Leonidas Mindrinos
Otmar Scherzer

Numerical Methods IV
Partial Differential Equations

Lecture Notes
Summer Semester 2017

Computational Science Center
University of Vienna
A-1090 Vienna, Austria

Preface

The lecture notes are mainly based on the books:

1. L. C. Evans, Partial Differential Equations, American Mathematical Society, 2nd edition, (2010).

2. K.W. Morton and D. Mayers, Numerical solution of partial differential equations. An introduction, Cambridge University Press, 2nd edition, (2005).

3. A. Quarteroni and A. Vali, Numerical approximation of partial differential equations, Springer series in Computational Mathematics, Springer-Verlag, (1997).

and on the lecture notes:

1. D. Leitner and O. Scherzer, Numerical Methods for the Solution of Differential Equations, University of Vienna, 2015.

2. O. Scherzer, Computational Mathematics IV, University of Vienna, 2014.

Contents

Chapter 1

Examples of ODEs and PDEs

1.1 Simple Examples of ODEs in Applications

1.1.1 Movement of a Falling Body

We describe the movement of a vertically falling body. Then its position at time t is determined by its height $h(t)$.

Newton's second law of motion implies that the acceleration of the body, that is, the change of its speed, is proportional to the forces acting on the body. In addition, the proportionality constant equals the mass of the body. That is, the equation

$$F = ma$$

holds, where F denotes the forces, m the mass of the body, and a its acceleration.

Now, the acceleration is the change of the speed, which itself is the change of the position of the body. Therefore

$$F = mh''(t) .$$

We still have to model the acting forces.

The main force is the gravity, which, for small heights h, equals approximately mg, where $g \approx 9.81\mathrm{m/s}^2$ is the gravitational acceleration at the earth's surface. The gravitation is acting downwards, and considering the above equation, we get

$$mh''(t) = -mg .$$

If either the body is very light or it is falling fast, it is necessary to take into account air friction as well, which will slow down the fall of the body. One possibility is to model air friction as a force proportional to the square of the body's velocity. Because friction always works against the current movement, the sign of the corresponding force will be opposite to the sign of h'. Thus we obtain the refined model

$$mh''(t) = -c\,\mathrm{sign}(h'(t))\,h'(t)^2 - mg\,,$$

where c is some material constant describing the drag of the body.

In order to obtain a complete description of the movement of the body, we need in addition a description of the state of the body at some initial time t_0, where we begin our considerations. More precisely, we will need its initial position h_0 and its initial velocity v_0. Then, assuming that this model is correct, the movement of the body is completely described by the *differential equation*

$$\begin{aligned} mh''(t) &= -c\,\mathrm{sign}(h'(t))\,h'(t)^2 - mg\,,\\ h(t_0) &= h_0\,,\\ h'(t_0) &= v_0\,. \end{aligned}$$

1.1.2 Population Dynamics

Now consider a simple model that describes the evolution of a population over some time period. That is, we know the population p_0 at some given time t_0, and we want to obtain an estimate $p(t)$ of the population at some future time $t > t_0$.

As a basic model, we assume that the *rate of change* of the population is given by some function $N(t,p)$ that depends only on the time and the size of the population. The time dependence can be used to model external influences on the population, for instance environmental changes, while size of the population influences the number of births and deaths, but can also be used to model overpopulation. Then, the function p that describes the population solves the differential equation

$$p'(t) = N(t,p(t)), \qquad p(t_0) = p_0\,. \tag{1.1}$$

One simple model assumes that the number of births and deaths within a certain amount of time is proportional to the size of the population that is, the birth and death rates are constant

$$N(t,p(t)) = (R-S)p(t),$$

where R and S denote the birth and death rate, respectively. Then (1.1) becomes

$$p'(t) = (R-S)\,p(t)\,.$$

Using the initial state $p(t_0) = p_0$, we obtain with this model the population dynamics

$$p(t) = p_0 e^{(R-S)(t-t_0)}\,.$$

That is, depending on the sign of $R-S$, either the population increases or decreases exponentially.

Now we try to introduce the effects of overpopulation into the model by assuming that the death rate depends on the size of the population. That is, instead of assuming a constant death rate $S > 0$, we assume that S is a function of p. The simplest model is to assume the death rate being proportional to p, setting

$$S(p) = \sigma p$$

for some constant $\sigma > 0$. Then we obtain the equation (the *logistic differential equation*)

$$p'(t) = (R - \sigma p(t))\, p(t)\,. \tag{1.2}$$

In the following, we will compute the analytic solution of this equation.

We note first that the derivative of p is positive if $R > \sigma p$ (and the population p is positive, which we tacitly assume), while it is negative if $R < \sigma p$. In other words, the population increases as long as $p < R/\sigma$, while it decreases for $p > R/\sigma$. In particular, this implies that the long term behavior of the population will be approximately stagnation at the value $p = R/\sigma$.

Separation of Variables

In order to solve the logistic differential equation (1.2), we define

$$\rho := R/\sigma\,,$$

and rewrite the equation as

$$\frac{1}{(\rho - p(t))\,p(t)}\,\frac{dp(t)}{dt} = \sigma\,.$$

Integrating (with an indefinite integral) both sides of this equation with respect to t, we obtain

$$\int_0^T \frac{1}{(\rho - p(t))\,p(t)}\,\frac{dp(t)}{dt}\,dt = \int_0^T \sigma\,dt.$$

Now, we change variables $t \to p := p(t)$, such that (formally)

$$\frac{dp(t)}{dt}dt = dp.$$

Using this identity, we obtain

$$\int_{p(0)}^{p(T)} \frac{1}{(\rho - p)p}\,dp = \sigma T, \quad \text{for all } T \geq 0.$$

Note that

$$\begin{aligned}\int_{p(0)}^{p(T)} \frac{1}{(\rho - p)p}\, dp &= \frac{1}{\rho}\int_{p(0)}^{p(T)} \left(\frac{1}{\rho - p} + \frac{1}{p}\right) dp \\ &= \frac{1}{\rho}\left(\ln\left|\rho - p(0)\right| - \ln\left|\rho - p(T)\right| + \ln|p(T)| - \ln|p(0)|\right) \\ &= \frac{1}{\rho}\ln\left|\frac{p(T)}{\rho - p(T)}\right| - C,\end{aligned}$$

where $C := \frac{1}{\rho}(\ln|p(0)| - \ln\left|\rho - p(0)\right|)$. Thus, the function p satisfies the equation

$$\frac{1}{\rho}\ln\left|\frac{p(T)}{\rho - p(T)}\right| = \sigma T + C, \quad \text{for all } T \geq 0.$$

Multiplying the equation with ρ and taking the exponential, it follows that

$$\left|\frac{p(T)}{\rho - p(T)}\right| = e^{\rho\sigma T}\, e^{\rho C},$$

which is equivalent to

$$\left|\frac{\rho - p(t)}{p(t)}\right| = e^{-\rho\sigma t}\, e^{-\rho C},$$

Now we define a new constant

$$D := \pm e^{-\rho C},$$

such that $D > 0$ if $\rho - p(t) > 0$, and else otherwise. Then, the last equation reads

$$\frac{\rho}{p(t)} - 1 = De^{-\rho\sigma t},$$

which implies that

$$p(t) = \frac{\rho}{1 + De^{-\rho\sigma t}}\,.$$

This is the general form of a solution of the differential equation (1.2). The specific solution satisfying $p(t_0) = t_0$ is obtained by determining D using

$$p_0 = p(t_0) = \frac{\rho}{1 + De^{-\rho\sigma t_0}}\,.$$

The method, we considered here, is called *separation of variables*.

1.2 ODEs with Separable Variables

Definition 1.2.1 *An ODE that can be brought into the form*

$$f(y(t))y'(t) = g(t), \tag{1.3}$$

where the function $f\colon \mathbb{R} \to \mathbb{R}$ *only depends on* y *and not on* t*, and the function* $g\colon \mathbb{R}_{\geq 0} \to \mathbb{R}$ *only depends on* t *and not on* y*, is called ordinary differential equation (of first order) with* separable variables.

The general strategy for solving differential equations of this kind, is to substitute $t \to y := y(t)$. Then, since

$$y' = \frac{dy}{dt},$$

we can *formally* multiply (1.3) with dt and obtain the *formal* equation

$$f(y)\,dy = g(t)\,dt\,.$$

Now we can apply indefinite integrals to both sides and obtain the equation

$$\int f(y)\,dy = \int g(t)\,dt + C\,,$$

where $C \in \mathbb{R}$ is some constant that appears due to the indefinite integration. Note, that the first integration is with respect to y and the right hand side reveals an integration with respect to t.

If it is possible to compute the integrals of f and g analytically, we obtain an equation that the solution necessarily has to satisfy. If, in addition, it is possible to solve this equation for y, we indeed obtain an analytic (general) solution of the differential equation.

Example 1.2.2 *Consider the ODE*

$$(T^2 - t^2)\,y'(t) + ty(t) = 0\,,$$

where $T > 0$ *is some given constant. This equation has separable variables, but in the form above they are not yet separated. In order to bring the equation in the form* (1.3)*, we rewrite the equation as*

$$\frac{y'(t)}{y(t)} = -\frac{t}{T^2 - t^2},$$

which is possible for $y \neq 0$ *and* $t \neq \pm T$*. We rewrite this formally as*

$$\frac{dy}{y} = -\frac{t}{T^2 - t^2}dt\,.$$

Now, integration of both sides of the equation leads to

$$\ln\left|y\right| = \frac{1}{2}\ln\left|T^2 - t^2\right| + C\,.$$

Taking the exponential of the equation, we obtain

$$\left|y\right| = e^C\sqrt{\left|T^2 - t^2\right|}\,.$$

Replacing the constant $e^C > 0$ *by the constant* $D \in \mathbb{R}$ *also encoding the sign of* y*, we get*

$$y(t) = D\sqrt{\left|T^2 - t^2\right|}\,. \tag{1.4}$$

The constant $D \in \mathbb{R}$ *still has to be determined using the initial condition* $y(t_0) = y_0$*. Inserting this condition into the general solution, we see that*

$$y_0 = y(t_0) = D\sqrt{\left|T^2 - t_0^2\right|},$$

and therefore

$$D = \frac{y_0}{\sqrt{\left|T^2 - t_0^2\right|}}\,. \tag{1.5}$$

Note that we have assumed during the computation of the solution of the ODE that $y_0 \neq 0$ *and* $t \neq \pm T$*. It can be easily seen, however, that the derivation above also covers the situation where* $y_0 = 0$ *and* $t_0 \neq \pm T$*. There, the constant function* $y = 0$ *is the unique solution of the ODE, at least until the time reaches one of the values* $\pm T$*.*

The case $t_0 = \pm T$*, however, is different. Then, if* $y_0 = 0$*, for* every *constant* $D \in \mathbb{R}$ *the function* (1.4) *satisfies the ODE and therefore is a solution. If, however,* $y_0 \neq 0$*, then the ODE has no solution at all—then the ODE and the initial conditions are inconsistent.*

Finally, note that all the solutions are valid only locally; that is, there exists at least a time interval $[t_0, t_0 + \epsilon)$ *for some* $\epsilon > 0$ *on which the solution exists and can be written as* (1.4) *with* D *given by* (1.5)*. For general ODEs, this is all that can be said about the solution. In this special case, one can specify the length of the interval on which the solution looks like* (1.4)*: If* $t_0 > T$*, then the formula* (1.4) *is valid on* $[t_0, +\infty)$*. If, however,* $-T < t_0 < T$*, then the solution is*

$$y(t) = \begin{cases} D_1\sqrt{\left|T^2 - t^2\right|} & \text{if } t \in [t_0, T], \\ D_2\sqrt{\left|T^2 - t^2\right|} & \text{if } t \in [T, +\infty), \end{cases} \qquad \text{with } \begin{cases} D_1 = y_0/\sqrt{T^2 - t_0^2}, \\ D_2 \in \mathbb{R} \text{ arbitrary.} \end{cases}$$

In particular, the solution is only unique up to time T*. Similarly, if* $t_0 < -T$*, then*

$$y(t) = \begin{cases} D_1\sqrt{\left|T^2 - t^2\right|} & \text{if } t \in [t_0, -T], \\ D_2\sqrt{\left|T^2 - t^2\right|} & \text{if } t \in [-T, T], \\ D_3\sqrt{\left|T^2 - t^2\right|} & \text{if } t \in [T, +\infty), \end{cases} \qquad \text{with } \begin{cases} D_1 = y_0/\sqrt{T^2 - t_0^2}, \\ D_2 \in \mathbb{R} \text{ arbitrary,} \\ D_3 \in \mathbb{R} \text{ arbitrary.} \end{cases}$$

1.2.1 Homogeneous ODEs

Definition 1.2.3 *An ODE of the form*

$$y' = f\left(\frac{y}{t}\right), \tag{1.6}$$

with $f : \mathbb{R} \to \mathbb{R}$, *is called of* homogeneous type.

If we are have a homogeneous ODE, we can solve it by starting with the substitution

$$z(t) = \frac{y(t)}{t}\,.$$

For the right hand side of (1.6) we are left with the term $f(z)$. For the left hand side of (1.6) we use the product rule and obtain

$$y'(t) = \frac{dy(t)}{dt} = \frac{d(tz(t))}{dt} = z(t) + t\frac{dz(t)}{dt} = z(t) + tz'(t).$$

Thus we have for the variable z the differential equation

$$z(t) + tz'(t) = f(z(t))\,.$$

Now it is easy to see that this ODE is of separable type: We can bring it in the form

$$\frac{z'}{f(z) - z} = \frac{1}{t}\,.$$

This ODE can now be solved by integration as in Section 1.2, and we obtain a solution $z(t)$. At the end, we obtain the solution y by $y(t) = tz(t)$.

Example 1.2.4 *Consider the ODE*

$$y' = \left(\frac{y}{t}\right)^2\,.$$

It is easy to see that this ODE is homogeneous with $f(y/t) = (y/t)^2$. *Using the substitution* $y = tz$ *we obtain*

$$z + tz' = z^2$$

and therefore

$$\frac{z'}{z^2 - z} = \frac{1}{t}\,.$$

Now, we follow Subsection 1.2 and reformulate this equation to

$$\frac{1}{z^2 - z}dz = \frac{1}{t}dt\,.$$

Integrating this equation, we obtain the indefinite integral equation

$$\int \frac{1}{z^2 - z}dz = \int \frac{1}{t}dt + C$$

or by calculating the integrals

$$\ln\left|\frac{z-1}{z}\right| = \ln|t| + C\,.$$

Now, assuming that $t > 0$ and $z \geq 1$, which depends on the initial condition, we get

$$\frac{z-1}{z} = Dt$$

for some constant $D = e^C \in \mathbb{R}^+$ depending on the initial value. Solving for z we obtain

$$z(t) = \frac{1}{1-Dt}$$

(note that z is greater than 1) and, after substitution of $y = tz$

$$y(t) = \frac{t}{1-Dt}\,.$$

For $z = 1$ and $z = 0$, we get the constant solutions $y = t$ and $y = 0$, respectively.

1.3 Linear ODEs

Definition 1.3.1 *An ODE that can be written as*

$$y'(t) + f(t)y(t) = g(t)$$

for some functions $f\colon \mathbb{R} \to \mathbb{R}$ and $g\colon \mathbb{R} \to \mathbb{R}$ is called linear ODE of first order.

Here, first order means that the highest derivative of the unknown function y that appears in the equation is one. Linear means that all the expressions are linear in the unknown y and its derivatives.

As in the case of linear algebraic equations, the linearity of an equation has some implications on the structure of its solutions. To that end we consider the *homogeneous* equation[1]

$$y'(t) + f(t)y(t) = 0\,.$$

If we are given two solutions y_1 and y_2 of this equation (with possibly different initial conditions), then

$$y_1'(t) + f(t)y_1(t) = 0\,,$$
$$y_2'(t) + f(t)y_2(t) = 0\,.$$

[1] Homogeneous means that the right hand side of the equation is zero, that is, $g = 0$

Consequently also

$$\frac{d}{dt}(y_1 + y_2) + f(t)(y_1 + y_2) = y_1' + f(t)y_1 + y_2' + f(t)y_2 = 0\,,$$

which shows that also $y_1 + y_2$ is a solution of the ODE. More general, if y_1 and y_2 solve the ODE and $c_1, c_2 \in \mathbb{R}$, then the linear combination $c_1 y_1 + c_2 y_2$ is also a solution.

In order to solve the (inhomogeneous) equation

$$y'(t) + f(t)y(t) = g(t) \tag{1.7}$$

we can proceed in two different (but similar) ways.

1.3.1 Integrating factor

We first observe that (1.7) is equivalent to

$$h(t)y'(t) + h(t)f(t)y(t) = h(t)g(t), \tag{1.8}$$

at least, if $h\colon \mathbb{R} \to \mathbb{R}$ is a function that is different from zero.

Now the idea is to choose the function h in such a way that the left hand side of (1.8) is itself a derivative of a different function. More precisely, we try to find $h\colon \mathbb{R} \to \mathbb{R}$ such that

$$h(t)y'(t) + h(t)f(t)y(t) = \frac{d}{dt}(hy) = h'y + hy'\,. \tag{1.9}$$

If (1.9) holds, then the equation (1.8) reads as follows,

$$\frac{d}{dt}(hy) = h(t)g(t),$$

which after integration becomes:

$$h(t)y(t) = \int^t g(s)h(s)\,ds + C\,. \tag{1.10}$$

For this reason, a function h satisfying (1.9) is called an *integrating factor* for the ODE (1.7).

Thus, (1.9) is satisfied, if h satisfies

$$h(t)f(t)y(t) = h'(t)y(t)\,.$$

Dividing this equation by y, we see that h has to satisfy the ODE

$$h' = f(t)h\,.$$

This ODE can be solved by separation of the variables, and we obtain the integrating factor

$$h(t) = D\exp\left(\int^t f(s)\,ds\right).$$

Inserting this integrating factor in (1.10), we obtain

$$y(t) = \frac{\int^t \left[g(s)D\exp\left(\int^s f(r)\,dr\right)\right]ds + C}{D\exp\left(\int^t f(s)\,ds\right)},$$

or, setting $\tilde{C} := C/D$,

$$y(t) = \frac{\int^t \left[g(s)\exp\left(\int^s f(r)\,dr\right)\right]ds + \tilde{C}}{\exp\left(\int^t f(s)\,ds\right)}. \tag{1.11}$$

1.3.2 Linear transformation

To obtain a solution of the inhomogeneous equation (1.7), we consider a linear transformation of the form

$$y(t) = h(t)Y(t),$$

for a non-zero and non-constant function h. Taking the derivative, we get

$$y'(t) = h'(t)Y(t) + h(t)Y'(t),$$

which together with (1.7) gives

$$h(t)Y'(t) + (h'(t) + f(t)h(t))Y(t) = g(t). \tag{1.12}$$

We choose h such that $h' + fh = 0$, meaning that h can be written as

$$h(t) = C_1\exp\left(-\int^t f(s)ds\right).$$

Then, equation (1.12) takes the form

$$h(t)Y'(t) = g(t) \quad\Rightarrow\quad Y'(t) = \frac{g(t)}{h(t)}$$

We integrate and we consider the special form of h to get

$$Y(t) = \int^t g(s)C_1^{-1}\exp\left(\int^s f(r)dr\right)ds + C_2.$$

Then, the general solution of (1.7) is given by

$$y(t) = h(t)Y(t) = C_1\exp\left(-\int^t f(s)ds\right)\left[\int^t g(s)C_1^{-1}\exp\left(\int^s f(r)dr\right)ds + C_2\right],$$

which is (1.11) for $\tilde{C} = C_1C_2$.

Example 1.3.2 *Solve the linear ODE*

$$y'(t) + y(t) = 2\cos(t), \quad y(0) = 1.$$

Example 1.3.3 *Solve the initial value problem*

$$y'(t) + 2y(t) = g(t), \quad y(0) = 0,$$

where

$$g(t) = \begin{cases} 1, & \text{for } t \in [0,1] \\ 0, & \text{for } t > 1 \end{cases}$$

Let us consider an initial value problem (IVP) of the form

$$\begin{aligned} y'(t) + f(t)y(t) = g(t), \quad t \in I, \\ y(t_0) = y_0, \quad t_0 \in I. \end{aligned} \tag{1.13}$$

We set

$$F(t) = \int^t f(s)ds$$

and we multiple (1.13) with $\exp(F)$ to obtain

$$(e^{F(t)}y(t))' = g(t)e^{F(t)}.$$

We integrate from t_0 to t

$$\int_{t_0}^t (e^{F(s)}y(s))'ds = \int_{t_0}^t g(s)e^{F(s)}ds$$

resulting to the solution

$$y(t) = exp(-\tilde{F}(t))\left[y_0 + \int_{t_0}^t g(s)e^{\tilde{F}(s)}ds\right],$$

where

$$\tilde{F}(t) = F(t) - F(t_0) = \int_{t_0}^t f(s)ds.$$

1.4 Non-linear ODEs of the first order

We consider differential equations of the form

$$y'(t) = f(t, y(t)), \quad y(t_0) = y_0.$$

The main question is when this problem has solution, what is its domain of definition and if it is unique. In the linear case, the answer is easy, provided that the functions f and g in (1.7) are continuous in an open set containing t_0.

Here, we don't have a form for the general solution and the domain may has nothing to do with the function f.

For example, the IVP

$$y'(t) = y^2(t), \quad y(0) = 1$$

admits the solution (using separable variables)

$$y(t) = \frac{1}{1-t},$$

which obviously is not defined for $t = 1$, something that is not clear from the initial equation. Remark here that for different initial condition, we get a different singular point.

Example 1.4.1 *Solve the non-linear ODE*

$$t^2 y'(t) = (2t + y(t))y(t).$$

Hint: set $z = y/t$ for $t \neq 0$.

1.5 Examples of Partial Differential Equations

Definition 1.5.1 *A partial differential equation (PDE) is an equation for more than two different derivatives of a function $u(x_1, x_2, \ldots, x_n)$ on a domain $\Omega \subseteq \mathbb{R}^n$.*

Example 1.5.2 *For instance*

$$y\frac{\partial^2 u}{\partial x^2} + \frac{\partial u}{\partial y} = x^2 yu \text{ or } yu_{xx} + u_y = x^2 yu\,, \tag{1.14}$$

which actually means

$$y\frac{\partial^2 u}{\partial x^2}(x,y) + \frac{\partial u}{\partial y}(x,y) = x^2 yu(x,y), \quad \text{for } (x,y) \in \Omega\,. \tag{1.15}$$

This is a PDE for a function $u(x,y)$ in two variables.

We use the notation

$$u_x = \frac{\partial u}{\partial x}, \quad u_y = \frac{\partial u}{\partial y},$$

and

$$u_{xx} = \frac{\partial^2 u}{\partial x^2} = \frac{\partial \frac{\partial u}{\partial x}}{\partial x}, \quad u_{xy} = \frac{\partial^2 u}{\partial x \partial y} = \frac{\partial \frac{\partial u}{\partial x}}{\partial y} = \frac{\partial \frac{\partial u}{\partial y}}{\partial x} = u_{yx}\,.$$

The variables x, y are called *independent* variables and u is called *dependent* variable.

The *order* of the differential equation is the order of the highest derivative of the dependent variables in the differential equation. The PDE (1.15) is a differential equation of second order. The differential equation

$$xu_x u_{xxy} + u_x^4 = 0$$

is of third order.

Most PDEs with relevance in practice are of first or second order.

Example 1.5.3

1. *The electrostatic potential* $u(x,y,z)$ *which is determined by a charge density* $\rho(x,y,z)$ *satisfies the* Poisson equation

$$\Delta u = u_{xx} + u_{yy} + u_{zz} = 4\pi\rho \, .$$

Δ *denotes the* Laplace *operator (in space dimension three).*

2. *The* wave equation *is the PDE*

$$\frac{1}{c^2} u_{tt} = \Delta u \, .$$

In air, $u(x,y,z,t)$ *denotes the density of air at a location* (x,y,z) *at time* t *and* c *denotes the sound speed.*

3. Heat *or* diffusion *equation:*

$$u_t = \alpha \Delta u \, ,$$

with some $\alpha > 0$.

4. *The velocity* $\vec{v}(\vec{x},t) = (v_1, v_2, v_3)(\vec{x},t)$ *and the pressure* $p(\vec{x},t)$ *of an incompressible fluid as a function of space* $\vec{x} = (x,y,z)$ *and time* t *satisfies the* Navier-Stokes-equations

$$\vec{v}_t + (\vec{v} \cdot \nabla)\vec{v} = \nu \Delta \vec{v} + \nabla p \, , \quad \nabla \cdot \vec{v} = 0 \, ,$$

for some constant ν. *In the above equation*

$$\nabla \cdot \vec{v} := (v_1)_x + (v_2)_y + (v_3)_z$$

denotes the divergence *and*

$$(\vec{v} \cdot \nabla)\vec{v} = \begin{pmatrix} v_1(v_1)_x + v_2(v_1)_y + v_3(v_1)_z \\ v_1(v_2)_x + v_2(v_2)_y + v_3(v_2)_z \\ v_1(v_3)_x + v_2(v_3)_y + v_3(v_3)_z \end{pmatrix} .$$

This is a system of four equations with four unknowns.

Definition 1.5.4 *A PDE is called* linear *if* u *and its derivative only appear* linearly. *More precisely a linear PDE has the form*

$$Lu = b\,,$$

where L *is a differential operator and* b *is a given function.*

Example 1.5.5 *Equation* (1.14) *is linear with*

$$L = y\frac{\partial^2}{\partial x^2} + \frac{\partial}{\partial y} - x^2 y\,, \qquad b = 0\,.$$

The Poisson equation is linear with $L = \Delta$ *and* $b = 4\pi\rho$*. The wave equation and the heat equation are linear, respectively. The Navier-Stokes equation is nonlinear. Another, frequently used nonlinear PDE is the* Burger's equation

$$u_x + u u_y = 0\,.$$

1.5.1 Conservation Principles

Differential equations are frequently derived from conservation of physical quantities like mass, energy, temperature and so on. To illustrate this we consider the temperature distribution in a homogeneous, non insulating slab of length L. We denote now by $u(x,t)$ the temperature in a point $x \in [0,L]$ at time $t \geq 0$.

We are modeling the following principles:

1. Conservation of energy: The time variation of thermal energy in every interval $[a,b] \subseteq [0,L]$ is equal to the heat flux across a and b.

2. The energy density (energy per length's unit) is $\rho c u$. Here ρ denotes the density, c is the specific heat. Both ρ and c are assumed to be constant here.

3. Fourier's law: The heat flux is proportional to the gradient of the temperature (heat is fluctuating from warm to cold), and the proportionality constant $k > 0$ is called heat conductivity. That is, we have

$$\frac{d}{dt}\int_a^b \rho c u(x,t)\,dx = -k\frac{\partial u}{\partial x}(a,t) + k\frac{\partial u}{\partial x}(b,t)\,. \tag{1.16}$$

The left hand side is the variation of the total energy in the slab. The right hand side is the energy, which migrates in and out of the slab per time unit. Equation (1.16) should hold for all intervals $[a,b]$. Thus by the fundamental theorem of integration we get from (1.16)

$$\int_a^b \left(\rho c\frac{\partial u}{\partial t}(x,t)\,dx - k\frac{\partial^2 u}{\partial x^2}(x,t)\right) dx = 0\,.$$

Because this holds for arbitrary intervals $[a, a+\varepsilon]$ we see that

$$0 = \frac{1}{\varepsilon}\int_a^{a+\varepsilon}\left(\rho c\frac{\partial u}{\partial t}(x,t)\,dx - k\frac{\partial^2 u}{\partial x^2}(x,t)\right)dx$$
$$\sim \left(\rho c\frac{\partial u}{\partial t}(a,t) - k\frac{\partial^2 u}{\partial x^2}(a,t)\right).$$

This should hold for all a, which gives that

$$0 = \frac{\partial u}{\partial t}(a,t) - \underbrace{\frac{k}{\rho c}}_{:=\alpha}\frac{\partial^2 u}{\partial x^2}(a,t).$$

This is the heat equation in $\mathbb{R}\times\mathbb{R}_0^+$.

In $\mathbb{R}^3$, the derivation from Fourier's law is similar and is based on the divergence theorem

$$\frac{d}{dt}\int_\Omega \rho c u\,dV = \int_{\partial\Omega} k\nabla u\cdot\vec{\eta}\,dS = k\int_\Omega \Delta u\,dV$$

resulting to

$$0 = \frac{\partial u}{\partial t}(a,t) - \underbrace{\frac{k}{\rho c}}_{:=\alpha}\Delta u(a,t),$$

for all $a\in\mathbb{R}^3$. Here Ω is an arbitrary closed subregion of the material slab and

$$\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}$$

denotes the Laplace operator.

Chapter 2

Numerical Solution of Partial Differential Equations as Systems of ODEs

We explain with an example how we can reformulate a partial differential equation as a system of ordinary differential equations. In this way we can apply all numerical methods for solving ordinary equations in order to solve this partial differential equation, such as Euler methods, Runge-Kutta methods and Adams-Bashforth method.

Example 2.0.6 *Let $u(x,t)$, $-1 \leq x \leq 1$, be the temperature distribution at time t in a slab of length $l = 2$. Assuming constant conductivity $\sigma = 1$, u satisfies the* heat conduction *equation:*

$$u_t = \sigma u_{xx} = u_{xx}\,, \quad -1 < x < 1\,, \quad 0 < t < T\,. \tag{2.1}$$

This is now a *partial differential equation* because it depends on derivatives of two variables x, t. By discretization of the x variable we can transform the partial differential equation in a system of ordinary differential equations.

Let $v : [-1,1] \to \mathbb{R}$ be an arbitrary function satisfying $v(-1) = v(1) = 0$, then we get by integration by parts

$$\int_{-1}^{1} u_t(t,x)v(x)\,dx = \int_{-1}^{1} u_{xx}(t,x)v(x)\,dx = -\int_{-1}^{1} u_x(t,x)v_x(x)\,dx\,. \tag{2.2}$$

We assume that the temperatures $u(-1,t) := u_0(t)$ and $u(1,t) := u_1(t)$ are measured. Then, for every $t > 0$, $u(t,x)$ can be approximated by a linear spline in space over the grid $\Delta = \{-1 = x_0 < x_1 < \ldots < x_n = 1\}$, that is

$$u(x,t) = \sum_{i=0}^{n} y_i(t)\Lambda_i(x)\,, \tag{2.3}$$

where Λ_i is a linear hat function with peak at x_i. Taking into account the boundary conditions we see that $y_0 = u_0(t)$ and $y_n = u_1(t)$. All other functions y_i are unknowns.

Inserting (2.3) in (2.2) we get a system of differential equations for the unknowns $y_1, ..., y_{n-1}$:

$$\sum_{i=0}^{n} y_i'(t) \int_{-1}^{1} \Lambda_i(x) v(x)\, dx = -\sum_{i=0}^{n} y_i(t) \int_{-1}^{1} \Lambda_i'(x) v_x(x)\, dx\,,$$

where we choose $v(x) \in \{\Lambda_j(x) : j = 1, ..., n-1\}$, meaning that v is a hat function, which satisfies homogeneous boundary conditions.

We denote by

$$G := [\langle \Lambda_i, \Lambda_j \rangle]_{1 \le i,j \le n-1} = \frac{h}{6} \begin{pmatrix} 4 & 1 & 0 & \cdots & \cdots & 0 \\ 1 & 4 & 1 & \ddots & & \vdots \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & 1 & 4 & 1 \\ 0 & \cdots & \cdots & 0 & 1 & 4 \end{pmatrix}$$

the element mass matrix and by

$$A := [\langle \Lambda_i', \Lambda_j' \rangle]_{1 \le i,j \le n-1} = h \begin{pmatrix} 2 & -1 & 0 & \cdots & \cdots & 0 \\ -1 & 2 & -1 & \ddots & & \vdots \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & -1 & 2 & -1 \\ 0 & \cdots & \cdots & 0 & -1 & 2 \end{pmatrix}$$

the element stiffness matrix. Then, we get a compact description of the system

$$Gy'(t) + Ay(t) = b(t), \tag{2.4}$$

where b is an appropriate vector, which depends on u_0 and u_1. To completely specify the system (2.4) we need also initial values for $y_1, ..., y_{n-1}$, which are typically determined from interpolation of the initial temperature $u(0, x)$.

The entries in the matrices G and A can be easily computed by specifying the form of the hat function Λ_i. Let

$$\Lambda_i(x) = \begin{cases} 0, & x < x_{i-1} \\ \frac{x - x_{i-1}}{h}, & x \in [x_{i-1}, x_i] \\ \frac{x_{i+1} - x}{h}, & x \in [x_i, x_{i+1}] \\ 0, & x > x_{i+1} \end{cases}$$

then, we compute for instance

$$G_{i,i-1} = \int_{-1}^{1} \Lambda_i(x)\Lambda_{i-1}(x)\,dx = \int_{x_{i-1}}^{x_i} \left(\frac{x - x_{i-1}}{h}\right)\left(\frac{x_i - x}{h}\right) dx = \frac{h}{6}$$

and

$$A_{i,i-1} = \int_{-1}^{1} \Lambda_i'(x)\Lambda_{i-1}'(x)\,dx = -\frac{1}{h}.$$

Chapter 3

Linear Partial Differential Equations

3.1 Classification of linear PDEs

To keep the presentation simple, we restrict ourselves to linear partial differential equations of second order with two variables. Such an equation for a function $u = u(x,y)$ reads as follows:

$$Au_{xx} + 2Bu_{xy} + Cu_{yy} + Du_x + Eu_y + Fu + G = 0 \, . \tag{3.1}$$

Here $A = A(x,y), \ldots, G = G(x,y)$ are functions.

Definition 3.1.1 *A PDE of the form* (3.1) *is called*

- elliptic *if* $AC - B^2 > 0$ *,*
- parabolic *if* $AC - B^2 = 0$ *, and*
- hyperbolic *if* $AC - B^2 < 0$.

Example 3.1.2

a) The wave equation

$$\frac{1}{c^2} u_{xx} - u_{yy} = 0$$

(note we changed the notation from t to y) is of the from (3.1) *with*

$$A = c^{-2}, \quad B = 0, \quad C = -1 \, .$$

Since $AC - B^2 = -c^{-2} < 0$, the equation is hyperbolic.

b) The Laplace equation

$$u_{xx} + u_{yy} = 0$$

is of the from (3.1) *with*

$$A = 1, \quad B = 0, \quad C = 1 \, .$$

Since $AC - B^2 = 1 > 0$, *the equation is elliptic.*

c) *The heat equation*

$$u_x - u_{yy} = 0$$

is of the from (3.1) *with*

$$A = 0, \quad B = 0, \quad C = -1 \, .$$

Here, $AC - B^2 = 0$. *Then, this equation is parabolic.*

Remark 3.1.3

1. *For the classification, we are interested only on the* main symbol
$$Au_{xx} + 2Bu_{xy} + Cu_{yy}.$$
These are the terms of the differential equation of highest order (in our case this is 2).

2. *Recall that* $AC - B^2$ *is the determinant of the symmetric matrix*
$$M = \begin{pmatrix} A & B \\ B & C \end{pmatrix} .$$
Denoting by $m_{11} = A$, $m_{12} = m_{21} = B$, $m_{22} = C$ *and* $x_1 = x$, $x_2 = y$ *the main symbol reads as follows*
$$\sum_{i,j=1}^{2} m_{ij} \frac{\partial^2 u}{\partial x_i \partial x_j} \, .$$

3. *If the coefficients* A, B, C *are not constant, but functions which depend on* x *and* y *in a non-trivial manner, then the type of the partial differential equation can be different for various points* (x, y).

   For instance the differential equation
$$xu_{xx} + u_{yy} = 0$$
is elliptic for $x > 0$, *parabolic for* $x = 0$, *and hyperbolic for* $x < 0$.

4. *The terminology elliptic, parabolic and hyperbolic is motivated from conic sections. A curve* $(X, Y(X))$, *which satisfies the equation*
$$AX^2 + 2BXY + CY^2 + DX + EY + F = 0 \qquad \text{(with constant coefficients)}$$
is either an ellipsis, parabola, or an hyperbola, depending on the sign of $AC - B^2$. *For instance*

- *For $A = C = 1$, $B = D = E = 0$ and $F = -R^2$ we get*

$$X^2 + Y^2 = R^2, \quad \textit{circle}$$

- *For $A = \frac{1}{a^2}$, $C = \frac{1}{b^2}$, $B = D = E = 0$ and $F = -1$ we get*

$$\frac{X^2}{a^2} + \frac{Y^2}{b^2} = 1, \quad \textit{ellipse}$$

- *For $A = \frac{1}{a^2}$, $C = -\frac{1}{b^2}$, $B = D = E = 0$ and $F = -1$ we get*

$$\frac{X^2}{a^2} - \frac{Y^2}{b^2} = 1, \quad \textit{hyperbola}$$

- *For $A = 1$, $B = C = D = F = 0$, and $E = -1$ we get*

$$Y = X^2, \quad \textit{parabola}$$

5. *A different terminology is used in linear algebra, where* quadratic forms *are investigated:*

$$Q(X, Y) = AX^2 + 2BXY + CY^2 \,.$$

We assume that one of the coefficients A, B, C is not identical zero.

- *Q is called positive (negative) definite, if $Q(X, Y) > 0$ ($Q(X, Y) < 0$) for all $(X, Y) \neq (0, 0)$.*
- *Q is called positive (negative) semi-definite, if $Q(X, Y) \geq 0$ ($Q(X, Y) \leq 0$) for all $(X, Y) \neq (0, 0)$.*
- *If Q is not semidefinite, then it is called* indefinite.

Theorem 3.1.4 *a) Q is positive or negative (semi-) definite if and only if $AC - B^2 > 0\,(\geq 0)$,*

b) Q is indefinite iff $AC - B^2 < 0$,

c) If $AC - B^2 = 0$, then Q is semi-definite.

Proof 3.1.5 *We first see that for $A \neq 0$*

$$\begin{aligned} Q(X, Y) &= AX^2 + 2BXY + CY^2 \\ &= A\left(X^2 + 2\frac{B}{A}XY + \frac{B^2}{A^2}Y^2\right) + \left(C - \frac{B^2}{A^2}\right)Y^2 \\ &= \frac{(AX + BY)^2}{A} + \frac{AC - B^2}{A}Y^2. \end{aligned} \tag{3.2}$$

a) ($\Rightarrow$) Let Q be positive definite (the other cases follow analogously). Then,

$$A = Q(1, 0) > 0, \quad C = Q(0, 1) > 0. \tag{3.3}$$

Now, we have two cases:

1. *If* $B = 0$, *then clearly* $AC > 0$.
2. *If* $B \neq 0$, *then*

$$0 < Q\left(1, -\frac{A}{B}\right) = (AC - B^2)\frac{A}{B^2}.$$

From (3.3), *follows that* $AC - B^2 > 0$.

($\Leftarrow$) *Let* $AC - B^2 > 0$. *Then also* $AC > 0$. *From* (3.2) *we get*

$$sign(A)Q(X,Y) \geq (AC - B^2)Y^2 > 0, \quad X \in \mathbb{R},\ Y \neq 0.$$

Thus, Q is either positive or negative definite, depending on the sign of A.

If $Y = 0$, *we have*

$$sign(A)Q(X,Y) = |A|X^2 > 0, \quad X \neq 0.$$

b) follows immediately from a), since Q is semi-definite is equivalent to $AC - B^2 \geq 0$. *Then, Q is indefinite if* $AC - B^2 < 0$.

c) Let $AC - B^2 = 0$. *We have two cases:*

1. *If* $B = 0$, *then* $A = 0$ *or* $C = 0$. *Without loss of generality, we assume that* $C = 0$. *Then,*

$$Q(X,Y) = AX^2 \begin{cases} \geq 0, & \text{if } A \geq 0 \\ \leq 0, & \text{if } A \leq 0. \end{cases}$$

That means that Q is semi-definite (positive or negative) depending on the sign of A.

2. *If* $B \neq 0$, *then also A and C are non-zero with the same sign. Then*

$$\begin{aligned} Q(X,Y) &= AX^2 + 2BXY + CY^2 \\ &= sign(A)\left(|A|X^2 + 2\sqrt{AC}XY + |C|Y^2\right) \\ &= sign(A)\left(\sqrt{|A|}X + \sqrt{|C|}Y\right)^2. \end{aligned}$$

That means that Q is semi-definite (positive or negative) depending on the sign of A.

Applying change of coordinates every quadratic form Q can be transformed to

- $X^2 + Y^2$ *if* $AC - B^2 > 0$,
- $X^2 - Y^2$ *if* $AC - B^2 < 0$,
- X^2 *if* $AC - B^2 = 0$.

For instance, let $A = C = 1$ and $B = 2$. Then,

$$\begin{aligned} Q(X,Y) &= X^2 + 4XY + Y^2 \\ &= (X+2Y)^2 - 4Y^2 + Y^2 \\ &= (X+2Y)^2 - 3Y^2 \\ &= (X')^2 - (Y')^2, \end{aligned}$$

where

$$X' = X + 2Y \text{ and } Y' = \sqrt{3}Y\ .$$

An important property of the classification, Definition 3.1.1, is that it is invariant under coordinate transformations. This means, that a coordinate transformation does not change the type of the differential equation.

Recall the main symbol

$$Au_{xx} + 2Bu_{xy} + Cu_{yy}. \tag{3.4}$$

We consider the change of coordinates:

$$\xi = \xi(x,y), \quad \eta = \eta(x,y),$$

with

$$J = \frac{\partial(\xi,\eta)}{\partial(x,y)} = \begin{pmatrix} \xi_x & \xi_y \\ \eta_x & \eta_y \end{pmatrix}.$$

We apply the chain rule

$$\begin{aligned} u_x &= u_\xi \xi_x + u_\eta \eta_x \\ u_y &= u_\xi \xi_y + u_\eta \eta_y \\ u_{xx} &= u_{\xi\xi}\xi_x^2 + 2u_{\xi\eta}\xi_x\eta_x + u_{\eta\eta}\eta_x^2 + \text{l.o.t.} \\ u_{xy} &= u_{\xi\xi}\xi_x\xi_y + u_{\xi\eta}(\xi_x\eta_y + \xi_y\eta_x) + u_{\eta\eta}\eta_x\eta_y + \text{l.o.t.} \\ u_{yy} &= u_{\xi\xi}\xi_y^2 + 2u_{\xi\eta}\xi_y\eta_y + u_{\eta\eta}\eta_y^2 + \text{l.o.t.} \end{aligned}$$

Then, the main symbol (3.4) takes the form

$$A'u_{\xi\xi} + 2B'u_{\xi\eta} + C'u_{\eta\eta},$$

with

$$\begin{aligned} A' &= A\xi_x^2 + 2B\xi_x\xi_y + C\xi_y^2 \\ B' &= 2A\xi_x\eta_x + 2B(\xi_x\eta_y + \eta_x\xi_y) + 2C\xi_y\eta_y \\ C' &= A\eta_x^2 + 2B\eta_x\eta_y + C\eta_y^2 \end{aligned}$$

To compute $A'C' - B'^2$ we consider the above equation in matrix form

$$\begin{pmatrix} A' & B' \\ B' & C' \end{pmatrix} = J \begin{pmatrix} A & B \\ B & C \end{pmatrix} J^\top$$

Computing the determinants, we have

$$A'C' - B'^2 = \det(J)^2(AC - B^2)$$

and thus the signs coincide. This shows that the type of a PDE does not change if we apply linear transformations.

3.2 IVP for the wave equation

One of the few cases, where all the solution of a PDE exist explicitly, is the wave equation. We consider the one-dimensional wave equation

$$u_{tt} = c^2 u_{xx}, \quad t \geq 0 \tag{3.5}$$

together with some initial conditions to be specified later. We present the d' Alembert solution of the above equation. The idea is to transform (3.5) into a simpler form that can be solved by a simple integration. To do so, we transform the independent variables. We change variables

$$\xi = x - ct, \quad \eta = x + ct.$$

We use the chain rule to compute

$$\begin{aligned} u_{xx} &= u_{\xi\xi} + 2u_{\xi\eta} + u_{\eta\eta} \\ u_{tt} &= c^2 u_{\xi\xi} - 2c^2 u_{\xi\eta} + c^2 u_{\eta\eta}. \end{aligned}$$

Then, (3.5) is transformed to ($c \neq 0$)

$$u_{\xi\eta} = 0.$$

It is easy to find the general solution of this equation by integrating twice. First we integrate with respect to η

$$u_\xi = C(\xi)$$

where the constant depends on ξ. Now we integrate with respect to ξ to get

$$u(\xi, \eta) = \int^{\xi} C(s)ds + G(\eta) = F(\xi) + G(\eta)$$

Then all solutions admit the general form

$$u(x,t) = F(x - ct) + G(x + ct), \tag{3.6}$$

for given $F, G \in C^2$. A solution of the wave equation is a superposition of two waves traveling at opposite directions. For example, the function $F(x - ct)$ describes one wave traveling to the right.

A typical initial values problem for the wave equation has the form

$$\begin{aligned} u_{tt}(x,t) &= c^2 u_{xx}(x,t), \quad t \geq 0 \\ u(x,0) &= f(x), \\ u_t(x,0) &= g(x). \end{aligned} \tag{3.7}$$

In order to find the solution of the IVP (3.7) for given f and g we consider (3.6) and the initial conditions

$$F(x) + G(x) = f(x), \quad c(-F'(x) + G'(x)) = g(x),$$

using that $\frac{\partial}{\partial t} F(x-ct) = F'(x-ct)(-c)$. We take the derivative of the first equation and we solve the system of equations for F' and G' to obtain

$$F'(x) = \frac{1}{2} f'(x) - \frac{1}{2c} g(x), \quad G'(x) = \frac{1}{2} f'(x) + \frac{1}{2c} g(x).$$

We integrate both equations

$$\begin{aligned} F(x) &= \frac{1}{2} f(x) - \frac{1}{2c} \int_0^x g(y)dy + C \\ G(x) &= \frac{1}{2} f(x) + \frac{1}{2c} \int_0^x g(y)dy - C, \end{aligned} \tag{3.8}$$

where the constant C is such that the initial condition $F(x) + G(x) = f(x)$ is satisfied. Combining (3.8) with (3.6) we have that the solution of the IVP (3.7) for the one-dimensional wave equation is given by

$$u(x,t) = \frac{1}{2}(f(x-ct) + g(x+ct)) + \frac{1}{2c} \int_{x-ct}^{x+ct} g(y)dy.$$

3.3 Characteristics

The following considerations make evident the importance of the classification of differential equations.

Firstly, we consider an ordinary differential equation:

$$u''(x) = F(x, u(x), u'(x)) \, .$$

Suppose that $u(x_0)$ and $u'(x_0)$ are known at a point x_0, then using Taylor series expansion

$$\begin{aligned} u(x_0 + \Delta x) &\approx u(x_0) + u'(x_0)\Delta x \, , \\ u'(x_0 + \Delta x) &\approx u'(x_0) + u''(x_0)\Delta x \\ &= u'(x_0) + F(x_0, u(x_0), u'(x_0))\Delta x \, . \end{aligned}$$

The above equations get exact as $\Delta x \to 0$. That is, the solution of the differential equation can be determined approximately from the values at x_0. In the following we apply this idea to partial differential equations: We use the linearization:

$$u(x_0 + \Delta x, y_0 + \Delta y) \approx u(x_0, y_0) + u_x(x_0, y_0)\Delta x + u_y(x_0, y_0)\Delta y \,.$$

But we need also approximations for higher order derivatives which can be derived by approximating u_x:

$$\begin{aligned} u_x(x_0 + \Delta x, y_0 + \Delta y) &\approx u_x(x_0, y_0) + u_{xx}(x_0, y_0)\Delta x + u_{xy}(x_0, y_0)\Delta y \,, \\ u_y(x_0 + \Delta x, y_0 + \Delta y) &\approx u_y(x_0, y_0) + u_{yx}(x_0, y_0)\Delta x + u_{yy}(x_0, y_0)\Delta y \,. \end{aligned}$$

This means that if you know the function value u at (x_0, y_0) and derivatives of up to second order at (x_0, y_0), then one knows also u and it's first derivatives in a neighborhood, that is at $(x_0 + \Delta x, y_0 + \Delta y)$.

This idea is generalized now to the PDE (3.1)

$$Au_{xx} + 2Bu_{xy} + Cu_{yy} = -Du_x - Eu_y - Fu - G,$$

We now assume that we know u_x, u_y and u on a curve (not just on (x_0, y_0)), with tangential directions $\vec{v}$ for every point on the curve. Then we also know the tangential derivatives of u_x and u_y in tangential direction $\vec{v}$ on the curve:

$$D_{\vec{v}} u_x = \lim_{t\to 0} \frac{u_x(x_0 + t\vec{v}) - u_x(x_0)}{t}, \quad \text{and} \quad D_{\vec{v}} u_y = \lim_{t\to 0} \frac{u_y(x_0 + t\vec{v}) - u_y(x_0)}{t} \,.$$

The tangential derivative can be expressed as

$$D_{\vec{v}} u_x = \nabla u_x \cdot \vec{v}, \quad D_{\vec{v}} u_y = \nabla u_y \cdot \vec{v} \,.$$

And therefore

$$\begin{aligned} v_1 u_{xx} + v_2 u_{xy} &= D_{\vec{v}} u_x \,, \\ v_1 u_{xy} + v_2 u_{yy} &= D_{\vec{v}} u_y \,. \end{aligned}$$

To summarize, we know that for a given curve with tangential direction $\vec{v}$ and given u_x, u_y and u on a part of the curve, the second order derivatives can be determined from the system

$$\begin{aligned} v_1 u_{xx} + v_2 u_{xy} \quad\quad\quad &= D_{\vec{v}} u_x \,, \\ v_1 u_{xy} + \quad v_2 u_{yy} &= D_{\vec{v}} u_y \,, \\ Au_{xx} + 2Bu_{xy} + \quad Cu_{yy} &= -Du_x - Eu_y - Fu - G \,. \end{aligned} \tag{3.9}$$

This linear system has a unique solution if

$$0 \neq \det \begin{pmatrix} A & 2B & C \\ v_1 & v_2 & 0 \\ 0 & v_1 & v_2 \end{pmatrix} = Av_2^2 - 2Bv_1 v_2 + Cv_1^2 =: Q(v_2, -v_1) \tag{3.10}$$

This means that we can solve the PDE (locally around a point on the curve) exactly.

Now, we reformulate the function Q:

$$\begin{aligned} Q(v_2,-v_1) &= Av_2^2 - 2Bv_1v_2 + Cv_1^2 \\ &= A\left(v_2^2 - 2\frac{B}{A}v_1v_2 + \frac{C}{A}v_1^2\right) \\ &= A\left((v_2 - \frac{B}{A}v_1)^2 - \left(\frac{B^2}{A^2} - \frac{C}{A}\right)v_1^2\right) \\ &=: A\left(\tilde{v}_2 + \frac{AC-B^2}{A^2}v_1^2\right) \end{aligned}$$

Depending on the sign of $AC - B^2$ we get different cases:

1. Elliptic: In this case $AC > B^2$, which in particular implies that $A \neq 0$. Then $Q(v_2,-v_1) = 0$ implies that $\tilde{v}_2 = v_1 = 0$, and consequently $v_1 = v_2 = 0$.

2. Hyperbolic: Now $AC < B^2$. If $A = 0$, every $\vec{v}$ solves $Q(v_2,-v_1) = 0$. If $A \neq 0$, then $\vec{v}$ is a solution of $Q(v_2,-v_1) = 0$ if

$$\tilde{v}_2 = \pm\sqrt{\frac{B^2}{A^2} - \frac{C}{A}}v_1 \,.$$

   Thus there exists a one-dimensional solution space, and thus the system (3.9) has nontrivial solutions too. That means that the solution of the PDE is not uniquely determined by the values u, u_x, u_y on the curve with tangent vector $\vec{v}$. Curves with such a property are called *characteristics*.

3. Parabolic: Here $AC = B^2$. If $A = 0$, every $\vec{v}$ is a solution of $Q(v_2,-v_1) = 0$. If $A \neq 0$, then $v_2 = \frac{B}{A}v_1$ is the solution of $Q(v_2,-v_1) = 0$.

Because $\vec{n} = (v_2,-v_1)$ is the normal vector to the tangent we get the following definition of characteristics:

Definition 3.3.1 *Let $AC - B^2 < 0$. A curve in the xy-plane is called* characteristics *of the hyperbolic PDE* (3.1) *if the normal vector $\vec{n} = (n_1, n_2)$ satisfies in every point the equation*

$$An_1^2 + 2Bn_1n_2 + Cn_2^2 = 0 \,.$$

Example 3.3.2 *The characteristics of the equation*

$$u_{xx} - u_{yy} = 0\,, \qquad (A = 1, C = -1, B = 0)$$

are the curves, where the normal vector satisfies

$$n_1^2 - n_2^2 = 0 \, .$$

That is $n_1 = \pm n_2$. *These are the lines* $x \pm y = const.$ *In this case the solution space* $Q(v_2, -v_1)$ *is a one dimensional linear space.*

If we identify y *with time* t, *this is the standard wave equation. The characteristics* $x \pm t = const$ *are the lines, where the waves propagate: If we consider waves, which move to the right, then this waves are given by*

$$u(x,t) = F(x-t) \, .$$

Note, that on the characteristics the value is constant, and from this value the solution cannot be determined in normal direction. The ones which move to the left are

$$u(x,t) = G(x+t) \, .$$

In general the waves are of the form

$$u(x,t) = F(x-t) + G(x+t) \, .$$

This example makes clear the role of characteristics.

Example 3.3.3 *Solve the PDE*

$$u_{xx} + 4u_{xy} + 4u_{yy} = 0$$

using the method of characteristics.

The following example shows that a initial value problem for a hyperbolic PDE, can be well defined only if the curve, where the initial conditions are given, is not a characteristic.

Example 3.3.4 *Consider the IVP for* $u(x,t)$:

$$\begin{aligned} u_{xt}(x,t) &= 0, \quad t \geq 0 \\ u(x,0) &= f(x), \\ u_t(x,0) &= g(x). \end{aligned}$$

Here $A = C = 0$ *and* $B = 1$. *The characteristics are the curves where* $n_1 n_2 = 0$. *The curve where the initial conditions are given is the* x*–axis with normal vector, for example,* $\vec{n} = (0,1)$. *Then, the* x*–axis is a characteristic and the IVP is not well defined (not unique solution).*

For example, let $f = g = 0$. *Obviously, one solution is the trivial* $u(x,t) = 0$, *but also the function* $u(x,t) = t^2$ *is a solution. In general all the functions of the form* $u(x,t) = F(t)$, *with* $F(0) = F'(0) = 0$ *are solutions.*

3.3.1 Characteristics for linear PDEs of the first order

We consider a PDE of the form:

$$Au_x + Bu_y + Cu + D = 0.$$

Similar to second order PDE, we want to see if the knowledge of u on a curve implies, together with the PDE, the knowledge of the solution in a neighborhood of the curve.

Thus, we compute the first derivatives at a point (x_0, y_0) using the PDE and similar to (3.9) we obtain the system

$$\begin{aligned} Au_x + Bu_y &= -Cu - D \\ v_1 u_x + v_2 u_y &= D_{\vec{v}} u. \end{aligned} \tag{3.11}$$

This system has a solution if $Av_2 - Bv_1 \neq 0$. Then the characteristics are defined by the equation $Av_2 - Bv_1 = 0$. We parametrize the curve $s \mapsto (x(s), y(s))$, and the tangential vector is $\vec{v} = (dx/ds, dy/ds)$. Then, the equation for the characteristic reads

$$\frac{dx}{dy} = \frac{A}{B}.$$

Chapter 4

Boundary Value Problems - Finite Difference Method

4.1 Ordinary Differential Equations

For motivating purposes we study first boundary value problems for ordinary differential equations:

$$\begin{aligned} L[u] = -u'' + b(x)u' + c(x)u = f(x) \quad \text{in } (0,1), \\ u(0) = u(1) = 0\,. \end{aligned} \tag{4.1}$$

It can be shown that this differential equation has a unique solution provided that

$$c(x) \geq 0\,, \qquad \forall x \in (0,1)\,,$$

and this will be assumed in the following. Here, for the numerical solution, we consider the *finite difference methods* (FDM). Later, we will investigate the *finite element methods* (FEM).

To simplify the presentation, we consider an equidistant grid

$$\Delta_h = \{x_i = ih : i = 1, \ldots, n-1, h = 1/n\} \subseteq (0,1)\,. \tag{4.2}$$

We denote by

$$\vec{u} = (u(x_1), \ldots, u(x_{n-1}))^t \in \mathbb{R}^{n-1} \tag{4.3}$$

the vector of the exact solution u of (4.1) on the grid Δ_h (4.2). In addition, we assume Dirichlet boundary conditions

$$0 = u(x_0) = u(x_n) = 0\,.$$

For the numerical solution we search for an approximating vector

$$\vec{u}_h = (u_1, \ldots, u_{n-1})^t \in \mathbb{R}^{n-1}\,. \tag{4.4}$$

For this purpose we discretize L from (4.1) by approximating the derivatives of u at the positions $x = x_i$ via *difference quotients*. Thereby we have several alternatives:

- *One-sided forward-difference operator:*

$$D_h^+[u](x) = \frac{u(x+h) - u(x)}{h} \sim u'(x)\,.$$

- *One-sided backward-difference operator:*

$$D_h^-[u](x) = \frac{u(x) - u(x-h)}{h} \sim u'(x)\,.$$

- *Central difference quotient:*

$$D_h[u](x) = \frac{u(x+h) - u(x-h)}{2h} \sim u'(x)\,. \tag{4.5}$$

Moreover, the second derivative can be approximated by a central difference quotient

$$D_h^2[u](x) = \frac{u(x+h) - 2u(x) + u(x-h)}{h^2} \sim u''(x)\,. \tag{4.6}$$

As we are going to see later the first two approximations admit a truncation error of order h and the last two of order h^2.

Example 4.1.1 *We study a simple case of* (4.1) *for* $b = c = 0$, *that is*

$$\begin{aligned} -u'' &= f, \quad \text{in } (0,1) \\ u(0) = u(1) &= 0. \end{aligned} \tag{4.7}$$

We approximate u'' *by* $D_h^2[u]$ *at the nodal points of* Δ_h. *Taking into account the Dirichlet boundary conditions* $u(x_0) = u(x_n) = 0$ *we get the discretized equation:*

$$\underbrace{\begin{pmatrix} f(x_1) \\ f(x_2) \\ \vdots \\ f(x_{n-1}) \end{pmatrix}}_{=:\vec{f}} = -\begin{pmatrix} u''(x_1) \\ u''(x_2) \\ \vdots \\ u''(x_{n-1}) \end{pmatrix} \simeq \underbrace{\frac{1}{h^2}\begin{pmatrix} 2 & -1 & & 0 \\ -1 & 2 & \ddots & \\ & \ddots & \ddots & -1 \\ 0 & & -1 & 2 \end{pmatrix}}_{=:L_h} \underbrace{\begin{pmatrix} u(x_1) \\ u(x_2) \\ \vdots \\ u(x_{n-1}) \end{pmatrix}}_{\vec{u}}.$$

Since we approximate $\vec{u}$ *by* $\vec{u}_h$, *we use the following linear equation to determine* $\vec{u}_h$*:*

$$L_h \vec{u}_h = \vec{f}\,. \tag{4.8}$$

The eigenvalues of L_h *are given by*

$$\lambda_k = -\frac{4}{h^2}\sin^2\left(\frac{kh\pi}{2}\right), \quad k = 1, \ldots, n-1.$$

The function $sinc(x) := \frac{\sin(x)}{x}$ *is monotonically decreasing in* $[0, \pi/2]$ *such that*

$$sinc(x) \geq sinc\left(\frac{\pi}{2}\right) = \frac{2}{\pi}, \quad \forall x \in [0, \pi/2],$$

which implies that:

$$\|L_h^{-1}\|_2 = \frac{1}{\lambda_{\min}(L_h)} = \max_{1 \leq k \leq n-1} \frac{h^2}{4\sin^2(kh\pi/2)} \leq \frac{1}{4}.$$

Consequently,

$$\begin{aligned} \|\vec{u} - \vec{u}_h\|_2 &= \|L_h^{-1}(L_h\vec{u} - \vec{f})\|_2 \\ &\leq \|L_h^{-1}\|_2 \|L_h\vec{u} - \vec{f}\|_2 \\ &\leq \frac{1}{4}\|L_h\vec{u} - \vec{f}\|_2 . \end{aligned} \tag{4.9}$$

Definition 4.1.2 *The FD scheme*

$$L_h[u](x_i) = f(x_i), \quad i = 1, ..., n-1$$

is said to be consistent with the PDE (4.1) *if for every smooth solution* u *the truncation error*

$$\epsilon_i := L_h[u](x_i) - f(x_i)$$

tends to zero as $h \to 0$*, meaning*

$$\lim_{h \to 0} \|\vec{\epsilon}_h\|_\infty = 0, \quad \textit{where} \quad \vec{\epsilon}_h = (\epsilon_1, \ldots, \epsilon_{n-1})^t$$

Moreover, if there exist a constant $C > 0$*, independent of* u, u', u'' *and a given* $h_0 > 0$ *such that for every* $h \in (0, h_0)$ *:*

$$\|\vec{\epsilon}_h\| \leq Ch^p, \quad p > 0,$$

the FD scheme is accurate of order p *for the norm* $\|\cdot\|$.

Then, if the FD scheme is consistent and there exists an estimate of the form (4.9), *consistency* implies *stability*. Consistency and stability imposes convergence.

In the following we determine error estimates for the different quotients presented before.

Lemma 4.1.3 *Let* $x \in [h, 1-h]$.

- *If* $u \in C^2[0,1]$*, for the one-sided difference quotients we have the estimate*

$$\left|D_h^{\pm}[u](x) - u'(x)\right| \leq \frac{1}{2}\|u''\|_\infty h .$$

- *If $u \in C^3[0,1]$, for a central difference quotient we have*

$$\left|D_h[u](x) - u'(x)\right| \le \frac{1}{6}\|u'''\|_\infty h^2 \,.$$

- *If $u \in C^4[0,1]$, the central difference quotient D_h^2 satisfies*

$$\left|D_h^2[u](x) - u''(x)\right| \le \frac{1}{12}\|u''''\|_\infty h^2 \,. \tag{4.10}$$

Proof 4.1.4 *We prove the assertion for the central difference quotient. Let $u \in C^3[0,1]$, then it follows from Taylor expansion around $x \in (0,1)$:*

$$\begin{aligned} u(x+h) &= u(x) + hu'(x) + \frac{1}{2}h^2u''(x) + \frac{1}{6}h^3u'''(\zeta_+), \\ u(x-h) &= u(x) - hu'(x) + \frac{1}{2}h^2u''(x) - \frac{1}{6}h^3u'''(\zeta_-), \end{aligned}$$

for some $\zeta_\pm$ satisfying $x-h < \zeta_- < x < \zeta_+ < x+h$. Therefore

$$u(x+h) - u(x-h) = 2hu'(x) + \frac{1}{6}h^3(u'''(\zeta_+) + u'''(\zeta_-)),$$

and thus

$$\left|\frac{u(x+h)-u(x-h)}{2h} - u'(x)\right| \le \frac{1}{12}h^2 \sup\left\{\left|u'''(\zeta)\right| : \zeta \in [0,1]\right\},$$

which gives the assertion.

Example 4.1.5 *Considering again the example 4.1.1 we find that, provided that the solution of the differential equation is four times continuously differentiable, that*

$$\|L_h\vec{u} - \vec{f}\|_\infty = \|D_h^2[\vec{u}] - \vec{f}\|_\infty \le \frac{1}{12}\|u''''\|_\infty h^2 = \frac{1}{12}\|f''\|_\infty h^2 \,.$$

In the following we discretize the operator L defined in (4.1). We use the discretization $D_h^2[u]$ for approximating u''. Moreover, the first derivative is approximated by one of the difference quotients $D_h^+[u]$, $D_h^-[u]$, $D_h[u]$, resulting to different tridiagonal matrices:

$$L_h = h^{-2}\begin{pmatrix} d_1 & s_1 & & 0 \\ r_2 & d_2 & \ddots & \\ & \ddots & \ddots & s_{n-2} \\ 0 & & r_{n-1} & d_{n-1} \end{pmatrix} \in \mathbb{R}^{(n-1)\times(n-1)}, \tag{4.11}$$

where

- if we use D_h^+, then

$$\begin{aligned} d_i &= 2 - hb(x_i) + h^2 c(x_i)\,, \\ r_i &= -1\,, \\ s_i &= -1 + hb(x_i)\,, \end{aligned} \tag{4.12}$$

- if we use D_h^-, then

$$\begin{aligned} d_i &= 2 + hb(x_i) + h^2 c(x_i)\,, \\ r_i &= -1 - hb(x_i)\,, \\ s_i &= -1\,, \end{aligned} \tag{4.13}$$

- and if we use D_h, then

$$\begin{aligned} d_i &= 2 + h^2 c(x_i)\,, \\ r_i &= -1 - hb(x_i)/2\,, \\ s_i &= -1 + hb(x_i)/2\,. \end{aligned} \tag{4.14}$$

The approximated solution is determined as the solution of the linear system (4.8).

Theorem 4.1.6 *Let the solution of the BVP* (4.1) *be four times continuously differentiable (which is for instance the case if b, c, f are twice continuously differentiable). Then the FD method* (4.8) *has the order of accuracy:*

- *$q = 2$, if the central difference quotient D_h is used for approximating u', or*
- *$q = 1$, if the forward or backward difference quotients $D_h^{\pm}$ are used for approximating u'.*

Remark 4.1.7 *Consistency depends on the choice of the norm. This happens because n is related to h. For example, the consistent FD scheme for D_h^2 is of order two in the infinity norm but it gives*

$$\|\vec{\epsilon}_h\|_1 = \mathcal{O}(h), \quad \textit{and} \quad \|\vec{\epsilon}_h\|_2 = \mathcal{O}(h^{3/2}).$$

Example 4.1.8 *a) We consider the one-dimensional BVP (elliptic)*

$$\begin{aligned} L[u] = -u'' + cu = f\,, \quad &\textit{in } \Omega = (0,1) \\ u(0) &= a \\ u(1) &= b, \end{aligned} \tag{4.15}$$

with Dirichlet boundary conditions, for $a, b \in \mathbb{R}$ and $c, f \in C[0,1]$. Using the above analysis we end up with the system

$$\frac{1}{h^2}\begin{pmatrix} 2 + h^2 c(x_1) & -1 & & 0 \\ -1 & 2 + h^2 c(x_2) & \ddots & \\ & \ddots & \ddots & -1 \\ 0 & & -1 & 2 + h^2 c(x_{n-1}) \end{pmatrix}\begin{pmatrix} u(x_1) \\ u(x_2) \\ \vdots \\ u(x_{n-1}) \end{pmatrix} = \begin{pmatrix} f(x_1) + \frac{a}{h^2} \\ f(x_2) \\ \vdots \\ f(x_{n-1}) + \frac{b}{h^2} \end{pmatrix}.$$

If $c \geq 0$, then we see that the symmetric matrix L_h is also positive definite, which means that the inverse exists.

b) In the case of Neumann boundary conditions, we have

$$
\begin{aligned}
L[u] = -u'' + cu = f, \quad & in\ \Omega = (0,1) \\
u'(0) = a & \\
u'(1) = b. &
\end{aligned}
\tag{4.16}
$$

Here, a solution exists if $c \geq c_0 > 0$ for all $x \in [0,1]$. To discretize the boundary conditions we consider the forward difference quotient for the BS at zero and the backward difference quotient for the BS at $x = 1$

$$u'(0) \simeq \frac{u(h) - u(0)}{h} = a, \quad u'(1) \simeq \frac{u(1) - u(1-h)}{h} = b,$$

resulting to

$$u(x_0) - u(x_1) = -ha, \quad -u(x_{n-1}) + u(x_n) = hb$$

at the grid points. Then, we obtain the augmented system

$$\frac{1}{h^2}\begin{pmatrix} 1 & -1 & & & 0 \\ -1 & 2 + h^2 c(x_1) & \ddots & & \\ & \ddots & \ddots & -1 & \\ & & -1 & 2 + h^2 c(x_{n-1}) & -1 \\ 0 & & & -1 & 1 \end{pmatrix} \begin{pmatrix} u(x_0) \\ u(x_1) \\ u(x_2) \\ \vdots \\ u(x_{n-1}) \\ u(x_n) \end{pmatrix} = \begin{pmatrix} -\frac{a}{h} \\ f(x_1) \\ f(x_2) \\ \vdots \\ f(x_{n-1}) \\ \frac{b}{h} \end{pmatrix},$$

where now $L_h \in \mathbb{R}^{(n+1)\times(n+1)}$. This FD scheme is also accurate but of first order. To improve consistence, we can use central differences at the boundary conditions to get consistence of the second order (in the infinity norm) but then we increase the number of unknowns.

4.2 Partial Differential Equations

4.2.1 One-dimensional FDM

We consider the one-dimensional heat equation for $u(x,t)$:

$$
\begin{aligned}
L[u] = \frac{\partial u}{\partial t} - a\frac{\partial^2 u}{\partial x^2} = 0, \quad & x \in \Omega = (0,1),\ t > 0 \\
u(x,0) = u_0(x), \quad & x \in \Omega \\
u(1,t) = u(0,t) = 0, &
\end{aligned}
\tag{4.17}
$$

where $a > 0$ and $u_0 \in C[0,1]$. We consider the same equidistant grid Δ_h for the spatial variable x and a regular grid for the time variable t

$$t^k = k\delta, \quad \text{for} \quad k \in \mathbb{N},\ \delta > 0.$$

Then the initial conditions read

$$u_j^0 := u(x_j, t^0) = u_0(x_j), \quad j = 0, \dots, n$$

and the boundary conditions take the form

$$u_0^k = u_n^k = 0, \quad k \in \mathbb{N}.$$

We consider the forward difference quotient for the time derivative and the central quotient for the second spatial derivative

$$\begin{aligned}\frac{\partial u}{\partial t}(x,t) &= \frac{u(x,t+\delta) - u(x,t)}{\delta} + \mathcal{O}(\delta)\\ \frac{\partial^2 u}{\partial x^2}(x,t) &= \frac{u(x+h,t) - 2u(x,t) + u(x-h,t)}{h^2} + \mathcal{O}(h^2).\end{aligned}$$

Then, the discretized version of the BVP (4.17) reads

$$\begin{aligned}\frac{u_j^{k+1} - u_j^k}{\delta} - a\frac{u_{j+1}^k - 2u_j^k + u_{j-1}^k}{h^2} &= 0, \qquad j = 1, \dots, n-1,\, k \geq 0\\ u_j^0 &= u_0(x_j), \quad j = 1, \dots, n-1\\ u_0^{k+1} = u_n^{k+1} &= 0, \qquad k \geq 0.\end{aligned} \tag{4.18}$$

The above equation can be written in the form

$$u_j^{k+1} = u_j^k + \frac{a\delta}{h^2}\left(u_{j+1}^k - 2u_j^k + u_{j-1}^k\right),$$

which is an explicit scheme (with respect to the time variable).

Definition 4.2.1 *The FD scheme* (4.18) *is consistent with the partial differential operator* $L = \partial_t - a\partial_{xx}$ *if for every smooth solution* u *the truncation error*

$$\epsilon_h[u](x,t) := \frac{u(x,t+\delta) - u(x,t)}{\delta} - a\frac{u(x+h,t) - 2u(x,t) + u(x-h,t)}{h^2} - L[u](x,t)$$

tends to zero as $\delta, h \to 0$, *independently.*

Moreover, if there exist a constant $C > 0$, *independent of* u *and its derivatives, such that*

$$\|\epsilon_h[u]\|_\infty \leq C(h^p + \delta^q), \quad p > 0,\, q > 0,$$

the scheme is said to be accurate of order p *in space and of order* q *in time.*

Theorem 4.2.2 *If* $u \in C^{4,2}([0,1],[0,T])$, *then the explicit scheme* (4.18) *is consistent, first order accurate in time and second order in space.*

We consider now an one-dimensional hyperbolic PDE. The simplest conservation equation is the transport equation

$$u_t = -au_x\,, \quad x \in \mathbb{R}\,, t \geq 0\,, \tag{4.19}$$

where we assume $a > 0$. The case $a < 0$ can be treated in an analogous manner by considering $\tilde{u}(x,t) = u(-x,t)$.

For a given initial value

$$u(x,0) = u_0(x) \tag{4.20}$$

the solution of (4.19) is given analytically by

$$u(x,t) = u_0(x - at)\,.$$

For the sake of simplicity we again assume an equidistant grid with step size h in space and δ in time, respectively. In this way we get a two-dimensional cartesian grid:

$$\{x_j = jh : j \in \mathbb{Z}\} \times \{t^k = k\delta : k \in \mathbb{N}_0\} \subseteq \mathbb{R} \times \mathbb{R}_0^+\,. \tag{4.21}$$

In each node we determine an approximate value $u_j^k \approx u(x_j, t^k)$. To approximate the time derivative we use the forward difference quotient

$$u_t(x_j, t^k) \approx \frac{u_j^{k+1} - u_j^k}{\delta}\,.$$

For the discretization of the spatial derivative we use one of the quotients D_h, D_h^+ and D_h^-. The initial condition reads

$$u_j^0 := u(x_j, t^0) = u_0(x_j), \quad j \in \mathbb{Z}$$

If we choose D_h^+, we obtain the explicit scheme

$$u_j^{k+1} = u_j^k - \frac{a\delta}{h}\left(u_{j+1}^k - u_j^k\right).$$

In general we obtain the following compact form of the iterative scheme

$$u^{k+1} = A_h u^k, \tag{4.22}$$

where

$$A_h = \begin{pmatrix} \ddots & \ddots & & & & 0 \\ \ddots & d & s & & & \\ & r & d & s & & \\ & & r & d & s & \\ & & & r & d & \ddots \\ 0 & & & & \ddots & \ddots \end{pmatrix},$$

with constant entries depending on the choice of the difference scheme:

- if we use D_h^+, then

$$d = 1 + \frac{a\delta}{h}, \quad r = 0, \quad s = -\frac{a\delta}{h},$$

- if we use D_h^-, then

$$d = 1 - \frac{a\delta}{h}, \quad r = \frac{a\delta}{h}, \quad s = 0,$$

- and if we use D_h, then

$$d = 1, \quad r = \frac{a\delta}{2h}, \quad s = -\frac{a\delta}{2h}.$$

Definition 4.2.3 *The FD scheme* (4.22) *is stable if for all* $T > 0$ *there exists a constant* $C_T > 0$ *such that*

$$\|A_h^k\|_\infty \le C_T, \quad \text{for all} \quad k, \delta, \ \text{with} \ 0 \le k\delta \le T.$$

The last case (D_h), results to a scheme with $\mathcal{O}(\delta + h^2)$ local truncation error but is unconditionally unstable. The first two cases are first order accurate in time and space but are stable under suitable conditions.

If $a \ge 0$, we use D_h^- and if $a < 0$ we use D_h^+. This algorithm is called *upwind -schemes* since the difference quotient to be used depend on the coefficient. The name comes form the fact that the transport equation also describes wind. Therefore, different difference quotients are used depending on the wind orientation.

The upwind scheme needs no numerical boundary condition and is stable if the Courant-Friedrichs-Lewy (CFL) condition

$$\delta \le \frac{h}{a},$$

is satisfied.

4.2.2 Two-dimensional FDM

We consider the two-dimensional Poisson equation for $u(x,y)$ in $\mathbb{R}^2$:

$$\begin{aligned} -\Delta u &= f, \quad x \in \Omega \subset \mathbb{R}^2 \\ u &= g, \quad x \in \partial\Omega, \end{aligned} \tag{4.23}$$

where Ω is a bounded, connected and open domain.

Theorem 4.2.4 *The Poisson problem* (4.23) *admits a unique solution, if it exists.*

For the homogeneous case $f \equiv 0$, the solution is given using separable variables. We specify the domain as the unit square: $\Omega = \{(x,y) : 0 < x,y < 1\}$ and the boundary conditions

$$\begin{aligned} u(0,y) = u(1,y) &= 0, \qquad 0 \le y \le 1 \\ u(x,0) &= 0, \qquad 0 \le x \le 1 \\ u(x,1) &= g(x), \quad 0 < x < 1, \end{aligned} \tag{4.24}$$

for $g \in C^1[0,1]$. We discretize the unit square using a set of grid points

$$\Omega_h = \{(x_j, y_k) = (jh, kh),\ 0 \le j,k \le n\},$$

with spacing $h = 1/n$. Then the set Ω_h contains $(n+1)^2$ grid points, $(n-1)^2$ are interior points and $4n$ points are located on the boundary.

We define the five-points discrete Laplace operator that acts at every interior point

$$\Delta_h[u](x_j, y_k) = \frac{1}{h^2}\left(-4u_{j,k} + u_{j-1,k} + u_{j+1,k} + u_{j,k-1} + u_{j,k+1}\right),$$

where $u_{j,k} := u(x_j, y_k)$.

Theorem 4.2.5 *Let $u \in C^4(\overline{\Omega})$. Then for every $1 \le j,k \le n-1$, we get*

$$|\Delta u(x_j, y_k) - \Delta_h[u](x_j, y_k)| \le \frac{h^2}{12} \max_{\overline{\Omega}} \left(\left| \frac{\partial^4 u}{\partial x^4} \right| + \left| \frac{\partial^4 u}{\partial y^4} \right| \right).$$

Thus, the discrete Laplacian Δ_h is a consistent second order approximation of the Laplace operator.

Then, the FD scheme of the problem (4.23) is to find $u_h : \overline{\Omega}_h \to \mathbb{R}$ as a solution of

$$\begin{aligned} -\Delta_h[u_h](x_j, y_k) &= f(x_j, y_k), \quad \text{for all } (x_j, y_k) \in \Omega_h \\ u_h(x_j, y_k) &= g(x_j, y_k), \quad \text{for all } (x_j, y_k) \in \partial\Omega_h \end{aligned} \tag{4.25}$$

This leads to a system of $(n-1) \times (n-1)$ linear equations for the $(n-1)^2$ unknowns (interior points). Now, in contrast to the one-dimensional case, u_h is a matrix and we have to rearrange its values as a column vector in order to obtain a linear system. To do so, we write the grid points according to the increasing order of the indices j,k (from the left to right and from the bottom to top).

The discrete problem takes the following form

$$A_h \tilde{u}_h = b_h,$$

where

$$A_h = \begin{pmatrix} L & -I & 0 & \cdots & 0 \\ -I & L & -I & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & -I & L & -I \\ 0 & \cdots & 0 & -I & L \end{pmatrix} \in \mathbb{R}^{(n-1)^2 \times (n-1)^2},$$

with 0 the $(n-1)\times(n-1)$ null matrix, I the $(n-1)\times(n-1)$ identity matrix and

$$L = \begin{pmatrix} 4 & -1 & 0 & \cdots & 0 \\ -1 & 4 & -1 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & -1 & 4 & -1 \\ 0 & \cdots & 0 & -1 & 4 \end{pmatrix} \in \mathbb{R}^{(n-1)\times(n-1)}.$$

Here

$$\tilde{u}_h = (u_{1,1}, \ldots, u_{1,n-1}, u_{2,1}, \ldots, u_{2,n-1}, \ldots, u_{n-1,n-1})^t \in \mathbb{R}^{(n-1)^2}$$

and

$$b_h = \begin{pmatrix} f_{1,1} + \frac{1}{h^2}(g_{0,1} + g_{1,0}) \\ f_{1,2} + \frac{1}{h^2} g_{0,2} \\ \vdots \\ f_{1,n-1} + \frac{1}{h^2}(g_{0,n-1} + g_{n-1,0}) \\ f_{2,1} + \frac{1}{h^2} g_{2,0} \\ f_{2,2} \\ \vdots \\ f_{2,n-1} + \frac{1}{h^2} g_{2,n-1} \\ \vdots \end{pmatrix} \in \mathbb{R}^{(n-1)^2},$$

where $f_{j,k} := f(x_j, y_k)$ and $g_{j,k} := g(x_j, y_k)$. The symmetric matrix A_h is again positive definite (well-posed problem) and the inverse is also bounded

$$\|A_h^{-1}\| \leq 1/2,$$

thus we have a convergent scheme.

The FD method can be extended to three dimensions without any further difficulty. However, the exponential increase of the matrix size makes this method not practical in more than two dimensions.

Chapter 5

PDEs - Finite Element Methods

In this chapter we consider *finite element methods* (FEM) for the solution of elliptic differential equations. Initially, we consider one-dimensional problems and then two-dimensional problems. The domain Ω, where we solve the partial differential equation is bounded and connected and has piecewise linear boundary Γ.

5.1 One-dimensional problems

The basis of finite element methods are *weak* solutions. We just give a short sketch of the basics of this theory:

Definition 5.1.1 *The space $H^1(\Omega)$ denotes the space of square integrable functions with square integrable derivatives and the inner product*

$$\langle u, v\rangle_{H^1(\Omega)} = \int_\Omega u'(x)v'(x)\,dx + \int_\Omega u(x)v(x)\,dx\,.$$

The associated norm is denoted by $\|\cdot\|_{H^1(\Omega)}$ and the semi-norm is denoted by

$$|u|^2_{H^1(\Omega)} = \int_\Omega (u'(x))^2\,dx\,.$$

The functions $u \in H^1(\Omega)$ are not necessarily continuous in Ω, but it is still possible to define boundary values. Here, we are only interested on point evaluations of them. Of particular importance is the set of zero-Dirichlet data:

$$H^1_0(\Omega) = \left\{u \in H^1(\Omega) : u|_\Gamma = 0\right\},$$

which is a closed linear subspace of $H^1(\Omega)$.

For functions in $H^1_0(\Omega)$ the Poincare - Friedrich inequality holds

$$\gamma_\Omega \|u\|_{H^1(\Omega)} \leq |u|_{H^1(\Omega)}\,, \qquad \forall u \in H^1_0(\Omega)\,. \tag{5.1}$$

After this clarification of notation we are investigating now elliptic differential equations of the form:

Let $\Omega = (0,1)$, $f \in L^2(\Omega)$ and $c \in L^\infty(\Omega)$. Find u that satisfies

$$\begin{aligned} -u''(x) + c(x)u(x) &= f(x), \quad x \in (0,1) \\ u(0) = u(1) &= 0. \end{aligned} \tag{5.2}$$

A classical solution (referring to standard theory) is one, where the second derivative is continuous. Here, we consider weak solutions, using integration by parts.

To obtain the weak formulation of (5.2) we consider

$$-\int_\Omega u''(x)v(x)dx + \int_\Omega c(x)u(x)v(x)dx = \int_\Omega f(x)v(x)dx, \quad \forall v \in H_0^1(\Omega)$$

and using integration by parts and that $v \in H_0^1(\Omega)$

$$\int_\Omega u'(x)v'(x)dx + \int_\Omega c(x)u(x)v(x)dx = \int_\Omega f(x)v(x)dx, \quad \forall v \in H_0^1(\Omega).$$

Let

$$\begin{aligned} a(u,v) &:= \int_\Omega u'v'\,dx + \int_\Omega cuv\,dx, \\ l(v) &:= \int_\Omega fv\,dx\,. \end{aligned}$$

The term a is called *bilinear* form because it is linear in every component on $V = H^1(\Omega)$. Moreover, l is a linear operator on V. With this notation we have a compact form of the weak formulation of (5.2): Find $u \in H_0^1(\Omega)$ such that

$$a(u,v) = l(v), \qquad \forall v \in H_0^1(\Omega)\,. \tag{5.3}$$

Note, that the space $H_0^1(\Omega)$ is designed such that the solution satisfies homogeneous Dirichlet conditions.

Definition 5.1.2 *Let V be Hilbert space. A bilinear form $a : V \times V \to \mathbb{R}$ is called*

- *symmetric, if $a(u,v) = a(v,u)$, for all $u,v \in V$,*
- *continuous, if there exists a number $a_\infty \in \mathbb{R}_+$ such that*

$$|a(u,v)| \le a_\infty \|u\|_V \|v\|_V\,, \qquad \forall u,v \in V\,,$$

- *V-elliptic, if there exists a constant $\alpha_0 > 0$ such that*

$$a(v,v) \ge a_0 \|v\|_V^2\,, \qquad \forall v \in V\,.$$

5.1.1 The Ritz and Galerkin methods

We write (5.3) in a general form:

$$\text{Find } u \in V \text{ such that } \quad a(u,v) = l(v), \quad \forall v \in V\,. \tag{5.4}$$

If $a : V \times V \to \mathbb{R}$ is symmetric, the above problem is equivalent to

$$\text{Find } u \in V \text{ such that } \quad J(u) = \inf_{v \in V} J(v) := \tfrac{1}{2} a(v,v) - l(v).$$

The second step is to derive the variational approximation of (5.4). In the Ritz method, we replace V by a finite dimensional subspace V_h, with $\dim V_h = n$. Then, we solve

$$\text{Find } u_h \in V_h \text{ such that } \quad J(u_h) = \inf_{v_h \in V_h} J(v_h).$$

Let $(\phi_j)_{1\le j\le n}$ be a basis of V_h, then we can decompose every $u_h \in V_h$ as

$$u_h(x) = \sum_{j=1}^{n} u_j \phi_j(x) \tag{5.5}$$

and the minimization problem admits the matrix form

$$\text{Find } U \in \mathbb{R}^n \text{ such that } \quad \mathcal{J}(U) = \inf_{V \in \mathbb{R}^n} \mathcal{J}(V) = \tfrac{1}{2} V^\top A_h V - V^\top f_h,$$

where $U = (u_1, \dots, u_n)^\top$ and $A_h = K_h + L_h$ with

$$\begin{aligned}
(K_h)_{ij} &= \int_\Omega \phi_j' \phi_i' dx \quad \text{(Stiffness matrix)}\\
(L_h)_{ij} &= \int_\Omega c\phi_j \phi_i dx \quad \text{(Mass matrix)}\\
(f_h)_i &= \int_\Omega f \phi_i dx
\end{aligned} \tag{5.6}$$

The matrix A_h is symmetric and positive definite and thus the functional $\mathcal{J}$ is quadratic. Then, we have existence and uniqueness of the solution for the minimization problem.

In the Petrov-Galerkin method, we consider two finite-dimensional spaces V_h and W_h with $\dim V_h = \dim W_h = n$, the approximation and the test space, respectively. Then the discretized version of (5.4) reads

$$\text{Find } u_h \in V_h \text{ such that } \quad a(u_h, v_h) = l(v_h), \quad \forall v_h \in W_h\,. \tag{5.7}$$

Let $(\phi_j)_{1\le j\le n}$ be a basis of V_h, and $(\psi_j)_{1\le j\le n}$ be a basis of W_h, then we can decompose every $u_h \in V_h$ and $v_h \in W_h$ as

$$u_h(x) = \sum_{j=1}^{n} u_j \phi_j(x), \quad v_h(x) = \sum_{j=1}^{n} v_j \psi_j(x) \tag{5.8}$$

and the problem (5.7) admits the matrix form

$$\text{Find } U \in \mathbb{R}^n \text{ such that} \quad A_h U = f_h,$$

where $U = (u_1, ..., u_n)^\top$ and now

$$(A_h)_{ij} = a(\phi_j, \psi_i), \quad (f_h)_i = \int_\Omega f \psi_i dx$$

If the bilinear form a is symmetric, the Galerkin and Ritz methods are equivalent.

5.1.2 The Finite Element Methods

We define $\mathbb{P}_k$ to be the vector space of polynomials in one variable degree less or equal to k :

$$\mathbb{P}_k = \left\{ p(x) = \sum_{j=0}^{k} a_j x^j, \quad a_j \in \mathbb{R} \right\}.$$

We consider a uniform mesh : $x_j = jh$, $j = 0, ..., n$ with $h = 1/n$. Then, for $0 = x_0 < x_1 < ... < x_n = 1$ we construct the intervals $K_j = [x_j, x_{j+1}]$.

We define the space of globally continuous affine functions on each interval:

$$V_h^1 = \{u_h \in C^0[0,1] : u_h|_{K_j} \in \mathbb{P}_1, 0 \le j \le n-1\}$$

and its subspace

$$V_{0,h}^1 = \{u_h \in V_h^1 : u_h(0) = u_h(1) = 0\}.$$

The space V_h^1 is a $(n+1)$–dimensional subspace of $H^1(\Omega)$. The space $V_{0,h}^1$ is a $(n-1)$–dimensional subspace of $H_0^1(\Omega)$. Moreover, every $v_h \in V_{0,h}^1$ can be uniquely written as

$$u_h(x) = \sum_{j=1}^{n-1} u_h(x_j)\phi_j(x), \quad \forall x \in [0,1], \tag{5.9}$$

for a given basis $(\phi_j)_{1 \le j \le n-1}$. Then the variational problem

$$\text{Find } u_h \in V_{0,h}^1 \text{ such that} \quad a(u_h, v_h) = l(v_h), \qquad \forall v_h \in V_{0,h}^1\,, \tag{5.10}$$

can be written as a linear system

$$\text{Find } U_h \in \mathbb{R}^{n-1} \text{ such that} \quad A_h U_h = f_h, \tag{5.11}$$

where now $U_h = (u_h(x_1), ..., u_h(x_{n-1}))^\top$, $A_h \in \mathbb{R}^{(n-1)\times(n-1)}$ and $f_h \in \mathbb{R}^{n-1}$ are as in (5.6). If the bilinear form a is V–elliptic, then the matrix $A_h = K_h + L_h$ is positive definite and we have existence and uniqueness of solution by the Lax-Milgram lemma.

Theorem 5.1.3 *Let $u \in H_0^1(\Omega)$ and $u_h \in V_{0,h}^1$, be the solutions of* (5.2) *and* (5.10)*, respectively. Then, the FEM (with Lagrange $\mathbb{P}_1$ polynomials) converges, meaning*

$$\lim_{h\to 0} \|u - u_h\|_{H^1(\Omega)} = 0.$$

We now compute the elements of the matrix A_h for a basis function of the form

$$\phi_j(x) = \begin{cases} \frac{x-x_{j-1}}{h}, & x \in [x_{j-1}, x_j] \\ \frac{x_{j+1}-x}{h}, & x \in [x_j, x_{j+1}], \end{cases}$$

such that $\phi_j(x_i) = \delta_{ij}$. We consider the interval $K_j = [x_j, x_{j+1}]$ where we have only two non-zero shape functions:

$$\begin{aligned} \phi_j|_{K_j} &= \frac{x_{j+1} - x}{h}, \qquad \phi_j'|_{K_j} = -\frac{1}{h} \\ \phi_{j+1}|_{K_j} &= \frac{x - x_j}{h}, \qquad \phi_{j+1}'|_{K_j} = \frac{1}{h}. \end{aligned} \tag{5.12}$$

We define then the 2×2 stiffness matrix $\mathcal{K}_j$ and mass matrix $\mathcal{L}_j$ by

$$\mathcal{K}_j = \begin{pmatrix} \mathcal{K}_{11}^j & \mathcal{K}_{12}^j \\ \mathcal{K}_{12}^j & \mathcal{K}_{22}^j \end{pmatrix}, \quad \mathcal{L}_j = \begin{pmatrix} \mathcal{L}_{11}^j & \mathcal{L}_{12}^j \\ \mathcal{L}_{12}^j & \mathcal{L}_{22}^j \end{pmatrix}, \tag{5.13}$$

where

$$\begin{aligned} \mathcal{K}_{11}^j &= \int_{x_j}^{x_{j+1}} (\phi_j'(x))^2 dx, & \mathcal{L}_{11}^j &= \int_{x_j}^{x_{j+1}} c(x)(\phi_j(x))^2 dx, \\ \mathcal{K}_{12}^j &= \int_{x_j}^{x_{j+1}} \phi_j'(x)\phi_{j+1}'(x) dx, & \mathcal{L}_{12}^j &= \int_{x_j}^{x_{j+1}} c(x)\phi_j(x)\phi_{j+1}(x) dx, \\ \mathcal{K}_{22}^j &= \int_{x_j}^{x_{j+1}} (\phi_{j+1}'(x))^2 dx, & \mathcal{L}_{22}^j &= \int_{x_j}^{x_{j+1}} c(x)(\phi_{j+1}(x))^2 dx. \end{aligned} \tag{5.14}$$

In the special case of c being constant, $c(x) = c_0$ for all $x \in \Omega$ we get

$$\mathcal{K}_j = \frac{1}{h}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}, \quad \mathcal{L}_j = \frac{c_0 h}{6}\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}. \tag{5.15}$$

For the right-hand side f_h we have to compute

$$(f_h)_i = \int_\Omega f\phi_i dx = \sum_{j=0}^{n-1} \int_{K_j} f\phi_i dx.$$

We decompose $f = \sum_{k=1}^{n-2} f_k\phi_k$ and then we have to compute the integrals

$$\int_{K_j} \phi_k(x)\phi_i(x) dx,$$

using, for instance, the trapezoidal or Simpson formulas.

The FEM method can also be applied to a Neumann boundary value problem. Let $\Omega = (0,1)$, $f \in L^2(\Omega)$ and $c \in L^\infty(\Omega)$, such that $c(x) \geq c_0 > 0$. Find u that satisfies

$$\begin{aligned} -u''(x) + c(x)u(x) &= f(x), \quad x \in (0,1) \\ u'(0) &= a \\ u'(1) &= b. \end{aligned}$$

Using integration by parts now we obtain

$$\int_\Omega u'(x)v'(x)dx - (u'(x)v(x))|_0^1 + \int_\Omega c(x)u(x)v(x)dx = \int_\Omega f(x)v(x)dx$$

which together with the Neumann boundary condition results to

$$\int_\Omega u'(x)v'(x)dx + \int_\Omega c(x)u(x)v(x)dx = \int_\Omega f(x)v(x)dx + bv(1) - av(0).$$

The variational problem takes the form

$$\text{Find } U_h \in \mathbb{R}^{n+1} \text{ such that } \quad A_h U_h = f_h,$$

where now $U_h = (u_h(x_0), ..., u_h(x_n))^\top$, $(A_h)_{ij} = a(\phi_j, \phi_i) \in \mathbb{R}^{(n+1)\times(n+1)}$ and

$$\begin{aligned} (f_h)_0 &= \int_\Omega f\phi_0 dx - a, \\ (f_h)_i &= \int_\Omega f\phi_i dx, \quad 1 \leq i \leq n-1 \\ (f_h)_n &= \int_\Omega f\phi_n dx + b \end{aligned}$$

5.2 Two-dimensional problems

Let now $\Omega \subset \mathbb{R}^2$, be open and bounded. We define

$$\langle u, v\rangle_{H^1(\Omega)} = \int_\Omega \nabla u(x) \cdot \nabla v(x)\, dx + \int_\Omega u(x)v(x)\, dx\,.$$

and the semi-norm

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

We consider the elliptic differential equation

$$L[u] := -\nabla \cdot (\sigma \nabla u) + cu = f, \quad \text{in } \Omega \tag{5.16}$$

with Dirichlet boundary conditions

$$u = 0, \quad \text{on } \Gamma . \tag{5.17}$$

Aside from some smoothness conditions (which we do not discuss in detail) essential conditions are the following:

$$0 < \sigma_0 \leq \sigma(x) \leq \sigma_\infty \text{ and } 0 \leq c(x) \leq c_\infty .$$

These conditions guarantee ellipticity.

Using Green's first identity (partial integration in $\mathbb{R}$) we get the weak formulation of (5.16)

$$\begin{aligned}\int_\Omega f v \, dx &= -\int_\Omega \nabla \cdot (\sigma \nabla u) v \, dx + \int_\Omega c u v \, dx \\ &= \int_\Omega \sigma \nabla u \cdot \nabla v \, dx + \int_\Omega c u v \, dx - \int_\Gamma v \sigma \frac{\partial u}{\partial n} \, ds .\end{aligned}$$

Definition 5.2.1 *A weak solution of the* homogeneous Dirichlet-problem, *that is of* (5.16) – (5.17), *solves*

$$\int_\Omega f v \, dx = \int_\Omega \sigma \nabla u \cdot \nabla v \, dx + \int_\Omega c u v \, dx, \qquad \forall v \in H_0^1(\Omega) . \tag{5.18}$$

Remark 5.2.2 *The weak solution is unique.*

The *inhomogeneous Dirichlet problem* consists in solving (5.16) together with boundary conditions:

$$u = g, \quad \text{on } \Gamma . \tag{5.19}$$

We extend the function g from Γ to Ω and denote such an extension by u_0. With u_0 we reduce (5.16) – (5.19) to a homogeneous Dirichlet problem. In fact $w := u - u_0$ solves the homogeneous Dirichlet problem

$$L[w] = f - L[u_0], \qquad w|_\Gamma = 0 .$$

The inhomogeneous *Dirichlet problem* has a unique solution:

Theorem 5.2.3 *Let σ, c and f be bounded functions satisfying*

$$0 \leq c(x) \leq c_\infty \text{ and } 0 < \sigma_0 \leq \sigma(x) \leq c_\infty .$$

Then, the Dirichlet problem has a unique weak solution.

The *Neumann problem* is to solve (5.16) together with boundary conditions

$$\sigma \frac{\partial u}{\partial n} = g, \quad \text{on } \Gamma\,. \tag{5.20}$$

The weak form of (5.16) is again derived by partial integration (using (5.20)):

$$\begin{aligned}\int_\Omega f v\,dx &= \int_\Omega \nabla\cdot(\sigma\nabla u)v\,dx + \int_\Omega cuv\,dx \\ &= \int_\Omega \sigma\nabla u\cdot\nabla v\,dx + \int_\Omega cuv\,dx - \int_\Gamma v\sigma\frac{\partial u}{\partial n}\,ds \\ &= \int_\Omega \sigma\nabla u\cdot\nabla v\,dx + \int_\Omega cuv\,dx - \int_\Gamma vg\,ds\,.\end{aligned}$$

Theorem 5.2.4 *Let σ, c and f be bounded functions satisfying*

$$0 < c_0 \le c(x) \le c_\infty, \quad \text{and} \quad 0 < \sigma_0 \le \sigma(x) \le c_\infty\,.$$

Then, the Neumann problem has a unique weak solution. For $c = 0$ we also have weak solutions, provided that

$$\int_\Omega f\,dx = -\int_\Gamma g\,ds\,. \tag{5.21}$$

In this case, however, the solution is not unique, and all solutions differ by a constant. Typically we select the one, which satisfies $\int_\Omega u\,dx = 0$.

As before we set

$$\begin{aligned} a(u,v) &:= \int_\Omega \nabla u\cdot\nabla v\,dx + \int_\Omega cuv\,dx\,, \\ l(v) &:= \int_\Omega f v\,dx \end{aligned}$$

and we write (5.16) and (5.17) in a compact form:

$$\text{Find} \quad u \in H_0^1(\Omega) \quad \text{such that} \quad a(u,v) = l(v), \quad \forall v \in H_0^1(\Omega)\,. \tag{5.22}$$

Proposition 5.2.5 *Let σ, c and f be bounded functions satisfying*

$$0 \le c(x) \le c_\infty, \quad \text{and} \quad 0 < \sigma_0 \le \sigma(x) \le \sigma_\infty\,.$$

Then a from (5.2) is symmetric and bounded on $V = H^1(\Omega)$ with

$$a_\infty = \max\{\sigma_\infty, c_\infty\}$$

and $H_0^1(\Omega)$-elliptic with constants

$$a_0 = \gamma_\Omega^2 \sigma_0$$

(γ_0 is the Poincare-Friedrich constant (5.1)). Moreover,

$$a(v,v) \ge 0\,, \qquad \forall v \in H^1(\Omega)\,.$$

Proof 5.2.6 *To prove boundness, we apply twice the Cauchy-Schwarz inequality for functions and numbers. Then, for all* $u, v \in H_0^1(\Omega)$ *we have*

$$\begin{aligned}|a(u,v)| &\leq \left|\int_\Omega \sigma \nabla u \cdot \nabla v \, dx\right| + \left|\int_\Omega cuv \, dx\right| \\ &\leq a_\infty \left(\int_\Omega |\nabla u \cdot \nabla v| \, dx + \int_\Omega |uv| \, dx\right) \\ &\leq a_\infty \left(\|\nabla u\|_{L^2(\Omega)} \|\nabla v\|_{L^2(\Omega)} + \|u\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)}\right) \\ &\leq a_\infty \left(\sqrt{\|u\|_{L^2(\Omega)}^2 + \|\nabla u\|_{L^2(\Omega)}^2} \cdot \sqrt{\|v\|_{L^2(\Omega)}^2 + \|\nabla v\|_{L^2(\Omega)}^2}\right).\end{aligned}$$

For $v \in H^1(\Omega)$ *we have:*

$$\begin{aligned}|a(v,v)| &= \int_\Omega \sigma \, |\nabla v|^2 \, dx + \int_\Omega cv^2 \, dx \\ &\geq \sigma_0 \int_\Omega \sigma \, |\nabla v|^2 \, dx \\ &\geq 0 \,.\end{aligned} \tag{5.23}$$

From (5.23) *it follows for* $v \in H_0^1(\Omega)$.

$$|a(v,v)| \geq \sigma_0 \, |v|_{H_0^1(\Omega)}^2 \geq \sigma_0 \gamma_\Omega^2 \|v\|_{H^1(\Omega)}^2 \,.$$

To determine an approximate solution we use again the Galerkin approach. We select a finite dimensional subspace $V_h \subseteq H_0^1(\Omega)$ and the variational problem reads

$$\text{Find } u_h \in V_{0,h}^1 \text{ such that} \quad a(u_h, v_h) = l(v_h), \qquad \forall v_h \in V_{0,h}^1 \,. \tag{5.24}$$

which can be written as a linear system

$$\text{Find } U_h \in \mathbb{R}^{n-1} \text{ such that} \quad A_h U_h = f_h, \tag{5.25}$$

where $U_h = (u_h(x_1), \ldots, u_h(x_{n-1}))^\top$, $A_h = (a(\phi_j, \phi_i))_{ij} \in \mathbb{R}^{(n-1)\times(n-1)}$ and $f_h = (l(\phi_i))_i \in \mathbb{R}^{n-1}$

Remark 5.2.7 *In order to solve the inhomogeneous Dirichlet problem, we determine a function* $\omega : \Omega \to \mathbb{R}$*, which extends the boundary data onto* Ω*: Then we are looking for a solution* $u^\dagger$ *of*

$$a(u_0 + \omega, v) = l(v)\,, \qquad \forall v \in H_0^1(\Omega)\,, \tag{5.26}$$

the solution of the inhomogeneous problem is then $u^\dagger + \omega$.

For the Neumann problem the right hand side reads as follows:

$$l(v) = \int_\Omega f v \, dx \; + \; \int_\Gamma gv \, ds$$

and the linear system takes the form

$$a(u,v) = l(v)\,, \qquad \forall v \in H^1(\Omega)\,.$$

As in the one-dimensional setting, we need an appropriate *ansatz space* $V_h \subseteq H^1(\Omega)$. Typically finite Element methods are based on *triangulations* of the domain Ω.

Definition 5.2.8 *A set of open triangles* $\Gamma = \{T_1, \dots, T_m\}$ *is called* regular triangulation *of* Ω, *if*

1. $T_i \cap T_j = \emptyset \quad \forall i \neq j$,
2. $\bigcup_{i=1}^{m} \overline{T_i} = \overline{\Omega}$,
3. *for* $i \neq j$ *we have either*
    (a) $\overline{T_i} \cap \overline{T_j} = \emptyset$,
    (b) $\overline{T_i} \cap \overline{T_j}$ *is a joint vertex of* T_i *and* T_j, *or*
    (c) *a common edge.*

The vertices of the triangle are called corners.

On a triangulation we define linear ansatz functions (in analogy to linear splines).

Theorem 5.2.9 *Let* Γ *be a regular triangulation of a polygonal domain* Ω *with nodal points* x_i, $i = 1, \dots, n$. *Then there exist continuous functions* $\Lambda_i : \Omega \to \mathbb{R}$, $i = 1, \dots, n$, *satisfying:*

1. $\Lambda_i(x_j) = \delta_{ij}\,, \quad i, j = 1, \dots, n\,,$
2. $\Lambda_i(x) = \beta_{ik} + \alpha_{ik} \cdot x$, *for* $x \in T_k$ *with* $\alpha_{ik} \in \mathbb{R}^2$, $\beta_{ik} \in \mathbb{R}$.

The set $V^\Gamma = span\{\Lambda_1, \dots, \Lambda_n\}$ *consists of piecewise linear functions with respect to* Γ. *The gradient of an element in* V^Γ *is piecewise constant and we have* $V^\Gamma \subseteq H^1(\Omega)$.

Definition 5.2.10 *The pair* (Γ, V^Γ) *is called* finite elements.

The analog of Lagrange interpolation for finite elements reads as follows:

Theorem 5.2.11 *Let* Γ *be a regular triangulation of* $\Omega \subseteq \mathbb{R}^2$ *with nodal points* $\{x_i : i = 1, \dots, n\}$. *Let* $\{y_i : i = 1, \dots, n\}$ *be given. Then* $\psi(x) = \sum_{i=1}^{n} y_i \Lambda_i(x) \in V^\Gamma$ *and*

$$\psi(x_i) = y_i\,, \quad i = 1, \dots, n\,.$$

5.3 Stiffness Matrix

We now focus on how to compute *stiffness matrix* A_h, given by

$$(A_h)_{ij} = a(\phi_i, \phi_j) = \int_\Omega \sigma \nabla\phi_i \cdot \nabla\phi_j + c\phi_i\phi_j \, dx \,. \tag{5.27}$$

We emphasize that the matrix is sparse. For every triangle $T_k \in \Gamma$ the matrix

$$S_k = \left(\int_{T_k} \sigma \nabla\phi_i \cdot \nabla\phi_j + c\phi_i\phi_j \, dx \right)_{ij} \in \mathbb{R}^{n\times n} \tag{5.28}$$

consists of all integrals over the triangle T_k. These matrices S_k are called *element stiffness matrices*. Because of

$$\begin{aligned} a(\phi_i, \phi_j) &= \int_\Omega \sigma \nabla\phi_i \cdot \nabla\phi_j + c\phi_i\phi_j \, dx \\ &= \sum_{k=1}^{m} \int_{T_k} \sigma \nabla\phi_i \cdot \nabla\phi_j + c\phi_i\phi_j \, dx \end{aligned}$$

we have

$$A = \sum_{k=1}^{m} S_k \,. \tag{5.29}$$

To determine the element stiffness matrices we consider the transformation $\Phi : \mathbb{R}^2 \to \mathbb{R}^2$ that maps the *reference triangle*

$$D = \left\{ \hat{a} = (\hat{x}, \hat{y})^\top : \hat{x} > 0, \hat{y} > 0, \hat{x} + \hat{y} < 1 \right\} \tag{5.30}$$

onto the triangle $T \in \Gamma$ with corners $a_i = (x_i, y_i)^\top$, $i = 1, 2, 3$, given by

$$\Phi(\hat{a}) = a_1 + B\hat{a} = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + \begin{pmatrix} x_2 - x_1 & x_3 - x_1 \\ y_2 - y_1 & y_3 - y_1 \end{pmatrix} \begin{pmatrix} \hat{x} \\ \hat{y} \end{pmatrix}.$$

Since T is not degenerated, there exists B^{-1} and Φ is a 1-1 mapping. We have in addition $|T| = |\det(B)||D| = \frac{1}{2}|\det(B)|$. For every $v \in T$ we define $\hat{v} \in D$ such that

$$v(a) = \hat{v}\left(\Phi^{-1}(a)\right) \quad \Rightarrow \quad v(a) = \hat{v}(\hat{a}).$$

Therefore, we have

$$\Phi'(\hat{a}) = \begin{pmatrix} x_2 - x_1 & x_3 - x_1 \\ y_2 - y_1 & y_3 - y_1 \end{pmatrix}.$$

These two vectors are linear independent. Thus $d := \det(\Phi') \neq 0$ and we compute

$$\Phi'^{-1}(\hat{a}) = \frac{1}{d} \begin{pmatrix} y_3 - y_1 & x_1 - x_3 \\ y_1 - y_2 & x_2 - x_1 \end{pmatrix}.$$

Example 5.3.1 *We calculate the element stiffness matrix $K_h = K_{ij}$, for $L[u] = -\Delta u$ and a triangle $T \in \Gamma$ with corners a_1, a_2 and a_3.*

We denote by Λ_i, $i = 1,2,3$ the hat functions, with nodal value 1 at x_i and 0 else. Recall the property $\hat{\Lambda}_i(\hat{a}) = \Lambda_i(a)$. Therefore,

$$\begin{aligned} K_{ij} &= \int_T \nabla_x \Lambda_i(x) \cdot \nabla_x \Lambda_j(x)\, dx \\ &= \int_D \nabla_x \Lambda_i(\Phi(\hat{x})) \cdot \nabla_x \Lambda_j(\Phi(\hat{x})) \left| det \Phi' \right| d\hat{x} \\ &= |d| \int_D \Phi'^{-\top} \nabla_{\hat{x}} \hat{\Lambda}_i(\Phi(\hat{x})) \cdot (\Phi'^{-\top} \nabla_{\hat{x}} \Lambda_j(\Phi(\hat{x}))\, d\hat{x}\,. \end{aligned}$$

The functions $\Lambda_i(\Phi(\cdot))$ are again hat functions over D with nodal value 1 at $\hat{x}_i$ and 0 else. Then

$$G := \begin{pmatrix} \nabla_{\hat{x}}(\Lambda_1(\Phi(\cdot))^\top \\ \nabla_{\hat{x}}(\Lambda_2(\Phi(\cdot))^\top \\ \nabla_{\hat{x}}(\Lambda_3(\Phi(\cdot))^\top \end{pmatrix} = \begin{pmatrix} -1 & -1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Therefore, the integrands K_{ij} are constant on D. Consequently,

$$\begin{aligned} \left(K_{ij}\right)_{ij} &= \frac{|d|}{2} G \Phi'^{-1} \Phi'^{-\top} G^\top \\ &= \frac{1}{2\,|d|} \begin{pmatrix} y_2 - y_3 & x_3 - x_2 \\ y_3 - y_1 & x_1 - x_3 \\ y_1 - y_2 & x_2 - x_1 \end{pmatrix} \begin{pmatrix} y_2 - y_3 & y_3 - y_1 & y_1 - y_2 \\ x_3 - x_2 & x_1 - x_3 & x_2 - x_1 \end{pmatrix} \in \mathbb{R}^{3\times 3} \end{aligned}$$

5.4 Parabolic Differential Equations

In this section we study the numerical solution of parabolic initial value problems. We consider the following problem

$$\begin{aligned} u_t + L[u] &= f, && \text{for } (x,t) \in \Omega \times \mathbb{R}^+, \\ u(x,t) &= 0, && \text{for } (x,t) \in \Gamma \times \mathbb{R}^+, \\ u(x,0) &= u_0(x), && \text{for } x \in \Omega\,. \end{aligned} \tag{5.31}$$

Here u_0 is the initial value and $L[u]$ is defined in (5.16). As before we associate L with the bilinear form:

$$a(v,w) = \int_\Omega \sigma \nabla v \cdot \nabla w + c v w\, dx\,. \tag{5.32}$$

5.4.1 Method of Lines

We determine for every fixed $t > 0$ an approximation $u_h(t) \in V_h$, which satisfies the equation

$$\int_\Omega u_h' w\, dx + a(u_h, w) = \int_\Omega f w\, dx\,, \qquad \forall w \in V_h\,. \tag{5.33}$$

The solution u_h can be determined again with a Galerkin approach. For this purpose we set

$$u_h(t,x) = \sum_{j=1}^{n} \eta_j(t)\Lambda_j(x). \tag{5.34}$$

Let $w = \Lambda_k$, $k = 1,\dots,n$ then from (5.33) we get the system of differential equations:

$$\sum_{j=1}^{n} \eta_j'(t) \int_\Omega \Lambda_j \Lambda_k \, dx + \sum_{j=1}^{n} \eta_j(t) a(\Lambda_j, \Lambda_k) = \int_\Omega f(t,x)\Lambda_k \, dx, \quad \forall k = 1,\dots,n. \tag{5.35}$$

We define the vectors

$$y(t) = [\eta_j(t)]_j \in \mathbb{R}^n \text{ and } b(t) = \left(\int_\Omega f(t)\Lambda_k \, dx \right)_k \in \mathbb{R}^n,$$

and we can reformulate the system (5.35) as

$$Gy' = b - Ay, \tag{5.36}$$

for $y(0) = [\eta_j(0)]_j \in \mathbb{R}^n$, where G denotes the Gram's matrix of the hat functions in $L^2(\Omega)$ and $A = (a(\Lambda_j, \Lambda_k))_{jk}$. An equivalent formulation is

$$y' = G^{-1}b - G^{-1}Ay. \tag{5.37}$$

This method is called *(vertical) method of lines*, because it reduces the PDE to a system of ordinary differential equations with respect to time.

5.4.2 Crank-Nicolson Method

For the sake of simplicity of presentation we consider an equidistant grid with step size $\tau > 0$:

$$\Delta_\tau = \{t_i = i\tau : i \in \mathbb{N}_0\}.$$

We define functions with respect to space

$$u_i = \sum_{j=1}^{n} \eta_{ij}\Lambda_j,$$

which approximate $u_h(t_i)$. The values η_{ij} approximate also the function values of u at the nodes $\{x_j : j = 1,\dots,n\}$ of the triangulation Γ:

$$\eta_{ij} \approx u_h(x_j, t_i) \approx u(x_j, t_i).$$

To solve (5.37) in a stable way we use the mid-point integration rule, which results to

$$\left(I + \frac{\tau}{2}G^{-1}A\right)y_{i+1} = \left(I - \frac{\tau}{2}G^{-1}A\right)y_i + \tau G^{-1}b_{i+1/2}, \qquad i = 0,1,2\dots, \tag{5.38}$$

where

$$b_{i+1/2} = \left(\int_\Omega f_{i+1/2} \Lambda_j \, dx \right)_j , \qquad \text{with } f_{i+1/2} = f(t_i + \tau/2) .$$

An equivalente formulation of this method is:

$$\left(G + \frac{\tau}{2} A \right) y_{i+1} = \left(G - \frac{\tau}{2} A \right) y_i + \tau b_{i+1/2} , \qquad i = 0, 1, 2 \ldots , \tag{5.39}$$

Since both G and A are positive definite, $G + \frac{\tau}{2} A$ is also for all $\tau > 0$. This recursion can be written as a finite dimensional variational problem for $u_{i+1} \in V_h$:

$$\int_\Omega u_{i+1} w \, dx + \frac{\tau}{2} a(u_{i+1}, w) = \int_\Omega u_i w \, dx - \frac{\tau}{2} a(u_i, w) + \tau \int_\Omega f_{i+1/2} w \, dx . \tag{5.40}$$

This is the Crank-Nicolson method.