

Numerical Methods for Differential Equations

Ferran Arqué

2018

Contents

Ordinary Differential Equations

1 Ordinary Differential Equations. Basic concepts	3
1.1 Introduction and some notation	3
1.2 Euler's method	3
1.3 Enhanced Euler's method	4
1.4 Final remarks	6
2 Runge-Kutta and Linear Multistep Methods	7
2.1 General Runge-Kutta methods	7
2.1.1 Embedded RK	8
2.2 Linear multistep methods	9
2.2.1 Generalities	9
2.2.2 Predictor-Corrector method	9
2.2.3 Richardson's extrapolation	9
2.2.4 Convergence of a linear multistep method	9
3 Stiff Problems	12

Partial Differential Equations

4 Partial Differential Equations. Generalities on their solution	17
4.1 Finite differences	17
4.1.1 Numerical derivatives	18
4.1.2 Forward Time Centered Space method (FTCS)	20
5 Numerical Solution of PDEs with the Finite Difference Method	21
6 Introduction to Boundary Value Problems	22
7 Quality Control of Solutions	23

ORDINARY DIFFERENTIAL EQUATIONS

1 Ordinary Differential Equations. Basic concepts

1.1 Introduction and some notation

Given $y' = f(x, y)$, where $\begin{cases} y(x) \in \mathbb{R}^n \\ f : \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n \end{cases}$

Definition. We denote by $y(x)$ the exact solution of the ODE system above.

Definition. y_k is the approximation of $y(x_k)$ (after k steps).

Objective. We want to approximate $y(x)$ within a given interval $[x_0, x_n]$.

We know $\begin{cases} x_0 \\ x_1 \\ x_2 \\ \vdots \\ x_n \end{cases}$ We'd like to know $\begin{cases} y(x_0) \\ y(x_1) \\ y(x_2) \\ \vdots \\ y(x_n) \end{cases}$ We find $\begin{cases} y_0 \\ y_1 \\ y_2 \\ \vdots \\ y_n \end{cases}$ (given by a method)

Definition. $\|y(x_n) - y_n\|$ is the **global error**.

Definition. We define the **local truncation error** as the error caused by one iteration, i.e.

$$LTE = \|y(x_k) - y_k\| \quad (\text{assuming the } \textit{localizing assumption}: y_{k-1} = y(x_{k-1}))$$

1.2 Euler's method

1.3 Enhanced Euler's method

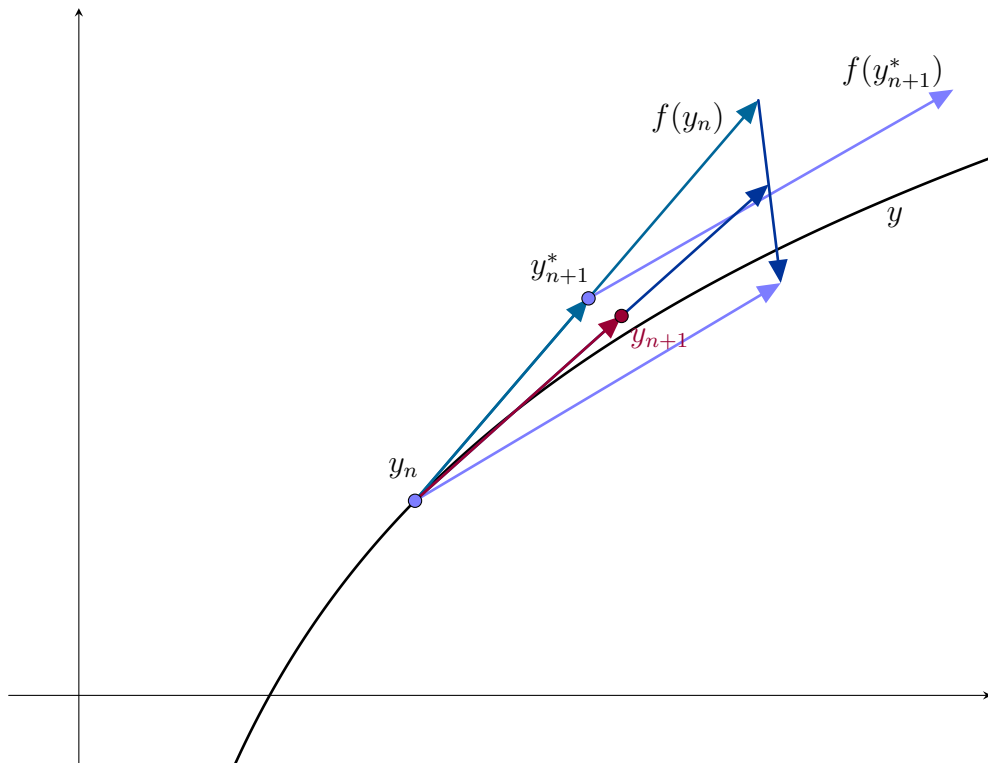


Figure 1.1: One step of the enhanced Euler's method

Given $y'(x) = f(y(x))$, $y : \mathbb{R} \rightarrow \mathbb{R}$, $f : \mathbb{R} \rightarrow \mathbb{R}$, we'll go through the steps to deduce the enhanced Euler's method with the help of the scheme in Figure 1.1.

The auxiliary point y_{n+1}^* can be found doing one step of the standard Euler's method, so

$$y_{n+1}^* = y_n + h \cdot f(y_n)$$

To get the point y_{n+1} we compute the average vector of $f(y_{n+1}^*)$ and $f(y_n)$, and with this new vector, we can apply again a step of Euler's method, ending up with our method

$$y_{n+1} = y_n + \frac{h}{2} \cdot (f(y_{n+1}^*) + f(y_n))$$

Let's find the LTE:

The Local Truncation Error is given by:

$$LTE = \|method - exact\ solution\|$$

Then

$$LTE = \|y_{n+1} - y(x_{n+1})\| = \left\| y_n + \frac{h}{2}f(y_n) + \frac{h}{2}f(y_{n+1}^*) - \underbrace{y(x_{n+1})}_{y(x_n+h)} \right\| = (*)$$

Applying Taylor on $f(y_{n+1}^*) = f(y_n + hf(y_n))$, we have

$$f(y_n + hf(y_n)) = f(y_n) + hf(y_n)f'(y_n) + \mathcal{O}(h^2)$$

and on $y(x_n + h)$ we have

$$y(x_n + h) = y(x_n) + h \cdot y'(x_n) + \frac{h^2}{2} \cdot y''(x_n) + \mathcal{O}(h^3)$$

(In this case, we expand to second order for later simplifications)

$$(*) = \left\| y_n + \frac{h}{2}f(y_n) + \frac{h}{2}(f(y_n) + hf(y_n)f'(y_n) + \mathcal{O}(h^2)) - \left(y(x_n) + h \cdot y'(x_n) + \frac{h^2}{2} \cdot y''(x_n) + \mathcal{O}(h^3) \right) \right\| \quad (1)$$

Now, given $y'(x) = f(y(x))$, we have

$$\begin{aligned} y''(x) &= f'(y(x)) \cdot y'(x) \\ &= f'(y(x)) \cdot f(y(x)) \end{aligned}$$

and we can rewrite the following expression as:

$$\frac{h^2}{2}f(y(x))f'(y(x)) = \frac{h^2}{2}y''(x)$$

With that and the localising assumption ($y_n = y(x_n)$), we can simplify most of the terms in (1) and we end up with

$$\left\| \cancel{y_n} + \cancel{\frac{h}{2}f(y_n)} + \cancel{\frac{h}{2}f(y_n)} + \cancel{\frac{h^2}{2}f(y_n)f'(y_n)} + \mathcal{O}(h^3) - \left(\cancel{y(x_n)} + \cancel{h \cdot y'(x_n)} + \cancel{\frac{h^2}{2} \cdot y''(x_n)} + \mathcal{O}(h^3) \right) \right\| = \mathcal{O}(h^3)$$

So $LTE = \mathcal{O}(h^3)$

Remark. Of course, this method also works for $y : \mathbb{R} \rightarrow \mathbb{R}^n$, $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$

1.4 Final remarks

- There's an improved Euler's method of order 2 similar to the previous one:

$$y_{n+1} = y_n + hf\left(\frac{y_{n+1}^* + y_n}{2}\right)$$

- If we have a method of order ≥ 2 , and we want the value of $y(x^*)$, with x^* off the mesh ($x^* \neq k \cdot h$), we have some options:

-Step back and take a step with the right h .

-Interpolate with the right order.

-Use a continuous Runge-Kutta method.

2 Runge-Kutta and Linear Multistep Methods

2.1 General Runge-Kutta methods

Runge-Kutta methods are a family of iterative methods, which include the previously seen Euler's Method. Let's define an RK method with s stages:

Given $x \in \mathbb{R}$, $y \in \mathbb{R}^n$, $y' = f(x, y)$

Definition. The **Butcher Tableau** is

$$\begin{array}{c|cccc} c_1 & a_{11} & a_{12} & \dots & a_{1s} \\ c_2 & a_{21} & a_{22} & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ c_s & a_{s1} & \dots & \dots & a_{ss} \\ \hline & b_1 & \dots & \dots & b_s \end{array}$$

With these coefficients given by the table, we can now define our Runge-Kutta method:

$$\left. \begin{array}{l} k_1 = f\left(x_n + c_1 h, y_n + h(a_{11}k_1 + a_{12}k_2 + \dots + a_{1s}k_s)\right) \\ k_2 = f\left(x_n + c_2 h, y_n + h(a_{21}k_1 + a_{22}k_2 + \dots + a_{2s}k_s)\right) \\ \vdots \\ k_s = f\left(x_n + c_s h, y_n + h(a_{s1}k_1 + a_{s2}k_2 + \dots + a_{ss}k_s)\right) \end{array} \right\} \begin{array}{l} \text{System of equations} \\ \text{with unknowns } k_1, k_2, \dots, k_s \end{array}$$

$$\boxed{y_{n+1} = y_n + h(b_1k_1 + b_2k_2 + \dots + b_s k_s)}$$

Cases:

- (1) Explicit ($a_{ij} = 0$ for $j \geq i$)
- (2) Semi-implicit ($a_{ij} = 0$ for $j > i$)
- (3) Implicit

Remark. (2) and (3) are used for stiff problems.

Theorem. An explicit s -stage Runge-Kutta method can't have order > 5 .

Theorem. There is no explicit 5-stage RK of order 5.

Theorem.

Let

$A = \text{order } p \text{ for } y' = f(y), f : \mathbb{R}^m \rightarrow \mathbb{R}^m, m > 1$

$B = \text{order } p \text{ for } y' = f(x, y), f : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$

$C = \text{order } p \text{ for } y' = f(y), f : \mathbb{R} \rightarrow \mathbb{R}$

Then

For $1 \leq p \leq 3$, $A \iff B \iff C$

For $p = 4$, $A \iff B \implies C$ but $C \not\Rightarrow B$

For $p \geq 5$, $A \implies B \implies C$ but $C \not\Rightarrow B$, $B \not\Rightarrow A$

2.1.1 Embedded RK

2.2 Linear multistep methods

2.2.1 Generalities

2.2.2 Predictor-Corrector method

2.2.3 Richardson's extrapolation

2.2.4 Convergence of a linear multistep method

Let's see an example of divergence using a linear multistep method:

Example

Given the method

$$y_{n+2} + a_1 y_{n+1} + a_0 y_n = h(b_1 f_{n+1} + b_0 f_n)$$

1) Find a_0, a_1, b_0, b_1 so that the method above has the highest possible order.

2) Try it on

$$\begin{cases} y' = -y \\ y(0) = 1 \end{cases} \quad (y_0 = 1, y_1 = e^{-h})$$

and prove the method diverges.

1) We want $y(x_n + 2h) - y_{n+2}$

We assume $y_{n+1} = y(x_n + h), y_n = y(x_n)$ (localizing assumption).

$$\begin{aligned} y(x_n + 2h) - y_{n+2} &= y(x_n + 2h) - \left[-a_1 y_{n+1} - a_0 y_n + h(b_1 f(y_{n+1}) + b_0 f(y_n)) \right] \underset{\text{loc.as.}}{=} \\ &= y(x_n + 2h) - \left[-a_1 y(x_n + h) - a_0 y(x_n) + h b_1 \underbrace{f(y(x_n + h))}_{y'(x_n+h)} + h b_0 \underbrace{f(y(x_n))}_{y'(x_n)} \right] \end{aligned}$$

As usual, we expand in powers of h . We'll expand to order 3

$$\begin{aligned} &y(x_n) + 2h y'(x_n) + \frac{4h^2}{2} y''(x_n) + \frac{8h^3}{6} y'''(x_n) + o(h^4) - \\ &- \left[-a_1 \left(y(x_n) + h y'(x_n) + \frac{h^2}{2} y''(x_n) + \frac{h^3}{6} y'''(x_n) + o(h^4) \right) - \right. \\ &\quad - a_0 y(x_n) \\ &\quad + h b_1 \left(y'(x_n) + h y''(x_n) + \frac{h^2}{2} y'''(x_n) + o(h^3) \right) + \\ &\quad \left. + h b_0 y'(x_n) \right] \end{aligned}$$

Let's group by powers of h and assume the right conditions to obtain the highest possible order:

$$h^0 \longrightarrow y(x_n) + a_1 y(x_n) + a_0 y(x_n) = 0$$

$$h^1 \longrightarrow 2hy'(x_n) + a_1 hy'(x_n) - hb_1 y'(x_n) - hb_0 y'(x_n) = 0$$

$$h^2 \longrightarrow 2h^2 y''(x_n) + a_1 \frac{1}{2} h^2 y''(x_n) - b_1 h^2 y''(x_n) = 0$$

$$h^3 \longrightarrow \frac{8h^3}{6} y'''(x_n) + a_1 \frac{h^3}{6} y'''(x_n) - b_1 h \left(\frac{h^2}{2} y'''(x_n) \right) = 0$$

With that, we get the system of equations

$$\begin{cases} 1 + a_1 + a_0 = 0 \\ 2 + a_1 - b_1 - b_0 = 0 \\ 2 + \frac{a_1}{2} - b_1 = 0 \\ \frac{8}{6} + \frac{a_1}{6} - \frac{b_1}{2} = 0 \end{cases}$$

And we end up with

$$a_0 = -5, \quad a_1 = 4, \quad b_0 = 2, \quad b_1 = 4$$

2) Our method is

$$y_{n+2} + 4y_{n+1} - 5y_n = h(4f_{n+1} + 2f_n)$$

and with

$$\begin{cases} y' = -y \\ y(0) = 1, y(h) = e^{-h} \end{cases} \quad (y(x) = e^{-x})$$

we have

$$y_{n+2} + 4y_{n+1} - 5y_n = h(-4y_{n+1} - 2y_n)$$

We'll find a solution of the form

$$y_n = c_1 ()^n + c_2 ()^n$$

and we'll see that it diverges.

$$\begin{aligned} \lambda^2 + 4\lambda - 5 + 4h\lambda + 2h &= 0 \\ \lambda^2 + (4(1+h))\lambda + (2h-5) &= 0 \end{aligned}$$

$$\lambda = \frac{-4(1+h) \pm \sqrt{4^2(1+h)^2 - 4(2h-5)}}{2}$$

Let's expand the discriminant

$$\begin{aligned} \sqrt{4^2(1+2h+h^2) - 8h + 20} &= \sqrt{36 + 24h + 16h^2} = 6\sqrt{1 + \frac{4}{6}h + \frac{4^2}{6^2}h^2} \stackrel{\text{Taylor}}{=} \\ &= 6 \left(1 + \frac{1}{2} \left(\frac{4}{6}h + \frac{4^2}{6^2}h^2 \right) + o(h^2) \right) = \\ &= 6 \left(1 + \frac{1}{3}h + o(h^2) \right) \end{aligned}$$

So

$$\lambda = \frac{-4 - 4h \pm (6 + 2h + o(h^2))}{2} = \begin{cases} 1 - h + o(h^2) \\ -5 - 3h + o(h^2) \end{cases}$$

$$\implies y_n = c_1(1 - h + o(h^2))^n + c_2(-5 - 3h + o(h^2))^n$$

Let's find c_1 and c_2 imposing the initial conditions

$$\begin{aligned} &\begin{cases} 1 = c_1 + c_2 \implies c_1 = 1 - c_2 \\ e^{-h} = c_1(1 - h + o(h^2)) + c_2(-5 - 3h + o(h^2)) \end{cases} \\ \implies e^{-h} &= (1 - h + o(h^2)) + c_2 \underbrace{(-5 - 3h + o(h^2) - 1 + h + o(h^2))}_{-6 - 2h + o(h^2)} \\ \implies c_2 &= \frac{e^{-h} - 1 + h + o(h^2)}{-6 - 2h + o(h^2)} \stackrel{\text{Taylor } e^{-h}}{=} \frac{1 - h + o(h^2) - 1 + h + o(h^2)}{-6 - 2h + o(h^2)} \simeq \frac{-1}{6 + 2h} \\ \implies c_1 &\simeq 1 + \frac{1}{6 + 2h} = \frac{7 + 2h}{6 + 2h} \end{aligned}$$

So

$$y_n = \frac{7 + 2h}{6 + 2h}(1 - h + o(h^2))^n + \frac{-1}{6 + 2h}(-5 - 3h + o(h^2))^n$$

and the term $(-5)^n$ will cause the solution to diverge.

3 Stiff Problems

In some ODEs, the step size taken by an adaptive method is forced to be unreasonably small even in regions where the solution curve is smooth. In these cases, it takes a large amount of steps to go through a short time interval.

These types of equations are called **stiff ODEs**.

Example (Van der Pol equation)

Given the Van der Pol equation

$$\ddot{x} - \mu(1 - x^2)\dot{x} + x = 0$$

the larger the constant μ , the stiffer is the problem.

Trying to solve it using an explicit adaptive stepsize method like Matlab's `ode45` yields

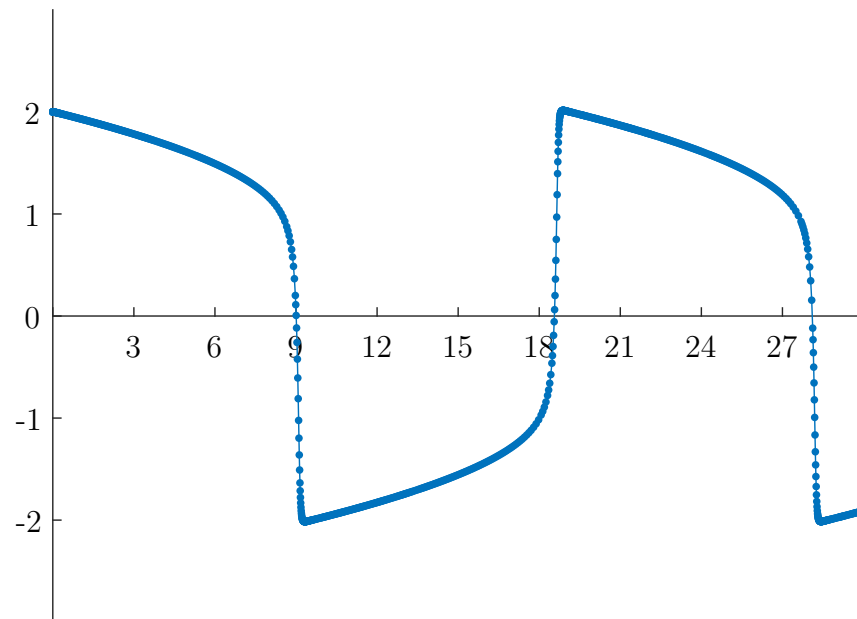


Figure 3.1: Van der Pol equation solution with `ode45` ($\mu = 10$)

With 873 steps needed, and a minimum stepsize of $2.5119 \cdot 10^{-5}$.

Now, using an implicit method like Matlab's `ode15s`, we have

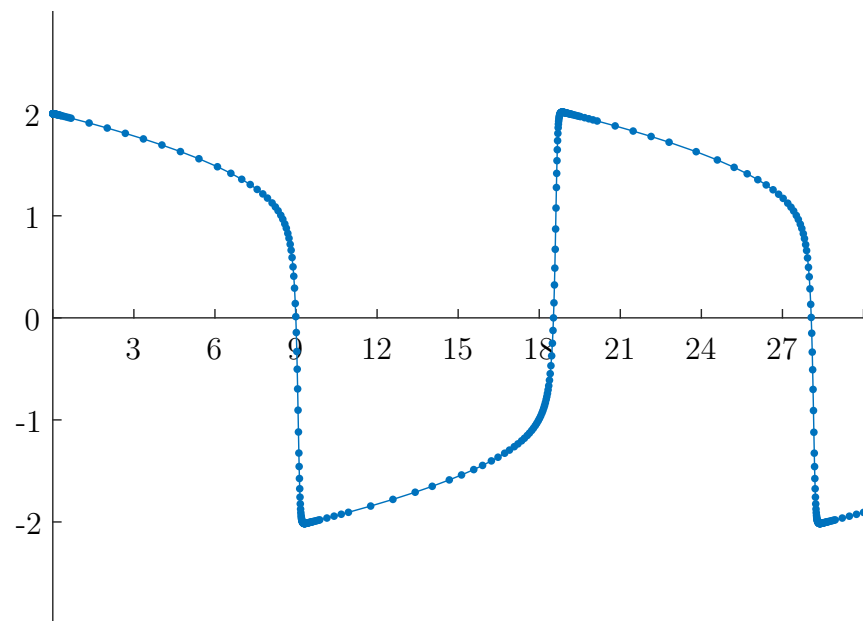


Figure 3.2: Van der Pol equation solution with `ode15s` ($\mu = 10$)

Which clearly uses less steps to pass through the stiff areas (a total of 326 with minimum stepsize 0.00014607).

If we where to solve it with a larger μ , for example $\mu = 1000$, the number of steps needed using `ode45` is 5.495.393 which is too much compared to the 586 needed with `ode15s`.

Definition. The set of values of the stepsize h such that $\lim_{h \rightarrow 0} y_n = 0$ is called the **absolute stability region**.

Let's see what happens with a general linear multistep method applied to a linear system:

Our method is

$$\underbrace{\sum_{j=0}^k \alpha_j y_{n+j}}_{\rho} = h \underbrace{\sum_{j=0}^k \beta_j f_{n+j}}_{\sigma}$$

We apply it to

$$y' = Ay$$

where we'll assume that the eigenvalues of A ($\lambda_1, \dots, \lambda_n$) are all different (so that it's diagonalizable) and have negative real parts.

$y = e^{Ax}$ is the fundamental solution, and if $y(0) = y_0$, the solution is

$$y(x) = e^{Ax} y_0$$

with

$$\lim_{x \rightarrow \infty} e^{Ax} = \begin{pmatrix} 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \end{pmatrix}$$

Now, with the system $y' = Ay$, our general method looks like

$$\sum_{j=0}^k \alpha_j y_{n+j} = h \sum_{j=0}^k \beta_j A y_{n+j}$$

Let's rewrite it:

$$\begin{aligned} & \sum_{j=0}^k (\alpha_j I - h \beta_j A) y_{n+j} = 0 \\ & \Downarrow \\ & \sum_{j=0}^k \left(\alpha_j I - h \beta_j \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} \right) y_{n+j} = 0 \\ & \Downarrow \\ & \sum_{j=0}^k (\alpha_j I - h \beta_j \lambda_i) y_{n+j} = 0 \quad (\forall i = 1, \dots, n) \end{aligned}$$

A finite difference equation bla bla bla

PARTIAL DIFFERENTIAL EQUATIONS

4 Partial Differential Equations. Generalities on their solution

4.1 Finite differences

Example (1D Heat equation)

Our problem is:

$$\begin{cases} u_t - u_{xx} = f \\ u(x, 0) = u_0(x) \\ \left. \begin{array}{l} u(a, t) = u_a \\ u(b, t) = u_b \end{array} \right\} \end{cases} \quad \begin{array}{l} \leftarrow \text{Initial condition (IC)} \\ \leftarrow \text{Boundary conditions (BC)} \end{array}$$

With $t \geq 0$, $x \in [a, b]$

We discretize x and t :



Idea: We'll impose $U_t(x_i, t^n) - U_{xx}(x_i, t^n) = f(x_i, t^n)$

4.1.1 Numerical derivatives



We'll use Taylor's expansion on f to find an approximation for the derivatives:

$$(1) \quad f_{i+1} = f_i + hf'_i + \frac{h^2}{2}f''_i + \frac{h^3}{3!}f'''_i + \dots$$

$$(2) \quad f_{i-1} = f_i - hf'_i + \frac{h^2}{2}f''_i - \frac{h^3}{3!}f'''_i + \dots$$

Approximation of the first derivative:

- Forward differences (1)

$$f'_i = \frac{f_{i+1} - f_i}{h} + \mathcal{O}(h)$$

- Backwards differences (2)

$$f'_i = \frac{f_i - f_{i-1}}{h} + \mathcal{O}(h)$$

- Centered differences (1)-(2)

$$f'_i = \frac{f_{i+1} - f_{i-1}}{2h} + \mathcal{O}(h^2)$$

Remark. With centered differences, it converges faster, but sometimes the other two methods are preferable for ease in computations.

Approximation of the second derivative (1)+(2)

$$f_i'' = \frac{f_{i+1} - 2f_i + f_{i-1}}{h^2} + \mathcal{O}(h^2)$$

There are higher order approximations, but these are the standard, and are more than enough.

Now, we'll present some notation to differentiate between the problem we want to solve and the numerical problem.

Definition.

The **continuous problem** is
$$\begin{cases} \mathcal{L}(u) = f & \text{in } \Omega, t > 0 \\ \mathcal{B}(u) = q & \text{in } \partial\Omega, t > 0 \\ u(x, 0) = u_0(x) \end{cases}$$

Where \mathcal{L} is any differential operator.

It's what we want to solve.

The **discrete problem** is
$$\begin{cases} L(u) + \tau = f \\ B(u) + \bar{\tau} = q \\ u(x, 0) = u_0(x) \end{cases}$$

The **numerical problem** is
$$\begin{cases} L(U) = f \\ B(U) = q \\ U(x, 0) = u_0(x) \end{cases}$$

Let's go back to our 1D heat equation:

The continuous problem is

$$\begin{cases} u_t - u_{xx} = f & t > 0, x \in (a, b) \\ u(a, t) = u_a \\ u(b, t) = u_b \\ u(x, 0) = u_0(x) \end{cases}$$

Notation: $u(x_i, t^n) = u_i^n$

$x_i = a + i\Delta x \quad i = 0, \dots, N \quad x_0 = a, x_N = b \quad (\text{sometimes } \Delta x = h).$

$t^n = n\Delta t \quad n \geq 0$

We have the following explicit method:

4.1.2 Forward Time Centered Space method (FTCS)

- Approximate u_t using forward differences (FT)
- Approximate u_{xx} using centered differences (CD)

Our discretized problem is

$$\frac{u_i^{n+1} - u_i^n}{\Delta t} + \mathcal{O}(\Delta t) - \frac{u_{i+1}^n - u_i^n + u_{i-1}^n}{\Delta x^2} + \mathcal{O}(\Delta x^2) = f_i^n$$

We define $\tau = \mathcal{O}(\Delta t, \Delta x^2)$

Neglecting τ , the numerical problem we want to solve is

$$\begin{cases} \frac{U_i^{n+1} - U_i^n}{\Delta t} - \frac{U_{i+1}^n - U_i^n + U_{i-1}^n}{\Delta x^2} = f_i^n \\ U_0^{n+1} = u_a(t^{n+1}) \\ U_N^{n+1} = u_b(t^{n+1}) \\ U_i^0 = u_0(x_i) \end{cases}$$

Now we define r as

$$r = \frac{\Delta t}{\Delta x^2}$$

So

$$U_i^{n+1} = rU_{i-1}^n + (1 - 2r)U_i^n + rU_{i+1}^n + \Delta t f_i^n \quad i = 1, \dots, N - 1$$

For $i = 1$,

$$U_1^{n+1} = (1 - 2r)U_1^n + rU_2^n + \Delta t f_1^n + ru_a^n$$

And for $i = N - 1$,

$$U_{N-1}^{n+1} = rU_{N-2}^n + (1 - 2r)U_{N-1}^n + \Delta t f_{N-1}^n + ru_b^n$$

Thus,

$$\overline{U^{n+1}} = \underset{\sim}{B} \overline{U^n} + \overline{F} + \overline{G}$$

Where

$$\overline{U^p} = (U_1^p, U_2^p, \dots, U_{N-1}^p)^T \quad \overline{F} = \Delta t (f_1^n, f_2^n, \dots, f_{N-1}^n)^T \quad \overline{G} = (ru_a^n, 0, \dots, 0, ru_b^n)^T$$

$$\underset{\sim}{B} = \begin{pmatrix} 1-2r & r & & & \\ r & 1-2r & r & & \\ & \ddots & \ddots & \ddots & \\ & & r & 1-2r & r \\ & & & r & 1-2r \end{pmatrix}$$

5 Numerical Solution of PDEs with the Finite Difference Method

6 Introduction to Boundary Value Problems

7 Quality Control of Solutions