Theoretical exercise sheet 3
Ordinary differential equations

Exercises 2, 3 and 5 (marked *) to be submitted via Moodle in pdf format (either handwritten and scanned, or typeset using LaTeX).
Some subset of these questions will be assessed.
Deadline: 10pm UK time Sunday 13th December.

EXERCISE 1 Consider the linear Cauchy problem:

$$\begin{cases} y'(t) = -e^t y(t), & t \in [0,1], \\ y(0) = 2. \end{cases}$$

(a) Write down the forward Euler method for the approximation of the solution $y(t)$.

(b) Let $h = \frac{1}{10}$. Compute the approximate solution at time $t_1 = t_0 + h$ (where $t_0 = 0$) using the forward Euler method.

(c) Now consider the nonlinear Cauchy problem:

$$\begin{cases} y'(t) = -e^t y^2(t), & t \in [0,1], \\ y(0) = 2. \end{cases}$$

Write down the backward Euler method for the approximation of the solution $y(t)$.

(d) Rewrite the backward Euler method in the form

$$F(u_{n+1}; t_n, u_n, h) = 0,$$

and write down Newton's method for solving this nonlinear equation.

EXERCISE 2(*) Consider the following differential equation:

$$\begin{cases} y'(t) = -C \arctan(ky), & t > 0 \\ y(0) = y_0 \ , \end{cases} \tag{1}$$

where C and k are given real positive constants.

(a) Write the backward Euler scheme for solving (1) in the form

$$u_{n+1} = g(u_n, u_{n+1}, h), \tag{2}$$

specifying the function g, where h is the timestep and u_n the approximation of $y(t_n)$.

(b) For each timestep one has to solve the nonlinear equation (2). Interpret this equation as a fixed point problem for the computation of u_{n+1}, and determine a condition on h which guarantees that there exists a unique fixed point to which the fixed point iteration $x^{k+1} = \phi(x^k)$ converges for any initial guess.

(c) Write down the Newton iteration for the solution of the nonlinear equation (2).

EXERCISE 3(*) For the numerical solution of the Cauchy problem

$$\begin{cases} y'(t) = f(t, y(t)), \quad t \in (0, T], \quad 0 < T < +\infty, \\ y(0) = y_0, \end{cases}$$

where f is assumed to be uniformly Lipschitz continuous with respect to its second argument, consider the following method for $n \geq 0$, with $u_0 = y_0$, and uniform step size $h = T/N$ for some $N \in \mathbb{N}$:

$$\begin{cases} u_{n+\frac{1}{2}} = u_n + \dfrac{h}{2} f(t_n, u_n), \\ p_n = f(t_n + \dfrac{h}{2}, u_{n+\frac{1}{2}}), \\ u_{n+1} = u_n + h p_n. \end{cases} \tag{3}$$

(a) Is the method given by (3) explicit or implicit ?

(b) Write the method in the form

$$u_{n+1} = u_n + h\Psi(t_n, u_n, u_{n+1}; h)$$

for an appropriate increment function Ψ.

(c) Prove that the method is zero-stable. Reference carefully any theorems you use from lectures.

(d) Write down the formula for the truncation error of the method, and show that the method is consistent, with $T_n = O(h^2)$ as $h \to 0$ for y and f sufficiently smooth. Under what smoothness assumptions on y and f is your analysis valid?

Hint: You may find it helpful to recall the "generalised chain rule"

$$\frac{d}{dh} F(T(h), Y(h)) = T'(h) \frac{\partial F}{\partial t}(T(h), Y(h)) + Y'(h) \frac{\partial F}{\partial y}(T(h), Y(h)).$$

(e) By quoting an appropriate theorem from lectures, combine the results of parts (c) and (d) to prove that the method is second order convergent.

EXERCISE 4 Consider the Cauchy problem

$$\begin{cases} y'(t) = 1 - y^2, \quad t > 0, \\ y(1) = (e-1)/(e+1). \end{cases}$$

(a) By deriving an exact solution and considering the behaviour of $\partial f/\partial y$ on the solution trajectory (or otherwise), determine an estimate for the critical value of h below which perturbations due to roundoff errors are controlled when the forward Euler method is used.

(b) Now do the same for Heun's method.

EXERCISE 5(*) Consider a two-by-two system of differential equations

$$\begin{cases} \mathbf{w}'(t) = A\mathbf{w}(t), \\ \mathbf{w}(0) = \mathbf{w}_0, \end{cases} \qquad \mathbf{w}(t) = \begin{bmatrix} w_1(t) \\ w_2(t) \end{bmatrix}, \tag{4}$$

where A is a given 2×2 matrix.

(a) Write down the forward Euler method, the backward Euler method and the Crank-Nicolson method for the system (4).

(b) Suppose that you are given the diagonalization $A = VDV^{-1}$, where D is a diagonal matrix containing the eigenvalues d_1 and d_2 of A, and V is an invertible matrix whose columns are the corresponding eigenvectors. Show how the problem (4) and the three schemes in part (a) can be rewritten in terms of the diagonal matrix D and the transformed unknowns $\mathbf{x} = V^{-1}\mathbf{w}$ and $\mathbf{x}^n = V^{-1}\mathbf{w}^n$.

(c) Using the transformations in (b), determine for which step sizes $h > 0$ the three schemes are absolutely stable, under the assumption that the eigenvalues satisfy $d_1, d_2 < 0$.

(d) Now consider the particular system of differential equations:

$$\begin{cases} w_1'(t) = w_2(t), & t > 0, \\ w_2'(t) = -\lambda\, w_1(t) - \mu\, w_2(t), & t > 0, \\ w_1(0) = w_{1,0}, & \\ w_2(0) = w_{2,0}, & \end{cases} \tag{5}$$

where λ and μ are two positive real numbers such that $\mu^2 - 4\lambda > 0$.

Write the system (5) in the form (4), specifying the matrix A. Use your results in (c) to determine the stability of the three schemes in this case. What is the stability condition for the forward Euler method in the special case $\lambda = 6$ and $\mu = 5$?

EXERCISE 6 Let A be the $(N-1)$-by-$(N-1)$ matrix defined by

$$A = \frac{1}{h^2}\begin{pmatrix} 2 & -1 & 0 & \cdots & \cdots & 0 \\ -1 & 2 & -1 & & & \vdots \\ 0 & -1 & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & & -1 & 0 \\ \vdots & & & -1 & 2 & -1 \\ 0 & \cdots & \cdots & 0 & -1 & 2 \end{pmatrix} + \begin{pmatrix} r_1 & 0 & \cdots & & \cdots & 0 \\ 0 & r_2 & 0 & & & \vdots \\ \vdots & 0 & \ddots & \ddots & & \\ & & \ddots & & 0 & \vdots \\ \vdots & & & 0 & r_{N-2} & 0 \\ 0 & \cdots & & \cdots & 0 & r_{N-1} \end{pmatrix},$$

where $h > 0$ and $r_n > 0$ for all $n = 1, \ldots, N-1$.

Show that, for any $\mathbf{0} \neq \mathbf{x} = (x_1, \ldots, x_N)^T \in \mathbb{R}^{N-1}$,

$$\mathbf{x}^T A \mathbf{x} \geq x_1^2 + x_{N-1}^2 + \sum_{n=2}^{N-1} (x_n - x_{n-1})^2,$$

and explain why this proves that the matrix A is invertible.

EXERCISE 7 Prove the following *discrete maximum principle*:

Let a_n, b_n, c_n, $n = 1, \ldots, N-1$ be positive real numbers such that

$$b_n \geq a_n + c_n, \tag{6}$$

and let U_n, $n = 0, \ldots, N$, be real numbers such that

$$-a_n U_{n-1} + b_n U_n - c_n U_{n+1} \leq 0, \qquad n = 0, \ldots, N. \tag{7}$$

Then

$$U_n \leq \max\{U_0, U_N, 0\}, \qquad n = 0, \ldots, N.$$

Hint: Let $U_r = \max_{n=0,\ldots,N} |U_n|$. *If* $r = 0$ *or* $r = N$, *or* $U_r \leq 0$, *then we are done. Otherwise, suppose that* $1 \leq r \leq N_1$ *and that* $U_r > 0$. *Use* (6) *and* (7) *to show that* $U_r = U_{r-1} = U_{r+1}$. *Then show by repeating this argument that we must have either* $U_r = U_0$ *or* $U_r = U_N$.