1.2 Ordinary Differential Equations

This project builds on material covered in the Part IA lectures on Computational Projects, see $http://www.maths.cam.ac.uk/undergrad/catam/part-ia-lectures$. The Part IA Differential Equations and Part IB Methods courses are also relevant.

1   Background Theory

This project is concerned with the numerical step-by-step integration of ordinary differential equations (ODEs) of the form

$$\frac{d\mathbf{y}}{dx} = \mathbf{f}(x, \mathbf{y}), \tag{1a}$$

where $\mathbf{y}$ and $\mathbf{f}$ are vectors of length m, subject to an initial condition

$$\mathbf{y} = \mathbf{y}_0 \quad \text{at } x = x_0 \tag{1b}$$

for some constants x_0 and $\mathbf{y}_0$. The exact solution of (1a)–(1b) will be denoted by $\mathbf{y} = \mathbf{y}_e(x)$. In the first part of the project (§2) the performance of three different numerical methods will be examined. A first-order equation ($m = 1$) has been chosen for which $y_e(x)$ has a known analytic form for comparison. In the second part (§3) one of the methods is extended to solve a second-order problem.

The numerical methods to be investigated are as follows.

(a) The **Euler** method, or more precisely the **forward Euler** method, is the simple scheme

$$\mathbf{Y}_{n+1} = \mathbf{Y}_n + h\,\mathbf{f}(x_n, \mathbf{Y}_n), \tag{2}$$

where $\mathbf{Y}_n$ denotes the numerical solution at $x_n \equiv x_0 + nh$, that is, at the nth step with step length h. The Euler method is called a *single-step* method since $\mathbf{Y}_{n+1}$ is determined solely by the value $\mathbf{Y}_n$ at the previous step.

Definition. The *global error* after the nth step is defined as

$$\mathbf{E}_n = \mathbf{Y}_n - \mathbf{y}_e(x_n). \tag{3a}$$

Definition. The *local error* of the first step is defined as $\mathbf{e}_1 = \mathbf{Y}_1 - \mathbf{y}_e(x_1)$. For subsequent steps the *local error* is defined as

$$\mathbf{e}_n = \mathbf{Y}_n - \tilde{\mathbf{y}}(x_n), \tag{3b}$$

where $\mathbf{y} = \tilde{\mathbf{y}}(x)$ is the exact solution to equation (1a) with $\mathbf{y} = \mathbf{Y}_{n-1}$ at $x = x_{n-1}$ (note that, in general, $\mathbf{Y}_{n-1} \neq \mathbf{y}_e(x_{n-1})$ for $n > 1$).

A numerical method is said to be p^{th}-*order accurate* if $\mathbf{e}_n$ is $O\left(h^{p+1}\right)$ as $h \to 0$. It can be shown that the Euler method is first-order accurate.

(b) The second-order-accurate **Adams-Bashforth** (AB2) method employs the scheme

$$\mathbf{Y}_{n+1} = \mathbf{Y}_n + h\left[\tfrac{3}{2}\mathbf{f}\left(x_n, \mathbf{Y}_n\right) - \tfrac{1}{2}\mathbf{f}\left(x_{n-1}, \mathbf{Y}_{n-1}\right)\right]. \tag{4}$$

This is a *two-step* method, using both $\mathbf{Y}_{n-1}$ and $\mathbf{Y}_n$ to obtain $\mathbf{Y}_{n+1}$, and the first step must be taken by a single-step method, such as the Euler method.

(c) The fourth-order-accurate **Runge–Kutta** (RK4) method employs the scheme:

$$\mathbf{Y}_{n+1} = \mathbf{Y}_n + \tfrac{1}{6}h\left[\mathbf{k}_1 + 2\mathbf{k}_2 + 2\mathbf{k}_3 + \mathbf{k}_4\right], \tag{5a}$$

where

$$\mathbf{k}_1 = \mathbf{f}(x_n,\ \mathbf{Y}_n), \tag{5b}$$
$$\mathbf{k}_2 = \mathbf{f}(x_n + \tfrac{1}{2}h,\ \mathbf{Y}_n + \tfrac{1}{2}h\mathbf{k}_1), \tag{5c}$$
$$\mathbf{k}_3 = \mathbf{f}(x_n + \tfrac{1}{2}h,\ \mathbf{Y}_n + \tfrac{1}{2}h\mathbf{k}_2), \tag{5d}$$
$$\mathbf{k}_4 = \mathbf{f}(x_n + h,\ \mathbf{Y}_n + h\mathbf{k}_3). \tag{5e}$$

The theoretical background for the stability and accuracy of these methods is set out in, for example, *An Introduction to Numerical Methods and Analysis* by J.F.Epperson, *An Introduction to Numerical Methods* by A.Kharab and R.B.Guenther and *Numerical Recipes* by Press *et al.*

2   Stability and accuracy of the numerical methods

The example to be studied in detail in this section is the scalar version of equation (1a) with

$$f(x, y) = -8y + 6e^{-2x} \tag{6a}$$

and initial condition

$$y = 0 \quad \text{at } x = 0.\,. \tag{6b}$$

This has the exact solution

$$y = y_e(x) \equiv e^{-2x} - e^{-8x}. \tag{7}$$

Programming Task: Write programs to apply each of the methods (a), (b) and (c) to this problem.

2.1   Stability

This subsection considers the stability of the AB2 method (with first step by Euler).

Question 1   Starting with $Y_0 = 0$, compute Y_n for x up to 3 with $h = 0.5$ (i.e. for n up to $3/h = 6$). Tabulate the values of x_n, the numerical solution Y_n, the analytic solution $y_e(x_n)$ from (7) and the global error $E_n = Y_n - y_e(x_n)$. You should find that the numerical result is unstable: the error oscillates wildly and its magnitude ultimately grows proportional to $e^{\gamma x}$, where the 'growth rate' γ is a positive constant which you should estimate.

Repeat with $h = 0.375$, 0.25, 0.125, 0.1 and 0.05 (you need only present a judicious selection of output to illustrate the behaviour). What effect does reducing h have on the instability and its growth rate?

Question 2
(i) Find the analytic solution of the AB2 *difference* equation

$$Y_{n+1} = Y_n + h\left[-12Y_n + 9\left(e^{-2h}\right)^n + 4Y_{n-1} - 3\left(e^{-2h}\right)^{n-1}\right] \tag{8}$$

with

$$Y_0 = 0 \;, \quad Y_1 = 6h \quad \text{(from the Euler method).} \tag{9}$$

(ii) Hence explain why and when instability occurs, and with what growth rate.

(iii) Show that in the limit $h \to 0$, $n \to \infty$ with $x_n \equiv nh$ fixed, the solution of the difference-equation problem (8)–(9) converges to the solution (7) of the differential-equation problem (1a), (6a), (6b).

How (if at all) would the conclusions in (ii) be altered if a more accurate method such as RK4 were used for the first step?

2.2 Accuracy

This subsection considers the accuracy of the three methods.

Question 3 For each of the three methods, integrate with $h = 0.08$ from $x = 0$ to $x = 2$. Tabulate and plot the numerical solution Y_n for each method against x_n, superposing a plot of the exact solution $y_e(x_n)$ given by equation (7).

Question 4 For each of the three methods, tabulate the global error E_n at $x_n = 0.16$ against $h = 0.16/n$ for $n = 2^k$, $k = 0, 1, 2, \ldots, 15$, and plot a log–log graph of $|E_n|$ against h over this range. Comment on the relationship of your results to the theoretical accuracy of the methods.

3 Numerical solutions of second-order ODEs

This section is concerned with small 'normal-mode' oscillations of a non-uniform string of length L under uniform positive tension T_0. The string's transverse displacement, $\eta(\xi, \tau)$, a function of longitudinal distance ξ and time τ, is assumed to satisfy the equation of motion (Newton's Second Law, linearised for small displacements and with gravity neglected)

$$\mu(\xi) \frac{\partial^2 \eta}{\partial \tau^2} = T_0 \frac{\partial^2 \eta}{\partial \xi^2} \,, \tag{10a}$$

where $\mu(\xi)$ is the mass per unit length of the string. If the ends $\xi = 0$ and $\xi = L$ are fixed, the appropriate boundary conditions are

$$\eta(0, \tau) = \eta(L, \tau) = 0 \quad \text{for all } \tau. \tag{10b}$$

Multiplying (10a) by $\partial \eta / \partial \tau$ leads to the energy conservation equation

$$\frac{\mathrm{d}}{\mathrm{d}\tau} \int_{\xi_1}^{\xi_2} \left[\tfrac{1}{2}\mu(\xi) \left(\frac{\partial \eta}{\partial \tau}\right)^2 + \tfrac{1}{2}T_0 \left(\frac{\partial \eta}{\partial \xi}\right)^2 \right] \mathrm{d}\xi = \left[\left(T_0 \frac{\partial \eta}{\partial \xi}\right) \frac{\partial \eta}{\partial \tau} \right]_{\xi_1}^{\xi_2} \quad \text{for any } \xi_1, \xi_2 \in [0, L], \tag{11}$$

i.e. the rate of change of total energy (kinetic plus potential) in the section $[\xi_1, \xi_2]$ is equal to the rate at which work is done at the ends (linearised transverse force times transverse velocity).

The problem admits 'normal-mode' solutions of the form

$$\eta(\xi, \tau) = L y(x) \cos(\omega \tau + \theta) \,, \tag{12}$$

where the angular frequency ω and phase θ are constants, and dimensionless variables are defined by

$$x = \frac{\xi}{L} , \quad m(x) = \frac{\mu(\xi)}{\mu(0)} . \tag{13}$$

From (10a) and (10b), it follows that the dimensionless $y(x)$ satisfies

$$\frac{d^2y}{dx^2} + p^2 m(x) y = 0 \quad \text{and} \quad y(0) = y(1) = 0 , \tag{14}$$

where $p = \omega L \sqrt{\mu(0)/T_0}$ is a dimensionless constant and, without loss of generality, $p \geqslant 0$. The system (14) is an example of a *Sturm-Liouville eigenproblem*: only for a discrete set of values of p, i.e. the *eigenvalues* $0 \leqslant p^{(1)} < p^{(2)} < \ldots$, are there non-zero *eigenfunction* solutions for y.

The remainder of this project specialises to a mass distribution of the form $m(x) = (1 + x)^{-\alpha}$ with α a constant, in which case (14) becomes

$$\frac{d^2y}{dx^2} + p^2 (1 + x)^{-\alpha} y = 0 , \tag{15a}$$

subject to

$$y(0) = y(1) = 0 . \tag{15b}$$

The eigenvalues and eigenfunctions can be found in explicit analytic form only for special values of α, one such being $\alpha = 2$.

> **Question 5**   Find the analytic solution $y = y_e(x)$ of equation (15a) with $\alpha = 2$ subject to the initial condition
>
> $$y = 0 , \quad \frac{dy}{dx} = 1 \quad \text{at } x = 0 \tag{16}$$
>
> for a *general* value of p. [The substitution $1 + x = e^z$ *may* be helpful.] Deduce *carefully* that the smallest (non-negative) eigenvalue of (15a)–(15b) with $\alpha = 2$ is
>
> $$p = p^{(1)} \equiv \left[\tfrac{1}{4} + \left(\frac{\pi}{\ln 2} \right)^2 \right]^{1/2} \tag{17}$$
>
> and write down the general eigenvalue $p^{(k)}$ and corresponding eigenfunction $y^{(k)}(x)$.

Equation (15a) can be solved numerically by noting that it is equivalent to

$$\frac{dy}{dx} = f(x, y, z) \equiv z , \quad \frac{dz}{dx} = g(x, y, z) \equiv -p^2 (1 + x)^{-\alpha} y \tag{18}$$

which has the form (1a) with $\mathbf{y} = (y, z)$ and $\mathbf{f} = (f, g)$, and a numerical approximation $\mathbf{Y_n} = (Y_n, Z_n)$ can be obtained using any of the methods described in §1. Here you should use the RK4 method.

> **Programming Task:** Write a program to compute an RK4 numerical approximation (Y_n, Z_n) to the solution of equations (18) with initial condition $Y_0 = 0$, $Z_0 = 1$ at $x_0 = 0$.

> **Question 6**   Taking $\alpha = 2$, run your program with $p = 4$ and $h = 0.1/2^k$ for $k = 0, 1, 2, \ldots, 12$ in turn, tabulating the numerical solution Y_n at $x_n = 1$ and the global error $Y_n - y_e(1)$ against $h \equiv 1/n$. Repeat with $p = 5$. Do the errors behave as expected?

Programming Task: Write a program to search for eigenvalues using the 'false position' method – a variant of the bisection method where if a root of $\phi(p) = 0$ has been located to the interval (p_1, p_2) with $\phi(p_1)$ and $\phi(p_2)$ having opposite signs, an estimate p_s for the root is calculated using the linear interpolant

$$\Phi(p) \equiv \phi(p_1)\left(\frac{p_2 - p}{p_2 - p_1}\right) + \phi(p_2)\left(\frac{p - p_1}{p_2 - p_1}\right) . \tag{19a}$$

Solving for $\Phi(p) = 0$ gives the estimate

$$p = p_s \equiv \frac{\phi(p_2)\, p_1 - \phi(p_1)\, p_2}{\phi(p_2) - \phi(p_1)} \tag{19b}$$

which is accepted if $|\phi(p_s)| < \epsilon$ where ϵ is a specified small value; if not, the process is iterated with p_1 or p_2 replaced by p_s such that $\phi(p_1)$ and $\phi(p_2)$ still have opposite signs. For the current task, $\phi(p)$ is the numerical (RK4) solution Y_n of (15a) and (16) at $x_n = 1$ obtained with a suitably small value of $h \equiv 1/n$.

Question 7   Taking $\alpha = 2$, run the program to obtain an approximation to the smallest (positive) eigenvalue $p^{(1)}$ of (15a)–(15b) with error no more than $\pm 5 \times 10^{-6}$, using $(4, 5)$ as the initial interval and tabulating all the iterates in your write-up. What values are you using for ϵ and h and why? (If you can justify them without reference to the analytic solution, so much the better.)

The final question addresses the case $\alpha = 10$, which has no simple analytic solution.

Question 8   Find approximations to the five smallest (non-negative) eigenvalues $p^{(k)}$, $k = 1, 2, 3, 4, 5$ of (15a)–(15b) with $\alpha = 10$ correct to within $\pm 5 \times 10^{-6}$, and plot the corresponding eigenfunctions $y^{(k)}(x)$ using the 'energy' normalisation

$$\int_0^1 (1 + x)^{-10} \left[p^{(k)} y^{(k)}(x)\right]^2 \mathrm{d}x = 1 . \tag{20}$$

(If you wish, this integral may be approximated using a black-box integration routine such as the MATLAB function `trapz`.) Explain carefully why you are satisfied that the eigenvalues found are indeed the smallest, and that they have the required accuracy.

Comment on the results, e.g. on the shape of the eigenfunctions, both from a mathematical point of view and in terms of the physical model.

Hint: it *may* be instructive to look at the form of the eigensolutions for larger p, and/or to note that when $p \equiv \delta^{-1}$ is 'large', equation (15a) has solutions of the form $\mathrm{e}^{S(x;\delta)}$ with complex exponent $S(x;\delta) = \delta^{-1} S_0(x) + S_1(x) + O\left(\delta\right)$ (the so-called Liouville-Green or WKB approximation.*)

Additional Reference

Boyce, W. E., and DiPrima, R. C., 2001, *Elementary Differential Equations and Boundary Value Problems*, 7th edition. Publ. John Wiley & Sons Inc.

Project 1.2: Ordinary Differential Equations

Marking Scheme and additional comments for the Project Report

The purpose of these additional comments is to provide guidance on the structure and length of your CATAM report. Use the same concepts to write the rest of the reports. To help you assess where marks have been lost, this marking scheme will be completed and returned to you during Lent Term. You are advised to keep a copy of your write-up in order to correlate your answers to the marks awarded.

Question no.	Marks available[1]	Marks awarded[2]
Programming task *Program:* for instructions regarding printouts and what needs to be in the write-up, refer to the introduction to the manual.		
Question 1 *Tables:* for presentation and layout, refer to the introduction. [*up to six tables and a quarter page of writing*][3]	C2	
Question 2 *Analytic solution:* do not include trivial steps in your working. [*half page*][3]	M4.5	
Question 3 *Graphs:* you may use one graph, or two, or three.	C1	
Question 4 *Graphs:* ditto. *Comments:* what can be said about how the global error E_n for each method varies with h? How is this reflected in the plots? [*quarter page*][3]	C1 M1	
Question 5 *Analytic solutions:* do not include trivial steps in your working [*half page*][3].	M1.5	
Question 6 *Analytic solution and numerical solution compared:* The reason for computing an analytic solution is to check that the program is working correctly ('validation'). [*couple of lines*][3].	C1	
Question 7 *Numerical approximations to the smallest eigenvalue:* Tabulate all the iterates; explain how you have chosen ϵ and h to ensure that the final approximation has the required accuracy. [*table, and quarter page of writing*][3]	C1 M1	
Question 8 *Numerical solutions*: explain how you have located the five smallest eigenvalues, and computed them to the required accuracy. *Comments:* first identify the salient features of the graphs. Then try to explain them using mathematical arguments; link to the theory of the physical system under investigation. [*one or two pages*][3]	C2 M2	
Excellence marks awarded for, *among other things*, mathematical clarity and good, clear output (graphs and tables) — see the introduction to the Project Manual.	E2	
Total Raw Marks	20	
Total Tripos Marks	**40**	

[1] C#, M# and E#: *Computational*, *Mathematical* and *Excellence* marks respectively.

[2] For use by the assessor.

[3] Your aim is to answer succinctly the questions including the graphs and tables, and to make all important points. The length specified here should be sufficient space for you to do this but is not a target.