that the series

$$f(x) = \sum_{n=1}^{\infty} A_n \sin \frac{n\pi}{l} x$$

converges, we wish to compute the coefficients A_n by mimicking the above reasoning for finite dimensional vector spaces with a scalar product. Applying on both sides of the above series expansion the scalar product with $\sin \frac{m\pi}{l}x$, we get

$$\int_0^l f(x) \sin \frac{m\pi}{l} x dx = A_m \int_0^l \sin^2 \frac{m\pi}{l} x dx = \frac{l}{2} A_m.$$

From here,

$$A_m = \frac{2}{l} \int_0^l f(x) \sin \frac{m\pi}{l} x dx.$$

Similarly, coefficients B_m in (35) are

$$B_m = \frac{2}{mc\pi} \int_0^l g(x) \sin \frac{m\pi}{l} x dx.$$

The same method is employed to deal with Fourier cosine series. In this case, the vector space V consists of functions $f : [0, l] \to \mathbb{R}$ which satisfy the Neumann boundary condition $f'(0) = f'(l) = 0$. Scalar product on V is the same as before

$$\langle f|g\rangle = \int_0^l f(x)g(x)dx.$$

[2]In linear algebra, a set $\mathcal{B}$ of linearly independent vectors in a vector space V, be it finite or infinite dimensional, forms a basis if every $v \in V$ is a *finite* linear combination of the elements in $\mathcal{B}$. As we intend to express functions from V as convergent series of its elements, the word basis is used in a loose sense.

The orthogonal "basis" of V is the set of functions $\{1, \cos\frac{n\pi}{l}x : n \in \mathbb{N}_{>0}\}$. Indeed, its elements are orthogonal as

$$\int_{x=0}^{l} \cos\frac{n\pi}{l}x \cos\frac{m\pi}{l}x dx = \begin{cases} l/2, & \text{if } m = n; \\ 0 & \text{otherwise.} \end{cases}$$
$$\int_{x=0}^{l} 1 \cdot \cos\frac{m\pi}{l}x dx = 0.$$

Convergence issues are once again postponed for later.

In the series

$$f(x) = \frac{1}{2} + \sum_{n=1}^{\infty} A_n \cos\frac{n\pi}{l}x,$$

we once again apply scalar products with the basis elements on both side to get

$$A_m = \frac{2}{l}\int_0^l f(x)\cos\frac{m\pi}{l}x dx$$

for $m > 0$, and

$$A_0 = \frac{2}{l}\int_0^l f(x)dx.$$

Coefficients B_m in (37) are

$$B_m = \frac{2}{mc\pi}\int_0^l g(x)\cos\frac{m\pi}{l}x dx$$

for $m > 0$, and

$$B_0 = \frac{2}{l}\int_0^l g(x)dx.$$

Denote by $u(x,t)$ be the solution of the heat equation with Neumann boundary conditions, which satisfies the initial condition $u(x,t) = f(x)$. In the time limit,

$$u(x,t) \xrightarrow{t\to\infty} \frac{1}{2}A_0 = \frac{1}{l}\int_0^l f(x)dx.$$

Indeed, if both ends of the heated rod are isolated, in time the heat will spread evenly throughout the rod, with the total amount of heat preserved ($\int_0^l u(x,t)dx = \int_0^l f(x)dx$).

5.3 Convergence

In all the considered initial-boundary value problems (including Robin, circular, and mixed boundary conditions which are dealt with in the exercise section), solution is determined by expanding initial condition(s) in a function series. Terms in that series are the solutions to the differential equation

$$X'' = \lambda X,$$

subject to the respective boundary conditions. A coefficient λ for which such solution exists is an *eigenvalue* of the differentiation operator d^2/dx^2 with the respective boundary condition, and the solution itself is the corresponding *eigenfunction.* Recall that with all the considered boundary conditions, eigenfunctions corresponding to different eigenvalues were orthogonal. There are deep reasons for this. Namely, it turns out that the differentiation operator d^2/dx^2 with all the considered boundary conditions is symmetric[3], and eigenfunctions of a symmetric operator corresponding to different eigenvalues are always orthogonal. Let us prove these facts!

Given functions $f, g : [0, l] \to l$,

$$\int_0^l f''(x)g(x)dx = f'(x)g(x)\Big|_0^l - \int_0^l f'(x)g'(x)dx = (f'(x)g(x) - f(x)g'(x))\Big|_0^l + \int_0^l f(x)g''(x)dx,$$

hence the operator d^2/dx^2 is symmetric if and only if

$$(f'(x)g(x) - f(x)g'(x))\Big|_0^l = 0. \tag{38}$$

Boundary conditions which satisfy the condition (38) are called *symmetric boundary conditions.* Clearly, all the boundary conditions which we consider are symmetric.

Now we prove that eigenfunctions of a symmetric operator corresponding to different eigenvalues are orthogonal. Let $A : V \to V$ be a symmetric operator, and $v_1, v_2 \in V$ such that $Av_1 = \lambda_1 v_1$, and $Av_2 = \lambda_2 v_2$, for $\lambda_1 \neq \lambda_2$.

$$(\lambda_1 - \lambda_2)\langle v_1, v_2\rangle = \langle \lambda_1 v_1, v_2\rangle - \langle v_1, \lambda_2 v_2\rangle = \langle Av_1, v_2\rangle - \langle v_1, Av_2\rangle = \langle Av_1, v_2\rangle - \langle Av_1, v_2\rangle = 0.$$

As $\lambda_1 - \lambda_2 \neq 0$, it follows that $\langle v_1, v_2\rangle = 0$.

With all the boundary conditions that we studied, we obtained an infinite sequence of eigenvalues of the differentiation operator d^2/dx^2. This was also not an accident. It turns out (and this time we will skip the proof) that this is the case with any symmetric boundary conditions. As for any given $\lambda \in \mathbb{R}$, the space of solutions to the differential equation $X'' = \lambda X$ is two dimensional, the eigenspace of any given eigenvalue is at most two dimensional. Fixing an orthogonal basis of every eigenspace, we get an orthogonal sequence

[3]Given a vector space V with a scalar product, a linear operator $A : V \to V$ is symmetric if $\langle Ax, y\rangle = \langle x, Ay\rangle$ for all $x, y \in V$

of eigenfunctions $(X_n)_{n\in\mathbb{N}}$. Given any function $f:[0,l]\to\mathbb{R}$ which satisfies the symmetric boundary condition in hands, we define its *Fourier coefficients* by the same formula as before

$$A_n = \frac{\langle f|X_n\rangle}{\langle X_n|X_n\rangle} = \frac{\int_0^l f(x)X_n(x)dx}{\int_0^l X_n^2(x)dx}.$$

The *Fourier series* of the function f is the series $\sum_{n=1}^{\infty} A_n X_n(x)$.

With this, we can move on to the issue of convergence. When it comes to functions, there are various type of convergence that one may encounter. Given a sequence of functions $(f_n:[0,l]\to\mathbb{R})_{n\in\mathbb{R}}$, we say that a function series $(\sum_{n=1}^{\infty} f_n:[0,l]\to\mathbb{R})_{n\in\mathbb{N}}$ converges towards the limit f

1. *pointwise*, if at each point $x\in[0,l]$,

$$\sum_{n=1}^{N} f_n(x) \xrightarrow{N\to\infty} f(x);$$

2. *uniformly*, if

$$\sup_{x\in[0,1]} |\sum_{n=1}^{N} f_n(x) - f(x)| \xrightarrow{N\to\infty} 0;$$

3. and in the L^2 norm, if

$$\int_0^l (\sum_{n=1}^{N} f_n(x) - f(x))^2 dx \xrightarrow{N\to\infty} 0.$$

We state the convergence theorem for Fourier series without a proof.

Theorem 5.1. *Given any symmetric boundary conditions, the Fourier series of a function $f:[0,1]\to\mathbb{R}$ which satisfies the boundary conditions converges to $f(x)$*

1. *uniformly, provided f is a C^2 function; and*

2. *in the L^2 norm provided $\int_0^l (f(x))^2 dx$ exists.*

Fourier sine series (34), Fourier cosine series (36), and full Fourier series (40) of a function f which satisfies the respective boundary conditions converge pointwise if the function f is continuous, and the function f' is piece-wise continuous (has at most finitely many discontinuities).

Finally, observe that we have so far studied the convergence only at the initial condition $t = 0$. A priori, there is no guarantee that a solution given by a function series in variables x and t which converges for $t = 0$ also converges for $t \neq 0$. This issue is simply ignored in most of the textbooks, and I also tend to ignore it in the class. However, for the sake of completeness, it is addressed in the following section which deals with the uniqueness of the solution for initial-boundary value problems.

5.4 Energy and uniqueness

The aim of this section is to prove that with Dirichlet, Neumann, and circular boundary conditions, initial value problems for both heat and wave equation have unique solutions. We already have candidates for the solutions, which are the function series (25,28,31,33,41,42). It remains to prove that, with coefficients calculated from the initial conditions, those series converge; that their sum solves the boundary value problem in hands; and that the resulting solution is unique. A common method for proving uniqueness is called the energy method.

Given a solution $u(x,t)$ of the wave equation, its energy is defined as

$$E[u](t) = \frac{1}{2}\int_0^l ((\partial_t u(x,t))^2 + c^2\partial_x u(x,t))^2 dx.$$

It is preserved for all the boundary conditions which satisfy

$$\partial_t u(x,t)\partial_x u(x,t)\Big|_{x=0}^{l} = 0, \tag{39}$$

in particular for all the boundary conditions listed above. Indeed, in this situation

$$\begin{aligned}
\frac{dE[u]}{dt}(t) &= \frac{1}{2}\frac{d}{dt}\int_0^l ((\partial_t u(x,t))^2 + c^2(\partial_x u(x,t))^2)dx \\
&= \frac{1}{2}\int_0^l (\partial_t(\partial_t u(x,t))^2 + c^2\partial_t(\partial_x u(x,t))^2)dx \\
&= \int_0^l (\partial_t^2 u(x,t)\partial_t u(x,t) + c^2\partial_t\partial_x u(x,t)\partial_x u(x,t))dx \\
&= \int_0^l (\partial_t^2 u(x,t)\partial_t u(x,t) - c^2\partial_t u(x,t)\partial_x^2 u(x,t))dx + \partial_t u(x,t)\partial_x u(x,t)\Big|_{x=0}^{l} \\
&= \int_0^l (\partial_t^2 u(x,t) - c^2\partial_x^2 u(x,t))\partial_t u(x,t)dx + \partial_t u(x,t)\partial_x u(x,t)\Big|_{x=0}^{l} = 0.
\end{aligned}$$

Theorem 5.2. *Assume $f : [0,l] \to \mathbb{R}$ to be a C^2 function, and $g : [0,l] \to \mathbb{R}$ to be C^1 function. Initial value problems for wave equation*

$$\begin{aligned}
&\partial_t^2 u(x,t) - c^2\partial_x^2 u(x,t) = 0, \quad 0 \le x \le l, t \in \mathbb{R}, \\
&u(x,0) = f(x); \quad \partial_t u(x,0) = g(x)
\end{aligned}$$

with Dirichlet, Neumann, and circular boundary conditions have unique solutions given respectively by the series (25,31,41), with the coefficient determined from the initial conditions.

Proof. First we prove that the considered initial value problems have at most one solution. Assume $u_1(x,t)$, and $u_2(x,t)$ are two solutions. Then $v = u_1 - u_2$ is a solution of the same boundary value problem, with the initial condition $v(x,0) = 0$, $\partial_t v(x,0) = 0$. Consequently, the energy is initially zero, but being constant, it is zero at all times

$$\int_0^l ((\partial_t v(x,t))^2 + c^2(\partial_x v(x,t))^2)dx = 0.$$

Observe that the integrand is a non-negative continuous function. However, integral of such function is zero if and only if the function itself is zero. This in turn implies that $\partial_t v(x,t) = \partial_x v(x,t) = 0$, hence $v(x,t)$ is a constant function. Since $v(x,0) = 0$, this constant value must be zero as well. We conclude that $u_1(x,t) = u_2(x,t)$.

We now prove the existence of the solution. Assume that the initial conditions $u(x,0) = f(x)$, $\partial_t u(x,0) = g(x)$ (for $0 \leq x \leq l$) satisfy the boundary condition in hands. We extend f and g to functions $\tilde{f}, \tilde{g} : \mathbb{R} \to \mathbb{R}$ (in the sense that $\tilde{f}|_{[0,l]} = f$, $\tilde{g}|_{[0,l]} = g$) as follows:

1. for Dirichlet boundary condition they are the unique odd periodic extensions of period $2l$;
2. for Neumann boundary condition they are the unique even periodic extensions of period $2l$;
3. for circular boundary conditions, extend f and g to periodic functions of period l.

In all the above cases, $\tilde{f}$ is still a C^2 function, and $\tilde{g}$ is still a C^1 function (assuming that that $f(x)$ and $g(x)$ satisfy the boundary condition in hands, check this!). d'Alambert's solution $\tilde{u}(x,t)$ to the wave equation on infinite string with initial conditions $\tilde{u}(x,0) = \tilde{f}$ and $\partial_t \tilde{u}(x,0) = \tilde{g}$ is a C^2 function which satisfies the wave equation. By the exercise 3.7, at each time t, d'Alambert's solution is again the extension (in the same sense as above) of a function $\tilde{u}(x,t)$ which satisfies the boundary condition in hands, hence its restriction to $[0,l]$ is a solution $u(x,t)$ to the respective initial-boundary value problem. It is proven by direct calculation that for any fixed $t \in \mathbb{R}$, the Fourier expansion of $u(x,t)$ in the variable x yields respectively the function series (25,31,41), with the coefficients determined from the initial conditions. The computations are rather tedious and perhaps not as interesting.

□

We now move our attention to the heat equation. Energy is defined as

$$E[u](t) = \int_0^l u^2(x,t)dx.$$

We have

$$\begin{aligned}\frac{dE[u]}{dt}(t) &= \int_0^l 2u(x,t)\partial_t u(x,t)dx = 2k\int_0^l u(x,t)\partial_x^2 u(x,t)dx \\ &= 2ku(x,t)\partial_x u(x,t)dx\Big|_{x=0}^l - 2k\int_0^l (\partial_x u)^2(x,t)dx.\end{aligned}$$

Hence, for all the symmetric boundary conditions $dE[u](t)/dt < 0$.

Theorem 5.3. *Initial value problems for heat equation*

$$\begin{aligned}&\partial_t u(x,t) - k\partial_x^2 u(x,t) = 0, \quad 0 \le x \le l, t > 0, \\ &\lim_{t\to 0} u(x,0) = f(x)\end{aligned}$$

with Dirichlet, Neumann, and circular boundary conditions have unique solutions given respectively by the series (28,33,42), with the coefficient determined from the initial condition.

Proof. First we prove that with any symmetric boundary conditions, initial value problem for the heat equation has at most one solution. Assume that $u_1(x,t)$ and $u_2(x,t)$ are solutions. Then $v = u_1 - u_2$ is a solution of the same boundary value problem, with the initial condition $v(x,0) = 0$. Consequently, the energy is initially zero, but being decreasing and positive, it is zero at all times. Once again, energy is the integral of a non-negative continuous function, so it is zero if and only if the function itself is zero. Hence, $v(x,t) = 0$ for all $x \in [0,l]$, $t > 0$.

It remains to prove the existence of a solution. We do this for Dirichlet boundary conditions, other cases are analogous. Let $f(x)$ be a continuous function on $[0,l]$ which satisfies Dirichlet boundary conditions, and whose Fourier sine series

$$f(x) = \sum_{n=1}^{\infty} A_n \sin\frac{n\pi}{l}x$$

converges pointwise. We will prove that the function series (28) with coefficients A_n converges, and that its sum solves the considered initial-boundary value problem. Observe that the sequence of numbers (A_n) is bounded. Indeed,

$$|A_n| = |\frac{2}{l}\int_0^l f(x)\sin(\frac{n\pi}{l}x)dx| \le \frac{2}{l}\int_0^l |f(x)\sin(\frac{n\pi}{l}x)|dx \le \frac{2}{l}\int_0^l |f(x)|dx =: M.$$

With this,

$$|A_n e^{-(\frac{n\pi c}{l})^2 t} \sin(\frac{n\pi}{l}x)| \leq M(e^{-(\frac{\pi c}{l})^2 t})^{n^2}.$$

As for $t > 0$, $e^{-(\frac{\pi c}{l})^2 t} < 1$, and for all positive $y < 1$, the series $\sum_{n=1}^{\infty} y^{n^2}$ converges (prove this!), the series (28) converges pointwise (absolutely).

Next we wish to verify that its sum satisfies the heat equation. As this is the case with each separated solution (i.e. each term of the series), it suffices to show that the series can be derived term-wise in both variables. That is the true as long as the series of its term-wise partial derivatives is uniformly convergent in a neighborhood of any given point. In particular, it suffices to show uniform convergence on the sets $(0, l) \times (t_0, \infty)$ for $t_0 > 0$. Let us prove the statement for $\partial_t u(x, y)$. First and second order derivatives with respect to x are dealt with analogously. For $y = e^{-(\frac{\pi c}{l})^2 t_0}$, the series $\sum_{n=1}^{\infty} n^2 y^{n^2}$ converges (prove this using the ratio test). Take $\epsilon > 0$, and let $N_0 > 0$ be such that $\sum_{n=N_0+1}^{\infty} n^2 y^n < \epsilon/(\frac{\pi c}{l})^2 M$. For $t > t_0$, $N \geq N_0$,

$$\begin{aligned} |\sum_{n=N+1}^{\infty} \partial_t A_n e^{-(\frac{n\pi c}{l})^2 t} \sin(\frac{n\pi}{l}x)| &= (\frac{\pi c}{l})^2 |\sum_{n=1}^{N} n^2 A_n e^{-(\frac{n\pi c}{l})^2 t} \sin(\frac{n\pi}{l}x)| \\ &\leq (\frac{\pi c}{l})^2 M \sum_{n=N+1}^{\infty} n^2 y^{n^2} < \epsilon, \end{aligned}$$

hence the series of termwise partial derivatives with respect to time converges uniformly as required.

The obtained solution $u(x, t)$ of the heat equation clearly satisfies the boundary conditions. It satisfies the initial condition due to the Abel's limit theorem. Indeed, for any fixed $x \in (0, l)$, the power series in the variable s

$$\sum_{n=1}^{\infty} A_n \sin(\frac{n\pi}{l}x) s^{n^2}$$

has the radius of convergence grater or equal to one, and for $s = 1$ it converges to $f(x)$. Abel's limit theorem tells us that

$$\lim_{s \to 1^-} \sum_{n=1}^{\infty} A_n \sin(\frac{n\pi}{l}x) s^{n^2} = f(x).$$

However, the limit on the left is equal to $\lim_{t \to 0^+} u(x, t)$. □

5.5 Exercises

Exercise 5.4. [Circular boundary conditions and full Fourier series] Function $f : [-l, l] \to \mathbb{R}$ satisfies the circular boundary condition if $f(-l) = f(l)$, and$f'(-l) = f'(l)$.

1. Solve the eigen-value problem for the operator d^2/dx^2 with circular boundary condition, and show that $\{1, \sin(n\pi x/l), \cos(n\pi x/l) : n \in \mathbb{N}\}$ is an orthogonal set of eigen-functions, for the scalar product given by $\langle f|g\rangle = \int_{-l}^{l} f(x)g(x)dx$.

2. Given a function f which satisfies circular boundary coefficients, determine the coefficients A_n B_n in the full Fourier series

$$f(x) = \frac{1}{2}B_0 + \sum_{n=1}^{\infty} A_n \sin(n\pi x/l) + B_n \cos(n\pi x/l). \tag{40}$$

3. Using separation of variables, show that general solutions to the wave equation on the interval $-l < x < l$ with circular boundary conditions $u(-l, t) = u(l, t)$, $\partial_x u(-l, t) = \partial_x u(l, t)$ is given by the function series

$$u(x, t) = A_0 + B_0 t + \sum_{n=1}^{\infty} (A_n \sin(\frac{nc\pi}{l} t) + B_n \cos(\frac{nc\pi}{l} t))(C_n \sin(\frac{n\pi}{l} x) + D_n \cos(\frac{n\pi}{l} x)). \tag{41}$$

   Determine the coefficients from the initial conditions $u(x, 0) = f(x)$, $\partial_t u(x, 0) = g(x)$.

4. Using separation of variables, show that general solutions to the heat equation on the interval $-l < x < l$ with circular boundary conditions $u(-l, t) = u(l, t)$, $\partial_x u(-l, t) = \partial_x u(l, t)$ is given by the function series

$$u(x, t) = B_0 + \sum_{n=1}^{\infty} e^{-(\frac{nk\pi}{l})^2 t} (A_n \sin(\frac{n\pi}{l} x) + B_n \cos(\frac{n\pi}{l} x)). \tag{42}$$

   Determine the coefficients from the initial condition $u(x, 0) = f(x)$.

Exercise 5.5 (Robin boundary conditions). Function $f : [0, l] \to \mathbb{R}$ satisfies the Robin boundary conditions if $f'(0) = -a_0 f(0)$, and$f'(-l) = -a_l f'(l)$. We will focus on the case $a_0, a_l > 0$.

1. Show that 0 is not an eigenvalue unless $a_0 = a_l = 0$.

2. Show that the positive eigenvalues $\lambda = \beta^2$ for the operator d^2/dx^2 with Robin boundary condition are the given by the solutions of the equation

$$\tanh \beta l = -\frac{(a_0 + a_l)\beta}{\beta^2 + a_0 a_l}.$$

For $a_0, a_1 > 0$, this equation doesn't have a solution. Hence, all the eigen-values are negative.

3. Show that the negative eigenvalues $\lambda = -\beta^2$ for the operator d^2/dx^2 with Robin boundary condition are the given by the solutions of the equation

$$\tan \beta l = -\frac{(a_0 + a_l)\beta}{\beta^2 - a_0 a_l}, \tag{43}$$

including, in the case $\cos(\sqrt{a_0 a_l} l) = 0$, an additional eigen-value $\beta = \sqrt{a_0 a_l}$. Show that the solutions to the equation (43) form a sequence $(\beta_n)_{n \geq 0}$ with

$$\frac{n\pi}{l} \leq \beta_n < \frac{(n+1)\pi}{l},$$

and that the corresponding eigenfunctions are

$$X_n(x) = C(\cos \beta_n x + \frac{a_0}{\beta_n} \sin \beta_n x).$$

Exercise 5.6. 1. Using separation of variables, find (in terms of a function series) the general solution of the wave equation with the mixed boundary conditions

$$\partial_t^2 u - c^2 \partial_x^2 u = 0, \;\; 0 \leq x \leq l$$
$$u(0,t) = 0, \;\; \partial_x u(l,t) = 0.$$

2. Given a flute (obeys Dirichlet boundary condition) and a clarinet (obeys mixed boundary conditions) of the same size, which will have the lower fundamental frequency ω_1? Which will have lower first overtone ω_2?

Exercise 5.7. Find the Fourier cosine series of the function $\sin(x)$ in the interval $[0, \pi]$. Use it to determine the sums

$$\sum_{n=1}^{\infty} \frac{1}{4n^2 - 1}, \text{ and } \sum_{n=1}^{\infty} \frac{(-1)^n}{4n^2 - 1}.$$

Exercise 5.8. Determine Fourier sine series of the function $f(x) = x$. Using the fact the Fourier series can be integrated term-wise together with the Fourier cosine series of $f(x) = x^2/2$, compute $\sum_{n=1}^{\infty} (-1)^{n+1}/n^2$.

Exercise 5.9. Using the maximum and minimum property of the heat equation, prove that the initial-boundary value problem for the heat equation

$$\partial_t u - k\partial_x^2 u = F(x,t), \quad 0 < x < l, \quad t > 0;$$
$$u(x,0) = \phi(t), \quad u(x,l) = \psi(t), \quad \lim_{t \to 0} u(x,t) = f(x)$$

has at most one solution. Prove the same statement using the energy method.

Exercise 5.10 (Stability). Let $u(x,t)$ and $v(x,t)$ be solutions to the Dirichlet boundary problem for the heat equation, for the initial conditions $\lim_{t\to 0} u(x,t) = f(x)$, and $\lim_{t\to 0} v(x,t) = g(x)$.

1. Using minimum and maximum property for the heat equation, show that for all $t > 0$,

$$\max_{x\in(0,l)} |u(x,t) - v(x,t)| \leq \max_{x\in(0,l)} |f(x) - g(x)|.$$

2. Using the energy method, show that for all $t > 0$,

$$\int_0^l (u(x,t) - v(x,t))^2 dx \leq \int_0^l (f(x) - g(x))^2 dx.$$

A1 Distributions

A function $\phi : \mathbb{R} \to \mathbb{R}$ is *compactly supported* if there exists (large enough) $M > 0$ such that $\phi(x) = 0$ if $|x| > M$. ϕ is *smooth* if it has derivatives of all the orders at all the points in $\mathbb{R}$.

Given smooth compactly supported functions ϕ_1, ϕ_2, and numbers $a, b \in \mathbb{R}$, the function

$$(a\phi_1 + b\phi_2)(x) = a\phi_1(x) + b\phi_2(x)$$

is again smooth and compactly supported. Hence, the set of all such functions forms an (infinite dimensional) vector space, called the space of *test functions* and denoted by $\mathcal{D}$. A linear function $f : \mathcal{D} \to \mathbb{R}$ is continuous if given a sequence $(\phi_n)_{n\in\mathbb{N}}$ of test functions which converges uniformly towards a function ϕ, the sequence of real numbers $(f(\phi_n))_{n\in\mathbb{N}}$ converges towards $f(\phi)$. A continuous linear function $f : \mathcal{D} \to \mathbb{R}$ is celled a *distribution*.

Example A1.1. Every function $f : \mathbb{R} \to \mathbb{R}$ which is Riemann integrable on closed intervals $[a, b]$, for $a, b \in \mathbb{R}$ defines a distribution

$$f(\phi) = \int_{-\infty}^{\infty} f(x)\phi(x)dx.$$

Example A1.2. Dirac delta distribution is defined by $\delta(\phi) = \phi(0)$.

A sequence of distributions $(f_n)_{n\in\mathbb{N}}$ *converges* towards a distribution f if for any test function $\phi \in \mathcal{D}$ the sequence of numbers $(f_n(\phi))_{n\in\mathbb{N}}$ converges towards $f(\phi)$.

Similarly, a given a family of distributions $(f_t)_{t>0}$, we say that $\lim_{t\to 0} f_t = f$ if for any test function $\phi \in \mathcal{D}$, $\lim_{t\to 0} f_t(\phi) = f(\phi)$. With this, it is an immediate consequence of the Theorem 4.2 that $\lim_{t\to 0} S(x,t) = \delta(x)$.

Corollary A1.3.

$$\lim_{t\to 0} S(x,t) = \delta(x).$$

Proof. Let $\phi \in \mathcal{D}$ be a test function. Denote $f_t(x) = S(x,t)$. Using Theorem 4.2, together with the fact that S is an even function in the first variable, we conclude that

$$f_t(\phi) = \int_{-\infty}^{\infty} S(y,t)\phi(y)dy = \int_{-\infty}^{\infty} S(0-y,t)\phi(y)dy \xrightarrow{t\to 0} \phi(0).$$

□

We end the appendix by introducing the distributional derivatives. The notion allows us to make sense of another claim from the Section 18, which is that Diract delta distribution is the derivative of step function

$$H(x) = \begin{cases} 0 & \text{for } x < 0, \text{ and} \\ 1 & \text{for } x > 0. \end{cases}.$$

Distributional derivatives are designed in such a way that they coincide with usual derivatives for distributions determined by a derivable function. Given a derivable function f and a test function ϕ, using integration by parts, we get

$$f'(\phi) = \int_{-\infty}^{\infty} f'(x)\phi(x)dx = f(x)\phi(x)\Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} f(x)\phi'(x)dx.$$

As ϕ is compactly supported, $f(x)\phi(x)\Big|_{-\infty}^{\infty} = 0$. Hence, $f'(\phi) = -f(\phi')$. *Distributional derivative* of a general distribution f is defined by the same formula $f'(\phi) = -f(\phi')$.

With this,

$$H'(\phi) = -\int_{0}^{\infty} \phi'(x)dx = -\phi(\infty) + \phi(0) = \phi(0) = \delta(\phi),$$

as claimed.

A2 Flutes and drums

Vibrations of the air pressure inside a flute of length l are governed by the wave equation with Dirichlet boundary condition $u(0,t) = u(l,t) = 0$. Separated solutions are standing waves (modes)

$$u_n(x,t) = X_n(x)T_n(t) = \sin(k_n x)(A_n \sin(\omega_n t) + B_n \cos(\omega_n t)) \tag{44}$$

of the wave number $k_n = \frac{n\pi}{l}$ and angular frequency $\omega_n = \frac{nc\pi t}{l} = n\omega_1$. When one blows in a flute, each standing wave u_n which forms in the instrument generates the sound of frequency ω_n. Overall sound consists of different frequencies $\omega_n = n\omega_1$, that is of multiples of ω_1. In music, we speak of a *harmonic series*, and ω_1 is called the *fundamental*. As it happens, human ear associates to a harmonic series the note which corresponds to its fundamental. In contrast to this, the spectrum of a drum-head happens to be far from the harmonic sequence (Figure 12). Consequently, the human ear will not associate a specific note to the sound of a drum. We say that drums are unpitched instruments. Still, a student knowledgeable in music might notice that this is not entirely true, as there are also pitched drums, the most notable being orchestral timpani and the Indian tabla.

Let us unpack and justify all this new information from the position of a mathematician/physicist!

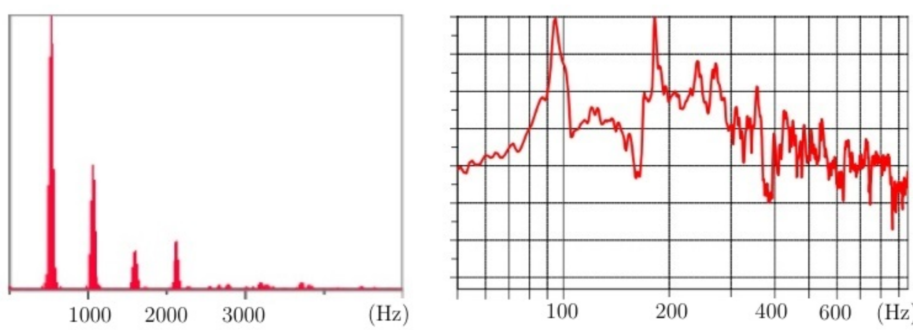


Figure 12: Sound spectrum of a flute (left) and of a drum (right)

We model a drum-head by a radius r_0 circular membrane of uniform tension T and density ρ. Its vibrations are governed by the two-dimensional wave equation subject to the Dirichlet boundary conditions, as the edge of the membrane is attached at all times to a still rim. Denote by $D(r_0, 0)$ the circle of radius r_0 centered at the origin, and by $\partial D(r_0, 0)$ its boundary. For $c^2 = T/\rho$, the boundary value problem in hands reads

$$\partial_t^2 u - c^2 \nabla^2 u = 0, \quad u|_{\partial D(r_0,0)\times\mathbb{R}} \equiv 0.$$

In polar coordinates, the wave equation reads

$$\partial_t^2 u(r,\theta,t) - c^2(\partial_r^2 u(r,\theta,t) + \frac{1}{r}\partial_r u(r,\theta,t) + \frac{1}{r^2}\partial_\theta^2 u(r,\theta,t)) = 0.$$

Let $u(x,t) = R(r)\Theta(\theta)T(t)$ be a separated solution. Substituting in the equation, and dividing by $c^2 R\Theta T$, we get

$$\frac{T''}{c^2 T} = \frac{R''}{R} + \frac{1}{r}\frac{R'}{R} + \frac{1}{r^2}\frac{\Theta'}{\Theta} =: \lambda.$$

An angular frequency $\omega = ck$ is included in the sound spectrum of our drum if the boundary value problem has a separated solution for $\lambda = -k^2$. For such λ, second equality (after multiplying by r^2 and reordering) reads

$$r^2\frac{R''}{R} + r\frac{R'}{R} + k^2 r^2 = \frac{\Theta'}{\Theta} =: \mu$$

As the variable θ is cyclic $(0 = 2\pi)$, function $\Theta(\theta)$ satisfies the circular boundary condition

$$\Theta'' = \mu\Theta, \quad \Theta(0) = \Theta(2\pi), \Theta'(0) = \Theta'(2\pi).$$

Solution of this eigenvalue problem (Exercise 5.4) is $\mu_m = -m^2$, for $n \in \mathbb{N}$, and $\Theta_m(\theta) = A_n \sin(m\theta) + B_n \cos(m\theta)$ (for $m = 0$, $\Theta_m \equiv B_0$.) With, this equation for R becomes

$$r^2 R'' + rR' + (k^2 r^2 - m^2)R = 0, \quad R(r_0) = 0. \tag{45}$$

Substituting $\chi = kr$,

$$\chi^2 R''(\chi) + \chi R'(\chi) + (\chi^2 - m^2)R(\chi) = 0.$$

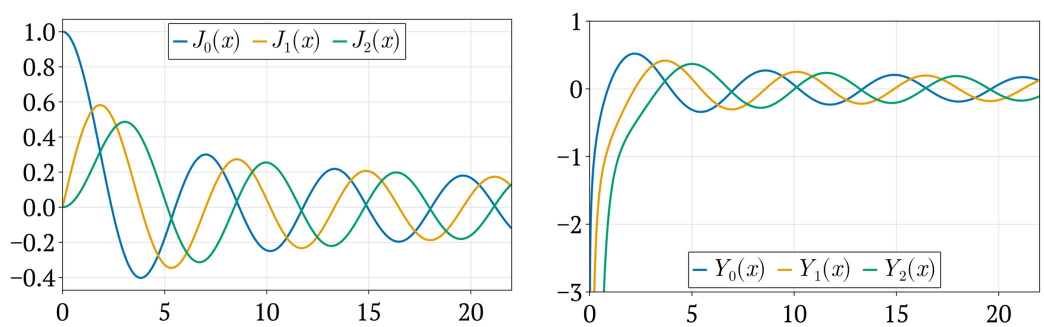


Figure 13: Bessel functions of the first kind (left) and of the second kind (right)

This is Bessel differential equation, and its solutions are linear combinations of Bessel functions of the first kind $J_m(\xi)$ and of the second kind $Y_m(\xi)$. Since Bessel functions of the second kind diverge at zero, they can not be part of a solution, so $R(r) = CJ_m(kr)$. From the boundary conditions, $J_m(kr_0) = 0$. Denoting by λ_{mn} the n-th positive root of J_m, the boundary value problem (45) has a nontrivial solution for every $n > 0$, which reads

$$R_{mn} = C_{mb} J_n(\lambda_{mn} r/r_0).$$

Finally, frequency spectrum of our drum consists of angular frequencies $\omega_{mn} = c\lambda_{mn}/r_0$, for $(m, n) \in \mathbb{N} \times \mathbb{N}_{>0}$. General solution of the wave equation on a circular membrane is the superposition of separated solutions

$$\begin{aligned} u(r, \theta, t) &= \sum_{m,n} R_{mn}(r)\Theta_m(\theta)T_{mn}(t) \\ &= J_m(\frac{\lambda_{mn} r}{r_0} x)(A_{mn}\sin(m\theta) + B_{mn}\cos(m\theta))(C_{mn}\sin(\omega_{mn}t) + D_{mn}\cos(\omega_{mn}t)), \end{aligned}$$

with constants $A_{mn}C_{mn}$, $A_{mn}D_{mn}$, $B_{mn}C_{mn}$, and $B_{mn}D_{mn}$ determined from the initial conditions.

To visualize the standing waves which correspond to all the different separated solutions, or better, to understand their shapes which may at first seem rather chaotic, it is useful to look at the nodal lines. Let us begin by revisiting the flute – one dimensional wave equation with Dirichlet boundary conditions. The shape (peak amplitude) of the standing wave (44) is, up to the re-scaling of amplitude, the graph of function $X_n(x) = \sin(n\pi x/l)$. Thus, the standing wave of the fundamental ω_1 does not have any fixed points except the boundaries $x = 0$ and $x = l$; the first overtone ω_2 has one fixed point in the middle of the flute; second overtone ω_2 has two fixed points at $x = l/3$ and $x = 2l/3$ and so on. For each mode, the points in between two neighboring nodal points vibrate in the same direction, and those on opposite sides of a nodal point move in the opposite directions.

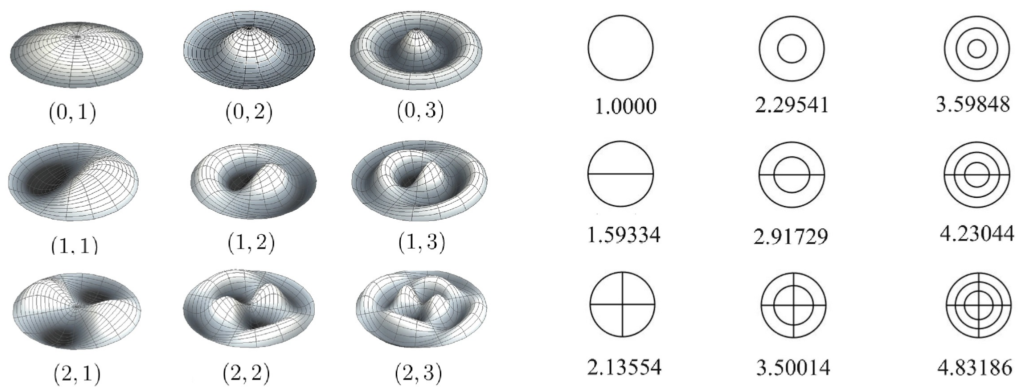


Figure 14: Vibration modes for $m = 1, 2, 3$, $n = 0, 1, 2$ (left); their fixed lines, and frequency ratio with respect to the fundamental $\omega_{mn}/\omega_{0,1}$ (right).

In the case of a drum, shape of the standing wave u_{nm} is, up to re-scaling of the amplitude, equal to the graph of the product of functions $R_{mn}(r)\Theta_m(\theta)$. A point (r, θ) is nodal if either $R_{mn}(r) = 0$, or $\Theta_m(\theta) = 0$. The function R_{mn} has n zero points between 0 and r_0, including the boundary r_0. In the standing wave, they manifest as nodal concentric circles. Being constant, Θ_0 does not provide additional nodal lines, hence standing waves u_{0n} have just n nodal concentric circles. As in the one dimensional situation, the points between two neighboring nodal circles all move in the same direction, and the points on the opposite sides of a nodal circle move in the opposite direction. For $m > 0$, the function Θ_m is, up to re-scaling of the amplitude and a phase-shift, equal to the function $\sin(mx)$. After identifying 0 and 2π, it has $2m$ evenly spaced zero points which manifest as $2m$ evenly spaced radial nodal lines, or better, m evenly spaced nodal diameters in the standing wave. Again, points in any area enclosed by neighboring diameters and neighboring nodal circles move in the same direction, and points on the opposite sides of any nodal line (away from the crossing) move in the opposite direction. Generally, nodal nodal lines of a vibrating membrane are called *Chladni patterns*. Experimentally,they are determined by placing sand on a membrane which is then subjected to the forcing of a resonant frequency. Look now at the front page!

Frequency spectrum of a circular membrane is far from harmonic (see the figure A2), so how can we create a drum with a harmonic spectrum - a pitched drum? As long as our drum is correctly modeled by the wave equation on a uniform circular membrane, this is hopeless. So, what should we change? Indian drum makers will tell you that the membrane shouldn't be uniform; western classical percussionists would tell you that the surrounding air should be used to harmonize the drum; while an African drummer will tell you that the sound of a drum is perfect the way it is, and that you shouldn't be obsessed by harmonizing it in the first place. All are correct, as each design gives the sound that fits perfectly with the music in which it is used.

The best known Indian pitched drum is called tabla. It is modeled by wave equation on a circular membrane of uniform tension T and *non-uniform density* $\rho(r)$. Consequently, $c^2 = T/\rho$ also depends on r. In practice, this is achieved by applying a special black paste (Syahi) on the membrane. Typically, a master craftsman would apply over 100 layers in gradually smaller concentric circles, polishing and verifying the sound at each step, until the drum becomes perfectly harmonic. The resulting spectrum is $\omega_{11} = 2\omega_{01}$, $\omega_{21} = \omega_{02} = 3\omega_{01}$, $\omega_{31} = \omega_{12} = 4\omega_{01}$, and $\omega_{41} = \omega_{03} = \omega_{22} = 5\omega_{01}$. Other modes, which remain non-harmonic, are damped by an additional layer of skin of annular shape (Keenar), attached at the very edge of the membrane.

There is a number of published scientific papers in which the wave equation is solved on a circular membrane of non-uniform density $\rho(r)$, with the intention to model tabla. Interestingly, none of the proposed functions $\rho(r)$ gives harmonicity as perfect as that achieved by a master craftsman.

So far, wave equations by which we modeled drums (ideal membrane and tabla) did not encode interaction with the surrounding air. As long as the diameter of the membrane is not too large, and volume of the air-filled space that surrounds it is not too small, the model is valid. In the case of timpani, neither of the assumptions are satisfied: the membrane is huge, while the volume of the air enclosed in the kettle below the drum is relatively small. Here is what happens:

The room in which the drum is situated provides a reservoir of air which interacts with the membrane from above, while the air enclosed within the kettle interacts with the membrane from below. Together, they alter the frequency spectrum in such a way that the frequency ratios of modes $(1,1)$, $(2,1)$, $(3,1)$, and $(4,1)$ become $1 : 1.5 : 1.99 : 2.44$, which is close enough to the harmonic sequence (with the missing fundamental) $1 : 1.5 : 2 : 2.5$ to create a strong sense of pitch. But what happens with the other dominant modes? Answer is that they magically disappear! It happens that the small volume of the kettle doesn't allow the modes $(0,1)$, $(0,2)$, and $(0,3)$ to resonate: those modes will receive little energy to begin with, and they will also fade out faster than the other modes.

The system of differential equation which governs this physical process is rather complex. On one side, difference of the air pressure on the two sides of the membrane is encoded by an inhomogeneous term in the two dimensional wave equation which governs vibrations of a circular membrane. On the other side, those two pressures serve as boundary conditions of a three dimensional wave equation which governs fluctuations of the air pressure within the two air reservoirs. There is a number of scientific papers in which that system is solved numerically. This time, results align perfectly with the sound spectrum of the modeled drums.