Lecture Note for Honor PDE

Liying LI

October 23, 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

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

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: $u(t,x):\mathbb{R}_+\times\mathbb{R}^d$, $d=2,3$,

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

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 General linear case

We consider the general linear case

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

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.

Rarefaction solution Suppose the initial condition is given by

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

By looking at the characteristics, we have

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

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

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

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.*

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:

$$\partial_t(\alpha u_1 + \beta u_2) = \alpha \partial_t u_1 + \beta \partial_t u_2,$$
$$\Delta(\alpha u_1 + \beta u_2) = \alpha \Delta u_1 + \beta \Delta u_2.$$

□

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

In this section we will demostrate how to use Fourier transform to solve the heat equation on the whole space. See also [Zho, Chap. 3.1.1] or [Eva, 4.3.1.a] and Section 3.2.7.

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 $\mathbb{F}$ and its inverse $\mathbb{F}^{-1}$.

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

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

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

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

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

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

□

Proof of (3.4): We have

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

□

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

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

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. □

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,

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

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

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

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 also [Lax, Chap 11.1] and [Eva, Appendix C.4]

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

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

Using the second condition and taking lim sup, we have

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

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

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

This part follows the presentation in [Eva, 2.3.1.a].

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

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

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

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

Integrating once, we obtain

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

where we fix the constant assuming

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

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

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

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

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

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, $\partial_{x_i x_j}$ 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

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

For I_2, we have

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

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

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

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 vaguely formulate the Duhamel's principle as follows.

Theorem 3.10 *Suppose that $x(t) = \Phi(t)x_0$ solves the equation*

$$x'(t) = Ax(t), \quad x(0) = x_0,$$

where $\Phi(t)$ and A are linear operators. Then

$$x(t) = \Phi(t)x_0 + \int_0^t \Phi(t-s)f(s)\,ds \tag{3.13}$$

solves the non-homogeneous equation

$$x'(t) = Ax(t) + f(t), \quad t > 0.$$

Proof: Let $x(t)$ be defined by (3.13). Differentiating in t yields

$$x'(t) = A\Phi(t)x_0 + \Phi(0)f(t) + \int_0^t \Phi'(t-s)f(s)\,ds.$$

Since $\Phi(t)x_0$ solves the homogeneous equation, we have $\Phi(0) = \mathrm{Id}$ and

$$\Phi'(t) = A\Phi(t).$$

Hence,

$$x'(t) = A\Phi(t)x_0 + f(t) + \int_0^t A\Phi(t-s)f(s)\,ds = A\Phi(t)x_0 + f(t) + A\int_0^t \Phi(t-s)f(s)\,ds = Ax(t) + f(t),$$

as desired. □

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

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

And $u(t,x) = \int_0^\infty ds \int_{\mathbb{R}^d} dy\, \Gamma(t,x;s,y) f(s,y)$ solves

$$(\partial_t - \Delta)u(t,x) = \int_0^\infty ds \int_{\mathbb{R}^d} dy \Big(\big[(\partial_t - \Delta)\Gamma\big](t,x;s,y) f(s,y) = \int \delta(t-s, x-y) f(s,y) = f(t,x).$$

The readers can find more on Duhamel's principle in [Eva, 2.3.1.c]

3.2.7 Note on Fourier transform

This section will sketch the Fourier transform theory involving generalized functions.

Fourier transform is first defined for functions. The *Fourier transform* of a function $g \in L^1(\mathbb{R})$ is defined by

$$(\mathbb{F}g)(\xi) := \int e^{i\xi x} g(x)\, dx. \tag{3.14}$$

The integrability condition $g \in L^1(\mathbb{R})$ is to ensure the integral in (3.14) to be defined.

In general, one needs to decide where to put constants and plus/minus signs in defining the Fourier transform; for example, more common definitions in harmonic analysis are

$$(\mathbb{F}g)(\xi) = \frac{1}{\sqrt{2\pi}} \int e^{-i\xi x} g(x)\, dx, \quad \text{or} \quad (\mathbb{F}g)(\xi) = \int e^{-2\pi i \xi x} g(x)\, dx.$$

But (3.14) agrees with the form of characteristic functions used in the probability theory so we will stick to it.

One can also define the *inverse Fourier transform* by

$$(\mathbb{F}^{-1}h)(x) := \frac{1}{2\pi} \int e^{-i\xi x} h(\xi)\, d\xi. \tag{3.15}$$

Note that like $\mathbb{F}$, the natural domain for $\mathbb{F}^{-1}$ are functions in $L^1(\mathbb{R})$. However, if $g \in L^1(\mathbb{R})$, then in general we merely have $\mathbb{F}g \in L^\infty(\mathbb{R})$, so $\mathbb{F}^{-1}$ is not a true "inverse" (but it will be after a proper generalization). When it happens that $\mathbb{F}g \in L^1(\mathbb{R})$, the map $\mathbb{F}^{-1}$ indeed takes $\mathbb{F}g$ back to g. Here, the form of $\mathbb{F}^{-1}$ in (3.15) depends on the choice we made in (3.14) to define $\mathbb{F}$.

Proposition 3.11 *If $g \in L^1(\mathbb{R})$ and $\mathbb{F}g \in L^1(\mathbb{R})$, then $(\mathbb{F}^{-1} \circ \mathbb{F})g = g$.*

The proof typically involves some integration tricks, and can be found in most analysis/PDE textbooks that present the Fourier transform. We skip the proof here since the most important thing for us is to know that the Fourier transform does have an inverse, at least in some sense.

The next question is that we need to define the Fourier transform for objects other than L^1 functions, like the probability measures. One can say that probability measures are like L^1 functions, but we will see below that the Fourier transform can even be defined for unbounded functions/measures. The key are the "Schwartz space" and its dual space, the "tempered distributions".

The *Schwartz space* contains smooth functions that decays fast at ∞; more precisely,

$$\mathcal{S} = \{g \in \mathcal{C}^\infty(\mathbb{R}) : \lim_{|x|\to\infty} |x^k| |g^{(m)}(x)| = 0, . \ \forall k, m \geq 0\}.$$

The functions in $\mathcal{S}$ are called *Schwartz functions.* We can talk about convergence in $\mathcal{S}$: $g_n \to g$ in $\mathcal{S}$ if for every $k, m \geq 0$, $\sup_x |x|^k |g_n^{(m)}(x) - g^{(m)}(x)| \to 0$. The convergence can also characterized by the metric

$$d(f, g) = \sum_{k,m=0}^{\infty} \frac{|f-g|_{k,m} \wedge 1}{2^{m+k}}, \quad |h|_{k,m} := \sup_x |x|^k |h^{(m)}(x)|.$$

A nice thing about the Fourier transform is that it turns differentiation ∂_x^k into multiplication $(-i\xi)^k$ and vise versa.

Proposition 3.12 *Let $g \in \mathcal{S}$. Then for $k \geq 1$,*

$$(\mathbb{F}g^{(k)})(\xi) = (-i\xi)^k (\mathbb{F}g)(\xi), \quad \mathbb{F}\big((-ix)^k g\big) = \mathbb{F}g^{(k)}.$$

Hence, the Schwartz space $\mathcal{S}$ is invariant under $\mathbb{F}$. By Proposition 3.11, it is a bijection on $\mathcal{S}$.

Proposition 3.13 *The Fourier transform $\mathbb{F} : \mathcal{S} \to \mathcal{S}$ is a bijection.*

Another obvious fact is that $\mathbb{F}$ is linear: $\mathbb{F}(f+g) = \mathbb{F}f + \mathbb{F}g$. It is natural to consider the action of $\mathbb{F}$ on the dual of $\mathcal{S}$, called the *tempered distribution*, defined by

$$\begin{aligned} \mathcal{S}' &:= \{\text{continuous, linear functional on } \mathcal{S}\} \\ &= \{\ell \text{linear} : \mathcal{S} \to \mathbb{R}, . \ |\ell(g)| \leq C_{m,k} |g|_{k,m}, \forall k, m \geq 0\}. \end{aligned}$$

The space $\mathcal{S}'$ contains all probability measures μ, identified with the linear functional

$$\ell_\mu(g) := \int g(x)\, d\mu(x).$$

It also contains $\mathcal{S}$ itself, identified with the linear functionals defined by taking L^2 inner product:

$$\ell_h(g) := \int g(x) h(x)\, dx, \quad h \in \mathcal{S}.$$

The Fourier transform can be defined on $\mathcal{S}'$ by duality:

$$(\mathbb{F}\ell)(g) := \ell(\mathbb{F}g).$$

For example, if μ is a probability measure on $\mathbb{R}$, then by Fubini's Theorem,

$$(\mathbb{F}\mu)(g) = \mu(\mathbb{F}g) = \int \Big[\int e^{i\xi x}\, dx\Big]\, d\mu(\xi) = \int \Big[\int e^{i\xi x}\, d\mu(\xi)\Big] g(x)\, dx = \int \varphi_\mu(x) g(x)\, dx, \quad \forall g \in \mathcal{S},$$

where φ_μ is the ch.f. of μ. Hence, the ch.f. φ_μ is $\mathbb{F}(\mu)$, when μ is treated as an element in $\mathcal{S}'$. Since $\mathbb{F} : \mathcal{S} \to \mathcal{S}$ is a bijection, it is also a bijection on $\mathcal{S}'$. Therefore, a probability measure is *uniquely* determined by its ch.f.

3.3 Heat equation on bounded domains, separation of variables

Some of the materials in this section can be found in [Zho, 3.2]

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

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

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

$$\begin{cases} -\Delta u = \lambda u, & x \in \Omega \\ \alpha u + \beta \dfrac{\partial u}{\partial n} = 0, & x \in \partial\Omega. \end{cases} \tag{3.16}$$

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 3.14 *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.

For more on the eigenvalue problem of $-\Delta$, the readers can refer to [Zho, Thm 3.6] and [Eva, 6.5.1].

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 (3.16) 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 3.14 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.

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

$$\begin{cases} \partial_t u = \partial_{xx} u, \quad t > 0, \ x \in (0, \ell), \\ u(0, x) = \phi(x), \quad x \in (0, \ell), \\ -\alpha_1 u'(0) + \beta_1 u(0) = 0, \\ \alpha_2 u'(\ell) + \beta_2 u(\ell) = 0. \end{cases} \tag{3.17}$$

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

$$\begin{cases} X''(x) + \lambda X(x) = 0, \quad x \in (0, \ell), \\ -\alpha_1 X'(0) + \beta_1 X(0) = 0, \\ \alpha_2 X'(\ell) + \beta_2 X(\ell) = 0. \end{cases} \tag{3.18}$$

Step 2:: solve the Sturm-Liouville problem (3.18). 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 (3.17) is given by

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

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

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

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

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

Solvable eigenvalue problems We list some cases where the eigenvalue problem

$$\begin{cases} -\Delta X_n(x) = \lambda X_n(x), & x \in \Omega, \\ \alpha X_n + \beta \dfrac{\partial X_n}{\partial n} = 0, & x \in \partial\Omega \end{cases} \tag{3.19}$$

is solvable.

$\Omega = [0, 2\pi]$ **(or any interval)**.

$$X_n(x) = a_n \cos(\sqrt{\lambda_n} x) + b_n \sin(\sqrt{\lambda_n} x) = c_n \sin(\sqrt{\lambda_n} x + \theta_n).$$

$\Omega = [0, 2\pi]^d$ **(or any rectangles)**.

$$X_{\vec{n}}(\vec{x}) = \sin(\sqrt{\lambda_{n_1}} x_1 + \theta_1) \cdots \sin(\sqrt{\lambda_{n_d}} x_d + \theta_d), \quad \vec{n} = (n_1, \ldots, n_d).$$

In particular, if $X_n = 0$ on the boundary, then

$$X_{\vec{n}}(x) = \sin(n_1 x_1) \cdots \sin(n_d x_d).$$

Periodic boundary condition $\Omega = \mathbb{T}^d$.

$$X_{\vec{n}}(\vec{x}) = e^{i\vec{n}\cdot\vec{x}}, \quad \vec{n} \in \mathbb{Z}^d.$$

Two-dimensional ball: $\Omega = B_1(0) \subset \mathbb{R}^2$. We can use separation of variables in the polar coordinate to find the eigenfunctions. We write

$$X(x) = \psi(r, \theta) = R(r)\Theta(\theta),$$

and $\lambda = k^2 \geq 0$. Then

$$\begin{aligned} 0 = \Delta\psi(r,\theta) + k^2\psi &= \partial_{rr}\psi + \frac{1}{r}\partial_r\psi_r + r^{-2}\partial_{\theta\theta}\psi \\ &= \big(R''(r) + r^{-1}R' + k^2 R\big)\Theta + r^{-2}R\Theta''. \end{aligned}$$

Hence,

$$\frac{r^2\big(R'' + r^{-1}R' + k^2 R\big)}{R} = -\frac{\Theta''}{\Theta} = \mu.$$

Since Θ is 2π-periodic, we have $\Theta(\theta) = e^{in\theta}$, and hence $\mu = n^2$. Then R solves

$$r^2 R'' + rR' + (r^2k^2 - n^2)R = 0.$$

We write $R(r) = J(kr)$. Then

$$x^2 J''(x) + xJ'(x) + (x^2 - n^2)J(x) = 0, \quad x = kr. \tag{3.20}$$

Solutions to (3.20) are Bessel functions $J_m(x)$, indexed by $m \geq 0$. The Dirichlet boundary condition requires $u\big|_{r=1} = 0$, which gives $J_m(k) = 0$, so k is a zero of J_m.

For the equation $\Delta u = 0$, $k = 0$ and R solves

$$r^2 R'' + rR' = n^2 R \implies R(r) = r^|n|.$$

Hence,

$$u(r, \theta) = \sum_{n=-\infty}^{\infty} a_n r^{|n|} e^{in\theta}.$$

3.3.4 Non-homogeneous equation and Green's function

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 $\{\sin \frac{n\pi x}{\ell}\}$:

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

where

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

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

$$G(t;x,y) = \frac{2}{\ell} \sum_{n=1}^{\infty} \sin \frac{n\pi x}{\ell} \sin \frac{n\pi y}{\ell} e^{-(\frac{n\pi}{\ell})^2 t}. \tag{3.21}$$

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

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

then we can rewrite the above Duhamel's principle as

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

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

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

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

Properties of the Green's function Using the explicity form of the Green's function (3.21), 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,

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

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

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

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

□

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:

$$\begin{cases}(\partial_t - \Delta)u = f, & t>0,\ x\in(0,\ell),\\ u(t,0) = g_1(t), & t>0,\\ u(t,\ell) = g_2(t), & t>0,\\ u(0,x) = \phi(x), & x\in(0,\ell).\end{cases} \tag{3.22}$$

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 (3.22) with

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

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

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

Here, in the second line we used

$$
\lim_{s\uparrow t} \int_0^\ell G(t,x;s,y)h(s,y)\,dy = \int_0^\ell \delta(x-y)h(t,y)\,dy = h(t,x),
$$

and in the fourth line we used $G(t,x,s,y) = 0$ for $x = 0, \ell$ and $\partial_{ss} h(s,y) = 0$.

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

For more on the maximum principle of the heat equation, the readers can refer to [Zho, 3.3.1] or [Eva, 2.3.3.a]

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*

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

Theorem 3.15 (Weak maximum principle) *Let $u \in \mathcal{C}^{1,2}(\Omega_T) \cap \mathcal{C}(\overline{\Omega_T})$. If*

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

then

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

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 3.16 *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

$$\begin{cases} \partial_t v = \Delta v, & \Omega_T, \\ v = 0, & \partial_p \Omega_T. \end{cases}$$

By weak maximum principle,

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

$\square$

Now let us prove the weak maximum principle.

Proof: First, let us assume that

$$\partial_t u - \Delta u < 0, \quad x \in \Omega_T. \tag{3.23}$$

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 (3.23).

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

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

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 3.17 *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

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

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

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

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 $[T,2T]$, $[2T,3T]$, ... to obtain that $u(t,x) = 0$ for all $t \ge 0$. $\square$

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

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

Theorem 3.18 (Comparison principle) *If $u, v \in \mathcal{C}^{1,2}(U_T) \cap \mathcal{C}(\overline{U_T})$ satisfy*

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

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

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

Theorem 3.19 *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

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

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

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

□

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 3.20 *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

$$\mathcal{L}w\geq 0,\quad w\big|_{t=0}\geq 0,\quad \frac{\partial w}{\partial n}+\beta w\big|_{\partial\Omega}\geq c, \tag{3.24}$$

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

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

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 $\min_{\overline{\Omega_T}} u \geq 0$. ☐

In the proof, the assumption on the domain is used solely to construct the function w satisfying (3.24). 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 3.21 *Let $u \in \mathcal{C}^{1,2}(\overline{\Omega_T}) \cap \mathcal{C}^{1,0}(\overline{\Omega_T})$ be a classical solution to*

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

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 (3.24) 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 (3.24) 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 3.20, $v(t,x) \geq 0$ on Ω_T and

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

The desired conclusion follows. ☐

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 3.22 *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 $\partial_t u - \mathcal{L}u \leq 0$ in Ω_T, then*

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

The following lemma from linear algebra is useful.

Lemma 3.23 *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 3.23, their Hadamard product

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

is negative semi-definite, and hence

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

where $\mathbb{1} = (1, 1, \ldots, 1)^T$. This implies $\partial_t u - \mathcal{L}u \geq 0$ 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 $\varepsilon \to 0+$.

□

The weak maximum principle about more general parabolic operators can be found in [Eva, 7.1.4.a]

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

Theorem 3.24 *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 3.25 *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 3.26 (Gronwall's inequlity) *Let G, F satisfy*

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

Then

$$G(t) \le e^t \int_0^t F(s)\,ds.$$

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

$$\int_\Omega u\partial_t u - u\Delta u = \int_\Omega uf. \tag{3.25}$$

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 3.26 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 3.27 *Let $u_i \in \mathcal{C}^2(\overline{\Omega_T})$, $i = 1, 2$, solve*

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

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,

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

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

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

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

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

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

$$-\Delta u(x) = 0,\ x \in \Omega, \quad u(x) = g(x),\ x \in \partial\Omega, \tag{4.1}$$

then formally we have

$$\begin{aligned} u(y) &= \int_\Omega \delta(x-y) u(x)\, dx \\ &= \int_\Omega (-\Delta G(x;y)) u(x)\, dx \\ &= \int_\Omega (-\Delta u) G(x;y)\, dx - \int_{\partial\Omega} \frac{\partial G}{\partial n} u\, dS + \int_{\partial\Omega} \frac{\partial u}{\partial n} G\, dS \\ &= \int_{\partial\Omega} \Big(-\frac{\partial G(x;y)}{\partial n} \Big) g(x)\, dS(x). \end{aligned} \tag{4.2}$$

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

$$G(x) = \Big(\frac{1}{4\pi^2 |\xi|^2} \Big)^{\wedge} = \int \frac{e^{2\pi i x \cdot \xi}}{4\pi^2 |\xi|^2}\, d\xi. \tag{4.3}$$

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. □

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 takes the form $\Phi(x) = c|x|^{-1}$. This form has significant implications in physics. In a static electric field, Maxwell's equation states that the electric field $\vec{E}$ satisfies

$$\nabla \cdot E = \frac{\rho}{\varepsilon_0}, \quad \nabla \times \vec{E} = 0,$$

where ρ is the charge density and ε_0 is a physical constant. The second equation implies that $\vec{E}$ is irrotational, so it can be expressed as $\vec{E} = \nabla\Phi$ for some scaler potential Φ. Plugging this into the first equation yields

$$\Delta\Phi = \frac{\rho}{\varepsilon_0}.$$

Let us consider the case of a point charge of strength q located at the origin, which corresponds to a charge density $\rho(x) = q\delta(x)$. The potential $\Phi(x)$ is then a multiple of the fundamental solution in $\mathbb{R}^3$. Specifically,

$$\Phi(x) = -\frac{q}{4\pi\varepsilon_0|x|}$$

Taking the gradient of this potential gives the electric field produced by the point charge:

$$\vec{E}(x) = \frac{q}{4\pi\varepsilon_0|x|^2} \cdot \frac{x}{|x|}.$$

This expression is recognized as Coulomb's law, which states that the electric force is inversely proportional to the square of the distance between two point charges. This inverse-square law can be observed as a consequence of the three-dimensional nature of the space.

4.1.2 Argument by spherical symmetry

Since $\delta(x)$ is rotationally invariant, we postulate that the fundamental solution is radially symmetric, that is, $\Phi(x) = \Phi(|x|)$.

We need a lemma about the Laplacian of a radially symmetric function.

Lemma 4.2 *Let $u(x) = u(r)$ where $r = |x|$. Then*

$$\Delta u = u''(r) + \frac{d-1}{r} u'(r).$$

Proof: We have

$$\begin{aligned}
\partial_{x_i} u(x) &= u'(r) \cdot \frac{x_i}{r}, \\
\partial_{x_i}^2 u(x) &= u''(r) \cdot (x_i/r)^2 + u'(r)(1/r - x_i/r \cdot x_i/r^2) \\
&= u''(r) \cdot \frac{x_i^2}{r^2} + u'(r)\Big(\frac{1}{r} - \frac{x_i^2}{r^2}\Big).
\end{aligned}$$

Summing over i and using $\sum_{i=1}^d x_i^2 = r^2$ give the desired result. $\square$

Since the fundamental solution satisfies $\Delta\Phi(x) = 0$ for $x \neq 0$, we have

$$\Phi''(r) + \frac{d-1}{r}\Phi'(r) = 0.$$

Integrating once gives

$$r\Phi' + (d-2)\Phi = \text{const}.$$

When $d = 2$, we obtain $u' = \text{const}/r$ and hence

$$u(r) = a \log r + b.$$

Apparently we should take $b = 0$ since $\Delta\text{const} = 0$. When $d \neq 2$, we can solve Φ from

$$\frac{d\Phi}{\text{const} - (d-2)\Phi} = \frac{dr}{r},$$

which gives

$$\Phi(r) = ar^{2-d} + b.$$

Again, $b = 0$. We can determine the constant using the relation

$$\int_{\partial B_r} \frac{\partial \Phi}{\partial n} = -1.$$

As a result, we have the fundamental solution for all dimensions:

$$\Phi(x) = \begin{cases} -\frac{1}{2}|x|, & d = 1, \\ -\frac{1}{2\pi}\log|x| & d = 2, \\ \frac{1}{d(d-2)\alpha_d}|x|^{2-d}, & d \geq 3. \end{cases}$$

4.2 Poisson kernel

In this section we want to construct Green's function on special domains. Recall that the Green's function satisfies the equation

$$-\Delta G(x;y) = \delta(x-y),\ x \in \Omega, \quad G(x;y) = 0,\ x \in \partial\Omega. \tag{4.4}$$

Case 1: Ω is the half-space

$$\Omega = \mathbb{R}^d_+ = \{(x_1, \ldots, x_d) : x_1 > 0\}.$$

We will check that

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

where $\bar{y} = (-y_1, y_2, \ldots, y_d)$ and $\Phi(x,y) = \Phi(x-y)$. Indeed, first,

$$-\Delta G(x;y) = -\Delta_x\Phi(x,y) + \Delta_x\Phi(x,\bar{y}) = -\Delta_x\Phi(x,y) = \delta(x-y),$$

since $x \in \Omega$ implies that $x \neq \bar{y} \in \Omega^c$; second,

$$G(x_0;y) = \Phi(|x_0 - y|) - \Phi(|x_0 - \bar{y}|) = 0, \quad x_0 \in \partial\Omega,$$

since y and $\bar{y}$ has equal distance to any point on $\partial\Omega$.

Writing $z = (z_1, \tilde{z})$, we can use (4.2) to write down the solution to the Laplace equation (4.1):

$$u(x) = \int_{\partial\Omega} (-\frac{\partial G}{\partial n}(z; x) g(z)\, dz) = \int_{\mathbb{R}^{d-1}} 2\partial_{z_1}\Phi(z; x)\big|_{z_1=0} g(\tilde{z})\, d\tilde{z}.$$

Writing $x = (x_1, \tilde{x})$ and $z = (z_1, \tilde{z})$, for $d \geq 3$, we have

$$\partial_{z_1}\Phi(z; x)\big|_{z_1=0} = \frac{1}{d(d-2)\alpha_d} \frac{1}{\left[(x_1 - z_1)^2 + |\tilde{x} - \tilde{z}|^2\right]^{\frac{d}{2}}} \cdot \Big(-\frac{d-2}{2}\Big) \cdot 2(z_1 - x_1)\big|_{z_1=0} = \frac{1}{d\alpha_d} \frac{x_1}{\left[x_1^2 + |\tilde{x} - \tilde{z}|^2\right]^{d/2}}.$$

Hence,

$$u(x) = \int_{\mathbb{R}^{d-1}} \frac{2x_1}{d\alpha_d} \cdot \frac{1}{\left[x_1^2 + |\tilde{x} - y|^2\right]^{d/2}} g(y)\, dy. \tag{4.5}$$

For $d = 2$,

$$\Phi(z; x) = -\frac{1}{4\pi} \log\Big((z_1 - x_1)^2 + (z_2 - x_2)^2\Big),$$

and

$$\partial_{z_1}\Phi(z; x)\big|_{z_1=0} = -\frac{1}{2\pi} \cdot \frac{z_1 - x_1}{(z_1 - x_1)^2 + (z_2 - x_2)^2}\Big|_{z_1=0} = \frac{1}{2\pi} \cdot \frac{x_1}{x_1^2 + (z_2 - x_2)^2},$$

so (4.5) also holds.

We call the function

$$K(x, y) = \frac{2x_1}{d\alpha_d} \frac{1}{\left[x_1^2 + |\tilde{x} - y|^2\right]^{d/2}}, \quad x \in \mathbb{R}^d,\ y \in \mathbb{R}^{d-1}, \tag{4.6}$$

the *Poisson kernel* on the half plane.

Case 2: $\Omega = B_R(0)$.

We start with a geometric lemma.

Lemma 4.3 (Apollonius) *Let $y \in \mathbb{R}^d \setminus \{0\}$ and $\bar{y} = \frac{R^2}{|y|^2} y$. Then for any $|x| = R$,*

$$\frac{|x - y|}{|x - \bar{y}|} = |y|/R.$$

In $d = 2$, B_R will be a circle of Apollonius *between y and $\bar{y}$.*

Proof: We have

$$\begin{aligned} |x - y|^2 - \frac{|y|^2}{R^2}|x - \bar{y}|^2 &= (1 - \frac{|y|^2}{R^2})|x|^2 + |y|^2 - \frac{|y|^2|\bar{y}|^2}{R^2} - 2x \cdot \Big(y - \frac{|y|^2}{R^2}\bar{y}\Big) \\ &= (R^2 - |y|^2) - (R^2 - |y|^2) = 0. \end{aligned}$$

□

Now let

$$G(x, y) = \Phi(|x - y|) - \Phi\Big(\frac{|y||x - \bar{y}|}{R}\Big).$$

Then by Lemma 4.3, $G(x, y) = 0$ for $|x| = R$, so the boundary condition is satisfied.

To obtain the solution to the Laplace equation, we need to compute $\partial G/\partial n$ on ∂B_R. We have

$$\begin{aligned}
\frac{\partial \Phi(x-y)}{\partial x_j} &= -\frac{1}{d\alpha_d}\frac{x_j-y_j}{|x-y|^d},\\
\frac{\partial \Phi}{\partial x_j}\Big(\frac{|y|(x-\bar{y})}{R}\Big) &= -\frac{1}{d\alpha_d}\frac{\frac{|y|^2}{R^2}(x_j-\bar{y})}{\left|\frac{|y|}{R}(x-\bar{y})\right|^d}\\
&= -\frac{1}{d\alpha_d}\frac{\frac{|y|^2}{R^2}x_j-y_j}{|x-y|^d}.
\end{aligned}$$

The unit normal vector at $x \in \partial B_R$ is $n = \frac{x}{R}$. Hence,

$$\begin{aligned}
\frac{\partial G}{\partial n}(x;y) &= \sum_{j=1}^{d}\frac{x_j}{R}\Big[\frac{\partial \Phi}{\partial x_j}(x-y) - \frac{\partial \Phi}{\partial x_j}\Big(\frac{|y|(x-\bar{y})}{R}\Big)\Big]\\
&= -\frac{1}{d\alpha_d|x-y|^d}\sum_{j=1}^{d}\frac{x_j}{R}\Big[(x_j-y_j) - (\frac{|y|^2}{R^2}x_j-y_j)\Big]\\
&= -\frac{1}{d\alpha_d|x-y|^d}\sum_{j=1}^{d}\frac{x_j^2}{R}\Big(1-\frac{|y|^2}{R^2}\Big)\\
&= -\frac{R^2-|y|^2}{d\alpha_d|x-y|^d\cdot R}.
\end{aligned}$$

Therefore, the solution to (4.1) is given by

$$u(y) = \frac{R^2-|y|^2}{d\alpha_d R}\int_{\partial B_R}\frac{g(x)\,dS(x)}{|x-y|^d}, \quad y \in B_R.$$

The kernel

$$K(x,y) = \frac{R^2-|y|^2}{d\alpha_d R|x-y|^d}$$

is called the Poisson kernel for B_R.

4.2.1 Laplace equation in special domains

Recall that if $G(x,y)$ is the Green's function in Ω and $u \in \mathcal{C}^2(\Omega)\cap\mathcal{C}(\bar{\Omega})$ solves

$$-\Delta u(x) = 0,\ x \in \Omega, \quad u(x) = g(x),\ x \in \partial\Omega, \tag{4.7}$$

then

$$u(x) = \int_{\partial\Omega}\Big(-\frac{\partial G}{\partial n}\Big)(x,y)g(y)\,dS(y). \tag{4.8}$$

We explicit compute $-\frac{\partial G}{\partial n}$ for $\Omega = \mathbb{R}^d_+$ and $B_R(0)$, which is also called the Poisson's kernel. For such special domains, the expression (4.8) indeed gives the unique solution to the Laplace equation (4.8). We formulate the following result for $\Omega = \mathbb{R}^d_+$.

Proposition 4.4 *Let $g \in \mathcal{C}(\mathbb{R}^{d-1})\cap L^\infty(\mathbb{R}^{d-1})$. Then*

$$u(x) = \int_{\mathbb{R}^{d-1}} K(x,y)g(y)\,dy,$$

where K is given in (4.6), is the unique bounded solution in $\mathcal{C}^\infty$ to (4.7), with

$$\lim_{\mathbb{R}^d_+ \ni x \to (0,y)} u(x) = g(y), \quad \forall y \in \mathbb{R}^{d-1}.$$

We will skip the proof uniqueness now since it follows from some form of maximum principle. Without the boundedness condition the statement is not true, since $\tilde{u}(x) = u(x) + kx_1$ is always another solution to (4.8).

Proof:

One can directly verify $\Delta u = 0$ in $\mathbb{R}^d_+$, and hence $u \in \mathcal{C}^\infty(\mathbb{R}^d_+)$ by Proposition 4.7.

Since $K(x,y) \geq 0$ and $\int_{\mathbb{R}^{d-1}} K(x,y)\,dy = 1$, we have

$$|u(x)| \leq \left| \int_{\mathbb{R}^{d-1}} K(x,y)g(y)\,dy \right| \leq \|g\|_{L^\infty},$$

so u is bounded.

Finally, we will verify the boundary condition. Let $y_0 \in \mathbb{R}^{d-1}$. We write $x = (x_1, \tilde{x}) \in \mathbb{R}^d$. For every $\varepsilon > 0$,

$$\begin{aligned}
|g(y_0) - u(x)| &\leq \left| \int_{\mathbb{R}^{d-1}} g(y_0)K(x,y)\,dy - \int_{\mathbb{R}^{d-1}} g(y)K(x,y)\,dy \right| \\
&\leq \int_{\mathbb{R}^{d-1}} K(x,y)|g(y) - g(y_0)|\,dy \\
&= \int_{|y-y_0|\geq\varepsilon} K(x,y)|g(y) - g(y_0)|\,dy + \int_{|y-y_0|<\varepsilon} K(x,y)|g(y) - g(y_0)|\,dy \\
&\leq 2\|g\|_{L^\infty} \int_{|y-y_0|\geq\varepsilon} K(x,y)\,dy + \sup_{|y-y_0|<\varepsilon} |g(y) - g(y_0)|.
\end{aligned}$$

As $\tilde{x} \to y_0$, we have

$$|\tilde{x} - y| \geq \frac{1}{2}|y_0 - y|, \quad \forall |y - y_0| \geq \varepsilon,$$

and hence

$$\begin{aligned}
\int_{|y-y_0|\geq\varepsilon} K(x,y)\,dy &\leq \int_{|y-y_0|\geq\varepsilon} \frac{2x_1}{d\alpha_d} \cdot \frac{1}{|\tilde{x} - y|^d}\,dy \\
&\leq \frac{2x_1 \cdot 2^d}{d\alpha_d} \int_{|y-y_0|\geq\varepsilon} \frac{1}{|y - y_0|^d}\,dy \\
&\leq Cx_1 \int_\varepsilon^\infty r^{-2}\,dr = C_\varepsilon x_1 \to 0
\end{aligned}$$

as $x_1 \to 0$. Hence, for every $\varepsilon > 0$,

$$\limsup_{x \to y_0} |g(y_0) - u(x)| \leq \sup_{|y-y_0|<\varepsilon} |g(y) - g(y_0)|.$$

Since $\varepsilon > 0$ is arbitrary, by the continuity of g, the LHS must be 0, and the conclusion follows. $\square$

Remark 4.2 For $\Omega = B_R$, since the explicit form of $K(x,y)$ is also known, we can verify

$$\lim_{x \to y_0} \int_{\partial B_R \cap \{|y-y_0|\geq\varepsilon\}} K(x,y)\,dy = 0.$$

How do we find Green's function for a general domain? Recall that the fundamental solution $\Phi(x,y)$ solves

$$-\Delta_x \Phi(x,y) = \delta(x-y).$$

The Green's function $G(x,y)$ needs to satisfy the boundary condition $G(x;y) = 0$ for $x \in \partial\Omega$. If we can solve the Laplace equation

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

then $G(x,y) = \Phi(x,y) - u(x)$ will be the desired Green's function.

4.2.2 Symmetry of the Green's function

If the Green's function satisfying (4.4) is known, then the solution to

$$-\Delta u = f,\ x \in \Omega, \quad u\big|_{\partial\Omega} = g$$

is given by

$$u(x) = \int_\Omega f(y)G(x,y)\,dy + \int_{\partial\Omega} \Big(-\frac{\partial G}{\partial n}(x;y)\Big) g(y)\,dS(y).$$

The Green's function is symmetric in x and y. We will give a formal proof below; for the rigorous proof one can see [Eva, 2.2, Theorem 13].

Theorem 4.5 *$G(x,y) = G(y,x)$ for $x \neq y$.*

Proof: Since $-\Delta_x G(x;y) = \delta(x-y)$ and $-\Delta_y G(y;x) = \delta(x-y)$, we have

$$\begin{aligned}
G(x;y) &= \int_\Omega \delta(x-z)G(z;y)\,dz \\
&= \int_\Omega -\Delta_z G(z;x)G(z;y)\,dz \\
&= \int_\Omega G(z;x)(-\Delta_z G(z;y))\,dz - \int_{\partial\Omega} \frac{\partial G}{\partial_z n}(z;x)G(z;y)\,dS(z) + \int_{\partial\Omega} \frac{\partial G}{\partial_z n}(z;y)G(z,x)\,dS(z) \\
&= \int_\Omega G(z;x)\delta(z-y)\,dz \\
&= G(y;x).
\end{aligned}$$

□

4.3 Harmonic function

A function $u \in \mathcal{C}^2(\Omega)$ is harmonic if $\Delta u = 0$ in Ω. In this section we present some properties of the harmonic functions. We follow [Eva, 2.2.3]

4.3.1 Mean-value property

Theorem 4.6 *If $u \in \mathcal{C}^2(\Omega)$ is harmonic, then*

$$u(x) = \fint_{\partial B_r(x)} u\,dS = \fint_{B_r(x)} u\,dx, \quad \forall B_r(x) \subset \Omega, \tag{4.9}$$

where for any domain D, $\fint_D := \frac{1}{|D|}\int$.

Conversely, if $u \in \mathcal{C}^2(\Omega)$ satisfies (4.9), then u is harmonic in Ω.

Proof:

Fix $x \in \Omega$. Let
$$\varphi(r) = \fint_{\partial B_r(x)} u\,dS = \fint_{\partial B_1(0)} u(x+ry)\,dS(y), \quad r > 0.$$
We have
$$\begin{aligned}\varphi'(r) &= \fint_{\partial B_1(0)} \nabla u(x+ry)\cdot y\,dS(y)\\ &= \fint_{\partial B_1(0)} \frac{\partial u}{\partial n}(x+ry)\,dS(y)\\ &= \frac{1}{|\partial B_1(0)|}\int_{B_1(0)} (\Delta u)\cdot 1\,dy = 0.\end{aligned}$$
Hence, φ is a constant, which equals to
$$\lim_{r\to 0+} \phi(r) = \fint_{\partial B_r(x)} u\,dS = u(x)$$
by the continuity of u at x.

For the second equality in (4.9), by the co-area formula, we have
$$\begin{aligned}\fint_{B_r(x)} u\,dx &= \frac{1}{|B_r(x)|}\int_0^r \int_{\partial B_{r'}(x)} u\,dSdr'\\ &= \frac{1}{|B_r(x)|}\int_0^r u(x)|\partial B_{r'}(x)|\,dr'\\ &= u(x).\end{aligned}$$

For the converse part, if $\Delta u(x) \neq 0$, assume without loss of generality that $\Delta u(y) > 0$ in $B_\delta(x)$. Then the previous computation implies that $\varphi'(r) > 0$, which contradicts with (4.9) that implies φ being a constant. $\square$

We can remove the assumption that $u \in \mathcal{C}^2(\Omega)$ for the converse part of Theorem 4.6.

Proposition 4.7 *If $u \in \mathcal{C}(\Omega)$ satisfies (4.9), then $u \in \mathcal{C}^\infty(\Omega)$ and is harmonic. In particular, harmonic functions are $\mathcal{C}^\infty$.*

Proof: Let $\eta \in \mathcal{C}^\infty(\mathbb{R}^d)$ be radially symmetric, non-negative such that $\operatorname{supp}\eta \subset B_1(0)$ and $\int_{B_1(0)} \eta(x)\,dx = 1$. For example, one can take
$$\eta(x) = \begin{cases} c\exp\left(\frac{1}{|x|^2-1}\right), & |x| < 1,\\ 0, & |x| \geq 1,\end{cases}$$
where $c > 0$ is a constant so that η integrates to 1. Let $\eta_\varepsilon(x) = \varepsilon^{-d}\eta(x/\varepsilon)$. Then $u_\varepsilon = u * \eta_\varepsilon \in \mathcal{C}^\infty(U_\varepsilon)$ where
$$U_\varepsilon = \{x \in U : \operatorname{dist}(x, \partial U) > \varepsilon\}.$$

On the other hand, we have

$$\begin{aligned}
u_\varepsilon(x) &= \int_{B_\varepsilon(0)} u(x-y)\eta_\varepsilon(y)\,dy \\
&= \int_0^\varepsilon dr \int_{\partial B_r(0)} u(x-y)\eta_\varepsilon(r)\,dS(y) \\
&= \int_0^\varepsilon |\partial B_r(0)|\eta_\varepsilon(r)u(x)\,dr \\
&= u(x)\cdot \int_{B_\varepsilon(0)} \eta_\varepsilon(y)\,dy = u(x),
\end{aligned}$$

where the third line follows from u having the mean-value property. Hence, $u(x) = u_\varepsilon(x) \in \mathcal{C}^\infty$. The rest follows from Theorem 4.6. □

4.3.2 Maximum principle

Theorem 4.8 (Maximum principle for harmonic functions) *Let $u \in \mathcal{C}^2(\Omega) \cap \mathcal{C}(\bar{\Omega})$ be harmonic in Ω where Ω is a bounded domain.*

1. *(Weak maximum principle)* $\max_{\bar{\Omega}} u = \max_{\partial\Omega} u$.
2. *(Strong maximum principle) Let Ω be connected. If $x_0 \in \Omega$ achieves the maximum of u over $\bar{\Omega}$, then u is a constant in Ω.*

Proof: The strong maximum principle implies the weak one, so we will only prove the strong one. However, the weak maximum principle has a proof similar to Theorem 3.22 and can be generalized to other elliptic operators, as we will see.

Let $M := u(x_0)$. By the mean-value property, if $u(x) = M$, then for any $B_r(x) \subset \Omega$,

$$u(x) = ⨍_{B_r(x)} u(y)\,dy = M,$$

so $u(y) = M$ for all $y \in B_r(x)$. This implies that

$$A = \{x \in \Omega : u(x) = M\}$$

is a (relatively) open set in Ω. The set A is also closed, since it is a pre-image of a singleton $\{M\}$ of a continuous function u. Since Ω is connected, the only sets that are both open and close are $\varnothing$ or Ω. Since $x_0 \in A$, A is non-empty, and hence $A = \Omega$.

□

Remark 4.3 A set Ω is connected if $\Omega = A \cup B$ where A, B are disjoint relatively open sets in Ω implies $A = \varnothing$ or $B = \varnothing$.

As a corollary, we have the uniqueness of solution to the Poisson equation.

Theorem 4.9 *Let Ω be a bounded domain. Then the solution $u \in \mathcal{C}^2(\Omega) \cap \mathcal{C}(\bar{\Omega})$ to*

$$-\Delta u = f,\ \Omega, \quad u = g,\ \partial\Omega,$$

is unique if it exists.

We can generalize the weak maximum principle to more general elliptic operators. Below is an example.

Proposition 4.10 *Let*

$$(\mathcal{L}u)(x) = -\Delta u(x) + c(x)u(x),$$

where $c \in \mathcal{C}(\Omega)$ and $c \geq 0$. If $u \in \mathcal{C}^2(\Omega) \cap \mathcal{C}(\bar{\Omega})$ satisfies $\mathcal{L}u \leq 0$ in Ω, then

$$\max_{\bar{\Omega}} u(x) \leq \max_{\partial\Omega} u^+(x), \quad u^+ = \max(u, 0).$$

Proof: Let $x_0 \in \operatorname{argmax}_{\bar{\Omega}} u(x)$. If $x_0 \in \partial\Omega$ or $u(x_0) \leq 0$, there is nothing to prove.

Assume that $u(x_0) > 0$ and $x_0 \in \Omega$. Also assume first that the strict inequality $\mathcal{L}u < 0$ holds in Ω. Then

$$-\Delta u(x_0) \geq 0, \quad c(x_0)u(x_0) \geq 0,$$

and hence $\mathcal{L}u(x_0) \geq 0$. This is a contradiction.

Assume that $\mathcal{L}u \leq 0$. Let

$$v_\varepsilon(x) = u(x) - \varepsilon|x|^2.$$

Then

$$\mathcal{L}v_\varepsilon = \mathcal{L}u - \varepsilon \cdot 2d < 0.$$

By what has been proved, we have

$$\max_{\bar{\Omega}} v_\varepsilon(x) \leq \max_{\partial\Omega} v_\varepsilon^+.$$

Since $v_\varepsilon \to u$ uniformly as $\varepsilon \to 0+$, the conclusion follows. $\square$

4.3.3 Mean-value property for heat equation

The heat equation also has a mean-value property, which leads to the strong maximum principle.

Let

$$E(t, x; r) = \left\{(s, y) \in \mathbb{R}^{n+1} : s \leq t,\ G_{t-s}(x - y) \geq \frac{1}{r^n}\right\},$$

where G is the heat kernel.

Theorem 4.11 *Let $u \in \mathcal{C}^{1,2}(\Omega_T)$ and $u_t = \Delta u$ in Ω. Then*

$$u(t, x) = \frac{1}{4r^n} \iint_{E(t,x;r)} u(s, y) \frac{|x - y|^2}{(t - s)^2}\, dyds$$

for all $E(t, s; r) \subset \Omega_T$.

For its proof, see [Eva, 2.3.2 Theorem 3].

4.3.4 Other properties of harmonic functions

Local estimate of derivatives

Theorem 4.12 *Let u be harmonic in $\Omega \subset \mathbb{R}^d$. Then for all $B_r(x_0) \subset \Omega$ and multi-index $|\alpha| = k$,*

$$|D^\alpha u(x_0)| \leq \frac{c_k}{r^{d+k}} \|u\|_{L^1(B_r(x_0))},$$

where $c_0 = \frac{1}{\alpha_d}$, $c_k = \frac{(2^{d+1}dk)^k}{\alpha_d}$.

We will give some applications of this result.

Application 1: Analycity of harmonic functions. Let u be harmonic in Ω. We want to show that the Taylor series expansion

$$u(x) = \sum_{\alpha} \frac{D^\alpha u(x_0)}{\alpha!}(x-x_0)^\alpha$$

has positive radius of convergence for every $x_0 \in \Omega$. Indeed, by Theorem 4.12, for $|x-x_0| \le s$, we have

$$\begin{aligned}\sum_{\alpha} \left|\frac{D^\alpha u(x_0)}{\alpha!}(x-x_0)^\alpha\right| &\le \sum_{k=0}^{\infty} \sum_{|\alpha|=k} \left|\frac{D^\alpha u(x_0)}{\alpha!}\right| s^k \\ &\le \sum_{k=0}^{\infty} \frac{s^k}{r^{d+k}} \cdot \frac{(2^{d+1}dk)^k}{\alpha_d} \sum_{|\alpha|=k} \frac{1}{\alpha!}.\end{aligned}$$

Since

$$d^k = (1+\cdots+1)^k = \sum_{|\alpha|=k} \frac{k!}{\alpha!},$$

and

$$\lim_{n\to\infty} \frac{k}{\sqrt[k]{k!}} = e,$$

we have

$$k^k \sum_{|\alpha|=k} \frac{1}{\alpha!} = \frac{d^k}{k!} \le c^k$$

for some $c > 0$. The conclusion follows.

Application 2:

Theorem 4.13 (Liouville's theorem) *If u is harmonic in $\mathbb{R}^d$ and bounded, then u is a constant.*

Proof: We have

$$|Du(x_0)| \le \frac{C}{r^{d+1}} \|u\|_{L^1(B_r(x_0))} \le \frac{C\|u\|_{L^\infty}}{r} \to 0, \quad r \to \infty.$$

Hence, $\nabla u \equiv 0$, and the conclusion follows. □

Application 3:

Proposition 4.14 *Let $f \in \mathcal{C}_c^2(\mathbb{R}^d)$, $d \ge 3$. Then all bounded solutions to $-\Delta u = f$ is given by*

$$u(x) = (\Phi * f)(x) + c,$$

where c is a constant.

Proof: For $d \ge 3$, $u = \Phi * f$ is bounded and solves $-\Delta u = f$. If $\tilde{u}$ is another solution, then $v = \tilde{u} - u$ is harmonic and bounded on $\mathbb{R}^d$, and the conclusion follows from Theorem 4.13. □

Now we will prove Theorem 4.12. **Proof:** We will prove by induction on k.

Let u be harmonic. Then $\partial_{x_j} u$ is also harmonic. By mean-value property,

$$\begin{aligned}
|(\partial_{x_j} u)(0)| &= \Big|\fint_{B_{r/2}} \partial_{x_j} u \, dx\Big| = \frac{1}{|B_{r/2}|}\Big|\int_{B_{r/2}} \partial_{x_j} u \, dx\Big| \\
&= \frac{1}{|B_{r/2}|}\Big|\int_{\partial B_{r/2}} u n_i \, dS(x)\Big| \\
&\le \frac{|\partial B_{r/2}|}{|B_{r/2}|}\|u\|_{L^\infty(B_{r/2})} \\
&\le \frac{|\partial B_{r/2}|}{|B_{r/2}|} \cdot \frac{1}{|B_{r/2}|}\|u\|_{L^1(B_r)} \\
&= \frac{d}{r/2} \cdot \frac{1}{\alpha_d}\left(\frac{2}{r}\right)^d \|u\|_{L^1(B_r)}.
\end{aligned}$$

So the results holds for $k = 1$.

Assume that

$$|D^\beta u(x_0)| \le \frac{c_{k-1}}{s^{d+k-1}}\|u\|_{L^1(B_s(x_0))}, \quad \forall s > 0, \ \forall |\beta| = k-1.$$

Let $D^\alpha = \partial_{x_j} D^\beta$. We have

$$\begin{aligned}
|(D^\alpha u)(x_0)| &\le \frac{d}{s}\|u\|_{L^\infty(B_s(x_0))} \\
&\le \frac{d}{s}\frac{c_{k-1}}{(r-s)^{d+k-1}}\|u\|_{L^1(B_r(x_0))}.
\end{aligned}$$

To optimize $s(r-s)^{d+k-1}$, we take $s = r/k$ and $r - s = (k-1)r/k$, and the conclusion follows. $\square$

Harnack's inequality

Theorem 4.15 *Let $u \ge 0$ be harmonic in Ω, and $V \subset \bar{V} \subset \Omega$, where V is bounded and connected. Then there exists a constant $C = C(V)$ such that*

$$\inf_V u \le C \cdot \sup_V u.$$

Proof: Let $r = \frac{1}{4}\operatorname{dist}(V, \partial\Omega)$. If $|x - y| \le r$ and $x, y \in \Omega$, then

$$u(x) = \fint_{B_{2r}(x)} u \, dz \ge \frac{1}{|B_{2r}|}\int_{B_r(y)} u \, dz = \frac{|B_r|}{|B_{2r}|}u(y) = \frac{1}{2^d}u(y).$$

Since $\bar{V}$ is compact, there exists $z_1, \ldots, z_N \in \bar{V}$ such that $\bar{V} \subset \bigcup_{j=1}^N B_r(z_j)$. For any $x, y \in V$, there exists a chain of points

$$w_0 = x, w_1, \ldots, w_m = y, \quad m \le 2N + 1,$$

such that $|w_i - w_{i+1}| \le r$. Hence, $u(x) \le 2^{d(2N+1)}u(y)$. The conclusion follows. $\square$

4.4 L^2-stability

Proposition 4.16 *Let $u \in \mathcal{C}^2(\Omega) \cap \mathcal{C}(\bar{\Omega})$ solve*

$$-\Delta u + cu = f \text{ in } \Omega, \quad u = 0, \text{ on } \partial\Omega, \tag{4.10}$$

where $c(x) \geq 0$ in Ω and $f \in L^2(\Omega)$. Then if Ω is bounded,

$$\int_\Omega |u|^2 + \int_\Omega |\nabla u|^2 \leq C \int_\Omega f^2. \tag{4.11}$$

If in addition $c(x) \geq c_0 > 0$, then for any Ω,

$$\int_\Omega |\nabla u|^2 + \frac{c_0}{2} \int_\Omega |u|^2 \leq C \int_\Omega f^2.$$

Proof: Multiplying u to both sides of (4.10), and using integration by parts, we have

$$\int_\Omega |\nabla u|^2 + c(x)|u|^2 = \int_\Omega fu. \tag{4.12}$$

If Ω is bounded, we have Theorem 4.28, and

$$\int_\Omega |\nabla u|^2 \leq \frac{1}{2\varepsilon} \int_\Omega f^2 + \frac{\varepsilon}{2} \int_\Omega u^2 \leq \frac{1}{2\varepsilon} \int_\Omega f^2 + \frac{\varepsilon K}{2} \int_\Omega |\nabla u|^2. \tag{4.13}$$

By choosing $\varepsilon > 0$ small enough, we have $\int_\Omega |\nabla u|^2 \leq C \int_\Omega f^2$, and using Theorem 4.28 again we obtain (4.11).

Now assume that $c \geq c_0$. We have

$$\int_\Omega |\nabla u|^2 + c_0 \int_\Omega |u|^2 \leq \frac{1}{2\varepsilon} \int_\Omega f^2 + \frac{\varepsilon}{2} \int_\Omega |u|^2. \tag{4.14}$$

Choosing $\varepsilon = c_0 > 0$, we obtain (4.11). □

4.5 L^∞-stability

Theorem 4.17 *Let Ω be a bounded domain. Let $u \in \mathcal{C}^2(\Omega) \cap \mathcal{C}(\bar{\Omega})$ solve*

$$-\Delta u + cu = f(x), \ x \in \Omega, \quad u(x) = g(x), \ x \in \partial\Omega,$$

where $c \geq 0$. Then there exists a constant C such that

$$\max_{\bar{\Omega}} |u| \leq C\big(\sup|f| + \max|g|\big).$$

Theorem 4.17 is a consequence of the following comparison principle.

Lemma 4.18 *Let $\mathcal{L} = -\Delta + c$. If $\mathcal{L}u \leq \mathcal{L}v$ in Ω and $u \leq v$ on $\partial\Omega$, then $u \leq v$ on $\bar{\Omega}$.*

As before, the comparison principle follows from the weak maximum principle.

Lemma 4.19 *If $\mathcal{L}u \leq 0$ in Ω, then*

$$\max_{\bar{\Omega}} u \leq \max_{\partial\Omega} u^+,$$

where $u^+ = \max(u, 0)$.

Proof: Assume first that $\mathcal{L}u < 0$ in Ω.

Let $x_0 = \operatorname{argmax}_{\bar{\Omega}} u(x)$. If $u(x_0) \le 0$ or $x_0 \in \partial\Omega$, there is nothing to prove. Assume now

$$x_0 \in \Omega, \quad u(x_0) > 0. \tag{4.15}$$

Then $\operatorname{Hess} u(x_0)$ is negative semi-definite, and hence $-\Delta u(x_0) \ge 0$. Since $c \ge 0$, we have $\mathcal{L}u(x_0) \ge 0$. This contradicts with (4.15), so (4.15) cannot happen.

For the general case $\mathcal{L}u \le 0$, we consider

$$u_\varepsilon(x) = u(x) + \varepsilon(e^{x_1} - L),$$

where L is sufficiently large so that $e^{x_1} - L \le 0$ for all $x \in \Omega$. Then

$$\mathcal{L}u_\varepsilon = \mathcal{L}u - \varepsilon e^{x_1} + c(x)\varepsilon(e^{x_1} - L) < 0.$$

By what have been proved,

$$\max_{\bar{\Omega}} u_\varepsilon \le \max_{\partial\Omega} u_\varepsilon^+.$$

Letting $\varepsilon \downarrow 0$ and using the fact $u_\varepsilon \to u$ uniformly in Ω, the inequality holds for u. □

Next we derive Theorem 4.17 from Lemma 4.18. **Proof:**

Let $w = G + F(M - \lambda e^{x_1})$ where

$$G = \max|g|, \quad F = \sup|f|.$$

We want to choose M and λ appropriately so that $\mathcal{L}w \ge f$ in Ω and $w \ge g$ on $\partial\Omega$. Indeed, suppose $\Omega \subset [-L, L] \times \mathbb{R}^{d-1}$; then

$$\mathcal{L}w = c\big(G + F(M - \lambda e^{x_1})\big) + F\lambda e^{x_1} \ge F,$$

provided that

$$\lambda e^{-L} \ge 1, \quad M \ge \lambda e^L,$$

which can be achieved by $\lambda = e^L$ and $M = e^{2L}$. For such λ and M, on $\partial\Omega$ we have

$$w \ge G \ge g.$$

Hence, $v = u - w$ satisfies

$$\mathcal{L}v \le 0 \text{ in } \Omega, \quad v \le 0 \text{ on } \partial\Omega,$$

and by Lemma 4.19,

$$\max_{\bar{\Omega}}(u - w) \le 0.$$

Therefore,

$$\max_{\bar{\Omega}} u \le \max_{\bar{\Omega}} w \le G + e^{2L}F.$$

Similarly, $z = -u - w$ satisfies

$$\mathcal{L}z \le 0 \text{ in } \Omega, \quad z \le 0 \text{ on } \partial\Omega,$$

and hence by Lemma 4.19

$$\max_{\bar{\Omega}}(-u) \le \max_{\bar{\Omega}} w \le G + e^{2L}F.$$

Combining these we obtain the desired conclusion. □

4.6 Perron's method

The goal of this section is to solve the Laplace equation

$$\Delta u(x) = 0,\ x \in \Omega, \quad u(x) = \phi(x),\ x \in \partial\Omega.$$

We follow [HL, Chap 6]

We first give two definitions for subharmonic functions.

Definition 4.1 *A function $v \in \mathcal{C}^2$ is subharmonic if $-\Delta v \leq 0$.*

Definition 4.2 *A continuous v is subharmonic if $v \leq w$ in B for every ball B and $w \in \mathcal{C}^2(B) \cap \mathcal{C}(\bar{B})$ harmonic such that $w \geq v$ on ∂B.*

In $d = 1$, harmonic functions are straight lines and subharmonic functions are convex functions. The advantage of the second definition is that it requires lower regularity.

Proposition 4.20 (Harmonic lifting) *Let $u \in \mathcal{C}(\bar{\Omega})$ be subharmonic and B be a ball such that $\bar{B} \subset \Omega$. Let w solve*

$$\Delta w(x) = 0,\ x \in B, \quad w(x) = u(x),\ x \in \Omega \setminus B.$$

Then w is subharmonic.

The function w is called a harmonic lifting of u in B. Such lifting exists since we can always solve the Laplace equation in a ball.

Proof: Let v be harmonic, $B_1 \subset \Omega$ and $v\big|_{\partial B_1} \geq w\big|_{\partial B_1}$. We want to show that

$$v \geq w, \text{ in } B_1.$$

Since $v\big|_{\partial B_1} \geq w\big|_{\partial B_1} \geq u\big|_{\partial B_1}$ and u is subharmonic, we have $v \geq u$ in B_1. Since $u = w$ in $B_1 \setminus B$, we have $v \geq w$ in $B_1 \setminus B$.

In $B_1 \cap B$, since $v \geq w$ on $\partial(B_1 \cap B) = (\partial B_1 \cap \bar{B}) \cup (\partial B \cap B_1)$ and v, w are harmonic, by comparison principle $v \geq w$ in $B_1 \cap B$.

This completes the proof. $\square$

Similarly, we can define superharmonic functions. We have the following comparison principle.

Lemma 4.21 *Let Ω be a bounded, connected domain. Let $u, v \in \mathcal{C}(\bar{\Omega})$. If u is subharmonic and v is superharmonic with $u \leq v$ on $\partial\Omega$, then $u \leq v$ in Ω.*

Proof: Let $M = \max_{\bar{\Omega}} u - v$ and

$$D = \{x \in \Omega : u(x) - v(x) = M\} \subset \Omega.$$

The set D is closed since it is the pre-image of a continuous function of a singleton $\{M\}$. We will show that D is also open in Ω, and hence $D = \varnothing$ or Ω. In both cases, the conclusion follows.

Let $x_0 \in D$. For $r > 0$ such that $B = B_r(x_0) \subset \Omega$, we consider the harmonic lifting of u, denoted by $\bar{u}$, and the harmonic lifting of v, denoted by $\bar{v}$, in the ball B. On the one hand, since $\bar{u}$ and $\bar{v}$ are harmonic lifting, we have

$$\bar{u}(x_0) - \bar{v}(x_0) \geq u(x_0) - v(x_0) = M.$$

On the other hand, for $x \in \partial B$,

$$\bar{u}(x) - \bar{v}(x) = u(x) - v(x) \leq M.$$

Hence,

$$(\bar{u}-\bar{v})(x_0) \geq \max_{\partial B}(\bar{u}-\bar{v}).$$

Since $\bar{u}-\bar{v}$ is a harmonic function in B, by the strong maximum principle, $\bar{u}-\bar{v} \equiv M$ in B, and therefore $u - v = M$ on ∂B. This holds for all $r > 0$ with $B_r(x_0) \subset \Omega$. So x_0 is in the interior of Ω. This proves the desired result. $\square$

Let $\phi \in \mathcal{C}(\partial\Omega)$. We define

$$\mathcal{S}_\phi = \{v \in \mathcal{C}(\bar{\Omega}) : \text{ subharmonic}, v \leq \phi \text{ on } \partial\Omega\}.$$

Note that if $u_1, u_2 \in \mathcal{S}_\phi$, then $\max(u_1, u_2) \in \mathcal{S}_\phi$. We will show that

$$\sup_{u \in \mathcal{S}_\phi} u$$

is harmonic with boundary condition ϕ. More precisely, we define

$$u_\phi(x) = \sup\{v(x) : v \in \mathcal{S}_\phi\}, \quad x \in \Omega.$$

Theorem 4.22 *$u_\phi(x) < \infty$ and u_ϕ is harmonic in Ω.*

Proof:

Since $w \equiv M := \max \phi$ is a (super)harmonic and

$$w(x) \geq v(x), \quad x \in \partial\Omega,$$

for all $v \in \mathcal{S}_\phi$, by Lemma 4.21, $w \geq v$ for all $v \in \mathcal{S}_\phi$. Hence, $u_\phi(x) \leq M$ for all $x \in \Omega$. In particular, u_ϕ is a well-defined function.

For every $x_0 \in \Omega$ and $B_{2r}(x_0) \subset \Omega$, we will show that u_ϕ is harmonic in $B = B_r(x_0)$.

First, by definition of u_ϕ, there exist $v_n \in \mathcal{S}$ such that $v_n(x_0) \uparrow u_\phi(x_0)$. We can assume that $v_n \geq m = \min \phi$; otherwise we can replace v_n by $\max(v_n, m) \in \mathcal{S}_\phi$. Let $\tilde{v}_n$ be the harmonic lifting of v_n in B. Then

$$v_n(x_0) \leq \tilde{v}_n(x_0) \leq u_\phi(x_0).$$

Since $\tilde{v}_n$ are bounded and harmonic, by Theorem 4.12 $\nabla\tilde{v}_n$ are uniformly bounded in $\bar{B}$. Hence, $\tilde{v}_n$ are uniformly bounded and equi-continuous in $\bar{B}$. By Azela–Ascoli, there exists a subsequence $\tilde{v}_{n_k}$ so that $\tilde{v}_{n_k} \to v_*$ uniformly in $\bar{B}$. Since the mean-value property is preserved under uniform convergence, v also has the mean-value property, and hence v_* is harmonic in B. Moreover, $v_*(x_0) = u_\phi(x_0)$.

Next, we will show that $v_* = u_\phi$ in B. Let $\bar{x} \in B$. There exist $w_n \in \mathcal{S}_\phi$ such that $w_n(\bar{x}) \uparrow u_\phi(\bar{x})$. We can assume $w_n \geq \tilde{v}_n$ in B; otherwise we can replace w_n by $\max(w_n, \tilde{v}_n)$. Let $\tilde{w}_n$ be the harmonic lifting of w_n. Again, there exists a further subsequence $\{n'_k\} \subset \{n_k\}$, so that $\tilde{w}_{n'_k} \to w_*$ uniformly in $\bar{B}$ and w_* is harmonic. Then $w_*(\bar{x}) = u_\phi(\bar{x})$. For any $x \in B$,

$$w_*(x) = \lim_{k\to\infty} w_{n'_k}(x) \geq \liminf_{k\to\infty} \tilde{v}_{n'_k}(x) = v_*(x).$$

On the other hand, $w_*(x_0) \leq u_\phi(x_0) = v_*(x_0)$. Hence, $w_*(x_0) = v_*(x_0)$. Therefore, the harmonic function $v_* - w_*$ achieves its maximum over $\bar{B}$ at $x_0 \in B$. By the strong maximum principle, $v_* - w_* = 0$ in B, and hence $w_*(\bar{x}) = v_*(\bar{x}) = u_\phi(\bar{x})$ for all $\bar{x} \in B$. This completes the proof. $\square$

The following result gives the boundary behavior of u_ϕ.

Theorem 4.23 *Let $x_0 \in \partial\Omega$. Assume that there exists w_{x_0} subharmonic such that*

$$w_{x_0}(x_0) = 0, \quad w_{x_0}(x) < 0, \ \forall x \in \partial\Omega \setminus \{x_0\}. \tag{4.16}$$

Then

$$\lim_{\Omega \ni y \to x_0} u_\phi(y) = \phi(x_0).$$

Proof: For every $\varepsilon > 0$, there exists $\delta > 0$ such that

$$|\phi(x) - \phi(x_0)| \le \delta, \quad \forall |x - x_0| \le \varepsilon, \ x \in \partial\Omega.$$

Since

$$\min_{x \in \partial\Omega \setminus B_\delta(x_0)} |w(x)| > 0,$$

there exists $K > 0$ such that

$$\frac{K}{2}|w(x)| \ge M := \max|\phi|, \quad \forall x \in \partial\Omega \setminus B_\delta(x_0).$$

Let $v(x) = \phi(x_0) - \varepsilon + Kw(x)$. Then $v(x)$ is subharmonic, and

$$v(x) \le \phi(x_0) - \varepsilon \le \phi(x), \quad x \in \partial\Omega \cap B_\delta(x_0),$$

$$v(x) \le \phi(x_0) - \varepsilon - 2M \le \phi(x), \quad \forall x \in \partial\Omega \setminus B_\delta(x_0).$$

Hence, $v(x) \in \mathcal{S}_\phi$, and thus by definition,

$$\phi(x_0) - \varepsilon + Kw(x) \le u_\phi(x), \quad \forall x \in \Omega.$$

Similarly, $\tilde{v}(x) = \phi(x_0) + \varepsilon - Kw(x)$ is super-harmonic and $\tilde{v}(x) \ge \phi(x)$ for $x \in \partial\Omega$, and hence $\tilde{v}(x) \ge u_\phi(x)$ for all $x \in \Omega$ by Lemma 4.21. Therefore,

$$|\phi(x_0) - u_\phi(x)| \le \varepsilon - Kw(x), \quad \forall x.$$

Letting $x \to x_0$, we have

$$\limsup_{x \to x_0} |\phi(x_0) - u_\phi(x)| \le \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, the LHS is 0, which completes the proof. $\square$

One sufficient condition for the existence of such function w such that (4.16) holds is the *exterior ball condition*, which holds for domain Ω with $\mathcal{C}^2$-boundary. Combining all these we have the following result.

Theorem 4.24 *Let Ω be a bounded domain with $\mathcal{C}^2$-boundary. For any $\phi \in \mathcal{C}(\bar{\Omega})$, there exists a function $u \in \mathcal{C}^\infty(\Omega) \cap \mathcal{C}(\bar{\Omega})$ solving*

$$-\Delta u = 0, \ \Omega, \quad u = \phi, \ \partial\Omega.$$

4.7 Green's function in general domains

We will use Theorem 4.24 to find the *Green's function.*

Let $y \in \Omega$. Recall that the Green's function $G(x, y)$ solves

$$\begin{cases} -\Delta_x G(x,y) = \delta(x-y), & x \in \Omega, \\ \quad G(x,y) = 0, & x \in \partial\Omega. \end{cases} \tag{4.17}$$

The term $\delta(x-y)$ is singular and thus problematic. We first use the fundamental function to remove it as follows. Recall that the fundamental solution $\Phi(x-y)$ solves

$$-\Delta_x \Phi(x-y) = \delta(x-y),$$

in the sense that $-\Delta(\Phi * f) = f$ for any bounded continuous function f. To find the Green's function, we write $G(x, y) = \Phi(x - y) - v(y)$, and look for v that solves

$$\begin{cases} \Delta v(x) = 0, & x \in \Omega, \\ \quad v(x) = \Phi(x-y), & x \in \partial\Omega. \end{cases} \tag{4.18}$$

The resulting G be a solution to (4.17) by the *principle of superposition.*

Using the explicit form of Φ, and that fact that $\operatorname{dist}(y, \partial\Omega) > 0$ for $y \in \Omega$, the boundary condition in (4.18) is $\mathcal{C}(\partial\Omega)$. Hence, Perron's method applies and there exists a classical solution $v \in \mathcal{C}^\infty(\Omega) \cap \mathcal{C}(\bar{\Omega})$ to (4.18).

Since $G(x, y) = \Phi(x - y) - v(x)$ and $\Phi(x - y)$ is smooth when $x \neq y$, we immediately know that $G(\cdot, y) \in \mathcal{C}^\infty(\Omega \setminus \{y\})$. Using the equation (4.17), by Theorem 4.5, the Green's function is symmetric, that is, $G(x, y) = G(y, x)$, and hence $G(x, y) \in \mathcal{C}^\infty(\Omega^2 \setminus \{x = y\})$.

Using the Green's function we can solve the Poisson equation

$$\begin{cases} -\Delta u = f, & \Omega, \\ \quad u = 0, & \partial\Omega, \end{cases} \tag{4.19}$$

whose solution is

$$u(x) = \int_\Omega G(x,y) f(y)\, dy,$$

as long as the source term f is nice enough so that the above integral makes sense; for example, $f \in \mathcal{C}(\Omega) \cap L^\infty(\Omega)$.

4.8 Dirichlet principle

In this section we present another way to solve the Poisson equation. Let I be a functional from $\mathcal{X}_g := g + \mathcal{C}_0^2(\Omega)$ to $\mathbb{R}$, defined by

$$I[u] := \int_\Omega \frac{1}{2}|\nabla u|^2 - fu, \tag{4.20}$$

where $f \in \mathcal{C}(\Omega) \cap L^2(\Omega)$ and $g \in \mathcal{C}(\partial\Omega)$. Assuming that there exists an extension of g to $\mathcal{C}^2(\Omega) \cap \mathcal{C}(\bar{\Omega})$, still denoted by g, we say that $u \in \mathcal{X}_g$ if $u - g \in \mathcal{C}_0^2(\Omega)$.

Here, we will be more careful about the distinction between $\mathcal{C}_0^k(\Omega)$, the space of functions that *vanish* on $\partial\Omega$, defined by

$$\mathcal{C}_0^k(\Omega) = \{v \in \mathcal{C}^k(\Omega) : \lim_{x \to \partial\Omega} |v(x)| = 0\},$$

and $\mathcal{C}_c^k(\Omega)$, the space of functions with *compact* support in Ω, defined by

$$\mathcal{C}_c^k(\Omega) = \{v \in \mathcal{C}^k(\Omega) : \exists \text{ compact } K \subset \text{s.t. } u = 0 \text{ in } K^c\}.$$

These two spaces are different; for example, for $\Omega = [-1, 1]$, the function $f = |x| - 1$ is in $\mathcal{C}_0^\infty(\Omega)$ but not $\mathcal{C}_c^\infty(\Omega)$, since $\operatorname{supp} u = [-1, 1] \not\subset (-1, 1)$. Although this distinction will not be so important later on, we will keep this in mind at this moment.

The *Dirichlet Principle* states that the "minimizer" of to the variation problem

$$\inf_{u \in \mathcal{X}_g} I[u] \tag{4.21}$$

will correspond to the solution to the Poisson equation

$$-\Delta u = f \text{ in } \Omega, \quad u = g \text{ on } \partial\Omega. \tag{4.22}$$

It is not obvious at all why $I[\cdot]$ has a minimizer in $\mathcal{X}_g$. However, in the rest of section we will explain why the problem of minimizing (4.20) is related to (4.22).

First, $I[\cdot]$ has a unique minimizer in $\mathcal{X}_g$.

Proof: We claim that

$$I[\frac{u_1 + u_2}{2}] \le \frac{1}{2} I[u_1] + \frac{1}{2} I[u_2]. \tag{4.23}$$

that is, $I[\cdot]$ is "convex" on its domain. Indeed, writing $w = (u_1 + u_2)/2$, we have

$$\begin{aligned}
\frac{1}{2} I[u_1] + \frac{1}{2} I[u_2] - I[w] &= \int_\Omega \frac{1}{4} |\nabla u_1|^2 + \frac{1}{4} |\nabla u_2|^2 - \frac{1}{8} |\nabla u_1 + \nabla u_2|^2 \\
&= \int_\Omega \frac{1}{8} |\nabla u_1 - \nabla u_2|^2 \ge 0.
\end{aligned}$$

The equality holds only if $|\nabla u_1 - \nabla u_2| \equiv 0$, since $|\nabla u_1 - \nabla u_2|^2$ integrates to 0 and is continuous. Since $u_1 - u_2 = 0$ on $\partial\Omega$, this implies $u_1 \equiv u_2$ on $\bar{\Omega}$.

Suppose that u_1 and u_2 are two minimizers of $I[\cdot]$ in $\mathcal{X}_g$, that is,

$$I[u_1] = I[u_2] = \inf_{u \in \mathcal{X}_g} I[u].$$

Then, by (4.23), we have $I[w] \le \inf_{\mathcal{X}_g} I[u]$, so w is also a minimizer, and the equality in (4.23) holds. Hence, we have $u_1 \equiv u_2$ on $\bar{\Omega}$, and this is the uniqueness. □

Second, if $u \in \mathcal{X}_g$ is a minimizer, then u solves (4.22).

To establish this, we need to understand the "derivative" of $I[\cdot]$, which is the so-called "calculus of variation". Recall that for a $\mathcal{C}^1$ function f, if $f(x_0)$ is the minimum, then by Fermat's lemma $f'(x_0) = 0$. So intuitively, if u is a minimizer of I, then $\frac{dI[u]}{du} = 0$.

But what is $\frac{dI}{du}$? The issue here is that $u \in \mathcal{X}_g$ and $\mathcal{X}_g$ is an infinite dimensional space, so much of our intuition for a function on $\mathbb{R}$ is useless. Let us consider instead a multivariate function $f : \mathbb{R}^d \to \mathbb{R}$. The gradient $\nabla f(x_0)$, is a vector, but it can also be seen as a linear map from $\mathbb{R}^d$ to $\mathbb{R}$, defined by

$$\big(\nabla f(x_0)\big)(h) = \nabla f(x_0) \cdot h = \frac{\partial f}{\partial h}(x_0) = \lim_{\varepsilon \to 0} \frac{f(x_0 + \varepsilon h) - f(x_0)}{\varepsilon}.$$

This motivates us to define some kind of "directional derivative" on $\mathcal{X}_g$.

Let $v \in \mathcal{C}_0^2(\Omega)$. Then $u + \varepsilon v \in \mathcal{X}_g$ for every ε. The function v will serve as the "direction".

Let $i(\varepsilon) = I[u+\varepsilon v]$. Let us compute $i'(\varepsilon)$. Note that everything is smooth so we can interchange the integral and differentiation. We have

$$i'(\varepsilon) = \int_\Omega \frac{d}{d\varepsilon}\Big[\frac{1}{2}|\nabla u + \varepsilon \nabla v|^2 - f(u+\varepsilon v)\Big] = \int_\Omega \nabla u \cdot \nabla v + \varepsilon |\nabla v|^2 - fv = \int_\Omega -\Delta u \cdot v + \varepsilon |\nabla v|^2 - fv,$$

where the boundary term $\int_{\partial\Omega} \frac{\partial u}{\partial n} v$ from the integration by parts in the last step is 0 since $v = 0$ on $\partial\Omega$. Hence,

$$i'(0) = \int_\Omega \Big(-\Delta u - f\Big)v. \tag{4.24}$$

The quantity (4.24) is called the *first variation* of $I[\cdot]$ (with respect to variation v). A necessary condition for u being a minimizer in $\mathcal{X}_g$ is that the first variation vanishes with respect to every variation $v \in \mathcal{C}_0^2(\Omega)$.

Since $-\Delta u - f \in \mathcal{C}(\Omega)$ and the first variation of $I[\cdot]$ is 0 for all v, by Lemma 4.25 below, we have

$$\Delta u(x) + f(x) = 0, \quad \forall x \in \Omega. \tag{4.25}$$

The equation (4.25) is the *Euler–Lagrange* equation associated with the variational problem (4.24). To summarize, a necessary condition for u to be a minimizer of a variation problem is that u solves the corresponding Euler–Lagrange equation.

Lemma 4.25 *Let $\varphi \in \mathcal{C}(\Omega)$ be such that*

$$\int_\Omega \varphi(x) v(x)\, dx = 0, \quad \forall v \in C_0^\infty(\Omega). \tag{4.26}$$

Then $\varphi \equiv 0$ in Ω.

Proof: We will prove by contradiction. If φ is not identically 0, without loss of generality we can assume that $\varphi(x_0) > 0$ for some $x_0 \in \Omega$. Since Ω is open and φ is continuous, there exist $\varepsilon, \delta > 0$ such that $\varphi(x_0) \geq \varepsilon$ in $B_\delta(x_0) \subset \Omega$. Let

$$v(x) = \delta^{-d}\eta\big(\delta^{-1}(x-x_0)\big), \quad \eta(x) = \begin{cases} e^{-\frac{1}{(1-|x|^2)}}, & |x| < 1, \\ 0, & |x| \geq 1. \end{cases} \tag{4.27}$$

Then

$$\int_\Omega \varphi(x) v(x)\, dx \geq \varepsilon \int_{B_\delta(x_0)} \delta^{-d}\eta\big(\delta^{-1}(x-x_0)\big) = \varepsilon \int_{B_1(0)} e^{-\frac{1}{1-|x|^2}} > 0, \tag{4.28}$$

which is a contradiction. □

However, a priori the variation problem (4.21) may not have a minimizer, and even if a minimizer exists, it can be outside of $\mathcal{X}_g$, since from the expression of $I[\cdot]$, its definition should require $\mathcal{C}^1$ differentiability at most, rather than $\mathcal{C}^2$.

To illustrate, let us consider the variation problem

$$\inf\Big\{\int_0^1 \big((\partial_x u)^2 - 1\big)^2\, dx : u \in \mathcal{C}^1[0,1],.\ u(0) = a,.\ u(1) = b\Big\}, \quad a < b < a+1. \tag{4.29}$$

Since $a \leq b < a+1$, the function

$$v(x) = \begin{cases} x + a, & 0 \leq x < \frac{b+1-a}{2}, \\ b + (1-x), & \frac{b+1-a}{2} \leq x \leq 1 \end{cases} \tag{4.30}$$

is well-defined and achieves the smallest possible infimum 0 in (4.29), except that it is not $\mathcal{C}^1$ at $x = x_0 := \frac{b+1-a}{2}$. But we can make change to v in an arbitrary small neighborhood around x_0, so that the resulting function is $\mathcal{C}^1$ and makes (4.29) arbitrarily close to 0. On the other hand, if a function $u \in \mathcal{C}^1$ taking slope ± 1, then by continuity of derivative, $\partial_x u \equiv 1$ or -1, so it cannot satisfy the boundary condition in (4.29). Combining all these together, we can say that (4.29) does not have a $\mathcal{C}^1$ minimizer.

But if we include piecewise $\mathcal{C}^1$ functions in the domain for (4.29), the minimizer will not be unique, since there are an infinite number of polygon curves with slope ± 1 connecting $(0, a)$ and $(1, b)$.

4.8.1 Weak derivatives and solutions

How do we obtain a minimizer to (4.21)? By definition of the infimum, there exists a sequence $(u_n) \subset \mathcal{X}_g$ such that $I[u_n] \to \inf I[u]$; such sequence is called a "minimizing sequence". We hope that there exists some limit point u_* of the minimizing sequence. However, as we have seen in (4.29), the limit point u_* may fall out of the original domain of the functional, due to lack of continuous derivative.

To overcome the above mentioned issue, we need to generalize our notion of derivatives, as well as our notion of solutions. This is done by the introduction of weak derivatives and weak solutions.

Recall the multi-index notion for derivative:

$$D^\alpha f := \partial_{x_1}^{\alpha_1} \cdots \partial_{x_d}^{\alpha_d} f, \quad \alpha = (\alpha_1, \alpha_2, \dots, \alpha_d).$$

Also recall that $L^1_{loc}(\Omega)$ is the space of functions that are absolutely integrable on any compact sets $K \subset \Omega$; for example, x^{-1} is in $L^1_{loc}(0,1)$ but not $L^1_{loc}(-1,1)$.

Let $u, v \in L^1_{loc}(\Omega)$. We say that $v = D^\alpha u$ in the weak sense, or v is the α-th *weak derivative* of u, if

$$\int_\Omega \varphi v = \int_\Omega (-1)^{|\alpha|} (D^\alpha \varphi) u, \quad \forall \varphi \in \mathcal{C}_c^\infty(\Omega). \tag{4.31}$$

To see the motivation, (4.15) is integration by parts (with no boundary terms since φ vanishes at the boundary), if v is a classical derivative of u. For the Poisson equation (4.22), we say that u is a weak solution if $-\Delta u = f$ holds in the weak sense, that it,

$$\int_\Omega (\Delta \varphi) u + \varphi f = 0, \quad \forall \varphi \in \mathcal{C}_c^\infty(\Omega).$$

As an example, let $u(x) = |x| \in L^1_{loc}(\mathbb{R})$. Then

$$u'(x) = \begin{cases} 1, & x > 0, \\ -1, & x < 0, \end{cases}$$

is the first-order weak derivative of u.

But u' is not further differentiable in the weak sense. Otherwise, suppose $v = u'$, then for any $\varphi \in \mathcal{C}_0^\infty(\mathbb{R})$,

$$\int \varphi(x) v(x)\, dx = -\int \varphi'(x) u'(x)\, dx. \tag{4.32}$$

For $a < b$ and any $n \geq 1$, it is not hard to construct $\varphi_n \in \mathcal{C}_c^\infty(\mathbb{R})$ so that

$$\varphi_n(x) \begin{cases} = 0, & x \notin (a, b), \\ = 1, & x \in [a + 1/n, b - 1/n], \\ \in (0, 1), & \text{otherwise}. \end{cases}$$

The function φ_n will approximate $\mathbb{1}_{(a,b)}$, the indicator function of the interval (a, b). Then taking $\varphi = \varphi_n$ in (4.32) and letting $n \to \infty$, we obtain in the limit

$$\int_a^b v(x)\,dx = -\lim_{n\to\infty} \int_a^{a+1/n} \varphi'(x)u'(x)\,dx + \int_{b-1/n}^b \varphi'(x)u'(x)\,dx = u'(b) - u'(a). \tag{4.33}$$

Now take $(a, b) = (-\varepsilon, \varepsilon)$ and let $\varepsilon \to 0$. On the one hand the right hand side in (4.33) is $1-(-1) = 2$, on the other hand since $|\mathbb{1}_{(-\varepsilon,\varepsilon)}v| \le |v|$ and v is locally integrable, by dominated convergence theorem

$$\lim_{\varepsilon\to 0+} \int_{-\varepsilon}^{\varepsilon} v(x)\,dx = \lim_{\varepsilon\to 0+} \int_{\mathbb{R}} \mathbb{1}_{(-\varepsilon,\varepsilon)}(x)v(x)\,dx = \int_{\mathbb{R}} \lim_{\varepsilon\to 0+} \mathbb{1}_{(-\varepsilon,\varepsilon)}(x)v(x)\,dx = \int_{\mathbb{R}} 0\,dx = 0.$$

This gives a contradiction.

In PDE theories, weak solutions allow more flexibility to obtain a solution, and after that there are other means to show that the so obtained solution has the desired smoothness, and thus the weak solution becomes the classical solution. These two parts will rely on different sets of tools. In this note we will focus on the existence part. The following result gives an example of the other part.

Proposition 4.26 *If $\Delta u = 0$ in the weak sense, then u is a harmonic function and $\mathcal{C}^\infty$.*

Proof: Let $\eta_\varepsilon \in \mathcal{C}^\infty(\mathbb{R}^d)$ be the standard smooth mollifiers. We will use the fact that (η_ε) is also an approximate identity, so that $\eta_\varepsilon * f \to f$ almost everywhere and in L^1_{loc} for any $f \in L^1_{loc}$.

Let $u_\varepsilon = u * \eta_\varepsilon$. Then $u_\varepsilon \in \mathcal{C}^\infty$ and for every $\varphi \in \mathcal{C}_c^\infty$,

$$\int (D^\alpha\varphi)u_\varepsilon = \int D^\alpha\varphi\cdot(u*\eta_\varepsilon) = \int (D^\alpha\varphi*\eta_\varepsilon)\cdot u = \int D^\alpha(\varphi*\eta_\varepsilon)u == \int (-1)^{|\alpha|}(\varphi*\eta_\varepsilon)D^\alpha u = \int (-1)^{|\alpha|}\varphi\cdot(D^\alpha u*\eta_\varepsilon),$$

where we use $\int f(g*h) = \int (f*h)g$. Hence, $D^\alpha u_\varepsilon = (D^\alpha u) * \eta_\varepsilon$ in the weak sense. But $u_\varepsilon \in \mathcal{C}^\infty$, so the weak derivative is strong derivative. In particular, $\Delta u_\varepsilon = 0$ and u_ε is harmonic.

Using Theorem 4.12, for any compact set K, there exists $K_1 \supset K$ and constant C depending on K, K_1, such that

$$\sup_K |u_\varepsilon(x)|, \sup_K |\nabla u_\varepsilon(x)| \le C|u_\varepsilon|_{L^1(K_1)} \le C|u|_{L^1(K_1)}.$$

Since u is locally integrable, (u_ε) is uniformly bounded and equi-continuous on K. By Arzelà–Ascoli, there exists a subsequence u_{ε_n} and u_* such that $u_{\varepsilon_n} \to u_*$ uniformly on K, and due to the mean-value property for harmonic function, the limiting function u_* is also harmonic. On the other hand, the sequence (u_ε) has a unique possible limit point which is u itself. Therefore, u is harmonic. □

4.8.2 Sobolev spaces and weak convergence

With the weak derivative, we can define the functional (4.21) on the largest possible domain. This leads to the introduction of certain *Sobolev spaces*.

For $k \ge 0$, let us define

$$H^k(\Omega) = \{u \in L^1_{loc}(\Omega) : D^\alpha u \in L^2(\Omega),.\ \forall|\alpha| \le k\}.$$

There is a natural norm on $H^k(\Omega)$:

$$\|u\|_{H^k(\Omega)} := \sum_{|\alpha|\le k} \|D^\alpha u\|_{L^2(\Omega)},$$

and under this norm, $H^k(\Omega)$ becomes a complete space, meaning that every Cauchy sequence under this norm admits a limit in $H^k(\Omega)$.

Next, we try to define the boundary condition on $H^k(\Omega)$. As the simplest example we will treat the zero boundary condition. We define

$$H_0^k(\Omega) = \text{closure of } \mathcal{C}_c^\infty. \text{ under } \|\cdot\|_{H^k(\Omega)}. \tag{4.34}$$

Note that $\mathcal{C}_0^\infty(\Omega) \subset H_0^k(\Omega)$, but there are more functions in (4.34). We say that $u \in g + H_0^k(\Omega)$ if $u - g \in H_0^k(\Omega)$, where $g \in \mathcal{C}^k(\Omega) \cap \mathcal{C}(\partial\Omega)$.

The function $I[\cdot]$ in (4.20) will make sense for all $u \in g + H_0^1(\Omega)$, where $f \in L^2(\Omega)$ and $g \in \mathcal{C}^1(\Omega) \cap \mathcal{C}(\partial\Omega)$: for the first term $\int |\nabla u|^2$, the gradient ∇u is a weak derivative and is in $L^2(\Omega)$; for the second term, by Cauchy–Schwartz, we have

$$\left| \int_\Omega fu \right| \le \left[\int_\Omega f^2 \right]^{1/2} \left[\int_\Omega u^2 \right]^{1/2},$$

so $u \mapsto \int_\Omega fu$ is a linear functional on $L^2(\Omega) \supset g + H_0^1(\Omega)$.

Weak convergence The next problem is how to extract limit points for a minimizing sequence. Recall that a sequence (x_n) in $\mathbb{R}^d$ has a limit point if and only if x_n are bounded. We can rephrase it as "a set $K \in \mathbb{R}^d$ is sequentially precompact if and only if K is bounded". One naturally expects similar results in H^k. Unfortunately, this is false.

As a counter-example, consider $\mathcal{X} = L^2(0, 2\pi) = H_0^0(0, 2\pi)$ and $f_n = \frac{1}{\sqrt{\pi}} \sin(nx)$. Note that f_n are orthonormal, so

$$\|f_n - f_m\|^2 = \int f_n^2 - 2f_n f_m + f_m^2 = \int f_n^2 + f_m^2 \equiv 2, \quad \forall n \ne m.$$

Hence f_n is bounded in $\mathcal{X}$ but cannot have any limit point since any of its subsequences fails to be Cauchy.

We need a more general notion of convergence. We say that u_n converges to $H^k(\Omega)$ *weakly*, denoted by $u_n \rightharpoonup u$, if

$$\lim_{n\to\infty} \int_\Omega \varphi D^\alpha u_n = \int_\Omega \varphi D^\alpha u, \quad \forall \varphi \in \mathcal{C}_c^\infty(\Omega),\ \forall |\alpha| \le k.$$

For weak convergence we have the following powerful result.

Theorem 4.27 *A set in $H^k(\Omega)$ is weakly sequentially precompact if and only if it is bounded in the $\|\cdot\|_{H^k(\Omega)}$ norm.*

In the previous counter-example, $f_n \rightharpoonup 0$. This follows from the *Riemann–Lebesgue Lemma*, which states for any $g \in L^1(\mathbb{R})$,

$$\lim_{n\to\infty} \int g(x) \sin(nx)\, dx = 0.$$

Poincaré inequality Recall that the $H_0^1(\Omega)$ norm is given by

$$\|f\|_{H_0^1(\Omega)}^2 = \int_\Omega |f(x)|^2 + |\nabla f(x)|^2\, dx.$$

Theorem 4.28 *Let Ω is bounded and $u \in H_0^1(\Omega)$. There exists a constant K depending on the diameter of Ω such that*

$$\int_\Omega |u(x)|^2\, dx \le K \int_\Omega |\nabla u|^2\, dx. \tag{4.35}$$

Proof: It suffices to establish (4.35) for $u \in \mathcal{C}_c^\infty(\Omega)$. Indeed, since $\mathcal{C}_c^\infty(\Omega)$ is dense in $H_0^1(\Omega)$, for any $u \in H_0^1(\Omega)$, there exist $u_n \in \mathcal{C}_c^\infty(\Omega)$ that converge to u in $H_0^1(\Omega)$. Then

$$\|u\|_{L^2(\Omega)} = \lim_{n\to\infty} \|u_n\|_{L^2(\Omega)} \le C \lim_{n\to\infty} \|\nabla u_n\|_{L^2(\Omega)} = C\|\nabla u\|_{L^2(\Omega)}.$$

Now assume that $u \in \mathcal{C}_c^\infty(\Omega)$. Without loss of generality, we assume that $\Omega \subset [0, L] \times \mathbb{R}^{d-1}$ for some $L > 0$. Then, there exists an extension of u to $\mathbb{R}^d$, still denoted by u. For $x_1 \in (0, L)$, by Cauchy–Schwartz, we have

$$\begin{aligned}
|u(x_1, x_2, \dots, x_d)|^2 &= |u(x_1, x_2, \dots, x_d) - u(0, x_2, \dots, x_d)|^2 \\
&\le \Big[\int_0^{x_1} \big|(\partial_1 u)(s, x_2, \dots, x_d)\big|\, ds \Big]^2 \\
&\le \int_0^{x_1} 1\, dx \cdot \int_0^{x_1} \big|(\partial_1 u)(s, x_2, \dots, x_d)\big|^2 ds \\
&\le L \cdot \int_0^L \big|\nabla u(s, x_2, \dots, x_d)\big|^2 ds.
\end{aligned}$$

Integrating over $(x_2, \dots, x_d) \in \mathbb{R}^{d-1}$, we obtain (4.35) with $K = \sqrt{L}$. □

Recall that we want to solve the equation (4.22). Let $g \in \mathcal{C}(\partial\Omega)$, $f \in \mathcal{C}(\bar{\Omega})$ and $\mathcal{X}_g = g + \mathcal{C}_0^2(\Omega)$. We can define the functional $I[u]$ by (4.20). The "Dirichlet principle" says that the minimizer of $I[u]$ in $\mathcal{X}_g$ will solve (4.22).

We find minimizers through a "minimizing sequence", as we did for continuous functions. Let $u_n \in \mathcal{X}_g$ be such that $I[u_n] \to \inf I[\cdot]$. We hope that there exists some u_* such that

$$u_n \to u_*, \tag{4.36a}$$

$$I[u_n] \to I[u_*]. \tag{4.36b}$$

There will be two issues.

Issue 1. The sequence u_n may have no limit point in $\mathcal{X}_g$, as in the variational problem (4.29). This is because the space $\mathcal{X}_g$ is too restrictive. For this reason we introduce the concepts of the weak convergence and weak solutions.

Issue 2. If u_* is a weak solution, is u_* a classical solution? The answer is yes in most cases, but we omit the discussion here. We presented an example in this direction, Proposition 4.26.

We point out that a special case of Theorem 4.27 is the following.

Proposition 4.29 *Let $u_n \in H^1(\Omega)$ be such that*

$$\int_\Omega |u|^2 + |\nabla u|^2 \le M, \quad \forall n \ge 1$$

for some $M > 0$. Then, there exists $u_* \in H^1(\Omega)$ and a subsequence (u_{n_k}) such that $u_{n_k} \rightharpoonup u_*$ in $H^1(\Omega)$, that is,

$$\int_\Omega u_{n_k} v \to \int_\Omega u_* v, \quad \int_\Omega \partial_{x_i} u_{n_k} v \to \int_\Omega \partial_{x_i} u_* v, \quad \forall v \in L^2(\Omega).$$

4.8.3 Existence of weak solution

Since $I[\cdot]$ includes the term $\int_\Omega |\nabla u|^2$ which is part of the $H^1(\Omega)$-norm, it is no hard to see that $I[\cdot]$ is continuous in the norm of $\|\cdot\|_{H^1}$, that is, $I[u_n] \to I[u]$ if $u_n \to u$ in $H^1(\Omega)$. But we cannot expect $I[\cdot]$ to be continuous w.r.t. the weak convergence. To show that the weak limit attains the minimum of $I[\cdot]$, we will establish the *weakly lower semi-continuity* of the functional.

Proposition 4.30 *[Lower semi-continuity in weak topology] If $u_m \rightharpoonup u$ in H^1, then*

$$\liminf_{m\to\infty} I[u_m] \geq I[u].$$

Proof: We have

$$\int_\Omega u_m f \to \int_\Omega uf,$$

since $u_m \rightharpoonup u$ in L^2.

For the other term, we have

$$\int |\nabla u_m|^2 - |\nabla u|^2 = \int |\nabla u_m - \nabla u|^2 + 2\nabla u \cdot \big(\nabla u_m - \nabla u\big) \geq \int 2\nabla u \cdot \big(\nabla u_m - \nabla u\big).$$

Since $\nabla u \in L^2$ and $\nabla u_m \rightharpoonup \nabla u$ in L^2, we have

$$\liminf_{m\to\infty} \int |\nabla u_m|^2 - |\nabla u|^2 \geq \lim_{m\to\infty} \int 2\nabla u \cdot \big(\nabla u_m - \nabla u\big) = 0.$$

This completes the proof. □

We are ready to prove the following result. We assume $f \in L^2(\Omega)$ in (4.20).

Proposition 4.31 *There exists a minimizer of $I[\cdot]$ in $\tilde{\mathcal{X}}_g = g + H_0^1(\Omega)$.*

Proof: Let $u_n \in \tilde{\mathcal{X}}_g$ be a minimizing sequence of $I[\cdot]$. Then $I[u_n] \leq M$ for some $M > 0$, and $v_n = u_n - u_1 \in H_0^1(\Omega)$.

To apply Proposition 4.29, we need to bound $\|v_n\|_{H^1}$ uniformly from above. By Poincaré inequality Theorem 4.28, it suffices to bound $|\nabla v_n|_{L^2}$.

Below C will stand for a generic constants independent of v_n, which may change from line to line. We have

$$\begin{aligned}
I[u_n] &= \int \frac{1}{2}\big|\nabla u_1 + \nabla v_n\big|^2 - f(u_1 + v_n) \\
&\geq \frac{1}{2}\int |\nabla u_1|^2 + |\nabla v_n|^2 - \int |\nabla u_1| \cdot |\nabla v_n| - \int f u_1 - \int |f| \cdot |v_n| \\
&\geq C + \frac{1}{2}\int |\nabla v_n|^2 - \frac{1}{2\varepsilon}\int |\nabla u_1|^2 - \frac{\varepsilon}{2}\int |\nabla v_n|^2 - \frac{1}{2\varepsilon}\int |f|^2 - \frac{\varepsilon}{2}\int |v_n|^2 \\
&\geq C + \Big(\frac{1}{2} - \frac{\varepsilon(1+K)}{2}\Big)\int |\nabla v_n|^2,
\end{aligned}$$

where we use $ab \leq \frac{1}{2\varepsilon}a^2 + \frac{\varepsilon}{2}b^2$ in the third line, and K in the last line is the constant from Theorem 4.28. By choosing $\varepsilon > 0$ small enough so that

$$\frac{1}{2} - \frac{\varepsilon(1+K)}{2} > 0,$$

we obtain

$$\int |\nabla v_n|^2 \leq C\big(I[u_n] + 1\big).$$

Since $I[u_n]$ is uniformly bounded from above, we have a uniform upper bound on $\|v_n\|_{H^1}$ as desired.

By Proposition 4.29, there exists v_* and a subsequence v_{n_k} such that $v_{n_k} \rightharpoonup v_*$ in H^1, and hence $u_{n_k} \rightharpoonup u_1 + v_* =: u_*$ in H^1.

By Proposition 4.30 we have

$$\liminf_{k\to\infty} I[u_{n_k}] \geq I[u_*].$$

But the LHS is $\inf I[\cdot]$ on $\tilde{\mathcal{X}}_g$, so $I[u_*]$ achieves the minimum of I. This completes the proof. $\square$

4.8.4 Free boundary condition

Next we brief discuss the Neumann boundary condition,

$$-\Delta u = f \text{ in } \Omega, \quad \frac{\partial u}{\partial n} = 0 \text{ on } \partial\Omega. \tag{4.37}$$

The first important thing is that a "compatibility condition" has to be satisfied for (4.37) to have any solutions at all.

Proposition 4.32 *There can exist a solution for (4.37) only if $\int_\Omega f = 0$.*

Proof: From integration by parts, we have

$$0 = \int_{\partial\Omega} \frac{\partial u}{\partial n} \cdot 1 = \int_\Omega (\Delta u) \cdot 1 = \int_\Omega -f.$$

$\square$

As a consequence, the functional $I[u]$ is invariant under addition of a constant to u, namely,

$$I[u + C] = I[u], \quad \forall C \in \mathbb{R}.$$

To define the variational problem, the functional I takes the same form, but the domain changes to $H^1(\Omega)$, that is, no boundary condition is imposed at all. That is why the boundary condition in (4.37) is also called "free boundary condition".

Proposition 4.33 *u is a minimizer of $I[u]$ in $\mathcal{C}^2(\Omega) \cap \mathcal{C}^1(\bar{\Omega})$ if and only if it solves (4.37).*

Proof: The "if" direction is similar as before. We will prove the "only if" part here.

Let u be a minimizer. Then for any $\varphi \in \mathcal{C}_0^\infty(\Omega)$, $u + \varphi \in \mathcal{C}^2(\Omega) \cap \mathcal{C}^1(\bar{\Omega})$ and hence

$$i(\varepsilon) = I[u + \varepsilon\varphi] \geq I[u], \quad \forall \varepsilon > 0.$$

As before, we can derive the first variation of $I[\cdot]$ by computing $i'(0)$:

$$i'(0) = \int_\Omega \nabla u \cdot \nabla \varphi - f\varphi = \int_\Omega (-\Delta u - f)\varphi + \int_{\partial\Omega} \frac{\partial u}{\partial n}\varphi. \tag{4.38}$$

Since $\varphi = 0$ on $\partial\Omega$, the second term is 0, so by Lemma 4.25, $\Delta u + f = 0$ in Ω.

Now let $\varphi \in \mathcal{C}^\infty(\bar{\Omega})$ be arbitrary. (4.38) still holds, but the first term is zero since $\Delta u + f = 0$ in Ω. Therefore,

$$\int_{\partial\Omega} \frac{\partial u}{\partial n}\varphi = 0, \quad \forall \varphi \in \mathcal{C}^\infty(\bar{\Omega}).$$

This will imply $\frac{\partial u}{\partial n} = 0$ on $\partial\Omega$, similar to Lemma 4.25. $\square$

As before, let u_n be a minimizing sequence. We want to use Proposition 4.29 to extract a convergent subsequence. But (4.35) cannot be true for any $u \in H^1(\Omega)$, since by adding a constant to u, the RHS is the same but the LHS can get arbitrarily large. On the other hand, by Proposition 4.32, the functional $I[u]$ is invariant under addition of constants. We may take advantage of that.

Proposition 4.34 *Let Ω be a bounded domain. There exists $K = K(\Omega)$ such that*

$$\int_\Omega |u - \bar{u}|^2 \le K \int_\Omega |\nabla u|^2, \quad \bar{u} := \frac{1}{|\Omega|} \int_\Omega u. \tag{4.39}$$

Proof: To illustrate the idea, we treat the case in one dimension.

Let $\Omega = (a, b)$. Then $H^1(a, b)$ coincides with the space of absolutely continuous function on (a, b) with $L^2(\Omega)$ derivative.

By the intermediate value theorem, there exists $x_0 \in (a, b)$ such that $u(x_0) = \bar{u}$. For any $x \in (a, b)$, by Cauchy–Schwartz, we have

$$|u(x) - u(x_0)|^2 \le \left[\int_{x_0}^x |u'(s)|\, ds\right]^2 \le (b - a) \int_a^b |u'(s)|^2\, ds.$$

Integrating over x we obtain (4.39) with $K = (b - a)^2$. $\square$

Now we can prove the existence of minimizer of $I[\cdot]$ in $H^1(\Omega)$.

Proposition 4.35 *There exists $u_* \in H^1(\Omega)$ such that*

$$I[u_*] = \inf_{u \in H^1(\Omega)} I[u].$$

Proof: Let u_n be a minimizing sequence. Since $I[\cdot]$ does not change after adding a constant to u, we can assume $\int_\Omega u_n = 0$, otherwise we can subtract $\bar{u}_n$ from u_n. Hence, by Proposition 4.34, $\|u_n\|_{L^2} \le K \|\nabla u_n\|_{L^2}$. The rest follows the same argument as in Proposition 4.31. $\square$

5 Wave equation

In this section, we study the wave equation

$$\begin{cases} u_{tt} - \Delta u = f(t, x), & t > 0,\ x \in \Omega, \\ u(0, x) = g(x), & x \in \Omega, \\ u_t(0, x) = h(x), & x \in \Omega, \end{cases} \tag{5.1}$$

plus some boundary condition. Note that the second-order t-derivative appears in the equation, so consequently, we need to impose another initial condition on $u_t(0, \cdot)$.

We introduce the D'Alembert notation

$$\square := \partial_{tt} - \Delta.$$

5.1 One dimension, D'Alembert's formula

5.1.1 Fourier transform

Consider the wave equation on the whole space:

$$\partial_{tt} u(t,x) = \Delta u(t,x), \quad x \in \mathbb{R}^d.$$

Let $\hat{u}(t,\xi) = \big[u(t,\cdot)\big]^{\wedge}$. Then $\hat{u}$ solves

$$u_{tt}\hat{u}(t,\xi) = -4\pi^2|\xi|^2\hat{u}(t,\xi),$$

with initial condition

$$\hat{u}(0,\xi) = \hat{g}(\xi), \quad \hat{u}_t(0,\xi) = \hat{h}(\xi).$$

The solution to this linear ODE is given by

$$\hat{u}(t,\xi) = \hat{g}(\xi)\cos(2\pi|\xi|t) + \frac{\hat{h}(\xi)}{2\pi|\xi|}\sin(2\pi|\xi|t).$$

Hence,

$$u(t,\cdot) = \big[\cos(2\pi|\xi|t)\big]^{\vee} * g + \left[\frac{\sin(2\pi|\xi|t)}{2\pi|\xi|}\right]^{\wedge} * h. \tag{5.2}$$

However, it is a non-trivial task to compute the inverse Fourier transform, except in dimension $d = 1$.

Since $\cos\theta = \frac{1}{2}(e^{i\theta} + e^{-i\theta})$, we have

$$\begin{aligned}\cos(2\pi|\xi|t)^{\vee} &= \frac{1}{2}\int_{\mathbb{R}} e^{2\pi i\xi(x+t)} + e^{2\pi i\xi(x-t)}\,d\xi \\ &= \frac{1}{2}\big[\delta(x+t) + \delta(x-t)\big].\end{aligned}$$

Noting that

$$\frac{\sin(2\pi|\xi|t)}{2\pi|\xi|} = \frac{\sin(2\pi\xi t)}{2\pi\xi} = \int_0^t \cos(2\pi\xi s)\,ds,$$

we have

$$\begin{aligned}\left[\frac{\sin(2\pi|\xi|t)}{2\pi|\xi|}\right]^{\vee} &= \int_0^t [\cos(2\pi\xi s)]^{\vee}\,ds \\ &= \int_0^t \frac{1}{2}\big[\delta(x+s) + \delta(x-s)\big]\,ds \\ &= \frac{1}{2}\big(\mathbb{1}_{\{0\le -x\le t\}} + \mathbb{1}_{\{0\le x\le t\}}\big) \\ &= \frac{1}{2}\mathbb{1}_{[-t,t]}(x).\end{aligned}$$

We obtain the D'Alembert formula which solves the wave equation on $\mathbb{R}^1$.

Theorem 5.1 (D'Alembert Formula) *Let $g \in \mathcal{C}^2(\mathbb{R})$ and $h \in \mathcal{C}^1(\mathbb{R})$. Then*

$$u(t,x) = \frac{1}{2}\big[g(x-t) + g(x+t)\big] + \frac{1}{2}\int_{x-t}^{x+t} h(y)\,dy, \tag{5.3}$$

is a classical solution to (5.1).

In particular,

$$\lim_{(t,x)\to(0,x_0)} u(t,x) = g(x_0), \qquad \lim_{(t,x)\to(0,x_0)} u_t(t,x) = h(x_0), \qquad \forall x_0 \in \mathbb{R}.$$

5.1.2 Method of characteristics

We will give another derivation of (5.3) using the method of characteristics.

Since
$$\Box = (\partial_t - \partial_x)(\partial_t + \partial_x) = (\partial_t + \partial_x)(\partial_t - \partial_x),$$
for any $F \in \mathcal{C}^1$ we have
$$0 = \Box F(x - t) = \Box F(x + t).$$
Thus, we postulate that $u(t, x)$ takes the form
$$u(t, x) = F(x + t) + G(x - t), \tag{5.4}$$
that is, the sum of two *travelling wave solutions*. From the initial conditions, we obtain
$$\begin{cases} F(x) + G(x) = g(x), \\ F'(x) - G'(x) = h(x), \end{cases} \implies \begin{cases} F(x) = \dfrac{1}{2}g(x) + \dfrac{1}{2}\displaystyle\int_0^x h(y)\, dy, \\ G(x) = \dfrac{1}{2}g(x) - \dfrac{1}{2}\displaystyle\int_0^x h(y)\, dy. \end{cases}$$
This gives (5.3). Indeed, this is the unique classical solution.

Uniqueness: Assume that $u(t, x)$ is a classical solution to (5.1) on $\mathbb{R}^1$. Since
$$(\partial_t - \partial_x)[(\partial_t + \partial_x)u] = 0,$$
we have
$$(\partial_t + \partial_x)u = F(x + t) \tag{5.5}$$
by method of characteristics. To solve (5.5), let
$$\eta(t) = u(t, x_0 + t).$$
Then
$$\begin{aligned} \eta(t) &= \eta(0) + \int_0^t f(x_0 + 2s)\, ds \\ &= u(0, x_0) + F(2t + x_0) - F(x_0) \\ &= g(x - t) + F(x + t) - F(x - t). \end{aligned}$$
Therefore, any classical solution must take the form of (5.4) $\square$

5.1.3 Application of D'Alembert formula

From (5.3), we see that $u(t, x_0) = 0$ if
$$g(x) = f(x) = 0, \quad \forall |x - x_0| \le t.$$
This can be interpreted as the initial data has effect on location of distance less than t at time t, that is, the wave speed is 1.

In general, for the wave equation
$$\partial_{tt}u = a^2 \Delta u, \quad a > 0,$$

the wave speed is a. As an example, we recall that from the Maxwell's equation in the vacuum

$$\nabla \cdot E = 0, \quad \nabla \times E = -\frac{\partial B}{\partial t}, \quad \nabla \cdot B = 0, \quad \nabla \times B = \mu_0 \varepsilon_0 \frac{\partial E}{\partial t},$$

one can derive

$$\partial_{tt} E = \mu_0 \varepsilon_0 \Delta E, \quad \partial_{tt} B = \mu_0 \varepsilon_0 \Delta B.$$

So the speed of electro-magnetic wave, the light speed, is $c = \sqrt{\mu_0 \varepsilon_0}$, where μ_0 and ε_0 are physical constants that can be measured from experiments.

We can also use Theorem 5.1 to solve wave equation on the half-line using the *reflection principle*. Consider the wave equation on $\mathbb{R}_+$ with Dirichlet boundary condition:

$$\begin{cases} u_{tt} = u_{xx}, & t > 0,\ x > 0, \\ u(0, x) = g(x), & x > 0, \\ u_t(0, x) = h(x), & x > 0, \\ u(t, 0) = 0, \quad t > 0. \end{cases} \tag{5.6}$$

Let

$$\tilde{u}(t, x) = \begin{cases} u(t, x), & x > 0, \\ -u(t, -x), & x < 0, \end{cases}$$

be the odd extension of $u(t, x)$. Then $\tilde{u}(t, x)$ solves the wave equation on $\mathbb{R}^1$ with initial conditions

$$\tilde{u}(0, x) = \tilde{g}(x), \quad \tilde{u}_t(0, x) = \tilde{h}(x),$$

where $\tilde{g}$ and $\tilde{h}$ are odd extensions of g and h. By Theorem 5.1,

$$u(t, x) = \begin{cases} \frac{1}{2}\big(g(x-t) + g(x+t)\big) + \frac{1}{2}\int_{x-t}^{x+t} h(y)\, dy, & x \geq t > 0, \\ \frac{1}{2}\big(g(x+t) - g(t-x)\big) + \frac{1}{2}\int_{t-x}^{x+t} h(y)\, dy, & 0 < x < t. \end{cases}$$

Similarly, we can use even extensions to treat the Neumann boundary condition

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

5.2 Method of spherical mean

For $d \geq 2$, we introduce the *method of spherical mean* to solve the wave equation. As we can see from (5.2), the spherical symmetry plays a role in the Fourier picture.

For $x \in \mathbb{R}^d$ and $t, r > 0$, we define

$$U(x; t, r) = \fint_{\partial B_r(x)} u(t, y)\, dS(y),$$

$$G(x; r) \fint_{\partial B_r(x)} g(y)\, dS(y),$$

$$H(x; r) \fint_{\partial B_r(x)} h(y)\, dS(y).$$

Lemma 5.2 (Euler–Poisson–Darboux Equation) *Fix $x \in \mathbb{R}^d$. The functions U, G, H solve the Euler–Poisson–Darboux equation*

$$\begin{cases} U_{tt} - U_{rr} - \dfrac{d-1}{r} U_r = 0, \\ U(0, \cdot) = G,\ U_t(0, \cdot) = H. \end{cases} \tag{5.7}$$

We remark that if ϕ is spherical symmetric, that is, $\phi(x) = \phi(|x|)$, then

$$\Delta\phi(x) = (\partial_{rr} + \frac{d-1}{r}\partial_r)\phi(r), \quad |x| = r.$$

Proof: We suppress the x-dependence in U. We can write

$$U(t,r) = \frac{1}{|\partial B_1|}\int_{\partial B_1} u(t, x+ry)\, dS(y).$$

Hence,

$$\begin{aligned}\partial_r U(t,r) &= \frac{1}{|\partial B_1|}\int_{\partial B_1} \nabla(t, x+ry)\cdot y\, dS(y)\\ &= \frac{1}{|\partial B_1|}\int_{B_1} \Delta u(t, x+ry)\, dy\\ &= \frac{r}{d}\fint_{B_r(x)} \Delta u(t,y)\, dy,\end{aligned}$$

where we use the divergence theorem in the second line with the fact that y is the normal vector on ∂B_1. In particular, since $u \in \mathcal{C}^2$, we see that

$$\lim_{r\to 0+} \partial_r U(t,r) = 0. \tag{5.8}$$

By the co-area formula,

$$\partial_r \int_{B_r} f(x)\, dx = \int_{\partial B_r} f(x)\, dS(x).$$

Hence,

$$\begin{aligned}\partial_{rr} U(t,r) &= \partial_r\Big[\frac{r}{d|B_r|}\int_{B_r(x)} \Delta u(y)\, dy\Big]\\ &= \Big(\frac{1}{d} - 1\Big)\fint_{B_r(x)} \Delta u(y)\, dy + \fint_{\partial B_r(x)} \Delta u(y)\, dS(y),\end{aligned}$$

where we use $|B_r| = \alpha_d r^d$.

Therefore,

$$U_{rr} + \frac{d-1}{r}U_r = \fint_{\partial B_r(x)} \Delta u(y)\, dS(y) = \fint_{\partial B_r(x)} \partial_{tt} u(y)\, dS(y) = U_{tt}.$$

The initial conditions are straightforward. $\square$

When $d = 3$, the Euler–Poisson–Darboux equation becomes

$$U_{tt} - (U_{rr} + \frac{2}{r}U_r) = 0.$$

Let $\tilde{U} = rU$. Then

$$\begin{aligned}\partial_r \tilde{U} &= U + r\partial_r U,\\ \partial_{rr}\tilde{U} &= 2\partial_r U + r\partial_{rr} U,\\ \partial_{tt}\tilde{U} &= r\partial_{tt} U.\end{aligned}$$

Hence, $\tilde{U}_{tt} = \tilde{U}_{rr}$. Combining with (5.8), $\tilde{U}$ soles the wave equation on $\mathbb{R}_+$ with Dirichlet boundary condition. For $r < t$, we have

$$\tilde{U}(x; r, t) = \frac{1}{2}\Big[\tilde{G}(r+t) - \tilde{G}(t-r)\Big] + \frac{1}{2}\int_{t-r}^{t+r} \tilde{H}(y)\, dy.$$

To recover u, we have

$$u(t, x) = \lim_{r \to 0+} \frac{\tilde{U}(x; r, t)}{r} = \tilde{G}'(t) + \tilde{H}(t),$$

We have

$$\begin{aligned}
\tilde{G}(t) &= \frac{\partial}{\partial t}\Big(t \fint_{\partial B_t(x)} g(y)\, dS(y)\Big) \\
&= \fint_{\partial B_t(x)} g(y)\, dS(y) + t\frac{\partial}{\partial t}\fint_{\partial B_t(x)} g(y)\, dS(y) \\
&= \fint_{\partial B_t(x)} g(y)\, dS(y) + \frac{\partial}{\partial t}\fint_{\partial B_1} g(x+ty)\, dS(y) \\
&= \fint_{\partial B_t(x)} g(y)\, dS(y) + \fint_{\partial B_1} \nabla g(x+ty) \cdot y\, dS(y) \\
&= \fint_{\partial B_t(x)} g(y)\, dS(y) + \fint_{\partial B_t(x)} \nabla g(y) \cdot \frac{y-x}{t}\, dS(y).
\end{aligned}$$

Plugging in, we obtain the *Kirchhoff's formula* for wave equation solution in $\mathbb{R}^3$:

$$u(t, x) = \fint_{\partial B_t(x)} \Big[th(y) + g(y) + \nabla g(y) \cdot (y - x)\Big]\, dS(y). \tag{5.9}$$

When $d = 2$, we cannot use a change of variable to reduce (5.7) to a wave equation. However, we can view a function in $\mathbb{R}^2$ as a projection of a function in $\mathbb{R}^3$. Precisely, let

$$\bar{u}(t, x_1, x_3, x_3) = u(t, x_1, x_2), \quad \bar{g}(x_1, x_2, x_3) = g(x_1, x_2), \quad \bar{h}(x_1, x_2, x_3) = h(x_1, x_2).$$

Then $\bar{u}$ solves the wave equation in $\mathbb{R}^3$, and hence by (5.9),

$$u(t, x) = \bar{u}(t, \bar{x}) = \frac{\partial}{\partial t}\Big[t\fint_{\partial \bar{B}_t(\bar{x})} \bar{g}\, d\bar{S}\Big] + t\fint_{\partial \bar{B}_t(\bar{x})} \bar{h}\, d\bar{S}.$$

We have

$$\fint_{\partial \bar{B}_t(\bar{x})} \bar{g}\, d\bar{S} = \frac{1}{4\pi t^2}\fint_{B_t(x)} 2g(y)\sqrt{1 + |\nabla\gamma|^2}\, dy,$$

where $\gamma(y) = \sqrt{t^2 - |y - x|^2}$ is the upper sphere centered at $(x, 0)$. Direct computation gives

$$\nabla\gamma = \frac{y - x}{\sqrt{t^2 - |y - x|^2}}, \quad 1 + |\nabla\gamma|^2 = \frac{t^2}{t^2 - |x - y|^2}.$$

Hence,

$$\begin{aligned}
t\fint_{\partial \bar{B}_t(\bar{x})} \bar{g}\, d\bar{S} &= \frac{1}{2\pi}\int_{B_t(x)} \frac{g(y)}{\sqrt{t^2 - |y - x|^2}}\, dy \\
&= \frac{t^2}{2}\fint_{B_t(x)} \frac{g(y)}{\sqrt{t^2 - |y - x|^2}}\, dy \\
&= \frac{t}{2}\fint_{B_1} \frac{g(x + ty)}{\sqrt{1 - |y|^2}}\, dy.
\end{aligned}$$

Its t-derivative is then given by

$$\begin{aligned}&\frac{1}{2}\fint_{B_1} \frac{g(x+ty)}{\sqrt{\sqrt{1-|y|^2}}}\,dy + \frac{t}{2}\fint_{B_1} \frac{\nabla g(x+ty)\cdot y}{\sqrt{1-|y|^2}}\,dy \\ =& \frac{t}{2}\fint_{B_t(x)} \frac{g(y)+\nabla g(y)\cdot(y-x)}{\sqrt{t^2-|x-y|^2}}\,dy.\end{aligned}$$

For the term involving $\bar{h}$, we have

$$t\fint_{\partial B_t(x)} \bar{h}\,d\bar{S} = \frac{t^2}{2}\fint_{B_t(x)} \frac{h(y)}{\sqrt{t^2-|x-y|^2}}\,dy.$$

Combining all these, we obtain the *Poisson's formula* that solves the wave equation in $\mathbb{R}^2$:

$$u(t,x) = \frac{1}{2}\fint_{B_t(x)} \frac{tg(y)+t\nabla g(y)\cdot(y-x)+t^2h(y)}{\sqrt{t^2-|x-y|^2}}. \tag{5.10}$$

In both (5.9) and (5.10), we need $g \in \mathcal{C}^3$ and $h \in \mathcal{C}^2$ to guarantee that $u \in \mathcal{C}^{2,2}$.

There is a key difference between (5.9) and (5.10). In (5.9), the solution $u(t,x)$ depends on the initial data over the boundary $\partial B_t(x)$, whereas in (5.10), it depends on the entire ball $B_t(x)$. Physically, this reflects that the phenomenon that three-dimensional waves have both a wavefront and a waveback, while two-dimensional waves do not.

5.3 Non-homogeneous problem

We will use the Duhamel's principle to solve the non-homogeneous problem

$$\begin{cases} u_{tt}u = \Delta u + f, \\ u(0,\cdot) = 0, \quad u_t(0,\cdot) = 0. \end{cases} \tag{5.11}$$

Theorem 5.3 *Let $v(t,x;s)$ solves*

$$\begin{cases} v_{tt}(t,x;s) = \Delta v(t,x;s), & t > s,\ x \in \mathbb{R}^d, \\ v(s,x;s) = 0,\ v_t(s,x;s) = f(s,x). \end{cases}$$

Then

$$u(t,x) = \int_0^t v(t,x;s)\,ds$$

solves (5.11).

Note that the non-homogeneous term f appears in the initial condition for v_t, not v. This is because formally we can write (5.11) as

$$\partial_t \begin{bmatrix} u \\ u_t \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ \Delta & 0 \end{bmatrix} + \begin{bmatrix} 0 \\ f \end{bmatrix}, \quad \begin{bmatrix} u \\ u_t \end{bmatrix}(t=0) = \begin{bmatrix} g \\ h \end{bmatrix}.$$

So f is associated with h.

Proof: We have

$$\begin{aligned}
\partial_t u &= v(t,x;t) + \int_0^t \partial_t v(t,x;s)\,ds \\
&= \int_0^t \partial_t v(t,x;s)\,ds, \\
\partial_{tt} u &= \partial_t v(t,x;t) + \int_0^t \partial_{tt} v(t,x;s)\,ds \\
&= f(t,x) + \int_0^t \Delta v(t,x;s)\,ds \\
&= f(t,x) + \Delta u.
\end{aligned}$$

□

We look at some examples. In dimension $d=1$, we have

$$v(t,x;s) = \frac{1}{2}\int_{x-t+s}^{x+t-s} f(s,y)\,dy,$$

so

$$u(t,x) = \frac{1}{2}\int_0^t \int_{x-t+s}^{x+t-s} f(s,y)\,dyds = \frac{1}{2}\int_0^t \int_{x-s}^{x+s} f(t-s,y)\,dyds.$$

The domain of dependence is the cone

$$\mathcal{D} = \{(s,y) \in \mathbb{R}_+ \times \mathbb{R} : 0 \le s \le t,\ |x-y| \le t-s\}.$$

In dimension $d=3$, we have

$$v(t,x;s) = (t-s)\fint_{\partial B_{t-s}(x)} f(s,y)\,dS(y),$$

and hence

$$\begin{aligned}
u(t,x) &= \int_0^t (t-s)\fint_{\partial B_{t-s}(x)} f(s,y)\,dS(y)ds \\
&= \int_0^t \frac{1}{4\pi}\int_{\partial B_{t-s}(x)} \frac{f(s,y)}{t-s}\,dS(y)ds \\
&= \frac{1}{4\pi}\int_{B_t(x)} \frac{f(t-|y-x|,y)}{|y-x|}\,dy.
\end{aligned}$$

Again, the domain of influence is a cone

$$\mathcal{D} = \{(s,y) \in \mathbb{R}_+ \times \mathbb{R}^3 : 0 \le s \le t,\ |x-y| \le t-s\}.$$

We will see a more clear picture in the next section.

5.4 Energy method

Consider the wave equation

$$\begin{cases}
\Box u = f, & (t,x) \in \Omega_T, \\
u = g, & (t,x) \in \partial_p\Omega_T, \\
u_t(0,x) = h(x), & x \in \Omega.
\end{cases} \tag{5.12}$$

Theorem 5.4 *The classical solution $u \in \mathcal{C}^2(\Omega_T)$ to (5.12) is unique.*

Proof:

Let u_1 and u_2 be two solutions. Then $v = u_1 - u_2$ solves (5.12) with $f = g = h = 0$. We consider the *energy*

$$e(t) = \frac{1}{2}\int_\Omega |v_t|^2 + |\nabla v|^2\, dx.$$

Then

$$\begin{aligned} e'(t) &= \int_\Omega v_{tt}v_t + \nabla v_t \cdot \nabla v\, dx \\ &= \int_\Omega (\Delta v)v_t - v_t\Delta v\, dx + \int_{\partial\Omega} \frac{\partial v}{\partial n} \cdot v_t\, dS \\ &= 0. \end{aligned}$$

Since $e(0) = 0$ and $e(t) \geq 0$, this implies $e(t) \equiv 0$. It follows that $v \equiv 0$. $\square$

Using a refined form energy, we have the following result about the domain of influence. Fix (t_0, x_0). Consider the cone

$$\mathcal{C} = \{(t,x) : 0 \leq t \leq t_0,\ |x - x_0| \leq t_0 - t\}.$$

Proposition 5.5 (Domain of influence) *If $u \in \mathcal{C}^2$, $\Box u = 0$ and $u = u_t = 0$ in $\{t = 0\} \times B_{t_0}(x_0)$, then $u \equiv 0$ in $\mathcal{C}$.*

Proof:

Let

$$e(t) = \frac{1}{2}\int_{B_{t_0-t}(x_0)} |u_t(t,x)|^2 + |\nabla u(t,x)|^2\, dx.$$

Then by co-area formula,

$$\begin{aligned} e'(t) &= \int_{B_{t_0-t}(x_0)} u_{tt} \cdot u_t + \nabla u \cdot \nabla u_t\, dx - \frac{1}{2}\int_{\partial B_{t_0-t}} |u_t|^2 + |\nabla u|^2\, dx \\ &= \int_{B_{t_0-t}(x_0)} u_t \cdot \Box u + \int_{\partial B_{t_0-t}(x_0)} \frac{\partial u}{\partial n}u_t - \frac{1}{2}\int_{\partial B_{t_0-t}} |u_t|^2 + |\nabla u|^2\, dx \\ &= \int_{\partial B_{t_0-t}(x_0)} \left[\frac{\partial u}{\partial n}u_t - \frac{1}{2}|u_t|^2 + |\nabla u|^2\right] \\ &\leq 0. \end{aligned}$$

Here, the last line follows from

$$\left|\frac{\partial u}{\partial n}u_t\right| \leq \frac{1}{2}\left|\frac{\partial u}{\partial n}\right|^2 + \frac{1}{2}|u_t|^2 \leq \frac{1}{2}|\nabla u|^2 + \frac{1}{2}|u_t|^2.$$

Since $e(0) = 0$, $e(t) \geq 0$ and $e'(t) \leq 0$, we have $e(t) \equiv 0$. It follows that $u \equiv 0$ in $\mathcal{C}$. $\square$

As a corollary, we have the energy estimate.

Proposition 5.6 *If $\Box u = f$, then*

$$e(t) \leq e^{t_0}\Big[\int_0^t \int_{B_{t_0-t}(x_0)} f^2(s,x)\, dsdx + e(0)\Big].$$

Proof: We have

$$e'(t) \le \int_{B_{t_0-t}(x_0)} u_t(t,x) f(t,x)\,dx \le \int_{B_{t_0-t}(x_0)} f^2(t,x) + e(t).$$

The conclusion follows from Gronwall's inequality.

$\square$

5.5 Separation of variables

As an example, we will solve the wave equation with Dirichlet boundary condition on $[0,\pi]$:

$$\begin{cases} u_{tt} - u_{xx} = 0, & t > 0,\ x \in (0,\pi), \\ u(t,0) = u(t,\pi) = 0, & t > 0, \\ u(0,x) = \phi(x), & x \in (0,\pi), \\ u_t(0,x) = \psi(x), & x \in (0,\pi). \end{cases} \tag{5.13}$$

Then

$$u(t,x) = \sum_{n=1}^{\infty} T_n(t) \sin(nx)$$

where T_n solves

$$T_n'' + n^2 T_n = 0, \quad T_n = \phi_n, \quad T_n'(0) = \psi_n,$$

with

$$\phi_n = \frac{2}{\pi}\int_0^{\pi} \phi(x)\sin(ny)\,dy, \quad \psi_n = \frac{2}{\pi}\int_0^{\pi} \psi(x)\sin(ny)\,dy.$$

Hence,

$$T_n(t) = \phi_n \cos(nt) + \psi_n \sin(nt).$$

We can also use odd and 2π-periodic extension to turn the equation into wave equation on $\mathbb{R}^1$. Let Φ and Ψ be the extensions of ϕ and ψ. Then by Theorem 5.1,

$$u(t,x) = \frac{1}{2}\big[\Phi(t+x) + \Phi(t-x)\big] + \frac{1}{2}\int_{x-t}^{x+t} \Psi(y)\,dy.$$

One can check

$$\sum_{n=1}^{\infty} \phi_n \cos(nt)\sin(nx) = \sum_{n=1}^{\infty} \phi_n \Big[\sin\big(n(t+x)\big) - \sin\big(n(x-t)\big)\Big] = \Phi(t+x) - \Phi(t-x),$$

and a similar identity for the ψ term.

Resonance. Let us include a non-homogeneous term in (5.13), that is,

$$u_{tt} - u_{xx} = f(t,x), \quad t > 0,\ x \in (0,\pi).$$

For simplicity we also set $\phi = \psi = 0$. Then T_n will solve

$$T_n'' + n^2 T_n(t) = f_n(t), \quad f_n(t) = \frac{2}{\pi}\int_0^{\pi} f(t,y)\sin(ny)\,dy.$$

Suppose $f(t,x) = e^{i\omega t}A(x)$, that is, the external force is period in time, with frequency ω. Then

$$T_n''(t) + n^2 T_n(t) = A_n \sin(\omega t).$$

By Duhamel's principle,

$$\begin{aligned}
T_n(t) &= A_n \int_0^t \sin(\omega s)\sin\big(n(t-s)\big)\,ds \\
&= \frac{A_n}{2}\int_0^t \Big[\cos\big(nt-(n+\omega)s\big) - \cos\big(nt-(n-\omega)s\big)\Big]\,ds \\
&= \begin{cases} \frac{C_1}{n+\omega} + \frac{C_2}{n-\omega}, & n \neq \omega, \\ \frac{tA_n}{2} + O(1), & n = \omega. \end{cases}
\end{aligned}$$

In particular, when $|\omega| = n$, that is, the frequency of the external force coincides with one of the frequency of the system, $\omega \in \{1,2,3,\dots\}$, the wave equation solution will tend to ∞ as $t \to \infty$. This is the phenomenon of resonance. In particular, there is no maximum principle for the wave equation, since the solution is not controlled by the initial and boundary data.

6 First-order evolution PDEs

We loosely follow [Eva, 7.1].

6.1 Set up

Let Ω be a bounded domain. Consider a general elliptic operator in the divergence form

$$\mathcal{L}u = -\sum_{i,j=1}^d \big(a_{ij}(t,x)u_{x_i}\big)_{x_j} + \sum_{i=1}^d b_i(t,x)u_{x_i} + c(t,x),$$

where $a_{ij} = a_{ji}$.

We want to develop solution theory for the linear evolution PDE

$$\begin{cases} u_t + \mathcal{L}u = f(t,x), & (t,x) \in \Omega_T, \\ u(t,x) = 0, & t > 0,\ x \in \partial\Omega, \\ u(0,x) = g(x), & x \in \Omega. \end{cases} \tag{6.1}$$

For example, when $a_{ij} = \delta_{ij}$ and $b = c = 0$, $\mathcal{L}u = -\Delta u$ and (6.1) becomes the heat equation.

Definition 6.1 (Uniform ellipticity) *$\mathcal{L}$ is uniformly elliptic if there exists $\theta > 0$ such that the smallest eigenvalue of the matrix $A(t,x) = \big(a_{ij}(t,x)\big)$ is at least θ for all $(t,x) \in \Omega_T$, or equivalently,*

$$\sum_{i,j=1}^d a_{ij}(t,x)\xi_i\xi_j \geq \theta|\xi|^2, \quad \forall (t,x) \in \Omega_T,\ \xi \in \mathbb{R}^d.$$

We impose the following assumptions.

1. $\mathcal{L}$ is uniformly elliptic.
2. $a_{ij}, b_i, c \in L^\infty(\Omega_T)$.
3. $f \in L^2(\Omega_T)$ and $g \in L^2(\Omega)$.

We will view $u(t,x)$ as a map from $t \in [0,T]$ to $u(t,\cdot) \in H_0^1(\Omega)$, where

$$H_0^1(\Omega) = \text{closure of } \mathcal{C}_c^\infty(\Omega)$$

under the H^1-norm

$$\|h\|_{H^1} := \int_\Omega |h(x)|^2 + |\nabla h(x)|^2 \, dx.$$

Roughly speaking, it consists of functions h such that $h, \nabla h \in L^2(\Omega)$ and vanishes on $\partial\Omega$; this means that the boundary condition is already encoded in the functional space.

6.2 Weak solutions

We need a notion of solutions that has minimum requirement on the regularity. Assume temporarily that u is smooth.

We define

$$[\bar{u}(t)](x) := u(t,x), \quad [\bar{f}(t)](x) := f(t,x).$$

We seek solution $\bar{u} : [0,T] \to H_0^1(\Omega)$. Note that by Fubini,

$$f \in L^2(\Omega_T) \iff \|f\|_{L^2(\Omega_T)} = \int_0^T \|\bar{f}(t)\|_{L^2(\Omega)} \, dt < \infty$$

Hence $\bar{f}(t) \in L^2(\Omega)$ for almost every t, and we can view $\bar{f}$ as a map from $[0,T]$ to $L^2(\Omega)$.

Let $v \in H_0^1(\Omega)$. Using integration by parts,

$$\int_\Omega v(x)(u_t + \mathcal{L}u) \, dx = \int_\Omega v(x) f(t,x) \, dx$$

becomes

$$(\bar{u}', v)_{L^2} + B[\bar{u}(t), v; t] = (\bar{f}(t), v)_{L^2}, \tag{6.2}$$

where

$$B[u,v;t] := \int_\Omega \Big[\sum_{i,j} a_{ij}(t,x) u_{x_i} u_{x_j} + \sum_i b_i(t,x) u_{x_i} v + c(t,x) uv \Big] \, dx.$$

How to define (6.2) in the weakest sense? For simplicity, let us assume $b = c \equiv 0$.

For the term $(\bar{f}, v)$ on the RHS, since $\bar{f}(t) \in L^2(\Omega)$ and $v \in H_0^1(\Omega) \subset L^2(\Omega)$, their inner product $(\bar{f}(t), v)$ is defined.

For the term $B[u,v;t]$, we have

$$|B[u,v;t]| \le \sup a_{ij} \|u\|_{H_0^1(\Omega)} \|v\|_{H_0^1(\Omega)}.$$

So we need $\bar{u} \in L^2([0,T]; H_0^1(\Omega))$.

Finally, we need to make sense of

$$\int_\Omega \phi(x) v(x) \, dx, \quad \phi = \bar{u}'(t)$$

for all $v \in H_0^1(\Omega)$. Then ϕ belongs to the dual space of $H_0^1(\Omega)$, denoted by $[H_0^1(\Omega)]^* = H^{-1}(\Omega)$. The space $H^{-1}(\Omega)$ is equipped with the norm

$$\|\phi\|_{H^{-1}(\Omega)} = \sup_{v \in H_0^1, \ \|v\|_{H_0^1} = 1} \langle \phi, v \rangle.$$

Later, we will use $\langle\cdot,\cdot\rangle$ to denote the pairing between a Banach space and its dual space, and $(\cdot,\cdot)$ for the L^2-inner product. From (6.2) we do know that $\bar{u}'(t) \in H^{-1}(\Omega)$, since for $v \in H_0^1(\Omega)$,

$$\langle \bar{u}'(t), v\rangle \le \|v\|_{H_0^1} \cdot \Big[\sup|a_{ij}|\|u\|_{H_0^1} + \|f\|_{L^2}\Big].$$

Definition 6.2 *A weak solution to (6.1) is a function $\bar{u} \in L^2\big(0,T; H_0^1(\Omega)\big)$ with $\bar{u}' \in L^2\big(0,T; H^{-1}(\Omega)\big)$, so that the following holds:*

1. *(6.2) holds for almost every $t \in [0,T]$ and all $v \in H_0^1(\Omega)$;*
2. *$\bar{u}(0) = g$.*

We will make a comment about the meaning of $\bar{u}'(t)$. Since $\bar{u} \in L^2\big(0,T; H_0^1(\Omega)\big)$, by definition for almost every $t \in [0,T]$, $\bar{u}(t) \in H_0^1(\Omega) \subset L^2(\Omega) \subset H^{-1}(\Omega)$. Then $\bar{u}'(t)$ is defined as the limit of

$$\lim_{h\to 0} \frac{\bar{u}(t+h) - \bar{u}(t)}{h}$$

in $H^{-1}(\Omega)$. We also have the following lemma that can make sense of the initial condition in Definition 6.2.

Lemma 6.1 *If $\bar{u} \in L^2\big(0,T; H_0^1(\Omega)\big)$ and $\bar{u}' \in L^2\big(0,T; H^{-1}(\Omega)\big)$, then $\bar{u} \in L^2\big(0,T; L^2(\Omega)\big)$ and*

$$\max_{0\le t\le T} \|\bar{u}(t)\|_{L^2} \le C\Big[\|\bar{u}\|_{L^2(0,T;H_0^1)} + \|\bar{u}'\|_{L^2(0,T;H^{-1})}\Big]. \tag{6.3}$$

for some constant $C = C(T)$.

Proof:

First, (6.3) holds for $\bar{u} \in \mathcal{C}_c^\infty\big(0,T; H_0^1\big)$ with $\bar{u}'(t) \equiv 0$ for $t < -1$. Indeed,

$$\begin{aligned}
\|\bar{u}(t)\|_{L^2}^2 &= 2\int_{-1}^t \langle \bar{u}(s), \bar{u}'(s)\rangle\, ds \\
&\le 2\int_{-1}^t \|\bar{u}(s)\|_{H_0^1} \cdot \|\bar{u}'(s)\|_{H^{-1}}\, ds \\
&\le \int_{-1}^t (\|\bar{u}(s)\|_{H_0^1}^2 + \|\bar{u}'(s)\|_{H^{-1}}^2)\, ds \\
&\le \|\bar{u}\|_{L^2(0,T;H_0^1)}^2 + \|\bar{u}'\|_{L^2(0,T;H^{-1})}^2.
\end{aligned}$$

In the general case, consider $\bar{u}_\varepsilon = \eta_\varepsilon * \bar{u}$. Then by the property of the mollifiers, we have

$$\bar{u}_\varepsilon \to \bar{u}, \text{ in } L^2(0,T;H_0^1) \text{ and for almost every } t \in [0,T]$$

and

$$\bar{u}_\varepsilon' \to \bar{u}', \text{ in } L^2(0,T;H^{-1}) \text{ and for almost every } t \in [0,T]$$

By (6.3), $\{\bar{u}_\varepsilon\}_{\varepsilon>0}$ forms a Cauchy sequence in $\mathcal{C}(0,T;L^2)$. Since $\bar{u}_\varepsilon(t) \to \bar{u}(t)$ for almost every $t \in [0,T]$, any limit point of the Cauchy sequence coincides with $\bar{u}$ except for a zero measure set of t. In other words, we can identify $\bar{u}$ with an element in $\mathcal{C}(0,T;L^2)$. $\square$

6.3 Uniqueness of weak solution

Theorem 6.2 *The unique weak solution to (6.1) with $g = f = 0$ is $\bar{u} \equiv 0$.*

The bilinear form B satisfies the following estimates.

Lemma 6.3 *1. There exists a constant $M > 0$ such that*

$$|B[u, v; t]| \leq M\|u\|_{H_0^1}\|v\|_{H_0^1}, \quad \forall t, \ \forall u, v \in H_0^1(\Omega).$$

2. There exist constant $\beta, \gamma > 0$ such that

$$B[u, u; t] \geq \beta\|u\|_{H_0^1}^2 - \gamma\|u\|_{L^2}^2, \quad \forall t, \ \forall u \in H_0^1(\Omega).$$

Proof:
Since $a, b, c \in L^\infty$ and $\|h\|_{L^2} \leq \|h\|_{H_0^1}$, by Cauchy–Schartz we have

$$\begin{aligned}
|B[u, v; t]| &\leq d^2 \sup|a_{ij}| \cdot \|u\|_{H_0^1}\|v\|_{H_0^1} + d \sup|b_i| \cdot \|u\|_{H_0^1}\|v\|_{L^2} + \sup|c| \cdot \|u\|_{L^2}\|v\|_{L^2} \\
&\leq M\|u\|_{H_0^1}\|v\|_{H_0^1}.
\end{aligned}$$

This proves the first part.

For the second part, by Poincaré's inequality (Theorem 4.28) there exists $\delta > 0$ such that

$$\|\nabla u\|_{L^2} \geq \delta\|u\|_{L^2}, \quad \forall u \in H_0^1(\Omega).$$

Using uniform ellipticity of (a_{ij}), we have

$$\begin{aligned}
B[u, u; t] &\geq \theta\|\nabla u\|_{L^2}^2 - d \sup|b| \cdot \|\nabla u\|_{L^2}\|u\|_{L^2} - \sup|c|\|u\|_{L^2}^2 \\
&\geq (\theta - \varepsilon)\|\nabla u\|_{L^2}^2 - (\sup|c| + d \sup|b|/4\varepsilon)\|u\|_{L^2} \\
&\geq \beta\|u\|_{H_0^1}^2 - \gamma\|u\|_{L^2}^2,
\end{aligned}$$

provided $\varepsilon > 0$ is choosen sufficiently small. □

Proposition 6.4 *If $\bar{u}$ is a weak solution to (6.1), then*

$$\int_0^T \langle \bar{u}'(t), \bar{v}(t) \rangle + B[\bar{u}(t), \bar{v}(t); t] \, dt = \int_0^T \big(\bar{u}(t), \bar{f}(t)\big)_{L^2} \, dt \tag{6.4}$$

for all $\bar{v} \in L^2\big(0, T; H_0^1(\Omega)\big)$.

Proof: Since $H_0^1(\Omega)$ is separable, it has an orthonormal basis $\{w_k\}_{k=1}^\infty$. By (6.2), (6.4) holds for

$$\bar{v}(t) = h_k(t) w_k$$

for any $h_k \in \mathcal{C}_0^\infty(\mathbb{R})$. By linearity, (6.4) holds for

$$\bar{v}(t) = \sum_{k=1}^N h_k(t) w_k, \quad h_k \in \mathcal{C}_0^\infty(\mathbb{R}). \tag{6.5}$$

Denote the difference of the LHS and the RHS of (6.4) by $F[\bar{v}]$. Then the conclusion follows from the fact that $\bar{v}$ of the form (6.5) is dense in $L^2(0, T; H_0^1)$ and

$$|F[\bar{v}_1] - F[\bar{v}_2]| \leq C(\bar{u}, \bar{f})\|\bar{v}_1 - \bar{v}_2\|_{H_0^1(\Omega)}.$$

$\square$

We are ready to prove Theorem 6.2.

Proof of Theorem 6.2:

In (6.4), letting

$$\bar{v}(s) = \mathbb{1}_{[0,t]}(s)\bar{u}(s),$$

we have

$$\int_0^t \langle \bar{u}'(s), \bar{u}(s)\rangle\, ds + \int_0^t B[\bar{u}(s), \bar{u}(s); s]\, ds = 0, \quad \forall t \in [0, T].$$

Note that the first term equals $\frac{1}{2}\|\bar{u}(t)\|_{L^2}^2$. By Lemma 6.3, we have

$$0 \geq \|\bar{u}(t)\|_{L^2}^2 - \gamma \int_0^t \|\bar{u}(s)\|_{L^2}^2\, ds.$$

By Gronwall's inequality applied to $t \mapsto \|\bar{u}(t)\|_{L^2}^2$, we obtain $\bar{u}(t) \equiv 0$. $\square$

6.4 Existence of linear equation: Galerkin approximation

Let $\{w_k\}_{k=1}^\infty$ be a basis in $H_0^1(\Omega)$ and $\Omega_m = \overline{\text{span}\{w_1, \dots, w_m\}}$. In practice, the basis can come from a mesh approxmation, finite-element method, or Fourier series approximation, etc. We want to find the function $\bar{u}_m \in \Omega_m$ that is closest to the solution to (6.1), and then take the limit $m \to \infty$. In other words, $\bar{u}_m$ satisfies the projection of (6.1) onto Ω_m:

$$\langle \bar{u}_m', w_j\rangle + B[\bar{u}_m, w_j; t] = (\bar{f}, w_j), \quad \forall 1 \leq j \leq m. \tag{6.6}$$

If we write

$$\bar{u}_m(t) = \sum_{k=1}^m d_m^k(t) w_k,$$

then

$$\sum_{k=1}^m (d_m^k)'(t)(w_k, w_j) + \sum_{k=1}^m d_m^k(t) B[w_k, w_j; t] = (\bar{f}, w_j).$$

Equivalently, the vector function $\bar{d}(t) = \big(d_m^1(t), \dots, d_m^m(t)\big)^T$ satisfies the ODE

$$A\bar{d}' + B\bar{d} = F, \quad \bar{d}(0) = g_m = \pi_{\Omega_m} g,$$

where

$$A^{ij} = (w_i, w_j), \quad B_{ij} = B[w_i, w_j; t], \quad F_j(t) = \big(\bar{f}(t), w_j\big).$$

The matrices A, B and F satisfy $A^{-1}, B \in L^\infty$, $F \in L^2$. Multiplying A^{-1} to the ODE, we obtain

$$\bar{d}' + A^{-1}B\bar{d} = A^{-1}F.$$

This is a linear ODE which always has a solution.

Next, we want to take the limit $\lim_{m\to\infty} \bar{u}_m$. Naturally we expect to extract the limit in the weak sense. We need an energy estimate.

Theorem 6.5 (Energy estimate) *There exists a constant C, independent of m, such that*

$$\max_{0\leq t\leq T} \|\bar{u}_m(t)\|_{L^2} + \|\bar{u}_m\|_{L^2(0,T;H_0^1)} + \|\bar{u}_m'\|_{L^2(0,T;H^{-1})} \leq C\Big[\|f\|_{L^2(0,T;L^2)} + \|g\|_{L^2}\Big]. \tag{6.7}$$

Proof: We have

$$\big(\bar{u}_m'(t), \bar{u}_m(t)\big) + B[\bar{u}_m(t), \bar{u}_m(t); t] = \big(\bar{f}(t), \bar{u}_m(t)\big).$$

By Lemma 6.3, we have

$$\frac{d}{dt}\frac{1}{2}\|\bar{u}_m(t)\|_{L^2}^2 + \beta\|\bar{u}_m(t)\|_{H_0^1}^2 \le \frac{1}{2}\|\bar{f}(t)\|_{L^2}^2 + (\frac{1}{2}+\gamma)\|\bar{u}_m(t)\|_{L^2}^2. \tag{6.8}$$

For the first term in the LHS of (6.7), let $\eta(t) = \|\bar{u}_m(t)\|_{L^2}^2$ and $\xi(t) = \|\bar{f}(t)\|_{L^2}^2$. From (6.8) we have

$$\eta'(t) \le C_1\eta(t) + C_2\xi(t),$$

and hence by Gronwall's inequality,

$$\eta(t) \le e^{C_1 t}\Big(\eta(0) + C_2\int_0^t \xi(s)\,ds\Big) \le e^{C_1 t}\Big(\|g\|_{L^2}^2 + C_2\|\bar{f}\|_{L^2(0,T;L^2)}\Big), \quad t \in [0,T].$$

This gives the upper bound of the first term.

For the second term, integrating (6.8) over $[0,T]$, we obtain

$$\beta\|\bar{u}_m\|_{L^2(0,T;H_0^1)}^2 \le \frac{1}{2}\|f\|_{L^2(0,T;L^2)}^2 + C(T+1)\max_{0\le t\le T}\|\bar{u}_m(t)\|_{L^2}^2.$$

This gives the desired upper bound of the second term.

For the third term, we recall that

$$\|\bar{u}_m'(t)\|_{H^{-1}} = \sup\big\{\langle \bar{u}_m'(t), v\rangle : v \in H_0^1,\ \|v\|_{H_0^1} \le 1\big\}.$$

Let $v \in H_0^1$ with $\|v\|_{H_0^1} \le 1$. Decompose v as

$$v = v_1 + v_2, \quad v_1 \in \Omega_m,\ v_2 \in \Omega_m^\perp.$$

Then $\|v_1\|_{H_0^1} \le 1$ and

$$\begin{aligned}\langle \bar{u}_m'(t), v\rangle = \langle \bar{u}_m', v_1\rangle &= (\bar{f}(t), v_1) - B[\bar{u}_m(t), v; t]\\ &\le C\Big(\|\bar{f}(t)\|_{L^2}^2 + \|\bar{u}_m(t)\|_{H_0^1}^2\Big).\end{aligned}$$

Hence,

$$\|\bar{u}_m'(t)\|_{H^{-1}} \le C\Big(\|\bar{f}(t)\|_{L^2} + \|\bar{u}_m(t)\|_{H_0^1}^2\Big).$$

Taking square of both sides, integrating over $[0,T]$ and using the bound on the second term yields the desired result. □

Theorem 6.6 *There exists a weak solution to* (6.1).

Proof: From the energy estimate Theorem 6.5, $\{u_m'\}$ is bounded in $L^2\big(0,T;H_0^1(\Omega)\big)$ and $\{\bar{u}_m'\}$ is bounded in $L^2\Big(0,T;H^{-1}(\Omega)\Big)$. Since both spaces are separable Hilbert spaces, by weak compactness (Theorem 4.27), there exists a subsequence, still denoted by $\{\bar{u}_m\}$, and a function u such that

$$\bar{u}_m \rightharpoonup \bar{u} \text{ in } L^2(0,T;H_0^1), \quad \bar{u}_m' \rightharpoonup \bar{u}' \text{ in } L^2(0,T;H^{-1}). \tag{6.9}$$

Let

$$\bar{v}(t) = \sum_{k=1}^{N} d^k(t) w_k \tag{6.10}$$

where $d^k(t)$ are smooth functions. Then $\bar{v} \in \mathcal{C}^1(0,T;\Omega_N)$. For $m \geq N$, by the construction of u_m we have

$$\int_0^T \langle \bar{u}_m(t), \bar{v}(t)\rangle + B[\bar{u}_m(t), \bar{v}(t); t]\, dt = \int_0^T \big(\bar{f}(t), \bar{v}(t)\big). \tag{6.11}$$

By (6.9), we have

$$\int_0^T \langle \bar{u}'(t), \bar{v}(t)\rangle + B[\bar{u}(t), \bar{v}(t); t]\, dt = \int_0^T \big(\bar{f}(t), \bar{v}(t)\big)\, dt. \tag{6.12}$$

Since functions of the form (6.10) is dense in $L^2(0,T;H_0^1)$, we have

$$\langle \bar{u}', \bar{v}\rangle + B[\bar{u}, \bar{v}; t] = (\bar{f}, \bar{v}), \qquad \text{for almost every } t \in [0,T].$$

To show that $\bar{u}$ is a weak solution, it remains to check $\bar{u}(0) = g$. In (6.12), taking $\bar{v} \in \mathcal{C}^1(0,T;H_0^1)$ with $\bar{v}(T) = 0$, and integrating by parts in t yields

$$\int_0^T -\langle \bar{u}, \bar{v}'\rangle + B[\bar{u}, \bar{v}; t]\, dt = \int_0^T (\bar{f}, \bar{v})\, dt + \big(\bar{u}(0), \bar{v}(0)\big),$$

while from (6.11) we can obtain

$$\int_0^T -\langle \bar{u}_m, \bar{v}'\rangle + B[\bar{u}_m, \bar{v}; t]\, dt = \int_0^T (\bar{f}, \bar{v})\, dt + \big(\bar{u}_m(0), \bar{v}(0)\big),$$

Taking $m \to \infty$ and using weak convergence (6.9), we obtain

$$\big(\bar{u}(0), \bar{v}(0)\big) = \big(g, \bar{v}(0)\big).$$

Since this holds for all $\bar{v}(0)$, we must have $\bar{u}(0) = g$. This completes the proof. $\square$

6.5 Existence of nonlinear equation: fixed point method

In this section we consider a nonlinear system

$$\begin{cases} u_t = -\mathcal{L}u + f(u), & \\ u(0,x) = g(x) \in L^2(\Omega), & x \in \Omega, \\ u(t,x) = 0, & x \in \partial\Omega, \end{cases} \tag{6.13}$$

where

$$\mathcal{L}u = -\sum_{i,j} (a_{ij}(x) u_{x_i})_{x_j},$$

and $f : \mathbb{R} \to \mathbb{R}$ is a Lipschitz function with Lipchitz constant L.

The tool we will use is the Banach Fixed Point Theorem.

Theorem 6.7 (Banach Fixed Point Theorem) *Let X be a Banach space (complete, normed linear space) and $A : X \to X$ a possible nonlineaer operator that satisfies*

$$\|A[u] - A[\tilde{u}]\| \leq \gamma \|u - \tilde{u}\|, \quad \forall u, \tilde{u} \in X,$$

for some constant $\gamma < 1$. Then A has a unique fixed point, that is, solution to the equation $Au = u$.

Proof: Pick any $u_0 \in X$. Then $\{A^n[u_0]\}$ is a Cauchy sequence in X since

$$\|A^n[u_0] - A^{n+m}[u_0]\| \le (\gamma^n + \gamma^{n+1} + \cdots + \gamma^{n+m})\|A[u_0] - u_0\| \le C\gamma^n \|A[u_0] - u_0\|.$$

Since X is a complete metric space, there exists u such that $A^n[u_0] \to u$, and hence by continuity,

$$u = \lim_{n\to\infty} A[A^n[u_0]] = A[\lim_{n\to\infty} A^n[u]] = A[u].$$

□

As a classical application, we consider the ODE

$$\dot{x}(t) = f(t, x), \quad x(0) = a,$$

where f is bounded and Lipchitz in x. The space $X = \mathcal{C}[0,T] \cap \{x(0) = a\}$ is a Banach space under the maximum norm. Let

$$A[x](t) = \int_0^t f\big(s, x(s)\big)\, ds.$$

Then

$$\big|A[x](t) - A[\tilde{x}](t)\big| \le \int_0^T \big|f\big(s, x(s)\big) - f\big(s, \tilde{x}(s)\big)\big|\, ds \le LT \sup_{0\le s\le T} |x(s) - \tilde{x}(s)|.$$

Hence,

$$\|A[x] - A[\tilde{x}]\| \le LT\|x - \tilde{x}\|.$$

Choosing $T < 1/L$ and applying Theorem 6.7, we obtain a solution to the ODE on $[0,T]$. Iterating this over $[T, 2T], [2T, 3T], \ldots$ yields global in t existence of solutions.

We will apply Theorem 6.7 to obtain a weak solution to (6.13).

Proof:

We take $X = \mathcal{C}\big(0, T; L^2(\Omega)\big)$, and $w = A[u]$ to be the weak solution of

$$w_t = -\mathcal{L}w + f(u), \quad w(0) = g,$$

so that $w \in L^2_t H^1_0$ and $w' \in L^2_t H^{-1}$. For Theorem 6.6 to apply, first we need to check that $h(t,x) = f\big(u(t,x)\big) \in L^2_t L^2_x$ provided that $u \in \mathcal{C}_t L^2_x$. Indeed, by Lipschtiz continuity we can assume

$$|f(z)|^2 \le C(1 + z^2)$$

for some constant C; we have

$$\begin{aligned}\int_0^T \int_\Omega h^2(t,x)\, dxdt &= \int_0^T \int_\Omega f^2\big(u(t,x)\big)\, dxdt \\ &\le \int_0^T \int_\Omega C\big(1 + u^2(t,x)\big) \\ &\le C(1 + \|u\|^2_{L^2_t L^2_x}).\end{aligned}$$

Next, to apply Theorem 6.7, we need to show that for T sufficiently small, $u \mapsto w = A[u]$ is a contraction. Let $w_1 = A[u]$ and $w_2 = A[\tilde{u}]$. By definition of weak solutions, for all $v \in H^1_0(\Omega)$ we have

$$\langle \bar{w}_1' - \bar{w}_2', v\rangle + B[\bar{w}_1 - \bar{w}_2, v] = \big(f(u) - f(\tilde{u}), v\big).$$

Using a similar argument to Proposition 6.4, we have

$$\begin{aligned}\frac{d}{dt}\|\bar{w}_1-\bar{w}_2\|_{L^2}^2+\beta\|\bar{w}_1-\bar{w}_2\|_{H_0^1}^2 &\le L\|u-\tilde{u}\|_{L^2}\|\bar{w}_1-\bar{w}_2\|_{L^2}\\ &\le \frac{L}{4\varepsilon}\|u-\tilde{u}\|_{L^2}^2+\varepsilon L\|\bar{w}_1-\bar{w}_2\|_{L^2}^2\\ &\le \frac{L}{4\varepsilon}\|u-\tilde{u}\|_{L^2}^2+\varepsilon L\delta\|\bar{w}_1-\bar{w}_2\|_{H_0^1}^2,\end{aligned}$$

where δ is the constant from the Poincaré's inequality in Ω. By choosing ε sufficiently small, we can make $\varepsilon L\delta \le \theta$, and hence

$$\frac{d}{dt}\|\bar{w}_1-\bar{w}_2\|_{L^2}^2 \le \frac{L}{4\varepsilon}\|u-\tilde{u}\|_{L^2}^2.$$

Integrating over $[0,t]$ yields

$$\|\bar{w}_1(t)-\bar{w}_2(t)\|_{L^2}^2 \le \frac{LT}{4\varepsilon}\|u-\tilde{u}\|_X^2, \quad \forall t\in[0,T].$$

If $LT/4\varepsilon<1$, then $u\mapsto w=A[u]$ is a contraction. In particular, there exists a solution $u=A[u]$ up to time $T_1=L/8\varepsilon$.

Since u is a weak solution up to time T_1, we have $u\in L^2(0,T;H_0^1)$ and hence $\|\bar{u}(t)\|_{H_0^1}<\infty$ for almost every $t\le T_1$. In particular, there exists $t_0\in[T_1/2,T_1]$ such that $\|\bar{u}(t_0)\|_{H_0^1}<\infty$. Using $\bar{u}(t_0)$ as the initial condition, we obtain a solution on $[T_1/2,3T_1/2]$. Interating this argument we obtain a solution for all time.

□

7 Hamilton–Jacobi equations

7.1 Set up and method of characteristics

Hamilton–Jacobi equation takes the form

$$u_t+H(u_x,x)=0. \tag{7.1}$$

For example, from the Burgers equation

$$v_t+v\cdot\partial_x v=0,$$

we can integrate once to get

$$u_t+\frac{1}{2}(u_x)^2=0, \quad v=u_x.$$

More generally, the RHS of (7.1) can be a function $F(t,x)$, which is equivalent to have a Hamiltonian $H(u_x,x,t)=H(u_x,x)-F(t,x)$.

Next, we will use the method of characteristic to solve

$$u_t+H(u_x,x,t)=0. \tag{7.2}$$

Comparing with the Burgers equation, we should make differentiate the equation in x to make (7.2) a semi-lineaer equation. Indeed, differentiating once in x and writing $v=u_x$, we obtain

$$v_t+\partial_p H(v,x,t)\cdot v_x+\partial_x H(v,x,t)=0.$$

Let $\eta(s)$ be the characteristic and

$$p(s) = v\big(s, \eta(s)\big).$$

We have

$$\dot{\eta}(s) = \partial_p H\big(p(s), \eta(s), s\big), \quad \dot{p}(s) = -\partial_x H\big(p(s), \eta(s), s\big). \tag{7.3}$$

(7.3) is called the *Hamiltonian's PDE.*

Proposition 7.1 (Preservation of Hamiltonian) *Assume $H(p, x, s) = H(p, x)$. Then $H\big(p(s), \eta(s), s\big)$ is constant.*

Proof: Let $H(s) = H\big(p(s), s, \eta(s)\big)$. We have

$$\begin{aligned}\dot{H}(s) &= \partial_p H \cdot \dot{p}(s) + \partial_x H \cdot \dot{\eta}(s) \\ &= \partial_p H \cdot (-\partial_x H) + (\partial_x H) \cdot (\partial_p H) = 0.\end{aligned}$$

□

Physically, $H(p, x, s) = H(p, x)$ means that the forcing $F(t, x) = F(x)$ is conservative, so the Hamiltonian, i.e., the energy, is perserved.

7.2 Calculus of variation and Lagrangian mechanics

We know that intersection of characteristics may cause problems, so we aim to try another representation of the solution.

We introduce the Lagrangian $L(q, x, t)$ and the variational problem

$$\inf_{w \in \mathcal{A}} I[w] = \inf_{w \in \mathcal{A}} \int_0^t L\big(\dot{w}(s), w(s), s\big)\, ds,$$

where the admissible set

$$\mathcal{A} = \{w \in \mathcal{C}^2[0, t] : w(0) = y,\ w(t) = x\}.$$

The idea of Lagrangian mechanics is that the minimizer of the variational problem $t \mapsto w(t)$ should give the trajectory of the system in the phase space.

Suppose that η is the minimizer. For any $v \in \mathcal{C}_0^2[0, t]$, the function $i(\varepsilon) = I[\eta + \varepsilon v]$ achieves the minimum at $\varepsilon = 0$, and hence $i'(0) = 0$. Direct computation gives

$$\begin{aligned}0 = i'(0) &= \frac{d}{d\varepsilon}\Big|_{\varepsilon=0} \int_0^t L(\dot{\eta} + \varepsilon \dot{v}, \eta + \varepsilon v, s)\, ds \\ &= \int_0^t \big[(\partial_q L) \cdot \dot{v} + (\partial_x L) v\big]\, ds \\ &= \int_0^t \Big[\frac{d}{ds}(-\partial_q L) + \partial_x L\Big] v\, ds.\end{aligned}$$

Since this holds for all $v \in \mathcal{C}_0^2[0, t]$, we must have

$$-\partial_x L + \frac{d}{ds}(\partial_q L) \equiv 0. \tag{7.4}$$

This is the *Euler–Lagrange equation.*

For example, to model a particle moving in a potential, we have

$$L(q, x, s) = \frac{1}{2} m q^2 - \Phi(s, x).$$

Then

$$\partial_x L = -\partial_x \Phi(s, x), \quad \partial_q L = mq.$$

(7.4) becomes

$$\Phi'\big(s, \eta(s)\big) + \frac{d}{ds}\big(m\dot{\eta}(s)\big) = 0 \implies m\ddot{\eta}(s) = -\Phi'\big(s, \eta(s)\big),$$

that is, the accelaration times mass is the the external force which is the negative gradient of the potential field.

To make a connection to the Hamilton's PDE (7.3), we define

$$p(s) = \partial_q L(\dot{\eta}, \eta, s).$$

Suppose $p = \partial_q L(q, x, s)$ can be solved by

$$q = q(p, x, s),$$

and let

$$H(p, x) = p \cdot q(p, x, s) - L\big(q(p, x), x, s\big).$$

Then the Euler–Lagrangian equation (7.4) becomes (7.3). Indeed, we have

$$\dot{\eta}(s) = q\big(p(s), \eta(s), s\big), \quad \dot{p}(s) = \partial_x L,$$

by the definition of $q(\cdot)$ and the Euler–Lagrange equation, while

$$\partial_p H = q(p, x, s) + p \cdot \frac{\partial q}{\partial p} - \partial_q L \cdot \frac{\partial q}{\partial p} = q,$$

and

$$\partial_x H\big(p(s), \eta(s), s\big) = p \cdot \frac{\partial q}{\partial x} - \frac{\partial L}{\partial q}\frac{\partial q}{\partial x} - \partial_x L = -\partial_x L.$$

7.3 Legendre transform

The functions L and H are associated by the so-called *Legendre transform*, or convex dual.

Let $L(q) : \mathbb{R}^d \to \mathbb{R}$ be a convex function with super-linear growth at ∞, that is,

$$\lim_{|q|\to\infty} \frac{L(q)}{|q|} = +\infty.$$

Its Legendre transform $L^*(p)$ is defined by

$$L^*(q) = \sup_{q\in\mathbb{R}^d} \{p \cdot q - L(q)\}.$$

If the function $L(q)$ is sufficiently nice so that the supremum can ba acheived, then the Legendre transform takes a simpler form. Indeed, suppose that q^* is the point of maximum; then

$$\frac{\partial}{\partial q}\Big|_{q=q^*}\big(p \cdot q - L(q)\big) = 0 \implies p = \partial_q L(q^*).$$

Hence,

$$L^*(q) = p \cdot q^* - L(q^*),$$

where q^* solves $p = \partial_q L(q^*)$.

The Legendre transform is really a duality.

Proposition 7.2 *The function $p \mapsto H(p)$ is convex with super-linear growth, and $L = H^* = L^{**}$.*

We recall a fact about convex functions.

Lemma 7.3 *Let $\{\varphi_\alpha\}_{\alpha \in A}$ be a family of convex functions. Then*

$$\varphi^*(x) = \sup_{\alpha \in A} \varphi_\alpha(x)$$

is also convex.

Proof: We recall that a function f is convex, if and only if for all x_0, y_0 satisfies $\varphi(x_0) > y_0$, there is a hyperplane $y = \ell(x)$ passing through y_0 such that $\varphi(x) \geq \ell(x)$ for all x.

If $\varphi^*(x_0) > y_0$, then by definition, there exists φ_α such that $\varphi_\alpha(x_0) > y_0$, and hence there is a hyperplane $y = \ell(x)$ such that $\varphi_\alpha \geq \ell$, so $\varphi^* \geq \varphi_\alpha \geq \ell$ as desired. $\square$

Proof:

Since $q \mapsto p \cdot q - L(q)$ are a family of convex (in fact, linear) functions, the convexity $p \mapsto H(p)$ follows from Lemma 7.3.

For $\lambda > 0$, we have

$$H(p) \geq \lambda|p| - L(\lambda) \geq \lambda|p| - \max_{|q| \leq \lambda} L(q).$$

Hence,

$$\liminf_{|p| \to \infty} \frac{H(p)}{|p|} \geq \liminf_{|p| \to \infty} \lambda - \frac{\max_{|q| \leq \lambda} L(q)}{|p|} \geq \lambda$$

for every λ. Therefore, H is super-linear.

By the definition, we have

$$H(p) + L(q) \geq p \cdot q, \quad \forall p, q \in \mathbb{R}.$$

So

$$L(q) \geq p \cdot q - H(p), \quad \forall p \implies L(q) \geq H^*(q).$$

On the other hand, since $L(q)$ is convex, there exists a hyperplane passing through q that is below $L(\cdot)$. We write this hyperplane as

$$\ell(r) = a \cdot (r - q) + L(q) \leq L(r).$$

Then

$$\begin{aligned}
H^*(q) &= \sup_{p \in \mathbb{R}^d} \{p \cdot q - H(p)\} \\
&= \sup_{p \in \mathbb{R}^d} \{p \cdot q - \sup_{r \in \mathbb{R}^d} \{p \cdot r - L(r)\}\} \\
&= \sup_p \inf_r \{p \cdot (q - r) + L(r)\} \\
&\geq \inf_r \{a \cdot (q - r) + L(r)\} \geq L(q).
\end{aligned}$$

This completes the proof. $\square$

7.4 Solve Hamilton–Jacobi via variational formula

We propose that the solution to

$$\partial_t u + H(\partial_x u, x, t) = 0, \quad u(0,x) = g(x), \tag{7.5}$$

is given by

$$u(t,x) = \inf\Big\{ \int_0^t L\big(\dot{w}(s), w(s), s\big)\, ds + g\big(w(0)\big) : w \text{ Lipschitz}, w(t) = x \Big\}, \tag{7.6}$$

where $L(q,x,t) = H^*(p,x,t)$. We will discuss in what sense (7.6) gives the solution to (7.5).

7.4.1 Hopf–Lax Formula

In this section we assume

$$H(p,x,t) = H(p), \quad L(q,x,t) = L(q),$$

where H and L are convex functions with super-linear growth and they are the Legendre transform of each other.

Lemma 7.4 (Hopf–Lax formula) *If g has at most linear growth, then u in (7.6) is given by*

$$u(t,x) = \min_y \Big\{ tL\Big(\frac{x-y}{t}\Big) + g(y) \Big\}. \tag{7.7}$$

Proof:

Considering the straight line

$$w_0(s) = y + \frac{s(x-y)}{t}, \quad s \in [0,t],$$

we have

$$u(t,x) \le \int_0^t L\big(\dot{w}_0(s)\big)\, ds + g\big(w_0(0)\big) = tL\Big(\frac{x-y}{t}\Big) + g(y).$$

On the other hand, since L is convex, for any Lipschitz function w, by Jensen's inequality,

$$L\Big(\frac{x-y}{t}\Big) = L\Big(\frac{1}{t}\int_0^t \dot{w}(s)\, ds\Big) \ge \frac{1}{t}\int_0^t L\big(\dot{w}(s)\big)\, ds.$$

Hence,

$$\inf\Big\{ \int_0^t L\big(\dot{w}(s)\big)\, ds + g\big(w(0)\big) : w \text{ Lipschitz}, w(t) = x \Big\} = \inf_y \Big\{ tL\Big(\frac{x-y}{t}\Big) + g(y) \Big\}.$$

The second infimum can be acheived since g has at most linear growth while L has super-linear growth. This completes the proof. □

Lemma 7.5 (Dynamic programming) *For $s \in (0,t)$,*

$$u(t,x) = \inf_y \Big\{ (t-s)L\Big(\frac{x-y}{t-s}\Big) + u(s,y) \Big\}.$$

Proof:

For every $\varepsilon > 0$, by Lemma 7.4, there exists z such that

$$u(s,y) \le sL\Big(\frac{y-z}{s}\Big) + g(z) \le u(s,y) + \varepsilon.$$

Hence,

$$\begin{aligned} u(t,x) &\le t\Big(\frac{x-z}{t}\Big) + g(z) \\ &\le (t-s)L\Big(\frac{x-y}{t-s}\Big) + sL\Big(\frac{y-z}{s}\Big) + g(z) \\ &\le u(s,y) + (t-s)L\Big(\frac{x-y}{t-s}\Big) + \varepsilon. \end{aligned}$$

Since $\varepsilon > 0$ is arbitrary, we obtain

$$u(t,x) \le \inf_y \Big\{(t-s)L\Big(\frac{x-y}{t-s}\Big) + u(s,y)\Big\}.$$

For the other direction, let

$$y = \frac{s}{t}x + \frac{t-s}{t}z.$$

Then

$$\begin{aligned} &(t-s)L\Big(\frac{x-y}{t-s}\Big) + u(y,z) \\ &\le (t-s)L\Big(\frac{x-y}{t-s}\Big) + sL\Big(\frac{x-z}{s}\Big) + g(z) \\ &= tL\Big(\frac{x-z}{t}\Big) + g(z) \\ &\le u(t,x) + \varepsilon. \end{aligned}$$

This completes the proof. □

Lemma 7.6 *Assume that g is Lipschtiz. Then u is Lipschitz in t and x and*

$$\lim_{t\to 0} u(t,x) = g(x).$$

Proof:

Suppose

$$u(t,x) = tL\Big(\frac{x-y}{t}\Big) + g(y).$$

Let $x' = x + h$. Then

$$\begin{aligned} u(t,x') &\le tL\Big(\frac{x'-(y+h)}{t}\Big) + g(y+h) \\ &\le u(t,x) + g(y+h) - g(y) \\ &\le u(t,x) + \mathrm{Lip}(g)\cdot|x-x'|. \end{aligned}$$

Switching the roles of x and x' we obtain

$$|u(t,x) - u(t,x')| \le \mathrm{Lip}(g)|x-x'|.$$

For the Lipchitz continuity in t, it suffices to show that if (7.7) holds, then

$$|u(t,x)-g(x)| \le Ct.$$

where C is uniform in a neighborhood of x. Then the Lipschitz continuity in t follows from this and Lemma 7.5.

For one direction, we have

$$u(t,x) \le tL(0)+g(x).$$

For the other direction, we have

$$\begin{aligned}
u(t,x) &\ge g(x)+\min_y\Big\{-\mathrm{Lip}(g)(x-y)+tL\Big(\frac{x-y}{t}\Big)\Big\} \\
&\ge g(x)-t\max_z\{\mathrm{Lip}(g)|z|-L(z)\} \\
&\ge g(x)-t\max_{|w|\le \mathrm{Lip}(g)} H(w).
\end{aligned}$$

The conclusion follows. □

The next result justifies calling (7.7) a solution to (7.5).

Theorem 7.7 *If u is differentiable at (t,x), then*

$$\partial_t u + H(\nabla u) = 0$$

holds at (t,x).

Proof:

For $h>0$ and $q\in\mathbb{R}^d$, we have

$$\begin{aligned}
u(t+h,x+hq)-u(t,x) &= \min_y\Big\{hL\Big(\frac{x+hq-y}{h}\Big)+u(t,y)\Big\}-u(t,x) \\
&\le hL(q).
\end{aligned}$$

Letting $h\downarrow 0$, we obtain

$$\partial_t u + q\cdot\nabla u \le L(q).$$

Hence,

$$\partial_t u + H(\nabla u) = \partial_t u + \max_q\{q\cdot\nabla u - L(q)\} \le 0.$$

Let $z\in\mathbb{R}^d$ be such that

$$u(t,x) = tL\Big(\frac{x-z}{t}\Big)+g(z).$$

Let $h>0$ and

$$s=t-h, \quad y=\frac{t-h}{t}x+\frac{h}{t}z.$$

(so that $(0,z),(s,y),(t,x)$ are coliear). Then

$$\begin{aligned}
u(t,x)-u(s,y) &\ge tL\Big(\frac{x-z}{t}\Big)+g(z)-sL\Big(\frac{y-z}{s}\Big)-g(z) \\
&= (t-s)L\Big(\frac{x-y}{t-s}\Big).
\end{aligned}$$

Letting $h\downarrow 0$, we obtain

$$\partial_t u + q\cdot\nabla u - L(q) \ge 0, \quad q=\frac{x-z}{t},$$

and hence

$$\partial_t u + H(\nabla u) = \partial_t u + \max_q\{q\cdot\nabla u - L(q)\} \ge 0.$$

This completes the proof. □

7.4.2 General case

In this section we consider the Hamilton–Jacobi equation

$$\partial_t u + H(\nabla u) = F(t,x), \quad u(0,x) = g(x). \tag{7.8}$$

We will prove that

$$u(t,x) = \inf\Big\{ g(w(0)) + \int_0^t \Big[L\big(\dot w(s)\big) + F\big(s, w(s)\big) \Big]\, ds : w \text{ [[roam:Strong convergence of an explicit numerical metho} \tag{7.9}$$

solves (7.9). When $F \equiv 0$, (7.9) reduces to (7.7).

We start with the dynamic programming principle.

Lemma 7.8 *Let $s \in (0,t)$. Then*

$$u(t,x) = \inf\Big\{ u\big(s, w(0)\big) + \int_s^t \Big[L\big(\dot w(r)\big) + F\big(r, w(r)\big) \Big]\, dr : w \textit{ Lipschitz},\ w(t) = x \Big\}.$$

Proof: Let

$$A_{z,y}^{s,t} = \inf\Big\{ \int_s^t \Big[L\big(\dot w(r)\big) + F\big(r, w(r)\big) \Big]\, dr : w \text{ Lipschitz},\ w(t) = y,\ w(s) = z \Big\}.$$

Then A is sub-additive, that is,

$$A_{x_1,x_2}^{t_1,t_2} + A_{x_2,x_3}^{t_2,t_3} \geq A_{x_1,x_3}^{t_1,t_3} = \inf_{x_2}\{A_{x_1,x_2}^{t_1,t_2} + A_{x_2,x_3}^{t_2,t_3}\}, \quad \forall t_1 < t_2 < t_3.$$

Also,

$$u(t,x) = \inf\{ g(y) + A_{y,x}^{0,t} : y \in \mathbb{R}^d \}. \tag{7.10}$$

We have for every y and z,

$$\begin{aligned} u(t,x) &\leq g(y) + A_{y,x}^{0,t} \\ &\leq g(y) + A_{y,z}^{0,s} + A_{z,y}^{s,t}. \end{aligned}$$

Taking the infimum in y and using (7.10), we have

$$u(t,x) \leq \inf\{ u(s,z) + A_{z,y}^{s,t} \}.$$

For the other direction, for every $\varepsilon > 0$, there exists y^* such that

$$g(y^*) + A_{y^*,x}^{0,t} \leq u(t,x) + \varepsilon,$$

and there exists z^* such that

$$A_{y^*,x}^{0,t} + \varepsilon \geq A_{y^*,z^*}^{0,s} + A_{z^*,x}^{s,t}.$$

Hence,

$$\begin{aligned} u(t,x) &\geq A_{y^*,x}^{0,t} + g(y^*) - \varepsilon \\ &\geq A_{y^*,z^*}^{0,s} + A_{z^*,x}^{s,t} + g(y^*) - 2\varepsilon \\ &\geq u(s,z^*) + A_{z^*,x}^{s,t} - 2\varepsilon. \end{aligned}$$

Since $\varepsilon > 0$ is arbitrary, this completes the proof. $\square$

Lemma 7.9 *If g and $F(t,\cdot)$ are Lipschitz, and F is bounded, then u given by (7.10) is Lipschitz in t and x.*

Proof:

Assume

$$|g(x)-g(\hat{x})|, |F(t,x)-F(t,\hat{x})| \le K|x-\hat{x}|.$$

For the Lipschitz continuity in x, it suffices to show that $A_{y,x}^{0,t}$ is Lipschitz in y and x. For every $\varepsilon > 0$, there exists w Lipschitz such that

$$\Big| \int_0^t \big[L\big(\dot{w}(s)\big) + F\big(s, w(s)\big)\big]\, ds - A_{y,x}^{0,t} \Big| \le \varepsilon.$$

For $\hat{x} \ne x$, we define

$$\hat{w}(s) = w(s) + \hat{x} - x,$$

so that

$$\hat{w}'(s) = w'(s), \quad \hat{w}(t) = \hat{x}, \quad \hat{w}(0) = \hat{y} := y + \hat{x} - x.$$

Then

$$\begin{aligned}
u(t,\hat{x}) &\le A_{\hat{y},\hat{x}}^{0,t} + g(\hat{y}) \\
&\le \int_0^t L\big(\hat{w}'(s)\big) + F\big(s,\hat{w}(s)\big)\, ds + g(\hat{y}) \\
&\le \varepsilon + K|x-\hat{x}| + A_{y,x}^{0,t} + \Big| \int_0^t F\big(s,\hat{w}(s)\big) - F\big(s,w(s)\big)\, ds \Big| + g(y) \\
&\le \varepsilon + (K + Kt)|x-\hat{x}| + A_{y,x}^{0,t} + g(y)
\end{aligned}$$

First taking the infimum over y, then letting $\varepsilon \downarrow 0$, we obtain

$$u(t,\hat{x}) \le K(t+1)|x-\hat{x}| + u(t,x).$$

Switching the roles of x and $\hat{x}$ we get the other inequality, and combining them we obtain the Lipchitz continuity of u in x.

For $t > s$, we have

$$u(t,x) - u(s,x) \le A_{x,x}^{s,t} \le (t-s)\|F\|_{L^\infty} + (t-s)L(0),$$

and

$$\begin{aligned}
u(t,x) - u(s,x) &\ge -K|x-y| + (t-s)L\Big(\frac{x-y}{t-s}\Big) \\
&\ge -(t-s)\sup_z\{K\cdot|z| - L(z)\} \\
&\ge -(t-s)\max_{|w|\le K} H(w).
\end{aligned}$$

This completes the proof. □

Theorem 7.10 *If $u(t,x)$ is differentiable at (t,x), then it satisfies (7.8) at (t,x).*

Proof: For $h > 0$ and $q \in \mathbb{R}^d$, we have

$$u(t+h, x+hq) - u(t,x) \le A^{t,t+h}_{x,x+hq} \le hL(q) + h \cdot \big(F(t,x) + o(1)\big).$$

Taking $h \to 0+$ and using the differentiability at (t,x), we have

$$\partial_t u + \nabla u \cdot q \le L(q) + F(t,x).$$

Since this inequality holds for all $q \in \mathbb{R}^d$, we have

$$\partial_t u(t,x) + H\big(\nabla u(t,x)\big) \le F(t,x).$$

For the other direction, let $s = t - h$, $h > 0$, and assume that the infimum in Lemma 7.8 can be acheived (this can be proved using that F is bounded and L has super-linear growth). Let $y \in \mathbb{R}^d$ be such that

$$u(t,x) = u(s,y) + A^{s,t}_{y,x}.$$

Writing $x - y = qh$, we have

$$u(t,x) - u(s,y) = A^{s,t}_{y,x} \ge hL(q) + h\Big(F(t,x) + o(1)\Big),$$

since for all w Lipschitz, by convexity of L the following inequality holds:

$$\int_s^t L\big(\dot{w}(r)\big)\, dr \ge (t-s)L\Big(\frac{w(t) - w(s)}{t-s}\Big).$$

Using differentiability at (t,x), we have

$$u(t,x) - u(s,y) = u(t,x) - u(t-h, x-qh) = h\partial_t u + \nabla u \cdot qh + o(h) \ge hL(q) + h\big(F(t,x) + o(1)\big).$$

Dividing both side by h, we obtain

$$\partial_t u(t,x) + \nabla u(t,x) \cdot q_h \ge L(q_h) + F(t,x) + o(1), \forall h > 0,$$

where q depends on h since y depends on h. Hence,

$$H\big(\nabla u(t,x)\big) = \sup_q \{\nabla u(t,x) \cdot q - L(q)\} \ge F(t,x) - \partial_t u(t,x).$$

This completes the proof. $\square$

7.4.3 Semi-concavity

We say that a function f is semi-concave if there exists a constant C such that

$$f(x+z) + f(x-z) - 2f(x) \le C|z|^2.$$

If $f \in \mathcal{C}^2$, this condition is equivalent to f'' being bounded from above.

The variational solution will be semi-concave.

Theorem 7.11 *If for some $\theta > 0$,*

$$\sum_{i,j} \partial_{p_i p_j} H \xi_i \xi_j \ge \theta |\xi|^2, \quad \forall \xi \in \mathbb{R}^d, \tag{7.11}$$

and F has bounded second derivative, then there exists $C > 0$ such that

$$u(t, x+z) + u(t, x-z) - 2u(t,x) \le \frac{C}{t}|z|^2, \quad z \in \mathbb{R}^d.$$

We need a lemma quatifying the convexity of H and L. We omit its proof here.

Lemma 7.12 *Assume (7.11). Then*

$$H\Big(\frac{p_1+p_2}{2}\Big) \le \frac{1}{2}H(p_1) + \frac{1}{2}H(p_2) - \frac{\theta}{8}|p_1-p_2|^2,$$
$$\frac{1}{2}L(q_1) + \frac{1}{2}L(q_2) \le L\Big(\frac{q_1+q_2}{2}\Big) + \frac{1}{8\theta}|q_1-q_2|^2.$$

Proof:

Let $z \in \mathbb{R}^d$. For any Lipschitz w with $w(0) = y$ and $w(t) = x$, we define

$$w_\pm(s) = w(s) \pm \frac{s}{t}\cdot z,$$

which moves the endpoint $w(t)$ to $x \pm z$. Since

$$\dot{w}_\pm(s) = \dot{w}(s) \pm \frac{z}{t},$$

by Lemma 7.12, we have

$$\int_0^t L\big(\dot{w}_-(s)\big)\,ds + \int_0^t L\big(\dot{w}_+(s)\big)\,ds - 2\int_0^t L\big(\dot{w}(s)\big)\,ds \le \frac{1}{8\theta}\int_0^t \frac{z^2}{t^2}\,ds = \frac{1}{8\theta}\frac{z^2}{t}.$$

Also,

$$\int_0^t F\big(s, w_-(s)\big)\,ds + \int_0^t F\big(s, w_+(s)\big)\,ds - 2\int_0^t F\big(s, w(s)\big)\,ds \le \int_0^t \Big|\frac{sz}{t}\Big|^2 \|F\|_{\mathcal{C}^2}\,ds = \frac{\|F\|_{\mathcal{C}^2}}{3}\frac{z^2}{t}.$$

Hence,

$$u(t, x+z) + u(t, x-z) \le 2\Big[\int_0^t L\big(\dot{w}(s)\big) + F\big(s, w(s)\big)\,ds + g\big(w(0)\big)\Big] + \frac{C}{t}z^2$$

for all Lipschitz w. Taking infimum over w we obtain the desired result. □

If we imposing the concavity condition, then any solution u that satisfies (7.9) almost everywhere will be unique. See [Eva, 3.3.3.b, Theorem 7].

7.5 Conservation laws and Rankine–Hugoniot condition

From the method of characteristics, we know that the solution to

$$\partial_t u + \big[H(u)\big])_x = 0, \quad u(0, x) = g(x) \tag{7.12}$$

may not non-smooth. (7.12) is the derivative of (7.5) with $H(p, x, t) = H(p)$. Equations of the form (7.12) is called *conservation law*.

We want to define a notion of weak solution to (7.12). Let $v \in \mathcal{C}_0^\infty(\mathbb{R}^2)$. Assuming u is smooth, using integration by parts we have

$$0 = \int_0^\infty dt \int_{\mathbb{R}} \Big(u_t + [H(u)]_x\Big) v\,dx$$
$$= -\int_0^\infty dt \int_{-\infty}^\infty u v_t\,dx - \int_{-\infty}^\infty g(x)v(0,x)\,dx - \int_0^\infty dt \int_{-\infty}^\infty H(u)v_x.$$

We say that u is a weak solution to (7.12) if for all $v \in \mathcal{C}_0^\infty(\mathbb{R}^2)$,

$$\int_0^\infty dt \int_{-\infty}^\infty (uv_t + H(u)v_t x)\, dxdt + \int_{-\infty}^\infty g(x)v(0,x)\, dx = 0. \tag{7.13}$$

Let u be a weak solution, and assume that u is $\mathcal{C}^1$ in $V_\ell \cap V_r$, but has discontinuity on the curve

$$C = V_\ell \cap V_r = \{(t,x) : x = s(t)\}.$$

Since u is $\mathcal{C}^1$ in $V_\ell \cup V_r$, it satisfies (7.12) in this region. We want to derive a condition on the jump of u along γ. Let $v \in \mathcal{C}_c^\infty(\mathbb{R}^2)$. Without loss of generality assume v vanishes outside $V_\ell \cup V_r$ and near $t = 0$.

Using integration by parts, we have

$$\iint_{V_\ell} uv_t + H(u)v_t\, dtdx = -\iint_{V_\ell} (u_t + H(u)_x)v\, dtdx + \int_C \nu_\ell \cdot (uv, H(u)v) = \int_C \nu_\ell \cdot \big(u_\ell v, H(u_\ell)v\big)\, dS.$$

Similarly,

$$\iint_{V_r} uv_t + H(u)v_t\, dtdx = \int_C \nu_r \cdot \big(u_r v, H(u_r)v\big)\, dS.$$

But on C, $\nu_r = -\nu_\ell$, so

$$\big(u_r - u_\ell, H(u_r) - H(u_\ell)\big) \cdot \nu_\ell = 0, \qquad \text{on } C.$$

Hence, we arrive at the *Rankine–Hugoniot condition*:

$$[[H(u)]] = \dot{s}[[u]], \tag{7.14}$$

where $[[f]]$ denotes the jump of f.

For Burgers equation, $H(u) = \frac{u^2}{2}$, and (7.14) becomes

$$\frac{u_\ell + u_r}{2} = \dot{s}.$$

7.6 Viscosity solution

For $\varepsilon > 0$, let u^ε solve

$$u_t^\varepsilon + H(\nabla u^\varepsilon, x) = \varepsilon \Delta u^\varepsilon + F(t,x), \quad u^\varepsilon(0,x) = g(x),$$

where F is a continuous function. This is a nonlinear parabolic equation, but using general theory (like those presented in the previous section), one can show that there exists a solution $u^\varepsilon \in \mathcal{C}^\infty\big((0,\infty)\times\mathbb{R}^d\big)$. We want to define the limit of u^ε as $\varepsilon \downarrow 0$ as the solution. The general picture is, $u^\varepsilon \to u$ uniformly on compact sets, but u_x^ε and u_t^ε are not controlled. Hence the corresponding conservation law equation (7.12) will have jumps.

We want to give a direct characterization of the vanishing viscosity limit $u = \lim_{\varepsilon\downarrow 0} u^\varepsilon$.

Let $v \in \mathcal{C}^\infty([0,\infty) \times \mathbb{R}^d)$ satisfy $u - v$ has strict local maximum at (t_0, x_0). First, we claim that for ε sufficiently small, $u^\varepsilon - v$ has a local maximum at $(t_\varepsilon, x_\varepsilon)$ and $(t_\varepsilon, x_\varepsilon) \to (t_0, x_0)$. Indeed, for r small enough,

$$(u - v)(t_0, x_0) > \max_{\partial B_r(t_0,x_0)} (u - v)(t,x),$$

and hence by uniform convergence,

$$(u^\varepsilon - v)(t_0, x_0) > \max_{\partial B_r(t_0,x_0)} (u^\varepsilon - v)(t,x).$$

Therefore, the maximum of $(u^\varepsilon - v)$ on $\overline{B_r(t_0,x_0)}$ is not on $\partial B_r(t_0,x_0)$. Let $(t_\varepsilon, x_\varepsilon)$ be the point of maximum in $B_r(t_0,x_0)$. Then this point will satisfy the claim.

Since $(t_\varepsilon, x_\varepsilon)$ is a local maximum for $u^\varepsilon - v$, at this point we have

$$\partial_t u^\varepsilon = \partial_t v, \quad -\Delta u^\varepsilon \geq -\Delta v, \quad \nabla u^\varepsilon = \nabla v.$$

Hence,

$$v_t(t_\varepsilon, x_\varepsilon) + H(\nabla v(t_\varepsilon, x_\varepsilon), x_\varepsilon) \leq \varepsilon \Delta v(t_\varepsilon, x_\varepsilon) + F(t_\varepsilon, x_\varepsilon).$$

Since v is smooth and $(t_\varepsilon, x_\varepsilon) \to (t_0, x_0)$, we have

$$v_t(t_0,x_0) + H\big(\nabla v(t_0,x_0), x_0\big) \leq F(t_0,x_0).$$

We can further replace the "strict local maximum" by "local maximum", since if $u - v$ has a local maximum at (t_0,x_0), then $u - \tilde{v}$ where

$$\tilde{v} = v + \delta\big(|x-x_0|^2 + |t-t_0|^2\big), \quad \delta > 0,$$

will have a strict local maximum at (t_0,x_0). To summarize, assuming $u^\varepsilon \to u$ on compact sets, then

$$v \in \mathcal{C}^\infty,\ u - v \text{ has local max at } (t_0,x_0) \implies v_t + H(\nabla v, x) \leq F(t,x) \text{ at } (t_0,x_0). \tag{7.15}$$

Similarly,

$$v \in \mathcal{C}^\infty,\ u - v \text{ has local min at } (t_0,x_0) \implies v_t + H(\nabla v, x) \geq F(t,x) \text{ at } (t_0,x_0). \tag{7.16}$$

Definition 7.1 *A bounded, uniformly continuous function u on $[0,\infty) \times \mathbb{R}^d$ is a viscosity solution to*

$$u_t + H(\nabla u, x) = F(t,x) \tag{7.17}$$

if (7.15) and (7.16) hold.

Theorem 7.13 *If u is a viscosity solution and u is differentiable at (t_0,x_0), then (7.17) holds at (t_0,x_0).*

We cite a result without proof.

Lemma 7.14 *If u is differentiable at (t_0,x_0), there exists $v \in \mathcal{C}^1$ such that*

$$(u-v)(t_0,x_0) = 0, \quad u - v \textit{ has a strictly local maximum at } (t_0,x_0).$$

Proof:

Let v be given by Lemma 7.14 and

$$v^\varepsilon = \eta^\varepsilon * v,$$

where η^ε is some $\mathcal{C}_c^\infty$ mollifiers. Then $v^\varepsilon \in \mathcal{C}^\infty$

$$v^\varepsilon \to v, \quad \nabla v^\varepsilon \to \nabla v, \quad v_t^\varepsilon \to v_t,$$

uniformly near (t_0,x_0). Moreover, there exist $(t_\varepsilon, x_\varepsilon)$ where $u - v^\varepsilon$ achieves the local maximum, with $(t_\varepsilon, x_\varepsilon) \to (t_0,x_0)$. Then, by the definition of viscosity solution we have

$$\partial_t v^\varepsilon(t_\varepsilon, x_\varepsilon) + H\big(\nabla v(t_\varepsilon, x_\varepsilon), x_\varepsilon\big) \leq F(t_\varepsilon, x_\varepsilon).$$

Taking the limit $\varepsilon \to 0+$, we obtain

$$v_t(t_0,x_0) + H\big(\nabla v(t_0,x_0), x_0\big) \leq F(t_0,x_0).$$

Since $u - v$ is differentiable at (t_0, x_0) and achieves a local maximum, we have

$$u_t(t_0, x_0) = v_t(t_0, x_0), \quad \nabla u(t_0, x_0) = \nabla v(t_0, x_0).$$

Hence,

$$u_t(t_0, x_0) + H\big(\nabla u(t_0, x_0), x_0\big) \leq F(t_0, x_0).$$

We can get the other direction of the inequality in a similar way. Therefore, u solves (7.17) at (t_0, x_0), as desired. $\square$

Proposition 7.15 *The variational formula (7.10) indeed gives a viscosity solution to (7.17).*

Proof: By Lemma 7.8, if $u - v$ has a local maximum at (t_0, x_0), then

$$v(t_0, x_0) + u(t_0 - h, x) - v(t_0 - h, x) \leq u(t_0, x_0) = \inf\big\{u(t_0 - h, y) + A_{y,x_0}^{t_0 - h, t_0}\big\} \leq u(t_0 - h, x) + A_{x,x_0}^{t_0 - h, t_0}.$$

Hence, for x close to x_0,

$$v(t_0, x_0) - v(t_0 - h, x) \leq A_{x,x_0}^{t_0 - h, t_0} \leq h \cdot L\Big(\frac{x_0 - x}{h}\Big) + hF(t_0, x_0) + o(h).$$

Let $x = x_0 + qh$. Then

$$v_t(t_0, x_0) + q \cdot \nabla v(t_0, x_0) - L(q) \leq F(t_0, x_0), \quad \forall q.$$

Therefore,

$$v_t(t_0, x_0) + H\big(\nabla v(t_0, x_0)\big) \leq F(t_0, x_0).$$

If $u - v$ has a local minimum at (t_0, x_0), then

$$u(t_0, x_0) - v(t_0, x_0) \leq u(t_0 - h, x) - v(t_0 - h, x), \quad \forall x.$$

Assuming that the infimum in Lemma 7.8 can be achieved, for every $h > 0$, there exists $q = q_h$ such that

$$u(t_0, x_0) = u(t_0 - h, x_0 - qh) + A_{x_0 - qh, x_0}^{t_0 - h, x_0}.$$

Hence,

$$A_{x_0 - qh, x_0}^{t_0 - h, t_0} \leq v(t_0, x_0) - v(t_0 - h, x - qh).$$

The LHS is

$$hL(q) + hF(t_0, x_0) + o(h),$$

while by the differentiability, the RHS is

$$v_t(t_0, x_0) \cdot h + v_x(t_0, x_0) \cdot qh + o(h).$$

Therefore,

$$F(t_0, x_0) \leq v_t(t_0, x_0) + v_x(t_0, x_0) \cdot q_h - L(q_h) + o_h(1) \leq v_t(t_0, x_0) + H\big(\nabla v(t_0, x_0)\big) + o_h(1).$$

Letting $h \downarrow 0$ yields the desired inequality. $\square$

Finally, we have the uniqueness of the solution.

Theorem 7.16 *Assume that $H(p,x)$ satisfies*

$$|H(p,x)-H(q,x)| \le K|p-q|, \quad |H(p,x)-H(p,y)| \le K|x-y|(1+|p|).$$

for some constant $K>0$. Then There exists at most one viscosity solution to (7.17).

Proof: Let u and $\tilde{u}$ be two viscosity solution. If $\tilde{u} \neq u$, let

$$\sigma := \sup(u-\tilde{u}) > 0.$$

We introduce

$$\Phi(t,s,x,y) = u(t,x) - u(s,y) - \lambda(t+s) - \frac{1}{\varepsilon^2}\big(|x-y|^2 + |t-s|^2\big) - \varepsilon\big(|x|^2+|y|^2\big), \tag{7.18}$$

with ε, λ small. For all $\varepsilon, \lambda \in (0,1)$, Φ achieves the maximum; let (t_0, s_0, x_0, y_0) be the point of maximum.

We have some bounds on t_0, s_0, x_0, y_0. First, due to the fourth term in (7.18), we have

$$|x_0-y_0|, |t_0-s_0| \le O(\varepsilon). \tag{7.19}$$

Second, (7.19) implies

$$\varepsilon\big(|x|^2+|y|^2\big) = O(1) \implies \varepsilon\big(|x|+|y|\big) = O(\varepsilon^{1/2}).$$

Since $\Phi(t_0,s_0,x_0,y_0) \ge \Phi(t_0,t_0,x_0,x_0)$, we have

$$\frac{1}{\varepsilon^2}\big(|x_0-y_0|^2+|t_0-s_0|^2\big) \le \tilde{u}(t_0,x_0) - \tilde{u}(s_0,y_0) + \lambda(t_0-s_0) + \varepsilon(x_0+y_0)\cdot(x_0-y_0).$$

By (7.19), the RHS goes to 0 as $\varepsilon \downarrow 0$, so we can improve (7.19) to

$$|x_0-y_0|, |t_0-s_0| \le o(\varepsilon).$$

Third, there exists $\mu > 0$ such that

$$t_0, s_0 \ge \mu$$

for ε, λ sufficiently small; indeed, for ε and λ sufficiently small, we can have

$$\Phi(t_0,s_0,x_0,y_0) \ge \sup_{[0,T]\times\mathbb{R}^n} \Phi(t,t,x,x) \ge \frac{\sigma}{2} > 0,$$

while

$$\limsup_{\mu\to 0+} \sup_{t,s\le\mu} \Phi(t,s,x,y) = 0,$$

since $u(0,x) = \tilde{u}(0,\tilde{x})$ and $u, \tilde{u}$ are uniformly continuous.

We now proceed to the proof of uniqueness. Since $(t,x) \mapsto \Phi(t,s_0,x,y_0)$ has a maximum at the point (t_0,x_0), we have

$$u - v \text{ has a maximum at } (t_0,x_0)$$

where

$$v(t,x) := \tilde{u}(s_0,y_0) + \lambda(t+s_0) + \frac{1}{\varepsilon^2}\big(|x-y_0|^2+|t-s_0|^2\big) + \varepsilon\big(|x|^2+|y|^2\big).$$

Since $v \in \mathcal{C}^\infty$ and u is a viscosity solution, we have

$$v_t(t_0,x_0) + H\big(\nabla v(t_0,x_0), x_0\big) \le F(t_0,x_0),$$

that is,

$$\lambda + \frac{2(t_0 - s_0)}{\varepsilon^2} + H\Big(\frac{2(x_0 - y_0)}{\varepsilon^2} - 2\varepsilon x_0, x_0\Big) \le F(t_0, x_0).$$

Similarly, we have

$$-\lambda + \frac{2(t_0 - s_0)}{\varepsilon^2} + H\Big(\frac{2(x_0 - y_0)}{\varepsilon^2} - 2\varepsilon y_0, y_0\Big) \ge F(s_0, y_0).$$

Taking the difference, and using the condition on H, we have

$$\begin{aligned} 2\lambda &\le F(t_0, x_0) - F(s_0, y_0) + H\Big(\frac{2(x_0 - y_0)}{\varepsilon^2} - 2\varepsilon y_0, y_0\Big) - H\Big(\frac{2(x_0 - y_0)}{\varepsilon^2} - 2\varepsilon x_0, x_0\Big) \\ &\le F(t_0, x_0) - F(s_0, y_0) + C|x_0 - y_0|\Big(\frac{2}{\varepsilon^2}|x_0 - y_0| + 2\varepsilon\big(|x_0| + |y_0|\big)\Big) + 2C\varepsilon\big(|x_0| + |y_0|\big). \end{aligned}$$

Taking $\varepsilon \downarrow 0$, we obtain $\lambda \le 0$, which is a contradiction. $\square$

Index

References

[Eva] Lawrence C. Evans. *Partial Differential Equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI.

[HL] Qing Han and Fanghua Lin. *Elliptic Partial Differential Equations*. Number v. 1 in Courant Lecture Notes. Courant Institute of Mathematical sciences ; American Mathematical Society, 2nd ed edition.

[Lax] Peter D. Lax. *Functional Analysis*. Wiley.

[Zho] Shulin Zhou. *Partial Differential Equations*. 1 edition.