



# Numerical Analysis and Simulation I: Ordinary Differential Equations (ODEs)

## Sheet 4 - Runge-Kutta Methods and Step Size Control

Return of the Exercise Sheet: Tuesday, November 18th before the lecture

### Exercise 7: Formulation of Runge-Kutta-Methods

For solving the IVP  $y' = f(x, y)$ ,  $y(x_0) = y_0$  ( $y : \mathbb{R} \rightarrow \mathbb{R}^n$ ) numerically, the formula of a general Runge-Kutta method with  $s$  stages reads

$$y(x+h) \doteq y_0 + h \sum_{i=1}^s b_i k_i$$
$$k_i = f\left(x_0 + c_i h, y_0 + h \sum_{j=1}^s a_{ij} k_j\right), \quad i = 1, \dots, s$$

for given coefficients  $c_i$ ,  $b_i$ ,  $a_{ij}$ . Another formulation using the same coefficients is

$$y(x+h) \doteq y_0 + h \sum_{i=1}^s b_i f(x_0 + c_i h, y_i)$$
$$y_i = y_0 + h \sum_{j=1}^s a_{ij} f(x_0 + c_j h, y_j), \quad i = 1, \dots, s.$$

- a) Show that the two formulations are equivalent, i.e. they produce the same approximation.
- a) What is the meaning of the values  $k_i$  in comparison to the  $y_i$  ?

### Exercise 8: Simple Embedded Schemes

For the numerical treatment of the initial value problem (IVP)  $y' = f(x, y)$ ,  $y(x_0) = y_0$ , we consider the embedded Runge-Kutta (RK) method

$$y_1 = y_0 + hK_1 \quad \text{and} \quad \hat{y}_1 = y_0 + h/2(K_1 + K_2)$$

with increments

$$K_1 = f(x_0, y_0)$$
$$K_2 = f(x_0 + h, y_0 + hK_1).$$

- a) State the corresponding Butcher tableaux. Verify that  $y_1$  defines a method of order 1 and  $\hat{y}_1$  a method of order 2.
- b) How many function evaluations of the right-hand side  $f(x, y)$  are necessary for this method (per step)? Which approximation  $y_1$  or  $\hat{y}_1$  is used to go on with computations? Please argue.
- c) Sketch an algorithm for the numerical approximation of an IVP applying the given method and including step size control.

**Exercise 9: Step-Size Control**

Consider the scalar IVP  $y'(x) = f(x, y(x))$ ,  $x \in [0, b]$ ,  $y(0) = y_0$  with a sufficiently smooth right-hand side  $f : [0, b] \times \mathbb{R} \rightarrow \mathbb{R}$ . Approximations  $y_n$  to  $y(x_n)$  should be computed by means of the *second order Collatz method* with variable step size

$$\begin{aligned} y_{n+1} &= y_n + h_n k_2, \\ k_1 &= f(x_n, y_n), \\ k_2 &= f\left(x_n + \frac{1}{2}h_n, y_n + \frac{1}{2}h_n k_1\right). \end{aligned}$$

*Kutta's third order rule*

$$\begin{aligned} \hat{y}_{n+1} &= y_n + \frac{1}{6}h_n(k_1 + 4k_2 + k_3), \\ k_1 &= f(x_n, y_n), \\ k_2 &= f\left(x_n + \frac{1}{2}h_n, y_n + \frac{1}{2}h_n k_1\right), \\ k_3 &= f\left(x_n + h_n, y_n - h_n k_1 + 2h_n k_2\right) \end{aligned}$$

should be used to be able to perform a step size control .

- a) State the Butcher-Tableaus for both methods.
- b) Estimate the local discretisation error  $\tau_{n+1}$  for the Collatz method. (You can just give a short formula  $\tau_{n+1} \approx e_{n+1} = \dots$  )
- c) Let  $\text{TOL}_n$  be a given tolerance for the local error in this step. What would have been an optimal step size  $h_{n,\text{opt}}$  with respect to the given tolerance? Why? Sketch an algorithm with step size control!
- d) How many evaluations of the right-hand side are necessary to compute the new approximation  $y_{n+1}$  together with a suggestion for the next step size  $h_{n+1}$  ?

**Homework 7: Step size control with ATOL and RTOL**

(10 Points)

Integration codes using step size control require the assignment of absolute and relative tolerances: **ATOL** and **RTOL**, respectively. Based on an estimate of the local error  $e^{\text{loc}}$ , we demand

$$|e_i^{\text{loc}}| \leq \text{ATOL} + WT_i \cdot \text{RTOL}, \quad WT_i = \max(|y_{0i}|, |y_{1i}|) \quad (\star)$$

for all vector components  $i = 1, \dots, n$ . Here,

$$y_0 = (y_{01}, y_{02}, \dots, y_{0n})^\top \quad \text{and} \quad y_1 = (y_{11}, y_{12}, \dots, y_{1n})^\top$$

denote the previously and newly computed approximation.

- a) Let  $y_0 = (500, 0.005)^\top, y_1 = (499, 0.008)^\top, e^{\text{loc}} = (3, 0.005)^\top$  be given. For which of the 4 combinations of  $\text{ATOL} = 0, \text{ATOL} = 0.01$  and  $\text{RTOL} = 0, \text{RTOL} = 0.01$  is the criterion  $(\star)$  fulfilled?
- b) For the step size control, one introduces the quantity

$$\text{ERR} := \sqrt{\frac{1}{n} \sum_{j=1}^n \left( \frac{e_j^{\text{loc}}}{\text{ATOL} + WT_j \cdot \text{RTOL}} \right)^2}.$$

Let the condition  $(\star)$  be fulfilled for all components. We assume  $\text{ATOL} \neq 0$ . Which property follows for ERR? For which value of ERR do we get an optimal step size (w.r.t. the given tolerances)?

- c) Let  $y_1$  and  $\hat{y}_1$  be numerical approximations of order  $p$  and  $p + 1$ , respectively. Formulate an algorithm for the calculation of the optimal step size based on ERR using the (only necessary) criterion obtained in part (b).

**Homework 8:** *Runge-Kutta-Fehlberg 1(2)*

(10 Points)

Our aim is to solve the ODE-IVP  $y' = f(x, y)$ ,  $y(x_0) = y_0$  numerically including a step size control. For this purpose, we use an embedded scheme based on explicit Runge-Kutta methods with  $s = 3$  stages. The corresponding Butcher tableau reads:

$c_1$	0	0	0
$c_2$	$a_{21}$	0	0
$c_3$	$a_{31}$	$a_{32}$	0
	$b_1$	$b_2$	0
	$\hat{b}_1$	$\hat{b}_2$	$\hat{b}_3$

The corresponding approximations are

$$y_1 = y_0 + h(b_1 k_1 + b_2 k_2),$$

$$\hat{y}_1 = y_0 + h(\hat{b}_1 k_1 + \hat{b}_2 k_2 + \hat{b}_3 k_3),$$

with the increments  $k_1, k_2, k_3$ .

- a) How many function evaluations are required in each step to calculate the approximations  $y_1$  and  $\hat{y}_1$ ?
- b) Choose the coefficients  $c_3, a_{31}, a_{32}$  (possibly in dependence on the other parameters) such that the function evaluation  $f(x_0 + c_3 h, z_3)$  of one step coincides with the function evaluation  $f(x_0, y_0)$  in the subsequent step. This technique is called FSAL (first same as last) and saves one function evaluation.
- c) Formulate the conditions such that  $y_1$  and  $\hat{y}_1$  become consistent approximations of order 1 and 2, respectively. Determine a set of coefficients satisfying these order conditions and the conditions from part (b) as well as the fundamental property

$$c_i = \sum_{j=1}^3 a_{ij} \quad \text{for each } i = 1, 2, 3.$$

(A feasible set of coefficients is not unique here. Try to find the free parameters.)