Lecture Note for Honor PDE

Liying LI

September 22, 2025

1 Introduction

1.1 Derivation of PDEs

Many partial differential equations originate from physical models. Understanding these models provides valuable insight into the intuition underlying the equations. In this section, we demonstrate the derivation of several common PDEs from fundamental physical principles.

1.1.1 Transport equation

Let $u(t,x) : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$ be the unknown function. The variable t is the time coordinate, and x is the space coordinate. The variable u can be the density of something, the velocity field, etc.

For illustration, suppose that we are modeling the traffic flow and $u(t,x)$ is the density of cars at (t,x). Let $a < b$. We first have the *conservation of mass* equation

$$\frac{d}{dt}\Big(\int_a^b u(t,x)\,dx\Big) = J(t,a) - J(t,b). \tag{1.1}$$

Here, the LHS is the rate of change of the total number of cars, and $J(t,x)$ is the *flux* at (t,x): the number of cars moving from the left of x to the right of x in unit time.

Assume that u and J is smooth enough, so that we can differentiation and interchange the order of differentiation and integration. Taking the t-derivative in (1.1) yields

$$\int_a^b \partial_t u(t,x)\,dx = J(t,a) - J(t,b) = -\int_a^b \partial_x J(t,x)\,dx.$$

Then

$$\int_a^b \big[\partial_t u(t,x) + \partial_x J(t,x)\big]\,dx = 0.$$

Since a and b are arbitrary, and the integrand is a continuous function, we must have the relation

$$\partial_t u(t,x) + \partial_x J(t,x) = 0. \tag{1.2}$$

This is the differential form of (1.1).

Next, we need to relate J to u to eliminate the unknown J in order to close the equation for u. Since u is the density, by the physical meaning of flux we have

$$J(t,x) = u \cdot V(t,x),$$

where $V(t,x)$ is the velocity field. It remains to determine how the velocity depends on the density; this may differ from one model from another. Here are some examples.

- $V(t,x) = \text{const}$. Then (1.2) reduces to

$$\partial_t u + c\partial_x u = 0.$$

One can check that the general solution is given by $u(t,x) = \phi(x - ct)$, that is, the initial density profile $\phi(\cdot)$ moves with constant speed c.

- $V(t,x) = 1 - u$. This is a more realistic model for the traffic jam: the velocity is decreasing as the density increases, and at maximum density $u = 1$ the traffic flow completely stops. The resulting equation is

$$\partial_t u + \partial_x \big(u(1-u)\big) = 0 = \partial_t u + \partial_x u - 2u \cdot \partial_x u = 0.$$

  Although this equation seems simple, it is a nonlinear PDE and exhibits nontrivial behaviors such as formation of shocks.

We have the general form of the transport equation

$$\partial_t u + \partial_x \big(uV(u)\big) = 0, \tag{1.3}$$

where $V : \mathbb{R} \to \mathbb{R}$ is a function that depends on the model.

We can further generalize (1.3) to dimension $d > 1$. We first guess the form of the equation by matching the dimension, and then we will derive it rigourously using the conservation of mass.

Since u is the density, it is a multi-variate function

$$u(t,x) : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}.$$

Since V gives the velocity, so V and $J = u \cdot V$ must be vector functions:

$$V(x) : \mathbb{R}^d \to \mathbb{R}^d, \quad J(t,x) = u \cdot V : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}^d.$$

Looking the LHS of (1.3), $\partial_t u$ takes value in $\mathbb{R}$, so the differential operator must turn $J(t,x)$ into a function that maps $\mathbb{R}^d$ to $\mathbb{R}$. The only such operator is the divergence operator $\nabla\cdot$ acting on a vector function $f = (f_1, \ldots, f_d)$

$$\nabla \cdot f = \nabla \cdot (f_1, \ldots, f_d) \coloneqq \sum_{i=1}^{d} \frac{\partial}{\partial x_i} f_i.$$

Hence, we obtain a reasonable guess of the generalization of (1.3) in an arbitrary dimension $d > 1$:

$$\partial_t u + \nabla \cdot \big(uV(u)\big) = 0, \tag{1.4}$$

where $u : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}$ is the unknown function and $V : \mathbb{R} \to \mathbb{R}^d$ is a given function depending on the model.

Next, we give a rigourous derivation using the conservation of mass. The key tool is the *Divergence Theorem.*

Theorem 1.1 *Let $\Omega \subset \mathbb{R}^d$ be a domain with piecewise $\mathcal{C}^1$-boundary. Let $F : \bar{\Omega} \to \mathbb{R}^d$ be $\mathcal{C}^1$. Then*

$$\int_{\partial\Omega} F \cdot \vec{n}\, dS = \int_{\Omega} \nabla \cdot F\, dx, \tag{1.5}$$

where dS denotes the surface element on $\partial\Omega$, and $\vec{n}$ is the outer unit normal vector on $\partial\Omega$.

(1.5) is also refered to as the *Stokes formula*, or simply integration by parts, since from right to left a differential operator is removed.

Now let Ω be an arbitrary $\mathcal{C}^1$-domain in $\mathbb{R}^d$. The total mass in Ω is given by $\int_\Omega u(t,x)\,dx$. By the conservation of mass, the rate of change of the total mass is a conseqeunce of the flux of mass across of the boundary. Thus we have

$$\frac{d}{dt}\int_\Omega u(t,x)\,dx = -\int_\Omega J\cdot\vec{n}\,dS.$$

To double check the RHS: if the direction of the flux is tangent to the boundary at some point, that is $J\cdot\vec{n}=0$, then there is no mass escaping from this point, justifying the form of the integrand. Also, if the flux J is point outwards and has the same direction as $\vec{n}$, this will result in a decrease of the mass, and hence the minus sign on the RHS.

Assuming u is smooth enough so that the order of differentiation and integration can be exchanged on the LHS, and using Theorem 1.1 on the RHS, we obtain

$$\int_\Omega (\partial_t u + \nabla\cdot J)\,dx = 0.$$

Since this holds for an arbitrary $\mathcal{C}^1$-domain, we have pointwise

$$\partial_t u + \nabla\cdot J = 0. \tag{1.6}$$

Plugging in $J = uV(u)$ we obtain (1.4).

For the last step we use the following simple result.

Lemma 1.2 • *If $f\in\mathcal{C}(\mathbb{R}^d)$ and $\int_\Omega f(x)\,dx = 0$ for any rectangle $\Omega\subset\mathbb{R}^d$, then $f\equiv 0$.*

- *If $f\in L^1_{\mathrm{loc}}(\mathbb{R})$ and $\int_\Omega f(x)\,dx = 0$ for any rectangle $\Omega\subset\mathbb{R}^d$, then $f=0$ almost everywhere.*

As we will see, the transport equation may develop singularity no matter how smooth the initial condition is, so (1.4) may not hold for every point, but it is at least safe to say that it holds almost everywhere.

1.1.2 Heat equation

In (1.6), we may interpret u as the temperature and J as the heat flux; then (1.6) follows from *the conservation of energy*, as confirmed by Joule's experiment. To close the equation, we need to relate J to u. Fourier's law states that the heat flux is proportional to the negative gradient of the temperature field, expressed as

$$J = -c\nabla u,$$

where the constant c denotes the *thermal conductivity*. Here, the gradient operator ∇ is defined by

$$(\nabla f)(x_1,\dots,x_d) = \big(\partial_{x_1}f(x_1,\dots,x_d),\dots,\partial_{x_d}f(x_1,\dots,x_d)\big).$$

Combined with (1.6), we obtain

$$\partial_t u = -\nabla\cdot(-c\nabla u) = \sum_{i=1}^d \partial_{x_ix_i}u =: c\Delta u. \tag{1.7}$$

The operator Δ is called the Laplacian operator. (1.7) is called the *heat equation*. Usually we set $c=1$.

The heat equation also models the phenomenon of *diffusion*. Let u represent the concentration of a substance within the fluid, analogous to the density. Particles of this substance may move under external forces, but even in the absence of such external forces, diffusion causes particles to move from regions of higher concentration to lower concentration. Specifically, *Fick's law* states that the flux J is proportional to the negative gradient of u, and hence the diffusion is modeled by the heat equation as well.

The heat equation is a second-order PDE since it involves second partial derivatives. It is classified as a *parabolic equation* since the time derivative is only first order, analogous to the *parabola equation* $t = x^2$.

1.1.3 Wave equation

The wave equation models the wave phenomena in elastic media. Let Ω be a domain representing an elastic object, such as a string, a rod, or membrane. For simplicity, we take $\Omega = (a, b)$ as an example. The unknown function $u(t, x) : \mathbb{R} \times \Omega \to \mathbb{R}$ under consideration is the displacement of the object from its equilibrium position. By Newton's second law, we have

$$\partial_{tt} u(t, x) = F(t, x),$$

where $F(t, x)$ is the force acting at position x. To determine this force, we invoke *Hooke's law*, which states that the elastic force is negatively proportional to the displacement

$$F = -k\Delta L.$$

We imagine there are two small springs on the intervals $(x - \Delta x, x)$ and $(x, x + \Delta x)$. The net force at (t, x) results from the combination of the elastic forces from these springs. Applying the Hooke's law gives

$$F(t, x) \approx F_1 + F_2 = -k\Big(u(t, x) - u(t, x - \Delta x)\Big) - k\big(u(t, x) - u(t, x + \Delta x)\big) \approx k(\Delta x)^2 \partial_{xx} u(t, x).$$

Combining all these and assuming $k(\Delta x)^2 \to c$ as $\Delta x \to 0$, we obtain the *wave equation*

$$\partial_{tt} u = c\partial_{xx} u.$$

The wave equation in dimensions $d > 1$ can be derived analogously or postulated as

$$\partial_{tt} u = c\Delta u.$$

This equation is classified as the *hyperbolic equation* since both the time and space derivatives are of second order and have the opposite signs, which resembles the *hyperbola equation* $t^2 = x^2$.

1.1.4 Laplace equation

Consider the heat equation in a domain Ω, with boundary condition $u\big|_{\partial\Omega} = \varphi$ and initial condition $u\big|_{t=0} = u_0$. From a physical perspective, if the temperature is fixed at the boundary, eventually the temperature field will reach an equilibrium state, that is, there is $u_* : \Omega \to \mathbb{R}$ such that $u(t, x) \to u_*(x)$ as $t \to \infty$, where u_* may or may not depend on u_0. Since $v(t, x) = u_*(x)$ also satisfies the heat equation as it is the equilibrium, we obtain

$$\Delta u_* = 0, \quad u_*\big|_{\partial\Omega} = \varphi. \tag{1.8}$$

This is known as the *Laplace equation*. It is classified as an *elliptic equation* since all second derivatives have the same sign, resembling the ellipse equation $ax^2 + by^2 = 1$.

We now derive the Laplace equation using the *calculus of variation*, a powerful tool to obtain PDEs. We consider the following minimization problem:

$$\inf_{u\big|_{\partial\Omega}=\varphi} \int_\Omega |\nabla u|^2(x)\, dx =: \inf_{u\big|_{\partial\Omega}=\varphi} I[u]. \tag{1.9}$$

The square bracket $[\cdot]$ stresses that I is a *functional*, that is, a "function" of functions. Assume that u_* achieves the minimum of (1.9), that is,

$$I[u_*] = \min_{u\big|_{\partial\Omega}=\varphi} I[u].$$

Intuitively, $I[u]$ is the L^2-norm of the heat flow corresponding to the temperature field u, and if the L^2-norm is minimized, the temperature field is at the equilibrium state.

Assuming u_* is the minimum function, let us derive conditions that u_* must satisfy. Let $v \in \mathcal{C}_0^\infty(\Omega)$ be arbitrary. We introduce perturbation of u_* as

$$u_\varepsilon = u_* + \varepsilon v, \quad \varepsilon \in \mathbb{R}$$

Since v vanishes at $\partial\Omega$, the function u_ε satisfies the boundary condition. Let $f(\varepsilon) = I[u_\varepsilon]$. Since f achieves minimum at $\varepsilon = 0$, we must have $f'(0) = 0$ provided that the derivative exist. We do not know if f is actually differentiable, but assuming that all functions are nice, this is indeed the case and we have:

$$0 = f'(0) = \frac{d}{d\varepsilon}\Big|_{\varepsilon=0} \int_\Omega |\nabla u_* + \varepsilon \nabla v|^2\, dx = 2\int_\Omega \nabla u_* \cdot \nabla v\, dx = 0. \tag{1.10}$$

To proceed, we use the following useful integration-by-part formula.

Lemma 1.3 *Let Ω be a $\mathcal{C}^1$-domain and $u, v \in \mathcal{C}^1(\bar\Omega) \cap \mathcal{C}^2(\Omega)$. Then*

$$\int_\Omega \nabla u \cdot \nabla v\, dx = -\int_\Omega u\Delta v\, dx + \int_{\partial\Omega} u\frac{\partial v}{\partial n}\, dS, \tag{1.11}$$

$$\int_\Omega \nabla u \cdot \nabla v\, dx = -\int_\Omega v\Delta u\, dx + \int_{\partial\Omega} v\frac{\partial u}{\partial n}\, dS, \tag{1.12}$$

$$\int_\Omega u\Delta v - v\Delta u\, dx = \int_{\partial\Omega} u\frac{\partial v}{\partial n} - v\frac{\partial u}{\partial n}\, dS. \tag{1.13}$$

Proof: The last identity follows from taking difference of the first two. Since the roles of u and v are symmetric, it suffices to prove (1.11). Indeed, consider the vector function $F = u\nabla v : \Omega \to \mathbb{R}^d$. Then $\nabla \cdot F = \nabla u \cdot \nabla v + u\Delta v$ and $F \cdot \vec{n} = u\frac{\partial v}{\partial n}$. Applying Theorem 1.1 to F yields the desired conclusion. □

Using Lemma 1.3, we can continue with (1.10) to obtain

$$0 = \int_\Omega \nabla u_* \cdot \nabla v\, dx = -\int_\Omega v\Delta u_*\, dx, \quad \forall v \in \mathcal{C}_0^\infty(\Omega). \tag{1.14}$$

There is no boundary term after integration by parts since v vanishes at the boundary. Since (1.14) holds for arbitrary $v \in \mathcal{C}_0^\infty(\Omega)$, a variant of Lemma 1.2 implies that $\Delta u_* = 0$ pointwise assuming its continuity. Hence we derive the Laplace equation again.

As another example, the variational problem

$$\inf_{u\big|_{\partial\Omega}=\varphi} \int_\Omega \sqrt{1 + |\nabla u|^2}\, dx$$

gives rise the *minimal surface equation*. This will be left as an exercise.

1.1.5 *Viscous Burgers equation and fluid equation

Let us consider the velocity field $u(t,x)$ of particles moving on $\mathbb{R}$. By Newton's law, we have

$$\text{acceleration of particles at } (t_0,x_0) = \text{friction} + \text{external force}. \tag{1.15}$$

First, let us express the acceleration field from the velocity field. The naive guess $\partial_t u$ is wrong, since the particles at position x are not the same for different t. To get the correct form of the acceleration, we must follow a fixed particle. Let $x(t)$ be the trajectory of the particle passing (t_0,x_0) (that is, $x(t_0)=x_0$). Then by definition

$$\dot{x}(t) = u\big(t, x(t)\big).$$

Hence,

$$\ddot{x}(t) = \frac{d}{dt}u\big(t,x(t)\big) = \partial_t u + \dot{x}\cdot\partial_x u = \partial_t u + u\cdot\partial_x u. \tag{1.16}$$

This gives the LHS of (1.15).

For the RHS, first, the friction force is modeled by $\partial_{xx}u$. To understand why second derivative appears, it suffices to note that if $\partial_{xx}u=0$, then u is linear and there is no friction in the sheer transform. Last, the external force is modeled by an arbitrary function $f(t,x)$. Combining all these, we obtain the full *viscous Burgers equation*:

$$\partial_t u + u\partial_x u = \partial_{xx}u + f(t,x).$$

Its multi-dimensional analogue is

$$\partial_t u + u\cdot\nabla u = \Delta u + f(t,x), \quad u:\mathbb{R}_+\times\mathbb{R}^d\to\mathbb{R}.$$

The Burgers equation is a mixture of the "transport term" $u\partial_x u$ and the "diffusion term" $\partial_{xx}u$.

The Burgers equation is a toy model for fluid dynamics. Here we also mention the celebrated Navier-Stokes equation, and by now we can understand the physical meaning of all the terms in the equation. Assuming the fluid is incompressible (meaning the density is constant), the Navier-Stokes equation reduced to an equation of the velocity field: $u(t,x):\mathbb{R}_+\times\mathbb{R}^d$, $d=2,3$,

$$\begin{cases}\partial_t u + u\cdot\nabla u + \nabla p = \Delta u + f,\\ \nabla\cdot u = 0.\end{cases}$$

Here, we recognize the material derivative term $\partial_t u + u\cdot\nabla u$, which is the acceleration field. All the other are forcing terms: the pressure term ∇p, the friction Δu, and the external force f. The divergence-free constraint comes from conservation of mass:

$$\frac{\partial\rho}{\partial t} + \nabla\cdot(\rho u) = 0 \implies \nabla\cdot u = 0.$$

Although the pressure p is also unknown and first equation seems under-determined, the divergence-free constraint can in fact eliminate the pressure term in the first equation.

1.1.6 *Maxwell equation

In this section we briefly look at the Maxwell's equation that models the electro-magnetic field. The unknowns are the electric field E and the magnetic field B, both are vector functions on $\mathbb{R}^3$. All of the four equations can be written down in the differential form and in the integral form.

- Gauss's law:
$$\nabla \cdot E = \frac{\rho}{\varepsilon_0}, \quad \int_{\partial\Omega} E \cdot \vec{n}\, dS = \frac{1}{\varepsilon_0}\int_\Omega \rho\, dx.$$
Here, ρ is the electric charge density, ε_0 is a physical constant, and Ω is an arbitrary domain.

- Gauss's law for magnetism:
$$\nabla \cdot B = 0, \quad \int_{\partial\Omega} B \cdot \vec{n}\, dS.$$

- Faraday's equation (electric generated from a changing magnetic field):
$$\nabla \times E = -\frac{\partial B}{\partial t}, \quad \oint_{\partial\Sigma} E \cdot d\ell = -\int_\Sigma \frac{\partial B}{\partial t} \cdot dA,$$
where Σ is any surface.

- Ampère's circuital law (magnetic field generated by currents):
$$\nabla \times B = \mu_0(J + \varepsilon_0 \frac{\partial E}{\partial t}), \quad \oint_{\partial\Sigma} B \cdot d\ell = \int_\Sigma \mu_0\big(J + \varepsilon_0 \frac{\partial E}{\partial t}\big) \cdot dA,$$
where J is the current.

It is well-known that electro-magnetic field related to waves. To see this from the equation, we consider the vacuum case where $\rho = J \equiv 0$. Then we have

$$\nabla \times (\nabla \times E) = \nabla(\nabla \cdot E) - \Delta E = -\Delta E = -\frac{\partial}{\partial t}(\nabla \times B) = -\mu_0\varepsilon_0 \frac{\partial^2 E}{\partial^2 t},$$

so E satisfies the wave equation, where the wave speed (i.e., the light speed) is $c = \sqrt{\mu_0\varepsilon_0}$. A similar calculation yields a wave equation for B.

1.2 key questions in this course

This course will focus on four elementary partial differential equations (PDEs), which model fundamental physical phenomena and serve as foundational components for more complex PDEs:

- the transport equation: $\partial_t u + \partial_x V(u) = f$;
- the Laplace equation: $\Delta u = f$;
- the heat equation: $\partial_t u = \Delta u$;
- the wave equation: $\partial_{tt} u = \Delta u$.

One part of the course is devoted to how to write down solutions of the PDEs, using techniques like Fourier analysis, separation of variables and etc. A more important part is to develop *well-posedness* theory without an explicit form of the solution. The well-posedness theory is three-fold:

- existence of solution, including suitable conditions on the boundary and initial condition, regularity requirement;
- uniqueness of solution
- stability: how sensitive the solution is to initial and boundary data.

For a rigorous well-posedness theory we must be accurate about the solution space. A key concept is the *classical solution*, where all the derivatives appearing in the PDE are continuous function so that the PDEs make sense pointwise. When there are both time and space derivative, we use $\mathcal{C}^{\alpha,\beta}$ to indicate the space of functions that has α-th order continuous derivative in t and β-th order continuous derivative in space. For example, classical solutions of the first order transport equation live in $\mathcal{C}^{1,1}$, for the heat equation $\mathcal{C}^{1,2}$, and for the wave equation $\mathcal{C}^{2,2}$. We may also spend some time discussing how to define weak solutions, solutions that have a lower regularity requirement.

2 First-order transport equation

In this section we study the first-order transport equation:

$$\begin{cases} \partial_t u + b(t,x,u) \cdot \partial_x u = f(t,x,u), \\ u(0,x) = \phi(x), \end{cases} \quad u(t,x) : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}. \tag{2.1}$$

2.1 Method of characteristics

2.1.1 Constant b

Suppose $b(t,x,u) = V$ is a constant and $f \equiv 0$. The equation becomes

$$\partial_t u + V \partial_x u = 0, \quad u(0,x) = \phi(x).$$

We introduce $U(t) = u(t, x_0 + Vt)$ where u is a solution and $x_0 \in \mathbb{R}$ is fixed. Then

$$\dot{U}(t) = \partial_t u(t, x_0 + Vt) + V \partial_x u(t, x_0 + Vt) = 0.$$

Hence,

$$U(t) \equiv U(0) = u(0, x_0) = \phi(x_0),$$

and we have

$$u(t,x) = \phi(x - Vt).$$

The curves $\eta(t) = x_0 + Vt$ are called *characteristics*. Intuitively, the initial data ϕ is propagating along these curves.

As a remark, if $\phi \in \mathcal{C}^1$, then $u = \phi(x - Vt) \in \mathcal{C}^{1,1}$ is a classical solution. But even if $\phi \notin \mathcal{C}^1$, this is still the only plausible solution to the PDE, despite being non-classical. From this example, we see that dealing with non-classical solutions is already inevitable even for very simple PDE,

2.2 A non-homogeneous example

We consider the following equation:

$$\begin{cases} \partial_t u + x \partial_x u = u + x, \\ u(0,x) = \phi(x). \end{cases} \tag{2.2}$$

We are seeking characteristics $\eta(t)$ so that

$$U(t) = u\big(t, \eta(t)\big)$$

solves a simple ODE. We clearly have

$$\dot{\eta}(t) = \eta \implies \eta(t) = C_1 e^t.$$

Plugging into U, we have

$$\dot{U}(t) = U(t) + C_1 e^t.$$

The general solution for this ODE is

$$U(t) = C_1 t e^t + C_2 e^t.$$

Finally, we need to determine the constants C_1 and C_2. We have

$$\eta(t) = C_1 e^t = x, \quad U(0) = C_2 = \phi(C_1) \implies C_1 = xe^{-t}, \quad C_2 = \phi(xe^{-t}).$$

Hence, the solution to the PDE is

$$u = xt + \phi(xe^{-t})e^t.$$

One can check by direct computation that it indeed solves the original PDE.

2.3 Genera linear case

We consider the general linear case

$$\begin{cases} \partial_t u + b(t,x)\partial_x u = f(t,x,u), \\ \qquad\qquad u(0,x) = \phi(x). \end{cases} \tag{2.3}$$

We state a well-posedness result.

Theorem 2.1 *Assume that $b \in \mathcal{C}^{0,1}$, $\phi \in \mathcal{C}^1$ and $f \in \mathcal{C}^{0,1,1}$. Then there exists a unique solution to* (2.3).

Proof: We consider the characteristic ODE

$$\dot{\eta}(t) = b\big(t, \eta(t)\big), \quad \eta(0) = x_0.$$

Since b is Lipschitz in x, by standard ODE theory, there is a unique solution for every initial condition x_0. Moreover, the solution map

$$\Phi_t : x_0 \mapsto \eta(t; x_0)$$

is a $\mathcal{C}^1$-diffeomorphism of $\mathbb{R}$, that is, both Φ_t and Φ_t^{-1} are in $\mathcal{C}^1$. Indeed, Φ_t' satisfies the ODE

$$\frac{d}{dt}\Big(\Phi_t'\Big) = \partial_x b\big(t, \Phi(t)\big)\Phi_t', \quad \Phi_0' = 1.$$

Let u_1 and u_2 be two $\mathcal{C}^{1,1}$-solutions of the PDE and let

$$w_i(t) = u_i\big(t, \eta(t)\big), \quad i = 1, 2. \tag{2.4}$$

Then w_i solves the ODE

$$\dot{w}_i(t) = f\big(t, \eta(t), w_i(t)\big), \quad w_i(0) = \phi\big(\eta(0)\big). \tag{2.5}$$

Since the above ODE has unique solution, we have $w_1 = w_2$. Hence the PDE has unique solution.

For the existence of the solution, let $w(t; w_0)$ be the solution to the ODE (2.5) with initial condition w_0. Then one can check that

$$u(t,x) = w\Big(t; \phi\big(\Phi_t^{-1}(x)\big)\Big)$$

is a $\mathcal{C}^{1,1}$-function that solves the PDE. The detailed computation will be omitted. For concrete equations, the justification will be more straightforward. $\square$

2.4 Nonlinear equation

In nonlinear transport equations, the function $b = b(t, x, u)$ also depends on u. In this case, we have to solve the ODEs of η and w together:

$$\begin{cases} \dot{\eta}(t) = b(t, \eta, w), \\ \dot{w}(t) = f(t, \eta, w). \end{cases}$$

2.4.1 Burgers equation

The Burgers equation is one of the simplest nonlinear PDEs. We start from the homogeneous equation ($f \equiv 0$).

$$\partial_t u + u\partial_x u = 0, \quad u(0, x) = \phi(x).$$

The characteristic ODE system is

$$\dot{\eta}(t) = w, \quad \dot{w}(t) = 0.$$

The second equation indicates that w is constant, implying that η is a linear function: $\eta(t) = x_0 + t\phi(x_0) = x$. Physically, this corresponds to particles moving at constant velocity due to the absence of external forcing f. The characteristics, representing particle trajectories, are therefore straight lines. To determine the velocity field at (t, x), one may identify the origin of the particle arriving at the point (t, x), and retrieve its velocity.

In nonlinear scenarios, however, characteristics may intersect, causing the correspondence $x \mapsto x_0$ to cease being one-to-one. If multiple characteristics pass through a point (t, x), it implies that particles carrying different velocity meet at (t, x), causing the velocity field at (t, x) is undetermined. On the other hand, one can check that a necessary and sufficient condition to avoid intersection is that ϕ is increasing, but in such case, certain points (t, x) may lack any passing characteristics, again leaving the velocity field undetermined.

Through two examples we will illustrate how to resolve these issues.

1. Rarefaction solution

   Suppose the initial condition is given by

   $$\phi(x) = \begin{cases} 0, & x \leq 0, \\ 1, & x > 0. \end{cases}$$

   By looking at the characteristics, we have

   $$u(t, x) = \begin{cases} 0, & x \leq 0 \\ 1, & x \geq t. \end{cases}$$

   There is no characteristic in the region $0 < x < t$, leaving the solution undetermined. Let us try to construct a reasonable solution. We notice that the initial condition ϕ is already discontinuous. We cannot expect our constructed solutions to be continuous, but we should make the discontinuous point as few as possible. One possible choice is

   $$u(t, x) = \begin{cases} 0, & x \leq kt, \\ 1, & x > kt, \end{cases}$$

   where $k \in (0, 1)$. The solution is only discontinuous along the line $x = kt$.

Are these solutions reasonable? From the point of view of differentiability it seems yes: apart from the curve $x = kt$, the function is continuously differentiable and satisfies the PDE. But it turns out that these are *non-physical* solution.

To obtain a physical solution, we note that the root issue is that the initial condition is not $\mathcal{C}^1$. Nonsmooth function is merely a pure mathematical object; we should think of the discontinuous function ϕ as an idealization of another function that has an abrupt near 0:

$$\phi_\varepsilon(x) = \begin{cases} 0, & x \le 0, \\ x/\varepsilon, & 0 \le x \le \varepsilon, \\ 1, & x \ge \varepsilon, \end{cases}$$

where ε is so small that make ϕ_ε look like discontinuous. With the initial condition $\phi_\varepsilon(x)$ one can check that characteristics fill the whole space, as by letting $\varepsilon \to 0$, we obtain another solution to the original PDE

$$u(x) = \begin{cases} 0, & x \le 0, \\ x/t, & ,0 < x < t, \\ 1, & x \ge t. \end{cases}$$

This is the so-called *rarefaction solution.*

2. Shocks Now we assume the initial condition takes the form

$$\phi(x) = \begin{cases} 1, & x < 0, \\ 0, & x \ge 0. \end{cases}$$

Since $\phi(x)$ is not increasing, characteristics will intersect. It is not hard to see that for any fixed $k \in (0, 1)$, the following function

$$u(t, x) = \begin{cases} 1, & x < kt, \\ 0, & x \ge kt. \end{cases}$$

is a solution, with singularity only on the curve $x = kt$.

Again, not all k corresponds to physical solution. The previous trick of smoothing ϕ no longer help. To determine the correct value of k, we need to understand the effect of collision, which is not quite modeled by this equation.

We will not dive deep into the theory at this moment, but we will mention two things.

First, the correct way of smoothing the PDE is to introduce the viscous term:

$$\partial_t u + u \cdot \partial_x u = \varepsilon \partial_{xx} u.$$

As we will see, the appearance of $\varepsilon \partial_{xx} u$ will make possible the existence of classical solution. The added term $\partial_{xx} u$ represents the friction force, a term ignored when deriving the Burgers equation but correctly handles the intersection of characteristics, the collision. By letting $\varepsilon \to 0$, one may get the unique physical solution.

Second, the correct answer is $k = 1/2$. The interface $x = kt$ is called shocks, where particle of velocity of 0 and 1 meet and stick together. Essentially by conservation of momentum, the shock will travel at velocity $\frac{1}{2}(1 + 0) = 1/2$, which gives the physically correct value of k.

A comprehensive study of the first-order transport equation needs a good understanding of the second-order diffusion equation. This should be a good motivation for the next section.

3 Heat equation

The heat equation takes the form

$$\begin{cases} \partial_t u = \Delta u, & t > 0, x \in \Omega, \\ u(0,x) = u_0(x), & x \in \Omega, \end{cases}$$

plus some boundary condition. We will take this opportunity to introduce three basic types of boundary condition in the PDE theory. In the context of the heat equation, all these boundary conditions have concrete physical meaning.

Dirichlet boundary condition

$$u\big|_{\partial\Omega} = \mu.$$

This means that the temperature at the boundary is fixed, like a thermal bath or in the ice water.

Neumann boundary condition

$$\frac{\partial u}{\partial n}\Big|_{\partial\Omega} = 0.$$

This models the insulation, where there is no heat flux across the boundary.

Mixed (Robin) boundary condition

$$-k\frac{\partial u}{\partial n} = H\big(u(t,x) - \mu(t,x)\big), \quad x \in \Omega, t > 0.$$

Physically the parameter k and H should be positive: the LHS is the heat flux across the boundary, the RHS is difference of the internal temperature and the surrounding temperature. In the limit $k \downarrow 0$, this converges to the Dirichlet boundary condition, where the heat transfer is instant and the internal and external temperature is identical. In the limit $H \downarrow 0$, there is no heat flux at the boundary and this is the Neumann boundary condition. A more general way to write the mixed boundary condition is

$$\alpha u + \beta\frac{\partial u}{\partial n} = \mu,$$

where $\alpha, \beta \in \mathbb{R}$.

Take the Dirichlet boundary condition as an example, we will present the definition of a *classical solution.*

Definition 3.1 *Let $\Omega \subset \mathbb{R}^d$ be a domain with continuous boundary. A classical solution to the PDE*

$$\begin{cases} \partial_t u = \Delta u, & t > 0, \ x \in \Omega, \\ u(t,x) = \mu(t,x), & t \geq 0, \ x \in \partial\Omega, \\ u(0,x) = u_0(x), & x \in \Omega \end{cases}$$

is a function $u \in \mathcal{C}([0,\infty) \times \bar{\Omega}) \cap \mathcal{C}^{1,2}((0,\infty) \times \Omega)$ that satisfies the equation and the boundary/initial condition.

For Neumann and mixed boundary conditions, the domain should have $\mathcal{C}^1$-boundary in order to define the normal derivative $\partial u/\partial n$.

In this section we will focus on the following aspects of the heat equations, each of which will lead to a set of tools to study the equation:

- linear equation,
- Fourier transform,
- smoothing effect of Δ,
- maximum principle/energy method.

3.1 Energy method: first proof of uniqueness

Consider the heat equation

$$\begin{cases} \partial_t u = \Delta u + f, & t > 0,\ x \in \Omega, \\ u(t,x) = g(x), & t > 0,\ x \in \partial\Omega, \\ u(0,x) = h(x), & x \in \Omega. \end{cases} \tag{3.1}$$

Using linearity, we have the *principle of superposition.*

Theorem 3.1 *If $u_i \in \mathcal{C}([0,\infty) \times \bar{\Omega}) \cap \mathcal{C}^{1,2}((0,\infty) \times \Omega)$, $i = 1,2$, are classical solutions to (3.1) with data (f_i, g_i, h_i), then $\alpha u_1 + \beta u_2$ is a classical solution to (3.1) with data $\big(\alpha f_1 + \beta f_2, \alpha g_1 + \beta g_2, \alpha h_1 + \beta h_2\big)$.*

Proof: It follows from the linearity of the operators ∂_t and Δ:

$$\begin{aligned} \partial_t(\alpha u_1 + \beta u_2) &= \alpha \partial_t u_1 + \beta \partial_t u_2, \\ \Delta(\alpha u_1 + \beta u_2) &= \alpha \Delta u_1 + \beta \Delta u_2. \end{aligned}$$

□

Next, we will give a proof of the uniqueness of the heat equation solution via the *energy method.*

Theorem 3.2 *Let Ω be a bounded $\mathcal{C}^1$ domain. Then (3.1) has a unique classical solution.*

Proof: By Theorem 3.1, it suffices to show that the only classical solution to

$$\begin{cases} \partial_t u = \Delta u, \\ u(0,x) = 0, u\big|_{\partial\Omega} = 0, \end{cases} \tag{3.2}$$

is $u \equiv 0$. Indeed, if u_1 and u_2 are classical solutions to (3.1), then $u = u_1 - u_2$ is a classical solution to (3.2).

Let $u \in \mathcal{C}^{1,2}((0,\infty) \times \Omega) \cap \mathcal{C}([0,\infty) \times \bar{\Omega})$ solve (3.2). Let

$$f(t) = \int_\Omega |\nabla u|^2(t,x)\, dx.$$

Since Ω is bounded, $f(t)$ is finite and well-defined. Then

$$\begin{aligned} f'(t) &= 2 \int_\Omega \nabla(\partial_t u) \cdot \nabla u \, dx \\ &= 2 \int_\Omega (-\partial_t u) \Delta u + \int_{\partial\Omega} \partial_t u \cdot \frac{\partial u}{\partial n} \, dS \\ &= -2 \int_\Omega |\partial_t u|^2 \le 0. \end{aligned}$$

But $f(0) = 0$ and $f(t) \ge 0$ by definition. Hence, $f(t) \equiv 0$ for $t \ge 0$. This implies $\nabla u(t,x) \equiv 0$ for all (t,x). Since Ω is connected, $u(t,\cdot)$ must be constant in Ω. Since $u(t,\cdot) \in \mathcal{C}(\bar{\Omega})$ and $u\big|_{\partial\Omega} = 0$, we have $u \equiv 0$. This completes the proof. □

If the domain Ω is unbounded, additional conditions need to be imposed to guarantee the uniqueness. Let us consider $\Omega = \mathbb{R}^d$. The question is whether $u = 0$ is the unique solution to the PDE

$$\begin{cases} \partial_t u = \Delta u, & t > 0,\ x \in \mathbb{R}^d, \\ u(0,x) = 0, & x \in \mathbb{R}^d. \end{cases}$$

For the energy method to go through, one needs $\nabla u \in L^2(\mathbb{R}^d)$. A weaker condition is that

$$|u(t,x)| \le e^{c|x|^2}, \quad \forall t > 0,$$

for some $c > 0$. This growth condition is optional since the so-called Tychonov solution will be a counter-example in the absence of the growth condition. But the proof can not be done with the energy method.

3.2 Heat equation on the whole space and Fourier transform

We recall that the operator Δ is a linear operator on functions:

$$\Delta(\alpha f + \beta g) = \alpha \Delta f + \beta \Delta g, \quad \forall f, g \in \mathcal{D}(\Delta).$$

Let us compare the heat equation $\partial_t u = \Delta u$ with the linear ODE system with constant coefficients:

$$\dot{x}(t) = Ax(t), \quad A \in \mathbb{R}^{d\times d}. \tag{3.3}$$

Assume that the matrix A can be diagonalized:

$$A = P^{-1}\Lambda P, \quad P \in O(d),\ \Lambda = \mathrm{diag}\{\lambda_1, \dots, \lambda_d\}.$$

Then $y(t) = Px(t)$ solves $\dot{y}(t) = \Lambda y(t)$, whose solution is given by

$$y(t) = \left(e^{\lambda_1 t} y_1(0), \dots, e^{\lambda_d t} y_d(0)\right)^T.$$

For the heat equation, the Laplacian Δ can be diagonalized by the Fourier transform $\mathcal{F}$ and its inverse $\mathcal{F}^{-1}$.

Definition 3.2 *Let $f \in L^1(\mathbb{R}^d)$. Its Fourier transform $\hat{f} = \mathcal{F}f$ and inverse Fourier transform $\check{f} = \mathcal{F}^{-1}f$ are*

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

3.2.1 Properties of Fourier transform

Linearity:

$$(\alpha f + \beta g)^\wedge = \alpha \hat{f} + \beta \hat{g}, \quad \forall f, g \in L^1(\mathbb{R}^d),\ \alpha, \beta \in \mathbb{C}.$$

Translation: for $k \in \mathbb{R}^d$,

$$\big(f(\cdot - k)\big)^\wedge = e^{-2\pi i \xi\cdot k}\hat{f}(\xi).$$

Proof: Denote the LHS by $\hat{g}(\xi)$. We have

$$\begin{aligned}\hat{g}(\xi) &= \int e^{2\pi i \xi\cdot x} f(x-k)\,dx \\ &= \int e^{-2\pi i \xi\cdot (y+k)} f(y)\,dy \\ &= e^{-2\pi i \xi\cdot k}\int e^{-2\pi i \xi\cdot y} f(y)\,dy = e^{-2\pi i \xi\cdot k}\hat{f}(\xi),\end{aligned}$$

as desired. $\square$

Dilation: for $k \in \mathbb{R}^d \setminus \{0\}$,

$$\big(f(k\cdot)\big)^\wedge = \frac{1}{|k|^d}\hat{f}\Big(\frac{\xi}{|k|}\Big).$$

Proof: Denote the LHS by $\hat{g}(\xi)$. We have

$$\begin{aligned}\hat{g}(\xi) &= \int_{\mathbb{R}^d} e^{-2\pi i \frac{\xi}{|k|}\cdot |k| x} f(|k|x)\, dx \\ &= \frac{1}{|k|^d}\int_{\mathbb{R}^d} e^{-2\pi i \frac{\xi}{|k|}\cdot y} f(y)\, dy \\ &= \frac{1}{|k|^d}\hat{f}(\xi/|k|).\end{aligned}$$

$\square$

Symmetry: $\big(f(x)\big)^\vee = \hat{f}(-\xi)$.

Derivative: $(d=1)$ if $f', f \in L^1(\mathbb{R})$ and $f' \in \mathcal{C}(\mathbb{R})$, then

$$(f')^\wedge = (2\pi i\xi)\hat{f}(\xi).$$

Proof: For $N > 0$, using integration by parts we have

$$\int_{-N}^{N} e^{-2\pi i\xi x} f'(x)\, dx = (2\pi i\xi)\int_{-N}^{N} e^{-2\pi i\xi x} f(x)\, dx + e^{-2\pi i\xi} f(x)\Big|_{-N}^{N}.$$

Since $f' \in \mathcal{C}(\mathbb{R}) \cap L^1(\mathbb{R})$, the limits $\lim\limits_{x\to\pm\infty} f(x)$ exists. Since $f(x) \in L^1(\mathbb{R}) \cap \mathcal{C}(\mathbb{R})$, it follows that $\lim\limits_{|x|\to\infty} f(x) = 0$, and hence the last term in the last display goes to zero as $N \to \infty$. The desired conclusion follows. $\square$

For $d > 1$, a similar argument shows that

$$(\partial_{x_j} f)^\wedge = (2\pi i\xi_j)\hat{f}(\xi).$$

We can generalize such results to higher-order derivatives. For this purpose we introduce the multi-index notation. Let $\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{N}^d$. We define

$$|\alpha| := \alpha_1 + \ldots + \alpha_d, \quad x^\alpha := x_1^{\alpha_1}\cdots x_d^{\alpha_d}, \quad D^\alpha f := \partial_{x_1}^{\alpha_1}\cdots\partial_{x_d}^{\alpha_d} f.$$

Then for all $f \in \mathcal{C}_c^\infty(\mathbb{R}^d)$ (smooth with compact supports),

$$(D^\alpha f)^\wedge = (2\pi i)^{|\alpha|}\xi^\alpha \hat{f}(\xi).$$

Convolution: for all $f, g \in L^1(\mathbb{R}^d)$,

$$(f * g)^\wedge = \hat{f}(\xi)\hat{g}(\xi), \tag{3.4}$$

where the convolution $f * g$ is defined as

$$(f * g)(x) := \int_{\mathbb{R}^d} f(x-y)g(y)\, dy = \int_{\mathbb{R}^d} f(y)g(x-y)\, dy.$$

First, if $f, g \in L^1(\mathbb{R}^d)$, then $f * g \in L^1(\mathbb{R}^d)$, as the following lemma shows.

Lemma 3.3 (Special case of Young's inequality) *Let $f, g \in L^1(\mathbb{R}^d)$. Then*

$$\|f * g\|_{L^1} \leq \|f\|_{L^1} \|g\|_{L^1}.$$

Proof: By Fubini, we have

$$\begin{aligned}
\int_{\mathbb{R}^d} |f * g(x)|\, dx &\leq \int_{\mathbb{R}^d} dx \int_{\mathbb{R}^d} |f(x-y)||g(y)|\, dy \\
&\leq \int_{\mathbb{R}^d} |g(y)| \cdot \int_{\mathbb{R}^d} |f(x-y)| \\
&= \int_{\mathbb{R}^d} dy\, |g(y)| \cdot \|f\|_{L^1} = \|g\|_{L^1} \|f\|_{L^1}.
\end{aligned}$$

$\square$ **Proof of (3.4):** We have

$$\begin{aligned}
(f * g)^\wedge(\xi) &= \int_{\mathbb{R}^d} e^{-2\pi i \xi \cdot x}\, dx \int_{\mathbb{R}^d} f(x-y) g(y)\, dy \\
&= \int_{\mathbb{R}^d} e^{-2\pi i \xi \cdot (x-y)} f(x-y)\, dx \cdot \int_{\mathbb{R}^d} e^{-2\pi i \xi \cdot y} g(y)\, dy \\
&= \hat{f}(\xi) \cdot \hat{g}(\xi).
\end{aligned}$$

$\square$

3.2.2 Fourier transform of Gaussians

We can explicitly compute the Fourier transform of some functions. Below is an important example.

$$\left(\frac{1}{\sqrt{4\pi}} e^{-\frac{x^2}{4}} \right)^\wedge = e^{-4\pi^2 \xi^2}. \tag{3.5}$$

More generally, for $a > 0$,

$$\left(\frac{1}{\sqrt{4\pi a}} e^{-\frac{x^2}{4a}} \right)^\wedge = e^{-4\pi^2 a^2 \xi^2}.$$

We recall that the density of the normal distribution $\mathcal{N}(0, \sqrt{2})$ is $(4\pi)^{-1/2} e^{-x^2/4}$, and hence

$$\int_{\mathbb{R}} \frac{1}{\sqrt{4\pi}} e^{-x^2/4}\, dx = 1. \tag{3.6}$$

To prove (3.5), we have

$$\int_{\mathbb{R}} \frac{1}{\sqrt{4\pi}} e^{-\frac{x^2}{4} - 2\pi i x \xi}\, dx = \int_{\mathbb{R}} \frac{1}{\sqrt{4\pi}} e^{-\frac{1}{4}(x + 4\pi i \xi)^2 - 4\pi^2 \xi^2}.$$

It suffices to show that

$$\int_{\mathbb{R}} \frac{1}{\sqrt{4\pi}} e^{-\frac{1}{4}(x+b)^2}\, dx = 1 \tag{3.7}$$

for $b = 4\pi i \xi$.

If $b \in \mathbb{R}$, then (3.7) follows from (3.6) by a change of variable $y = x + b$. For a complex number b, we need to use some complex analysis to justify this identity.

Let

$$g(z) = \frac{1}{\sqrt{4\pi}} e^{-z^2/4}, \quad z \in \mathbb{C}.$$

Then $g(z)$ is an analytic function on the complex plane. We consider the closed contour

$$\gamma^L = \gamma_1 \cup \gamma_2 \cup \gamma_3 \cup \gamma_4,$$

where

$$\gamma_1 = \{x : x \in [-L, L]\}, \qquad \gamma_2 = \{L + iy : y \in [0, 4\pi\xi]\},$$
$$\gamma_3 = \{x + 4\pi\xi i : x \in [-L, L]\}, \qquad \gamma_4 = \{-L + iy : y \in [0, 4\pi\xi]\}.$$

By Cauchy Theorem, $\int_{\gamma^L} g(z)\, dz$ for all L. Letting $L \to \infty$, we have

$$\int_{\gamma_1} g(z)\, dz \to \int_{-\infty}^{\infty} \frac{1}{\sqrt{4\pi}} e^{-x^2/4}\, dx = 1,$$
$$\int_{\gamma_3} g(z)\, dz \to -\int_{-\infty}^{\infty} \frac{1}{\sqrt{4\pi}} e^{-(x+4\pi i\xi)^2/4}\, dx,$$

and

$$\int_{\gamma_2} g(z)\, dz, \int_{\gamma_4} g(z)\, dz \to 0.$$

This proves (3.7).

We have the following corollary in dimension $d > 1$.

Lemma 3.4 *For $a > 0$,*

$$\left(\frac{1}{(4\pi a)^{d/2}} e^{-\frac{|x|^2}{4}}\right)^{\wedge} = e^{-4\pi^2|\xi|^2 a}.$$

Proof: We have

$$\begin{aligned}
\int_{\mathbb{R}^d} \frac{1}{(4\pi a)^{d/2}} e^{-|x|^2/4} e^{-2\pi i x\cdot\xi}\, dx &= \prod_{j=1}^{d} \int_{\mathbb{R}} \frac{1}{\sqrt{4\pi a}} e^{-x_j^2/4} e^{-2\pi i x_j \xi_j}\, dx_j \\
&= \prod_{j=1}^{d} e^{-4\pi^2 \xi_j^2 a} \\
&= e^{-4\pi^2|\xi|^2 a}.
\end{aligned}$$

□

3.2.3 Cauchy problem of the heat equation

To solve the heat equation on the whole space

$$\partial_t u = \Delta u, \quad u(0, x) = \phi(x),$$

we consider the Fourier transform of the solution in the x-variable

$$\hat{u}(t, \xi) = \mathcal{F}[u(t, \cdot)].$$

Then $\hat{u}$ solves

$$\begin{cases} \partial_t \hat{u}(t, \xi) = -4\pi^2|\xi|^2 \hat{u}(t, \xi), \\ \hat{u}(0, \xi) = \hat{\phi}(\xi). \end{cases}$$

For a fixed ξ, $\hat{u}(t,\xi)$ solves a linear ODE, whose solution is given by

$$\hat{u}(t,\xi) = \hat{\phi}(\xi)e^{-4\pi^2|\xi|^2 t}.$$

Hence,

$$u(t,\cdot) = G_t * \phi,$$

where

$$G_t(x) = \frac{1}{(4\pi t)^{d/2}} e^{-|x|^2/4t}.$$

The function $G_t(x)$ is called the *fundamental solution* of the heat equation.

Consider the Cauchy problem of the heat equation on the whole space:

$$\begin{cases} \partial_t u = \Delta u, & t > 0,\ x \in \mathbb{R}^d, \\ u(0,x) = \phi(x), & x \in \mathbb{R}^d. \end{cases} \tag{3.8}$$

Using Fourier transform, with some extra effort we can show that $G_t * \phi$ solves (3.8), provided that $\phi \in L^1(\mathbb{R})$. Since $G_t(\cdot)$ decays very fast at ∞, this still holds for more general ϕ. Below is an example of such result, for which we will give a direct proof.

Theorem 3.5 *Assume $\phi \in \mathcal{C}(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$. Let $u(t,x) = (G_t * \phi)(x)$. Then*

1. *$u \in \mathcal{C}^\infty\big((0,\infty) \times \mathbb{R}^d\big)$ and $\partial_t u = \Delta u$ for $t > 0$ and $x \in \mathbb{R}^d$.*
2. *For all $x^0 \in \mathbb{R}^d$,*

$$\lim_{t \downarrow 0, x \to x^0} u(t,x) = \phi(x^0). \tag{3.9}$$

We will cite the following result from real analysis without proof.

Lemma 3.6 (Dominated Convergence Theorem) *Let $f_n \in L^1(\mathbb{R}^d)$ satisfying $|f_n| \le g$ for some $g \in L^1(\mathbb{R}^d)$. If $f_n \to f$ a.e., then*

$$\lim_{n\to\infty} \int_{\mathbb{R}^d} f_n(x)\,dx = \int_{\mathbb{R}^d} \lim_{n\to\infty} f_n(x)\,dx = \int_{\mathbb{R}^d} f(x)\,dx.$$

Lemma 3.7 *Let $f, g \in L^1(\mathbb{R}^d)$. If $\partial_{x_j} f \in \mathcal{C}(\mathbb{R}^d) \cap L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$, then $\partial_{x_j}(g * f) = g * (\partial_{x_j} f)$.*

Proof: Let e_j be the unit vector in the j-th coordinate. We have

$$\frac{(f*g)(x+he_j) - (f*g)(x)}{h} = \int_{\mathbb{R}^d} \frac{1}{h}[f(x+he_j-y) - f(x-y)]g(y)\,dy.$$

By Mean Value Theorem, The integrand is bounded by

$$\sup|\partial_{x_j} f| \cdot g(y) \in L^1(\mathbb{R}^d).$$

Then by Lemma 3.6, we have

$$\begin{aligned} \partial_{x_j}(f*g)(x) &= \lim_{h\to 0} \frac{(f*g)(x+he_j) - (f*g)(x)}{h} \\ &= \int_{\mathbb{R}^d} \lim_{h\to 0} \frac{1}{h}[f(x+he_j-y) - f(x-y)]g(y) \\ &= (\partial_{x_j} f) * g(x), \end{aligned}$$

as desired. $\square$

By direct computation one can check $f = D^\alpha G_t$ satisfies the condition in Lemma 3.7, and hence $D^\alpha u = (D^\alpha G_t) * g$. This implies $u \in \mathcal{C}^\infty$.

To show that $\partial_t u = \Delta u$, by Lemma 3.7, it suffices to show that $(\partial_t - \Delta)G = 0$. Indeed,

$$\partial_t G_t(x) = \big(-\frac{d}{2}t + \frac{|x|^2}{4t^2}\big)G_t(x),$$
$$\partial_{x_j} G_t(x) = -\frac{x_j}{2t} \cdot G_t(x),$$
$$\partial_{x_j x_j} G_t(x) = \Big[-\frac{1}{2t} + \frac{x_j^2}{4t^2}\Big]G_t(x),$$

so $G_t(x)$ satisfies the heat equation.

Finally, let us show that $u = G_t * \phi$ satisfies the initial condition in the sense of (3.9). Noting that $\int_{\mathbb{R}^d} G_t(x)\,dx = 1$, we have

$$\Big|u(t,x) - \phi(x^0)\Big| = \Big|\int_{\mathbb{R}^d} G_t(y)\big(\phi(x-y) - \phi(x^0)\big)\,dy\Big|$$
$$\int_{|y|\geq\varepsilon} G_t(y)(|\phi(x-y)| + |\phi(x^0)|)\,dy + \int_{\{|y|\leq\varepsilon\}} G_t(y)|\phi(x-y) - \phi(x^0)|\,dy$$
$$\leq 2\sup|\phi| \cdot \int_{|y|\geq\varepsilon} G_t(y)\,dy + \sup_{|y|\leq\varepsilon}|\phi(x-y) - \phi(x^0)|.$$

Since $G_t(y) = t^{-d/2}G_1(y/\sqrt{t})$, we have

$$\int_{|y|\geq\varepsilon} G_t(y)\,dy = \int_{|z|\geq\varepsilon/\sqrt{t}} G_1(z)\,dz \to 0, \quad t\downarrow 0.$$

Therefore,

$$\limsup_{t\downarrow 0,\ x\to x^0} |u(t,x) - \phi(x^0)| \leq \sup_{|y|\leq\varepsilon}|\phi(x^0 - y) - \phi(x^0)|.$$

Since ε is arbitrary and ϕ is continuous, the LHS limit must be zero. This completes the proof.

In fact, we have use a general result about the *approximate identity* in the proof above.

Lemma 3.8 *Let $\{k_n(x)\}$ be non-negative and continuous functions. Assume that*

1. $\int_{\mathbb{R}^d} k_n(x)\,dx = 1$ *for all n;*
2. $\lim_{n\to\infty}\int_{|x|\geq\varepsilon} k_n(x)\,dx = 0$ *for all $\varepsilon > 0$.*

Then for all g bounded and continuous, we have

$$\lim_{n\to\infty}\int_{\mathbb{R}^d} k_n(x)g(x)\,dx = g(0).$$

We call such $\{k_n(x)\}$ an approximate identity, which gives a mathematical meaning to the Dirac δ-function. The two conditions are also in fact necessary conditions, see Lax, Chap 11.

Proof: Assume that $|g(x)| \leq M$. We have

$$\Big|\int_{\mathbb{R}^d} k_n(x)g(x)\,dx - g(0)\Big| = \Big|\int_{\mathbb{R}^d} k_n(x)\big(g(x) - g(0)\big)\,dx\Big|$$
$$\leq \int_{|x|\leq\varepsilon} k_n(x)|g(x) - g(0)|\,dx + \int_{|x|>\varepsilon} k_n(x)\big(|g(x)| + |g(0)|\big)\,dx$$
$$\leq \sup_{|x|\leq\varepsilon}|g(x) - g(0)| + 2M\int_{|x|>\varepsilon} k_n(x)\,dx.$$

Using the second condition and taking $\limsup$, we have

$$\limsup_{n\to\infty} \text{LHS} \le \sup_{|x|\le\varepsilon} |g(x) - g(0)|.$$

Since ε is arbitrary, and g is continuous at 0, the limit at LHS must be 0. $\square$

The heat kernel $\{G_t(x)\}_{t>0}$ gives an approximate identity as $t \downarrow 0$.

3.2.4 Fundamental solution and derivation from scaling symmetry

The fundamental solution $G(t,x) = G_t(x)$ solves the following Cauchy problem

$$\begin{cases} \partial_t G = \Delta G, \\ G(0,x) = \delta(x), \end{cases} \tag{3.10}$$

where δ is a *generalized function* that satisfies

$$\delta * \phi = \phi, \quad \forall \phi \in \mathcal{C}(\mathbb{R}).$$

Physists usually think of the function $\delta(x)$ as

$$\delta(x) = \begin{cases} \infty, & x = 0, \\ 0, & x \ne 0, \end{cases}$$

so that $\int_{\mathbb{R}^d} \delta(x)\,dx = 1$. For us, we could think of $\delta(x)$ as the limit of an approximate identity. Indeed,

$$\lim_{t\downarrow 0} G_t(x) = \begin{cases} \infty, & x = 0, \\ 0, & x \ne 0. \end{cases}$$

We will give a second derivation of the solution to (3.10), using scaling symmetry of the heat equation. For simplicity let us assume $d = 1$.

We seek a solution to (3.10) invariant under the transform

$$u(t,x) \mapsto u_\lambda := \lambda^\alpha u(\lambda t, \lambda^\beta x).$$

Letting $u = u_\lambda$, we obtain

$$u(t,x) = \frac{1}{t^\alpha} v(\frac{x}{t^\beta}), \quad v(y) = u(1,y).$$

Using the expression, we have

$$\partial_t u = -\alpha t^{-\alpha-1} v(xt^{-\beta}) - \beta t^{-\alpha} x t^{-\beta} v'(xt^{-\beta}), \quad \partial_{xx} u = t^{-\alpha} \cdot t^{-2\beta} v''(xt^{-\beta}).$$

From the heat equation $\partial_t u = \partial_{xx} u$, the power in t must agree, so we have

$$\alpha + 1 = \alpha + 2\beta \implies \beta = 1/2.$$

and the function v must solve

$$-\alpha v(r) - \beta r v'(r) = v''(r).$$

To fix α, we assume the initial condition is also invariant, that is

$$u_\lambda(0,x) = \lambda^\alpha u(0,x) = \lambda^\alpha \delta(\lambda^{1/2} x) = \delta(x).$$

Using the fact that $\int \delta(x)\,dx = 1$, we obtain $\alpha = 1/2$.

So v solves

$$\frac{1}{2}v + \frac{1}{2}rv' + v'' = 0.$$

Integrating once, we obtain

$$\frac{1}{2}rv + v' = \text{const} = 0,$$

where we fix the constant assuming

$$\lim_{r\to\infty} v(r) = \lim_{r\to\infty} v'(r) = 0.$$

Finally, from $v' = -\frac{1}{2}rv$, we obtain

$$v = Ce^{-\frac{r^2}{4}}.$$

3.2.5 Understand the fundamental solution

The solution to the heat equation on the $\mathbb{R}$ can be written as

$$u(t,x) = \int G_t(x-y)\phi(y)\,dy.$$

Let $\Gamma(t,x;s,y) = G_{t-s}(x-y)$. For fixed (s,y), $\Gamma(\cdot,\cdot;s,y)$ solves

$$\begin{cases} \partial_t\Gamma - \Delta\Gamma = 0, & t > s,\ x \in \mathbb{R}, \\ \lim\limits_{t\downarrow s} \Gamma(t,x;s,y) = \delta(x-y). \end{cases}$$

The solution $\Gamma(\cdot,\cdot;s,y)$ can be thought of as the HE solution of placing a heat source at location y at time s. $\Gamma(t,x;s,y)$ is called the *fundamental solution*.

We can also understand the role of Γ from the principle of superposition. If the heat equation has initial condition

$$\phi(x) = \int \delta(x-y)\phi(y)\,dy,$$

that is, a "linear combination" of δ-functions with weights given by ϕ, then the solution is also a linear combination of Γ with the same weights:

$$u(t,x) = \int \Gamma(t,x;s,y)\phi(y)\,dy.$$

We can compare this terminology with the same one from the ODE theory. Recall that for a constant coefficient linear ODE

$$\dot{x}(t) = Ax(t),$$

its fundamental solution is

$$\Phi(t) = e^{At} = \sum_{n=0}^{\infty} \frac{(At)^n}{n!},$$

so that for any initial condition x_0, the solution is given by $\Phi(t)x_0$.

From the explicit form of the solution, we have a few more observations. First, for all $\phi \geq 0$, we have

$$(G_t * \phi)(x) > 0, \quad \forall x,\ \forall t > 0.$$

This indicates that diffusion has infinite speed of propagation: consider $\phi(x)$ representing the density of particles, and $\phi(x) = \mathbb{1}_{(-\infty,0]}(x)$; then at $t > 0$, the particles immediately spread to the whole real line.

Second, the heat equation has a smoothing effect on the initial condition. For a general function f, the decay in its Fourier transform $\hat{f}$ implies differentiability in f, since

$$D^\alpha f = \left[(2\pi\xi)^\alpha \hat{f}\right]^\vee.$$

The heat equation, in the Fourier space, turns any function $\hat{\phi}$ into

$$\hat{\phi}(\xi) \mapsto e^{-4\pi^2|\xi|^2 t}\hat{\phi}(\xi),$$

which has super-exponential decay as long as $t > 0$. In other words, the heat equation solution becomes smooth at any positive time.

3.2.6 Duhamel's principle

We consider the non-homogeneous problem

$$\begin{cases} \partial_t u = \Delta u + f, & t > 0,\ x \in \mathbb{R}^d, \\ u(0,x) = \phi(x), & x = 0. \end{cases} \tag{3.11}$$

Again, we look at the analogous linear ODE system first.

To solve the non-homogeneous ODE system

$$\dot{x}(t) = Ax(t) + f(t), \quad x(0) = x_0,$$

we use *variation of constant*, writing the candidate solution as

$$x(t) = \Phi(t)c(t), \tag{3.12}$$

where $c(t)$ is a function to be determined, and $\Phi(t) = e^{At}$ is the fundamental solution, which solves the matrix equation

$$d\Phi(t) = A\Phi = \Phi A.$$

Differentiating the expression (3.12), we obtain

$$\dot{x}(t) = \dot{\Phi}(t)c(t) + \Phi(t)\dot{c}(t) = A\Phi(t)c(t) + \Phi(t)\dot{c}(t),$$

so c solves

$$\dot{c}(t) = \left[\Phi(t)\right]^{-1} f(t) = \Phi(-t)f(t).$$

Therefore,

$$c(t) = \int_0^t \Phi(-s)f(s)\,ds + c(0),$$

and hence

$$x(t) = \Phi(t)x_0 + \Phi(t)\int_0^t \Phi(-s)f(s)\,ds = \Phi(t)x_0 + \int_0^t \Phi(t-s)f(s)\,ds.$$

That is, the final solution is a combination of the effect of the non-homogeneous terms $f(s)$ from all times, where the source at time s evolves for a duration of $t - s$. This form of solution holds for much more general linear systems, and is refereed to as the *Duhamel's principle*.

For (3.11) we can formulate the following result.

Theorem 3.9 *Let $f \in \mathcal{C}^{1,2}\big((0,\infty)\times\mathbb{R}^d\big) \cap \mathcal{C}_c([0,\infty)\times\mathbb{R}^d)$ and $\phi \in \mathcal{C}_c(\mathbb{R}^d)$. Then*

$$u(t,x) = \int_{\mathbb{R}^d} \Gamma(t,x;0,y)\phi(y)\,dy + \int_0^t \Gamma(t,x;s,y)f(s,y)\,dyds \in \mathcal{C}^{1,2}\big((0,\infty)\times\mathbb{R}^d\big),$$

solves (3.11) with

$$\lim_{(t,x)\to(0,x^0)} u(t,x) = 0, \quad \forall x^0 \in \mathbb{R}^d.$$

Proof: Without loss of generality we can assume $\phi = 0$. Since $\Gamma(t,x;s,y) = G_{t-s}(x-y)$, using a change of variable $s \mapsto t-s$, $y \mapsto x-y$, we can write

$$u(t,x) = \int_0^t \int_{\mathbb{R}}^d G_s(y)f(t-s,x-y)\,dsdy.$$

Therefore, ∂_t, $\partial_{x_ix_j}$ can be passed to f under the integral, and hence $u \in \mathcal{C}^{1,2}\big((0,\infty)\times\mathbb{R}^d\big)$.

We have

$$\begin{aligned}\partial_t u - \Delta u &= \int_{\mathbb{R}^d} G_t(y)f(0,x-y)\,dy + \int_0^t \int_{\mathbb{R}^d} G_s(y)(\partial_t - \Delta_x)f(t-s,x-y)\,dsdy \\ &= K + \int_\varepsilon^t \dots + \int_0^\varepsilon \dots = K + I_1 + I_2.\end{aligned}$$

For I_2, we have

$$|I_2| \le \varepsilon\big(\|\partial_t f\|_{L^\infty} + \|D^2 f\|_{L^\infty}\big) \le C\varepsilon.$$

For I_1, using integration by parts we have

$$\begin{aligned}I_1 &= \int_\varepsilon^t \int_{\mathbb{R}^d} G_s(y)(-\partial_s - \Delta_y)f(t-s,x-y)\,dsdy \\ &= \int_\varepsilon^t \int_{\mathbb{R}^d} \big[\partial_s G_s(y) - \Delta_y G_s(y)\big] f(t-s,x-y)\,dsdy - K + \int_{\mathbb{R}^d} G_\varepsilon(y)f(t-\varepsilon,x-y)\,dy.\end{aligned}$$

The first term is 0 since $G_s(y)$ solves the HE for $s > 0$. The last term can be written as $(G_\varepsilon * f(t-\varepsilon,\cdot))(x)$, and we have

$$\begin{aligned}\|(G_\varepsilon * f(t,\cdot)) - f(t,\cdot)\|_{L^\infty} &\to 0, \\ \|G_\varepsilon * f(t-\varepsilon,\cdot) - G_\varepsilon * f(t,\cdot)\|_{L^\infty} &\le \|f(t-\varepsilon,\cdot) - f(t,\cdot)\|_{L^\infty}\|G_\varepsilon\|_{L^1} \\ &\to 0, \quad \varepsilon \downarrow 0.\end{aligned}$$

Here, the first line is due to that $\{G_\varepsilon\}$ is an approximate identity, the second line is by the Young's inequality, and the last line follows from the f is a continuous function with compact support.

Therefore,

$$\lim_{\varepsilon\to 0}(G_\varepsilon * f(t-\varepsilon,\cdot))(x) = f(t,x),$$

uniformly in x, and $\partial_t u = \Delta u + f$ for $t > 0$, $x \in \mathbb{R}^d$.

Finally,

$$\|u(t,\cdot)\|_{L^\infty} \le t\|f\|_{L^\infty} \to 0,$$

so the initial condition is satisfies. ∎

We can also use the Duhamel's principle in the Fourier space. Again assume $\phi = 0$. Let $\hat{u}(t,\xi) = [\mathcal{F}u(t,\cdot)](\xi)$ and $\hat{f}(t,\xi) = [\mathcal{F}f(t,\cdot)](\xi)$. Then $\hat{u}$ solves

$$\partial_t \hat{u}(t,\xi) = -4\pi^2|\xi|^2 t \cdot \hat{u}(t,\xi) + \hat{f}(t,\xi), \quad \hat{u}(t,0) = 0.$$

Then we have

$$\hat{u}(t,\xi) = \int_0^t e^{-4\pi^2|\xi|^2(t-s)}\hat{f}(s,\xi)\,ds,$$

so

$$u(t,x) = \int_0^t \big[G_{t-s} * f(s,\cdot)\big](x)\,ds.$$

Let

$$\Gamma(t,x;s,y) = \begin{cases} G_{t-s}(y-x), & t > s, \\ 0, & t \le s. \end{cases}$$

Then $\Gamma(\cdot,\cdot;s,y)$ solves

$$\begin{cases} (\partial_t - \Delta)\Gamma = \delta(t-s, x-y), \\ \Gamma(0,\cdot;x,y) = 0. \end{cases}$$

Hence $u(t,x) = \Gamma *_{(t,x)} f$ solves

$$(\partial_t - \Delta)u(t,x) = \Big(\big[(\partial_t - \Delta)\Gamma\big] *_{(t,x)} f\Big)(t,x) = \int \delta(t-s, x-y) f(s,y) = f(t,x).$$

3.3 Heat equation on bounded domains, separation of variables

3.3.1 Motivation

Consider the linear ODE system

$$\dot{x}(t) = Ax(t).$$

Assume that the matrix $A \in \mathbb{R}^{d\times d}$ can be diagonalized as $A = P\Lambda P^{-1}$, then $y = P^{-1}x$ solves $\dot{y} = \Lambda y(t)$, and hence

$$y(t) = \big(e^{\lambda_i t} y_i(0)\big).$$

Plugging in, we obtain

$$x(t) = [v_1 \cdots v_d]\Lambda[c_1 \dots c_d]^T = \sum_{i=1}^d c_i e^{\lambda_i t} \dot{v}_i.$$

As the principle of superposition suggest, if $x_i(t) = e^{\lambda_i t}\vec{v}_i$ are solutions, then their linear combinations

$$\sum_{i=1}^d c_i e^{\lambda_i t}\vec{v}_i$$

are also solutions. Here, $(\lambda_i, \vec{v}_i)$ are eigen-pairs of the matrix A.

Returning to the heat equation $\partial_t u = \Delta u$. What are the "eigenfunctions" of the Laplacian operator Δ? We have seen that

$$\Delta e^{2\pi i \xi\cdot x} = -4\pi^2|\xi|^2 e^{2\pi i\xi\cdot x}.$$

Thus, $(-4\pi^2|\xi|^2, e^{2\pi i\xi\cdot x})$ may be interpreted as eigenpairs. By applying the Fourier transform, the heat equation solution can be expressed as

$$u(t,x) = \int \hat{u}(t,\xi)e^{2\pi i\xi\cdot x}\,d\xi = \int \hat{u}(0,\xi)e^{-4\pi^2|\xi|^2 t}e^{2\pi i\xi\cdot x}\,d\xi,$$

which is analogous to the linear ODE solution $\sum_{j=1}^d c_j e^{\lambda_j t}\vec{v}_j$.

3.3.2 Spectral theory of the Laplacian operator

The characterization of $(-4\pi^2|\xi|^2, e^{2\pi i\xi\cdot x})$ as an eigenpair of the Laplacian on $\mathbb{R}^d$ is not entirely accurate, since the function $x \mapsto e^{2\pi i\xi\cdot x}$ does not belong to $L^2(\mathbb{R}^d)$, the standard space for spectral theory of linear operators. In fact, $\lambda = -4\pi^2|\xi|$ belongs to the *continuous spectrum* of the Laplacian, whereas eigenpairs are associated with the point spectrum of an operator.

One way to characterize the continuous spectrum is the following. For any $\lambda > 0$, for every $\varepsilon > 0$, there is $f_\varepsilon \in L^2(\mathbb{R}^d) \cap \mathcal{C}^2(\mathbb{R}^d)$ such that

$$\|\Delta f_\varepsilon + \lambda f_\varepsilon\|_{L^2} \leq \varepsilon.$$

The appearance of the continuous spectrum is essentially due to the infinite-dimensional nature of the operator.

However, the spectral theory of Δ on a bounded domain is much simpler. As an example, let Ω be a bounded $\mathcal{C}^1$-domain and consider the eigenvalue problem

$$\begin{cases} -\Delta u = \lambda u, & x \in \Omega \\ \alpha u + \beta \dfrac{\partial u}{\partial n} = 0, & x \in \partial\Omega. \end{cases} \tag{3.13}$$

Here, α and β are constants such that $\alpha^2 + \beta^2 > 0$. If (λ, u) satisfies the above equation, then it is called an eigenpair. The following holds.

Theorem 3.10 *1. $-\Delta$ is symmetric in $L^2(\Omega)$: $(-\Delta u, v)_{L^2} = (u, -\Delta v)_{L^2}$.*

2. All eigenvalues of $-\Delta$ are real; and if $\alpha, \beta \geq 0$, they are non-negative.

3. If (λ_1, u) and (λ_2, v) are two eigenpairs such that $\lambda_1 \neq \lambda_2$, then $(u, v)_{L^2} = 0$.

4. The eigenvalues are countable, and if ordered as $0 \leq \lambda_1 \leq \lambda_2 \leq \cdots$, then $\lim_{n\to\infty} \lambda_n = \infty$.

5. The eigenfunctions $\{w_k\}$ forms an orthonormal basis for $L^2(\Omega)$, that is, for any $f \in L^2(\Omega)$ satisfying the boundary condition, there are c_k such that

$$f(x) = L^2\text{-}\sum_{k=1}^{\infty} c_k w_k(x).$$

Remark 3.1 We will prove the first three items, which is analogous to the spectral theory of semi-positive definite matrices. The last two items requires deeper results from functional analysis.

Proof: We have

$$\int_\Omega (u\Delta v - v\Delta u)\, dx = \int_{\partial\Omega} (u\frac{\partial v}{\partial n} - v\frac{\partial u}{\partial n})\, dS.$$

Since (α, β) is a non-trivial solution of the linear system

$$\alpha u + \beta\frac{\partial u}{\partial n} = \alpha v + \beta\frac{\partial v}{\partial n} = 0,$$

we have

$$\det\begin{bmatrix} u & \frac{\partial u}{\partial n} \\ v & \frac{\partial v}{\partial n} \end{bmatrix} = u\frac{\partial v}{\partial n} - v\frac{\partial u}{\partial n} = 0.$$

This proves the first item.

Let (λ, u) be an eigenpair. Then

$$\lambda \int_\Omega |u|^2 = \lambda \int_\Omega \bar{u}(-\Delta u) = \int_\Omega (-\Delta \bar{u})u = \bar{\lambda} \int_\Omega |u|^2.$$

Hence $\lambda = \bar{\lambda}$, which implies $\lambda \in \mathbb{R}$. If $\alpha, \beta \geq 0$, then on $\partial\Omega$, we have $u \cdot \frac{\partial u}{\partial n} \geq 0$. Therefore,

$$\lambda \int_\Omega |u|^2 = \int_\Omega (-\Delta u)u = \int_\Omega \|\nabla u\|^2 - \int_{\partial\Omega} u\frac{\partial u}{\partial n} \geq 0.$$

So $\lambda \geq 0$.

Let (λ_1, u) and (λ_2, v) be two eigenpairs with $\lambda_1 \neq \lambda_2$. We have

$$\lambda_1 \int_\Omega uv = \int_\Omega (-\Delta u)v = \int_\Omega v(-\Delta u) + \int_{\partial\Omega} (u\frac{\partial v}{\partial n} - v\frac{\partial u}{\partial n}) = \lambda_2 \int_\Omega uv.$$

Hence, $\int_\Omega uv = 0$.

□

3.3.3 Separation of variables

Exact solutions to the eigenvalue problem (3.13) are difficult to obtain, unless the domain Ω is sufficiently simple such as an interval or a rectangle. In such cases, the final two items of Theorem 3.10 can be directly verified by theory of Fourier series. The method of solving linear PDEs using eigenfunction expansion is also known as *separation of variables*, which is typically formulated in a different way.

1. Set up

   We will illustrate how to solve the following PDE using separation of variables:

$$\begin{cases} \partial_t u = \partial_{xx} u, & t > 0,\ x \in (0, \ell), \\ u(0, x) = \phi(x), & x \in (0, \ell), \\ -\alpha_1 u'(0) + \beta_1 u(0) = 0, \\ \alpha_2 u'(\ell) + \beta_2 u(\ell) = 0. \end{cases} \tag{3.14}$$

   Here, $\ell > 0$ so that $\Omega = (0, \ell)$, and α_i, β_i are constants.

   Separation of variables usually consists of the following steps.

   Step 1: consider nontrivial solution of the form (where t, x are separated):

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

   Plugging it into the equation, we obtain

$$T'(t)X(t) = T(t)X''(x).$$

   Hence,

$$\frac{T'(t)}{T(t)} = \frac{X''(x)}{X(x)} =: -\lambda,$$

   which is a constant since the expression is independent of t and x. Given λ, the functions T and X satisfy different equations. For T, it solves

$$T'(t) = -\lambda T(t),$$

while for X, combined with the boundary condition, it solves the *Sturm-Liouville* problem

$$\begin{cases} X''(x) + \lambda X(x) = 0, \quad x \in (0, \ell), \\ -\alpha_1 X'(0) + \beta_1 X(0) = 0, \\ \alpha_2 X'(\ell) + \beta_2 X(\ell) = 0. \end{cases} \tag{3.15}$$

Step 2:: solve the Sturm-Liouville problem (3.15). This is the eigenvalue problem for ∂_{xx} on $(0, \ell)$ with the given boundary condition. Denote by (λ_n, X_n) all its eigenpairs, and let $T_n(t) = e^{-\lambda_n t} T_n(0)$.

Step 3: by principle of superposition, any linear combination of

$$u_n(t, x) = T_n(t) X_n(x)$$

will satisfy the equation and the boundary condition. We need to determine a correct combination so that the initial condition is also satisfied. This is possible since $\{X_n(x)\}$ is a basis in $L^2(0, \ell)$, so we have the decomposition

$$\phi(x) = \sum_{n=1}^{\infty} \phi_n X_n(x).$$

With this decomposition, the final solution to (3.14) is given by

$$u(t, x) = \sum_{n=1}^{\infty} \phi_n e^{-\lambda_n t} X_n(x).$$

2. Examples

Example 1: Consider the equation

$$\begin{cases} \partial_t u = \partial_{xx} u, \quad x \in (0, \ell), \\ u(t, \ell) = u(t, 0) = 0, \\ u(0, x) = \phi(x). \end{cases}$$

The correspondign Sturm-Liouville problem is

$$X''(x) + \lambda X = 0, \quad X(0) = X(\ell) = 0.$$

The general solution to the S-L problem is

$$X(x) = C_1 \sin(\sqrt{\lambda} x) + C_2 \cos(\sqrt{\lambda} x).$$

To match the boundary condition, we must have $C_2 = 0$ and $\sqrt{\lambda}\ell = n\pi$, so that $\lambda = \left(\frac{n\pi}{\ell}\right)^2$. Hence,

$$X_n(x) = \sin \frac{n\pi x}{\ell}.$$

The solution is then given by

$$u(t, x) = \sum_{n=1}^{\infty} c_n e^{-\left(\frac{n\pi}{\ell}\right)^2 t} \sin \frac{n\pi x}{\ell}.$$

The constants c_n are determined by

$$\phi(x) = \sum_{n=1}^{\infty} c_n \sin \frac{n\pi x}{\ell}.$$

As the theory of the Fourier series suggests, we should use the orthogonality of the basis function to find out c_n, that is,

$$\int_0^\ell \sin \frac{n\pi x}{\ell} \phi(x)\, dx = \int_0^\ell \sin \frac{n\pi x}{\ell} \sum_{m=1}^{\infty} c_m \sin \frac{m\pi x}{\ell}\, dx = c_n \int_0^\ell \sin^2 \frac{n\pi x}{\ell}\, dx = c_n \cdot \frac{\ell}{2}.$$

Hence,

$$c_n = \frac{2}{\ell} \int_0^\ell \sin \frac{n\pi x}{\ell} \phi(x)\, dx.$$

Example 2: Consider the equation

$$\begin{cases} \partial_t u = \partial_{xx} u, & x \in (0, \ell), \\ \partial_x u(t, \ell) = \partial_x u(t, 0) = 0, \\ u(0, x) = \phi(x). \end{cases}$$

The correspondign Sturm-Liouville problem is

$$X''(x) + \lambda X = 0, \quad X'(0) = X'(\ell) = 0.$$

The general solution to the S-L problem is

$$X(x) = C_1 \sin(\sqrt{\lambda} x) + C_2 \cos(\sqrt{\lambda} x).$$

To match the boundary condition, we must have $C_1 = 0$ and $\lambda = \left(\frac{n\pi}{\ell}\right)^2$. Hence,

$$X_n(x) = \cos \frac{n\pi x}{\ell}.$$

The solution is then given by

$$u(t, x) = \sum_{n=1}^{\infty} c_n e^{-\left(\frac{n\pi}{\ell}\right)^2 t} \cos \frac{n\pi x}{\ell}.$$

The constants c_n are given by

$$c_n = \begin{cases} \frac{1}{\ell} \int_0^\ell \phi(x)\, dx, & n = 0, \\ \frac{2}{\ell} \int_0^\ell \sin \frac{n\pi x}{\ell} \phi(x)\, dx, & n \geq 1. \end{cases}$$

Example 3: Consider the equation

$$\begin{cases} \partial_t u = \partial_{xx} u, & x \in (0, \ell), \\ u(t, \ell) = \partial_x u(t, 0) = 0, \\ u(0, x) = \phi(x). \end{cases}$$

The correspondign Sturm-Liouville problem is

$$X''(x) + \lambda X = 0, \quad X(0) = X'(\ell) = 0.$$

The general solution to the S-L problem is

$$X(x) = C_1 \sin(\sqrt{\lambda}x) + C_2 \cos(\sqrt{\lambda}x).$$

To match the boundary condition, we have $C_2 = 0$ and

$$\sqrt{\lambda}\ell = (n + 1/2)\pi \implies \lambda_n = \frac{(n+1/2)^2\pi^2}{\ell^2}, \quad n \geq 0.$$

Hence,

$$X_n(x) = \sin \frac{(n+1/2)\pi x}{\ell}.$$

The solution is then given by

$$u(t,x) = \sum_{n=0}^{\infty} c_n e^{-\left(\frac{(n+1/2)\pi}{\ell}\right)^2 t} \sin \frac{(n+1/2)\pi x}{\ell}.$$

The constants c_n are given by

$$c_n = \left[\int_0^\ell \sin^2 \frac{(n+1/2)\pi x}{\ell}\, dx\right]^{-1} \int_0^\ell \sin \frac{(n+1/2)\pi x}{\ell} \phi(x)\, dx = \frac{2}{\ell}\int_0^\ell \sin \frac{(n+1/2)\pi}{\ell}\phi(x)\, dx.$$

3.3.4 Non-homogeneous equation and Green's function

1. Duhamel's principle Consider the equation

$$\begin{cases} \partial_t u = \partial_{xx} u + f, & t > 0,\ x \in (0,\ell), \\ u(t,0) = u(t,\ell) = 0, & t > 0, \\ u(0,x) = \phi(x), & x \in (0,\ell). \end{cases}$$

To solve it, we expand $u(t,\cdot)$, $f(t,\cdot)$ and $\phi(\cdot)$ in the basis $\{\sin \frac{n\pi x}{\ell}\}$:

$$\begin{aligned} u(t,x) &= \sum_{n=1}^{\infty} T_n(t) \sin \frac{n\pi x}{\ell}, \\ f(t,x) &= \sum_{n=1}^{\infty} f_n(t) \sin \frac{n\pi x}{\ell}, \\ \phi(x) &= \sum_{n=1}^{\infty} \phi_n \sin \frac{n\pi x}{\ell}, \end{aligned}$$

where

$$\begin{aligned} f_n(t) &= \frac{2}{\ell}\int_0^\ell f(t,y) \sin \frac{n\pi y}{\ell}\, dy, \\ \phi_n &= \frac{2}{\ell}\int_0^\ell \phi(y) \sin \frac{n\pi y}{\ell}\, dy, \end{aligned}$$

Then $T_n(t)$ solves the ODE

$$T_n'(t) + \left(\frac{n\pi}{\ell}\right)^2 T_n(t) = f_n(t), \quad T_n(0) = \phi_n.$$

From linear ODE theory or Duhamel's principle, we have

$$T_n(t) = \phi_n e^{-(\frac{n\pi}{\ell})^2 t} + \int_0^t e^{-(\frac{n\pi}{\ell})^2(t-s)} f_n(s)\, ds.$$

Hence,

$$u(t,x) = \int_0^\ell \phi(y) G(t; x, y)\, dy + \int_0^t ds \int_0^\ell f(s,y) G(t-s; x, y)\, dy,$$

where

$$G(t; x, y) = \frac{2}{\ell} \sum_{n=1}^{\infty} \sin \frac{n\pi x}{\ell} \sin \frac{n\pi y}{\ell} e^{-(\frac{n\pi}{\ell})^2 t}. \tag{3.16}$$

The function G is called the *Green's function*. If we define

$$\Phi(t) : h(\cdot) \mapsto \big[\Phi(t) h\big](x) = \int_0^\ell G(t; x, y) h(y)\, dy,$$

then we can rewrite the above Duhamel's principle as

$$u(t, \cdot) = \Phi(t)\phi + \int_0^t \Phi(t-s) f(s, \cdot)\, ds.$$

2. Green's function We introduce

$$G(t, x; s, y) = \begin{cases} G(t-s; x, y), & t > s, \\ 0, & t \le s. \end{cases}$$

Then formally the Green's function $G(t, x; s, y)$ solves

$$\begin{cases} \partial_t G - \partial_{xx} G = \delta(t-s, y-x), & t > 0,\ x \in (0, \ell), \\ u(t,0) = u(t,\ell) = 0, & t > 0, \\ u(0,x) = 0. \end{cases}$$

The Green's function is similar to the fundamental solution, but the former satisfies additional some boundary conditions.

3. Properties of the Green's function

Using the explicity form of the Green's function (3.16), we will prove some of its properties. These properties still holds true for general domain and boundary condition, but the proof will be more difficult.

Symmetry: $G(t, x; s, y) = G(t, y; s, x)$. One can say that the influence of x at y is the same as the influence of y at x.

Smoothness: $G(t, x; s, y) \in \mathcal{C}^\infty$ and

$$(\partial_t - \partial_{xx}) G = (\partial_s + \partial_{yy}) G = 0.$$

Singularity at $t = 0$: for some constant $C > 0$,

$$|G(t, x; s, y)| \le \frac{C}{\sqrt{t-s}}.$$

For the fundamental solution $G_t(x) = \frac{1}{\sqrt{4\pi t}} e^{-\frac{x^2}{4t}}$, the same bound holds.

Proof: It suffices to show that

$$\frac{2}{\ell} \sum_{n=1}^{\infty} e^{-(\frac{n\pi}{\ell})^2 t} \le \frac{C}{\sqrt{t}}.$$

Indeed,

$$\begin{aligned}\sum_{n=1}^{\infty} e^{-an^2} &= \sum_{\sqrt{a}n \le 1} e^{-an^2} + \sum_{\sqrt{a}n >} e^{-an^2} \\ &\le \frac{1}{\sqrt{a}} + \sum_{n > 1/\sqrt{a}} e^{-n\sqrt{a}} \\ &\le \frac{1}{\sqrt{a}} + C \le \frac{C}{\sqrt{a}},\end{aligned}$$

provided that $a < 1$. Letting $a = -\pi^2 t/\ell^2$ completes the proof. $\square$

Initial condition: if $\phi \in \mathcal{C}^1[0, \ell]$ and $\phi(0) = \phi(\ell) = 0$, then

$$\lim_{t \to 0+} \int_0^\ell G(t; x, y)\phi(y)\, dy = \phi(x).$$

Proof: We have

$$\begin{aligned}\int_0^\ell G(t; x, y)\phi(y)\, dy &= \int_0^\ell \sum_{n=1}^{\infty} \frac{2}{\ell} \sin\frac{n\pi x}{\ell} \sin\frac{n\pi y}{\ell} e^{-(\frac{n\pi}{\ell})^2 t}\phi(y)\, dy \\ &= \sum_{n=1}^{\infty} \frac{2}{\ell} \int_0^\ell \sin\frac{n\pi x}{\ell} \sin\frac{n\pi y}{\ell} \phi(y)\, dy \\ &= \sum_{n=1}^{\infty} \phi_n \sin\frac{n\pi x}{\ell} e^{-(\frac{n\pi}{\ell})^2 t}.\end{aligned}$$

One can show that

$$\sum_{n=1}^{\infty} \phi_n \sin\frac{n\pi x}{\ell}$$

converges since $\phi \in \mathcal{C}^1$, and $e^{-(\frac{n\pi}{\ell})^2 t}$ is monotone in t. By Abel's test, the whole series converges uniformly in t, and hence

$$\lim_{t \to 0+} \sum_{n=1} \phi_n \sin\frac{n\pi x}{\ell} e^{-(\frac{n\pi}{\ell})^2 t} = \sum_{n=1} \phi_n \sin\frac{n\pi x}{\ell} \lim_{t \to 0+} e^{-(\frac{n\pi}{\ell})^2 t} = \phi(x).$$

$\square$ We only use the fact that the Fourier series conveges. A weaker sufficient condition may be ϕ being absolute continuous.

3.3.5 Non-homogeneous boundary conditions

Let us consider a heat equation on $(0,\ell)$ where the boundary condition is time-dependent and thus non-homogeneous:

$$\begin{cases} (\partial_t - \Delta)u = f, & t>0,\ x\in(0,\ell), \\ u(t,0) = g_1(t), & t>0, \\ u(t,\ell) = g_2(t), & t>0, \\ u(0,x) = \phi(x), & x\in(0,\ell). \end{cases} \tag{3.17}$$

Let

$$\tilde{u}(t,x) = u(t,x) - \frac{x}{\ell}g_2(t) + \frac{\ell-x}{\ell}g_1(t) =: u(t,x) - h(t,x).$$

Then $\tilde{u}$ solves (3.17) with

$$\tilde{f}(t,x) = f(t,x) - \left[\frac{x}{\ell}g_2'(t) + \frac{\ell-x}{\ell}g_1'(t)\right] = f(t,x) - \partial_t h(t,x),$$
$$\tilde{\phi}(x) = \phi(x) - \left[\frac{x}{\ell}g_2(0) + \frac{\ell-x}{\ell}g_1(0)\right] = \phi(x) - h(0,x).$$

For simplicity we assume $f=\phi=0$. Then

$$\begin{aligned} u(t,x) &= h(t,x) - \int_0^\ell G(t,x;0,y)h(0,y)\,dy - \int_0^t\int_0^\ell G(t,x;s,y)\partial_s h(s,y)\,dyds \quad = h(t,x) + \int_0^t\int_0^\ell \partial_s G(t,x;s,y)h(s \\ &= h(t,x) + \int_0^t\int_0^\ell -\partial_{yy}G(t,x;s,y)h(s,y)\,dsdy \\ &= h(t,x) + \int_0^t \partial_y G(t,x;s,y)\partial_y h(s,y)\big|_0^\ell\,ds \\ &= h(t,x) + \int_0^t \partial_y G(t,x;s,y)g_1(s) - \partial_y G(t,x;s,\ell)g_2(s)\,ds. \end{aligned}$$

We point out that the final solution is also a linear functional of the boundary data g_1 and g_2.

3.4 Maximum principle

3.4.1 Bounded domain

For a domain Ω and $T>0$, we introduce the *parabolic interior*

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

and the *parabolic boundary*

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

Theorem 3.11 (Weak maximum principle) *Let $u\in\mathcal{C}^{1,2}(\Omega_T)\cap C(\overline{\Omega_T})$. If*

$$\partial_t u(t,x) - \Delta u(t,x) \le 0, \quad x\in\Omega_T,$$

then

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

that is, the maximum on $\overline{\Omega_T}$ is acheived on the parabolic boundary.

The most application of the maximum principle is the uniqueness of the solution to the heat equation.

Theorem 3.12 *There is at most one solution $u \in \mathcal{C}^{1,2}(\Omega_T) \cap \mathcal{C}(\overline{\Omega_T})$ to the PDE*

$$\begin{cases} \partial_t u = \Delta u, & \Omega_T, \\ u\big|_{\partial\Omega} = g(t), & \partial\Omega, \\ u\big|_{t=0} = \phi. \end{cases}$$

Proof: Let u_1, u_2 be two solutions. Then $v = u_1 - u_2 \in \mathcal{C}^{1,2}(\Omega_T) \cap \mathcal{C}(\overline{\Omega_T})$ solves

$$\begin{cases} \partial_t v = \Delta v, & \Omega_T, \\ v\big|_{\Gamma_T} = 0. \end{cases}$$

By weak maximum principle,

$$\begin{cases} \max_{\overline{\Omega_T}} v \le \max_{\Gamma_T} v = 0, \\ \max_{\overline{\Omega_T}}(-v) \le \max_{\Gamma_T}(-v) = 0, \end{cases} \implies v \equiv 0 \text{ in } \overline{\Omega_T}.$$

$\square$

Now let us prove the weak maximum principle. **Proof:** First, let us assume that

$$\partial_t u - \Delta u < 0, \quad x \in \Omega_T. \tag{3.18}$$

Assume on the contrary that u achieves the maximum at $(t^*, x^*) \in \Omega_T$. Since $u(t^*, x^*) \ge u(t, x^*)$ for all $t < t^*$, we have $\partial_t u(t^*, x^*) \ge 0$. Also, since $u(t^*, x^*) \ge u(t^*, x)$ for all x, the Hessian of $u(t^*, \cdot)$ at $x = x^*$ is negative, and hence

$$\Delta u(t^*, x^*) \le 0.$$

Combination of these two inequalities contradicts with (3.18).

If the inequality is non-strict, for every $\varepsilon > 0$, let us consider $u_\varepsilon(t, x) = u(t, x) - t\varepsilon$. Then

$$\partial_t u_\varepsilon - \Delta u_\varepsilon = \partial_t u - \Delta u - \varepsilon < 0, \quad x \in \Omega_T,$$

so the weak maximum principle for u_ε implies

$$\max_{\overline{\Omega_T}} u_\varepsilon \le \max_{\Gamma_T} u_\varepsilon.$$

Taking $\varepsilon \to 0+$ we obtain the desired result. $\square$

3.4.2 Unbounded domain

We can use the maximum principle to obtain uniqueness of heat equation solution on unbounded domain.

Theorem 3.13 *If $u \in \mathcal{C}^{1,2}((0, \infty) \times \mathbb{R}^d) \cap \mathcal{C}([0, \infty) \times \mathbb{R}^d)$ solves*

$$\partial_t u = \Delta u, \quad u(0, x) = 0,$$

and satisfies

$$|u(t, x)| \le Ce^{Ax^2}$$

for some $A > 0$. Then $u = 0$.

Proof: Let $\Omega_{L,T} = (0,T] \times B_L(0)$, where $T < \frac{1}{4A}$. Let $\varepsilon > 0$ be such that $T + \varepsilon < \frac{1}{4A}$. We consider

$$v(t,x) = u(t,x) - \frac{\mu}{(T+\varepsilon-t)^{d/2}} e^{\frac{|x|^2}{4(T+\varepsilon-t)}}, \quad t \in [0,T].$$

Then $\partial_t v - \Delta v = 0$. On $\overline{\Omega_{L,T}}$ the weak maximum principle applies, so

$$\max_{\overline{\Omega_{L,T}}} v = \max_{\Gamma_{L,T}} v.$$

We notice two things. First, $v(0,x) \le u(0,x) = 0$. Second, for every $\mu > 0$, when $|x| = L$,

$$\begin{aligned} v(t,x) &\le Ce^{AL^2} - \frac{\mu}{(T+\varepsilon-t)^{d/2}} e^{\frac{L^2}{4(T+\varepsilon-t)}} \\ &\le Ce^{AL^2} - \frac{\mu}{(T+\varepsilon)^{d/2}} e^{\frac{L^2}{4(T+\varepsilon)}} \\ &\le 0, \end{aligned}$$

provided L is sufficiently large, since $\frac{1}{4(T+\varepsilon)} > A$. Hence, $v \le 0$ on $\Gamma_{L,T}$. This implies

$$u(t,x) \le \frac{\mu}{(T+\varepsilon-t)^{d/2}} e^{\frac{|x|^2}{4(T+\varepsilon-t)}}, \quad \forall t \in [0,T], \ \forall x.$$

Since $\mu > 0$ is arbitrary, we obtain $u(t,x) \le 0$ when $t \in [0,T]$. Similarly, we have $-u(t,x) \le 0$. Therefore, $u(t,x) = 0$ for $t \in [0,T]$.

Finally, we can iterate the argument on $[T,2T]$, $[2T,3T]$, $\dots$ to obtain that $u(t,x) = 0$ for all $t \ge 0$.
□

Remark 3.2 When the growth condition is not satisfied, a counterexample, known as Tychonoff's solution can be constructed; see Firtz John 7.1.

3.4.3 Comparison principle and stability in maximum norm

1. Dirichlet boundary condition, bounded domain As a corollary of the weak maximum principle, we have the following result.

   Theorem 3.14 (Comparison principle) *If $u, v \in \mathcal{C}^{1,2}(U_T) \cap \mathcal{C}(\overline{U_T})$ satisfy*

   $$\begin{cases} (\partial_t - \Delta)u \ge (\partial_t - \Delta)v, & U_T, \\ \qquad u \ge v, & \Gamma_T, \end{cases}$$

   then $u \ge v$ in $\overline{U_T}$.

   From this we can derive a stability result in the L^∞ norm.

   Theorem 3.15 *Let $u_i \in \mathcal{C}^{1,2}(\Omega_T) \cap \mathcal{C}(\overline{\Omega_T})$, $i = 1,2$, be classical solutions to*

   $$\begin{cases} \partial_t u_i = \Delta u_i + f_i, & \Omega_T, \\ u_i\big|_{\partial\Omega} = g_i, \\ u_i\big|_{t=0} \phi_i, \end{cases}$$

   where f_i, g_i, ϕ_i are continuous in their respective domains. Then

   $$\max_{\Omega_T} |u_1 - u_2| \le T\|f_1 - f_2\|_{L^\infty} + \|g_1 - g_2\|_{L^\infty} + \|\phi_1 - \phi_2\|_{L^\infty}.$$

We can interpret this result as the map from data to solution:

$$(f, g, \phi) \mapsto u,$$

is continuous (stable) in the space of continuous functions, which is equipped with the maximum norm.

Proof: Let

$$v(t,x) = u_1(t,x) - u_2(t,x), \quad w(t,x) = t\|f_1 - f_2\|_{L^\infty} + \|g_1 - g_2\|_{L^\infty} + \|\phi_1 - \phi_2\|_{L^\infty}.$$

Then one can check

$$\begin{cases} \partial_t w - \Delta w \geq \partial_t v - \Delta v, & \Omega_T, \\ w \geq v, & \Gamma_T. \end{cases}$$

Hence, $w \geq u$ in $\overline{U_T}$, and

$$\max_{\overline{\Omega_T}} |u_1 - u_2| \leq \|w\|_{L^\infty} = T\|f_1 - f_2\|_{L^\infty} + \|g_1 - g_2\|_{L^\infty} + \|\phi_1 - \phi_2\|_{L^\infty}.$$

□

2. mixed boundary condition, bounded domain We can also formulate similar results for more general mixed boundary conditions. First, we need a version of comparison principle. For simplicity, we will assume $\Omega = (0, \ell)$, but the result holds for general bounded domains as well.

Proposition 3.16 *Let $u \in \mathcal{C}^{1,2}(\Omega_T) \cap \mathcal{C}^{0,1}(\overline{\Omega_T})$ satisfy*

$$\begin{cases} \mathcal{L}u = \partial_t u - \Delta u \geq 0, & \Omega_T, \\ u(t,x) \geq 0, & \Omega, \\ \dfrac{\partial u}{\partial n} + \beta u\big|_{\partial U} \geq 0, & t \in [0,T], \end{cases}$$

where $\beta : \partial\Omega \to [0, \infty)$. Then $u \geq 0$ on $\overline{\Omega_T}$.

Proof:

First let us assume the strict inequality on the boundary:

$$\frac{\partial u}{\partial n} + \beta u > 0, \quad x \in \partial\Omega.$$

By weak maximum principle, $\min_{\overline{\Omega_T}} u$ is achieved on $\partial_p \Omega_T$. Let (t^*, x^*) be the point of maximum. We claim that $u(t^*, x^*) \geq 0$. Indeed, if $(t^*, x^*) \in \{t = 0\} \times \Omega$, then $u(0, x^*) \geq 0$ due to the initial condition; if $(t^*, x^*) \in [0,T] \times \partial\Omega$, then $\frac{\partial u}{\partial n} \leq 0$ on $\partial\Omega$. Since $\beta \geq 0$, we have $u(t^*, x^*) \geq 0$. This proves the claim.

Next, we assume the non-strict inequality. Let

$$w(t,x) = 2t + (x - \ell/2)^2.$$

Then

$$\mathcal{L}w \geq 0, \quad w\big|_{t=0} \geq 0, \quad \frac{\partial w}{\partial n} + \beta w\big|_{\partial\Omega} \geq c, \tag{3.19}$$

where $c > 0$ is a constant. Also,

$$\max_{\overline{\Omega_T}} |w| \leq C_1(T+1)$$

by direct computation. We consider

$$u_\varepsilon(t,x) = u(t,x) + \varepsilon w(t,x),$$

then u_ε satisfies the strict inequality on the boundary. Hence, we have

$$\min_{\overline{\Omega_T}} u_\varepsilon \geq 0 \implies \min_{\overline{\Omega_T}} u \geq -\varepsilon \max_{\overline{\Omega_T}} \big(2t + (x-\ell/2)^2\big).$$

Letting $\varepsilon \to 0+$, we obtain $\min_{\overline{\Omega_T}} u \geq 0$. $\square$

In the proof, the assumption on the domain is used solely to construct the function w satisfying (3.19). For a general bounded domain, such function still exists; however, its existence relies on the theory of elliptic equations.

We can formulate the L^∞-stability for mixed boundary condition.

Theorem 3.17 *Let $u \in \mathcal{C}^{1,2}(\overline{\Omega_T}) \cap \mathcal{C}^{1,0}(\overline{\Omega_T})$ be a classical solution to*

$$\begin{cases} \mathcal{L}u = \partial_t u - \Delta u = f(t,x), & (t,x) \in \Omega_T, \\ u(0,x) = \phi(x), & x \in \Omega, \\ \dfrac{\partial u}{\partial n} + \beta u\big|_{\partial\Omega} = g(t,x), & (t,x) \in \partial_p \Omega_T. \end{cases}$$

Then

$$\max_{\overline{\Omega_T}} \leq C(T+1)\Big(\|f\|_{L^\infty(\Omega_T)} + \|\phi\|_{L^\infty(\Omega)} + \|g\|_{L^\infty(\partial_p\Omega_T)}\Big)$$

for some constant $C = C(\Omega)$.

Proof:

Remark 3.3 If $\beta > 0$, then we do not need w.

Let w satisfy (3.19) such that

$$\max_{\overline{\Omega_T}} |w| \leq C_1(T+1).$$

We consider

$$v(t,x) = Ft + \Phi + \frac{G}{c} \pm u(t,x),$$

where c is the constant in (3.19) and

$$F = \|f\|_{L^\infty}, \quad \Phi = \|\phi\|_{L^\infty}, \quad G = \|g\|_{L^\infty}.$$

Then

$$\begin{aligned} \mathcal{L}v &= F \pm \mathcal{L}u + G\mathcal{L}w \geq 0, && (t,x) \in \Omega_T, \\ v(0,x) &\geq \Phi \pm \phi(t,x) \geq 0, && x \in \Omega, \\ \frac{\partial v}{\partial n} + \beta v &\geq G + g \geq 0, && x \in \partial\Omega. \end{aligned}$$

Hence, by Proposition 3.16, $v(t,x) \geq 0$ on Ω_T and

$$\max_{\overline{\Omega_T}} |u(t,x)| \leq FT + \Phi + \frac{\|w\|_{L^\infty}}{c} G.$$

The desired conclusion follows. $\square$

3.4.4 Weak maximum principle for general parabolic operators

The weak maximum principle also holds for general parabolic operators $\partial_t - \mathcal{L}$, where

$$\mathcal{L}f = \sum_{i,j} a_{ij}(x)\partial_{ij}f(x) + \sum_i b_i(x)\partial_i u.$$

Theorem 3.18 *Let Ω be a bounded domain. Suppose $A(x) = \big(a_{ij}(x)\big)$ is positive semi-definite for every $x \in \Omega$. If $u \in \mathcal{C}^{1,2}(\Omega_T) \cap \mathcal{C}(\overline{U_T})$ satisfies $\partial_t u - \mathcal{L}u \le 0$ in Ω_T, then*

$$\max_{\overline{U_T}} u = \max_{\partial_p \Omega_T} u.$$

The following lemma from linear algebra is useful.

Lemma 3.19 *If two symmetric $d \times d$ matrices (a_{ij}) and (b_{ij}) are positive semi-definite, then their Hadamard product $(c_{ij}) = (a_{ij}b_{ij})$ is also positive semi-definite.*

Proof: Assume first that $\partial_t u - \mathcal{L}u < 0$ in Ω_T. Assume the contrary, that is, the point of maximum of u over $\overline{\Omega_T}$, (t^*, x^*) is in Ω_T. Then we have

$$\partial_t u\big|_{(t^*,x^*)} > 0,$$
$$\nabla u(t^*, x^*) = 0 \implies \sum_i b_i(x^*)\partial_i u(t^*, x^*) = 0,$$

Also, the matrix

$$H = (\mathrm{Hess} u(t^*, x^*)) = (\partial_{ij} u(t^*, x^*))$$

is negative semi-definite. Since $(a_{ij}(x^*))$ is positive semi-definite, by Lemma 3.19, their Hadamard product

$$M = (a_{ij}(x^*)\partial_{ij}u(t^*, x^*))$$

is negative semi-definite, and hence

$$\sum a_{ij}(x^*)\partial_{ij}u(t^*, x^*) = \mathbb{1}^T M \mathbb{1} \le 0,$$

where $\mathbb{1} = (1, 1, \ldots, 1)^T$. This implies $\partial_t u - \mathcal{L}u \ge 0$ at (t^*, x^*), which contradicts with the assumption.

For the non-strict inequality, we can consider $u_\varepsilon(t, x) = u(t, x) - \varepsilon t$ and then let $\varepsilon \to 0+$.

□

Last, we will say a few words about the strong maximum principle, which is formulated as follows.

Theorem 3.20 *Let u be a classical solution to the heat equation. If Ω is connected and $\exists (t_0, x_0) \in \Omega_T$ such that*

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

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

To prove this, we need a more powerful tool: the mean-value property for the heat equation solution. This property can also be employed to show that $u \in \mathcal{C}^\infty(\Omega_T)$, a result which we have derived for solutions on the whole space but not yet for general domains. Although we omit the proof here, a parallel development exists for harmonic functions — those satisfying $\Delta u = 0$.

3.5 Energy estimates

Theorem 3.21 *Let $u \in \mathcal{C}^{1,2}(\Omega_T) \cap \mathcal{C}(\overline{\Omega_T})$ solve*

$$\partial_t u = \Delta u + f, \quad u\big|_{t=0} = \phi, \quad u\big|_{\partial\Omega} = 0.$$

Then there exists a constant $C = C(T)$ such that

$$\sup_{0\le t\le T} \|u(t,\cdot)\|^2_{L^2(\Omega)} + 2\int_0^T \|\nabla u(t,\cdot)\|^2_{L^2(\Omega)}\,dt \le C\Big(\|\phi\|^2_{L^2(\Omega)} + \int_0^T \|f(t,\cdot)\|^2_{L^2(\Omega)}\,dt\Big).$$

This result states that the solution map $(\phi, f) \mapsto u$ is continuous in the L^2-norm.
We will need the following Gronwall's inequality.

Lemma 3.22 (Gronwall's inequlity) *Let G, F satisfy*

$$G'(t) \le G(t) + F(t), \quad G(0) = 0.$$

Then

$$G(t) \le e^t F(t).$$

Proof: Multiplying the equation by u and integrating over Ω, we have

$$\int_\Omega u\partial_t u - u\Delta u = \int_\Omega uf. \tag{3.20}$$

Using integration by parts and noting that $u\big|_{\partial U} = 0$, the LHS is

$$\frac{1}{2}\int_\Omega u^2(t,x)\,dx + \int_\Omega |\nabla u|^2(t,x)\,dx,$$

while the RHS is bounded by

$$\frac{1}{2}\int_\Omega u^2(t,x)\,dx + \frac{1}{2}\int_\Omega f^2(t,x)\,dx.$$

Integrating over $[0,t]$, we obtain

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

Applying Lemma 3.22 with

$$G(t) = \int_0^t \|u(s,\cdot)\|^2_{L^2}\,ds, \quad F(t) = \|\phi\|^2_{L^2} + \int_0^t \|f(s,\cdot)\|^2_{L^2}\,ds,$$

we obtain

$$\|u(t,\cdot)\|^2_{L^2} \le e^t F(t).$$

Therefore,

$$\sup_{0\le t\le T} \|u(t,\cdot)\|^2_{L^2} \le e^T F(T) = e^T\Big(\|\phi\|^2_{L^2} + \int_0^T \|f(s,\cdot)\|^2_{L^2}\,ds\Big),$$

and

$$2\int_0^T \|\nabla u(t,\cdot)\|^2_{L^2} \le F(T) + G(T) \le (e^T+1)F(T) = (e^T+1)\Big(\|\phi\|^2_{L^2} + \int_0^T \|f(s,\cdot)\|^2_{L^2}\,ds\Big).$$

This completes the proof. □

3.6 Backward heat equation

It may happen that the PDE solution is unique, but is not stable with respect to the data. One such example is the backward heat equation.

Proposition 3.23 *Let $u_i \in \mathcal{C}^2(\overline{\Omega_T})$, $i = 1, 2$, solve*

$$\partial_t u_i = \Delta u, \text{ in } \Omega_T, \quad u_i\big|_{\partial\Omega} = g.$$

Assume that $u_1(T, \cdot) = u_2(T, \cdot)$. Then $u_1 \equiv u_2$ on $\overline{\Omega_T}$.

Here, the stability cannot hold. Let φ solve

$$-\Delta\varphi = \lambda\varphi, \quad \varphi\big|_{\partial U} = 0.$$

Then $u_\lambda(t, x) = e^{(T-t)\lambda}\varphi(x)$ solves

$$\partial_t u_\lambda = \Delta u_\lambda, \quad u_\lambda(T, \cdot) = \varphi(x),$$

with

$$\|u_\lambda(0, \cdot)\|_{L^2} = e^{T\lambda}\|\varphi\|_{L^2}.$$

But for the eigenvalue problem, there exists eigenpair (λ, φ) with λ arbitrarily large, so the solution cannot be controlled by the data φ in any sense. In terms of physics, this can be interpreted as the irreversibility of a thermodynamical system.

Proof: Let $w = u_1 - u_2$ and $e(t) = \|w(t, \cdot)\|^2_{L^2(\Omega)}$. We have

$$\begin{aligned}
\dot{e}(t) &= -2\int_\Omega |\nabla w|^2(t, x)\, dx = 2\int_\Omega w\Delta w, \\
\ddot{e}(t) &= -4\int_\Omega \nabla w \cdot \nabla w_t \\
&= 4\int_\Omega (\Delta w)w_t - 4\int_{\partial\Omega} \frac{\partial w}{\partial n} w_t \\
&= 4\int_\Omega |\Delta w|^2.
\end{aligned}$$

Hence, by Cauchy-Schwartz,

$$|\dot{e}(t)|^2 = 4\Big|\int_\Omega w\Delta w\Big|^2 \le 4\int_\Omega w^2 \int_\Omega |\Delta w|^2 = e(t)\ddot{e}(t).$$

We claim that if a non-negative $\mathcal{C}^2$-function f satisfies

$$|f'(t)|^2 \le f(t)f''(t), \quad 0 \le t \le T,$$

and $f(T) = 0$, then $f(t) \equiv 0$ for $t \in [0, T]$. Indeed, if $f(t) \not\equiv 0$, then there exists an interval $[a, b]$ such that $f(t) > 0$ on $[a, b)$ and $f(b) = 0$. Let $g(t) = \log e(t)$. Then

$$g''(t) = \frac{f''(t)f(t) - [f'(t)]^2}{g(t)} \ge 0, \quad t \in (a, b),$$

so $g(t)$ is a convex function on (a, b). On the other hand,

$$\lim_{t\to b-} g(t) = -\infty.$$

This is impossible for a convex function, and thus leads to a contradiction. $\square$

4 Elliptic equation

In this section, we will study the Laplace equation $\Delta u = 0$ and the Poisson's equation $-\Delta u = f$. Similar to the heat equation, the fundamental solution and the Green's function play important roles in the solution theory. A function Φ is called the *fundamental solution* if it solves

$$-\Delta u(x) = \delta(x), \quad x \in \mathbb{R}^d,$$

and a function $G(x; y)$ is called the *Green's function* for the domain Ω, where $y \in \Omega$, if

$$\begin{cases} -\Delta G(x; y) = \delta(x-y), & x \in \Omega, \\ \quad G(x; y) = 0, & x \in \partial\Omega. \end{cases}$$

We can use the Green's function to solve the Laplace equation on a domain. Indeed, if $u \in \mathcal{C}^2(\Omega) \cap \mathcal{C}(\bar{\Omega})$ solves

$$-\Delta u(x) = 0,\ x \in \Omega, \quad u(x) = g(x),\ x \in \partial\Omega,$$

then formally we have

$$\begin{aligned} u(y) &= \int_\Omega \delta(x-y)u(x)\,dx \\ &= \int_\Omega (-\Delta G(x; y))u(x)\,dx \\ &= \int_\Omega (-\Delta u)G(x; y)\,dx - \int_{\partial\Omega} \frac{\partial G}{\partial n} u\,dS + \int_{\partial\Omega} \frac{\partial u}{\partial n} G\,dS \\ &= \int_{\partial\Omega} \Big(-\frac{\partial G(x; y)}{\partial n}\Big) g(x)\,dS(x). \end{aligned}$$

4.1 Fundamental solution

4.1.1 Method of Fourier transform

Let $\Phi(x)$ be the fundamental solution. Then its Fourier transform satisfies

$$4\pi^2|\xi|^2\hat{\Phi}(\xi) = 1,$$

and hence

$$G(x) = \Big(\frac{1}{4\pi^2|\xi|^2}\Big)^{\wedge} = \int \frac{e^{2\pi i x\cdot\xi}}{4\pi^2|\xi|^2}\,d\xi. \tag{4.1}$$

This integral does not exist in the classical sense unless $d \geq 3$.

We have two observations. First, the function G is radially symmetric, that is, $G(x) = G(|x|)$. Indeed, for any orthogonal transform $O : \mathbb{R}^d \to \mathbb{R}^d$,

$$G(Ox) = \int \frac{1}{4\pi^2|\xi|^2} e^{2\pi i(Ox\cdot\xi)}\,d\xi = \int \frac{1}{4\pi^2|\xi|^2} e^{2\pi i x\cdot O^T\xi}\,d\xi = G(x),$$

since $|O\xi| = |\xi|$. Second, there is scaling relation: for $\lambda > 0$,

$$G(\lambda x) = \int \frac{1}{4\pi^2|\xi|^2} e^{2\pi i x\cdot(\lambda\xi)}\,d\xi = \frac{1}{\lambda^{d-2}} \int \frac{1}{4\pi^2|\lambda\xi|^2} e^{2\pi i x\cdot(\lambda\xi)}\,d(\lambda\xi) = \frac{1}{\lambda^{d-2}} G(x).$$

Therefore, $G(x) = c_d|x|^{d-2}$ for $d \geq 3$.

To determine the constant c_d, we use the following argument. For any domain $\Omega \ni 0$, we have

$$1 = \int_{\Omega} \delta(x)\,dx = \int_{\Omega} (-\Delta\Phi) \cdot 1 = -\int_{\partial\Omega} \frac{\partial \Phi}{\partial n} \cdot 1.$$

Taking $\Omega = B_r(0)$, we have

$$1 = \int_{\partial B_r(0)} (d-2)\frac{c_d}{r^{d-1}}\,dS = (d-2)\frac{|\partial B_r|}{r^{d-1}}.$$

Let $|B_r| = \alpha_d r^d$. Then $|\partial B_r| = d\alpha_d r^{d-1}$, and hence

$$c = \frac{1}{(d-2)d\alpha_d}.$$

It is known that

$$\alpha_d = |B_1| = \frac{\pi^{d/2}}{\Gamma(d/2+1)}.$$

To demystify the appearance of the δ-function, we formulate the following result.

Theorem 4.1 *Let $f \in \mathcal{C}_c^2(\mathbb{R}^d)$, $d \geq 3$. Then*

$$u(x) = (\Phi * f)(x) = \frac{1}{d(d-2)\alpha_d}\int_{\mathbb{R}^d} \frac{1}{|x-y|^{d-2}} f(y)\,dy$$

is in $\mathcal{C}^2(\mathbb{R}^d)$ and solves $-\Delta u = f$.

Proof: Direct computation shows $\Delta\Phi(x) = 0$ for $x \neq 0$, and $\partial_{ij}(\Phi * f) = \Phi * (\partial_{ij} f)$ exists so $u \in \mathcal{C}^2$. Let $x \in \mathbb{R}^d$ and R be sufficiently large so that $B_R(x)$ contains the support of f. We have

$$\Delta u(x) = \int_{|y|\leq R} \Phi(y)\Delta_x f(x-y)\,dy = \int_{\varepsilon\leq|y|\leq R} \Phi(y)\Delta_x f(x-y)\,dy + \int_{|y|<\varepsilon} \Phi(y)\Delta_x f(x-y)\,dy =: I_1 + I_2.$$

For I_1, since f vanishes on ∂B_R, integration by parts yields

$$\begin{aligned} I_1 &= -\int_{\varepsilon\leq|y|\leq R} \Delta_y \Phi(y) f(x-y)\,dy - \int_{\partial B_\varepsilon} \frac{\partial f}{\partial n}(x-y)\Phi(y)\,dS + \int_{\partial B_\varepsilon} f(x-y)\frac{\partial \Phi}{\partial n}\,dS \\ &= 0 + I_{11} + I_{12}. \end{aligned}$$

We have

$$|I_{11}| \leq \|Df\|_{L^\infty} \int_{\partial B_\varepsilon} \Phi(y) \leq \|Df\|_{L^\infty} c\varepsilon^{d-1} \cdot \frac{1}{\varepsilon^{d-2}} \to 0, \quad \varepsilon \to 0+,$$

and

$$\begin{aligned} |I_{12} - f(x)| &= \left| \int_{\partial B_\varepsilon} f(x-y)\frac{\partial \Phi}{\partial n} - \int_{\partial B_\varepsilon} f(x)\frac{\partial \Phi}{\partial n} \right| \\ &\leq \sup_{|y|\leq\varepsilon} |f(x-y) - f(x)| \to 0, \quad \varepsilon \to 0, \end{aligned}$$

where we used $-\frac{\partial \Phi}{\partial n} \geq 0$ and integrates to 1 on ∂B_ε. For I_2, we have

$$|I_2| \leq \|D^2 f\|_{L^\infty} \int_{|y|\leq\varepsilon} G(\varepsilon) \leq \|D^2 f\|_{L^\infty} c\varepsilon^d \frac{1}{\varepsilon^{d-2}} \to 0, \quad \varepsilon \to 0+.$$

Combining all these we prove the desired conclusion. $\square$

Remark 4.1 In the proof, besides $\Delta\Phi = 0$ at $x \neq 0$, we have use two things:

$$\int_{\partial B_r} \frac{\partial \Phi}{\partial n} = -1, \quad \forall r > 0,$$

and

$$\int_{B_r} G(y)\,dy, \int_{\partial B_r} G(y)\,dS(y) \to 0, \quad r \to 0+\,.$$

These two facts still hold for fundamental solution in $d = 1, 2$, as we will see.

In $d = 3$, the fundamental solution has the form $\Phi(x) = c|x|^{-1}$. This has deep implication in physics. Recall that in a static magnetic field, by the Maxwell's equation,