Math 6008 Numerical PDEs–Lecture 1

Preliminaries: basics of PDEs

Instructor: Lei Li, INS, Shanghai Jiao Tong University;
Email: leili2010@sjtu.edu.cn

1 Preliminaries: Introduction to PDEs

A differential equation is an equation that relates an unknown function with its derivatives. If the function is a one-variable function, we say the differential equation is an ordinary differential equation (ODE). If the function is a multi-variable function and its partial derivatives are involved, then the equation is called a partial differential equation (PDE).

- $y''(x) - 2y'(x) = 3x$: ODE
- $y'(t) = e^t y$: ODE
- $u_x + u_{yy} = x^2$: PDE
- $u_x(x,y) - 2u_{xx}(x,y) = y^2 u(x,y)$. Formally, this is a PDE but it is intrinsically an ODE because y can be regarded as a parameter.

Of course, PDEs sometimes can be reduced to ODEs and then be solved exactly. For example, $u_t(t,x) + u_x(t,x) = u(t,x)$. If one view this equation along $x(t) = t + x_0$, then

$$\frac{d}{dt}u(t,x(t)) = u(t,x(t))$$

This is an ODE about $v(t) := u(t,x(t))$.

1.1 Basic concepts

The **order** of a PDE is the largest order of the partial derivatives involved. For example, $u_t(t,x) + u_x(t,x) = 0$ is a first order equation while $u_{xy}(x,y) + u^2 = x^2$ is a second order equation.

Suppose that a PDE about the unknown function $u(x)$ where $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ can be rewritten as

$$L(u, x) = f(x),$$

so that

$$L(c_1 u_1 + c_2 u_2, x) = c_1 L(u_1, x) + c_2 L(u_2, x)$$

then the PDE is called a linear PDE. In particular, it can be written as

$$\sum_{|\alpha| \le k} a_\alpha(x) D^\alpha u = f(x)$$

where D^α represents terms like $\partial_{x_1}^2 \partial_{x_3}$ etc. If $f = 0$, then it is called homogeneous linear PDE.

- $u_{xx} - 2u_{xy} + \frac{1}{2} u_{yy} + 3x^2 u = e^{-x^2 - y^2}$ is linear, nonhomogeneous.
- $uu_{xx} - 2u_{xy} + \frac{1}{2} u_{yy} + 3x^2 u = e^{-x^2 - y^2}$ is nonlinear.

When there is the concept of time, the equation is called a **evolutionary equation** (发展方程) (examples include the hyperbolic and parabolic equations as we see below). In some equations (like the elliptic equations), we do not have the concept of time. Even if you call one variable (like x_1) the "time", you cannot determine the solution for bigger x_1 using the information for smaller x_1–there is no causal effects.

Remark 1. *If one component x_i means the time, we often rewrite it to t. If $x \in \mathbb{R}^2$, we usually rewrite x_1 to x and x_2 to y.*

A k-th order system of PDEs is a system consists of m coupled PDEs and the highest order is k. Formally, we can introduce vector-valued functions $\boldsymbol{u} = (u_1, \ldots, u_m)$ so that the k-th order system is a PDE about the vector-valued function $\boldsymbol{u}$.

To determine the solutions of PDEs, there should be conditions given, called the definite condition (定解条件). For equations without the concept

of time (like elliptic equations), the conditions include the boundary conditions. For evolutionary equations, there could be initial conditions and boundary conditions. The evolutionary equation with initial conditions is called the initial value problem, or the *Cauchy problem*.

With suitable definite conditions, one may solve the PDEs out. There are several notions of solutions, including classical solutions, strong solutions, weak solutions, mild solutions etc. Those who are interested in these concepts may refer to a theoretic PDE book. In our course, we will often talk about classical solutions with enough smoothness. In some cases, the solutions can be solved out analytically. However, often, we are not able to find the expressions for the solutions. However, we can prove the existence and uniqueness with the definite conditions for some cases. Moreover, we may also prove the regularity (smoothness) of the solutions.

1.2 Some typical examples of PDEs

Recall the Laplace operator in Cartesian coordinates

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

Note that $\Delta = \nabla \cdot \nabla$ where

$$\nabla = \sum_i \boldsymbol{e}_i \frac{\partial}{\partial x_i}.$$

The equation

$$-\Delta u := -\sum_{i=1}^{n} \partial_{x_i}^2 u = f(x)$$

is called the **Poisson** equation. If $f = 0$, then it is called the **Laplace** equation. A solution of the Laplace equation is called a **harmonic** function.

The Poisson equation can have variable coefficient

$$-\sum_i \partial_{x_i}(k_i(\boldsymbol{x})\partial_{x_i} u) = f.$$

Below, we mostly only consider $k_i \equiv 1$.

- The Poisson equation is usually used to model stationary problems(定常问题).
- The fundamental solution for general $\mathbb{R}^n$ is

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

- The unknown function inside a bounded domain U can be *stably* determined by the value on the boundary ∂U and f.
- Any change on the boundary value affects all the interior values.

The Poisson equation is a typical example of **elliptic** equations.

The equation

$$u_t + b(t,x)\cdot\nabla u = 0$$

is called the transport equation.

- The equation can be solved by **characteristics** $\dot{x} = b(t,x)$.
- The solution can be determined by the values given at $t = 0$.
- The speed of propagation is finite. Consequently, the solution at some location at later time only depends on the initial data on some interval.
- The energy $E(u) = \int u^2\,dx$ is stable.

The equation

$$u_{tt} = a^2\Delta u + f(x,t)$$

is the **wave equation**. We take the simplest case $u_{tt} = \Delta u$ as the example.

- The solution can be determined by u and u_t given at $t = 0$.

- The speed of propagation is finite. Consequently, the solution at some location at later time only depends on the initial data on some interval.
- The energy $E(u) = \int u_t^2 + |\nabla u|^2 \, dx$ is stable.

Both the transport equations and wave equations are examples of **hyperbolic** equations.

Remark 2. *If we change the sign on the right, we have*

$$u_{tt} = -u_{xx}.$$

We can see that this equation is indeed the Laplace equation. The initial value problem for this equation is not well-posed. In other words, we can not solve u stably given u and u_t at $t = 0$. From here, we see some difference between elliptic equations and hyperbolic equations: for the elliptic equation, the solution should be determined by the boundary values all at once while for hyperbolic equations, we can determine the solutions time after time using the 'initial values'.

The equation

$$\partial_t u = \sum_i \partial_{x_i}(k_i \partial_{x_i} u) + f(x,t)$$

where $k_i(x) > 0$ is called the diffusion equation or the heat equation. Consider the special case

$$u_t = \Delta u.$$

- The heat equation is used to model the *diffusion* effects in nature.
- The solution can be determined by the values at $t = 0$. The fundamental solution is

$$\Phi(x,t) = \frac{1}{(4\pi t)^{n/2}} \exp(-\frac{|x|^2}{4t})$$

- The speed of propagation is infinity. The trajectory of a particle is the Brownian motion and the scales satisfy $[x]^2 \sim [t]$.

- The energy $E(u) = \int |u|^2 \, dx$ is decreasing. There is dissipation.

The heat equation is an example of **parabolic** equations.

2 Classification of second order quasilinear PDEs (free reading)

Consider the second order quasilinear PDEs for $u : \mathbb{R}^n \to \mathbb{R}$:

$$\sum_{i,j} a_{ij}(u, Du, x)\partial^2_{ij} u + F(x, u, Du) = 0.$$

where $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ and $a_{ij} = a_{ji}$.

$A = [a_{ij}]$. Just like quadratic forms, if $det(A) = 0$, and the PDE can't be reduced to one with a lower dimension, then the equation is said to be parabolic; If A is definite, the equation is said to be elliptic; If A is indefinite, the equation is said to be hyperbolic.

Examples

- The equation $Lu = -\sum_{i=1,j=1}^{n} a_{ij}\partial^2_{ij} u + \sum_i b_i \partial_i u + cu = f$ ($\sum_{ij} a_{ij}\xi_i\xi_j \geq \theta|\xi|^2$) is elliptic. Typical example is the Poisson equation $\Delta u = f$.

- Letting $x_n = t$, the equation $u_t + Lu = f$ where $Lu = -\sum_{i=1,j=1}^{n-1} a_{ij}\partial^2_{ij} u$ and $\sum_{i=1,j=1}^{n-1} a_{ij}\xi_i\xi_j \geq \theta|\xi|^2$ is parabolic. Typical example is the heat equation $u_t = a^2\Delta u$.

- The wave equation $u_{tt} - a^2 u_{xx} = 0$ is hyperbolic.

From here on, we only talk about hyperbolic equations where A has only one positive e-value or only one negative e-value.

Some general intuition:

- we cannot talk about initial value problems from elliptic equations because that is not well-posed.Instead, we can talk about boundary value problems.

- For hyperbolic and parabolic equations, we should talk about initial value problems and we have the notion of time. The boundary value problem for these equations are not well-posed. Such types of equations are therefore called **evolutionary** equations, because they are causal: the state at $t = t_2$ (let's say $t = x_n$ without loss of generality) has nothing to do with the states at $t > t_2$.

- Hyperbolic equations are energy stable, while parabolic equations are dissipating.

- These second order hyperbolic equations have characteristics, i.e. the curves for propagation of information. The solutions can be determined locally. The parabolic equations here usually have infinite speed of propagation.

3 Hyperbolic first-order quasi-linear system of equations (free reading)

Now, let $\boldsymbol{u} : \mathbb{R}^n \to \mathbb{R}^m$. Consider the system

$$\boldsymbol{u}_t + \sum_{j=1}^{n} B_j(x, \boldsymbol{u}) \boldsymbol{u}_{x_j} = \boldsymbol{f}$$

where B_j is $m \times m$.

The system is said to be hyperbolic if for any linear combination of B_j, i.e. $B = \sum_j y_j B_j$ has m **real** eigenvalues and is diagonalizable.

Example: the convection equation $u_t + a u_x = 0$ is hyperbolic.

Usually, some hyperbolic second order equations can be reduced to first-order hyperbolic system. For $u_{tt} - a^2 u_{xx} = 0$, if we introduce $v = u_t + a u_x$, then $v_t - a v_x = 0$.

Remark 3. *In the one-dimensional space case, the system $\boldsymbol{u}_t + B(x, \boldsymbol{u})\boldsymbol{u}_x = f$ is said to be elliptic if B has no real eigenvalues.*

An important class of hyperbolic systems is the conservation law $\boldsymbol{u}_t + \boldsymbol{f}(\boldsymbol{u})_x = 0$. $f(u)$ is called the flux function.

Consider the one equation case: u can be understood as the density of the material. $f(u)_x$ is the divergence and it describes how the the material flows. Integrating on x,

$$\frac{d}{dt}\int_{x_1}^{x_2} u dx = f(u)|_{x_2} - f(u)|_{x_1}.$$

If we integrate on the whole axis, we see $\int u dx = const$.

If there is source due to reaction, the equation will be modified to $\boldsymbol{u}_t + \boldsymbol{f}(\boldsymbol{u})_x = \psi(x, \boldsymbol{u})$.

4 Dispersion relation for linear equations with constant coefficients

For linear evolutionary equations

$$u_t = Lu,$$

where L does not depend on x, $e^{i\xi\cdot x}$ is an eigenfunction of L:

$$L(e^{i\xi\cdot x}) = -i\omega(\xi)e^{i\xi\cdot x}.$$

Apply the Fourier transform: $\hat{u} = \int u(x)e^{-i\xi\cdot x}dx$. Then,

$$u(x,t) = \frac{1}{(2\pi)^d}\int \hat{u}(\xi,t)e^{i\xi\cdot x}\, d\xi.$$

It is then not hard to find that the Fourier transform of the equation is

$$\hat{u}_t = -i\omega(\xi)\hat{u}.$$

Then, $\hat{u} = c(\xi)\exp(-i\omega(\xi)t)$ and

$$u = \frac{1}{(2\pi)^d}\int c(\xi)e^{-i\omega(\xi)t}e^{i\xi\cdot x}d\xi.$$

The solution is the superposition of different Fourier modes. The properties of the solutions are determined by $\omega(\xi)$ which is called the dispersion relation. In practice, we simply plug in $e^{i(\xi\cdot x-\omega t)}$ to find the dispersion relation. The phase velocity is given by

$$v_p(\xi) = \frac{\omega(\xi)}{\xi},$$

which is the velocity for the point that has a specific phase $\phi_0 = \xi \cdot x - \omega t$. Another related concept is the group velocity $\omega'(\xi)$, which we omit here.

If ω is real, the amplitude of each mode doesn't decay and different mode has different speed, the equation is said to be dispersive (色散的). The equation shows properties of hyperbolic equations. If ω is not real and imaginary part is negative, the amplitude decays. The equation is said to be dissipative (耗散的).

Example: Consider the heat equation $u_t = a^2 u_{xx}$: $-i\omega = a^2(-\xi^2)$ and $\omega = -ia^2\xi^2$. Hence, for each mode, it evolves like

$$\hat{f}(\omega)\exp(i\xi \cdot x - i\omega t) = \hat{f}(\omega)\exp(i\xi \cdot x - a^2\xi^2 t).$$

Clearly, the mode decays and the energy of the system is consumed. This is why it is *dissipative*.

Example: $u_{tt} = a^2 u_{xx}$: $-\omega^2 = a^2(-\xi^2)$ and $\omega = \pm a\xi$. The mode is like

$$\hat{f}(\omega)\exp(i(\xi \cdot x \pm a\xi t)).$$

Clearly, this is a planar wave.

Example: Schrodinger equation: $iu_t = -u_{xx}$. $\omega = \xi^2$. Schordinger equation is wave-like and shares properties with hyperbolic equations. Different modes have different speeds. Hence, this is dispersive.

- *What if the equation is nonlinear? We can't find the dispersion relation. However, if we have a stationary solution, we can linearize the equation around the stationary solution and then do the linear stability analysis.*

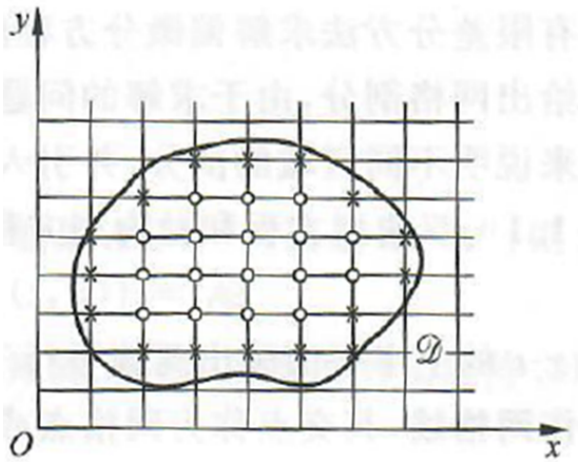

图 1: Illustration of the grid points for a mesh

- *What if L depends on x? $e^{ix\xi}$ is not an eigenfunction, but we can fix x and see the local behavior there.*

5 Finite difference approximations

(See Chapter 2 of the book or Chap. 1 in Leveque.)

Finite Difference Method means that we replace the derivatives in the differential equations with finite difference approximations, given the values at discrete points.

5.1 Mesh generation(网格生成，网格剖分)

Often, we impose a square grid/mesh on the domain, parallel to the axes (including the axis for each spatial dimension and time) as shown in Fig. 5.1. The intersections of the grid lines (网格线) are called the grid points (网格点).

If the neighbors of a grid point all belong to the domain $\overline{D}$, then it is called an interior (grid) point (内部节点), otherwise a grid point belonging

to $\overline{D}$ that is not an interior point is called a boundary (grid) point (边界节点) [In some references, these points are further divided into the so-called "boundary points" and "near-boundary points"].

Take the 2D case as the example. Given a reference point (x_0, y_0) (often $(0,0)$), we may define

$$(x_i, y_j) = (x_0 + ih, y_0 + jh),$$

where h is the spatial step size (空间步长) for the grid/mesh.

5.2 Some basic notations

Consider that h is a step size., which is small. Suppose that we want to approximate $u'(x)$.

The one sided differences (偏心差分):

$$u'(x) \approx \frac{u(x+h) - u(x)}{h} = D_+ u(x),$$
$$u'(x) \approx \frac{u(x) - u(x-h)}{h} = D_- u(x)$$

Centered differences (中心差分):

$$u'(x) \approx \frac{u(x+h) - u(x-h)}{2h} = D_0 u(x) = \frac{1}{2}(D_+ + D_-)u(x).$$

Using these basic operators, we can construct higher derivative. For second order derivatives, we may for example have:

$$u''(x) \approx \frac{u(x+h) - 2u(x) + u(x-h)}{h^2} = D^2 u(x) = D_+ D_- u(x) = D_- D_+ u(x).$$

Another way is to introduce

$$\hat{D}_0 u(x) = \frac{1}{h}(u(x + \frac{h}{2}) - u(x - \frac{h}{2})).$$

It can be checked easily that $D^2 = \hat{D}_0^2$.

For $u'''(x)$, we may for example have

$$u'''(x) \approx D_+ D^2 u(x) = \frac{1}{h^3}(u(x+2h) - 3u(x+h) + 3u(x) - u(x-h))$$

$$u'''(x) \approx D_0 D_+ D_- u(x) = \frac{1}{2h^3}(u(x+2h) - 2u(x+h) + 2u(x-h) - u(x-2h))$$

5.3 Approaches to derive finite difference approximations

There in general two ways to derive: (1). Taylor expansion; (2). Interpolation.

Note that

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h},$$

where h can be either positive or negative. Hence, consider

$$\frac{u(x_{j+1}) - u(x_j)}{h}.$$

Assume that u is smooth (meaning that it is infinitely differentiable). Then, by Taylor expansion, we can have

$$\frac{u(x_{j+1}) - u(x_j)}{h} = u'(x_j) + O(h).$$

Hence, $\frac{u(x_{j+1})-u(x_j)}{h}$ can be used as an approximation to $u'(x_j)$ with an error $O(h)$.

Let us use an example to illustrate how to obtain the finite difference method.

Example: We consider the following problem:

Given $x_1, x_2, x_3,\ x_2 - x_1 = h_1, x_3 - x_2 = h_2$. Use the values of u at these points to approximate $u'(x_2)$.

- Method of undetermined coefficients by Taylor expansion

  Let

  $$au(x_1) + bu(x_2) + cu(x_3) \approx u'(x_2).$$

  $$au(x_1) + bu(x_2) + cu(x_3) = [ae^{-h_1 \frac{d}{dx}} + b + ce^{h_2 \frac{d}{dx}}]u(x_2)$$

  Using the expansion for exponentials (intrinsically, we just do Taylor expansion), we need

  $$a + b + c = 0$$
  $$-h_1 a + c h_2 = 1$$
  $$\frac{1}{2} a h_1^2 + \frac{1}{2} c h_2^2 = 0$$

Solving yields $a = \frac{-h_2}{h_1(h_1+h_2)}, b = \frac{h_2-h_1}{h_1h_2}, c = \frac{h_1}{(h_1+h_2)h_2}$. The error is $O(ah_1^3 + ch_2^3) = O(h^2)$.

- Interpolation. We use polynomial interpolation here. Different methods will yield the same polynomial. Here we choose Lagrange interpolation.

$$p(x) = u(x_1)\frac{(x-x_2)(x-x_3)}{(x_1-x_2)(x_1-x_3)} + u(x_2)\frac{(x-x_1)(x-x_3)}{(x_2-x_1)(x_2-x_3)} + u(x_3)\frac{(x-x_1)(x-x_2)}{(x_3-x_1)(x_3-x_2)}$$

Then,

$$\begin{aligned} u'(x_2) \approx p'(x_2) &= u(x_1)\frac{x_2-x_3}{(x_1-x_2)(x_1-x_3)} \\ &+ u(x_2)\frac{x_2-x_3+x_2-x_1}{(x_2-x_1)(x_2-x_3)} + u(x_3)\frac{x_2-x_1}{(x_3-x_1)(x_3-x_2)} \\ &= u(x_1)\frac{-h_2}{h_1(h_1+h_2)} + u(x_2)\frac{h_2-h_1}{h_1h_2} + u(x_3)\frac{h_1}{(h_1+h_2)h_2} \end{aligned}$$

5.4 Local Truncation error

Suppose that $\mathcal{D}^m$ is some finite difference operator for $u^{(m)}(x)$ at point x, with step sizes $h_1, h_2, \ldots, h_p$. Let $h = \max_i h_i$. The so-called truncation error for a smooth function $u(x)$ is

$$TE = \mathcal{D}^m u - u^{(m)}(x).$$

If for any smooth function u, there exists C_u which is not zero for some u such that

$$TE = C_u h^p + o(h^p),$$

we say the *order of accuracy* is p and $\mathcal{D}^m$ is called a p-th order approximation.

To determine the order, we do Taylor expansion:

$$u(x+h) = \sum_{n=0}^{\infty} \frac{h^n u^{(n)}(x)}{n!} = e^{h\frac{d}{dx}} u(x)$$

By Taylor expansion, it's clear to see that the one-sided difference is first order accuracy:

$$TE = D_+u(x) = \frac{1}{h}[u(x+h)-u(x)] = \frac{1}{h}(hu'(x)+\frac{1}{2}h^2u''(x)+\ldots) = u'(x)+O(h)$$

The centered difference D^2 is of second order accuracy ($TE = O(h^2)$).

Example: Show that

$$D_3u(x) = \frac{1}{6h}[2u(x+h) + 3u(x) - 6u(x-h) + u(x-2h)]$$

gives a third order accuracy approximation.

Sample Code Presentation: Comparison of the three different finite differences.